



HAL
open science

Jalons pour une théorie du questionnaire dans les sciences de l'homme

Jean-Paul Grémy

► **To cite this version:**

Jean-Paul Grémy. Jalons pour une théorie du questionnaire dans les sciences de l'homme. Psychologie. Université René Descartes - Paris V, 1980. Français. NNT : . tel-00655203

HAL Id: tel-00655203

<https://theses.hal.science/tel-00655203>

Submitted on 27 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ RENÉ DESCARTES

SCIENCES HUMAINES – SORBONNE

THESE POUR LE DOCTORAT ES LETTRES ET SCIENCES HUMAINES

présentée par

Jean-Paul GRÉMY

***Jalons pour une théorie
du questionnaire***

dans les sciences de l'homme

Tome 1

Sous la Direction
de Monsieur le Professeur Roger DAVAL

UNIVERSITÉ RENÉ DESCARTES

*Laboratoire d'Étude des Méthodes et Techniques
de l'Analyse Sociologique*

Jean-Paul GRÉMY

JALONS POUR UNE THÉORIE DU QUESTIONNAIRE
DANS LES SCIENCES DE L'HOMME

Thèse présentée devant l'Université de Paris V
– Le 7 Juillet 1980 –

LEMTAS Avril 1980

Je tiens à remercier toutes celles et tous ceux qui ont favorisé l'avancement de ce travail, et en premier lieu Monsieur le Professeur Roger DAVAL, qui a accepté la direction de cette thèse et en a encouragé la réalisation. En outre, mes remerciements s'adressent plus particulièrement à (dans l'ordre du plan de l'exposé) :

Mesdames Solange LARCEBEAU et Françoise AUBRET, qui m'ont aimablement donné accès à la testothèque du Service de Recherches de l'INOP ;

Madame Janina LAGNEAU, qui m'a fait connaître les travaux des méthodologues polonais et m'a permis de découvrir, à travers ceux-ci, les possibilités d'application aux questionnaires de la logique érotétique ;

Mademoiselle Anne DARSKA et de nouveau Madame LAGNEAU, pour m'avoir aidé de leurs compétences linguistiques à mieux comprendre les textes de KUBIŃSKI ;

Mademoiselle PAULHIAN, grâce à qui j'ai pu avoir en communication certains textes, introuvables en France, de HAMBLIN et d'ÅQVIST ;

Monsieur Nuel D. BELNAP, qui m'a communiqué son texte sur la logique des questions datant de 1963, et devenu pratiquement introuvable depuis, ainsi d'ailleurs que des tirés à part d'articles plus récents ;

Monsieur Bernard MONJARDET, qui m'a indiqué la solution du problème des permutations à positions interdites ;

Monsieur Claude-François PICARD et Mademoiselle Bernadette BOUCHON, avec qui j'ai eu de nombreux et fructueux échanges de vues sur la formalisation des questionnaires ;

Messieurs Marc BARBUT, Jean-Michel GRÉMY, et Henri ROUANET, qui m'ont aidé et encouragé dans la définition des relations \mathcal{P} et \mathcal{D} , et la recherche de leurs propriétés ;

Monsieur Mohamed CHERKAOUI, à qui je dois d'avoir eu la possibilité de présenter certains résultats de cette recherche lors du Colloque Franco-Allemand de Sociologie en 1978 ;

Madame Roselyne CANTAREL, qui a assuré avec beaucoup de compétence la tâche ingrate de transformer mon manuscrit en un texte agréable à voir, sinon à lire.

Jean-Paul GRÉMY

I N T R O D U C T I O N

L'objet de ce travail est de présenter quelques-uns des axes de recherches envisageables sur le questionnaire et son utilisation dans les sciences de l'homme. Le terme même de *questionnaire* évoque en général, non plus le sens vieilli d' "ouvrage ou partie d'ouvrage dans lequel se trouvent formulées les questions que l'on peut adresser à des élèves sur les objets de leurs études", mais plutôt une "série de questions que l'on pose pour servir de guide à une enquête" (Émile LITTRÉ, *Dictionnaire de la langue française*). C'est donc avant tout à l'outil de l'enquête économique ou sociale que se réfère l'usage courant de ce mot.

Si, encore actuellement, le questionnaire d'enquête n'a guère fait l'objet de recherches théoriques, son utilisation intensive a permis d'accumuler une expérience considérable sur l'art de formuler les questions et de construire un questionnaire, sur la manière de former les enquêteurs et de réduire les biais liés à la situation d'enquête, et sur la validation et l'interprétation des réponses recueillies. Diverses recherches expérimentales, conduites principalement aux États-Unis, ont d'ailleurs permis de corroborer et de préciser les intuitions des praticiens, et d'en tirer des règles empiriques concernant le bon usage du questionnaire. En réalité, le savoir d'un praticien du questionnaire d'enquête est encore avant tout un savoir-faire, acquis sur le tas, et constitué d'un ensemble de recettes dont la seule justification réside, pour la plupart d'entre elles, dans leur efficacité pratique. C'est d'ailleurs pourquoi ce savoir-faire est le plus souvent soit ignoré, soit tourné en dérision par le sociologue en chambre. Or, peut-être le moment est-il venu d'entreprendre une systématisation des acquis empiriques des

sociologues de terrain, afin de constituer une véritable théorie du questionnaire d'enquête. En effet, les progrès récents de la linguistique, de la logique appliquée, et de certaines branches des mathématiques (telles que la "théorie du questionnaire" de Claude-François PICARD), peuvent fournir une assise théorique solide à une présentation formalisée du questionnaire d'enquête.

La difficulté d'une telle formalisation prend sa source, nous semble-t-il, moins dans la complexité de l'objet de la recherche (questionnaire, situation d'enquête, analyse et interprétation des données), que dans l'actuelle division des sciences et dans ses conséquences sociales (répartition des disciplines universitaires et formation des chercheurs). Le questionnaire d'enquête en effet s'insère dans le processus général de la recherche, qui ressortit à l'épistémologie des sciences de l'homme, voire à la philosophie des sciences dans son ensemble ; ce qui n'est d'ailleurs que l'un des aspects de la philosophie de la connaissance. En tant que procédure d'interrogation verbale ou écrite, le questionnaire est un phénomène linguistique susceptible d'une analyse morpho-syntaxique et sémantique, voire d'une formalisation logique. En tant que moyen de recueillir de l'information, l'enquête est un processus interactif de communication qui est justiciable à la fois d'une analyse psycho-sociale, et d'une formalisation dans le cadre des théories mathématiques ou statistiques de l'information. Dans la mesure où la forme du questionnaire détermine la nature et la forme des informations recueillies, celui-ci peut être étudié dans l'optique des "théories des données" (pour reprendre la terminologie de Clyde H. COOMBS), mais aussi dans la perspective de la statistique descriptive ou inférentielle, puisque certaines des propriétés formelles des données ainsi déterminées peuvent interdire ou au contraire privilégier le choix de certaines techniques d'analyse statistique. Enfin, l'enquête par questionnaire est un phénomène proprement sociologique, non seulement par l'impact de la publication des résultats de sondage sur la vie politique et économique, mais aussi en tant que reflet d'une conception particulière de la vie en société. C'est pourquoi, compte-tenu du fait social que constitue l'actuelle division du travail de recherche telle que la traduit l'organisation des filières universitaires, toute recherche sérieuse sur le questionnaire apparaît comme nécessairement pluridisciplinaire.

Aussi ne doit-on pas considérer le titre peu euphonique que nous avons donné à ce travail comme l'expression d'une feinte modestie, mais au contraire comme l'affirmation d'un parti-pris fortement restrictif. Au lieu de ce que pourrait (et devrait) être une théorie du questionnaire dans les sciences de l'homme, nous ne présentons que quelques éléments de théorisation, portant évidemment sur les points qui nous ont paru offrir le moins de résistance à la formalisation. Encore cette théorisation mériterait-elle d'être poussée plus loin que nous ne l'avons fait. Par contre, nous avons cru bon d'étendre le champ d'application de ces éléments d'une théorie du questionnaire à certaines formes d'épreuves psychologiques, telles que tests d'aptitudes ou de niveau, et inventaires de personnalité. Il nous a semblé en effet que ces épreuves étaient formellement identiques (à quelques réserves mineures près) aux questionnaires d'enquête. C'est pourquoi nous utiliserons généralement le mot *questionnaire* dans un sens plus large que son sens usuel dans les sciences sociales, puisque étendu aux épreuves psychologiques utilisant des questions standardisées.

Notre fil conducteur, dans la détermination du champ de cette recherche, a été l'optique de l'*analyse secondaire* des données d'enquête. La raison majeure de ce choix réside dans un pronostic optimiste sur l'influence des progrès de l'informatique sur le développement des sciences de l'homme, et sur les conséquences possibles de cette influence. En effet, l'introduction de l'informatique dans les sciences de l'homme s'est traduite par le développement de techniques automatisées de traitement statistique (regroupées pour l'essentiel sous le vocable d' "analyse des données") et d'analyse lexicographique et linguistique ("analyse de données textuelles"), ainsi que par l'essor de la modélisation (techniques de "simulation" ou de "synthèse"). Ces applications de l'informatique consomment des informations ("données") ; leur extension a donc favorisé l'accroissement de banques de données lisibles par ordinateur (bases de "données ordinolingues", ou "archives informatiques"), et la création de banques de données nouvelles. La mise à la disposition des chercheurs en sciences de l'homme d'un volume important d'informations, corrélativement à des moyens de traitement automatique puissants et relativement bon marché, favorise la ré-utilisation de données recueillies à d'autres fins

par d'autres chercheurs. Dans les sciences sociales, cette forme de traitement, qui a reçu le nom d' "analyse secondaire", peut être utilisée à la fois comme instrument de recherche et comme outil pédagogique. Il est vrai qu'en France, l'utilisation de l'informatique par les sociologues se heurte à des obstacles relativement solides (cf. p.ex. [Grémy 1975 : 374-376]) ; mais la tendance générale nous semble reproduire, avec un certain retard, l'évolution, dans ce domaine technique, de la sociologie américaine. Un symptôme de cette tendance est la création récente, auprès du Conseil National de la Statistique d'une part, au Centre National de la Recherche Scientifique et à l'Institut National de Recherche en Informatique et Automatique d'autre part, de commissions et de groupes de travail sur les banques de données numériques et textuelles pour les sciences de l'homme. On peut donc raisonnablement penser que l'analyse secondaire de données d'enquêtes se développera au cours de la prochaine décennie, suscitant par là même un intérêt pour l'analyse logico-sémantique des questionnaires rédigés par d'autres chercheurs.

Ce parti-pris de ne traiter du questionnaire que dans la perspective de l'analyse secondaire des données nous conduit donc à prendre pour point de départ le questionnaire (ou le test) tel qu'il a été élaboré. Nous évitons ainsi le difficile problème des relations entre le texte du questionnaire et les hypothèses du chercheur, et donc de la genèse du questionnaire. S'il est en définitive assez facile de reconstituer les ultimes hypothèses de l'auteur d'un questionnaire à partir de l'ensemble des questions posées, l'analyse de la démarche de pensée qui aboutit à la rédaction du questionnaire nous paraît encore trop mal connue et probablement trop complexe pour pouvoir faire l'objet d'une tentative sérieuse de formalisation. Par contre, l'analyse linguistique d'un questionnaire, préalablement au traitement statistique des informations qu'il a permis de recueillir, semble pouvoir être formalisée relativement aisément.

Dans cette introduction, nous commençons par rappeler la place du questionnaire dans la recherche en sciences sociales et dans les applications de la psychologie. Ce rappel débouche sur une distinction importante entre le questionnaire *empirique*, objet de la présente recherche, et le questionnaire *théorique*, correspondant aux objectifs de l'auteur ou de l'utilisateur du questionnaire empirique ; la notion de questionnaire théorique joue un rôle important dans l'analyse des données. Nous abordons ensuite le problème des banques de données et de l'analyse secondaire, afin de mieux cerner le contexte dans lequel nous nous situons, et de définir plus précisément notre problématique. Nous présentons enfin les objectifs de la recherche entreprise, et le plan selon lequel en seront exposés les résultats.

0.1. LE QUESTIONNAIRE DANS LES SCIENCES DE L'HOMME

La place du questionnaire d'enquête dans la démarche du sociologue et du politologue, de même que le rôle des tests dans la pratique professionnelle du psychologue, ont été suffisamment décrits dans la littérature à orientation méthodologique pour qu'un long développement sur ces points ne soit pas nécessaire ; nous nous contenterons donc d'un simple rappel. Par contre, nous nous efforcerons de montrer, exemples à l'appui, qu'il n'y a pas de différence fondamentale entre un questionnaire d'enquête et une épreuve psychologique standardisée ; les seules différences notables ayant trait aux objectifs poursuivis, et aux conditions d'utilisation. Enfin, nous tenterons en quelques pages de situer le questionnaire dans la démarche scientifique en général.

a) Le questionnaire dans l'enquête sociologique

D'après une communication de Bernard-Pierre LÉCUYER au séminaire d'histoire de la sociologie empirique en France, il semble que la paternité des enquêtes avec questionnaire en France soit imputable à l'abbé Louis de

DANGEAU et à Sébastien LE PRESTRE, seigneur de VAUBAN. En effet, si les premières enquêtes connues sous l'ancien régime remontent au seizième siècle, il faut attendre les enquêtes de 1694 et de 1697 pour voir l'emploi sur une grande échelle d'un questionnaire uniforme (rédigé selon les principes exposés par VAUBAN dans sa *Méthode générale et facile pour faire le dénombrement des peuples*, publiée en 1686 ; cf. [Esmonin 1954 : 508]) [Lécuyer 1963 : 14-19], et la seconde moitié du dix-huitième siècle pour qu'apparaisse le questionnaire imprimé [Lécuyer 1963 : 5]. Le texte du questionnaire adressé par VAUBAN aux intendants en 1697 (cf. [Lécuyer 1963 : 20-22]) est d'ailleurs plus un guide d'entretien structuré qu'un questionnaire au sens technique du mot : la plupart des questions y sont largement ouvertes, et leur formulation s'apparente plutôt à la description d'un thème d'enquête qu'à une interrogation véritable (à l'exception des questions 30 et 32, qui comprennent des sous-questions à forme interrogative). En outre, les données qu'il vise à obtenir ne sont pas des informations individuelles, mais des données agrégatives concernant une région administrative dans son ensemble. On remarquera toutefois que "l'ordre adopté dans le questionnaire est un ordre logique [...]. D'abord la description physique et naturelle du pays (dix questions), ensuite la statistique générale (quatre questions), les questions administratives et l'état moral des populations et des individus (douze questions), puis les questions économiques et financières (quatorze questions)" [Lécuyer 1963 : 19] ; et que les thèmes abordés comprennent les mentalités et les coutumes des habitants (question 13), et la réputation des ecclésiastiques (question 24) et des magistrats (question 28).

Toujours d'après LÉCUYER, l'enquête ordonnée en 1745 par le Contrôleur Général ORRY est la première enquête effectuée en France où apparaisse la préoccupation de connaître l'opinion publique [Lécuyer 1963 : 33-34]. La procédure suivie pour ce premier sondage d'opinion est d'ailleurs fort intéressante, puisqu'il est demandé aux intendants de susciter de fausses rumeurs afin d'en observer les effets dans la population. Le texte de la question 8 est en effet le suivant (d'après [Dainville 1952 : 50-51], en modernisant l'orthographe) : "Vous ferez semer le bruit dans les villes franches de votre département d'une augmentation d'1/3 sur les droits des entrées. Vous y ferez

aussi semer les bruits, ainsi que dans le plat pays de la levée d'une future milice de deux hommes dans chaque paroisse ; le fort portant le faible, vous recueillerez avec soin ce qu'en diront les habitants et vous en ferez mention dans l'État que le Roi vous demande". Cette procédure associe par conséquent l'expérimentation à l'observation dans l'étude de l'opinion publique ; elle demande d'autre part à l'expérimentateur de dégager la tendance centrale des informations ainsi recueillies ("le fort portant le faible" = "toute compensation faite" \approx "en moyenne").

Les étapes ultérieures de l'histoire des recensements et des sondages par questionnaire sont suffisamment connues pour qu'il ne soit pas nécessaire de les rappeler ici (cf. p.ex. : [Moser 1958 : 18-38], [Stoetzel 1973 : 37-55]). En ce qui concerne la place du questionnaire dans la démarche du chercheur en sciences sociales, notre expérience personnelle (tant en psychosociologie appliquée que dans la recherche universitaire en sciences sociales) recoupe le point de vue de la plupart des auteurs ayant traité du sujet. En schématisant beaucoup, la démarche-type du chercheur peut être décomposée en douze étapes :

① *Définition du problème.* Le point de départ de la recherche est en général une question de type *pourquoi* ? [Boudon 1969 : 31], ou parfois, en sociologie appliquée, une question de type *comment* ? ("Comment faire pour ... ?") qui implique la résolution préalable d'un problème théorique de type *pourquoi* ? La première phase de la recherche est une réflexion *a priori* sur le problème posé, ses présuppositions, et ses implications. De cette analyse conceptuelle doit en principe sortir un faisceau de sous-questions spécifiques issues de la question initiale, définissant les objectifs de la recherche, ainsi que la nature et la source des informations à recueillir. De plus, cette étape permet au chercheur d'explicitier l'ensemble de ses idées préconçues (ce que DURKHEIM appelle ses *prénotions* [Durkheim 1895 : 22-23], et BACON : *anticipations* [Bacon 1620 : I § 26-28], ou : *idoles* [Bacon 1605 : v ch.IV ; 1620 : I § 38-44]), se donnant ainsi la possibilité d'en prendre conscience ultérieurement en les confrontant aux observations.

② *Recensement et analyse des travaux antérieurs* sur le sujet et dans les domaines connexes. Il peut s'agir de théories générales dont certains aspects sont susceptibles d'éclairer les phénomènes étudiés ; d'explications *ad hoc*, sortes de théories locales dont la portée est volontairement limitée à un petit nombre d'observations, mais qu'il est néanmoins possible d'étendre ou de généraliser ; de descriptions de type ethnographique (monographies, biographies, protocoles ou analyses du contenu d'entretiens peu structurés, etc.) ; de résultats d'analyse de données statistiques obtenues par recensement ou sondage ; de données d'enquête brutes, susceptibles d'être ré-analysées ; etc. Selon la nature des matériaux dont on dispose, les techniques d'analyse sont évidemment de types extrêmement variés.

③ *Prise de contact* avec l'univers social concerné par l'objet de la recherche. Le terme d' "expédition de reconnaissance" ("*scouting*"), utilisé il est vrai par Daniel KATZ dans une acception plus large englobant l'étape ⑤ [Katz 1953 : 81-90], décrit parfaitement le rôle de cette phase : donner au chercheur l'occasion de se familiariser avec un cadre géographique, un environnement, une culture, un milieu, des mentalités, des modes de vie, etc., que sa propre histoire personnelle ne lui a pas permis de connaître directement et de manière approfondie.

④ *Élaboration d'un premier embryon de théorie*, ensemble structuré d'hypothèses et de contre-hypothèses ("hypothèses de rechange") opératoires, préalable à tout travail systématique sur le terrain. Cette ébauche de théorie est une élaboration des sous-questions énoncées à l'issue de la première étape, en fonction des acquis de la réflexion et des informations apportées par les étapes ② et ③. Elle définit en particulier les thèmes de l'enquête exploratoire sur le terrain.

⑤ *Enquête exploratoire* de type "qualitatif", caractérisée par une grande souplesse dans le choix des techniques de recueil d'information et dans la sélection des informateurs. L'objectif de cette enquête est de compléter, d'affiner, de préciser l'embryon de théorie esquissé dans la phase ④ ; c'est par conséquent le développement de cette ébauche qui doit guider le chercheur. Il est possible, selon les besoins, de recourir à l'observation du cadre de vie ou des comportements, à l'expérimentation *in situ* ou en laboratoire, à

l'entretien individuel ou de groupe, etc. Toutefois, la pratique courante actuelle donne la primauté aux techniques d'entretien, dont les inconvénients ont été bien analysés (cf. p. ex. [Cannell 1968 : 535-552]), et peuvent être maîtrisés ou compensés par des chercheurs expérimentés. L'élaboration d'un ensemble cohérent d'hypothèses commence, en principe, par l'examen et la consolidation des hypothèses les plus générales. Aussi, les premiers entretiens doivent-ils être le plus ouverts et le moins structurés possible, afin d'atténuer les effets des prénotions des chercheurs. Puis, au fur et à mesure que la théorie se développe, les thèmes de l'enquête exploratoire deviennent plus spécifiques, et les entretiens plus focalisés et plus structurés. L'enquête exploratoire se distingue par conséquent de l'enquête extensive de la phase ⑨ en ce qu'elle est relativement étendue dans le temps (il faut analyser le contenu des entretiens réalisés avant d'en entreprendre d'autres), que son déroulement et ses objectifs peuvent être modifiés à tout moment (en fonction des informations recueillies), que l'analyse des protocoles d'entretiens est centrée sur la compréhension de chaque cas individuel, et enfin que la sélection des informateurs ne vise pas à la représentativité, mais plutôt à la diversité des cas analysés (échantillons des cas théoriquement envisageables plutôt qu'échantillon d'individus).

⑥ *Re-définition du problème* à partir du bilan des phases antérieures. Il est commode, parvenu à ce stade, de rédiger un rapport de l'enquête exploratoire présentant une synthèse des informations recueillies, énumérant les principaux problèmes encore en suspens et les hypothèses envisageables à leur propos, et décrivant les procédures à mettre en oeuvre pour aboutir à une solution (souvent partielle et provisoire) de ces problèmes.

⑦ *Préparation de l'enquête extensive*. Le pré-rapport rédigé à l'étape précédente peut évidemment préconiser l'approfondissement qualitatif, ou au contraire l'arrêt des recherches. Dans la plupart des cas, si les étapes ① à ⑥ ont été menées à leur terme, ce pré-rapport constitue plutôt une introduction à l'enquête extensive qui est la suite obligée et le prolongement naturel de la phase exploratoire. Il fournit en principe des éléments de réponse à trois questions : "Qui interroger ?" (plan d'observation ou de sondage), "A quel sujet ?" (liste des thèmes à aborder dans le questionnaire), "Comment

exploiter les réponses ?" (premier plan de dépouillement des données).

⑧ *Mise au point du questionnaire.* Le questionnaire sert à recueillir, auprès des personnes interrogées, une certaine information dont le chercheur a besoin pour étayer, chiffrer, ou infirmer ses hypothèses. La rédaction du questionnaire d'enquête est un art difficile, nécessitant un apprentissage sur le tas, une bonne connaissance du terrain, et une solide formation théorique. Le chercheur peut d'ailleurs, pour formuler ses questions, tirer parti d'une analyse linguistique du contenu des protocoles d'entretiens de la phase ⑤, et bien entendu des questionnaires rédigés par ses prédécesseurs. Quels que soient sa compétence et les efforts qu'il déploie, il ne peut échapper à la nécessité de l'essai préalable ("*pretesting*") du questionnaire, afin d'améliorer la formulation et l'ordre de présentation des questions. Dans certains cas, il peut être indispensable de procéder à une enquête pilote ("*pilot survey*"), par exemple pour évaluer les taux de non réponses suscitées par une question donnée (cf. [Moser 1958 : 46-47]).

⑨ *Enquête extensive.* Elle ne se déroule pas de la même manière selon qu'il s'agit de questionnaires postaux, d'interviews téléphoniques, de questionnaires remplis collectivement par les personnes interrogées, ou de questionnaires administrés par enquêteurs. Mais dans tous les cas, il s'agit d'une forme de travail de type industriel (par opposition au caractère artisanal des phases précédentes), dans laquelle les délais et la standardisation de l'outil jouent un rôle capital.

⑩ *Contrôle des données.* Cette phase comprend la relecture des questionnaires remplis, la codification des questions ouvertes, le report des réponses sur cartes perforées ou sur bande magnétique, le nettoyage de ces données, et les premiers dépouillements destinés à vérifier le respect du plan d'observation ou de sondage, l'application des consignes données aux enquêteurs, etc. Selon le résultat des diverses opérations de contrôle, les actions correctives envisageables peuvent être le renvoi à l'enquêteur d'un questionnaire à refaire ou à compléter, le remplacement d'un enquêteur défaillant, la re-définition des catégories d'analyse de contenu, la modification du plan de sondage et le lancement d'une vague d'enquête supplémentaire, etc.

⑪ *Analyse des données.* Cette étape a pour but d'extraire des informations recueillies des éléments de réponses aux questions du chercheur telles qu'elles étaient exprimées à l'issue de la phase ⑥ . L'analyse des données est un processus de condensation des informations apportées par l'enquête afin de mettre en évidence certaines régularités dans les réponses enregistrées (liaisons entre variables, typologies des répondants, etc.). La procédure suivie est nécessairement interactive, puisque l'analyse des données a pour première fonction de dégager des présomptions ou des arguments pour étayer certaines des hypothèses du chercheur ; en cas d'échec ou d'incertitude, le chercheur est amené à modifier ses hypothèses, et à procéder à d'autres analyses. En outre, s'il s'agit d'un sondage, la perspective du chercheur est nécessairement de type inférentiel, puisque sa préoccupation est la "validité" de ses hypothèses sur la population dans son ensemble. Cette démarche est matérialisée par les plans de dépouillement successifs des données de l'enquête extensive.

⑫ *Synthèse des résultats.* A partir de l'ensemble des matériaux recueillis ou analysés au cours des étapes précédentes, le chercheur s'efforce de fournir une réponse précise et argumentée à la question initiale qui a motivé la recherche.

Cette présentation du déroulement-type d'une recherche sociologique utilisant l'enquête sur le terrain est avant tout le reflet simplifié de notre expérience personnelle. Nous avons toutefois annoncé qu'à peu de choses près, notre description est compatible avec celle qu'en donnent les spécialistes qui ont traité des méthodes d'enquête. En effet, les divergences que nous avons relevées ne portent guère que sur la manière de découper ce processus en phases distinctes, et non sur le processus lui-même. Par exemple, Daniel KATZ ne considère que six grandes étapes : le projet initial (correspondant aux phases ①, ② et ④), l'expédition de reconnaissance (phases ③ et ⑤), l'élaboration du plan de recherches (phases ⑥ et ⑦), l'épreuve préalable des instruments et des procédures (phase ⑧), l'enquête sur le terrain (phase ⑨), et l'analyse des matériaux (phases ⑩ à ⑫) [Katz 1953 : 79-115]. De même, Raymond BOUDON distingue quatre grands moments de la recherche : la formulation des hypothèses (phases ① à ⑥), la construction du plan d'observation (phases ⑦ à

(10)), la construction des variables (fin de la phase (8) et premier moment de la phase (11)), et l'analyse des relations entre variables (phases (11) et (12)) [Boudon 1969 : 31-71]. Ces deux exemples révèlent de légères différences dans l'ordre de présentation de certaines étapes, ou dans la subdivision possible d'une même phase. Sur ce deuxième point, nous pourrions citer également l'éclatement du problème du plan de sondage dans la phase (7) en détermination de l'univers de l'enquête et échantillonnage proprement dit, et la division de la phase (8) en mise à l'épreuve du projet de questionnaire et rédaction définitive de celui-ci [Mucchielli 1967 : 9-10]. Chaque problème peut ainsi être divisé, selon la procédure cartésienne (*Règle XIII* [Descartes : X 430]), en problèmes plus élémentaires (cf. p.ex. [Javeau 1971 : 4-5]), sans pour autant altérer le sens de cette description.

En ce qui concerne l'ordre de présentation des étapes, deux remarques sont à faire. La première est que l'ordre "logique" d'exposition n'est pas nécessairement l'ordre chronologique. Comme le soulignent entre autres JAVEAU, et GHIGLIONE et MATALON, les différentes étapes ne se déroulent pas selon un ordre linéaire et constant ([Javeau 1971 : 5], [Ghiglione 1978 : 21]). Il y a le plus souvent chevauchement et interaction entre les étapes, et par exemple les phases (2), (4) et (5) peuvent être poursuivies simultanément sans que la recherche ait à en souffrir. La seconde remarque est que, dans une recherche concrète, certaines phases peuvent se trouver omises et d'autres au contraire réitérées. Une recherche consistant en une relecture d'un grand auteur, ou en une analyse secondaire de données empiriques, ne comporte que les phases (1), (2) et (12). Une étude qualitative se limite aux phases (1) à (6), ou, selon sa conclusion, aux phases (1) à (5) suivies de la phase (12). Par contre, une recherche portant sur l'évolution dans le temps d'un phénomène social (étude sur panel, "baromètre" d'opinion) comporte une réitération, à divers moments, des phases (9) à (11). Il peut de même arriver qu'à l'issue de la onzième étape, le comportement jugé "déviant" d'une sous-population donnée impose le retour à une approche qualitative de cette sous-population, c'est-à-dire une réitération de la phase (5), éventuellement suivie de la phase (6). A titre d'exemple, Bernard S. PHILLIPS décrit une recherche menée en deux temps : une première enquête correspondant aux phases (1) à (11), suivie d'une seconde correspondant aux phases (6) à (12) [Phillips 1966 : 5-25]. On trouverait sans peine d'autres illustrations de la diversité des stratégies de la recherche en sciences sociales,

sans toutefois remettre en cause le schéma décrit ci-dessus. Ajoutons qu'il est bien connu que le processus de recherche est essentiellement cyclique, les conclusions d'une recherche donnée fournissant un point de départ pour les recherches ultérieures.

A l'issu de ce survol, la place du questionnaire dans l'enquête sociologique apparaît comme clairement précisée, du moins si l'on s'en tient au questionnaire de la phase extensive (phase ⑨). Il peut toutefois arriver que l'on ait également recours à un véritable questionnaire à la fin de la phase exploratoire (phase ⑤), ce questionnaire étant cependant plus "ouvert" en principe que celui de la phase extensive. Le rôle du questionnaire ainsi défini est de servir au recueil normalisé d'informations auprès des personnes interrogées, en vue de fournir une description précise et détaillée du phénomène social étudié ; une telle description sociographique ayant pour fonction d'apporter des éléments de réponse aux questions que se pose le chercheur. Cette dernière considération suggère d'ailleurs une remarque : pourquoi ne pas considérer également l'ensemble des questions que le chercheur se pose comme une sorte de questionnaire ? Dans une telle perspective, il est nécessaire de bien distinguer cette dernière forme, que nous appellerons questionnaire *théorique*, du questionnaire des phases ⑤ et ⑨, que nous qualifierons d'*empirique*. Nous aurons l'occasion de revenir sur ces notions dans la suite de cette introduction, ainsi qu'au chapitre 3. Notons seulement que, dans une recherche du type que nous avons décrit, il n'y a pas *un* questionnaire théorique, mais une série, matérialisée par les étapes successives de l'élaboration et de la re-définition du problème initial, et par les divers plans de dépouillement des données recueillies.

b) Le questionnaire en psychologie appliqué

Un grand nombre d'épreuves psychologiques portent le nom de *questionnaire* : questionnaires d'intérêts (de type SRONG ou KUDER), et questionnaires de personnalité (de BERNREUTER, de BERGER, de CATELL, etc.). La méthode des questionnaires "consiste à faire remplir un questionnaire avec, soit un choix entre plusieurs réponses, soit une réponse par oui ou

par non, soit, plus rarement, par une réponse libre pour un ensemble de questions qui ont été l'objet d'une élaboration spéciale, et souvent d'un étalonnage. La méthode est utilisée pour la détermination des traits de caractère, des goûts, des intérêts" (Henri PIÉRON, *Vocabulaire de la psychologie*). Déjà au début du vingtième siècle, Théodule RIBOT notait, à propos de l'emploi de questionnaires en psychologie, que "malgré quelques protestations isolées, ce moyen d'investigation s'est tellement répandu et est devenu d'un si constant usage qu'il est utile d'en examiner la portée et la valeur" [Ribot 1904 : 1]. Or, la plupart des auteurs s'accordant à en faire remonter l'usage à Francis GALTON peu avant 1880 (cf. [Ribot 1904 : 1], [Reuchlin 1957 : 19]), il a suffi par conséquent d'un quart de siècle pour que cette technique s'impose auprès des psychologues.

Si l'on se réfère aux travaux sur l'usage des questionnaires en psychologie, la frontière n'est pas facile à tracer entre les questionnaires psychologiques et les tests. Il arrive que les mots soient interchangeables, et, comme le fait remarquer Paul ALBOU, la confusion terminologique entre *test*, *questionnaire*, *inventaire*, *formulaire*, *indice*, *échelle*, etc., n'est pas aisée à dissiper [Albou 1968 : 11]. La définition adoptée par l'Association Internationale de Psychotechnique pour le mot *test* est la suivante : "C'est une épreuve définie, impliquant une tâche à remplir, identique pour tous les sujets examinés, avec [une] technique précise pour l'appréciation du succès ou de l'échec, ou pour la notation numérique de la réussite. La tâche peut comporter une mise en oeuvre, soit de connaissances acquises (*test pédagogique*), soit de fonctions sensori-motrices ou mentales (*test psychologique*)" (Cité par Henri PIÉRON, *Vocabulaire de la psychologie*). Cette définition englobe par conséquent les questionnaires de connaissances scolaires ; mais les commentaires d'Henri PIÉRON l'étendent aux tests de *caractère* ou de *personnalité* (y compris aux tests *projectifs*), alors que l'on ne peut évidemment pas, dans ce type d'épreuve, parler de réussite. Formellement, tous les tests auxquels le sujet donne une réponse linguistique par écrit sont des questionnaires au sens de PIÉRON. Si l'on accepte d'étendre la définition aux questionnaires dont les réponses sont exprimées oralement, les seules formes d'épreuves psychologiques qui ne

peuvent pas être assimilées aux questionnaires sont celles dans lesquelles la réponse résulte d'un ensemble de gestes (test myokinétique de MIRA Y LOPEZ, "waltzing test" de HIRSCH) ou de la manipulation d'objets divers (test du tourneur, test de PIORKOWSKY, puzzles, peg-boards, "wiggly block", cubes de KOHS, etc.), ainsi que celles dans lesquelles la réponse écrite ou verbale est dénuée de toute signification linguistique utile pour l'évaluation de la performance (tests de barrage, épreuves de temps de réaction à réponse verbale). Encore doit-on reconnaître que, dans ce cas également, la frontière n'est pas facile à tracer entre le questionnaire psychologique au sens large, et les tests qui ne peuvent pas être considérés comme des questionnaires (exemples de cas limites : tests de mémoire immédiate avec répétition de mots ou de phrases ; épreuves de dessin, soit pour évaluer le niveau mental, comme le "test du bonhomme" de Florence GOODENOUGH, soit pour l'étude de la personnalité, comme le "test de l'arbre" de Karl KOCH).

En donnant au mot *questionnaire* le sens très général de "série de questions présentées dans l'ordre dans lequel elles doivent être posées", nous considérerons par conséquent comme des questionnaires aussi bien les épreuves psychologiques individuelles administrées oralement que les tests collectifs papier-crayon. Les propriétés communes à l'ensemble de ces questionnaires psychologiques concernent d'une part le caractère standardisé de l'épreuve, des conditions de passation, de l'enregistrement des réponses (idéalement, l'épreuve est rigoureusement la même pour tous les sujets testés), et d'autre part l'utilisation qui est faite des réponses fournies (système de codification extrêmement précis et rigide, calcul d'une note globale). Ces caractères sont imposés par l'objectif même du test, qui est de "classer le sujet examiné, soit quantitativement, soit typologiquement" par rapport à d'autres individus placés dans la même situation [Pichot 1954 : 5] ; ce que Lee J. CRONBACH résume en définissant le test comme "une procédure systématique pour comparer le comportement de deux personnes ou plus" [Cronbach 1949 : 11]. Nous reviendrons sur cette importante définition au chapitre 3, à propos des fondements de l'analyse des données.

Pour évaluer le rôle que jouent ces questionnaires dans l'exercice du métier de psychologue, nous décrivons les principales étapes dans l'emploi d'un test standardisé, en nous appuyant à la fois sur notre expérience personnelle (malheureusement plus limitée qu'en matière d'enquêtes psychosociales, et portant surtout sur les examens d'enfants et d'adolescents), et sur les commentaires relevés dans les ouvrages spécialisés [(Goodenough 1949 : 297-307, 455-503], [Cronbach 1949 : 20-83], [Remmers 1955 : 49-72, 121-175]). On peut ainsi considérer qu'un examen psychologique individuel courant, en milieu professionnel ou scolaire, peut être réduit aux onze étapes suivantes :

① *Définition des objectifs* de l'examen. Ils varient naturellement selon le contexte institutionnel dans lequel celui-ci se situe, selon qu'il est demandé par le sujet ou qu'il lui est imposé, et selon les motifs ou les buts de l'examen (dépistage psychiatrique, orientation scolaire, sélection en vue d'une embauche, conseil pour une reconversion professionnelle, etc.).

② *Analyse du dossier* du sujet : biographie, examens psychologiques antérieurs, informations médicales, résultats scolaires et professionnels, etc., afin de se faire une première idée du cas. A l'issue de cette analyse, le psychologue a en principe défini les grandes lignes de la stratégie de l'examen.

③ *Premier entretien* avec le sujet. Cet entretien de prise de contact est pour le psychologue l'occasion d'expliquer le but et les modalités de l'examen, et de répondre aux questions que le sujet se pose à ce propos ; d'établir un climat de confiance encourageant le sujet à faire de son mieux au cours des épreuves ; de recueillir sur la vie du sujet, ses intentions, ses motivations, ses goûts, ses réussites et ses échecs antérieurs, etc., une série d'informations confirmant, corrigeant, ou complétant les informations contenues dans le dossier ; et enfin de porter un diagnostic provisoire sur la personnalité et les capacités du sujet.

④ *Organisation et choix des épreuves.* Le diagnostic provisoire formulé par le psychologue et les objectifs de l'examen déterminent le choix et l'ordre d'administration des tests psychologiques. Dans la plupart des situations réelles, le cadre institutionnel et les préférences techniques du psychologue (dues pour l'essentiel à sa formation et à son expérience professionnelles) réduisent fortement l'univers de choix des épreuves. Le premier élément de décision dans le choix d'un test est naturellement son adéquation au problème posé (Que "mesure" le test ? Quelle est sa valeur de diagnostic et de pronostic ? Dispose-t-on d'un étalonnage adapté au sujet ? etc.) ; entrent également en ligne de compte les qualités internes du test (fidélité, validité, finesse discriminative, etc.), et les considérations pratiques (durée de l'épreuve, facilité d'administration et de correction, disponibilité du matériel et du local, etc.).

⑤ *Administration des épreuves.* En principe, l'administration des tests obéit à des règles extrêmement strictes, et demande au psychologue qui les utilise une formation poussée, et une grande familiarité avec les épreuves et le matériel utilisé.

⑥ *Codage des réponses.* Dans la plupart des tests papier-crayon et des questionnaires de personnalité ou d'intérêts, ce codage est effectué sur la feuille de réponse par le sujet lui-même (questions fermées). Dans les épreuves individuelles administrées oralement (tests de niveau du type BINET-SIMON, TERMAN, WECHSLER, etc., et tests projectifs tels que RORSCHACH, T.A.T., C.A.T., Fables de DUSS, etc.), chaque réponse doit être interprétée, et codée selon une grille d'analyse de contenu pré-établie.

⑦ *Calcul des scores.* A partir du dénombrement de chacun des types de réponses relevés, le psychologue calcule soit une note brute globale par test (dans le cas par exemple de la plupart des tests de développement ou de niveau intellectuel), soit une série de notes brutes correspondant chacune à l'une des dimensions "mesurées" par le test (dans le cas par exemple des questionnaires de personnalité, ou des tests dits "factoriels").

⑧ *Evaluation des résultats*, en comparant les scores obtenus par le sujet à la distribution des scores d'une population de référence (étalonnage).

⑨ *Synthèse des résultats* en tenant compte également des éléments du dossier, des apports de l'entretien, des indications sur les conditions de passation des épreuves (Le sujet était-il à l'aise, ou contracté ? Était-il suffisamment motivé pour faire de son mieux ? Y-a-t-il eu des incidents perturbant le bon déroulement d'une épreuve ? etc.), et des observations qualitatives faites au cours de celles-ci (méthode utilisée pour parvenir à la réponse ; primauté à la rapidité ou à la précision ; etc.).

⑩ *Entretien final*, au cours duquel les résultats des épreuves sont présentés et expliqués au sujet (ou à ses parents ou tuteurs s'il s'agit d'un jeune enfant, d'un retardé mental, ou d'un malade jugé irresponsable). Cet entretien permet au psychologue d'avoir l'explication "d'anomalies" dans les résultats (consigne mal comprise ou mal interprétée, blocage, fatigue, etc.), de recueillir des informations complémentaires sur la personnalité du sujet et son environnement, et d'explorer avec lui le champ des décisions possibles à l'issue de l'examen.

⑪ *Rédaction du compte-rendu d'examen*, ainsi que des pièces administratives annexes, nécessaires aux prises de décisions concernant le sujet, et à leur exécution.

Comme pour le déroulement-type d'une enquête sociologique, ce schéma d'un examen psychologique individuel appelle de nombreux correctifs. En premier lieu, l'examen peut être très étendu dans le temps, divisé en périodes entre lesquelles peuvent prendre place des phases d'activité scolaire ou professionnelle, des périodes de formation ou d'essai, des séances ou des cures thérapeutiques ; chaque période peut alors comporter un entretien avant et un entretien après les tests proprement dits. A l'inverse,

l'examen peut se limiter à une partie seulement de notre schéma ; par exemple, à une série de tests administrés collectivement, suivie du compte-rendu (phases ⑤ à ⑪), avec ou sans l'entretien final ⑩). D'autre part, suivant le cadre institutionnel dans lequel se déroule l'examen, la liberté de choix des épreuves par le psychologue est plus ou moins grande ; dans la plupart des cas dont nous avons eu connaissance, il existe pour chaque type de problème une batterie de tests minimale et quasi-obligatoire, que le psychologue peut compléter s'il le désire par des épreuves de son choix. Enfin, le psychologue a, le plus souvent, la possibilité de modifier la stratégie de l'examen en cours de route, au vu des premiers résultats.

Cette description de la place des questionnaires et des tests en psychologie appliquée appelle toutefois une remarque importante : elle ne traite pas de l'élaboration et de la mise au point de ces épreuves. En effet, il est relativement rare de voir le psychologue construire lui-même les outils dont il a besoin. La raison en est que le test, instrument de mesure et de comparaison, doit posséder certaines des qualités que l'on exige des appareils servant aux mesures physiques : sensibilité, mobilité, fidélité, précision (cf. [Pérard 1947 : 17-23]). Sa réalisation, lente et difficile, reste l'affaire d'un petit nombre de spécialistes ; sa mise au point et son étalonnage, préalables à son utilisation systématique, demandent un travail collectif d'expérimentation étalé souvent sur plusieurs années. En outre, les épreuves psychologiques subissent deux influences perturbatrices dont les instruments de mesure physique sont en principe beaucoup moins affectés : avec le temps, elles se dérèglent et doivent être ré-étalonnées ; et, dans certains cas, elles finissent par mesurer autre chose que ce qu'elles étaient censées mesurer à l'origine.

Le problème du ré-étalonnage ne se pose pas dans les mêmes termes que pour les mesures physiques. En effet, dans ce dernier cas, il existe une valeur de référence, appelée *étalon*, et supposée constante ; cette valeur permet d'une part de définir l'unité de mesure, et d'autre part de régler les instruments servant à mesurer. Dans le cas des épreuves psychologiques, l'étalon est défini en termes statistiques, en fonction de normes sociales particulières ; si l'on admet qu'il n'existe pas d'épreuve psychologique

"indépendante du type de civilisation" ("*culture-free*" test), l'évolution de la société peut entraîner un changement de la correspondance entre les résultats aux tests et la valeur de l'étalon originel. Par exemple, d'après le premier étalonnage du BINET-SIMON (1911), un enfant de quatre ans réussissait en moyenne neuf épreuves, et un enfant de sept ans, vingt-quatre ; trente-cinq ans plus tard (en 1946), René ZAZZO constatait qu'un enfant de quatre ans réussissait en moyenne quatorze épreuves, et un enfant de sept ans, vingt-neuf (cf. [Zazzo 1966 : I 57]). Le test était donc devenu plus facile (du moins pour les jeunes enfants), et devait être ré-étalonné. Parallèlement, les résultats scolaires moyens des enfants des classes primaires s'étaient améliorés (cf. le "barème d'instruction" publié dans [Binet 1909 : 35]).

Le problème de l'évolution de la signification même du test, de "ce qu'il mesure", tient également à l'évolution de la civilisation. D'ailleurs, dans l'exemple du BINET-SIMON, René ZAZZO a tout d'abord publié un ré-étalonnage du test ([Zazzo 1949]), pour finalement aboutir à la création d'une "nouvelle échelle métrique de l'intelligence" ([Zazzo 1966]) intégrant certains items du BINET-SIMON. Ce qui est vrai pour un changement temporel de civilisation l'est également pour un changement géographique : on connaît les multiples adaptations anglo-saxonnes du BINET-SIMON (notamment par TERMAN, YERKES, et leurs collaborateurs ; cf. [Goodenough 1949 : 59-67]). Un autre aspect du phénomène de vieillissement d'un test est l'incidence de sa vulgarisation (les "bonnes réponses" sont connues d'avance par les sujets), et surtout de la vulgarisation des principes qui ont présidé à sa construction. Ainsi, aux Etats-Unis, où les tests ont acquis une valeur sociale importante (épreuves psychologiques pour l'admission à l'Université, prestige social du Q.I., etc.), il est normal de s'entraîner pour obtenir un score élevé à certains tests (cf. p.ex. [A. 1965]) ; de plus, en réaction contre la "testomanie", quelques psychologues américains ont porté à la connaissance du grand public le contenu de certaines épreuves ; enfin, de nombreux jeux diffusés par la radio, la télévision, et les hebdomadaires, s'inspirent d'épreuves psychologiques (de niveau ou de personnalité) classiques. La conséquence en est qu'il est possible de développer une aptitude particulière, l'aptitude à

réussir aux tests d'efficience et à "bien répondre" aux épreuves de personnalité, qui fausse évidemment la signification des résultats telle qu'elle était définie dans la version primitive de l'épreuve.

Il y a donc, pour les psychologues, nécessité à la fois de conserver les tests éprouvés, et de les remplacer par d'autres plus récents. Il est donc naturel que de nouvelles épreuves soient expérimentées chaque année, pour qu'un petit nombre d'entre elles puisse être finalement adoptées. En outre, les progrès des connaissances de base en psychologie conduisent à la création d'épreuves tout à fait nouvelles. Ainsi, les progrès de l'analyse factorielle ont permis à Raymond B. CATTELL la mise au point et la validation d'un questionnaire mesurant seize facteurs de la personnalité. Dans un tout autre domaine, les recherches de Jean PIAGET ont servi de point de départ pour la conception par François LONGEOT d'une échelle de développement de la pensée logique chez l'enfant, qui n'a pas d'équivalent parmi les épreuves psychologiques mises au point antérieurement. D'autre part, il semble que les pédagogues américains aiment mieux fabriquer leurs propres tests de connaissances scolaires qu'utiliser des épreuves standardisées (cf. [Remmers 1955 : 70]) ; aussi existe-t-il une littérature nombreuse sur la construction des épreuves psychologiques (cf. par exemple [Conrad 1948 : 631-657], [Goodenough 1949 : 97-293], [Remmers 1955 : 73-120, 176-192], etc.). En nous inspirant de cette littérature (et tout particulièrement de [Tinkelman 1971] et [Henrysson 1971]), mais également de l'historique de certains tests ([Binet 1907 : 97-99 ; 1909 : 103-125], [Terman 1937 : 3-51], [Zazzo 1966 : I 9-66] pour le BINET-SIMON ; [Adorno 1950 : 222-262] pour l'échelle de F ; etc.), nous décrivons succinctement ci-après les étapes de l'élaboration d'une épreuve psychologique :

① *Définition des objectifs* de l'épreuve (détecter ou évaluer certains traits de personnalité ; mesurer telle aptitude ou telles connaissances ; définir les aspirations, les intérêts, les goûts ; etc.), et aussi de la population à laquelle elle doit être appliquée, et du contexte institutionnel dans lequel les résultats seront utilisés.

② *Recherche des indicateurs possibles* des traits ou des aptitudes que l'on désire évaluer. Les trois sources principales d'indicateurs sont l'observation et l'analyse du comportement dans la vie quotidienne, les conséquences logiquement déductibles des théories psychologiques existantes, les items déjà utilisés dans d'autres épreuves (et d'ailleurs susceptibles de correction, de notation, et d'interprétations différentes).

③ *Sélection des indicateurs* jugés les plus intéressants *a priori* ("échantillon" dans l'univers des indicateurs possibles) et rédaction des questions ou des items correspondants.

④ *Essai des items sélectionnés* auprès d'individus appartenant à la population à laquelle l'épreuve est destinée, afin de mettre au point la formulation des items et l'ensemble de la procédure de l'épreuve.

⑤ *Analyse statistique interne* du test, afin de détecter les items ne mesurant pas la même chose que l'ensemble des items du test (ou d'un sous-test). Les procédures les plus usuelles pour effectuer cette détection sont l'analyse de la variance (évaluation de la contribution de chaque item dans la variance de la note globale), l'analyse factorielle (comparaison des saturations des items dans chacun des facteurs), et l'analyse hiérarchique (élimination des items "non scalables") ; mais il existe bien d'autres techniques utilisables pour une telle analyse (cf. p.ex. [Stanley 1971]). Cette phase débouche naturellement sur la "purification" du test, par réduction du nombre d'items utilisables, et la mise au point de la procédure de notation.

⑥ *Mise au point de la version provisoire* du test, comprenant la reformulation et le reclassement des items sélectionnés, ainsi que la rédaction des consignes d'application, de correction, et de notation de l'épreuve.

⑦ *Epreuve de la version provisoire* du test sur un échantillon de la population, afin de déterminer la difficulté des items, la durée moyenne de passation, etc., et de vérifier la cohérence interne de l'épreuve et la valeur de la procédure de correction et de notation (p.ex., par comparaison de la note obtenue à chaque item et de la note globale).

⑧ *Validation* du test. Le problème est ici de vérifier si l'épreuve mesure bien ce pour quoi elle a été conçue. Si l'objectif principal du test est le pronostic (par exemple, pronostic de réussite scolaire ou professionnelle), il est en principe nécessaire d'attendre que les sujets examinés aient eu le temps de faire leurs preuves (cf.p.ex. [Conrad 1948 : 649-651]) ; à moins que l'on se contente, provisoirement, de valider le test sur une population de personnes ayant eu l'occasion de réussir ou d'échouer avant de subir le test. Si l'objectif est le diagnostic, il faut disposer d'indicateurs, distincts des items du test, de ce que le test est censé mesurer (observation de comportements, résultats obtenus à d'autres épreuves psychologiques, etc.). Dans tous les cas, ce que l'on recherche est une corrélation entre les notes au test (ou aux sous-tests) et les indicateurs ou les comportements liés à ce que le test est supposé mesurer.

⑨ *Mise au point de la version définitive* du test. Cet essai "en grandeur réelle" de la version provisoire du test conduit généralement à un remaniement de l'épreuve : sélection des items les plus pertinents, modification des règles d'administration et de correction, ré-ajustement de la notation (pondération différentielle des items) et du mode de calcul de la note globale, etc. Ce remaniement permet d'obtenir la corrélation maximale entre la note globale, et le trait, le comportement, l'aptitude, que le test doit évaluer ou pronostiquer.

⑩ *Étalonnage* du test. La version définitive est administrée à des échantillons importants de la population étudiée, afin de déterminer statistiquement la signification relative des notes obtenues, pour chaque strate considérée de la population à laquelle l'épreuve est destinée.

Cette description souffre des défauts inhérents à toute schématisation. Par exemple, il est clair que le problème de la validation de l'épreuve est présent à l'esprit de l'auteur du test dès la deuxième phase, et que cette préoccupation intervient chaque fois qu'il s'agit de sélectionner, de modifier, ou d'éliminer un item, ou encore de perfectionner le système de correction et de notation. D'autre part, la mise au point d'un test peut s'étaler sur de longues périodes d'essais, et c'est un peu abusivement que nous avons qualifié la version de la phase ⑨ de "définitive" : en pratique, les étapes ② à ⑧ peuvent être réitérées de nombreuses fois, éventuellement suivies de la phase ⑩. La condition d'arrêt dans le perfectionnement d'une épreuve étant une sorte de compromis, ou de situation d'équilibre entre les aspirations de l'auteur au progrès de l'outil, et la nécessité pour les utilisateurs de disposer d'un instrument définitif. Enfin, nous l'avons déjà signalé, la phase ⑩ doit être périodiquement réitérée pour tenir compte du vieillissement du test, et de l'évolution de la société.

Il est toutefois une objection plus grave à la validité de cette description : c'est l'existence de tests psychologiques dont la réalisation a suivi une toute autre voie. Par exemple, il serait facile de montrer que l'historique du "test des taches d'encre" d'Hermann RORSCHACH s'accorde mal avec la genèse d'une épreuve psychologique telle que nous l'avons présentée. Il semble en effet que l'élaboration de ce test ne soit pas le fruit d'une recherche systématique en vue de construire un test de personnalité, mais bien plutôt le résultat d'un échec dans la mise au point d'un test d'imagination (cf. [Anzieu 1960 : 16-24]). La démarche du chercheur se trouve par conséquent ici complètement renversée par rapport à celle décrite dans notre schéma. Pour RORSCHACH, le problème n'était plus de concevoir des situations expérimentales telles que les réponses des sujets puissent servir d'indicateurs des traits ou des aptitudes qu'il désirait déceler (puisque sa tentative pour évaluer les capacités imaginatives avait échoué) ; il était au contraire de déterminer quels traits ou quelles aptitudes pouvaient bien révéler les différences dans les réponses recueillies pour un stimulus donné, dans une situation expérimentale donnée (cf. [Goodenough 1949 : 97 n1]). Par conséquent, les étapes ① à ③ de notre description n'ont pas

d'équivalent ici ; il faudrait leur substituer une phase d'observation et d'expérimentation sur les comportements suscités par un stimulus constant (et relativement peu structuré), suivie d'une phase d'analyse des liens entre ces comportements et chacun des traits ou aptitudes décelables chez les sujets par d'autres moyens. En outre, il n'est pas non plus certain que l'étape ⑩ ait un sens, encore que l'on puisse considérer que la détermination des "réponses banales" constitue une sorte d'étalonnage, puisque ces réponses, qui "représentent la participation [du sujet] à la manière de concevoir de la collectivité" [Rorschach 1921 : 226], sont définies en pratique comme "les interprétations que l'on retrouve dans 1 protocole sur 3 environ" [Rorschach 1921 : 217]. Il convient toutefois de noter que cet exemple concerne la découverte d'un type d'épreuve radicalement nouveau ("interprétation libre de formes fortuites"), événement somme toute assez rare, et naturellement rebelle aux tentatives de classification et de systématisation. Remarquons d'ailleurs que la démarche suivie, tant pour la constitution par RORSCHACH de séries de planches parallèles aux planches originales (cf. [Rorschach 1921 : 43-44]), que pour l'élaboration d'autres tests projectifs à stimulus visuel (cf. p.ex. [Morgan 1938 : 531-534] pour l'historique du T.A.T.), obéit assez largement au schéma que nous avons proposé.

Il semble donc, à première vue, que la division du travail entre chercheurs et praticiens de la psychologie donne aux épreuves psychologiques un statut assez différent de celui du questionnaire d'enquête psychosociale. Nous allons, dans le paragraphe suivant, comparer systématiquement ces deux formes de questionnaire.

c) Questionnaire d'enquête et test psychologique.

Si nous comparons la réalisation d'une enquête psychosociale et la construction d'une épreuve psychologique, nous pouvons de prime abord déceler un certain parallélisme entre les deux démarches. Ainsi, les phases ② à ⑧ pour le questionnaire d'enquête, et les phases ② à ⑨ pour le test psychologique, correspondent à une recherche et une mise en forme des indicateurs

pertinents, et à leur présentation en une série ordonnée de questions, assortie de consignes d'administration. De même, la phase ⑨ pour l'enquête est assez proche de la phase ⑩ pour le test, à ceci près que l'enquête extensive proprement dite ne dure en principe que quelques jours, alors que l'administration extensive du test à des fins d'étalonnage peut être très étendue dans le temps (plusieurs années pour certaines épreuves individuelles). Si nous comparons d'autre part un questionnaire d'enquête et un test psychologique, nous ne relevons pas de différence importante dans leur présentation ou dans leurs modalités d'application : dans les deux cas, il s'agit d'une suite de questions ordonnée ; ces questions peuvent être ouvertes ou fermées, accompagnées ou non de présentation de matériel ; l'administration du questionnaire peut être collective ou individuelle ; la feuille de réponse peut être remplie par l'interviewé lui-même ou par la personne qui interroge ; etc.

Il existe cependant, entre ces deux formes de questionnaire, des différences notables, qui tiennent aux objectifs visés, et à la signification attachée aux réponses. Nous avons vu que l'objectif de l'enquête sociologique est en général l'explication d'un fait social, à partir de la description de la population étudiée, et des caractéristiques des individus qui la composent ; alors qu'un test psychologique sert à situer un individu dans une population donnée, relativement à certaines aptitudes ou certains traits de comportement. Cette divergence fondamentale explique que l'épreuve psychologique conserve une forme figée à travers toute la durée de son utilisation (ainsi, le BINET-SIMON et le RORSCHACH n'ont pas varié depuis plus d'un demi-siècle), alors que le questionnaire sociologique est le plus souvent conçu pour une enquête unique ; que les problèmes de validation soient cruciaux pour l'épreuve psychologique et fassent l'objet de travaux théoriques importants, alors qu'ils ne sont guère abordés que pour être résolus empiriquement dans l'enquête sociologique ; et enfin que les conditions d'utilisation des tests après leur mise au point définitive soient définies avec autant de précautions (standardisation minutieuse des conditions d'application, formation poussée des opérateurs, etc.). D'autre part, la signification attachée aux réponses dans les épreuves psychologiques paraît moins directe, moins transparente pour le sujet, que dans les questionnaires d'enquête. Ceci apparaît nettement dans les procédures de correction

ou de codification des réponses aux tests, qui portent souvent sur un contenu latent non perçu par le répondant, et dans le calcul d'indices synthétiques correspondant aux réponses à plusieurs questions, voire à l'ensemble du test ; alors que l'exploitation des réponses à un questionnaire d'enquête se limite souvent à la détection de relations de covariation ou de concomitance entre diverses formes de réponses manifestes. En outre, les tests de connaissances ou de niveau se caractérisent par l'existence de bonnes et de mauvaises réponses (ou de réponses vraies et de réponses fausses), alors que cette caractéristique semble exclue *a priori* des questionnaires sociologiques.

Il ne saurait être question de minimiser les différences relatives aux objectifs et aux conditions d'utilisation des épreuves psychologiques et des questionnaires sociologiques. Nous voudrions toutefois montrer que ces objectifs divergents ne sont pas incompatibles, puisqu'il arrive que des enquêtes sociologiques aient pour conséquence (sinon pour but) la mise au point d'une procédure permettant de situer un individu dans une population, et aussi que des épreuves psychologiques fassent l'objet de véritables enquêtes de type sociologique. Le premier cas est assez fréquent dans les études de marché qui visent à définir une ou plusieurs "cibles" de l'action promotionnelle en faveur d'un produit. Ces "cibles" sont des sous-populations qui, au travers de leurs réponses au questionnaire, sont jugées devoir être particulièrement réceptives à l'action envisagée ; pour des raisons pratiques, elles sont définies en termes de variables socio-démographiques (âge, sexe, revenu, etc.), permettant d'identifier sans difficulté les individus appartenant à une sous-population "cible" (Nous connaissons même des exemples d'enquêtes commerciales, qui ont été suivies de sessions de véritables groupes de travail rassemblant des responsables techniques de l'enquête et des agents commerciaux de la firme qui l'avait financée, afin d'élaborer en commun un ensemble de règles concrètes, servant à situer chaque client, actuel ou potentiel, dans une typologie déterminant la stratégie commerciale à adopter à son égard). Réciproquement, il existe des épreuves psychologiques qui ont fait l'objet d'enquêtes par sondage, évidemment distinctes dans leurs objectifs de la phase d'étalonnage (phase ⑩ de notre schéma). L'exemple français le plus connu est certai-

nement celui des enquêtes nationales sur le niveau intellectuel des enfants d'âge scolaire réalisées en 1944 et en 1965 (cf. [Ined 1950 ; 1954 ; 1969 ; 1973]). L'historique de ces deux enquêtes (tel qu'il est décrit par George HEUYER, Madame Henri PIERON, Raymond MANDE, et René GILLE pour la première [Ined 1950 : 13-39, 63-99], et par Alain GIRARD et Paul CLERC pour la seconde [Ined 1969 : 3-32]) s'insère tout à fait dans le schéma-type que nous avons proposé, à ceci près que les phases ③ à ⑥ ont évidemment été assez réduites, puisque l'arrière-plan théorique avait déjà été bien étudié au cours de multiples recherches antérieures [Ined 1950 : 41-62]. L'originalité de ces enquêtes réside dans l'utilisation d'un test de niveau, conjointement au questionnaire proprement dit. Dans l'enquête de 1944, il s'agit du "test mosaïque" de René GILLE, constitué d'épreuves papier-crayon parfois originales, le plus souvent empruntées à des tests de développement mental existants ; le questionnaire destiné à compléter le test était rempli par l'instituteur. Dans l'enquête de 1965, l'"échelle collective de niveau intellectuel" comprend des épreuves verbales et des épreuves non verbales, regroupées par niveau scolaire ; à ce test sont associés deux questionnaires destinés à être remplis, l'un par le maître, l'autre par les parents de l'enfant. Le plan de dépouillement des deux enquêtes était relativement simple, la note obtenue au test étant la variable à expliquer (variable dépendante), et les renseignements apportés par les réponses au(x) questionnaire(s) jouant le rôle de variables explicatives (variables indépendantes).

Si l'on considère les autres traits par lesquels les questionnaires d'enquête sociologique et les épreuves psychologiques semblent s'opposer, on constate de même que les différences relevées ne sont guère, lorsqu'elles existent, que des différences de degré. Cette remarque vaut en particulier pour :

- la *standardisation* des questions. Il est vrai que les items des épreuves psychologiques sont toujours standardisés. Mais d'une part, il existe des enquêtes périodiques pour lesquelles le questionnaire est fixé définitivement jusque dans ses détails dès la première administration ; d'autre part, bien que la majeure partie des questionnaires sociologiques soient élaborés "au coup par coup", en fonction des objectifs spécifiques d'une enquête

donnée, les sociologues spécialistes de l'enquête s'efforcent de mettre au point des formulations standards pour les questions les plus souvent utilisées (variables socio-économiques ; comportements de consommation et d'achat, de lecture des journaux et périodiques, de pratique religieuse ; échelles d'attitudes ou d'opinion politiques ; etc.).

- la *validation* des questions. Pour le psychologue, il s'agit surtout de perfectionner le test afin d'accroître sa valeur pour le diagnostic ou le pronostic individuels. Pour le sociologue, le problème est non seulement d'affiner les questions afin d'obtenir des réponses plus précises et plus circonstanciées, mais aussi d'identifier les sources de biais dans les réponses (cf. p.ex. l'impressionnante énumération de [Sudman 1974 : 21-23]), afin soit de les réduire, soit d'en corriger les effets ; c'est d'ailleurs souvent ce dernier souci qui prévaut, et tout professionnel de l'enquête sociologique en France sait de quel coefficient il faut multiplier le montant des ressources déclarées par les ménages, ou le kilométrage que les automobilistes disent avoir parcouru dans l'année, pour obtenir les valeurs réelles qui leur correspondent (en moyenne, pour une sous-population donnée). La différence dans les méthodes de validation dépend surtout de la multiplicité des facteurs qui interviennent dans la situation d'enquête, beaucoup moins standardisée que la situation d'examen psychologique. Signalons d'ailleurs que le problème des biais dans les réponses existe aussi pour le psychologue. Par exemple, dans le test de RORSCHACH, la fatigue ou la réticence à passer le test modifient le rapport entre le nombre de réponses kinesthésie et le nombre de réponses couleur (cf. [Anzieu 1960 : 54]) ; la variation dépressive de l'humeur diminue le nombre des réponses globales et réduit le nombre des réponses kinesthésie, alors qu'une humeur euphorique augmente le nombre de réponses globales et de réponses originales [Rorschach 1921 : 67-68].

- la *formation* des enquêteurs et des utilisateurs de tests. En règle générale, les psychologues utilisant des épreuves psychologiques standardisées ont reçu une formation beaucoup plus complète que les enquêteurs chargés d'administrer des questionnaires standardisés. Cela se conçoit en raison de l'objectif même des tests, qui est de permettre de formuler un jugement sur

un individu. Il est donc nécessaire que le psychologue soit capable de maîtriser la situation d'examen, et d'interpréter correctement non seulement les réponses du sujet, mais aussi l'ensemble de son comportement ; alors que l'enquêteur n'a, en principe, qu'à recueillir les réponses à des questions le plus souvent fermées ou pré-codées. Pourtant, sans toutefois être comparable à celle du psychologue qui réalise des entretiens en profondeur au cours de la phase exploratoire d'une recherche, la formation de l'enquêteur par sondage peut être poussée très loin : par exemple, au *Survey Research Center* de l'Université du Michigan, où les enquêteurs reçoivent des consignes générales extrêmement détaillées concernant la manière de remplir un questionnaire (cf. [Src 1976]), une technique de contrôle a été mise au point, qui permet, à partir de l'analyse du contenu d'enregistrements de l'enquêteur sur le terrain, de porter un jugement non seulement sur la pertinence ou la neutralité (non directivité) des propos tenus et sur la fidélité aux formulations des questions ou des relances, mais également sur l'intonation adoptée, la vitesse du débit, etc. (cf. [Cannell 1975]). D'autre part, même une formation très poussée du psychologue ou de l'enquêteur ne garantit pas la disparition des biais dans les réponses liés à la personne de l'interviewer. Hermann RORSCHACH avait déjà noté que, pour la détermination des réponses kinesthésie, "une disposition kinesthésique trop forte ou trop faible chez l'examineur peut intervenir d'une façon gênante" [Rorschach 1921 : 15]. Divers auteurs ont exploré systématiquement les variations, dans les réponses au test des taches d'encre, qui paraissent dépendre des relations entre le sujet et l'examineur (cf. la revue de ces travaux dans [Miller 1968 : 478-481]) ; ainsi, dans le compte-rendu d'une expérience concernant l'influence des conditions d'administration du test de RORSCHACH, Edith LORD arrive à la conclusion que c'est la personnalité du psychologue qui constitue la source la plus importante des variations dans les réponses des sujets (après les différences de personnalité des sujets ; cf. [Lord 1950]).

- la *transparence* des questions. Il est vrai que l'exemple des questionnaires de personnalité, et surtout des tests projectifs, illustre particulièrement bien la non transparence pour le sujet de certaines épreuves psychologiques. Ainsi, pour la première planche du RORSCHACH, seront codées de la

même manière (D K+ H) des réponses aussi diverses que : "Deux sportives", "Dieu avec des ailes", "Une femme effrayée", "Une danseuse", "Deux St Nicolas avec des balais sous le bras", "Forme féminine avec les jambes étroitement serrées", "Figure diabolique les mains pendantes" [Rorschach 1921 : 136-193, *passim*] ; toutes ces interprétations de la partie centrale de la première planche, tenue droite (non renversée), sont en effet fondées sur un grand détail (D), dans lequel le sujet voit une figure humaine en entier (H), à laquelle il prête un mouvement (K) qui est perçu assez fréquemment (+) par les sujets jouissant de leur intégrité mentale (cf. [Rorschach 1921 : 10-39], [Anzieu 1960 : 30-62]). En outre, la connaissance de cette codification n'apporterait en elle-même au sujet aucune information sur l'interprétation qui sera faite de ses réponses. Ces remarques peuvent s'appliquer, avec quelques correctifs, aux autres techniques projectives, et même aux questionnaires de personnalité : s'il est vrai que le contenu de réponses proposées au répondant, dans un questionnaire de type BERGER ou CATTELL, est directement perceptible par celui-ci (d'autant que de nombreuses questions font appel à l'introspection), la manière dont le questionnaire sera dépouillé et les informations qui en seront extraites ne peuvent être devinées par le sujet (qui trouve ses réponses "naturelles", c'est-à-dire sans signification différentielle). La non transparence du questionnaire pour le répondant est donc liée surtout à l'usage que le psychologue fait de ses réponses (un test de RORSCHACH présenté sous forme de questions fermées [Eysenck 1947 : 241-242] reste opaque pour le sujet). Par contre, les tests de connaissances ou d'efficacité sont en principe toujours transparents, puisque la signification, pour le psychologue, des réponses aux questions, est généralement perçue correctement par le répondant, et que l'on a de plus pris soin d'annoncer au début de l'épreuve quel était l'objectif réel du test. De prime abord, il semble que les questionnaires sociologiques soient toujours transparents pour la personne interrogée : il y a souvent des questions purement factuelles, et même les questions d'opinion ont, pour la plupart des répondants, un sens immédiatement perceptible. Il est clair pourtant que les personnes interrogées n'ont aucune idée, non seulement de la manière dont leurs réponses seront exploitées (exemple des échelles d'opinion ou d'attitudes extrayant un contenu sous-jacent aux contenus manifestes), mais surtout des informations que le

sociologue tirera des relations statistiques observées entre les réponses à deux questions ou plus (présomptions de liens de causalité, utilisables pour "expliquer" le phénomène étudié d'une manière littéralement inconcevable pour la grande majorité des répondants).

- *l'utilisation d'indices numériques*. Pour pratiquement toutes les épreuves psychologiques, les résultats sont exprimés à l'aide de notes ou de coefficients servant à situer le sujet au sein d'une population de référence. Il est vrai que le calcul d'une note individuelle est proportionnellement moins fréquent en sociologie. Toutefois, ce cas se présente chaque fois que le sociologue essaye de construire une échelle d'opinion ou d'attitudes, voire une "échelle de standing" ; en outre, on trouve, dans la littérature sociologique, beaucoup d'autres exemples d'indices numériques par unité de sondage, comme les "indices d'éminence" des universitaires, calculés par Paul LAZARFELD et Wagner THIELENS [Lazarsfeld 1958b], ou le coefficient d'immunité ou de préservation relativement au suicide proposé par J. BERTILLON fils (cf. [Durkheim 1897 : 175]), ou encore l'indice de distance globale entre conjoints proposé par Alain GIRARD pour caractériser chacun des couples mariés interrogés [Girard 1964 : 87-94]. En revanche, l'usage d'indices globaux paraît assez répandu chez les sociologues : indices de coût de la vie, indices de mobilité sociale, etc. (cf. sur l'usage des indices en sociologie : [Hagood 1941 : 216-244], [Zeisel 1968 : 76-102]).

- *l'existence de "bonnes réponses" et de "mauvaises réponses"* (ou de réponses vraies et de réponses fausses). Ceci semble caractériser les tests de connaissances ou de niveau, et ne se rencontrer ni dans les questionnaires de personnalité, ni dans les questionnaires d'enquête sociologique. Il est exact que la notion de vérité de la réponse est fondamentale pour l'évaluation des performances des sujets, encore que ces derniers ne soient jamais informés, en cours de passation, de la valeur de leurs réponses. Mais il arrive toutefois que soient glissées, dans les questionnaires de personnalité, des questions de contrôle destinées à estimer la sincérité du sujet (cf. p. ex. les échelles L et K du M.M.P.I.) ; et d'autre part, on peut rencontrer dans certains questionnaires d'enquête soit des questions de connaissances (du type : "Quel est, à votre avis, le prix d'un litre de ce

produit ?", ou : "Quel est le nom de votre député ?" cf. p.ex. [Hyman 1972 : 200-204]), soit même des questions-pièges, portant sur des objets ou des produits qui n'existent pas.

- l'*homogénéité* des questions. Les questionnaires de personnalité sont en général remarquablement homogènes : toutes les questions sont fermées, et le sujet doit choisir entre deux (ou trois) modalités de réponses. Les questionnaires d'enquête, par contre, sont extrêmement hétérogènes : on y trouve à la fois des questions ouvertes, des questions précodées, et des questions fermées ; des questions à réponse unique et des questions à réponses multiples ; des questions d'opinion et des questions factuelles ; des questions posées à tous, et des questions destinées à certains interviewés seulement ; etc. Cette hétérogénéité est l'une des caractéristiques du questionnaire d'enquête. Mais elle n'est guère moindre dans les échelles composites d'intelligence du type BINET-SIMON, TERMAN, ou WECHSLER.

A l'issue de cette mise en parallèle des questionnaires ou des tests psychologiques, et des questionnaires d'enquête psycho-sociale, il apparaît donc que pratiquement aucune des différences relevées n'appartient en propre à l'un de ces deux types de questionnaire, et que "le questionnaire peut être assimilé à un test auquel tous les sujets de la population sont soumis dans les mêmes conditions" [Daval 1963 : I 190]. On peut même affirmer que, formellement, deux épreuves servant à "mesurer" la même chose peuvent être plus différentes l'une de l'autre que deux épreuves sans objectif commun : par exemple, le test original de RORSCHACH est un test individuel à questions ouvertes auxquelles le sujet répond oralement ; le test Z , qui s'en inspire, est une épreuve collective à questions ouvertes auxquelles le sujet répond par écrit [Zulliger 1953 : 19-22] ; l'adaptation du RORSCHACH par Mollie R. HARROWER-ERICKSON est une épreuve collective à questions fermées, dans laquelle le sujet sélectionne les réponses correspondant à sa propre conception (questions à réponses multiples ; cf. [Eysenck 1947 : 238], [Wittenborn 1949 : 150]) ; l'adaptation par H.J. EYSENCK de cette dernière forme est une épreuve collective à questions fermées, dans laquelle le sujet ordonne les réponses proposées

selon qu'elles lui paraissent convenir plus ou moins à l'image qui lui est présentée.

Aussi considèrerons-nous que les différences de degré que l'on décèle souvent entre les tests et les questionnaires ne constituent jamais des différences de nature. Tout au plus traduisent-elles soit des disparités dans les objectifs ou les contextes d'utilisation, soit l'avance technique de certaines branches de la psychologie sur les autres sciences de l'homme. C'est pourquoi désormais nous appellerons indifféremment *questionnaire* toute série de questions standardisées, quel que soit son domaine d'application.

d) La place du questionnaire dans la recherche

S'il est vrai que "la Science porte sur la quantité et s'exerce au moyen de la mesure" (BAUER, cité par [Ullmo 1969 : 23]), la psychologie et la sociologie sont encore loin d'avoir acquis le statut de sciences : toutes deux manquent en effet d'unités de mesure [Piaget 1970 : 71, 78] (comme d'ailleurs la plupart des sciences de l'homme), "soit que l'on n'ait pas encore réussi à les constituer, soit que les structures en jeu, tout en pouvant fort bien être de nature logico-mathématique (algébrique, ordinale, topologique, probabiliste, etc.), ne présentent pas de caractères proprement numériques" [Piaget 1970 : 69]. Par contre, dans la mesure où la psychologie et la sociologie font partie d' "un groupe de disciplines ayant pour objet les activités de l'homme et se donnant pour but la recherche de « lois » [...]", et dont les méthodes "consistent soit en observations systématiques ou en expérimentations exprimables toutes deux en termes statistiques, soit en déductions mais réglées ou réglables par des algorithmes rigoureux (mathématiques ou logiques), soit en combinaison de la déduction et de l'expérience" [Piaget 1967 : 1116], elles sont parmi les sciences de l'homme les plus proches du "modèle" que constituent pour les philosophes (à tort ou à raison) les sciences physico-mathématiques. C'est pourquoi les réflexions concernant l'épistémologie des sciences physiques nous offrent un cadre commode pour décrire la place du questionnaire dans la démarche scientifique.

Actuellement, les spécialistes de l'épistémologie admettent généralement que "la science recherche ses objets, elle les construit, elle les élabore, elle ne les trouve pas « tout faits », tout donnés dans la perception ou l'expérience immédiate" [Ullmo 1969 : 23]. Certes, "si toutes nos connaissances commencent *avec* l'expérience, il n'en résulte pas qu'elles dérivent toutes *de* l'expérience" [Kant 1781 : I 37] ; et le contenu de toute sensation ne peut être perçu qu'à travers les formes *a priori* de l'espace et du temps [Kant 1781 : I 61-63] (même si l'espace et le temps sont d'une certaine manière construits eux aussi ; cf. [Brunschvicg 1922 : 414]). Mais en outre, la sensation ne peut être pensée qu'à l'aide des mots et des concepts issus de nos expériences et de nos connaissances antérieures. A l'origine de nos connaissances scientifiques, il y a l'observation de relations répétables entre des phénomènes ; de la synthèse de relations de ce type résultent les "lois" scientifiques, qui sont "l'expression [...] de la validité permanente escomptée d'une relation répétable constatée dans les phénomènes naturels" [Ullmo 1969 : 53]. "Nous appelons *objet* une partie des phénomènes observables qui donne lieu à la manifestation, dans des relations répétables, de la valeur constante d'un certain nombre de paramètres" ; "on peut dire aussi qu'il n'est qu'une *intersection de relations répétables*" [Ullmo 1969 : 49, 56]. Les lois scientifiques suggèrent des structures, c'est-à-dire des ensembles de rapports, considérés comme caractéristiques, entre éléments jusqu'à un certain point quelconques ; "ces structures *représentent* les objets scientifiques nés des lois, c'est-à-dire qu'elles en donnent une *description* en termes de rapports constants" [Ullmo 1969 : 102], et l'on peut ainsi considérer qu'un objet scientifique est "une théorie cristallisée" [Ullmo 1969 : 97]. La théorie étant soumise à vérification par l'observation ou l'expérimentation (les lois sont des "instruments de recherche et d'exploration de la nature" [Ullmo 1969 : 55]), l'échec de ses prévisions peut conduire à la définition d'objets scientifiques nouveaux.

Si l'on essaye de schématiser et de classer les questions qu'un chercheur peut se poser lorsqu'il recourt à l'observation ou à l'expérimentation, on peut distinguer quatre formes principales d'interrogation, selon qu'elles ont pour objectif :

- la validation d'une théorie ("Les conséquences prévisibles et non encore observées de la théorie se produisent-elles, sont-elles observables ?") ;

- l'extension ou la généralisation d'une théorie ("Lorsque changent les conditions dans lesquelles se produisent les phénomènes dont l'observation est à l'origine de la théorie, celle-ci reste-t-elle toujours valable ?") ;

- le choix entre deux théories concurrentes ("Que se passe-t-il effectivement lorsque se trouve réalisée une situation dans laquelle les prévisions de chacune des théories ne concordent pas ?") ;

- la compréhension d'un phénomène inattendu ("Quelle théorie nouvelle ou remaniée permettrait de rendre compte d'une observation apparemment non déductible des théories actuellement admises ?").

Les deux premières formes constituent si l'on peut dire les questions de routine que tout chercheur est conduit à se poser dans le cadre de ses activités. La troisième correspond à l' "expérience cruciale" ("*instantia crucis*"), dans laquelle le chercheur hésite entre plusieurs explications possibles [Bacon 1620 : II § 36]. L'histoire des sciences a retenu un grand nombre d'expériences cruciales, comme celle de Florin PÉRIER (qui a permis de choisir entre l' "horreur du vide" et l' "équilibre des liqueurs"), ou celle de Léon FOUCAULT (qui a justifié le rejet de la théorie de l'émission de la lumière, et l'adoption de la théorie ondulatoire). Remarquons cependant que, dans l'esprit de François BACON, l'*instantia crucis* ne nécessitait pas toujours l'expérimentation proprement dite, mais pouvait également se fonder sur l'observation : le premier exemple qu'il en donne est d'ailleurs celui de l'explication des marées, où l'expérimentation n'est évidemment pas possible. Remarquons en outre que le mécanisme et la genèse d'une expérience cruciale sont moins simples que la description qu'en donne BACON [Brunschvicg 1922 : 428-431]. La quatrième forme est bien illustrée par les recherches qu'ont suscitées les résultats imprévus de l'expérience d'Albert MICHELSON, sur la vitesse de propagation de la lumière selon deux directions perpendiculaires. Mais

il arrive souvent que cette forme présente toutes les apparences de la découverte fortuite : c'est Claude BERNARD découvrant qu' "à jeun tous les animaux se nourrissent de viande" à partir de l'observation que l'urine des lapins que l'on venait de lui apporter était claire et acide, au lieu d'être trouble et alcaline [Bernard 1865 : 216-218] ; ou encore Henri BECQUEREL constatant "par hasard" la radio-activité d'un sel d'uranium. En réalité, la découverte naît ici de la contradiction perçue entre une observation, et les conséquences des connaissances et des théories scientifiques du chercheur (c'est "Iris fille de Thaumas", selon le mot de PLATON, *Le Théétète*, 155d) ; elle n'aurait donc pas été possible sans l'acceptation préalable (au moins de façon implicite), par l'observateur, de la théorie que l'observation paraît infirmer.

Les formes de questions que nous venons de décrire débouchent naturellement sur un plan d'observation ou d'expérience, mais sans toutefois se confondre avec celui-ci. Sous une forme plus élaborée, issue de la décomposition du problème initial en questions plus spécifiques et plus précises, l'interrogation du chercheur constitue ce que nous avons appelé un *questionnaire théorique*. C'est ce type de "questionnaire" qui motive la recherche et en définit la problématique ; c'est lui également qui détermine l'exploitation qui sera faite des résultats de l'observation ou de l'expérimentation (au moins dans un premier temps). Le questionnaire théorique ainsi défini évolue évidemment selon les progrès de la réflexion du chercheur ; il se matérialise en particulier dans les projets de recherches. Le passage du projet à l'acte impose un pas de plus : la conception et la réalisation d'un appareillage d'observation ou d'expérimentation ("L'expérimentation, et même la simple observation, sont des théories en acte" [Ullmo 1969 : 97]). Pour les sciences physiques, David Mac KAY a proposé une description de la procédure expérimentale en deux étapes [Mac Kay 1950 : 291-292] :

① Mise au point d'un appareil d'expérimentation, et d'un système de classification des phénomènes que l'on se propose d'observer, tel que l'on dispose d'un nombre suffisant de catégories indépendantes pour décrire les résultats de l'expérience. Par exemple, si l'on envisage d'enregistrer des fluctuations dont la fréquence varie de une à cent par seconde, l'appareil devra avoir un temps de réponse de l'ordre du 1/100 de seconde

(c'est-à-dire qu'il devra pouvoir fournir cent informations dépendantes par seconde).

② Réalisation de l'expérience, en utilisant l'appareil pour classer les événements produits ou observés, dans la zone choisie de l'espace - temps (déterminée par le "questionnaire théorique"). Le résultat de l'expérience se traduit par l'association d'un nombre (mesure d'une quantité, par exemple) avec chacune des catégories définies par la structure de l'expérience (appareillage et système de classification des phénomènes).

L'intérêt du schéma proposé par Mac KAY est double : il peut sans grande modification être appliqué aux procédures de recherches dans les sciences de l'homme [Grémy 1970a : 9-10], et il fournit une base théorique pour la description des techniques d'analyse des données ; nous examinerons ce deuxième point au chapitre 3. Le questionnaire d'enquête ou le test psychologique peuvent être assimilés à un dispositif expérimental au sens de Mac KAY dans la mesure où ils permettent de classer et de dénombrer des observations. Le passage du "questionnaire théorique" au *questionnaire empirique* (questionnaire d'enquête ou épreuve psychologique standardisée) obéit d'ailleurs aux mêmes règles dans les sciences humaines que dans les sciences physiques. Dans ces dernières en effet, le passage de la théorie à l'expérimentation s'effectue par la définition opératoire du concept : "Une définition opératoire est une définition qui comporte la description d'un procédé régulier pour repérer, mesurer, plus généralement *atteindre* et identifier le concept défini [...]. Toute notion ainsi introduite recouvre une expérience virtuelle" [Ullmo 1969 : 24-25]. Or, Paul LAZARFELD a décrit une démarche analogue pour traduire les concepts abstraits des sciences sociales en indices empiriques, qui en sont les équivalents opératoires [Lazarsfeld 1958a : 100-105] ; la sélection d'une batterie d'indicateurs (c'est-à-dire la construction d'un questionnaire empirique) étant ici comparable à la réalisation d'un dispositif expérimental. Il est d'ailleurs possible de pousser plus loin encore le parallèle entre sciences physiques et sciences de l'homme, puisque certains traits que l'on considère souvent comme spécifiques des sciences humaines se rencontrent également dans les sciences physiques ; comme par

exemple, le fait que les "données" ne sont pas données mais fabriquées [Brunschvicg 1922 : 430-431], ou que toute observation soumet nécessairement l'objet observé à une interaction avec un autre système, et par conséquent le modifie [Ullmo 1969 : 117-118].

En conclusion, le questionnaire empirique apparaît comme l'instrument d'une procédure standardisée de recueil d'informations destinées à fournir des éléments de réponse aux interrogations qui ont présidé à sa conception (questionnaire théorique). Cette procédure a pour caractère spécifique de consister à poser à des informateurs sélectionnés une série de questions prévue à l'avance et identique pour tous, et à classer les réponses obtenues dans des catégories également identiques pour tous (codage). L'objet des questions peut être l'informateur lui-même, ses caractéristiques, ses goûts, ses aptitudes, ses opinions, etc. ; ce peut être également une tierce personne, un groupe social, un organisme, une institution, une entité géographique, etc. La procédure est par construction destinée à être répétée auprès d'individus différents, ou à des moments différents. L'administration d'un questionnaire empirique correspond, dans le cycle de la recherche (théorie/observation/analyse/théorie), à la phase d'observation [Flament 1976 : 44-45] ; mais l'observation ne se limite pas nécessairement à l'administration d'un questionnaire [Loubet del Bayle 1978 : 34-35].

0.2. L'ANALYSE SECONDAIRE ET LES BANQUES DE DONNÉES

Une banque de données est un "ensemble d'informations archivées dans des mémoires accessibles à l'ordinateur en vue de permettre le traitement des diverses applications prévues pour elles" (cité par [Chaumier 1976 : 5]). Une *donnée* est une notion ou un fait, représenté sous une forme conventionnelle convenant à la communication, l'interprétation, ou le traitement ; elle se compose de deux éléments : une *rubrique* qui identifie la nature de l'information, et une *valeur* définie dans un système d'unités ou de codification déterminé [Chaumier 1976 : 7-8]. Les données sont organisées selon une

structure de fichier dépendant de leur nature et des informations dont elles sont le support. Selon la nature des données archivées et le type d'information qu'elles contiennent, on peut distinguer (cf. [Durieux 1975 : 28-31], [Chaumier 1976 : 22-23]) : les banques de données bibliographiques, qui renseignent l'utilisateur sur les sources possibles de l'information qu'il recherche (exemple : le Centre de Documentation Sciences Humaines du CNRS) ; les banques de données textuelles, qui fournissent à l'utilisateur le document en langage naturel dont il a besoin, et effectuent éventuellement des traitements lexicographiques ou proprement linguistiques sur ce document (exemple : le Trésor de la Langue Française) ; et enfin les banques de données numériques, dans lesquelles l'information est soit le résultat de mesures ou de dénombrements, soit codée sous forme numérique (exemple : toute collection de résultats d'enquête par sondage ou de recensement).

Dans les sciences de l'homme, ces trois types de banques de données se rencontrent également ; la répartition en est toutefois très différente selon les disciplines, puisque l'histoire, la sociologie et l'économie utilisent surtout des données numériques, alors que l'archéologie, le droit, et la linguistique traitent plutôt des données textuelles [Durieux 1975 : 31]. Nous décrivons ci-dessous les principales caractéristiques des banques de données numériques qu'utilisent les sciences sociales, et les grandes lignes de la procédure d'analyse secondaire de ces données ; nous concluons sur la place du questionnaire dans les banques de données numériques.

a) Les banques de données pour les sciences sociales

Karl DEUTSCH rattache la première utilisation des banques de données sociales aux tous débuts des sciences de la société : "Ceci remonte à HÉRODOTE et à ARISTOTE, qui fit procéder, par ses assistants de recherches, à la comparaison des constitutions de 120 cités autonomes" [Deutsch 1970 : 19]. Dans le tableau de l'évolution des besoins et des ressources en données sociales qu'il dresse ensuite, il distingue quatre étapes ("the four data revolutions") : la première est marquée par le souci d'accumuler des informations statistiques à des fins administratives (impôts, conscription) dans la seconde moitié du

XVIIe siècle ; la seconde étape correspond à l'émergence, vers 1840, des préoccupations touchant à l'évolution et au devenir des sociétés (Herbert SPENCER, Karl MARX), qui conduit ainsi à la recherche et l'accumulation de données historiques ; la troisième étape voit l'apparition, à partir de 1935, de nouvelles méthodes de collecte et de traitement des données ; la quatrième étape enfin est celle de la constitution, à partir de 1960 environ, de grandes banques de données complexes (c'est-à-dire correspondant à des thèmes, des populations, des zones géographiques, etc. très variés) [Deutsch 1970 : 20-23]. Cette quatrième étape est contenue en germe dans le rapport, rédigé en 1957 pour l'Université de Columbia, par LUCCI, ROKKAN, et MEYERHOFF, sur les moyens de sauvegarder les données d'enquêtes réalisées aux États-Unis et en Europe de l'Ouest (cf. [Heim 1977]).

Aux États-Unis, la création de la première banque de données en sciences sociales (Le Roper Public Opinion Research Center) remonte à 1946 ; mais il faudra attendre 1969 pour que les Archives Nationales américaines créent une Division consacrée aux données lisibles par ordinateur. En France, de nombreux organismes publics (INSEE, INED, etc.) et privés (IFOP) ont pris l'habitude de conserver, au moins pendant un certain temps, les données qu'ils avaient recueillies ; mais il ne semble pas que l'on y ait défini, et encore moins mis en oeuvre, une véritable politique d'archivage (sélection et conservation des données à l'usage des chercheurs). Toutefois, depuis quelques années, le souci de conservation des données sociales semble s'être développé, et avoir touché des organismes tels que l'INSEE ou les Archives Nationales ; il s'est traduit, en outre, par la création de banques de données liées assez étroitement à la recherche universitaire (CEVIPOF, CERAT, CES, GEMAS, CRH, etc.). Enfin, assez récemment, les pouvoirs publics ont suscité des groupes de travail et de réflexion touchant à ces problèmes (ainsi que nous l'avons signalé au début de cette introduction).

Pour les chercheurs en sciences sociales, le phénomène nouveau des années soixante est "l'explosion des données" [Nasatir 1973 : 7] : les progrès parallèles de l'informatique et des techniques d'enquête par sondage, joints à l'augmentation de la demande en informations socio-économiques des "décideurs" publics et privés, ont entraîné un accroissement accéléré du

volume de données recueillies annuellement, en particulier en matière de consommation et d'opinion publique (cf. la progression du volume des données nécessaires à la recherche en sciences politiques, dans [Deutsch 1970 : 30]) ; "le résultat a été un torrent de données [...] qui, dans de nombreux cas, a paralysé et immobilisé les chercheurs et les décideurs, plus qu'il ne les a aidés dans leur quête de savoir" [Nasatir 1973 : 8]. L'expansion rapide et anarchique de ces "galaxies de données" [Ennis 1964 : 5], hétérogènes dans leur contenu et disséminées à travers l'espace des lieux et des institutions, a entraîné un accroissement de l'effet des "facteurs datacides" (selon le mot d'Allen BARTON). Aux causes naturelles de destruction ou d'altération des supports physiques de l'information s'ajoutait déjà le coût du stockage des documents, entraînant la destruction massive (par les organismes d'enquêtes en particulier) de volumes importants de données irremplaçables ; dans les cas de destruction sélective, les critères de choix des données à sauvegarder s'avéraient souvent assez différents de ce qu'auraient souhaité les chercheurs en sciences sociales (historiens, politologues, sociologues). L'informatisation des procédures d'archivage a, paradoxalement, engendré un nouveau facteur de destruction d'informations : celui-ci se manifeste par exemple dans la conservation de données, à l'origine "qualitatives" (réponses à des questions ouvertes par exemple), sous une forme codée, qui non seulement sélectionne une faible partie de l'information contenue dans les réponses en langage naturel, mais surtout la fixe dans un moule conceptuel qui dépend d'un schéma de pensée par essence temporaire (lié à des conditions sociales et historiques particulières ; cf. par exemple l'inadéquation de notre forme de pensée à celle des chercheurs scientifiques de la Renaissance, pourtant historiquement relativement proches). Enfin, moins grave à long terme que la destruction d'informations, mais analogue dans ses effets immédiats, il faut citer l'insuffisance des outils intellectuels d'exploitation de grands volumes de données hétérogènes, et les lacunes de la formation actuelle des utilisateurs potentiels.

Ces conséquences prévisibles de "l'explosion des données" avaient été décrites par Philip H. ENNIS dès 1964 [Ennis 1964 : 11-12] ; la définition d'une politique cohérente d'archivage des données en sciences sociales répond au souci de réduire autant que possible ces effets négatifs. C'est pourquoi toute banque de données en sciences sociales doit en principe viser trois objectifs : la collecte des données, leur conservation, et leur mise à la disposition des utilisateurs.

La *collecte* des données archivables dépend des sources de ces données. En France, par exemple, les principaux producteurs de données sont les grands instituts de recherches économiques ou démographiques dépendant de l'Etat (INSEE, INED, CREDOC, CEREP, SEDES, etc.) ; les sociétés de sondage privées (IFOP-ETMAR, SOFRES, SECED, etc.) ; et enfin les laboratoires rattachés aux Universités ou au CNRS. A ces sources en quelque sorte officielles s'ajoutent les chercheurs isolés, qui réalisent (par exemple dans le cadre d'une recherche de troisième cycle) des enquêtes pouvant porter sur plus de quatre cents individus ; les apports qui résultent d'échanges internationaux ; les données historiques (recension du contenu d'archives, ou données numériques existant sous forme de tableaux imprimés). En raison du volume global des données éventuellement disponibles, toute banque de données doit nécessairement se spécialiser dans une discipline, un thème de recherche, une région, une période historique, un milieu social ; et ce d'autant plus que ses moyens sont plus limités. D'ailleurs, les banques existantes se sont en général créées autour d'un laboratoire, pour répondre aux besoins spécifiques de celui-ci.

Les possibilités et les conditions d'obtention des données varient considérablement selon les sources : données payantes, données couvertes par le secret statistique ou le secret commercial, rétention des données par le chercheur qui les a recueillies tant qu'il n'a pas achevé ses publications, etc. L'absence d'un dépôt légal obligatoire et l'exigence de confidentialité de la part des fournisseurs de données compliquent évidemment la tâche de prospection des archivistes. D'un autre côté, ceux-ci doivent sélectionner les données à archiver en tenant compte non seulement de leur objet, mais

aussi de leur qualité (qui n'est pas toujours très élevée, si l'on en croit l'expérience du politologue Angus CAMPBELL, qui est également l'un des fondateurs de l'I.C.P.S.R. en 1962 [Campbell 1970 : 44-45]). C'est à partir de l'ensemble de ces facteurs, mais aussi de l'évaluation du coût de l'archivage et de la prévision du taux d'utilisation de chacun des fichiers (en fonction par exemple de l'originalité des informations qu'ils contiennent [Chaumier 1976 : 17] et de l'intérêt théorique de celles-ci), que la politique de collecte des données peut être définie sous la forme d'une hiérarchie des priorités (cf. [Nasatir 1973 : 37-39]).

La *conservation* des données acquises soulève plusieurs problèmes. En premier lieu, certaines données ne se présentent pas sous une forme lisible par l'ordinateur, et doivent par conséquent être reportées sur cartes mécanographiques ou sur bande magnétique (en l'état actuel du développement technologique) ; tel est le cas par exemple des résultats des recensements du siècle dernier (pour se limiter aux données numériques ; pratiquement, aucune donnée textuelle ne se présente sous une forme lisible par ordinateur, malgré les efforts entrepris en liaison avec les recherches effectuées dans l'imprimerie sur la composition assistée par ordinateur). Un problème voisin est celui des fichiers mécanographiques anciens, qui n'ont pas toujours été stockés dans de bonnes conditions de température et d'hygrométrie, et que certains lecteurs de cartes ne peuvent absorber sans dommage. En pratique, actuellement, tous les fichiers sont reportés sur bande magnétique ; il faudra donc assurer la protection des fichiers contre tout effacement imputable à une erreur de manipulation (verrouillage), et procéder périodiquement au recopiage des bandes [Nasatir 1973 : 41], pour prévenir la dégradation de l'information due aux effets de la démagnétisation.

A ces problèmes matériels s'ajoute l'obligation de nettoyer les données, de les présenter sous un format standard, de décrire leur contenu et leur origine, et enfin de classer les fichiers. Le nettoyage des données consiste d'abord à débarrasser celles-ci des erreurs qui ont pu s'y glisser : dans le cas de données individuelles (données de recensement ou d'enquêtes, déjà vérifiées en principe avant utilisation), ce sont surtout les codes sans signification ("wild codes") résultant d'erreurs de perforation ; dans

les données agrégées tirées de documents imprimés, il peut s'agir en outre d'erreurs imputables à la source elle-même (cf. p.ex. [Grémy 1978 : 38-40]). Ensuite, on fait généralement entrer dans la procédure de nettoyage la recodification des perforations non numériques (hors-textes, et signes divers), le remplacement des perforations multiples par des ensembles de perforations uniques, l'uniformisation des codes correspondant à "non-réponse", "ne sait pas ou sans opinion", "question non posée", etc. [Nasatir 1973 : 39, 49-50]. En dernier lieu, le nettoyage inclut le contrôle de la concordance des résultats obtenus sur le fichier dont on dispose, avec les résultats fournis ou publiés antérieurement par les chercheurs qui l'ont constitué ou utilisé. La présentation des données sous un format standard, identique pour tous les fichiers, répond à une exigence de certains systèmes d'analyse des données (les formats les plus courants étant l'image-carte, et les formats OSIRIS et SPSS). La description du contenu du fichier comprend d'une part le *dictionnaire*, qui fournit la signification des codes utilisés pour chaque variable (texte de la question et des réponses possibles, ou intitulé de la rubrique ; zone de l'enregistrement et codes numériques correspondants), et d'autre part les caractéristiques de la procédure suivie pour le recueil des données (population visée, technique d'échantillonnage, dates de l'enquête, etc. ; cf. [Nasatir 1973 : 88]). Enfin, le classement des fichiers impose la conception d'un système de classification, qui permette de retrouver facilement les données dont on a besoin ; le problème se pose ici dans pratiquement les mêmes termes que pour les bibliothèques ou les dépôts d'archives.

La *mise à la disposition des utilisateurs* recouvre des tâches extrêmement variées, depuis la diffusion d'informations sur les ressources de la banque de données et la fourniture de copies des fichiers, jusqu'à l'assistance technique à l'analyse des données et la formation méthodologique des utilisateurs. Parmi celles qui incombent logiquement à une banque de données, les deux plus importantes sont la recherche de l'information demandée, et la fourniture des documents correspondants (cf. [Bell 1970]). La première tâche des archivistes est donc la mise au point d'un système d'indexation des fichiers se référant à leur contenu, et d'une procédure permettant à l'utilisateur de déterminer dans quels fichiers et sous quelle forme se trouve

l'information qu'il recherche ; cette tâche ne se confond pas avec la conception du système de classement, puisqu'un même fichier peut recevoir plusieurs identificateurs (alors qu'il ne doit être classé qu'en un seul endroit), et que les critères de classement doivent tenir compte de considérations matérielles (nature du support physique, encombrement) ou historiques (source, date d'acquisition) autant ou plus que des besoins des utilisateurs. L'indexation peut se faire selon un schème pré-établi (comme le système de classification décimale utilisé dans certaines bibliothèques) ou par mots-clés ; au niveau du fichier ou au niveau des rubriques ou des questions ; à la main ou automatiquement (par utilisation du système KWIC par exemple ; cf. [Wilcox 1967 : 227-229]). La conception du système d'indexation est étroitement liée à celle d'un langage documentaire adapté aux besoins des utilisateurs (cf. sur ces problèmes [Phillipson 1968], [Nasatir 1973 : 41-42], [White 1977]). Lorsque les documents dont l'utilisateur a besoin ont été identifiés, il reste à les lui procurer. S'agissant de données lisibles sur ordinateur, le problème ne se pose pas dans les mêmes termes que pour des livres ou des documents d'archives : les documents fournis à l'utilisateur sont ici un exemplaire de la documentation décrivant les données, plus une copie des fichiers demandés. L'opération de copie de l'information autorise diverses opérations optionnelles : mise des données sous un autre format, extraction de sous-fichiers, voire recodification de variables, calcul d'indices, constitution d'un fichier de données agrégées à partir de données individuelles, etc. La nature des services de ce type rendus aux utilisateurs varie considérablement d'une organisation à l'autre.

Certaines banques de données ont dû prendre en charge diverses tâches incombant en principe aux universités ou aux laboratoires de recherche : formation des utilisateurs, fourniture de logiciel d'analyse des données, assistance technique pour l'utilisation de ce logiciel, formation d'archivistes de données en sciences sociales. Les banques de données sont en effet souvent sous-utilisées faute d'utilisateurs compétents. Pour pallier les lacunes de la formation dispensée dans les universités, de grandes banques de données ont donc organisé des écoles d'été d'initiation à l'analyse secon-

daire des données ; par exemple, l'I.C.P.S.R. (Ann Arbor) dispense une telle formation depuis 1963 (basée sur le système OSIRIS), et le S.S.R.C. Survey Archive (Colchester) depuis 1967 (basée sur le SPSS). Le développement et la multiplication de ces stages montre bien qu'ils correspondent à un besoin. En France, de tels stages sont régulièrement organisés par (entre autres) le LISH (Paris) et le CERAT (Grenoble).

Il n'est pas commode d'établir un bilan du développement des banques de données en sciences sociales. Certes, diverses listes des centres existants ont été établies, en particulier pour les pays occidentaux (cf. p.ex. : [Hyman 1972 : 331-333], [Nasatir 1973 : 85-87], [Gerhan 1975 : 31-44], et l'énumération des organismes qui collaborent à la revue *SS Data*) ; mais la comparaison de ces listes révèle de nombreuses discordances, ainsi que certaines lacunes. Il est toutefois certain que, dans la plupart des pays, des actions d'archivage sont entreprises : en font foi les échos répercutés par les associations et les comités spécialisés (I.A.S.S.I.S.T., I.F.D.A., C.E.S.S.D.A., etc.), ainsi que par le Conseil International des Sciences Sociales de l'UNESCO. Cependant, les obstacles à ce développement sont nombreux. Les deux obstacles majeurs nous paraissent être le coût de la conservation des données et de leur mise à la disposition des utilisateurs (on trouvera un aperçu du budget d'une grande banque de données dans [Grémy 1978 : 14-17]) ; et l'insuffisance de la formation technique et méthodologique des utilisateurs potentiels, qui les conduit à méconnaître l'intérêt de l'analyse secondaire (cf. [Hyman 1972 : 2]).

b) Analyse secondaire et théorie du questionnaire

L'analyse secondaire des données est un vocable nouveau pour désigner une pratique assez ancienne dans les sciences sociales. On peut en effet considérer que les travaux d'Adolphe QUÉTELET sur les facteurs expliquant la gravité des condamnations subies par les délinquants, ou ceux d'Émile

DURKHEIM sur les causes du suicide, constituent déjà des exemples d'analyse secondaire de données, puisque celles-ci avaient été recueillies antérieurement à leur recherche, et pour d'autres objectifs. L'émergence de cette terminologie après la seconde guerre mondiale (le mot ne figure pas dans le *Dictionary of Sociology* de Henri Pratt FAIRCHILD, publié en 1944) résulte apparemment de l'intérêt croissant pour la ré-analyse de données d'enquêtes, peu ou mal exploitées au moment de leur collecte ("analyse primaire") ; intérêt qui trouve sa source dans l'augmentation du volume des enquêtes par sondage, le développement des techniques de traitement statistique sur ordinateur, mais également dans le souci de mieux comprendre l'évolution politique dans le monde (d'où l'essor des études comparatives inter-culturelles, et des recherches sur le changement social).

Les avantages de l'analyse secondaire sur l'analyse primaire sont à la fois d'ordre économique et technique. Sur le plan économique, il est évident que le coût d'acquisition des données peut, pour le chercheur, être pratiquement négligeable : la copie du fichier correspondant à une enquête revient moins cher qu'une seule interview [Stewart 1967 : 13-14]. Sur le plan technique, les possibilités nouvelles offertes au chercheur en sciences sociales par l'analyse secondaire sont particulièrement intéressantes (cf. [Wilcox 1967 : 225], [Stewart 1967 : 15-16], [Hyman 1972 : 20-24], [Nasatir 1973 : 8], etc.) :

- amélioration de la précision des estimations, grâce à l'augmentation de la taille de l'échantillon correspondant à la population étudiée (par concaténation de fichiers) ;

- distinction, dans une population, de sous-groupes plus nombreux et plus variés que lorsque l'on travaille sur un échantillon restreint ;

- extraction d'un échantillon suffisamment important d'une sous-population de très faible effectif, à partir d'une série de sondages portant sur l'ensemble de la population ;

- meilleur contrôle des biais dus à la technique de l'enquête (possibilité d'identifier et d'annuler certains artéfacts) ;

- plus grande qualité du dépouillement des données, en raison du temps dont on dispose (à la fois pour essayer diverses techniques, et pour en comparer et en exploiter plus complètement les résultats), et de la possibilité de recourir à des méthodes de traitement statistique mises au point postérieurement à l'analyse primaire des données ;

- validation d'un modèle ou d'une théorie élaborée à partir d'un extrait des données, en comparant les conséquences déductibles de la théorie à d'autres éléments des données.

L'ensemble de ces possibilités permet d'expliquer le fait que l'analyse secondaire constitue la procédure privilégiée pour les recherches en histoire contemporaine, pour les études comparatives entre pays, et pour la mise au point et la validation de modèles du changement social [Hyman 1972 : 11-20]. En outre, l'analyse secondaire de fichiers peu volumineux, et de structure simple, peut permettre l'initiation des étudiants en sciences sociales à l'élaboration d'hypothèses opératoires et à leur vérification ; cet aspect pédagogique de l'analyse secondaire se développe depuis 1974 environ (cf. [Grémy 1978 : 18-34]).

Les problèmes techniques de l'analyse secondaire sont tout d'abord ceux de l'analyse des données en général : contrôle et redressement de l'échantillon, construction des variables, vérification et quantification des hypothèses. Mais il s'y ajoute deux problèmes spécifiques, l'un lié à la discordance entre les objectifs des diverses enquêtes dont on dispose et les objectifs propres à l'analyse secondaire entreprise, et l'autre à l'hétérogénéité des fichiers traités (cf. [Hyman 1972 : 25-31]).

Le fait de n'être pas l'auteur des questionnaires qui ont servi à recueillir les informations qu'il utilise représente, pour le chercheur, un handicap certain. Il lui faut en effet sélectionner, dans l'ensemble des données dont il dispose, les fichiers qui conviennent aux objectifs de sa recherche ; dans

chaque fichier, les questions ou les items pertinents ; pour chaque question, les réponses qui lui apportent l'information dont il a besoin. Ces sélections successives nécessitent une bonne connaissance non seulement de la langue et des habitudes de pensée des populations touchées par les enquêtes, mais aussi de l'impact sur ces populations des diverses formulations possibles d'une même question. Depuis les recherches expérimentales de MENEFEE, de BLANKENSHIP, de CANTRIL, de RUGG, etc. sur la formulation des questions (recherches dont nous reparlerons au chapitre 3), les sociologues connaissent l'existence d'influences, sur les taux de réponses, de légères modifications dans le libellé des questions ; mais les explications avancées par les chercheurs ne permettent que très imparfaitement de prévoir le sens et l'amplitude de ces différences. C'est pourquoi la connaissance, préalablement à l'analyse secondaire des données, de l'ensemble des informations dont a disposé l'auteur du questionnaire, est indispensable au chercheur ; et plus particulièrement, les matériaux recueillis au cours de l'enquête exploratoire et le résultat des réflexions qui en sont issues (phases ⑤ et ⑥ de notre description de la démarche-type du chercheur en sciences sociales). Ces informations sont irremplaçables ; et, en leur absence, les publications de l'auteur de la recherche originale ne semblent guère aider l'analyste sur ce point (cf. [Bell 1970 : 56-57]). A défaut de pouvoir entreprendre une nouvelle enquête exploratoire (qui, lorsqu'elle est possible, ne saurait de toute manière apporter les mêmes informations que l'enquête exploratoire originale), le chercheur doit donc se livrer à une analyse sémantique approfondie du questionnaire (libellé des questions, codification des réponses, ordre de présentation et enchaînement des questions, consignes aux enquêteurs, etc.), afin de reconstituer la problématique qui était celle de l'auteur du questionnaire au moment de la rédaction et de la mise au point de cet instrument de recueil des données, et d'en conjecturer la signification des réponses fournies par les personnes interrogées.

D'autre part, l'hétérogénéité des fichiers traités constitue un obstacle considérable à la fusion des fichiers, c'est-à-dire à leur concaténation pour constituer une base de données plus importante ou plus complète. Cet obstacle présente trois aspects distincts : un problème de traitement de fichiers complexes (à plusieurs niveaux d'information, et de formats

d'enregistrement variés ; cf. p.ex. [Miller 1970]), en principe résolu par certains logiciels ; un problème de compatibilité entre fichiers provenant d'organismes dont les fonctions, et les exigences en matière de qualité des données, sont très différentes (cf. p.ex. [Martinotti 1978]) ; et enfin un problème de recherche d'équivalences sémantiques entre des questions dont les formulations ne sont pas rigoureusement identiques. Ce dernier point est évidemment lié au problème de l'analyse sémantique des questionnaires. La méthode de recherche d'équivalences entre items la plus connue est celle qu'ont adoptée les chercheurs du *Projet SIMULMATICS*, pour réaliser la fusion d'une cinquantaine de fichiers de données individuelles (100.000 interviews environ) concernant divers problèmes politiques (un court résumé de cette procédure se trouve dans [Grémy 1970b : 176-178]). Dans une première étape, les chercheurs se sont efforcés d'identifier, dans chaque enquête, les items qui paraissaient correspondre à une même attitude politique : les critères utilisés étaient les ressemblances dans le contenu du libellé des questions, et l'identité des relations observées ("corrélations") entre les réponses aux questions et les principales variables socio-économiques. Dans une seconde étape, ils ont cherché à établir une hiérarchie entre les réponses aux items retenus, de manière à constituer une échelle de type GUTTMAN pour chaque attitude. Cette procédure paraît avoir donné satisfaction à leurs auteurs pour la construction d'indices numériques mesurant l'intensité des attitudes. Remarquons toutefois que, par construction, ces indices ne pouvaient pas s'appliquer aux individus, puisqu'il aurait fallu pour cela que toutes les questions sélectionnées figurent dans toutes les enquêtes utilisées : il s'agit donc ici d'indices destinés à caractériser des sous-populations bien déterminées (types d'électeurs, définis selon les variables socio-économiques usuelles et les comportements politiques et religieux).

Ce rapide survol des principaux aspects de l'analyse secondaire des données en sciences sociales nous permet de préciser les relations qui peuvent exister entre une théorie du questionnaire et les banques de données en sciences sociales. Avoir les moyens de comprendre ce que signifient les perforations

d'une carte mécanographique ou les signes binaires enregistrés sur une bande magnétique, ce n'est pas seulement disposer d'un dictionnaire établissant la correspondance entre les signaux lisibles par l'ordinateur et les informations qu'ils sont censés représenter. Certes, la connaissance de ce dictionnaire est indispensable pour traiter les données enregistrées : dans le cas de données agrégées, il précise le nom ou les caractéristiques des objets, des entités, ou des individus dénombrés ; dans le cas de données individuelles, il indique la réponse que la personne interrogée a choisie, ou la catégorie d'analyse de contenu à laquelle sa réponse a été rattachée. Mais il reste au chercheur à appréhender la signification des rubriques ou des énoncés figurant dans le dictionnaire. Nous avons vu que, dans l'analyse primaire de données d'enquêtes, ce problème ne se pose pas, puisque l'analyste est aussi l'auteur du questionnaire. D'ailleurs, Paul LAZARFELD a bien décrit le passage des notions et des concepts du chercheur à leur traduction opératoire en termes d'indicateurs, qui sont les réponses des personnes interrogées à des questions judicieusement choisies et formulées ; mais il manque une description du processus inverse, c'est-à-dire du passage de l'indicateur au concept. Espérer parvenir à une description opératoire de ce processus est certainement utopique : il n'existe pas d'indicateur univoque (chaque réponse fournie peut servir d'indicateur à des concepts différents, selon la nature des autres indicateurs auxquels il sera associé), et les indicateurs utilisés ne suffisent pas à cerner le concept sous-jacent avec certitude (ils ne sont qu'un échantillon non aléatoire de l'ensemble des indicateurs possibles de ce concept). Mais une meilleure connaissance du questionnaire peut contribuer à diminuer l'incertitude du sociologue sur ce que recouvrent les réponses enregistrées.

Lorsqu'il comporte l'énumération de l'ensemble des réponses possibles à chaque question (y compris aux questions ouvertes), le questionnaire d'enquête constitue la meilleure base pour l'élaboration du dictionnaire correspondant à un fichier de données individuelles. C'est pourquoi, dans certaines banques de données (telles que l'*Inter-University Consortium for Political and Social Research* d'Ann Arbor, ou l'*International Data Library and Reference Service* de Berkeley), le texte complet du questionnaire et de l'ensemble des

réponses possibles, ou des catégories d'analyse de contenu retenues pour les réponses aux questions ouvertes, est archivé sous une forme lisible par ordinateur. Or, ce questionnaire est un document linguistique ; il peut par conséquent faire l'objet d'une analyse automatique de type linguistique ou lexicographique. A notre connaissance, cette possibilité n'a guère été utilisée que pour élaborer des répertoires thématiques des questions posées dans des enquêtes (cf. [Phillipson 1968 : 217-220]). Toutefois, les progrès des techniques de l'analyse linguistique nous font espérer que le questionnaire d'enquête ainsi présenté pourra être étudié non seulement du point de vue de la sémantique des questions considérées isolément, mais aussi en tant que processus discursif global (cf. p.ex. [Harris 1962]). Outre son intérêt pour la méthodologie de l'enquête, une telle analyse présenterait l'avantage de permettre de tenir compte, dans l'interprétation des réponses au questionnaire, des effets de contexte induits par les questions posées précédemment ; ces effets pouvant varier d'un interviewé à l'autre, dans une même enquête, par le simple jeu des questions-filtres.

Avant de conclure sur l'analyse secondaire des données dans les sciences de l'homme, remarquons qu'il ne semble pas que les psychologues se soient beaucoup préoccupés de l'archivage des résultats de tests à des fins d'exploitation extérieure. Peut-être le souci de réaliser des épreuves psychologiques stables, peu sensibles aux influences perturbatrices de facteurs étrangers à ce qu'elles mesurent, ne dispose-t-il pas les psychologues à des recherches sur les variations interculturelles des réponses aux épreuves, ou sur le vieillissement des tests. Pourtant, la constitution de banques de données psychologiques favoriserait par exemple des recherches comparatives sur les différentes formes de tests ou d'épreuves de personnalité, voire des recherches pluridisciplinaires sur la personnalité de base, ou les variations dans les aptitudes (telles qu'elles sont évaluées par les tests) selon les facteurs socio-économiques (comme celles entreprises par l'INED et l'INOP). Dans l'application, la constitution de banques de données, analogues à celle du *Projet TALENT* réalisé aux Etats-Unis, déboucherait sur un ré-étalonnage permanent des épreuves, et sur un suivi scolaire et professionnel des sujets examinés. Il est vrai qu'en pratique, certains laboratoires font de l'analyse

secondaire sans la nommer ainsi, et disposent de leur propre stock de données ; il ne s'agit cependant pas de véritables banques de données, dans la mesure où les soucis de conservation et de mise à la disposition d'usagers extérieur n'y sont pas prédominants.

En conclusion, les bénéfices escomptés, pour l'analyse secondaire des données, d'une théorie du questionnaire, sont de trois sortes :

- possibilité de continuer à améliorer la qualité générale des questionnaires d'enquête, en particulier en évitant certaines ambiguïtés dans la formulation des questions ou des consignes, en développant la normalisation des questions usuelles (consommation, comportements politiques ou religieux, caractéristiques socio-économiques), en maîtrisant mieux les biais dans les réponses induits par la situation d'enquête. Il est clair qu'une atténuation des défauts les plus couramment relevés dans les questionnaires et qu'une standardisation accrue des techniques de recueil d'information, si elle ne profite pas immédiatement à l'analyse secondaire, facilitera cependant l'archivage et l'exploitation ultérieure des données recueillies.

- meilleure exploitation des informations contenues dans les données grâce à une analyse linguistique approfondie du texte du questionnaire, permettant de mieux comprendre la signification des questions et des réponses. Si l'on admet, en paraphrasant Jean ULLMO, que "le questionnaire est une théorie cristallisée", l'identification de l'ensemble des présupposés des questions et des réponses par une analyse appropriée (cf. le chapitre 2) conduit à une reconstitution au moins partielle de la problématique de l'auteur du questionnaire, et ainsi à un approfondissement du sens des réponses recueillies. Si les progrès de la linguistique continuent sur leur lancée, peut être sera-t-il possible à moyen terme de coupler l'analyse sémantique du questionnaire et l'analyse statistique des données dans une même procédure automatisée.

- clarification des rapports entre le "questionnaire empirique" ayant servi à recueillir les données, et les multiples "questionnaires théoriques" qui président à leur exploitation. Nous avons vu que, pour chaque enquête, le questionnaire empirique est une traduction opératoire du questionnaire théorique du chercheur qui entreprend l'enquête, et que les plans de dépouillement successifs sont des traductions opératoires des questionnaires théoriques qu'il élabore au fur et à mesure du développement de ses réflexions sur les données. Dans le cas d'une analyse secondaire d'une base de données complexe, résultant de la fusion de plusieurs fichiers, l'analyse linguistique du questionnaire empirique correspondant à chacun des fichiers conduit à la reconstitution d'un questionnaire théorique par fichier. D'autre part, pour la sélection des items pertinents, et leur combinaison dans la construction de nouvelles variables, le chercheur qui entreprend l'analyse secondaire utilise au moins implicitement un questionnaire théorique qui lui est propre, et dont le plan d'analyse est une traduction opératoire ; les étapes successives de l'analyse concrétisant les questionnaires théoriques successifs élaborés par le chercheur. Une meilleure compréhension des relations formelles qui unissent les questionnaires théoriques aux questionnaires empiriques, et les questionnaires théoriques entre eux, peut contribuer au perfectionnement des langages d'interrogation de fichiers de données d'enquête.

0.3. OBJECTIFS DE LA RECHERCHE ENTREPRISE

Nous avons par conséquent pris le parti de considérer comme point de départ de notre travail le fichier d'enquête archivé dans une banque de données, comportant à la fois le texte du questionnaire et des réponses codifiées, et les résultats de l'enquête (le tout sous une forme lisible par ordinateur : fichier de classe I, dans la terminologie adoptée à l'I.C.P.S.R. ; cf. [Grémy 1978 : 9]), et de nous placer dans la situation générale de l'analyse secondaire de ces résultats. En d'autres termes, l'objet

de notre recherche est, en définitive, moins le questionnaire empirique concret, tel qu'il a été utilisé sur le terrain par les enquêteurs, que le *questionnaire formel classificatoire* servant à organiser les réponses dans le fichier archivé (cf. [Flament 1976 : 52-58]). Il est toutefois évident que les informations ainsi classées et conservées tirent leur signification du questionnaire empirique qui a servi à les recueillir. C'est pourquoi le problème central, dans cette perspective, est l'interprétation des réponses recueillies, en ne disposant toutefois que de la description de la technique d'échantillonnage, du texte du questionnaire (et éventuellement des consignes aux enquêteurs), et de la liste des réponses enregistrées pour chaque question. La seule procédure applicable dans ces conditions consiste à remonter, par une démarche inductive, des indicateurs aux concepts : non nécessairement d'ailleurs aux concepts qui furent ceux de l'auteur du questionnaire, mais à un échantillon des concepts possibles qui satisfasse le chercheur.

Une telle démarche est étroitement liée à la nature et au contenu des données, ainsi qu'aux centres d'intérêt et aux hypothèses de l'analyste. Il serait donc difficile, alors que l'analyse secondaire n'en est encore qu'à ses débuts, d'en faire une description générale ; il serait en tout cas prématuré de tenter de la formaliser, ou de la réduire à une démarche-type. Par contre, si l'on considère que l'aboutissement de cette procédure est l'analyse statistique des résultats d'enquête afin d'en tirer des conclusions théoriques, il est possible d'énumérer quelques-unes des étapes qui préparent l'analyse des données. Par exemple, sur un fichier déterminé considéré isolément, la préparation de l'analyse statistique comporte au moins quatre grandes phases :

① Sélection des indicateurs. L'analyste, à ce stade, se demande quels étaient les objectifs du responsable de l'enquête : que cherchait-il ? Comment a-t-il procédé ? Qu'a-t-il obtenu en réalité ? Cette série d'interrogations débouche sur une question pratique : compte-tenu des objectifs de l'analyste, quelles informations va-t-il extraire du fichier ? Ou, plus pragmatiquement, quelles informations peut-il écarter sans qu'il soit nécessaire de les soumettre à un examen plus approfondi ? Cette première étape

se traduit donc par un élagage sommaire du questionnaire, des séries entières de questions étant écartées parce que considérées *a priori* comme extérieures aux préoccupations de l'analyste.

② Analyse des ensembles d'indicateurs non rejetés. Toute question désigne un ensemble d'indicateurs, qui est l'ensemble des réponses enregistrées à cette question. Pour chaque question provisoirement retenue, l'analyste se demande : Que signifie l'énoncé de la question pour les interviewés ? Quelles autres formulations auraient été possibles ? Pourquoi n'ont-elles pas été adoptées ? Les expressions choisies risquent-elles d'avoir modifié la réponse que l'interviewé avait l'intention de donner ? Les questions posées précédemment donnent-elles à cette question un éclairage particulier, qu'elle n'aurait pas eu dans un autre contexte ? Qu'attendait-on comme genre de réponses ? Qu'a-t-on obtenu en réalité ? Pourquoi ? La consigne de la question était-elle précise et sans ambiguïté ? Quel est le contenu des réponses effectivement enregistrées ? Quels autres systèmes de classification des réponses aurait-on pu réaliser ? Quelles significations auraient-ils permis de dégager ? De quoi les catégories retenues pour l'analyse du contenu des réponses peuvent-elles être l'indicateur ? En quoi cela concerne-t-il les préoccupations de l'analyste ? A l'issue de cette seconde phase, certains ensembles d'indicateurs sont retenus pour l'étape suivante.

③ Reconstruction et combinaison des indicateurs retenus. La manière dont les réponses ont été codées initialement ne convient généralement pas à l'analyste, puisque, par définition, sa problématique et ses objectifs sont distincts de ceux du promoteur de l'enquête. Il est donc souvent nécessaire de supprimer une précision devenue inutile en regroupant (par recodification) certaines des réponses enregistrées à une même question, ou au contraire d'augmenter la précision de l'information en combinant plusieurs réponses obtenues à des questions différentes. Nous retrouvons ici le processus de construction de variables décrit par LAZARSELD.

④ Conception d'un plan d'analyse des données ainsi restructurées. Cette étape est l'équivalent du dépouillement d'enquête dans l'analyse primaire. Elle implique une bonne connaissance de l'ensemble des méthodes mathématiques d'analyse de données, et de leurs articulations avec les problèmes de la recherche dans les sciences de l'homme.

Les deux dernières étapes ne sont pas spécifiques de l'analyse secondaire. En outre, leur analyse a déjà fait l'objet, en langue française, de travaux importants (dont les plus notables sont d'ailleurs des thèses de doctorat : [Reuchlin 1962], [Boudon 1967], [Flament 1976], [Bacher 1978]). Les deux premières étapes, en revanche, constituent un champ de recherches non totalement vierge, certes, mais actuellement encore peu exploré par les psychologues ou les sociologues. Notre propos a donc été de mettre l'accent sur l'analyse formelle du questionnaire, en tant qu'objet linguistique, et la connexion entre cette analyse sémantique et l'analyse statistique des réponses obtenues au questionnaire.

L'exposé est centré sur trois grands thèmes :

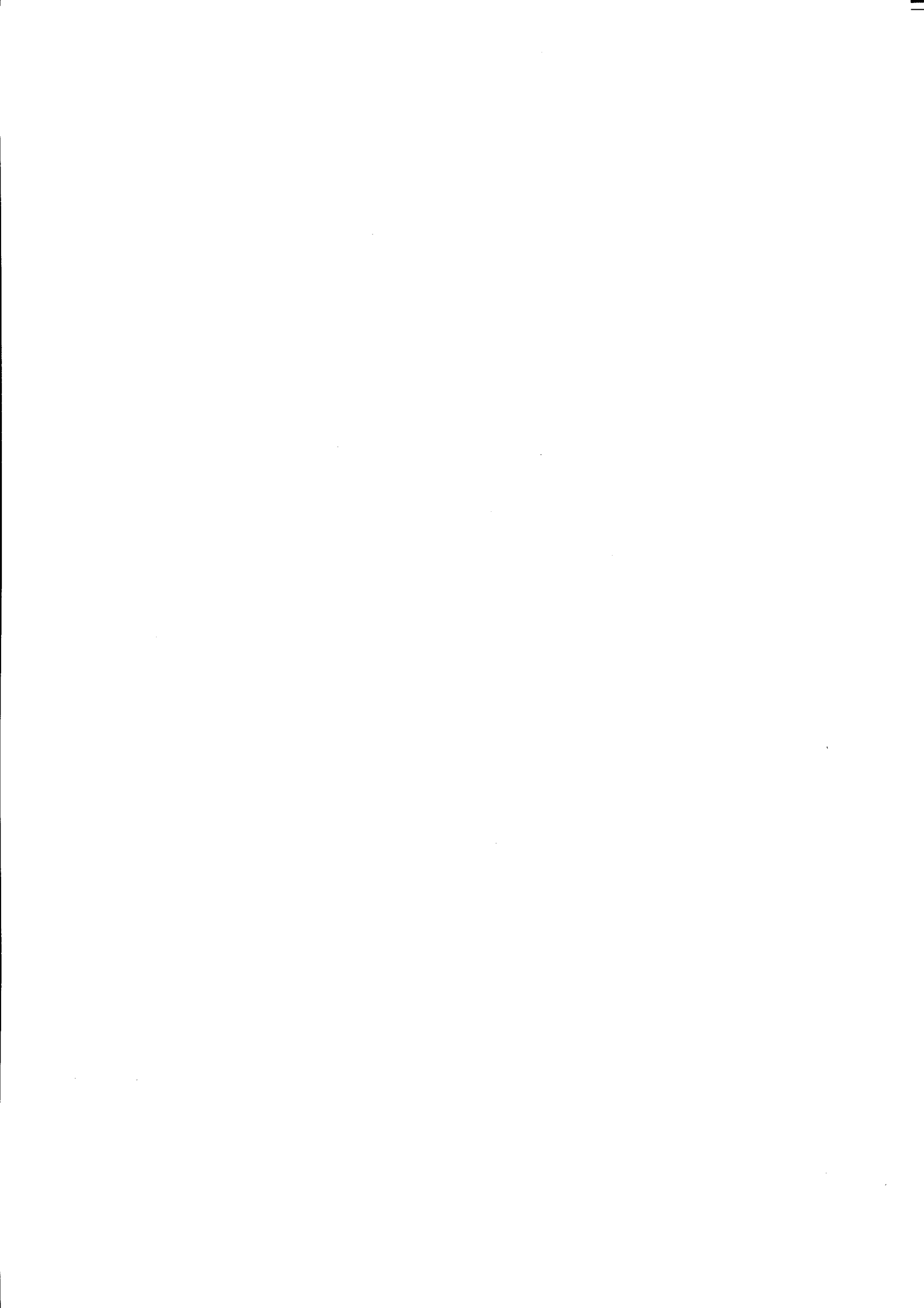
① *Qu'est-ce qu'une question ?* A partir de l'analyse systématique d'un corpus de 2.000 énoncés de questions, tirés de tests ou de questionnaires divers, nous commençons par montrer l'extrême diversité morpho-syntaxique de ce que psychologues et sociologues appellent une question. Nous examinons ensuite, à la lumière des acquis de la linguistique contemporaine, la possibilité de réduire à une forme canonique tout énoncé à valeur de question quel qu'il soit, sans altérer son sens. Nous sommes ainsi conduit à mesurer l'étroitesse de la relation qui lie une question à l'ensemble de ses réponses.

② *Questions et réponses.* Les relations syntaxiques et sémantiques qui unissent une question à l'ensemble de ses réponses possibles ont été surtout étudiées par les logiciens. Dans ce domaine, l'apport de la logique érotétique nous a paru surpasser très largement celui de la linguistique. C'est pourquoi l'analyse formelle que nous présentons en est largement inspirée, ainsi d'ailleurs que les systèmes de classification des questions et

des réponses aux questions que nous décrivons.

③ *Questionnaire et analyse des données.* La structure du questionnaire détermine dans une certaine mesure les limites de l'analyse statistique des réponses aux questions qui le composent. L'analyse formelle des relations entre questions et réponses, et la réduction du questionnaire tout entier à une forme canonique, débouchent naturellement sur la définition des informations élémentaires qu'apporte l'administration du questionnaire. Si l'on pousse l'analyse un peu plus loin, on peut déterminer l'existence de sortes de quanta d'information, et montrer que la plupart des techniques statistiques usuelles sont fondées sur un simple dénombrement des quanta d'information apportés par le questionnement.

En guise de *conclusions*, nous présentons enfin une liste de questions demeurées sans réponse satisfaisante pour le moment, et nous suggérons quelques directions de recherches qui nous ont parues prometteuses.



1. QU'EST-CE QU'UNE QUESTION ?

Nous commencerons par poser provisoirement qu'une question est un énoncé, verbal ou écrit, ayant pour fonction de susciter un apport d'information de la part de l'interlocuteur. Nous laisserons de côté pour l'instant les problèmes tels que : "qui interroge qui ?", "pour quoi faire ?", "dans quel contexte ?", ou "comment se fait-il que la personne interrogée réponde effectivement ?". Nous y reviendrons dans les chapitres suivants. Notre propos, dans ce premier chapitre, est de fournir une description morpho-syntaxique des questions, et, par cela même, de définir les traits linguistiques qui les caractérisent.

Pour cela, nous avons constitué un corpus représentant environ 2.000 énoncés de questions variés, empruntés à des questionnaires d'enquête (administrés individuellement par enquêteur ou collectivement sur table, ou adressés par voie postale), et à diverses épreuves psychologiques (inventaires de personnalité, tests d'aptitudes ou de connaissances, etc.). Tous ces énoncés sont en langue française, et tous (même ceux qui sont traduits d'une langue étrangère) ont été effectivement utilisés sous cette forme. On trouvera en annexe I la liste des documents recensés.

Au premier abord, on ne peut qu'être frappé par l'extraordinaire variété des énoncés relevés. Bien que la situation de test ou d'entretien d'enquête impose une relative standardisation des formulations, et interdise en principe certaines formes d'interrogation (comme par exemple "l'interrogation élevée à la puissance deux" [Jespersen 1924 : 433], possible par contre dans un entretien moins structuré), nous avons trouvé des énoncés de questions qui n'entrent pas dans les catégories que les grammairiens ont assignées aux interrogations. C'est

pourquoi nous avons consacré la première partie de ce chapitre à décrire succinctement cette variété, à partir des classifications les plus communément admises dans les grammaires traditionnelles.

Dans la seconde partie du chapitre 1, nous nous efforçons de réduire à une forme canonique les multiples formes d'énoncés exprimant une question. Nous verrons que la grammaire transformationnelle permet, par des chaînes de transformations paraphrastiques élémentaires appliquées aux énoncés, d'aboutir à un type d'énoncé unique caractérisant la question. Cette forme canonique servira ultérieurement de point de départ à une analyse logico-sémantique des questions (chapitre 2).

1.1. POLYMORPHIE DES ÉNONCÉS EXPRIMANT UNE QUESTION

Dans une recherche comparative portant sur l'expression de l'interrogation en français (littéraire) contemporain, un linguiste fait remarquer l'"étonnante polymorphie instrumentale" de l'interrogation française [Wandruszka 1970 : 65], sans équivalent dans les langues anglo-saxonnes ou dans les autres langues romanes. Si l'on consulte une grammaire française à prétention normative, on y apprend que l'on interroge en français soit à l'aide d'un mot interrogatif placé en tête de la phrase, soit en inversant le pronom sujet, soit en recourant à la périphrase interrogative *est-ce que*, soit par la seule intonation ; ces différents moyens pouvant d'ailleurs être combinés entre eux sous certaines conditions (cf. p.ex. : [Bloch 1945 : 217-218] ; [Georgin 1952 : 291-293] ; [Mauger 1968 : 379-381]). Les grammaires à visée descriptive confirment pour l'essentiel cette présentation, en ne relevant pas d'autres formes d'interrogation directe que celles exprimées ci-dessus (cf.p.ex. : [Guiraud 1962 : 89-91] ; [Wagner 1962 : 527-532]). S'il est vrai que, dans notre corpus, l'interrogation directe est la forme de questionnement que l'on rencontre le plus fréquemment, nous verrons que l'on utilise pour interroger bien d'autres formes de phrases.

Comme l'objet de cette première partie du chapitre I est de présenter les multiples facettes des énoncés exprimant une question, en utilisant pour cela les catégorisations que proposent les grammaires usuelles, ce sera pour nous l'occasion d'introduire des distinctions formelles que nous retrouverons au stade de l'analyse des énoncés.

1.1.1. Fausses interrogatives, et vraies questions sans interrogation

La forme interrogative est l'une des catégories utilisées par les linguistes pour caractériser les phrases simples ; chacune de ces catégories, ou *modalités*, est en principe définie à partir de critères uniquement morpho-syntaxiques. Or, il se trouve que les catégories grammaticales permettant de classer les phrases simples ne coïncident pas tout à fait avec les fonctions que ces phrases remplissent dans le discours. Nous verrons en particulier qu'il existe des phrases au tour interrogatif qui n'ont pas pour fonction de recueillir de l'information auprès de l'interlocuteur.

Après ce préalable théorique, l'essentiel de notre propos sera de montrer les différentes formes que revêtent, dans notre corpus, les questions n'utilisant pas la forme interrogative, directe ou indirecte.

a) Modalités de la phrase et fonctions de l'énoncé

Les modalités "expriment la façon dont le locuteur envisage et présente l'énoncé" [Guiraud 1962 : 88]. Parmi les modalités de base, les grammairiens distinguent les formes :

- *interrogative*, qui traduit en principe une question ;
- *énonciative* (appelée également : présentative, déclarative, assertive, ou affirmative), qui correspond à "la *présentation* pure et simple d'un fait, d'un événement ou d'un jugement" [Veyrenc 1968 : 104] ;
- *injonctive* (appelée également : impérative ou jussive), qui exprime un ordre.

Pour certains auteurs, ces trois catégories épuisent l'ensemble des modalités de phrase possibles. D'autres distinguent par exemple la forme *désidérative*, dont l'impératif et l'optatif sont des sous-catégories [Guiraud 1962 : 88-89], ou la

forme *exclamative*, incluant l'impératif [Wagner 1962 : 524] ; d'autres utilisent une classification en cinq types, mettant les formes *exclamative* et *optative* sur le même plan que les trois formes énumérées ci-dessus [Bloch 1945 : 217-220] [Veyrenc 1968 : 103-107] ; on trouve enfin une typologie en sept catégories, ajoutant aux cinq types précédents l'expression du *doute* ou de la possibilité, et la *négation* [Mauger 1968 : 369-386].

Pour notre propos, nous nous limiterons aux trois modalités de base acceptées par la quasi-totalité des grammairiens, puisque nous n'avons pas rencontré dans notre corpus d'énoncé prenant une forme optative, ou exclamative (au sens strict). En outre, cette catégorisation présente l'avantage de correspondre aux trois formes de l'élément obligatoire du constituant de phrase dans la syntaxe générative et transformationnelle du français de type CHOMSKY (cf. p. ex. : [Dubois 1970 : 133]).

Les linguistes ont relevé la difficulté qu'il pouvait y avoir à ranger certaines phrases dans une catégorie : "En fait, une phrase manifestement interrogative comme *Veux-tu venir ici ?* [...] peut être équivalente sémantiquement à la phrase impérative *Viens ici, veux-tu ?* [...] et on a du mal à la classer dans une modalité plutôt qu'une autre" [Lyons 1968 : 236]. Cette difficulté naît des désaccords entre les critères morpho-syntaxiques généralement utilisés pour caractériser la modalité d'une phrase, et les critères logico-sémantiques permettant d'en définir la signification. Les fonctions que les énoncés remplissent dans le discours sont, d'une certaine manière, liées à leurs formes syntaxiques ; mais elles ne se confondent pas avec les modalités de la phrase. En effet, "le désaccord entre la grammaire et la logique consiste en ce que les catégories logiques et les catégories grammaticales se recouvrent très rarement ; presque jamais le nombre des unes et des autres ne coïncide" [Vendryes 1923 : 132]. Il semble toutefois que les grammairiens qui ont tenté de distinguer modalités et fonctions n'y soient pas toujours parvenu. Tel nous paraît être le cas d'Otto JESPERSEN : se proposant d'établir "une classification *purement notionnelle* des énoncés, sans faire intervenir la forme grammaticale qu'ils prennent", il définit deux classes fondamentales, selon que le locuteur désire ou ne désire pas "exercer une influence par le seul intermédiaire

de l'énoncé sur la volonté de celui à qui il est destiné" [Jespersen 1924 : 430-431] ; les questions, qui ont pour but de susciter une réponse de l'interlocuteur, se rangent dans la première classe. Mais la description que JESPERSEN donne ensuite de celles-ci n'englobe aucune phrase qui, grammaticalement, ne soit une interrogative.

La catégorisation des fonctions remplies par les énoncés a été menée à bien par les philosophes du langage. Après l'inventaire, par Ludwig WITTGENSTEIN, des divers "jeux de langage" [Wittgenstein 1953 : § 23], c'est probablement John AUSTIN qui a fourni les concepts les plus féconds pour analyser les fonctions du langage. L'idée fondamentale d'AUSTIN nous paraît être que toute énonciation est essentiellement une action, un "acte de discours" [Austin 1962 : 147], visant à produire un effet ; d'où le titre de l'ouvrage où il développe ce point de vue : *How to do Things with Words* (Remarquons que cette idée, inspirée de WITTGENSTEIN [Wittgenstein 1953 : § 25], se trouve également chez le linguiste BLOOMFIELD, quoique dans un contexte théorique sensiblement différent. Cf. [Bloomfield 1961 : 27-28]). AUSTIN distingue trois aspects des actes de discours : un aspect *locutoire*, qui se réduit à l'acte d'énonciation ; un aspect *illocutoire*, qui correspond à la signification de l'acte locutoire selon le contexte dans lequel il est effectué (telle locution *a valeur de* question, d'ordre, de constatation, d'avertissement, de promesse, etc.) ; et un aspect *perlocutoire*, traduisant l'objectif réellement visé par le locuteur (flatter, effrayer, tromper, etc.) [Austin 1962 : 98-103]. Si l'on analyse la forme des actes illocutoires, on remarque que ces énoncés présentent un verbe à la voix active, à la première personne du singulier du présent de l'indicatif : "je demande", "j'accepte", "je parie", etc. Ce sont des phrases *performatives* (cf. [Austin 1962 : 5-7]).

Il est possible de classer les actes illocutoires et les verbes performatifs correspondant. A la classification d'AUSTIN, que lui-même ne considère d'ailleurs pas comme entièrement satisfaisante [Austin 1962 : 150-162], on préférera par exemple celle de SEARLE, qui distingue la *requête* (dont l'*ordre* est un cas particulier), l'*assertion*, la *question*, le *remerciement*, le *conseil*, l'*avertissement*, la *salutation* et la *félicitation* [Searle 1969 : 66-67]. Non que cette dernière typologie soit beaucoup plus satisfaisante : comme l'auteur

le fait remarquer, il y a des recouvrements entre certains types d'actes illocutoires [Searle 1969 : 69]. Mais cette classification présente le mérite à nos yeux d'accorder une place bien définie à l'acte de questionnement.

L'acte de questionnement peut prendre deux formes. L'une, que l'on peut appeler *question réelle* (SEARLE), ou *interrogation authentique* (SEMERARI), correspond à la situation dans laquelle le locuteur a besoin de l'information qu'il suppose détenue par l'interlocuteur ; l'incertitude du locuteur est donc réelle. L'autre forme est la *question d'examen*, qui est destinée à assurer le locuteur que l'interlocuteur possède bien une information donnée, que lui-même possède également ; son incertitude est donc feinte (cf. [Searle 1969 : 66] ; [Semerari 1968 : 27-29]). Nous retrouverons cette distinction ultérieurement, avec les questions du type "test de connaissances", et celles du type "questionnaire psychosocial".

b) Fausse questions

Il n'y a donc pas toujours coïncidence entre la forme de l'énoncé et sa fonction dans le discours. Cela signifie qu'il peut y avoir des questions qui n'empruntent pas le tour interrogatif (nous en verrons des exemples aux paragraphes suivants), et des interrogatives qui ne sont pas des questions. Nous avons distingué quatre variétés d'énoncés interrogatifs constituant de "fausses questions" : l'interrogation rhétorique, l'interrogation orientée, l'interrogation impositive, et l'interrogation injonctive.

L'*interrogation rhétorique* est une figure de style dont on se sert non pour interroger, mais "pour délibérer, pour prouver, pour décrire, pour accuser, pour blâmer, pour exciter, pour encourager, pour dissuader, enfin pour mille divers usages" [Fontanier 1827 : 370]. Cette figure "consiste à prendre le tour interrogatif, non pas pour marquer un doute et provoquer une réponse, mais pour indiquer au contraire la plus grande persuasion, et défier ceux à qui l'on parle de pouvoir nier ou même répondre" [Fontanier 1827 : 368]. Ainsi, la phrase : "Qui étudie encore Aristote de nos jours ?" est une manière emphatique de dire : "Personne n'étudie plus Aristote de nos jours !" [Sadock 1974 : 79]. De même : "Où ne voyage-t-il pas ?" signifie : "Il finira par connaître le monde entier"

[Wagner 1962 : 539] ; et : "Je ne t'ai point aimé, cruel ?" a une valeur exclamative plutôt qu'interrogative [Georgin 1952 : 292, n.2]. Ce procédé oratoire est parfois appelé *interrogation apparente* [Brunot 1926 : 487].

L'*interrogation orientée* est également un artifice de style. Mais, à la différence de l'interrogation rhétorique, elle invite l'interlocuteur à répondre. Il s'agit cependant bien d'une fausse question, car le locuteur dicte en quelque sorte sa réponse à l'interlocuteur. Ainsi, lorsque X demande à Y : "N'est-ce pas que Z a bien agi cette fois-ci ?", X sait bien que Y considère comme lui que Z a bien agi [Semerari 1968 : 28]. Toutefois, Y ne pourra pas se dispenser de répondre : "Certainement, Z a agi d'une façon remarquable". De même, la pseudo-question : "Ne viendrez-vous pas ?" appelle une réponse positive : "Si, je viendrai" [Wagner 1962 : 539].

Contrairement aux vraies questions, cette forme d'interrogation intentionnellement biaisée présente des caractéristiques morpho-syntaxiques précises. En français, présence d'une locution ou d'un terme tel que : "dites", "n'est-ce pas", "j'espère" (*questions insistantes*) ou fausse négation (*questions dirigées*) [Wagner 1962 : 538-539]. En latin, en grec ancien, en anglais, en russe, il existe des tournures différentes selon que l'on attend une réponse positive ou négative (cf. [Collart 1975 : 107] ; [Bizos 1949 : 112-113] ; [Thomson 1960 : 62, § 106] ; [Ferrand 1967 : 142, § 126]). Et même, pour le grec ancien, selon que l'on attend un *non* tout en craignant de recevoir un *oui* [Ragon 1889 : 164, § 227] ; pour le russe, selon que l'on attend une réponse négative, ou que l'on souhaite une réponse négative [Tesnière 1959 : 208].

L'*interrogation impositive* a pour fonction d'exprimer un souhait du locuteur : "Pourquoi ne fermez-vous pas la porte ?" ; ou : "Si nous mettions un disque ?" [Sadock 1974 : 114-118]. En français, elle peut également prendre des formes telles que : "Que diriez-vous d'aller au cinéma ?" ; ou : "Pourquoi perdre notre temps ?" ; ou encore : "A quoi bon continuer ainsi ?".

L'*interrogation injonctive* sert à exprimer un ordre. On considère le plus souvent qu'il s'agit d'une forme atténuée de commandement, dont on use par

politesse : "Voulez-vous entrer ?" ou : "Pourriez-vous fermer la fenêtre ?". Cependant, il arrive que le tour interrogatif, loin d'atténuer l'expression de l'ordre, serve au contraire à la renforcer : "Allez-vous vous taire ?" ou : "Quand vous déciderez-vous enfin à fermer la porte ?".

Le fait qu'il s'agisse de fausses questions est mis en évidence par l'incongruité qu'il y aurait, dans le dialogue, à les traiter comme des questions véritables. Ainsi, lorsque Clytemnestre dit à Agamemnon (Jean RACINE : *Iphigénie*, IV, 4, v. 1258) : "Pensez-vous par des pleurs prouver votre tendresse ?" (interrogation rhétorique), il serait mal venu à celui-ci de répliquer que telle est bien en effet son intention. De même, quand Alice demande à Humpty Dumpty perché sur son mur (Lewis CARROLL : *De l'autre côté du miroir*, chap. 6, début) : "Ne croyez-vous pas que sur le sol, vous seriez plus en sécurité ?" (interrogation orientée), la réponse de ce dernier : "Bien sûr que je ne le crois pas !" est perçue par le lecteur comme un manquement aux règles implicites de la conversation. Ceci est encore plus évident pour l'interrogation injonctive ; on imagine mal l'interlocuteur répondant sans faire un geste : "effectivement, il me serait possible de fermer la fenêtre". Seule, l'interrogation impositive peut, à la rigueur, susciter une véritable réponse chez l'interlocuteur ; mais le contexte permettrait alors de trancher puisque la réponse constituerait non un simple apport d'information, mais une véritable prise de décision.

c) Injonctives à valeur de question.

En français, les injonctions se caractérisent par le mode et le temps du verbe principal : impératif, indicatif futur, infinitif, subjonctif [Georgin 1952 : 109-110]. La dernière forme (qui a d'ailleurs plutôt un sens optatif) ne se rencontre pas dans notre corpus. Dans les énoncés de type injonctif que nous avons relevés, l'injonction porte presque toujours sur la manière d'exprimer la réponse plus que sur la réponse elle-même : dire, montrer, décrire, énumérer, classer, citer, etc. D'autre part, bien que cette forme de question soit relativement plus fréquente dans les tests d'aptitudes, elle se rencontre également dans les questionnaires d'enquête.

Les *injonctives* à l'*impératif* peuvent n'avoir pour complément qu'un substantif sans proposition relative. Par exemple, dans les tests d'aptitudes :

- "Montre-moi la plus grande ligne" (TERMAN : 4 ans, 1°)
- "Dis-moi les jours de la semaine" (TERMAN : 7 ans, supplément 1)
- "Classez les 12 éléments contenus dans les cases en deux groupes de 6, ces 6 éléments ayant un caractère commun" (ABS)
- "Effectuez ces opérations, sans les poser et en allant le plus vite possible" (CC 8: I)
- "Voici le même message que tout à l'heure écrit au moyen d'un autre code. Étudiez-le, essayez d'en comprendre le système, et écrivez le même mot : connu, avec ce nouveau code" (TMI : 27)

Dans les questionnaires d'enquête :

- "Indiquez avec le plus de précision possible la profession de votre père" (OCDE-IUT : 10a)
- "Donnez par ordre d'importance les 3 principales raisons parmi celles énumérées ci-dessous" (OCDE-IUT : 53)
- "Nommez votre emploi et décrivez-le de la manière la plus détaillée possible" (CEREQ-SEU : 29a)
- "Si vous avez éliminé votre premier choix, précisez les raisons" (F.ING 69 : VI 12)
- "Si vous envisagez de poursuivre d'autres études, dès la fin de vos études à l'I.U.T., précisez le type d'études" (OCDE-IUT : 51a)

Lorsque le verbe à l'*impératif* a un complément déterminé par une relative, trois cas sont à distinguer, selon que le complément est un substantif ou un pronom, et, dans cette dernière hypothèse, selon le sens du verbe. En effet, si le verbe (ou la locution verbale) a le sens de *dire*, *demander*, *voir*, *savoir*, *comprendre*, etc., il introduit une interrogative indirecte (cf. [Wagner 1962 : 568-569]). Nous avons classé à part les interrogations indirectes, puisqu'il s'agit de véritables interrogatives. C'est pourquoi nous n'avons retenu ici que les énoncés correspondant aux deux premiers cas. Par exemple, dans les tests psychologiques (d'aptitudes ou projectifs) :

- "Regardez chaque groupe et cherchez la valeur du domino qui manque" (D 48)
- "Lisez attentivement toute la ligne, et trouvez le mot qui manque, le plus vite possible" (ECO 5 : 3)
- "Regardez les mots qui se trouvent sur la ligne de droite et choisissez celui qui ne semble pas être de la même espèce que les autres" (CAT 16PF : 122)
- "Maintenant, dites-moi tous les mots que vous pouvez qui riment avec jouer" (TMI : 13)
- "Vous avez devant vous une tache d'encre, essayez d'imaginer tout ce à quoi elle ressemble, écrivez seulement le plus grand nombre de choses auxquelles cela vous fait penser" (TACHES)

Dans les questionnaires d'enquête :

- "Citez les revues que vous lisez au moins une fois par mois" (BNIST : 16)
- "Classez par ordre d'importance les raisons qui vous incitent à exercer une activité professionnelle" (F.ING 69 : IX 2)
- "Si vous n'appartenez pas à l'une des 14 catégories ci-contre, écrivez dans le cadre profession (n° 8) les renseignements qui vous concernent" (REL 55 : page 2)

Les injonctives à l'indicatif futur ont l'apparence d'énonciatives: leur valeur d'injonctive tient à la personne du verbe principal, qui désigne l'interlocuteur, et au sens du verbe, qui exprime une action. Une variante de ce type d'énoncé présente la tournure idiomatique *aller* au présent de l'indicatif suivi d'un verbe à l'infinitif, tournure qui exprime un futur proche. Exemples relevés dans les tests d'aptitudes :

- "Voici un proverbe. Vous allez me dire ce qu'il signifie" (TMI : 29)
- "[...] je vais te dire trois mots, et tu feras une phrase dans laquelle il y aura les trois mots" (TERMAN : 9 ans,59)
- "J'aimerais savoir combien de mots tu connais, écoute-moi, et quand j'aurai dit un mot tu me diras ce qu'il signifie" (TERMAN : vocabulaire)

- "Maintenant je voudrais savoir combien de mots tu peux dire en trois minutes. Quand je dirai ça y est, tu commenceras, et tu diras les mots aussi vite que tu pourras [...]".
(BS Z : 44)
- "Vous trouverez ci-dessous des pensées accompagnées de quatre ou six phrases. Vous ferez une croix dans les cercles correspondant aux deux phrases dont le sens vous paraît se rapprocher le plus (ou s'éloigner le moins) de la pensée".
(BV 50)

Dans les questionnaires d'enquête :

- "Pour chacun des magazines que je vais vous citer, vous me direz s'il est plutôt moderne ou plutôt vieillot" (SOFRES 686 : 32a)
- "Je vais vous citer quelques points et vous allez me dire ceux qui vous semblent véritablement un bien et ceux qui, au contraire, vous semblent un mal" (SOFRES 686 : 18)

Les injonctives à l'infinitif ne se rencontrent que dans des documents destinés à être lus par l'interlocuteur : tests papier-crayon, et questionnaires postaux ou administrés collectivement sur table. Par exemple :

- "Trouver la réponse aussi vite que possible" (NB : 5)
- "Dans les phrases suivantes il manque un mot. Trouver le mot qui convient le mieux, parmi ceux qui sont en-dessous de la phrase"
(ECO 5 : 2)
- "Classer de 1 à 4 les activités qui permettent de normaliser le vocabulaire scientifique et technique" (BNIST : 25)

d) Énonciatives à valeur de question

C'est à propos des énonciatives faisant fonction de question que le problème de la distinction entre l'aspect morpho-syntaxique des énoncés et leur aspect logico-sémantique se pose avec le plus d'acuité. On peut en effet admettre qu'"une interrogation est aussi une sorte d'ordre, c'est-à-dire l'ordre de dire quelque chose à celui qui pose la question" [Jespersen 1924 : 431]. Par contre, il est plus difficile d'assimiler interrogatives et énonciatives. D'ailleurs, aucune grammaire française ne signale la possibilité de questionner en

utilisant une assertion (Par contre, elle est mentionnée pour l'anglais dans [Sadock 1974 : 10]). Pour ce survol des énoncés utilisés dans les tests et les questionnaires, nous avons donc admis le point de vue de la primauté de la forme sur la fonction, selon lequel par exemple "on ne doit pas confondre l'énoncé postulé (optatif ou impératif) avec des phrases du type : *je souhaite, je désire, j'ordonne, etc., qu'il vienne* ; formellement ce sont des déclarations dont les valeurs sont purement lexicales" [Guiraud 1962 : 89].

Dans notre corpus, les énonciatives à valeur de question peuvent être classées en deux grandes catégories : celles qui expriment une *consigne*, une procédure à suivre pour apporter de l'information au locuteur, et celles qui sont formulées en termes de *réponse*, à compléter ou à élaguer par l'interlocuteur.

Les *énonciatives-consigne* apparaissent comme des injonctives déguisées. Aussi ne les rencontre-t-on guère que dans les épreuves psychologiques. La forme la plus courante rappelle l'impératif catégorique ; elle est du type : "il faut", "vous devez". Par exemple :

- "Il faut trouver parmi les 5 dessins placés sur la même ligne, à droite, celui qui manque dans le coin inférieur droit du grand carré" (EAI 5/2 : Ma.)
- "Dans chaque série, vous devez indiquer les trois activités que vous préférez et les trois que vous aimez le moins, en répondant uniquement sur la feuille de réponses" (QIA/m)
- "Chaque ligne comprend 5 mots placés dans un ordre quelconque, et vous devez les ordonner logiquement en commençant par celui qui porte le numéro 1" (EAI 5/1 : G.V.)
- "Je vais vous montrer un dessin pendant dix secondes, puis je recouvrirai la carte et vous devrez dessiner de mémoire ce que vous aurez vu" (TMI : 6)

Une forme légèrement atténuée se présente comme une description de la tâche à accomplir par l'interlocuteur :

- "On vous donne maintenant huit phrases. Il s'agit de classer ces huit phrases en autant de façons que possible" (ABS : V6)

- "Chaque problème se compose d'une série de dessins qui se transforment selon une ou plusieurs règles. Votre travail consiste à découvrir ces règles, et à les appliquer en traçant les dessins qui manquent dans les cases vides" (KRX 25)
- "Vous regardez bien à gauche la position du point marqué et vous cherchez parmi les 5 cases de droite celle où vous pouvez faire un point dans les mêmes conditions" (CB 2 : 4)
- "Je vais maintenant te lire une phrase qui n'est pas tout à fait finie. Tu n'auras qu'à la terminer" (BS Z : 48)

Une autre variété d'énonciatives-consigne fait apparaître le locuteur, soit sous une forme anonyme ("on vous demande"), soit sous une forme plus personnelle ("je") :

- "On vous donne un proverbe tel que celui-ci : il n'y a pas de roses sans épines, et on vous demande de trouver parmi les cinq phrases écrites à la suite celle qui ressemble le plus au proverbe" (EAI 5/1 : Pr.)
- "On vous donne deux membres de phrases, précédés des lettres (a) et (b) : on vous demande de choisir, en faisant une croix sur la lettre qui le précède, celui qui correspond à ce que vous êtes, à ce que vous aimez, à ce que vous avez l'habitude de faire, de penser ou de sentir" (E-I/6)
- "Regarde ces boîtes [...] Il n'y en a pas deux qui pèsent la même chose. Eh bien j'aimerais que tu trouves la plus lourde ..." (TERMAN : 9 ans, 2°)
- "Lorsque j'aurai fini mon travail, je le cacherai, et je vous demanderai de faire exactement la même chose que moi" (TMI : 19)

Enfin, dans un type assez rare de formulation, c'est l'interlocuteur lui-même qui parle :

- "Je fais une croix sous le dessin qui ressemble le moins au dessin qui est dans le cadre tout seul" (NIV.INT : II)
- "Je souligne le mot qui ne va pas avec les autres" (NIV.INT : II)

Notons que nous n'avons relevé, dans l'ensemble des questionnaires d'enquête

de notre corpus qu'un seul exemple d'énonciative-consigne :

- "Nous allons résumer les qualités que vous recherchez pour le genre de voiture que vous voulez acheter. Pour cela, vous pouvez classer ces cartons en 3 groupes : d'un côté les points très importants, de l'autre les points peu importants, et au milieu, ceux qui sont entre les deux pour vous" (SOFRES 769 : 75)

Les énonciatives-réponse expriment ce que doit être la réponse de l'interlocuteur. Dans une première variété, celui-ci doit sélectionner sa propre réponse dans une liste de réponses possibles. Cette variété se rencontre dans les tests papier-crayon :

- "Si en me promenant je rencontre un crapaud ...
 - . je continue mon chemin
 - . je me sauve
 - . je l'observe
 - . je lui jette des pierres". (EAG 7/6 : 31)
- "Si je dois sortir seul la nuit ...
 - . je me munis d'une lampe
 - . je prends mes jambes à mon cou
 - . je me couvre bien
 - . je vais voir les vitrines illuminées" (EAG 7/6 : 38)

On la trouve également, avec une fréquence plus grande, dans les questionnaires d'enquête remplis directement par le répondant. Par exemple :

- "Depuis que vous exercez une activité professionnelle, des stages, cours, sessions de perfectionnement
 - . ont été suivis par vous
 - . sont actuellement en cours
 - . sont en projet précis
 - . sont souhaités mais non en voie de réalisation
 - . ne sont pas souhaités" (F.ING 69:IV 1)
- "A votre connaissance, votre établissement fait partie d'un groupe plus vaste (groupe industriel, grosse banque, administration ...) employant au moins 5.000 personnes" (CEREQ-SEU : 32)
- "Dans le travail, une difficulté majeure apparaît :
 - . j'essaie de la résoudre immédiatement moi-même
 - . je laisse les choses se décanter
 - . je contourne la difficulté
 - . j'en parle autour de moi
 - . j'en réfère à mon responsable direct
 - . je me procure des documents" (BNIST : 2)

- "Si vous aviez pu faire un autre métier, vous l'auriez choisi du type : médecin, artisan, technico-commercial, agriculteur, le même, travailleur social, administratif, autre" (BNIST : 38)
- "Si vous deviez imaginer l'Information Scientifique et Technique par un animal, ce serait : une hirondelle, une araignée, un éléphant, un singe, un chat, un serpent, un scorpion, un mille-pattes, une abeille, un pélican, autre". (BNIST : 60)

Une seconde variété d'énonciatives-réponse se présente comme une phrase inachevée, que l'interlocuteur doit terminer lui-même. Cette forme de question, connue sous le nom de *phrase à compléter*, est utilisée principalement dans les tests papier-crayon :

- "L'année dernière j'avais 1 an de plus que mon camarade. Cette année, il a que moi" (ECO 5 : 6,3)
- "La plupart de mes amis ne savent pas que je crains de ...". (STEIN : 22)
- "Je pourrais être certainement heureux si ...". (STEIN : 18)

Lorsqu'on la rencontre dans un test administré oralement, la suspension de la phrase peut être suivi d'une question, comme dans le second exemple ci-après :

- "Une personne qui parle beaucoup est loquace, une personne qui parle peu est ...". (TMI : 31)
- "Une personne qui se promenait dans la forêt s'est arrêtée tout à coup très effrayée et a couru chez le commissaire de police le plus voisin pour l'avertir qu'elle venait de voir à une branche d'arbre un un quoi ?" (TERMAN : 14 ans, 4^oc)

Enfin, certains questionnaires d'enquête présentent une forme qui peut s'apparenter aux énonciatives. Il s'agit de phrases sans verbe, substituées par économie dans la langue parlée à une proposition construite qui serait plus lourde [Wagner 1962 : 507]. Ce sont en général des questions concernant des faits aisés à mesurer ou à vérifier, et pour lesquelles on estime que les différences dans la formulation n'introduisent pas de biais dans les réponses. Lorsque le questionnaire est administré par un enquêteur, on peut en effet penser que celui-ci reformule la question pour la poser oralement. Exemples :

- "Domicile habituel" (REL 55 : 7)
- "Activités de votre mère durant votre enfance" (F.ING 69 : 18)
- "Nombre de personnes de l'établissement dans lequel vous travaillez" (CEREQ-SEU : 93)
- "Marque et modèle des voitures actuellement possédées par le foyer ?" (SOFRES 769 : 11)
- "Kilométrage parcouru avec cette voiture depuis que vous l'avez achetée ?" (SOFRES 769 : 14)

Par contre, la série de questions ci-dessous a pu éventuellement être énoncée telle quelle, grâce à l'avertissement qui la précède :

- "Je vais pour commencer vous demander quelques renseignements concernant votre voiture. Marque. Puissance [...]" (SOFRES 340 : 1a et b)

e) Pseudo-interrogatives à valeur de question

Cette dernière catégorie d'énoncés relevés dans les tests ou questionnaires est quelque peu paradoxale. Il est en effet possible de questionner en utilisant une fausse interrogative. Pour s'en convaincre, il suffit de considérer l'énoncé suivant :

- "Sais-tu dire : maman ?" (TERMAN : 3 ans, 6°)

S'il s'agissait d'une vraie interrogative, l'enfant pourrait se contenter de répondre : "oui, je sais", et l'on passerait alors à l'item suivant. Or, le jeune sujet comprend généralement fort bien que, ce que le psychologue lui demande, c'est de montrer qu'il est capable de répéter sans erreur les mots qu'on lui propose (ce qui lui sera d'ailleurs confirmé par les items suivants, qui sont des propositions injonctives). Le sens de cet énoncé est en réalité : "Dis : maman !". C'est une impérative exigeant un comportement linguistique, et assimilable de ce fait à une question (cf. [Katz 1964 : 136, n.19]).

L'interrogative injonctive à valeur de question se rencontre surtout dans les questionnaires d'enquête administrés par enquêteur. Par exemple :

- "Pouvez-vous me dire la marque d'essence qui y est vendue ?"
(SOFRES 340 : 6a)
- "Voudriez-vous me dire, pour chaque cas, la solution qui vous paraît la meilleure ?"
(IFOP 58-11 : 17)
- "Voudriez-vous, tout d'abord, me préciser le modèle de la voiture que vous avez achetée en octobre/novembre 1965 ?"
(SOFRES 967 : 1)
- "Parmi les explications suivantes, pouvez-vous citer les deux qui vous paraissent les plus justes ; les deux qui vous paraissent les plus injustes ?"
(IFOP 58-12 : 10)
- "Voici plusieurs points importants pour le confort d'une voiture. Voulez-vous les classer en trois groupes : les points très importants d'un côté, les points moins importants de l'autre, et au milieu, ceux qui sont moyennement importants"
(SOFRES 769 : 52)

1.1.2. Interrogatives directes

"Une interrogative est dite *directe* quand la question posée constitue une phrase indépendante. Cette phrase de termine toujours par un point d'interrogation (?), marque du ton sur lequel elle doit être prononcée" [Wagner 1962 : 525]. Comme l'on pouvait s'y attendre, les interrogatives directes sont majoritaires dans notre corpus, tant dans les questionnaires d'enquête que dans les épreuves psychologiques. Cette forme syntaxique est en effet la manière la plus simple de poser une question en français.

L'interrogation directe présente de nombreuses espèces. Nous donnons ci-après quelques exemples des variétés d'énoncés que nous avons relevés, en utilisant pour les présenter les principaux critères de classification auxquels les grammairiens ont habituellement recours : les marques de l'interrogation, la portée de l'interrogation, et la disjonctivité de l'énoncé.

a) Marques morphologiques de l'interrogation directe.

Au commencement de la première partie de ce chapitre, nous avons énuméré les quatre marques de l'interrogation selon les grammaires françaises usuelles : mot interrogatif, postposition du pronom sujet, particule *est-ce que*, intonation ; nous avons en outre signalé que certaines de ces marques pouvaient être présentes simultanément dans un même énoncé.

Si toutes les configurations se rencontraient dans la langue, on aurait $2^4 = 16$ formes d'interrogation directe. En fait, on n'observe que six des seize formes théoriquement envisageables. En effet, l'intonation (ou son substitut graphique, le point d'interrogation) est toujours présente, ce qui réduit à huit le nombre des formes théoriquement possibles. L'intonation caractéristique du questionnement, qui consiste à prononcer sur une note plus haute la dernière syllabe tonique d'un mot particulier (variable selon les cas : cf. [Wagner 1962 : 527-528]), peut par conséquent être considérée comme la marque de l'interrogation directe.

D'autre part, la périphrase interrogative *est-ce que* ne se rencontre jamais en même temps que la postposition du pronom sujet. Les linguistes expliquent ce phénomène à partir de l'histoire de la langue. Sous l'influence du tour germanique, les populations bilingues de la Gaule Franque auraient adopté l'ordre verbe-sujet pour l'interrogation [Lejeune 1937 : 13]. Au douzième siècle, "l'ordre verbe-sujet était une des constructions favorites du français d'alors, et le tour interrogatif qui reproduit cet ordre n'en était qu'un cas particulier" [Foulet 1921 : 262]. La tournure "est-ce que" avait alors un sens emphatique, et respectait les règles de concordance des temps. Au quinzième siècle, la langue tend à adopter comme ordre fondamental la succession sujet-verbe ; la valeur de la tournure *est-ce que*, devenue progressivement invariable, tient alors à ce que, "tout en donnant à la phrase un tour interrogatif, elle y maintenait l'ordre normal" [Foulet 1921 : 263]. De nos jours, s'il arrive à ce tour d'ajouter (dans la langue littéraire) certaine coloration affective [Wandruszka 1970 : 71-72], il ne modifie guère le sens de la question dans la langue courante. Tel n'est pas le cas pour les autres survivances de la tournure emphatique (par exemple : "serait-ce que ... ?"). C'est pourquoi, "si le français était une langue exotique et qu'elle fut notée pour la première fois par des missionnaires n'ayant aucun moyen d'en connaître la préhistoire, on lui

aurait attribué un mot interrogatif préposé *esk* '(ə)' [Tesnière 1959 : 210].

Cette incompatibilité réduit donc à six le nombre des formes générales de l'interrogation directe, selon que celle-ci est ou non introduite par un mot interrogatif (pronom, adjectif, adverbe), et selon qu'elle est marquée par l'intonation seule, la postposition du sujet, ou la particule interrogative *est-ce que*.

Les interrogations marquées par l'intonation seule sont assez rares dans notre corpus. Elles apparaissent comme des énonciatives suivies d'un point d'interrogation. Par exemple :

- "Vous comprenez ?" (TMI : 5)
- "Vous y êtes ?" (TMI : 26)
- "Par conséquent, aujourd'hui, vous êtes très satisfait, plutôt satisfait, ou mécontent ?" (IFOP 58-11 : 3b)

Les interrogations marquées par la postposition du sujet peuvent prendre deux formes, selon que le sujet est un pronom personnel ou l'un des pronoms *on* ou *ce* (inversion simple), ou bien un autre pronom indéfini, un pronom possessif, ou un nom (inversion complexe) [Mauger 1968 : 299]. L'inversion simple est certainement la forme la plus fréquente dans notre corpus (avec ou sans mot interrogatif). Sans mot interrogatif préposé, les énoncés sont de la forme :

- "Etes-vous sujet à la rêverie ?" (BERN 35 : 2)
- "Avez-vous une tendance à être lent et réfléchi dans vos actions ?" (CAT 16PF : 4)
- "Croyez-vous qu'il y ait en France un danger de dictature ?" (IFOP 58-9 : 22a)
- "A votre avis, a-t-on intérêt à acheter les versions les plus chères ?" (SOFRES 769 : 48)
- "Parmi les personnes de même rang hiérarchique que vous, avec qui vous êtes amenée à travailler, se trouve-t-il uniquement des hommes, surtout des hommes, des hommes et des femmes, surtout des femmes, uniquement des femmes ?" (F.ING 69 : III 10)

L'*inversion complexe* se caractérise par la présence d'un pronom personnel postposé au verbe, servant à reprendre un sujet nominal placé en tête de la phrase [Wagner 1962 : 529]. Par exemple :

- "Votre famille était-elle très favorable ou favorable, ou plutôt défavorable ou très défavorable à votre mariage ?"
(INED 081 : fiançailles 5)
- "L'un des conjoints (ou les deux) a-t-il reçu une dot ?"
(INED 081 : mariage 6)
- "Cette voiture sera-t-elle neuve ou d'occasion ?"
(SOFRES 769 : 38)
- "Cette voiture remplacera-t-elle une voiture actuellement possédée par le foyer ?"
(SOFRES 769 : 43)
- "Dans quel secteur votre mari exerce-t-il (exerçait-il) son activité ?"
(F.ING 69 : V 1)

Les interrogations marquées par "est-ce que" sont de la forme :

- "Est-ce qu'il arrive que vos émotions vous fassent passer pour quelqu'un de déraisonnable, même à vos propres yeux ?"
(CAT 16PF : 93)
- "Est-ce que la radio vous manque ?"
(ORTF 65 : 3bis)
- "Si l'on doit pratiquer une politique de participation dans les entreprises, à votre avis, est-ce que cela doit plutôt être une participation des cadres aux décisions importantes concernant l'entreprise ?"
(IFOP 69-5 : 11)
- "Dans l'ensemble, et compte-tenu de tous vos déplacements, est-ce que vous considérez qu'au cours de l'année vous roulez ...
. surtout en ville
. surtout sur route
. autant l'un que l'autre ?"
(SOFRES 340 : C)

Les énoncés du type *mot interrogatif + intonation seule* sont la seule forme préconisée par les grammaires lorsque le mot interrogatif est un pronom sujet du verbe (ou un adjectif rattaché au substantif sujet) :

- "Lequel des objets énumérés à droite n'est pas de la même espèce que les deux autres ?"
(CAT 16PF : 182)

- "Qui s'occupera du choix et de l'achat de cette voiture ?"
(SOFRES 769 : 9)
- "Vous personnellement, laquelle de ces deux solutions vous paraît la meilleure ?"
(SOFRES 769 : 61)
- "A votre avis, quelles sortes de gens ont voté non ?"
(IFOP 58-11 : 10)

Dans les autres cas, l'énoncé est au style parlé (et le mot interrogatif n'est pas nécessairement placé en début de phrase) :

- "C'est une image de quoi ?" (TERMAN : images 7 ans)
- "Dans quelle branche d'activité vous exerceriez ?"
(OCDE-IUT : 52a)
- "Cela vous aurait amené à envisager quel modèle de voiture neuve ?"
(SOFRES 967 : 3bis)
- "Pourquoi cette solution vous paraît préférable ?"
(SOFRES 659 : 37)
- "Dans quelle région vous comptez travailler ?"
(OCDE-IUT : 50e)

Les énoncés du type *mot interrogatif + postposition du sujet* que nous avons relevés dans notre corpus ne présentent que l'inversion simple (sans toutefois qu'il y ait d'impossibilité grammaticale pour l'inversion complexe à une exception près [Mauger 1968 : 381, § 820]). Par exemple :

- "Quel fut le premier succès de Charles pendant sa tournée aux Etats-Unis ?"
(TMI : 10a)
- "De quoi a-t-il peur ?" (FABLES : V)
- "Quel sera, à votre avis, le résultat de ce référendum dans l'ensemble de la France métropolitaine ?"
(IFOP 58-9 : 17)
- "De quel parti ou courant politique vous sentez-vous le plus proche ?"
(GEMAS 76 : 34)
- "Sur quelle chaîne suivez-vous généralement les actualités télévisées de la soirée ?"
(ORTF 65 : 39)

- "Pourquoi avez-vous entrepris des études supérieures ?" (F.ING 69 : VI 5)
- "Comment vous êtes-vous connus ?" (INED 081 : fiançailles 1)

Les interrogations de la forme *mot interrogatif* + "est-ce que" sollicitent le plus souvent la définition ou la désignation d'un objet ou d'un concept. Par exemple :

- "Qu'est-ce que la justice ?" (BS Z : 42)
- "Qu'est-ce que c'est qu'une chaise ?" (BS Z : 13-34)
- "Qu'est-ce que c'est que cela ?" (TERMAN : 3 ans, 2°)
- "Qu'est-ce qui manque à cette figure-là ?" (BS Z : 21)
- "Qu'est-ce qu'il a de changé ?" (FABLES : VI)
- "Qu'est-ce que le général de Gaulle a fait de mieux depuis qu'il est au pouvoir ?" (IFOP 58-11 : 37)
- "Qu'est-ce qui, à votre avis, faciliterait la reprise d'activité après une interruption ?" (F.ING 69 : III 19)

b) Portée de l'interrogation

La notion de portée de l'interrogation est une notion fondamentale, tant sur le plan syntaxique que sur le plan sémantique. La quasi-totalité des linguistes qui traitent de l'interrogation distingue deux grandes classes de questions, selon que l'interrogation porte sur tout ou partie de l'énoncé. Mais, bien qu'il y ait accord sur ce que recouvre cette dichotomie, chaque linguiste apporte et son propre vocabulaire, et sa propre définition des deux espèces d'interrogatives (on trouvera un relevé impressionnant des termes utilisés dans [Jespersen 1924 : 432]). Les uns s'appuient essentiellement sur le *sens* de la question :

- "Pour les grammairiens allemands, il y a deux types d'interrogatives : la *Entscheidungsfrage* ou interrogation globale et la *Ergänzungsfrage*, interrogative ne portant que sur un seul élément du groupe verbal" [Philipp 1974 : 103].

- "Les interrogations totales sont celles qui portent sur le fait central lui-même : « Viens-tu ? » [...] Les interrogations partielles sont celles qui ne portent que sur une circonstance du fait : « *Quand* viens-tu ? - *D'où* viens-tu ? »" [Damourette 1931 : IV, 309].

- "L'interrogation est *totale* quand elle porte sur l'ensemble de la phrase (thème ou sujet + prédicat). [...] L'interrogation est *partielle* quand elle porte précisément sur un des termes de la phrase" [Wagner 1962 : 526-527].

- "*Interrogation particulière* : Nous appellerons ainsi l'interrogation qui n'intéresse qu'un terme de la phrase [...] *Interrogation générale* : Nous appellerons ainsi l'interrogation qui intéresse la phrase entière". L. LE BAULT cité par [Tesnière 1959 : 204].

- "L'énoncé interrogatif peut être une requête limitée portant sur l'identité d'un individu ou objet inclus dans un procès, ou bien sur le moment, le lieu, le motif d'une action qui, de toute façon, est affirmée [...] Mais l'énoncé interrogatif peut être aussi une enquête globale sur l'existence ou la non-existence d'un fait" [Collart 1975 : 106-107].

D'autres linguistes prennent comme point de départ *la phrase énonciative* qui correspond à l'interrogative :

- "Il existe deux types d'interrogatives : le premier est de la forme : *Did he say that ?* [...] et le second de la forme : *What did he say ?* [...]. Dans le premier cas, c'est toujours un nexus qui est l'objet de la question : en effet, le locuteur cherche à savoir s'il est correct ou non de mettre en relation tel sujet et tel prédicat. On est donc en droit de parler dans ce cas de « question nexale ». Dans l'autre cas, l'interrogation porte sur une quantité inconnue, au sens où on l'entend dans les équations algébriques ; [...] on peut appeler « question sur x » la question qui vise à établir la nature de x " [Jespersen 1924 : 431-432].

- "Chaque phrase énonciative donne lieu à autant d'interrogations nucléaires qu'elle comporte de nucléus, parce que chaque nucléus est susceptible d'être révoqué en doute. C'est pourquoi, dans toute interrogation nucléaire, il y a au moins un nucléus sémantiquement vide [...]. Mais il existe, en face des mêmes phrases énonciatives, des interrogations d'un type tout différent, et qui se reconnaissent aisément à ce que tous les nucléus sont pleins. [...] C'est donc sur la connexion de nucléus pleins que porte ici l'interrogation. C'est pourquoi nous donnerons à ce type d'interrogation tout différent du précédent, le nom d'*interrogation connexionnelle*". [Tesnière 1959 : 203].

Certaines classifications se fondent essentiellement sur le *contenu de la réponse*:

- "En ce qui concerne les questions, il y a une différence de schéma tonal entre une question à laquelle on répond par oui ou par non, telle que *Il est dix heures ? Avez-vous vu telle pièce ?* et une question à compléter ("supplément question") à laquelle on doit répondre par un mot ou par une phrase particulière,

exemple : *Quelle heure est-il ?* ou, *Qui a vu la pièce ?* avec une montée moindre à la fin" [Bloomfield 1961 : 109].

- "Question-réponse si/no/forse ω può darsi. La question porte alors sur la phrase comme ensemble [...] *Phrases appelant en réponse une phrase de structure identique à SN* [syntagme nominal] *différemment réalisé*. La question porte alors sur une seule partie de la phrase, nommément sur l'un des SN constituants de phrase, de GV [groupe verbal] ou de SP [syntagme prépositionnel]" [Genot 1973 : 100-101].

Enfin, on rencontre également des *typologies mixtes*, utilisant à la fois la forme de la réponse et la forme de l'énoncé interrogatif :

- "Questions entraînant une réponse par *Yes/No* : elles ne font pas appel à un élément interrogatif. Dans ce cas, la question porte sur la totalité de l'énoncé [...] Questions de type « *Wh-* » : elles utilisent les interrogatifs *who(m)*, *whose*, *which*, *what*, *when*, *where*, *why* (groupe auquel on intègre *how*, bien qu'il ne soit pas morphologiquement de type *wh-*). De telles questions sont incidentes à une partie seulement de l'énoncé" [Tellier 1974 : 106-107].

- "Il existe en français deux types principaux de phrases interrogatives : (1) *Pierre viendra-t-il ?* (2) *Qui est venu ?* [...] Dans le cas de la phrase (1) la réponse est *oui* ou *non* (ou diverses autres formules qui relèvent du même principe) ; dans le cas de la phrase (2) la réponse est un syntagme nominal correspondant à celui qui est formé avec la forme *qui*" [Dubois 1970 : 208].

Les *interrogations totales* ne comportent pas de mot interrogatif. Elles appellent une réponse par *oui* ou par *non*. Nous en avons vu divers exemples au paragraphe précédent. Elles sont de la forme :

- "Vous faites-vous facilement de nouveaux amis ?" (BERN 35 : 62)
- "Sa maman lui demande de le donner, il est libre, le donnera-t-il ?"
(FABLES : VII)
- "Y a-t-il dans votre foyer un poste de radio en état de fonctionner ?"
(ORTF 65 : 3)
- "Si le projet de Constitution n'est pas adopté, à votre avis risque-t-on la guerre civile ?"
(IFOP 58-9 : 21)

Les *interrogations partielles* peuvent prendre des formes très variées, selon la nature du mot qui les introduit (pronom, adjectif, adverbe), et selon la fonction que ce mot remplit dans la phrase.

Lorsque le mot interrogatif est un *pronom*, celui-ci peut pratiquement remplir toutes les fonctions dans la phrase. Par exemple :

- Sujet : - "*Qu'est-ce qui vous plaisait surtout dans cette Constitution ?*"
(IFOP 58-9 : 7)
- "*A votre avis, qu'est-ce qui vous fera prendre votre décision ?*"
(IFOP 58-9 : 14b)
- "*Qu'est-ce qui vous gêne particulièrement ?*" (SOFRES 659 : 9)

Attribut :

- "*Qu'est-ce que l'obéissance ?*" (TMI : 2)

Complément d'objet :

- "*Que se passe-t-il donc chez mon voisin ?*" (BS Z : 48)
- "*Qui avez-vous consulté avant de prendre la décision d'entrer à l'I.U.T. ?*" (OCDE-IUT : 28)
- "*De ces deux pages de modèles, laquelle préférez-vous ?*"
(SOFRES 686 : 51)

Complément déterminatif (d'un adjectif, d'un verbe) :

- "*Voici deux opinions contraires que nous avons entendues chez les personnes que nous avons interrogées. Avec laquelle êtes-vous le plus d'accord ?*" (SOFRES 769 : 46)
- "*Avec qui étiez-vous ?*" (SOFRES 571 : 22d)
- "*A quoi attribuez-vous cette possibilité ?*" (GEMAS 76 : 14)

Complément circonstanciel :

- "*En quoi se ressemblent surtout : agriculture et industrie ?*"
(TMI : 25a)
- "*En quoi la nouvelle Assemblée Nationale ne reflète-t-elle pas l'opinion du pays ?*" (IFOP 58-12 : 8b)

- "En ce qui concerne le choix de cette voiture, dans lequel des cas suivants vous situez-vous ?" (SOFRES 769 : 41)

De même, l'adjectif interrogatif peut soit être attribut, soit être épithète d'un substantif remplissant n'importe quelle fonction.

Sujet : - "Quelle solution vous paraît préférable ?" (SOFRES 571 : 37)

Attribut :

- "Quelle est la principale différence entre féliciter et flatter ?" (TMI : 18c)
- "Quelle est la fraction de la ligne à droite qui n'appartient pas à la même catégorie que les autres ?" (CAT 16PF : 106)
- "Quel était le degré d'instruction de votre père ?" (IFOP 58-9 : 43)
- "Quelle était, à cette époque, l'opinion de votre entourage sur les études supérieures pour les femmes ?" (F.ING 69 : VI 6)
- "Vous même, avant de vous décider pour la ... (modèle de voiture achetée récemment), quels étaient les différents modèles de voitures auxquels vous avez d'abord pensé ?" (SOFRES 967 : 3)

Complément d'objet :

- "Quelle direction auriez-vous devant vous si vous aviez le nord à votre droite ?" (TMI : 17a)
- "Quel type d'emploi occupez-vous ?" (F.ING 69 : III 3)
- "Quel instrument de musique écoutez-vous avec le plus de plaisir ?" (ORTF 65 : 22)
- "S'il arrivait que le parti ou la tendance que vous préférez ne puisse présenter un candidat, quelle solution choisiriez-vous ?" (IFOP 58-11 : 30a)

Complément déterminatif (d'un nom, d'un adjectif, d'un verbe) :

- "Au(x) modèle(s) sport de quelle(s) marque(s) songez-vous plus particulièrement ?" (SOFRES 967 : 23)
- "A quelles marques étrangères, par contre, êtes-vous le moins favorable ?" (SOFRES 967 : 17a)
- "A quel secteur appartient l'entreprise qui vous emploie ?" (F.ING 69 : III 1)
- "Pour quels partis accepteriez-vous alors de voter ?" (IFOP 58-11 : 30b)
- "Toujours s'il venait à disparaître, par quel magazine le remplaceriez vous ?" (SOFRES 686 : 29)
- "Dans quelles rues êtes-vous d'avis d'interdire la circulation ?" (SOFRES 659 : 26)

Complément circonstanciel :

- "En quelle année eut-il lieu ?" (TMI : 10b)
- "En semaine, à partir de quelle heure ces enfants regardent-ils la TV dans la soirée ?" (ORTF 65 : 33)
- "A quel endroit l'avez-vous acheté ?" (SOFRES 571 : 24)
- "Dans quels domaines parmi les suivants souhaitez-vous que l'Etat dépense moins ?" (IFOP 69-5 : 5b)
- "Pour quelle raison principale ne cherchez-vous pas d'emploi ?" (CEREQ-SEU : 25)

Les adverbess interrogatifs devraient en principe n'introduire que des questions portant sur le complément circonstanciel. Il faut cependant faire une exception pour deux d'entre eux : *combien*, et *comment*. L'adverbe de quantité *combien* peut déterminer un substantif [Wagner 1962 : 404] remplissant les fonctions de sujet [Wagner 1962 : 533] ou de complément d'objet (nous n'avons pas relevé d'énoncés dans lesquels le substantif complément de *combien* était complément déterminatif) :

- "Combien de spectateurs l'ont admiré ?" (TMI : 10c)
- "Actuellement, combien de personnes dans votre foyer ont le permis de conduire ?" (SOFRES 769 : 6)
- "Le plus souvent, combien de personnes au total transportez-vous à l'arrière ?" (SOFRES 769 : 59)
- "Combien de temps consacrez-vous en moyenne chaque semaine à la lecture ?" (BNIST : 59)
- "Finalement, combien de partis croyez-vous qu'il devrait y avoir ?" (IFOP 58-9 : 25a)

L'adverbe de manière *comment* nous paraît remplir dans certains cas une fonction d'attribut (avec un verbe principal pouvant introduire une telle fonction. cf. [Bloch 1945 : 199] ; [Mauger 1968 : 57-59]), bien que cette possibilité ne soit jamais envisagée par les grammaires usuelles que nous avons consultées. On s'en convaincra en envisageant les réponses attendues aux questions suivantes :

- "Comment étiez-vous classé au point de vue travail scolaire en classe de 6e ?" (OCDE-IUT : 6)
- "Dans l'ensemble, comment trouvez-vous les émissions de radio destinées aux enfants de 6 à 15 ans ?" (ORTF 65 : 13)
- "Comment voyez-vous l'avenir de votre entreprise ?" (F.ING 69 : VIII 5)
- "Comment croyez-vous que va voter l'Algérie ?" (IFOP 58-9 : 19)

Par ailleurs, on retrouve ces deux adverbes dans l'ensemble des adverbes interrogatifs, pour introduire des questions portant sur le complément circonstanciel du verbe principal :

- "Combien pèsera maintenant le tout ?" (TERMAN : adulte moyen, supplément 2b)
- "Comment traversa-t-il le Niagara pour la seconde fois ?" (TMI : 10d)
- "Comment êtes-vous entrée dans la société qui vous emploie ?" (F.ING 69 : VII 12)
- "Comment à votre avis, serait-il souhaitable de désigner ceux qui commandent ?" (GEMAS 76 : 84)

- "Pourquoi pardonne-t-on plutôt une mauvaise action exécutée avec colère qu'une mauvaise action exécutée sans colère ?"
(BS Z : 39)
- "Pourquoi n'avez-vous pas acheté une voiture d'occasion ?"
(SOFRES 769 : 21)
- "Quand avez-vous prévenu votre famille ?"
(INED 081 : fiançailles 4)
- "Depuis quand avez-vous pris une décision dans ce sens ?"
(IFOP 58-9 : 13b)
- "Où a-t-elle été vendue ?"
(SOFRES 769 : 31)
- "Où suivez-vous ce genre d'émission ?"
(ORTF 65 : 26)

La notion de portée de l'interrogation se fonde bien, nous semble-t-il, sur des caractères morphologiques facilement identifiables, et dépourvus d'ambiguïté. En outre, elle correspond à des différences sémantiques certaines. Cependant, si l'on désire affiner la classification des interrogatives partielles, il paraît difficile, en l'état actuel des sciences du langage, de s'en tenir aux seuls critères morfo-syntaxiques. En effet, comme nos exemples le montrent abondamment, il n'y a pas de correspondance simple entre la forme et la signification des interrogatives partielles ; ceci est d'ailleurs facile à vérifier en reformulant une des questions ci-dessus de toutes les manières possibles qui n'en altèrent pas le sens. D'autre part, si l'on utilise des caractères sémantiques (ce qui serait en contradiction avec la perspective adoptée dans ce premier chapitre), on aboutit soit à une dichotomie assez élémentaire des interrogations partielles (portant sur les modalités de l'action/portant sur l'identité ou la qualité du sujet ou du complément [Mauger 1968 : 380-381]), soit à une multiplication des catégories sémantiques excédant facilement la dizaine (cf. par exemple [Brunot 1926 : 108-848, *passim*]). Enfin, comme nous le verrons plus loin, certains auteurs dénie toute existence, sur le plan sémantique, à l'interrogation générale (cf. [Zuber 1972]).

c) Énoncé simple ou disjonctif

Ce troisième critère de classification des interrogatives est moins fréquemment utilisé par les grammairiens. Il présente cependant un grand intérêt pour l'étude des tests et des questionnaires, en raison de la parenté qui existe entre cette dichotomie et la distinction que font les psychologues entre question ouverte et question fermée.

La présentation la plus rigoureuse de cette dichotomie se trouve à notre connaissance dans la *Syntaxe grecque* de BIZOS. L'*interrogation disjonctive* est marquée en grec ancien par le couple adverbe-conjonction "πότερον...ἢ ... ;" ("est-ce que ... ou bien ... ?" ; cf. latin : "utrum ... an ... ?"). Exemple : "Te semble-t-il vouloir imiter la nature elle-même ou les oeuvres des artisans ?" (PLATON, *République*, 598a) [Bizos 1949 : 115]. Cette forme d'énoncé correspond au cas dans lequel le locuteur propose explicitement à l'interlocuteur un choix entre deux ou plusieurs termes. Elle s'oppose à l'*interrogation simple* [Bizos 1949 : 112], dans laquelle le locuteur ne présente pas d'alternative (et qui peut prendre des formes variées).

Il est intéressant de confronter cette définition strictement morphologique avec une autre dichotomie fondée sur la manière de répondre à la question : "On peut distinguer deux formes d'interrogation en chinois : *exclusive*, où le locuteur pose un choix, le plus souvent une alternative, auquel on répondra ceci/ou/cela, oui/ou/non ; *substitutive*, où la question porte sur un indéfini, qui sera remplacé dans la réponse par une forme lexicale « pleine » " [Alleton 1973 : 31]. Malgré l'apparence, les deux dichotomies ne coïncident pas. Certes, on est tenté de considérer comme synonymes d'une part *disjonctive* et *exclusive* ("est-ce X qui est là ou Y ?"), d'autre part *simple* et *substitutive* ("Qui est là ?"). Il n'est cependant pas difficile d'imaginer des contre-exemples, soit de disjonctive-substitutive ("Qui est là : X, ou Y ?), soit de simple-exclusive ("Est-ce que X est là ?"). Certains grammairiens toutefois pourraient faire remarquer que, dans l'énoncé disjonctif-substitutif, l'alternative est une apposition au mot interrogatif (comme si le pronom interrogatif avait un complément déterminatif sous forme d'alternative : "Qui, de X ou de Y, est là ?") ; et que par conséquent il s'agit en fait d'une interrogation simple-substitutive. On trouve d'autre part la thèse selon laquelle l'énoncé simple-exclusif (à réponse *oui* ou

non) est en réalité une abréviation d'un énoncé disjonctif-exclusif ("Est-ce que X est là, ou non ?") [Langacker 1972 : 188].

A ce stade de notre analyse, seule nous importe la forme sous laquelle les énoncés se présentent effectivement ; nous réserverons donc notre jugement jusqu'à la seconde partie de ce chapitre, et nous nous en tiendrons à la dichotomie *simple/disjonctif*, en considérant le mode de réponse comme un procédé de classement interne à chaque catégorie.

Les énoncés *simples* de type substitutif sont tous des interrogations partielles. Par exemple :

- "Si les mots étaient bien placés, quelle phrase est-ce que cela ferait ?" (BS Z : 43)
- "Qu'y-a-t-il de bête sur cette image ?" (TMI : 4)
- "Si j'achetais pour 20 francs de bonbons et que je donne 50 francs, combien me rendrait-on ?" (TERMAN, 9 ans : 3^{oa})
- "En quelle année l'avez-vous achetée ?" (SOFRES 340 : 1c)
- "Parmi ces stations de radio, lesquelles écoutez-vous ?" (ORTF 65 : 4)
- "De quelle somme devrait, selon vous, disposer un couple avec deux enfants pour n'avoir pas trop de problèmes d'argent ?" (GEMAS 76 : 45)

Ceux de type exclusif sont des questions à réponse par *oui* ou *non* :

- "Arrive-t-on parfois à vous abuser ?" (BERN 35 : 77)
- "Aimeriez-vous avoir le genre de situation qui offre le changement les voyages, la variété (malgré d'autres inconvénients) ?" (CAT 16PF : 60)
- "A-t-on déjà possédé une voiture dans votre foyer ?" (SOFRES 769 : 2)
- "Y-a-t-il, dans votre circonscription, un candidat qui correspondait à vos opinions, au premier tour ?" (IFOP 58-12 : 14)

Notons enfin que certains énoncés simples (non disjonctifs) peuvent donner l'impression de proposer une alternative à l'interlocuteur, en raison de l'emploi de la conjonction *ou*. Si cet emploi n'introduit aucune ambiguïté dans les deux premiers exemples ci-après (interrogations partielles), il n'en est pas de même pour les deux autres (interrogations totales) ; seul, l'examen de la codification des réponses (figurant sur le questionnaire) permet d'établir avec certitude qu'il s'agit bien d'interrogations non disjonctives.

- "Quel devrait être ce ou ces partis ?" (IFOP 58-9 : 25b)
- "De quel instrument jouez-vous ou avez-vous joué ?" (ORTF 65 : 23)
- "Est-ce qu'un évènement, ou quelqu'un, ou quelque chose vous a aidé à prendre une décision ?" (IFOP 58-9 : 13c)
- "Depuis votre venue ici, vous est-il arrivé d'acheter un réfrigérateur, une télé ou une machine à laver ?" (SOFRES 571 : 23)

Les énoncés disjonctifs de type substitutif sont très rares dans notre corpus. En voici un exemple :

- "Pour parvenir aux résultats économiques et sociaux que vous souhaitez vous, à qui de M. POMPIDOU ou de M. POHER, faites-vous le plus confiance ?" (IFOP 69-5 : 8)

Parmi les énoncés disjonctifs de type exclusif, certains linguistes distinguent, sous le nom d'*interrogation multiple* [Brunot 1926 : 492] ou *interrogation alternative* [Damourette 1931 : IV 369], la succession de deux phrases interrogatives complètes s'excluant mutuellement (et pouvant n'être pas de la même forme) ; comme par exemple :

- "Est-ce que :
 - a) vous restez indifférent si les gens ne montrent pas qu'ils vous apprécient suffisamment ?
 - b) ou bien êtes-vous démoralisé si vous n'êtes pas assez apprécié ?" (CAT 16PF : 95)
- "Après un acte de colère (ou, si vous ne vous mettez jamais en colère, après avoir subi un affront), êtes-vous immédiatement réconcilié (tout à fait comme auparavant, sans plus y songer) ?
Ou restez-vous quelque temps de mauvaise humeur ?
Ou êtes-vous difficile à réconcilier (rancune persistante) ?" (BERGER : 53)

- "Poursuiviez-vous des études ou suiviez-vous une session de formation en 1975-76 ?" (CEREQ-SEU : 16)
- "Dans l'ensemble, est-ce que vous pourriez dire, lorsque vous participez à une conversation sur le référendum que c'est vous qui écoutez davantage ou que c'est vous qui essayez davantage de convaincre votre entourage ?" (IFOP 58-9 : 9)
- "Est-ce que vous prenez indifféremment ... (marques citées à Q.12) ou y en a-t-il une que vous prenez de préférence aux autres ?" (SOFRES 340 : 13)
- "Si un tel centre existait, pensez-vous que vous iriez y faire les courses que vous allez faire d'habitude à M. ou bien iriez-vous quand même faire vos achats à M. ?" (SOFRES 571 : 43)

Toutefois, sous sa forme la plus fréquente, l'interrogation disjonctive-exclusive ne propose à l'interlocuteur qu'un choix entre deux ou plusieurs termes (adverbes, adjectifs, substantifs, etc.). Ces termes sont en principe coordonnés par *ou* ou *ou bien* :

- "Avez-vous pris votre décision facilement ou non ?" (IFOP 58-11 : 9)
- "Cette opinion est-elle vraie ou fausse ?" (INED 081 : opinions 7a)
- "Ce parti conseillera-t-il de voter oui ou non ?" (IFOP 58-9 : 40d)
- "A votre avis, est-il très important, ou sans importance qu'une jeune fille se garde jusqu'au mariage ?" (INED 081 : opinions 7b)
- "Dans votre famille, est-ce qu'on a plutôt la même position ou pas la même position que vous sur le projet de constitution ?" (IFOP 58-9 : 8a)
- "Lorsque vous remplacerez votre ... [voiture], pensez-vous que vous allez probablement reprendre le même modèle, ou bien reprendre un autre modèle ?" (SOFRES 967 : 20)

Mais on trouve également l'ensemble des réponses proposées simplement énumérées sans conjonction de coordination :

- "Préfèreriez-vous être :
 - a) ingénieur ?
 - b) professeur de sociologie ?" (CAT 16PF : 156)

- "Préfèreriez-vous un emploi à temps complet, mi-temps, horaire aménagé, journée continue, domicile (?)" (F.ING 69 : IX 4)
- "Au niveau de votre entreprise, ne pas s'informer sur les sciences et techniques, représente-t-il une erreur, un danger, un risque ?" (BNIST : 48)
- "Aimez-vous beaucoup, un peu, pas du tout les chansons ?" (ORTF 65 : 24)
- "La dernière fois que vous avez acheté de l'essence, était-ce ...
 - dans un garage
 - sur la piste d'une station-service
 - au bord du trottoir, à une pompe ?" (SOFRES 340 : 7b)
- "Dans l'ensemble, trouvez-vous ces émissions régionales excellentes, très bonnes, moyennes, mauvaises ?" (ORTF 65 : 7)

1.1.3. Interrogatives indirectes

"Une interrogative est dite *indirecte* quand elle est subordonnée à un terme principal qui lui sert d'appui" [Wagner 1962 : 526]. "Une interrogation indirecte est une question posée au moyen d'une proposition qui contient l'interrogation et d'une subordonnée qui contient l'objet de cette interrogation" [Bloch 1945 : 225]. Toutes les grammaires usuelles françaises présentent les interrogations indirectes comme une transposition d'interrogatives directes ayant le même sens. Nous partirons donc des caractéristiques des interrogatives directes pour analyser les marques de l'interrogation indirecte, et décrire les grandes catégories d'interrogatives indirectes. Nous examinerons ensuite la nature de la phrase introduisant l'interrogative subordonnée.

a) Les marques morphologiques de l'interrogation indirecte.

L'interrogative indirecte commence toujours par un mot interrogatif, et ne prend jamais l'intonation interrogative (ou son équivalent graphique, le point d'interrogation). Si l'on considère l'interrogative indirecte comme une dérivation de l'interrogative directe, les principales règles de passage au style indirect peuvent être résumées comme suit (cf. [Mauger 1968 : 148-149, 300, 311-312, 379-381]) :

- ① L'intonation interrogative est supprimée au style indirect.
- ② S'il n'y a pas de mot interrogatif au style direct, l'interrogative indirecte commence par *si*.
- ③ S'il y a un mot interrogatif au style direct, il est conservé sans changement, sauf si ce mot est un pronom neutre ; dans ce cas, il devient *ce qui* s'il est sujet, *ce que* s'il est complément direct d'objet ou attribut, et il ne change pas s'il est complément indirect ou déterminatif.
- ④ La particule *est-ce que* peut éventuellement être conservée au style indirect dans la langue parlée, sauf lorsque l'interrogative directe ne comporte pas de mot interrogatif. Lorsqu'elle est combinée avec le pronom neutre sujet, attribut, ou complément d'objet direct ("qu'est-ce qui", "qu'est-ce que"), elle disparaît également (cf. règle 3).
- ⑤ S'il y a postposition du sujet dans l'interrogative directe, l'ordre sujet-verbe est rétabli dans le style indirect. Toutefois, l'inversion simple peut être utilisée au style indirect dans deux cas :

- l'interrogative indirecte a pour sujet un nom, et elle est introduite par *ce que* en fonction de complément direct d'objet ("je me demande ce que fait cet homme"), ou *quoi* en fonction de complément indirect d'objet ("je me demande à quoi pense cet individu") ;

- l'interrogative indirecte a pour sujet un nom, un pronom possessif, un pronom démonstratif autre que *ce*, ou un pronom indéfini autre que *on* ; elle est introduite par l'un des adverbes interrogatifs *où*, *quand*, *comment* ("je me demande où est le mien" ; "... quand reviendra celui-ci" ; "... comment font certaines").

Ces distinctions morpho-syntaxiques permettent de rendre compte de la portée des interrogatives indirectes. Toutes les interrogatives indirectes *totales* commencent par la conjonction *si*. Exemples d'interrogation totale non disjonctive :

- "Parmi les matières enseignées en secondaire , pourriez-vous dire s'il y en avait que vous préféreriez ?" (OCDE-IUT : 8)

- "Est-ce que je peux vous demander enfin si vous vous intéressez à la politique ?" (IFOP 58-12 : 41)

Exemples d'interrogation totale disjonctive (multiple, avec *ou*, sans *ou*) :

- "Pour chaque point que je vais vous citer, dites-moi si vous estimez que les difficultés rencontrées étaient principalement la faute du Gouvernement ou si c'était des difficultés normales ?" (IFOP 69-5 : 3)
- "Voudriez-vous me donner votre opinion, en disant si vous vous estimez d'accord ou non ?" (IFOP 58-11 : 43)
- "On nous a donné un certain nombre d'opinions sur ... (localité). J'aimerais que vous me disiez si, à votre avis, c'est vrai ou faux ?" (SOFRES 571 : 5)
- "Voici quelques opinions concernant les rapports de la TV avec la vie familiale et les enfants. Dites-moi si, pour chaque opinion, vous êtes tout à fait d'accord, à peu près d'accord, pas d'accord du tout" (ORTF 65 : 29)

Toutes les interrogatives indirectes *partielles* commencent par un pronom, un adjectif, ou un adverbe interrogatif. L'inversion du sujet y est très rare (cf. règle 5) :

- "Dites-moi en quoi se ressemblent : serpent, vache, moineau" (TMI : 11)

Exemples d'interrogation partielle introduite par un pronom :

- "Pèse-les et dis-moi laquelle est la plus lourde" (TERMAN : 5 ans, 1°)
- "J'aimerais que tu écoutes bien, et que tu me dises ce qu'il y a de bête" (TERMAN : 10 ans, 2°)

- "Pouvez-vous préciser ce que vous avez fait entre l'obtention de ce diplôme et votre entrée à l'IUT ?" (OCDE-IUT : 32)
- "Si la décision finale était d'abattre certaines rues ou quartiers, voulez-vous me dire lesquels, parmi ceux que je vais vous citer, seraient à abattre, lesquels seraient à conserver ?" (SOFRES 659 : 38)
- "Je vais vous demander maintenant dans laquelle de ces catégories de revenus entre votre foyer ..." (SOFRES 967 : RS 11)
- "Parmi ce qu'on entend dire à ce sujet, pouvez-vous me dire sur quoi vous êtes d'accord et sur quoi vous n'êtes pas d'accord ?" (IFOP 69-5 : 7)

Exemples d'interrogation partielle introduite par un adjectif :

- "Tu vois ces lignes. Dis-moi quelle est la plus grande ..." (BS Z : 6)
- "Maintenant, je vais te lire une fable. Tu écouteras bien et quand j'aurai fini, je te demanderai de me dire quelle leçon cette fable nous enseigne" (TERMAN : 12 ans, 5^o)
- "Indiquez dans quelle catégorie d'âge vous vous situez" (OCDE-IUT : 3)
- "Voulez-vous me dire à l'aide de cette liste, par quels moyens d'information votre attention a été attirée sur la [modèle cité à Q3] ?" (SOFRES 967 : 4)

Exemples d'interrogation partielle introduite par un adverbe :

- "Regardez bien et dites-moi comment la phrase doit être lue" (TMI : 7)
- "Supposez que vous ne l'ayez pas encore trouvé, montrez-moi partout où vous iriez pour le chercher" (TMI : 12)
- "Pouvez-vous me dire environ combien vous voulez dépenser pour l'achat de cette voiture ?" (SOFRES 769 : 42)

- "Est-ce que je pourrais vous demander comment vous avez voté ?"
(IFOP 58-11 : 7)

Nous avons vu, à propos des interrogatives directes, qu'elles avaient en commun une caractéristique, l'intonation, marquée par le point d'interrogation. Peut-on caractériser de même les interrogatives indirectes ? Il ne semble pas que les grammairiens soient parvenus à identifier un trait morphologique qui soit présent dans toutes les interrogatives indirectes, et dans elles seules. Ainsi, DAMOURETTE et PICHON font remarquer que les interrogatives indirectes ont la plupart du temps une construction formellement identique à celle des autres subordonnées, et que seul le sens permet de les en distinguer [Damourette 1931 : IV, 116-117]. Par exemple, pour interpréter correctement les deux phrases suivantes : "Je demandais quand elle est partie" et : "Je dormais quand elle est partie", il est indispensable de se référer à la signification du verbe principal. Cela explique la manière dont les grammaires usuelles présentent l'interrogation indirecte ; on ne peut pour autant en inférer que celle-ci n'est qu'une forme dérivée de l'interrogation directe.

b) Modalité de la phrase exprimant l'interrogation

La phrase dont dépend l'interrogative subordonnée peut prendre l'une quelconque des trois modalités que nous avons distinguées. Ce peut être une *énonciative* (à valeur optative ou injonctive, en position de proposition principale ou de subordonnée) :

- "Je vais vous nommer deux choses, et vous me direz en quoi elles se ressemblent et en quoi elles sont différentes"
(TMI : 1)
- "Pour chacune d'elles je voudrais que vous me disiez si vous la trouvez très importante, assez importante, ou pas importante du tout ?"
(SOFRES 340 : 23)

Ce peut être une *injonctive* (toujours proposition principale) :

- "Dis-moi ce que tu vois sur cette image ?"
(TERMAN : images, 3 ans)
- "Je vais te lire des phrases dans lesquelles il y a des bêtises ; écoute bien attentivement, et dis-moi chaque fois ce qu'il y a de bête"
(BS 2 : 35)
- "Maintenant, comptez-les et dites-moi combien il y en a dans chaque tas en commençant par ici"
(TMI : 3)
- "Indiquez quels ont été à ce sujet vos principaux informateurs ou moyens d'informations ?"
(OCDE-IUT : 40)
- "Regardez bien ces deux voitures R et S et dites-moi, à votre avis dans laquelle des deux il y a plus de place pour les passagers ?"
(SOFRES 769 : 77)

Ce peut être une *interrogative* (toujours une proposition principale à valeur injonctive) :

- "Pouvez-vous me dire si ce projet vous plaît plutôt ou vous déplaît plutôt ?"
(IFOP 58-9 : 6)
- "Pouvez-vous me dire, pour chacune d'elles, si vous êtes d'accord ou pas d'accord ?"
(SOFRES 340 : 21)
- "Voulez-vous me dire pour chacun d'eux si cela vous fait plaisir ou non ?"
(IFOP 58-12 : 3)
- "Voulez-vous me dire laquelle de ces formes S, H ou F vous conviendrait le mieux pour le genre de voiture que vous comptez acheter ?"
(SOFRES 769 : 62)
- "Est-ce que je pourrais savoir ce que c'est ?"
(IFOP 59-9 : 10)
- "Est-ce que je pourrais vous demander s'il y a des chances pour que vous votiez plutôt oui ou plutôt non au référendum ?"
(IFOP 58-9 : 13a)

Notons que, dans l'ensemble de notre corpus, le verbe qui introduit la subordonnée interrogative indirecte est le plus souvent le verbe *dire*. Nous avons relevé également les verbes *demander*, *indiquer*, *montrer*, *préciser*, *savoir*.

1.1.4. Autres caractères morpho-syntaxiques

Il nous reste à décrire rapidement certaines caractéristiques morpho-syntaxiques des énoncés exprimant une question. Ces caractéristiques, communes à toutes les formes d'énoncés que recèle notre corpus, présentent un intérêt pour l'analyse sémantique ultérieure du contenu de la question. Toutefois, ces notions ne sont pas propres aux énoncés exprimant une question, mais s'appliquent au contraire à l'ensemble des énoncés. C'est pourquoi nous n'avons pas jugé utile de leur donner un grand développement. Nous n'avons retenu que trois caractéristiques : la présence d'une proposition circonstancielle, le mode et le temps du verbe exprimant le contenu de l'interrogation, et les marques morphologiques désignant le locuteur ou la personne interrogée.

a) Propositions subordonnées circonstancielles

Parmi l'ensemble des formes de propositions dépendantes circonstancielles que distinguent les grammaires usuelles, deux seulement se rencontrent fréquemment dans notre corpus : les temporelles, et les hypothétiques.

Les *propositions dépendantes temporelles* sont des propositions subordonnées exprimant des relations d'antériorité, de simultanéité, ou de postériorité entre l'action principale et l'action subordonnée. Ces propositions sont introduites par des conjonctions (*quand, lorsque*), des locutions conjonctives (*avant que, depuis que, pendant que*), ou des locutions formées d'un substantif et du pronom relatif *où* (*au moment où*) [Wagner 1962 : 574-575]. Par exemple :

- "Lorsque l'on vous demande de faire du travail supplémentaire après que vous avez accompli votre tâche et une pleine journée de travail, est-ce que :
 - a) vous ne montrez jamais aucun ressentiment ?
 - b) vous montrez clairement que vous n'êtes pas content ?"

(CAT 16PF : 119)
- "Quand vous êtes déprimé, essayez-vous de trouver quelqu'un pour vous remonter le moral ?"

(BERN 35 : 78)
- "A quelle distance étais-je de mon point de départ quand je me suis arrêté ?"

(TMI : 30b)
- "Lisez-vous (X) chaque fois que vous lisez (Y) ?" (SOFRES 666 : 36)

- "Depuis que vous êtes électeur, avez-vous toujours voté pour le même parti ?" (IFOP 58-11 : 50)
- "Lorsque vous écoutez la radio chez vous, est-ce :
 - uniquement sur un poste branché sur le secteur ?
 - uniquement sur un poste à transistors ?
 - sur l'un et l'autre ?"
 (ORTF 65 : 14)

Les propositions dépendantes hypothétiques expriment d'idée que "la réalisation du fait exprimé dans la proposition principale y dépend d'un autre fait conçu comme éventuel et pose soit comme tel, soit comme objet d'une hypothèse. Dans l'un et l'autre cas, la proposition dépendante traduit un acte de l'esprit par lequel ou bien on recrée le passé autrement qu'il n'a eu lieu, ou bien on révoque une actualité présente, ou bien on construit l'avenir en imagination" [Wagner 1962 : 588]. Le plus fréquemment, l'hypothèse est introduite par la conjonction *si*. Par exemple :

- "Si tu vois ta maison brûler, que faut-il faire ?" (TERMAN : 6 ans, 4°)
- "Si vous êtes témoin d'un accident, vous hâtez-vous de porter secours ?" (BERN 35 : 41)
- "Si les chances semblent être contre votre réussite dans une entreprise, croyez-vous que cela vaille la peine de prendre un risque ?" (CAT 16PF : 29)
- "Si vous avez besoin d'essence et que vous ne trouvez pas tout de suite de station de votre marque, est-ce que :
 - . vous prenez ce qui se présente ?
 - . ou vous attendez jusqu'au dernier moment pour essayer de trouver votre marque ?"
 (SOFRES 340 : 14)
- "Si l'on vous dit que, pour améliorer la situation du pays en vue de l'avenir, il est nécessaire qu'aujourd'hui vous acceptiez, vous personnellement, de réduire votre niveau de vie, seriez-vous prêt à vous restreindre ?" (IFOP 58-9 : 33a)
- "Si vous trouviez cette possibilité en France, est-ce que cela vous intéresserait ?" (SOFRES 340 : 22)

b) Mode et temps du verbe de l'interrogation.

Les modes utilisés dans la partie de nos énoncés correspondant au contenu de l'interrogation sont l'indicatif, et (dans la mesure où l'on peut l'en distinguer, cf. [Wagner 1962 : 295]) le mode conditionnel. Les temps utilisés le plus souvent sont l'indicatif présent, le passé indéfini, et le conditionnel présent.

Le *présent de l'indicatif* sert en principe à évoquer un fait qui constitue l'actualité du locuteur (*présent ponctuel*) ; en pratique, on l'emploie également pour exprimer (entre autres choses) des caractères permanents ou durables ou des faits répétés [Wagner 1962 : 336-337]. D'où en particulier son emploi dans les questionnaires de personnalité. Exemples d'utilisation de l'indicatif présent :

- "Combien y a-t-il d'objets ici ?" (BS Z : 16)
- "Etes-vous très confus s'il vous arrive de saluer quelqu'un par erreur ?" (BERN 35 : 65)
- "Lorsque vous êtes malade, aimez-vous beaucoup être l'objet d'attentions de la part de vos amis ?" (BERN 35 : 70)
- "Lorsque vous avez un rival ou un concurrent, pouvez-vous soutenir énergiquement la compétition sans ressentir des sentiments d'hostilité ou de jalousie ?" (CAT 16PF : 172)
- "Depuis combien de temps travaillez-vous dans l'entreprise qui vous emploie actuellement ?" (F.ING 69 : VII 10)
- "Dans quel but poursuivez-vous ces études ?" (CEREQ-SEU : 19a)
- "Est-ce que vous vous intéressez beaucoup, moyennement, peu ou pas du tout aux prochaines élections ?" (IFOP 58-11 : 18c)
- "Lorsqu'un programme de TV ne vous plaît pas, que faites-vous le plus souvent ?" (ORTF 65 : 37)
- "Normalement, vous remplacez votre voiture tous les combien ?" (SOFRES 967 : 26)

Le *passé indéfini* (ou *passé composé*) est la forme utilisée le plus généralement pour exprimer un fait passé, qu'il soit ponctuel ou étendu. Par exemple :

- "Avez-vous déjà organisé spontanément des groupements, des équipes, ou d'autres associations ?" (BERN 35 : 40)
- "Avez-vous jamais essayé de bluffer vis-à-vis d'un gardien ou d'un portier pour entrer dans une enceinte tout en sachant que vous aviez le droit de vous y trouver ?" (CAT 16PF : 57)
- "A quel âge avez-vous terminé vos études ?" (GEMAS 76 : 102)
- "A quel moment avez-vous pris la décision de faire des études supérieures ?" (OCDE-IUT : 26)
- "Depuis que vous possédez cette voiture, avez-vous changé de numéro minéralogique, par exemple parce que vous avez déménagé ?" (SOFRES 769 : 22)
- "Combien de mois avez-vous cherché cet emploi ?" (CEREQ-SEU : 23a)
- "Combien de temps vous êtes-vous connus et fréquentés avant le mariage ?" (INED 081 : fiançailles 2)

Mais on rencontre également parfois l'imparfait de l'indicatif, "en fonction de présent dans le passé" [Wagner 1962 : 353], par exemple pour décrire un état durable repéré par référence à un événement ponctuel passé :

- "Que faisiez-vous quand vous avez commencé votre vie professionnelle ?" (IFOP 58-9 : 31)

Le futur de l'indicatif sert à situer un événement dans l'avenir. Comme notre corpus ne contient pas de questionnaire à visée proprement prospective, cette forme y est relativement rare. Exemples :

- "Quand prendrez-vous votre décision ?" (IFOP 58-9 : 14a)
- "Lorsque vous voterez pour un candidat aux élections des 23 et 30 novembre, le ferez-vous pour exprimer votre confiance en un homme dont les qualités vous plaisent, ou pour exprimer votre confiance dans le parti dont il sera le représentant ?" (IFOP 58-11 : 24)
- "Pour chacune des dépenses suivantes, dites-moi si vous y ferez très attention, ou, au contraire, peu attention, ou si c'est entre les deux" (SOFRES 769 : 50)

Le *conditionnel présent* s'emploie souvent avec une proposition hypothétique. Par exemple :

- "Si vous voyiez deux enfants que vous connaissez se battre pour un jouet, que feriez-vous ?" (CAT 16PF : 63)
- "Vous sentiriez-vous sûr de vous, si vous deviez lancer une idée pour amorcer un débat dans un groupe ?" (BERN 35 : 125)
- "Si ce magazine venait à disparaître, en seriez-vous affectée ?" (SOFRES 686 : 29)
- "S'il y avait un tel endroit près de chez vous, est-ce que vous y prendriez systématiquement votre essence ?" (SOFRES 340 : 7)

Plus généralement, le conditionnel est le mode de l'éventuel, et peut être employé sans proposition conditionnelle :

- "Aimeriez-vous vous livrer aux occupations suivantes ?" (QI Filles)
- "Préfèreriez-vous travailler en tant que :
 - a) orienteur professionnel pour des jeunes qui cherchent une carrière ?
 - b) directeur d'une entreprise ?" (CAT 16PF : 107)
- "Dans quel ordre classeriez-vous ces 3 raisons en commençant par la plus importante pour vous ?" (SOFRES 686 : 21)
- "Conseilleriez-vous actuellement à une jeune fille de s'engager dans la même voie professionnelle, de s'orienter dans la même direction que vous ?" (F.ING 69 : VII 23)

C'est également le cas pour le *conditionnel passé* :

- "Auriez-vous souhaité ne pas interrompre votre activité professionnelle ?" (F.ING 69 : III 16)
- "Tout bien considéré, auriez-vous aimé avoir plus d'enfants, ou moins d'enfants que vous n'en avez, ou bien le même nombre ?" (INED 081 : opinions 13b)

c) Désignation du locuteur et de l'interlocuteur.

Dans l'interrogation indirecte, la proposition qui contient l'interrogation (c'est-à-dire celle à laquelle la proposition interrogative est subordonnée) peut faire référence à la fois au locuteur et à l'interlocuteur ("je vous demande", "vous me direz") ; au locuteur seul ("je voudrais savoir") ; à l'interlocuteur seul

("décrivez", "indiquez") ; enfin, à aucun des deux ("il faut montrer", "préciser").

Si l'on s'intéresse aux seules propositions interrogatives (ou aux autres types de propositions ayant valeur de question), on ne relève dans notre corpus que deux cas de figure : ou bien l'interlocuteur est représenté, ou bien il est absent de la proposition ; mais le locuteur, lui, n'est jamais présent. Remarquons d'ailleurs qu'il y a là un problème d'usage (lié à la fonction remplie par les questionnaires et les tests les plus usuels), et non un problème de langue. En effet, rien n'interdit en principe d'imaginer, dans un test d'aptitude ou de niveau, un item du genre : "Es-tu plus grand ou plus petit que *moi* ? " ; ou bien, dans un questionnaire de personnalité : "Accepteriez-vous de *me* le prêter ?".

Lorsque l'interlocuteur est présent dans l'interrogation, il est en général représenté par un pronom personnel ou un adjectif possessif, en position de sujet ou de complément. Par exemple :

Pronom sujet :

- "Dans quelle direction allez-vous maintenant ?" (TMI : 17b)
- "A quelle date avez-vous commencé à travailler ?" (CEREQ-SEU : 23b)
- "Voterez-vous de préférence pour un député sortant ou pour un candidat nouveau ?" (IFOP 58-11 : 25)
- "Quand êtes-vous allé au cinéma pour la dernière fois ?" (SOFRES 571 : 22a)

Pronom complément (déterminatif ou d'objet) :

- "Est-ce facile pour vous de vous reposer et de "ne pas vous en faire", quand cela vous est possible ?" (CAT 16PF : 140)
- "Vous est-il arrivé d'être ennuyé par des propos malveillants tenus à votre sujet derrière votre dos, lorsque ces propos ne reposaient absolument sur rien ?" (CAT 16PF : 79)
- "Les railleries vous humilient-elles, même quand vous avez conscience d'être dans le vrai ?" (BERN 35 : 46)
- "Vous personnellement, vous intéresserait-il que votre prochaine voiture soit équipée d'une boîte de vitesses automatique ?" (SOFRES 967 : 34)

Adjectif possessif (se rapportant au sujet ou à un complément) :

- "Lorsque vous avez terminé vos études secondaires, votre mère exerçait-elle une activité professionnelle ?" (F.ING 69 : VI 1)
- "Sur quels points l'entretien courant de votre ... (modèle cité à Q.1) diffère-t-il des recommandations du constructeur ?" (SOFRES 967 : 29)
- "Dans quelle mesure les raisons ci-dessous ont-elles joué dans votre choix de faire des études à l'I.U.T. ?" (OCDE-IUT : 34)

Notons que, dans un énoncé, il peut y avoir plus d'un mot (pronom ou adjectif) se référant à l'interlocuteur. Par exemple :

- "Est-ce que votre conscience vous pousse en fait à être trop consciencieux ?" (CAT 16PF : 37)
- "Est-ce que votre réserve naturelle est en général pour vous un obstacle quand vous désirez entamer une conversation avec une personne du sexe opposé, que vous ne connaissez pas, mais qui est très séduisante ?" (CAT 16PF : 103)
- "Dans votre famille, on dit que vous êtes :
 - . gai
 - . entreprenant
 - . méditatif".
 (BNIST : 53)

Les énoncés dont l'interlocuteur est absent ne comportent évidemment aucun pronom ou adjectif possessif à la seconde personne :

- "Pourquoi est-ce bête ?" (TMI : 4)
- "A quelle heure le cambriolage a-t-il eu lieu ?" (TMI : 8)
- "Combien y-a-t-il de boîtes en tout en comptant la plus grande aussi ?" (TERMAN : adulte moyen, 4°)
- "Jean a deux frères. Combien y a-t-il de garçons dans sa famille ?" (ECO 5 : 6, ex.)
- "A quelles marques ces automobilistes pensaient-ils ?" (SOFRES 340 : 17)
- "A combien devrait-on fixer le salaire mensuel maximum ?" (GEMAS 76 : 40)

Ou bien, si l'énoncé complet compte un pronom ou un adjectif possessif à la seconde personne, celui-ci est extérieur à l'interrogation proprement dite :

- "On vous donne une série de nombres qui se suivent dans un certain ordre et on vous demande de trouver deux nombres qui continuent la série" (EAI 5/1 : S.N.)
- "A votre avis, qu'y aurait-il encore à changer ou à améliorer ...?" (SOFRES 571 : 32)
- "A votre avis, combien gagnent, en moyenne par mois, les personnes exerçant les professions indiquées ci-dessous ?" (GEMAS 76 : 36)
- "A votre avis, y aura-t-il dans ce sens :
 - une majorité écrasante (plus de 75 %)
 - une forte majorité (de 60 à 74 %)
 - une faible majorité (de 50 à 59 %) ?" (IFOP 58-9 : 18)

1.1.5. Bilan de la description du corpus.

Ce survol descriptif des énoncés utilisés, dans les épreuves psychologiques et les questionnaires d'enquête, pour poser des questions, n'est évidemment qu'un premier pas dans l'analyse morpho-syntaxique de ces énoncés. Après en avoir rappelé les limites et les insuffisances, et avoir marqué par cela même les perspectives qui sont ouvertes aux recherches ultérieures, nous montrerons que la diversité de ce type d'énoncés est vraisemblablement plus grande encore qu'il ne nous paraît. Nous concluons ainsi cet aperçu sur la polymorphie de l'interrogation.

a) Limites de ce survol descriptif.

Cette première étape avait pour but de recenser, classer, et étiqueter les différentes formes d'énoncés de notre corpus. Dans la mesure du possible, nous nous en sommes tenu à l'aspect purement grammatical, en laissant de côté l'aspect sémantique de l'énoncé. Nous avons ainsi renoncé délibérément à toute une série de recherches possibles sur le vocabulaire ; par exemple, sur le langage des épreuves psychologiques et le langage des enquêtes, comparé au langage de la conversation courante ; sur les formes pouvant se substituer l'une à l'autre sans altérer

le sens de la phrase (nous en verrons une illustration au paragraphe suivant) ; ou sur le style de l'énoncé, selon qu'il s'agit d'un questionnaire administré individuellement ou collectivement, s'adressant à des enfants ou à des adultes, visant à recueillir des opinions ou à évaluer une aptitude, etc. Il s'agit là de phénomènes immédiatement perceptibles (on peut même déceler un style propre à chaque institut de sondage), qui mériteraient à eux seuls plusieurs recherches importantes.

En restant dans le cadre que nous nous sommes fixé au départ, il nous faut souligner certaines insuffisances de cette description. Tout d'abord, nous sommes implicitement resté au niveau de phrases relativement simples : le plus haut niveau de complexité auquel nos catégories s'appliquent est celui qui correspond à l'interrogative indirecte avec subordonnée circonstancielle. Nous avons rencontré, dans notre corpus, des énoncés dont la structure grammaticale était plus complexe ; comme par exemple :

- *"Si au second tour, vous avez le choix entre M. Poher et M. Pompidou, voulez-vous me dire quel est celui pour lequel il y a les plus grandes chances que vous votiez ?"* (IFOP 69-5 : 20)

Mais nous avons estimé hors de notre propos de pousser plus avant l'analyse de phrases de ce type, nous intéressant principalement au contenu de la question et aux circonstances pouvant le modifier ; par conséquent, une telle analyse reste à faire.

D'autre part, nous avons renoncé à développer (voire à aborder) certains problèmes grammaticaux paraissant propres à la langue française, et ne nous semblant pas d'une importance fondamentale pour notre classification. En voici deux exemples :

- Existe-t-il des interrogatives indirectes en *ce qui* ou *ce que*, ou bien s'agit-il de subordonnées relatives [Bloch 1945 : 226, §402, Remarque], ou de propositions complément d'objet du verbe principal [Wagner 1962 : 545, § 648, 1 c]? Nous avons tranché en faveur de leur existence chaque fois que le passage au style direct engendrait une interrogative.

- Dans une question de la forme : "Qu'y a-t-il sur cette image ?", le pronom *que* est-il sujet ou complément d'objet direct ? Une solution consiste à

distinguer entre sujet apparent (*il*) et sujet réel (*que*) (cf. : [Académie 1932 : 28] ; [Georgin 1952 : 351]) ; encore qu'il ne soit pas aisé de trancher dans certains cas (comparer l'analyse de : "il pleut" et celle de "il pleut des hallebardes" [Tesnière 1959 : 106, § 7 ; 241, § 6]). Mais cette solution n'est pas acceptée par tous les grammairiens (cf.p.ex. [Le Bidois 1968 : I, 175-178]). Sans prendre position sur le fond, nous avons jugé commode de traiter de la même manière toutes les interrogations utilisant un tour impersonnel ("Y-a-t-il", "Arrive-t-il"), et de considérer le sujet grammatical comme étant le sujet réel. Par conséquent, dans cette perspective, les interrogatives avec "y-a-t-il" portent sur le complément direct d'objet du verbe *avoir*.

En outre, nous avons rencontré quelques cas d'énoncés difficiles à classer. Les uns parce qu'ils présentaient simultanément deux formes différentes (énonciative ou interrogative, et injonctive) :

- "*Vous occupez vos temps de loisirs à (classer de 1 à 4) : ...*"
(BNIST : 35)
- "*A votre avis, qu'est-ce qui compte le plus ? Et ensuite ? Classez les autres qualités par ordre d'importance décroissante*"
(INED 081 : opinions 3)

D'autres parce qu'ils posaient plusieurs questions en une seule phrase :

- "*Quand les deux familles se sont-elles rencontrées et où ?*"
(INED 081 : fiançailles 7)
- "*Quand et sous quelle forme pourrait se manifester ce danger ?*"
(IFOP 58-9 : 22b)
- "*Est-ce que je pourrais vous demander pour quel candidat et pour quel parti vous avez l'intention de voter ?*"
(IFOP 58-11 : 33)
- "*Vous arrive-t-il d'écouter ces feuilletons et les aimez-vous ?*"
(ORTF 65 : 11)
- "*Je vous demande maintenant de me dire à quel moment et pendant combien de temps vous avez écouté la radio, ou regardé la TV, pendant la journée d'hier ...*"
(ORTF 65 : 28)

Ces deux séries d'exemples illustrent la difficulté qu'il y a à s'en tenir aux limites de phrase si l'on désire analyser les questions. De même que la phrase peut contenir plusieurs questions, de même il arrive qu'une question unique soit posée en plusieurs phrases

- *"Parmi les 4 dessins indiqués par des lettres, 3 se ressemblent par quelque chose. Mais un seul dessin est différent des 3 autres. Lequel ?"* (ECO 5 : 7)

Ou, ce qui revient au même, que la question n'ait de sens qu'en fonction des questions qui la précèdent :

- *"Et maintenant, combien y en a-t-il ?"* (BS Z : 20)
- *"Dans quel type d'établissement et à quel niveau ?"* (CEREQ-SEU : 17)
- *"Combien a-t-elle été vendue ?"* (SOFRES 769 : 32)

Ces considérations conduisent tout naturellement à préconiser d'une part la subdivision de la phrase en autant d'unités qu'il y a d'interrogations, d'autre part la prise en compte, pour l'analyse du sens des questions, du contexte amont de celles-ci.

b) Variété observée et variété potentielle.

Pour classer les énoncés de notre corpus, nous avons considéré les grandes catégorisations acceptées par la majorité des grammairiens, ou celles qui nous paraissaient les plus importantes en vue d'une formalisation ultérieure. Cela nous a conduit par conséquent à négliger certaines catégories (par exemple, les propositions dépendantes circonstancielles autres que les temporelles et les hypothétiques), ou même ne pas prendre en compte certaines dimensions pourtant fondamentales. Par exemple, nous n'avons pas classé à part les énoncés à la voix passive, de la forme :

- *"Au total, et en comptant cette nouvelle voiture, dans les 3 prochaines années, combien de voitures vont être achetées dans votre foyer ?"* (SOFRES 769 : 44)

ni les tournures négatives, comme :

- *"Pourquoi n'avez-vous pas la radio ?"* (ORTF 65 : 3ter)

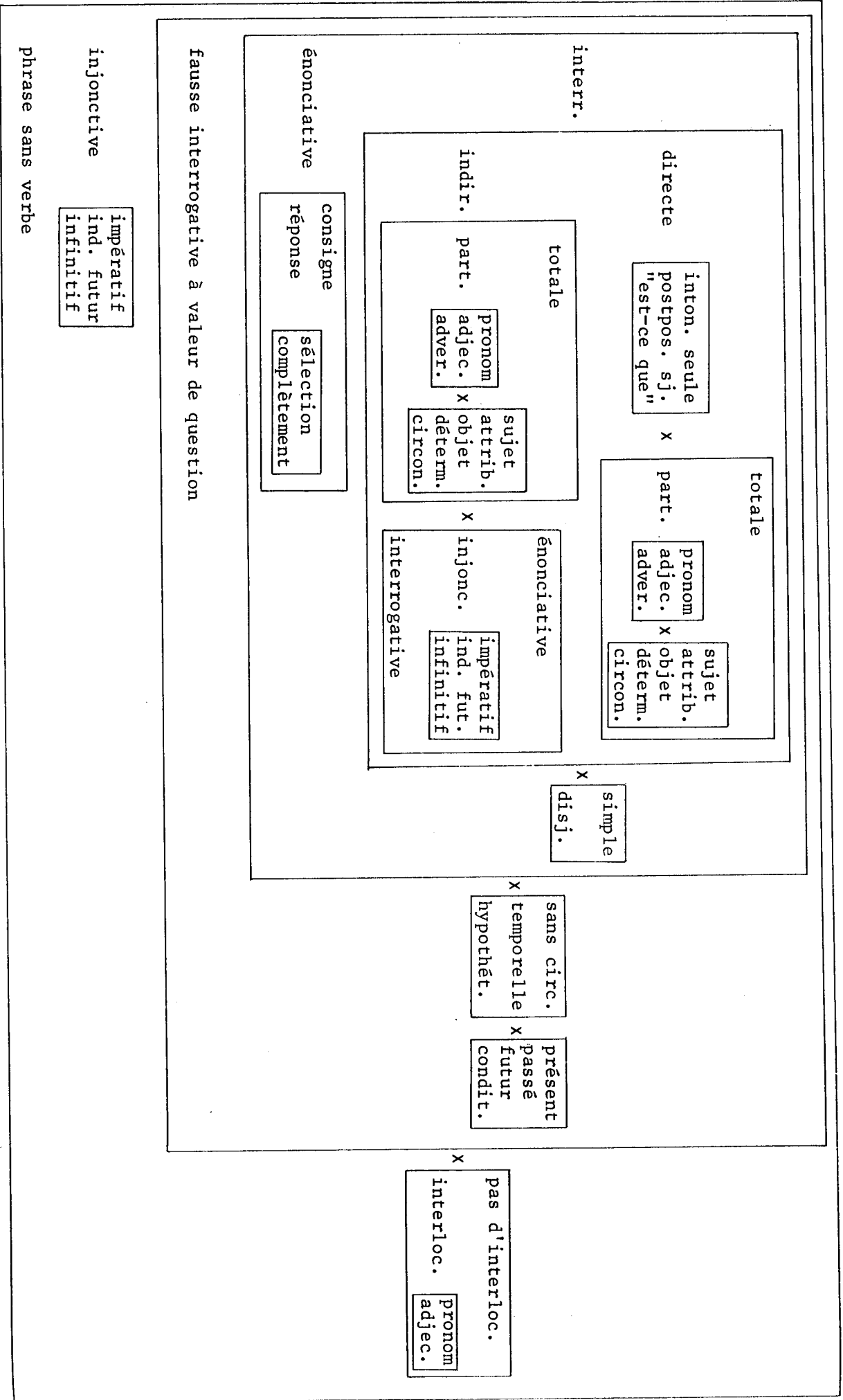


Schéma n° 1.1.1 : Classification des énoncés à valeur de question d'après leur caractéristiques morpho-syntaxiques.

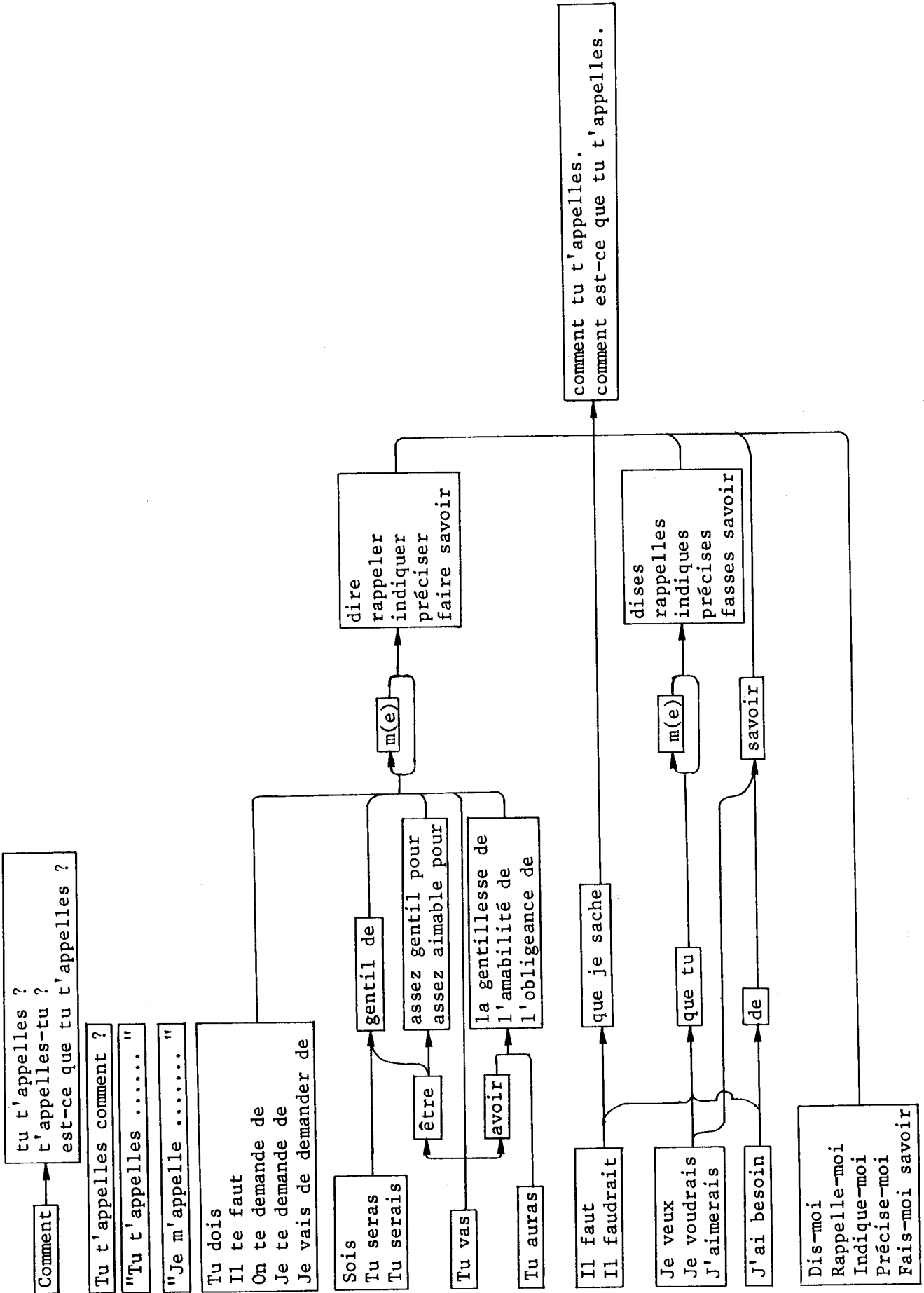
- "Pourquoi n'avez-vous pas voté pour lui ?" (IFOP 58-12 : 20a)

Il est évident que si nous avons pris ces caractéristiques en considération (bien que les exemples en soient assez rares dans notre corpus), cela aurait augmenté le nombre des formes d'interrogation recensées.

D'autre part, nous nous sommes borné le plus souvent à présenter chaque dimension indépendamment des autres, au lieu de tenir compte de l'ensemble des combinaisons possibles des caractéristiques servant à décrire les énoncés. Cela n'est pas dû à la seule facilité d'exposition. En effet, nous présentons, sur le schéma n° 1.1, un résumé simplifié du système de classification que nous avons adopté. Ce schéma tient compte de l'ensemble des incompatibilités : par exemple, dans une injonctive, l'interlocuteur est toujours sujet (éventuellement marqué par la seule personne du verbe) ; une fausse interrogative à valeur de question ne peut pas être disjonctive ; etc. En nous limitant à ces quelques dimensions, nous arrivons à un total de plus de dix mille formes possibles (11.662). Il est clair que, pour parvenir à illustrer chacune de ces combinaisons, ou pour être à même d'affirmer que l'absence d'une forme est due à sa rareté dans les tests ou les questionnaires en général, un corpus d'environ cent mille énoncés ne serait pas suffisant. Avec une classification un peu plus détaillée, qui prendrait en compte des phrases plus complexes, et considèrerait quelques critères morpho-syntaxiques supplémentaires (complément déterminatif d'un nom/d'un adjectif/d'un adverbe ; adverbe interrogatif circonstanciel de lieu/de temps/de cause/de manière/de quantité ; etc.), nous dépasserions largement les cent mille formes possibles. Quelle devrait être, dans ces conditions, la dimension de notre corpus ?

Nous atteignons là les limites des recherches sur corpus : toute collection d'énoncés, quelle que soit sa taille, ne saurait présenter qu'un tableau partiel des possibilités du langage [Crystal 1971 : 104]. L'intérêt d'un tel corpus nous paraît résider surtout dans la possibilité qu'il offre au chercheur de recenser les dimensions qui lui permettront ensuite de décrire la variété potentielle du type d'énoncés qu'il étudie.

Pour illustrer sur un exemple concret la polymorphie potentielle d'une interrogation, nous avons choisi une interrogation partielle de quatre mots, le



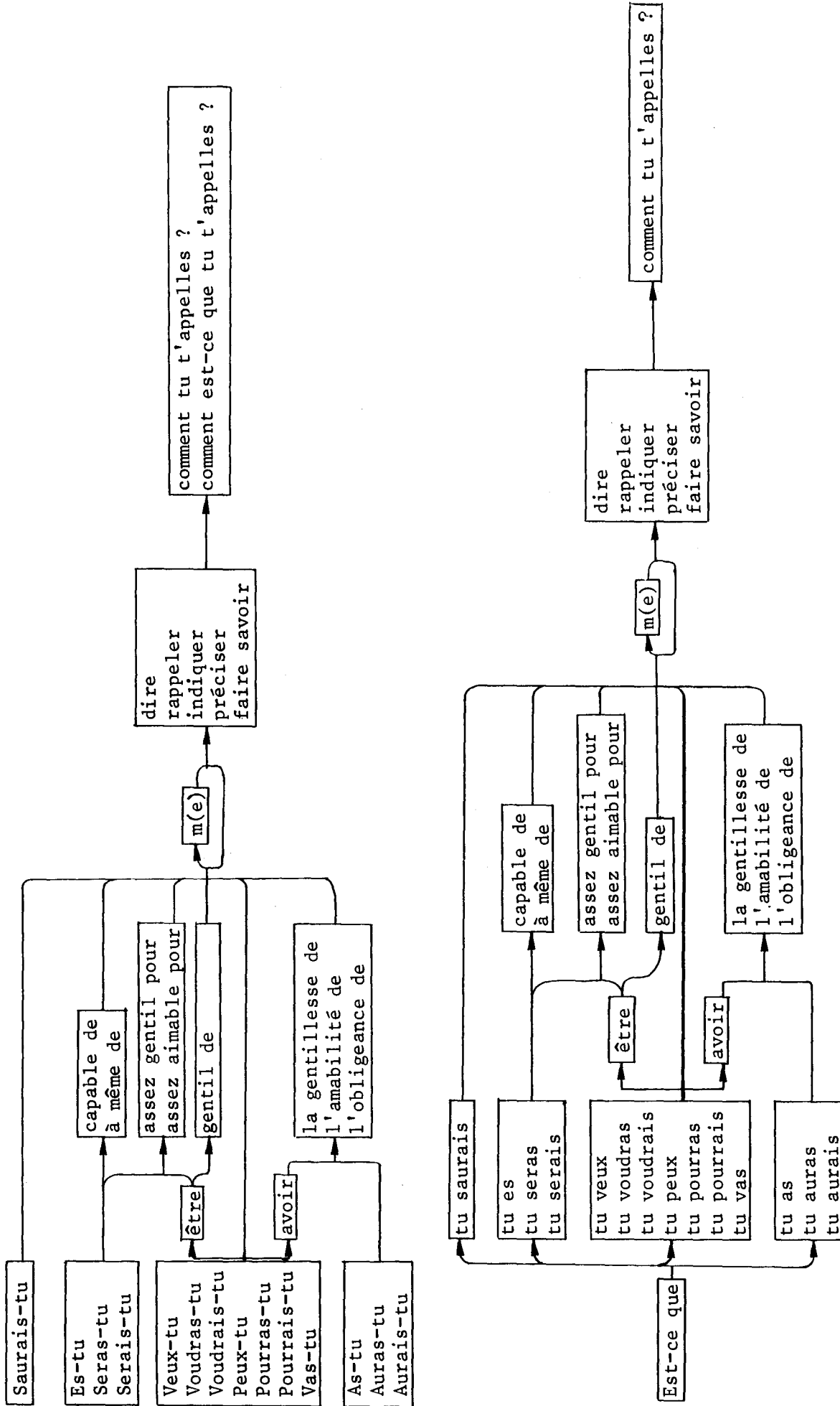


Schéma n° 1.2 : Deux mille six cent trente-huit manières de demander : "comment t'appelles-tu ?".

premier item du *Binet-Simon* : "Comment t'appelles-tu ?". Nous avons cherché de combien de manières cette question pouvait être posée, en conservant toutefois le vocabulaire de l'énoncé original. Pour cela, nous avons considéré, parmi l'ensemble des formes d'interrogation que nous avons décrites, celles que l'on peut appliquer à ce type d'énoncé sans en altérer le sens. En outre, pour construire les phrases introduisant la forme interrogative indirecte, nous avons également tenu compte de l'existence de tournures sémantiquement équivalentes, mais utilisant des mots différents ; nous ne nous sommes donc pas limité aux seuls aspects morpho-syntaxiques. Toutefois, nous n'avons pas pris en considération l'ensemble des relations de synonymie possibles ; au contraire, nous nous sommes borné à un petit nombre de verbes ou de locutions verbales équivalentes dans la langue parlée, sans tenir compte de la présence ou de l'absence d'adverbes ("je vais te demander *maintenant*..." ; "je voudrais *bien*..." ; "j'aimerais *beaucoup*..." ; "tu serais *très* gentil..." ; etc.). En outre, nous avons écarté des formulations plus proches de la langue écrite que de la langue parlée ("Indique (ici) comment tu t'appelles"), puisque le *Binet-Simon* est administré oralement. Enfin, nous avons délibérément omis certains énoncés, soit que leur usage nous ait paru trop invraisemblable dans une situation de test ("Serais-tu d'accord pour me dire comment tu t'appelles ?" ; "Et si tu me disais (maintenant) comment tu t'appelles ?"), soit que l'on puisse les considérer comme des phrases indépendantes (Parataxe : "Dis-moi : comment t'appelles-tu ?").

Malgré ces contraintes, nous avons dénombré, dans le schéma 1.2, plus de *deux mille cinq cents* formulations possibles. Encore serait-il facile de construire des énoncés ne s'y trouvant pas, comme par exemple : "Vas-tu pouvoir me dire comment tu t'appelles ?" ; ou bien : "Sois assez gentil pour dire comment tu t'appelles" ; ou encore : "Voudrais-tu me déclarer comment tu t'appelles ?" ; etc. Si l'on accepte de prendre en compte les identités sémantiques à l'intérieur de la question elle-même ("comment te nommes-tu ?" ; "comment t'appelle-t-on ?" ; "comment te nomme-t-on ?"), on dépasse sans peine les *dix mille* formulations. En acceptant des variantes à peine moins proches ("Quel est ton nom ?"), et en introduisant quelques adverbes, nous parviendrions à engendrer les (plus de) "cent mille manières de demander son nom à quelqu'un en français ...".

La variété potentielle de l'interrogation apparaît donc comme pratiquement illimitée. Dans ces conditions, comment réduire l'ensemble quasi-infini des formes observables à une forme unique, qui soit toutefois susceptible de rendre compte de l'ensemble des propriétés caractéristiques des questions ?

1.2. RÉDUCTION DE L'INTERROGATION À UNE FORME CANONIQUE

D'après les grammaires usuelles, les énoncés à valeur de question utilisés dans les tests ou les questionnaires se rangent dans des catégories morpho-syntaxiques assez disparates (même si les formes proprement interrogatives y sont largement majoritaires). A cette polymorphie incontestable s'oppose une non moins incontestable identité de sens : tous ces énoncés sont interprétés comme des interrogatives par des interlocuteurs connaissant bien la langue et jouissant de toutes leurs facultés. Cette identité d'interprétation tient-elle essentiellement au contexte (par exemple, à la situation d'examen psychologique, ou d'entretien d'enquête), ou bien peut-on identifier les caractéristiques communes à l'ensemble de ces énoncés, et à eux seuls ? Dans cette dernière hypothèse, doit-on recourir à des traits sémantiques, ou existe-t-il un ensemble suffisant de caractéristiques morpho-syntaxiques pour rendre compte de toutes les variétés d'énoncés à valeur de question ?

La grammaire traditionnelle, nous l'avons vu, est mal armée pour résoudre ce type de problèmes. C'est pourquoi nous nous sommes tourné vers des aspects plus récents de la théorie linguistique, correspondant à des domaines de recherche encore incomplètement explorés, mais dont les premiers résultats paraissent prometteurs. Toutes ces théories ont en commun d'analyser les mécanismes de la langue en termes de *transformations* (ou de *translations*) qui *engendrent* les énoncés effectivement observés. Il paraît donc possible de découvrir, sous la variété impressionnante des énoncés à valeur de question, un petit nombre de formes originelles dont ces énoncés sont dérivés. En outre, la plupart de ces théories (et en particulier la totalité de celles auxquelles nous nous référons) a donné lieu à des travaux de recherche ou d'application débouchant sur une automatisation des processus de transformation des énoncés. Bien qu'il soit encore trop tôt pour établir un bilan de ces travaux, on peut espérer les voir aboutir à (au moins) une automatisation partielle du traitement des énoncés en langue naturelle, qui nous permette d'analyser automatiquement la formulation des énoncés à valeur de question utilisés dans les tests et les questionnaires.

Nous nous sommes volontairement limité à l'examen de trois courants de la linguistique contemporaine : la syntaxe structurale de TESNIÈRE, les grammaires génératives inspirées des travaux de CHOMSKY, et la grammaire transformationnelle de HARRIS (appliquée au français par GROSS en particulier). Dans la perspective d'une réduction à une forme canonique des énoncés exprimant une question, l'intérêt de ces travaux est d'une part, de mettre en évidence la parenté (voire l'identité) de la structure d'énoncés de même signification, mais de forme différente ; d'autre part, de décrire certaines procédures permettant soit de passer d'une forme d'énoncé à une autre, soit d'engendrer tous les énoncés possibles à partir d'une même structure sous-jacente. Bien que la place accordée, dans ces recherches, à l'analyse de l'interrogation, ne soit généralement pas la place centrale, les éléments qu'ils nous apportent nous conduisent à conclure qu'il est possible de ramener la partie essentielle (le *noyau*) de tout énoncé interrogatif à une forme unique (*structure profonde* ou *énoncé distingué*).

2.1.1. La syntaxe structurale de Lucien TESNIÈRE

Les conceptions de TESNIÈRE en matière de syntaxe sont exposées dans son ouvrage posthume *Éléments de syntaxe structurale* [Tesnière 1959], auquel nous nous référons pour décrire la manière dont il rend compte de l'interrogation. L'objectif de TESNIÈRE semble avoir été de fonder une syntaxe générale, valable pour toutes les langues naturelles, et directement applicable aux problèmes de traduction et à ceux d'enseignement des langues étrangères.

L'auteur présente l'interrogation indirecte comme une dérivation de l'interrogation directe, et celle-ci comme une dérivation de la phrase verbale énonciative. Nous commencerons donc par exposer les notions de base indispensables pour comprendre la structure d'une phrase énonciative simple ; nous appli-

querons ensuite l'analyse structurale aux principales formes de phrases servant à poser une question ; nous conclurons sur l'apport de Lucien TESNIÈRE à la réduction des phrases interrogatives à une forme commune.

a) La phrase verbale énonciative simple

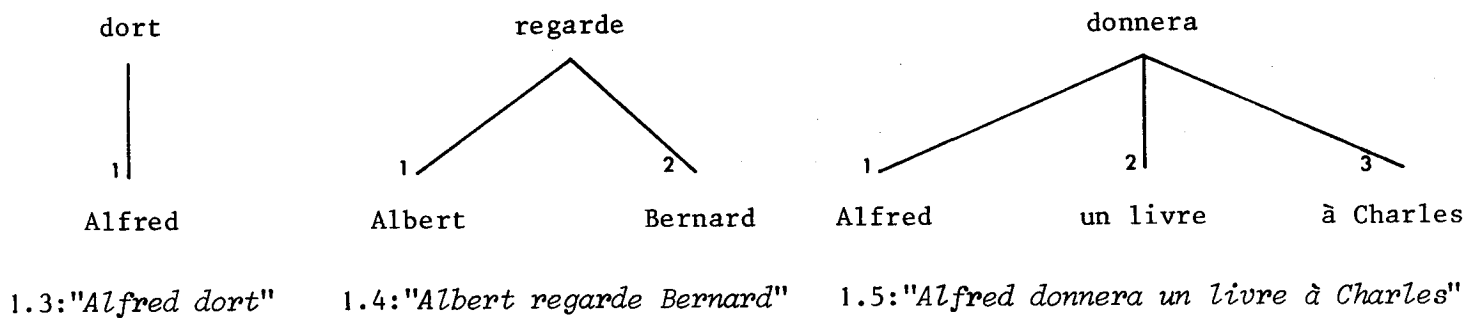
"La phrase est un *ensemble organisé* dont les éléments constitutifs sont les *mots*" [Tesnière 1959 : 11]. Une phrase comme "Alfred parle" est composée de *trois* éléments : "Alfred", "parle", et la connexion qui les unit ; cette connexion donne un sens à la phrase, et permet à l'interlocuteur de comprendre non seulement qu'il y a un homme qui s'appelle Alfred et que quelqu'un parle, mais aussi que ce quelqu'un qui parle est Alfred. Par conséquent, "comprendre une phrase, c'est *saisir l'ensemble des connexions* qui en unissent les différents mots" [op.cit. : 12].

Il est commode de représenter l'ensemble des connexions qui constitue la structure syntaxique de la phrase par un graphe orienté (*stemma*) dont les noeuds représentent les mots porteurs de sens, et les arcs, les connexions syntaxiques. Les mots porteurs de sens, ou *mots pleins*, se classent en quatre grandes catégories : d'une part, selon qu'ils expriment l'idée d'une *substance* (chose perçue par les sens et conçue par l'esprit comme douée d'une existence distincte) ou celle d'un *procès* (état ou action par lequel la substance manifeste son existence) ; d'autre part, selon qu'ils expriment des *notions concrètes* ou des *attributs abstraits*. Le nom (et le symbole général) donné par l'auteur à chaque espèce de mots pleins figure dans le tableau ci-dessous [op.cit. : 63] :

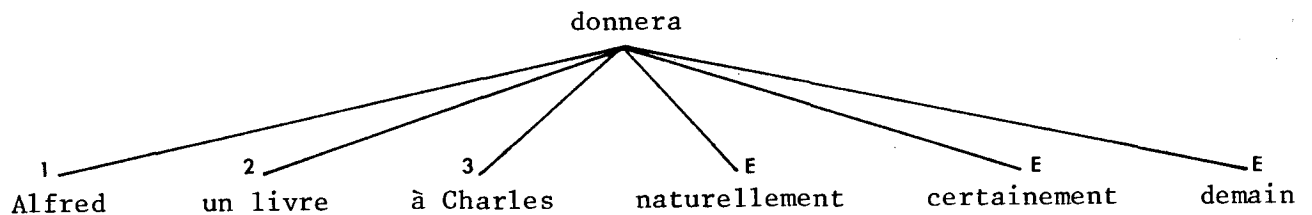
	Substance	Procès
Concret	<i>Substantif</i> (O)	<i>Verbe</i> (I)
Abstrait	<i>Adjectif</i> (A)	<i>Adverbe</i> (E)

A ces quatre espèces, définies par le contenu sémantique des mots, correspondent des structures syntaxiques bien spécifiques. Ainsi, le verbe peut *régir* (avoir pour subordonnés) jusqu'à trois substantifs (les *actants* du verbe), et un nombre variable d'adverbes (les *circonstants*). Le nombre maximum d'actants

est une propriété stable des verbes (*valence*) : on distingue les verbes *avalents*, à zéro actant (impersonnels ; ex. : "il pleut") ; *monovalents*, à un actant (intransitifs ; ex. : "l'homme tombe") ; *divalents*, à deux actants (transitifs ; ex. : "Albert lit un livre" ; "je me lave") ; et *trivalents*, à trois actants (ex. : "Armand demande un renseignement à Claude"). La fonction sémantique des actants permet de les caractériser avec précision : le *prime actant* est celui qui fait l'action (sujet) ; le *second actant* est celui qui supporte l'action (dans la voix active) ; le *tiers actant* est celui au bénéfice ou au détriment duquel se fait l'action. Exemples (schémas n° 1.3 à 1.5) :



Dans ses stemmas, TESNIÈRE n'utilise pas le symbole usuel de la flèche pour figurer un arc orienté : le sens de la relation syntaxique *régissant* → *subordonné* y est marqué par la position relative des deux mots, le régissant occupant un niveau supérieur à celui du subordonné. Lorsque le noeud verbal comporte des circonstants, ceux-ci sont placés au même niveau que les actants. De même que la fonction des actants est symbolisée par un nombre indiquant leur rang, la fonction des circonstants peut être marquée sur le stemma par un E (adverbe). Exemple (schéma n° 1.6) :



1.6 : "Naturellement, Alfred donnera certainement un livre à Charles demain".

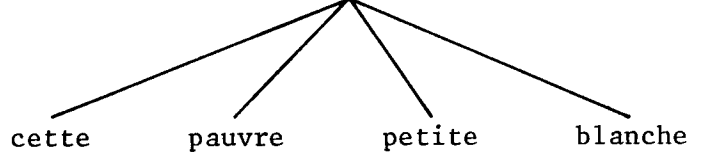
Les substantifs ne peuvent avoir comme subordonnés que des adjectifs (en nombre variable). Exemples d'adjectifs épithètes (schémas n° 1.7 et 1.8) :

un livre



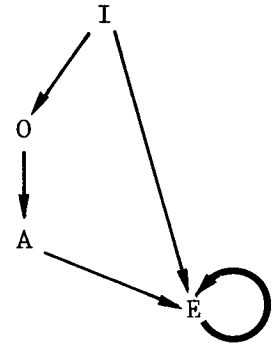
1.7 : "Un beau livre"

maison



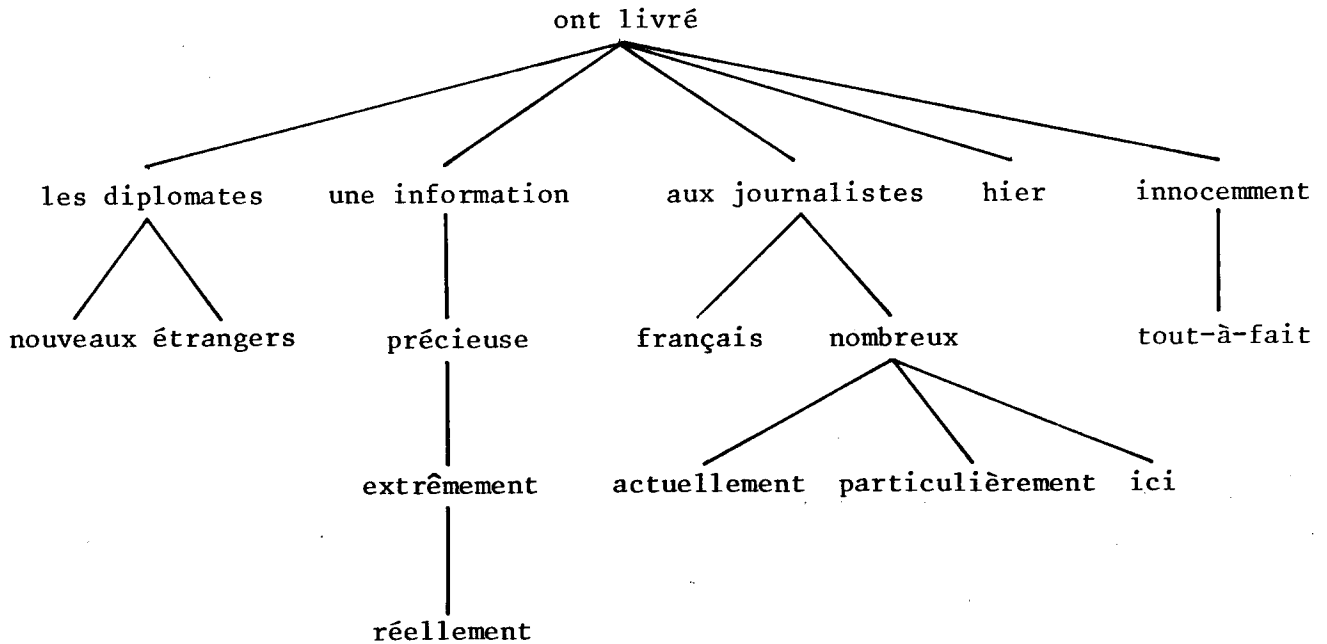
1.8 : "Cette pauvre petite maison blanche"

Un adjectif ne peut régir qu'un adverbe (ex. : "très beau" ; "plus grand"). Un adverbe, enfin, ne peut régir qu'un autre adverbe ("assez facilement" ; "relativement bien"). Le schéma n° 1.9 ci-contre résume l'ensemble des relations de subordination possibles entre les différentes espèces de mots-pleins. Comme on le constate, il y a une sorte de hiérarchie sur le plan syntaxique, hiérarchie dont le verbe occupe le sommet et l'adverbe la base.



1.9. Relations de subordination entre les mots-pleins

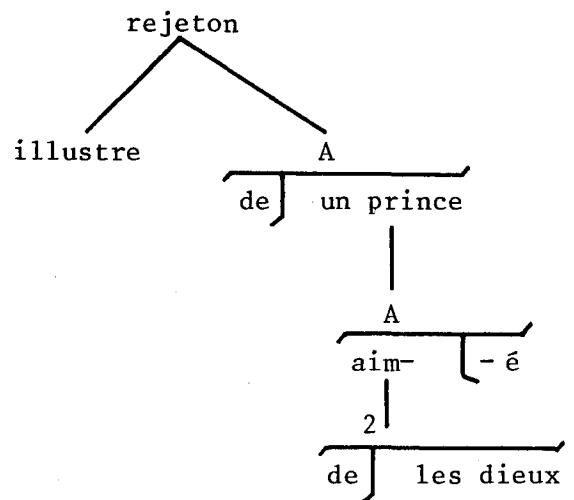
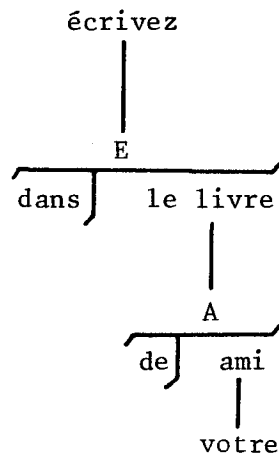
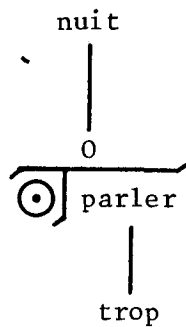
Ce type de connexion permet de rendre compte de la structure syntaxique de phrases relativement longues, comme le montre l'exemple suivant (schéma n° 1.10) :



1.10 : "Hier, les nouveaux diplomates étrangers ont livré tout-à-fait innocemment, aux journalistes français actuellement particulièrement nombreux ici, une information réellement extrêmement précieuse".

Toutefois, il suffit de modifier légèrement la forme de cet énoncé pour s'apercevoir que ces notions ne permettent pas de décrire n'importe quelle phrase simple. Par exemple, dans la phrase : "les diplomates récemment arrivés ont livré par inadvertance une information d'une grande importance", nous avons un verbe en position d'adjectif ("arrivés"), un substantif en position d'adverbe ("inadvertance"), et un substantif en position d'adjectif ("importance"). Pour rendre compte de ce phénomène, TESNIÈRE introduit la notion de *translation*, changement de nature syntaxique qui transfère un mot plein d'une catégorie grammaticale dans une autre. Ce changement de catégorie a pour conséquence un changement de fonction ; par exemple, l'adjectif transféré dans la catégorie du nom ("un jeune" ; "les autres") acquiert de ce fait la possibilité d'être actant d'un verbe.

La translation peut être marquée par un *translatif*, mot vide qui est une préposition, un préfixe, ou un suffixe ; elle peut également ne pas être marquée morphologiquement (⊙). Dans les stemmas, la translation est symbolisée par un signe en forme de T. Exemples :

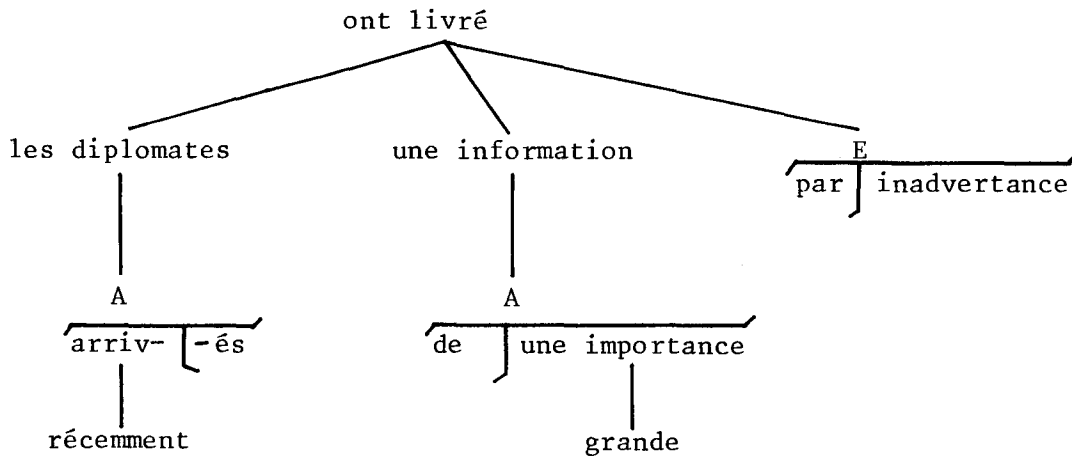


1.11: "Trop parler nuit"
(verbe → substantif: I > O)

1.12: "Ecrivez dans le
livre de votre ami"
(O > E, O > A)

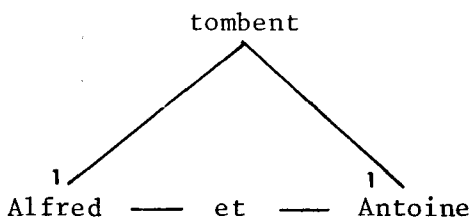
1.13: "Illustre rejeton d'un prince
aimé des dieux"
(O > A, I > A)

Cette notion de changement de catégorie grammaticale permet de représenter la structure syntaxique de la phrase citée précédemment (schéma n° 1.14) :

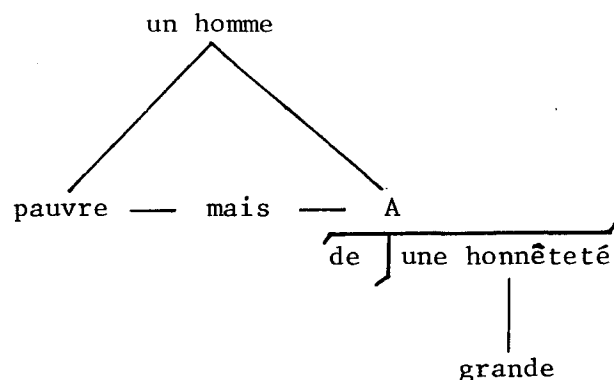


1.14 : "Les diplomates récemment arrivés ont livré par inadvertance une information d'une grande importance".

Nous venons de signaler que les translatifs sont des mots vides. Les *mots vides* sont de simples outils grammaticaux (les "struments" de DAMOURETTE et PICHON : [Damourette 1931 : I,80]), dont la fonction est d'indiquer les rapports établis entre les mots pleins. TESNIÈRE distingue deux catégories de mots vides : les *jonctifs* et les *translatifs*. Les jonctifs ont pour fonction d'unir des mots pleins qui jouent le même rôle dans la phrase. Par exemple, l'énoncé : "Alfred et Antoine tombent" correspond à une sorte d'addition [Tesnière 1959 : 106-107] des deux phrases : "Alfred tombe" et : "Antoine tombe". Il y a dédoublement du prime actant. Exemples de jonctions :



1.15. Jonction de deux primes actants (0 — 0) : "Alfred et Antoine tombent"



1.16: "Un homme pauvre mais d'une grande honnêteté" (Jonction A — A).

Les translatifs, nous l'avons vu, permettent de modifier la catégorie grammaticale d'un mot. Les *indices* sont une variété de translatifs, qui se bornent à indiquer la fonction du mot sans la modifier. Les principales espèces d'indices

sont l'article, l'indice personnel ("pronom" personnel indiquant le nombre et la personne du verbe), et les terminaisons des différentes flexions.

Nous disposons maintenant des principaux concepts indispensables pour décrire la structure syntaxique d'une phrase énonciative simple (sans proposition subordonnée). Toutefois, pour aborder l'étude des énoncés interrogatifs, il nous manque certaines notions relatives à la structure sémantique de la phrase. En principe, la fonction de la structure syntaxique est d'exprimer la structure sémantique de l'énoncé. En principe toujours, chaque noeud (syntaxique) est le support matériel d'un *nucléus*, centre sémantique regroupant, autour du mot plein qui constitue le centre structurel, tous les autres éléments (mots-vides) qui lui sont associés ; et chaque connexion syntaxique est le support d'une relation sémantique. Dans de nombreux cas, les deux structures sont parallèles. Par exemple (schéma 1.17) :

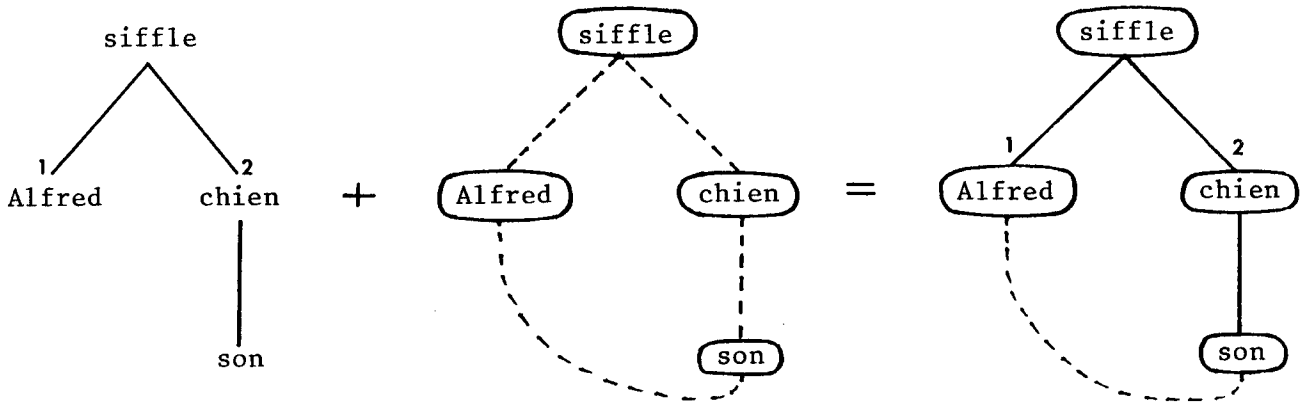


1.17 : Identité des plans syntaxique et sémantique.

Il faut toutefois signaler une différence fondamentale entre les deux structures : l'importance hiérarchique des mots sur le plan sémantique est en raison inverse de leur importance sur le plan syntaxique. "Plus un mot est bas situé sur l'échelle structurale, plus il a de chances d'être essentiel pour le sens de la phrase" [Tesnière 1959 : 43]. Tel est le cas par exemple du marquant de la négation (*ne... pas*).

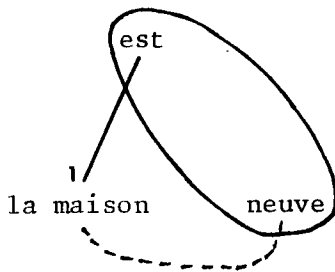
Il arrive que les deux structures ne coïncident pas. Ainsi, certaines connexions sémantiques peuvent ne pas avoir de support syntaxique. Cela est très net dans le cas de deux phrases indépendantes qui se suivent ; dans l'énoncé : "l'homme s'arrêta. Il regarda à ses pieds", la liaison anaphorique unissant *il* à

L'homme n'a pas d'équivalent syntaxique. Mais ce cas se rencontre également à l'intérieur d'une phrase (cf. schéma n° 1.18). Dans le stemma correspondant, on figurera les connexions anaphoriques par des traits tiretés. D'autre part, le

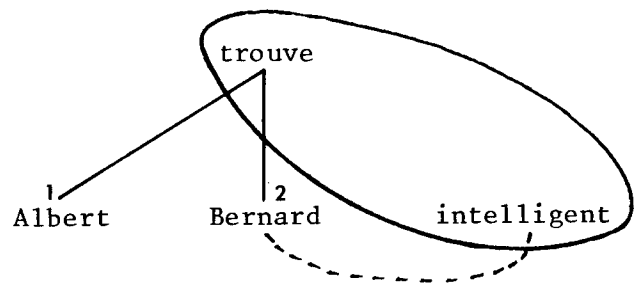


1.18 : Superposition des relations syntaxiques et sémantiques.

support syntaxique d'un nucléus peut être formé de deux mots pleins, l'un étant le noeud syntaxique et l'autre le centre du nucléus. Le cas se présente en particulier pour l'adjectif attribut (schémas n° 1.19 et 1.20) :



1.19: "La maison est neuve"



1.20: "Albert trouve Bernard intelligent"

b) L'interrogation directe.

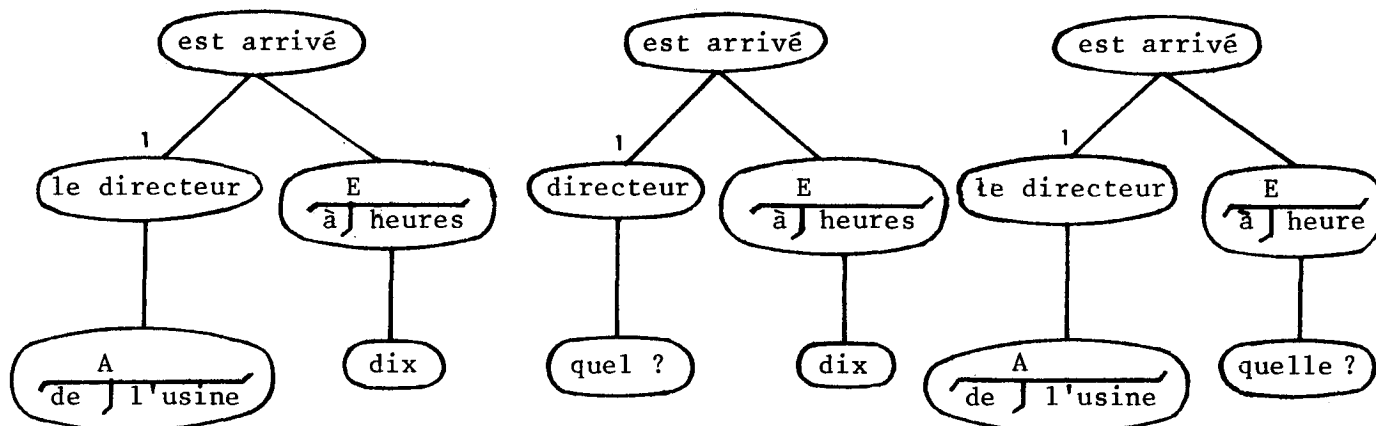
Soit la phrase verbale énonciative simple : "Le directeur de l'usine est arrivé à dix heures" ; cette phrase énonce un ensemble de faits qui sont tous présentés comme vrais. "Mais chacun de ces faits peut être mis en doute, et à chaque doute possible correspond une interrogation" [Tesnière 1959 : 191] ;

par exemple :

- "Qui est arrivé ?"
- "Quel directeur est arrivé ?"
- "Qu'a fait le directeur de l'usine ?"
- "A quelle heure le directeur de l'usine est-il arrivé ?"

Ce type de question, qui ne porte que sur une unité sémantique élémentaire, est appelé interrogation *nucléaire*. Par ailleurs, il existe un autre type de question qui porte sur la totalité de l'énoncé, et que TESNIÈRE appelle l'interrogation *connexionnelle*. Sur l'exemple précédent, la question correspondante serait : "Le directeur de l'usine est-il arrivé à dix heures ?".

L'interrogation *nucléaire* a la même structure que la phrase énonciative qui lui correspond, à ceci près que l'un des nucléus (voire deux en anglais, en russe, en latin, en grec ancien, etc.) est vide (cf. schémas 1.21 à 1.23).



1.21: "Le directeur de l'usine est arrivé à dix heures"

1.22: "Quel directeur est arrivé à dix heures ?"

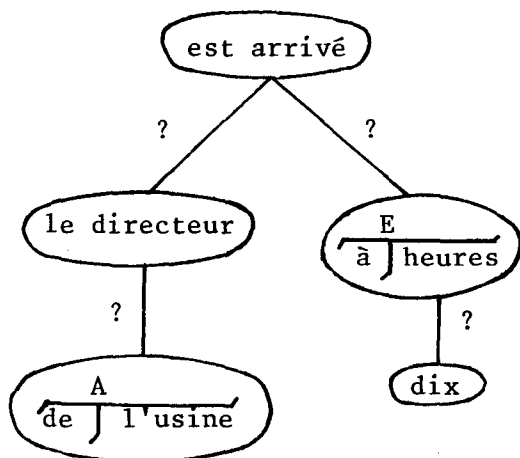
1.23: "A quelle heure est arrivé le directeur de l'usine ?"

En posant la question, le locuteur demande à l'interlocuteur de remplir le nucléus vide. Ce nucléus, sur lequel porte l'interrogation, est marqué par un *mot interrogatif*. Les mots interrogatifs appartiennent à une espèce hybride, celle des *mots pleins généraux*. Jusqu'ici, tous les mots pleins que nous avons utilisés dans nos exemples étaient des *mots pleins particuliers*, ayant à la fois un contenu

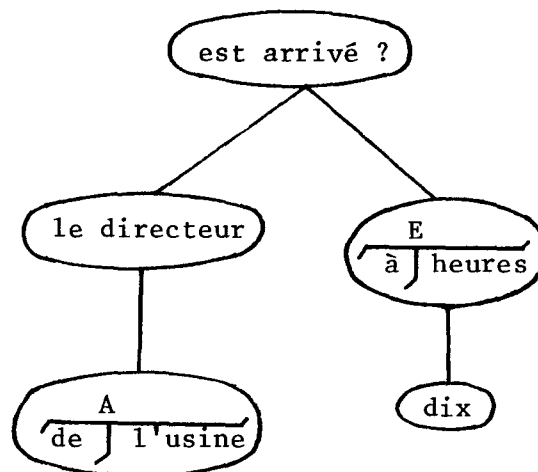
sémantique (ex. : le mot cheval peut servir à désigner un cheval particulier) et un contenu catégorique (ex. : ce même mot s'applique à l'espèce "cheval" en général). "Les mots pleins généraux au contraire n'expriment que les catégories générales qui permettent de saisir les idées particulières, à l'exclusion de ces idées mêmes. Ils ont donc un contenu *catégorique*, mais pas de contenu proprement *sémantique*" [Tesnière 1959 : 60]. Exemples de mots pleins généraux : "quelqu'un", "personne", "moi", "tel", "ce", etc., ainsi que certains emplois du verbe "faire". C'est pourquoi TESNIÈRE qualifie également l'interrogation nucléaire d'interrogation *catégorique*. A chaque catégorie de mots pleins particuliers correspond une catégorie de mots interrogatifs :

- *substantifs* généraux interrogatifs (*actanciels*) : "qui ?", "que ?", "quoi ?" ;
- *adverbes* interrogatifs (*circonstanciels*) : "où ?", "par où ?", "quand ?", "comment ?", "pourquoi ?", etc. ;
- *adjectif* interrogatif : "quel ?" ;
- *verbe* interrogatif : verbe "faire" avec comme second actant l'interrogatif actanciel "que ?" ("que fait Albert ?").

L'*interrogation connexionnelle* est exprimée par une phrase dans laquelle tous les nucléus sont pleins. Ce qui n'est pas donné, ce sur quoi porte l'interrogation, c'est la possibilité de mettre bout à bout la totalité de ces nucléus, c'est-à-dire leurs connexions. Le doute porte par conséquent non plus sur le contenu des nucléus, mais sur la structure de la phrase ; d'où le nom d'interrogation *fonctionnelle* donné également à l'interrogation connexionnelle. Pour représenter l'interrogation connexionnelle, on devrait en toute rigueur affecter un point d'interrogation à chaque connexion (cf. schéma n° 1.24). En pratique, le noeud central des phrases verbales étant le noeud verbal, il suffit de placer la marque de l'interrogation après le verbe principal (schéma n° 1.25) :



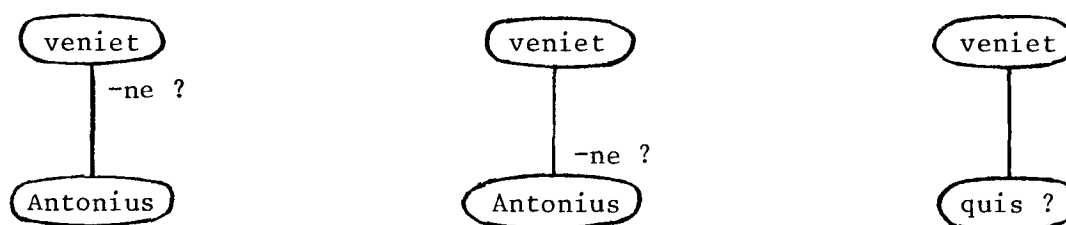
1.24: Représentation théorique de l'interrogation connexionnelle (interrogation marquée par l'intonation seule).



1.25: "Le directeur de l'usine est arrivé à dix heures ?"

En posant la question, le locuteur demande qu'on lui dise s'il y a ou non connexion unissant la totalité des nucléus. Pour répondre, on utilise deux mots conventionnels, *oui* qui signifie : "il y a connexion", et *non*, qui signifie : "il n'y a pas connexion". Ce sont des *mots anaphoriques*, mots vides qui se remplissent automatiquement du contenu de l'interrogation à laquelle ils répondent. Dans la question prise comme exemple, "oui" signifie : "le directeur de l'usine est arrivé à dix heures".

On remarquera que l'interrogation connexionnelle peut ne pas porter sur l'ensemble des connexions qui constituent la structure de la phrase, mais seulement sur l'une de ces connexions, voire sur l'une des extrémités de l'arc qui représente la connexion. Cela est particulièrement net en latin et en russe. A l'interrogation centrée sur l'arrivée de la connexion (pôle du régissant) : "venietne Antonius ?" et "придѣт ли Антон?" ("est-ce qu'Antoine viendra ?"), s'oppose l'interrogation centrée sur l'arrivée de la connexion (pôle du subordonné) : "Antoniusne veniet ?" et "Антон ли придѣт ?" ("est-ce Antoine qui viendra ?"). Ces formes se distinguent bien de l'interrogation nucléaire ; par exemple, l'interrogation portant sur le nucléus du subordonné sera : "quis veniet ?" et : "кто придѣт?" ("qui viendra ?") :



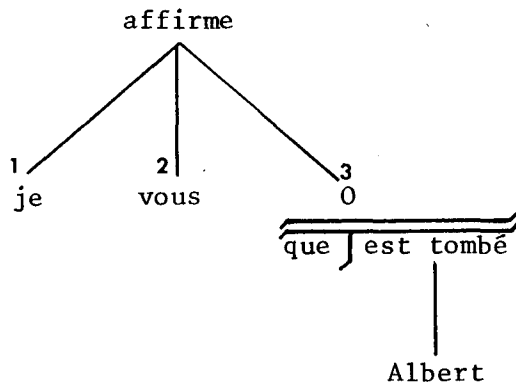
1.26: "Venietne Antonius ?" 1.27: "Antoniusne veniet ?" 1.28: "Quis veniet ?"

En français, seule l'interrogation renforcée permet d'introduire ces nuances. Or, si l'on peut admettre que la locution *est-ce que* est un mot interrogatif au même titre que *-ne* ou *ли*, et ranger "est-ce qu'Antoine viendra ?" dans les phrases verbales simples, on doit par contre considérer "est-ce Antoine qui viendra ?" comme une phrase complexe.

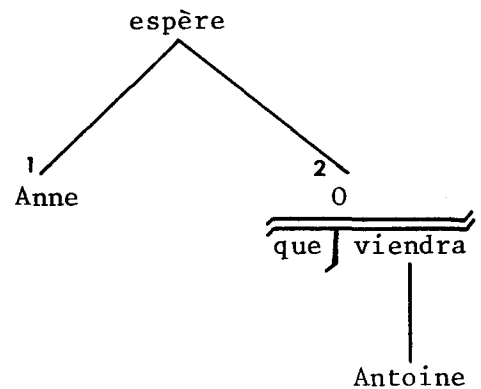
c) L'interrogation indirecte et la phrase complexe.

La phrase interrogative indirecte étant une subordonnée, il nous faut tout d'abord examiner comment TESNIÈRE rend compte des phrases complexes. Cela nous conduit à introduire la notion de *translation du second degré*. Dans les exemples de translation que nous avons présentés jusqu'ici, le mot transféré dans une autre catégorie n'était qu'un élément de la phrase simple ; il s'agissait alors de translation du premier degré. Dans la translation du second degré (ou *translation secondaire*), le mot transféré est un noeud verbal avec tous ses subordonnés, c'est-à-dire une phrase entière.

Il y a trois cas possibles de translation du second degré : transfert d'une phrase verbale dans la catégorie du substantif (I >> O), de l'adjectif (I >> A), ou de l'adverbe (I >> E). La substantivation du noeud verbal correspond au passage au style indirect, à l'aide du translatif du second degré *que* (conjonction de subordination) : "Je vous affirme : Albert est tombé" devient "Je vous affirme qu'Albert est tombé" (cf. les schémas 1.29 et 1.30) :

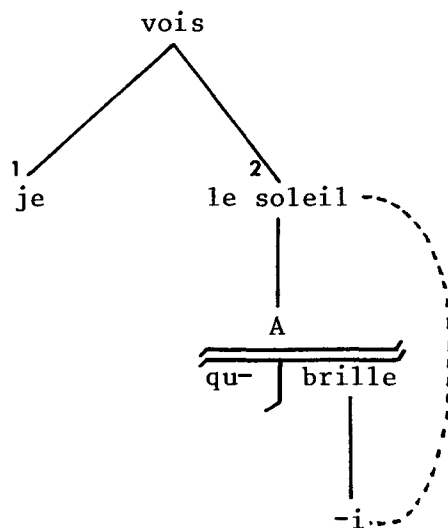


1.29: "Je vous affirme qu'Albert est tombé"

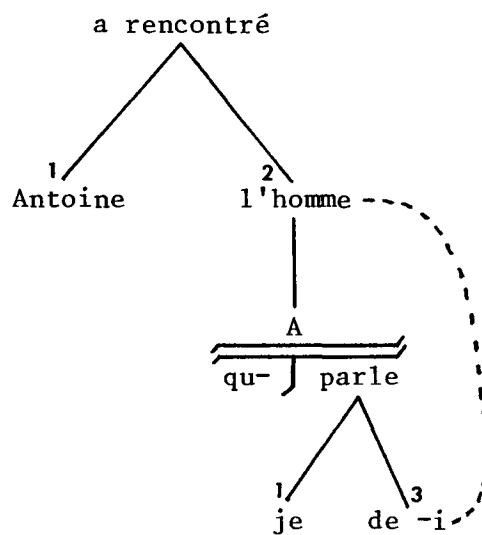


1.30: "Anne espère qu'Antoine viendra"

Le transfert du noeud verbal dans la catégorie de l'adjectif donne naissance à une subordonnée relative. En français, le marquant de la translation (pronom relatif) joue également un rôle d'actant du verbe subordonné ; on peut en conséquence le considérer comme formé de deux éléments agglutinés, un élément translatif proprement dit, invariable (*qu-*), et un élément anaphorique, variable (*-i, -e*), jouant un rôle analogue au substantif personnel (pronom personnel) (cf. schéma n° 1.31). On notera qu'en français la tournure *de qui* prend la forme synthétique *dont* : **"l'homme que je parle de lui"* devient *"l'homme de qui je parle"*, puis : *"l'homme dont je parle"* (schéma n° 1.32).

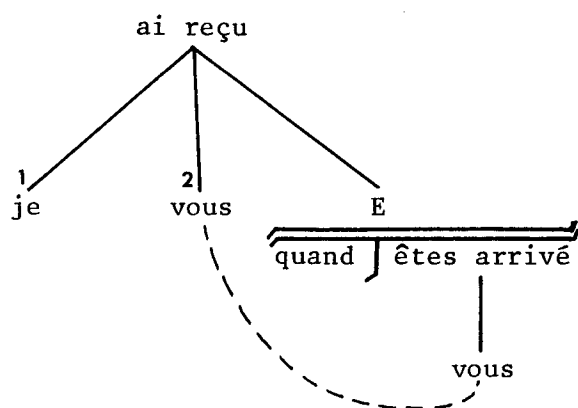


1.31: "Je vois le soleil qui brille"

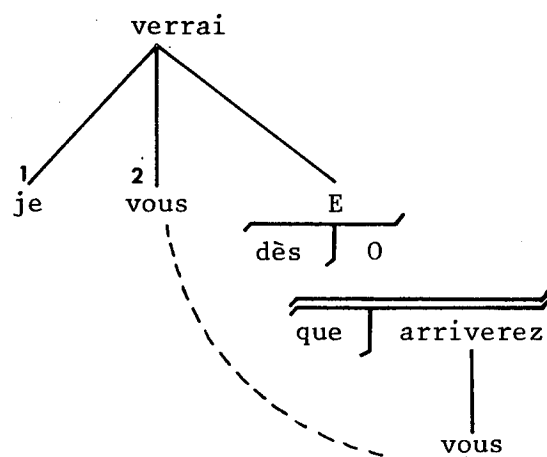


1.32: "Antoine a rencontré l'homme dont je parle"

La translation du noeud verbal en adverbe introduit une subordonnée circonstancielle : si mon interlocuteur est arrivé hier, "je vous ai reçu quand vous êtes arrivé" (schéma n° 1.33) équivaut à "je vous ai reçu hier". En français, on trouve souvent, au lieu de la translation simple $I > E$, la translation double $I >> O > E$; cette translation double est réalisée au moyen de translatifs doubles : *alors que, pendant que, par où*, etc. Soit en effet la phrase : "je vous verrai dès votre arrivée" ; le substantif "arrivée" est transféré dans la catégorie adverbiale ($O < E$) par la préposition "dès". Dans la phrase "je vous verrai dès que vous arriverez", "que vous arriverez" correspond à "votre arrivée" ($I >> O$). Il y a donc bien eu deux translations successives (schéma n° 1.34).



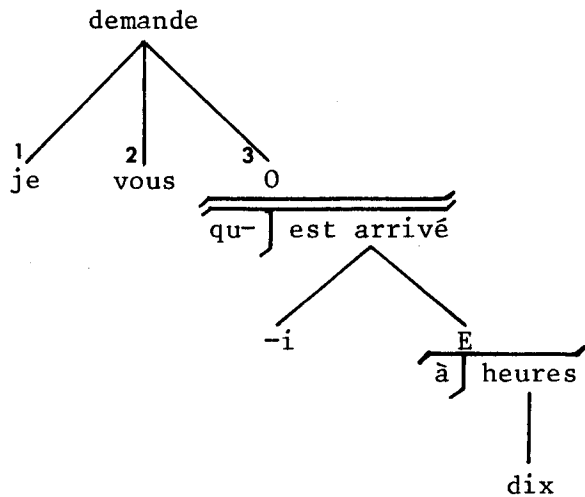
1.33: "Je vous ai reçu quand vous êtes arrivé"



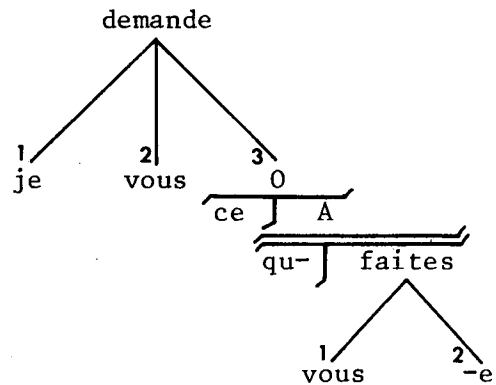
1.34: "Je vous verrai dès que vous arriverez"

La question contenue dans l'interrogation directe comporte un noeud verbal indépendant ; dans l'interrogation indirecte, ce noeud verbal devient le second actant d'un verbe de la forme : *je vous demande, je me demande, j'ignore*, etc., plaçant la phrase interrogative en position de subordination. Ainsi, l'interrogation nucléaire directe : "Qui est arrivé à dix heures ?" devient : "je vous demande qui est arrivé à dix heures" (schéma n° 1.35). Dans cet exemple, la structure de la phrase interrogative subordonnée ne change pas. Par contre, la question : "Que faites-vous ?" devient : "je vous demande ce que vous faites". Il y a eu deux translations successives : une translation du second degré ($I >> A$) plaçant l'interrogative directe en position d'adjectif (proposition relative : "que vous faites"), suivie d'une translation du premier degré ($A > O$) marquée par

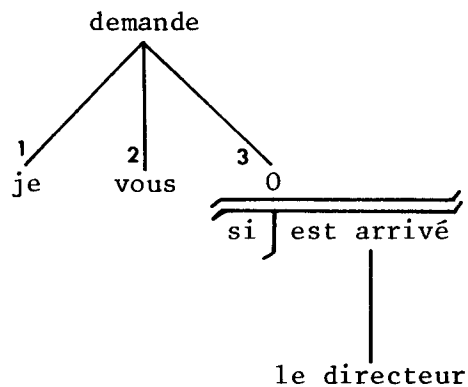
le retranslatif *ce*, transformant l'adjectif en substantif (schéma 1.36). Pour l'interrogation connexionnelle indirecte, il y a toujours (en français) changement de structure, par adjonction du marquant *si*. La question : "Le directeur est-il arrivé ?" devient : "je vous demande si le directeur est arrivé" (schéma n° 1.37).



1.35: "Je vous demande qui est arrivé à dix heures"



1.36: "Je vous demande ce que vous faites" (translation double I >> A > 0).

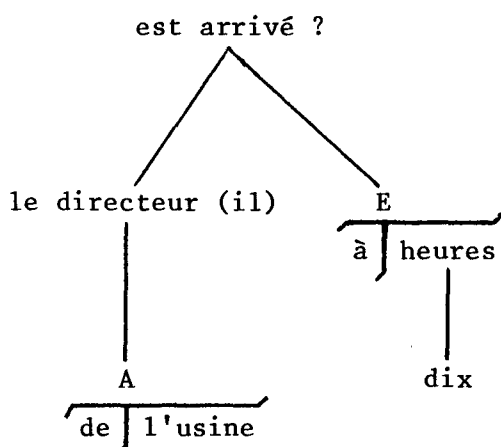


1.37: "Je vous demande si le directeur est arrivé"

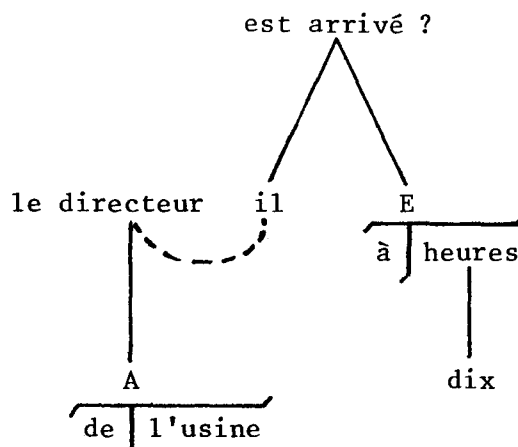
d) Vue d'ensemble de l'interrogation selon TESNIÈRE.

Lucien TESNIÈRE rend compte des différentes formes interrogatives à partir des diverses transformations possibles d'une phrase énonciative sur laquelle porte l'interrogation. Pour ces transformations, l'auteur recourt à une remarquable économie de moyens : quatre catégories de mots, quatre connexions syntaxiques seulement entre ces catégories, et deux types de translations ; le premier comportant pratiquement tous les changements possibles de catégorie, le second limité à trois formes de transfert seulement.

En reprenant l'exemple utilisé précédemment ("le directeur de l'usine est arrivé à dix heures", schéma n° 1.21), nous allons voir que la syntaxe structurale de TESNIÈRE permet de relier à la phrase énonciative qui constitue le noyau de l'interrogation, pratiquement toutes les formes de questions que nous avons décrites dans la première partie de ce chapitre. Nous avons déjà vu une version de l'interrogation connexionnelle : "Le directeur de l'usine est arrivé à dix heures ?" (schéma n° 1.25). D'autres formes sont possibles. Par exemple : "Le directeur de l'usine est-il arrivé à dix heures ?" (schéma n° 1.38), dans laquelle l'indice anontif prime actant (le pronom personnel de la troisième personne sujet du verbe), en connexion anaphorique avec le substantif prime actant, est postposé au verbe ; cette forme est très proche de : "Il est arrivé à dix heures, le directeur de l'usine ?" (schéma n° 1.39), dans laquelle il y a

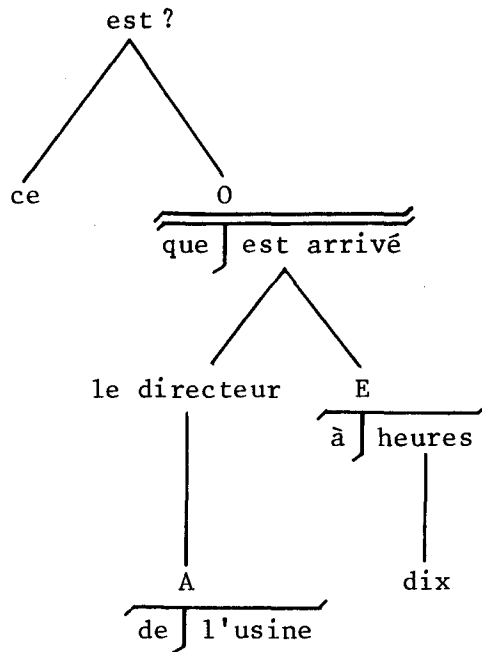


1.38: "Le directeur de l'usine est-il arrivé à dix heures?"

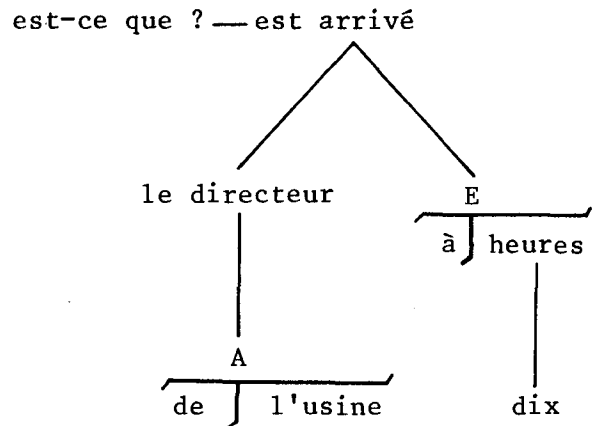


1.39: "Il est arrivé à dix heures, le directeur de l'usine ?"

projection du prime actant [Tesnière 1959 : 172-174]. L'interrogation : "Est-ce que le directeur est arrivé à dix heures ?" est susceptible de deux représentations, selon que l'on considère l'étymologie de l'expression (schéma n° 1.40), ou que l'on assimile *est-ce que* à une particule interrogative (schéma n° 1.41).

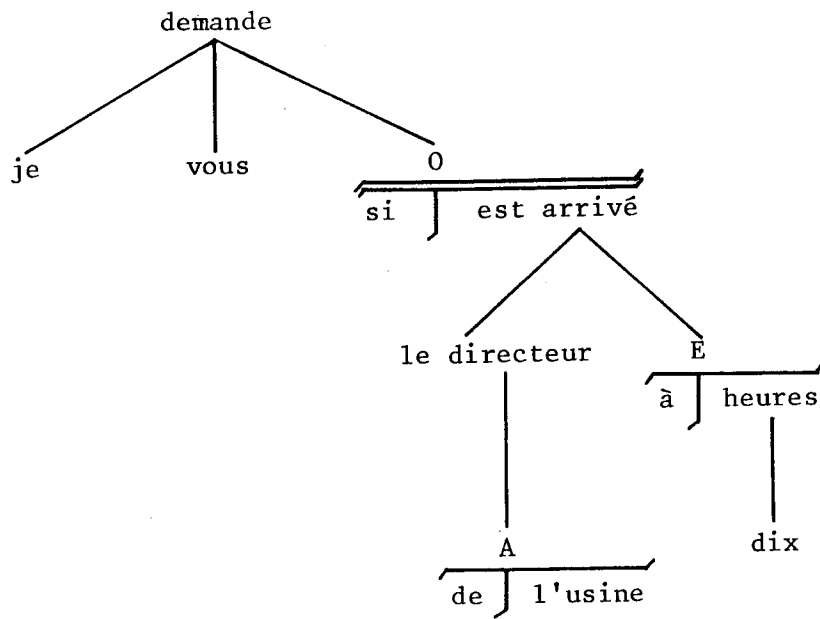


1.40: "Est-ce que le directeur de l'usine est arrivé à dix heures ?" (étymologie)

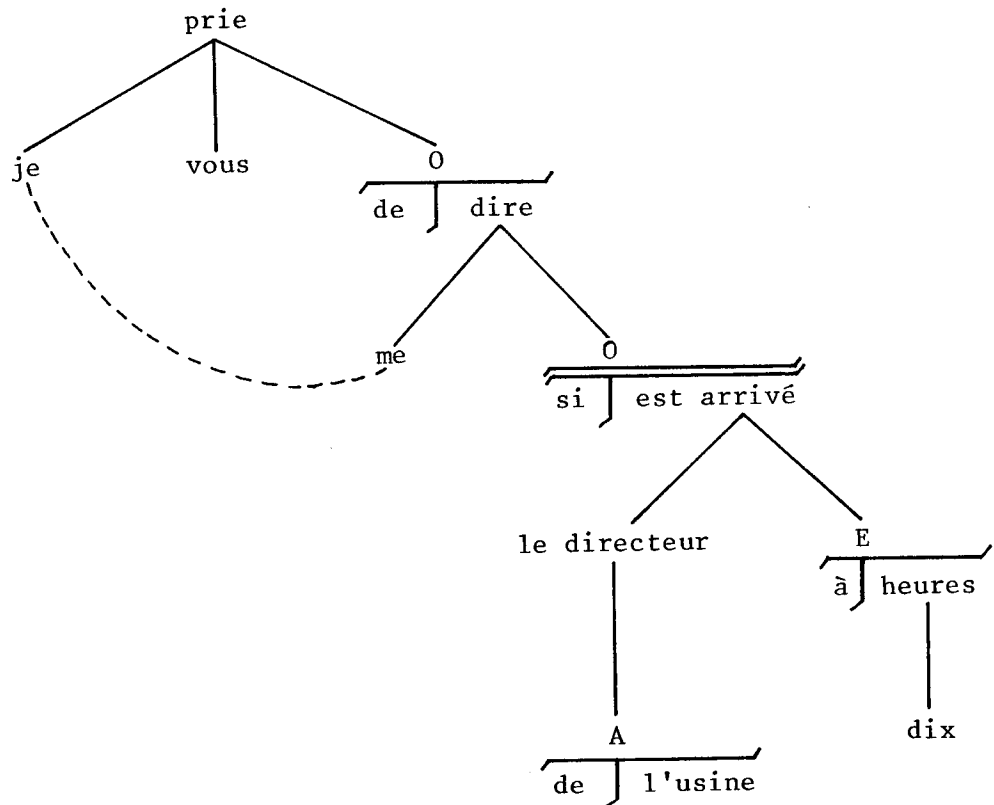


1.41: "Est-ce que le directeur de l'usine est arrivé à dix heures ?" (en considérant "est-ce que" comme un mot interrogatif).

L'interrogation indirecte correspondante est introduite par *si* : "Je vous demande si le directeur de l'usine est arrivé à dix heures" (schéma n° 1.42) ; "Je vous prie de me dire si le directeur de l'usine est arrivé à dix heures" (schéma n° 1.43). On construirait de même l'interrogation connexionnelle subordonnée à une proposition injonctive, ou à une fausse interrogative à valeur injonctive.

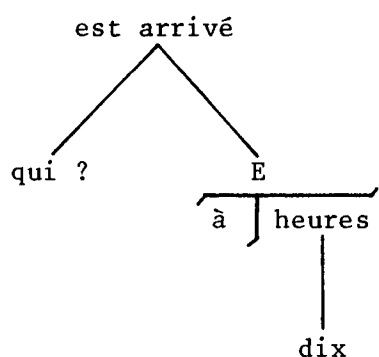


1.42: "Je vous demande si le directeur de l'usine est arrivé à dix heures"

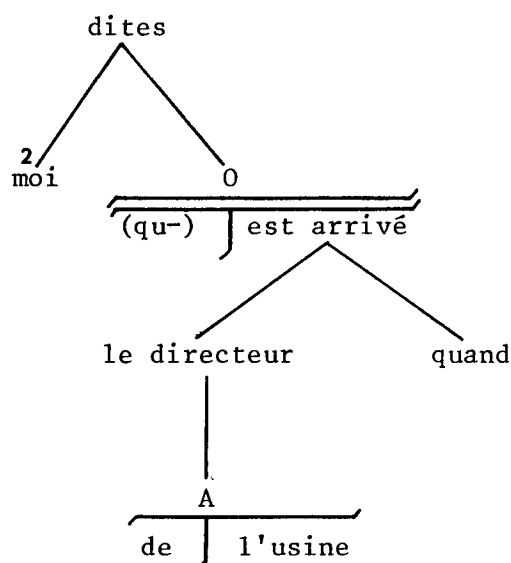


1.43: "Je vous prie de me dire si le directeur de l'usine est arrivé à dix heures".

Dans l'interrogation nucléaire, nous n'avons considéré jusqu'à présent que les cas où la question portait sur les nucléus terminaux de l'arborescence (schémas n° 1.22, 1.23, 1.35). Si l'on fait porter l'interrogation sur un autre nucléus, par exemple sur le prime actant (sujet) ou le circonstant (complément circonstanciel) du verbe, les nucléus subordonnés sont effacés dans l'énoncé correspondant. L'interrogation est alors du type : "Qui est arrivé à dix heures ?", ou "Dites-moi quand le directeur de l'usine est arrivé" (schémas n° 1.44 et 1.45). Il n'est pas tout à fait exact, dans de tels cas, de dire que l'énonciative, et l'interrogation nucléaire qui lui correspond, ont la même structure.

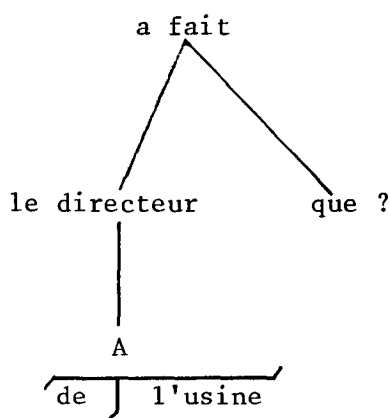


1.44: "Qui est arrivé à dix heures ?"

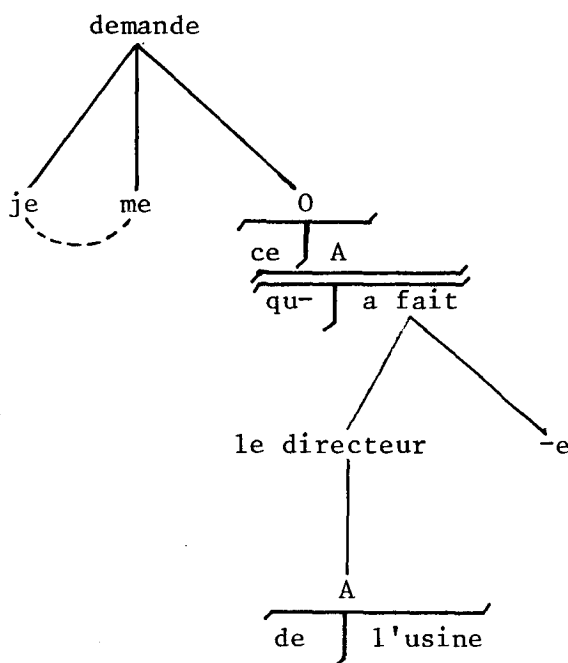


1.45: "Dites-moi quand le directeur de l'usine est arrivé".

Lorsque l'interrogation porte sur le nucléus verbal, on utilise le verbe général *faire* (cf. [Tesnière 1959 : 73, § 13]) en adjoignant au noeud verbal le substantif interrogatif actanciel *que* ou *quoi* : "Qu'a fait le directeur de l'usine ?" (schéma n° 1.46) ; au style indirect, on a recours à la double translation I >> A > O : "Je me demande ce qu'a fait le directeur de l'usine" (schéma n° 1.47).



1.46: "Qu'a fait le directeur de l'usine ?"

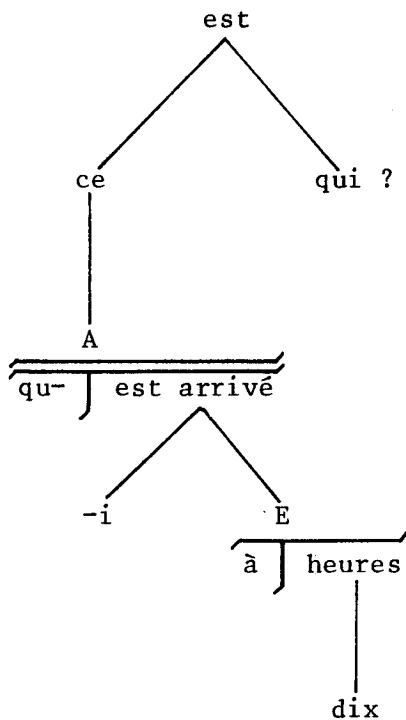


1.47: "Je me demande ce qu'a fait le directeur de l'usine"

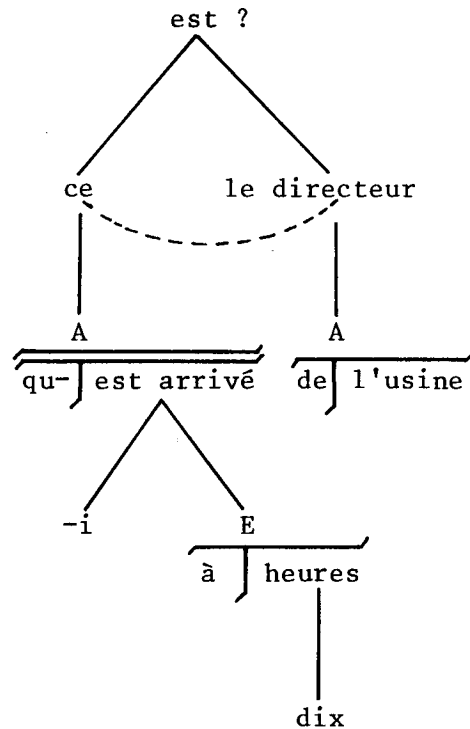
L'usage de la particule *est-ce que* dans l'interrogation nucléaire ne suscite pas de remarque particulière. On trouve des formes telles que : "Qui est-ce qui est arrivé à dix heures ?", "A quelle heure est-ce que le directeur de l'usine est arrivé ?", etc. Dans la langue parlée, qui est le plus souvent celle des tests et des questionnaires, on peut rencontrer par exemple les interrogations indirectes : "Dites-moi quand est-ce que le directeur de l'usine est arrivé", ou : "Je me demande qui est-ce qui est arrivé à dix heures".

Par contre, il existe une forme d'interrogation qui, tout en étant une interrogation connexionnelle du point de vue syntaxique, permet en fait d'interroger sur le contenu d'un nucléus. Il s'agit de questions du type "Est-ce ... qui (ou : que) ... ?". Cette forme d'interrogation n'est pas aisée à analyser ; d'autant que TESNIÈRE ne la mentionne pas dans ses *Eléments de syntaxe structurale*. Soit l'énoncé : "Est-ce le directeur de l'usine qui est arrivé à dix heures ?". C'est bien une interrogation connexionnelle (la réponse sera *oui* ou *non*), mais qui vise en fait à faire préciser par l'interlocuteur le contenu du nucléus du prime actant de la phrase énonciative correspondante. La difficulté de l'analyse tient à ce que l'antécédent de la proposition relative :

"qui est arrivé" paraît être "le directeur" ; d'ailleurs, le verbe s'accorde en nombre et, éventuellement, en genre avec cet antécédent (ex. : "... elles qui sont arrivées"). C'est d'ailleurs ainsi que l'auteur analyse les phrases du type : "c'est...qui..." [Tesnière 1959 : 635, stemma 359]. Or, il nous paraît exister un parallèle entre une question nucléaire comme : "Qui est-ce qui est arrivé à dix heures ?" (schéma n° 1.48), et la question connexionnelle : "Est-ce le directeur de l'usine qui est arrivé à dix heures ?" (schéma n° 1.49) ; c'est à dire que, dans les deux cas, l'antécédent de la relative est *ce*.

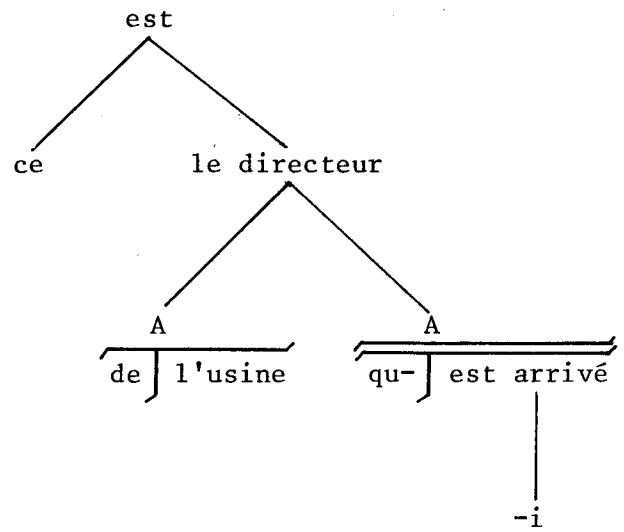
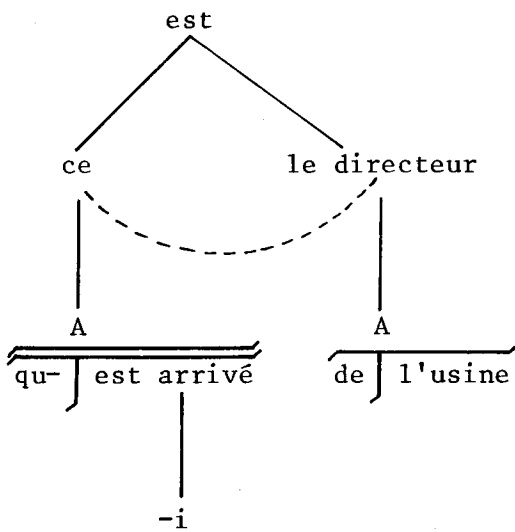


1.48: "Qui est-ce qui est arrivé à dix heures ?"



1.49: "Est-ce le directeur de l'usine qui est arrivé à dix heures ?"

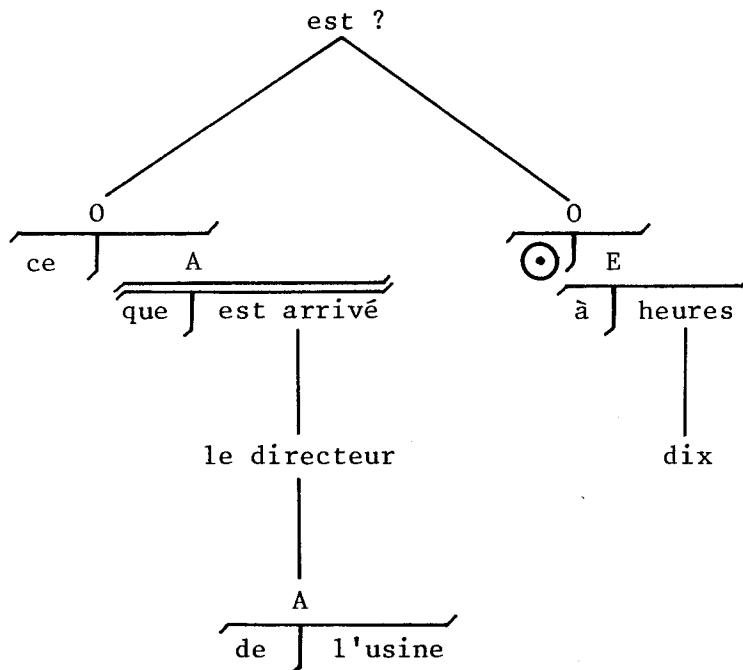
D'ailleurs, la réponse à cette dernière question serait de la forme : "Oui, c'est bien le directeur de l'usine qui est arrivé à dix heures" ; la réponse porte donc sur la connexion entre "directeur" et "est arrivé". On peut s'en convaincre en comparant la structure de la phrase : "c'est le directeur de l'usine qui est arrivé", selon qu'elle constitue une réponse à l'interrogation précédente (schéma n° 1.50), ou qu'elle répond à la question : "Quel est ce bruit ?" (schéma n° 1.51).



1.50: Réponse possible à la question :
 "Qui est-ce qui est arrivé ?" ou:
 "Est-ce le directeur de l'usine
 qui est arrivé ?"

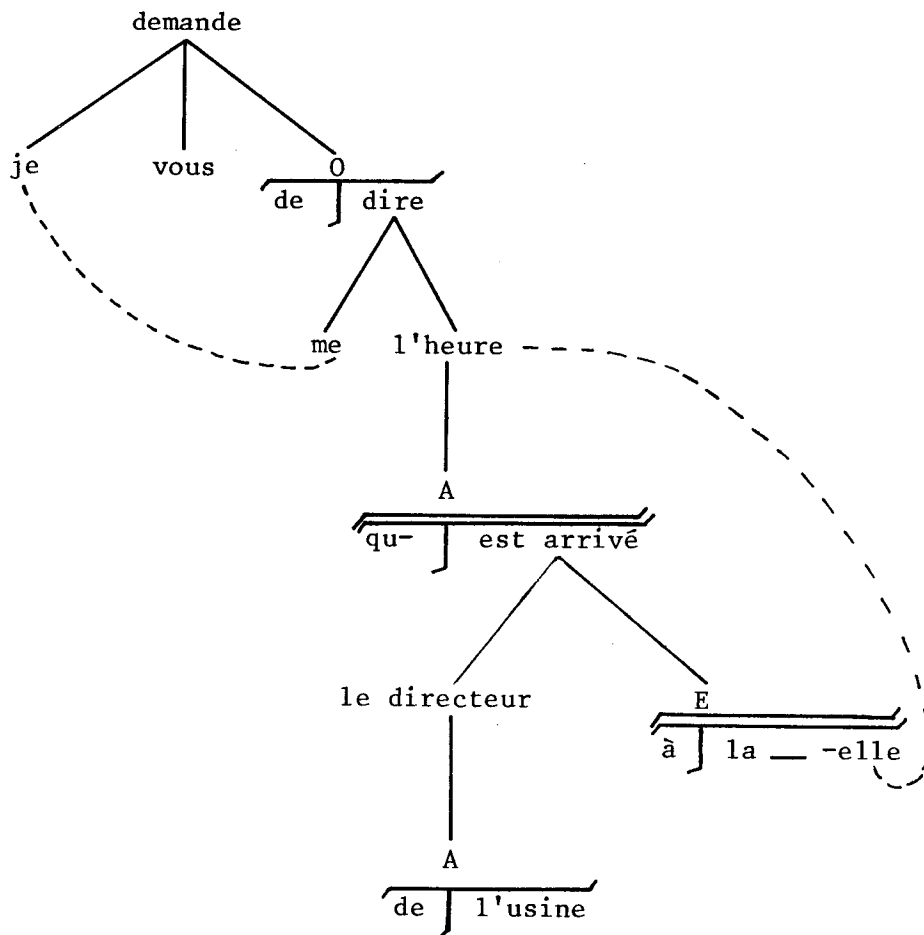
1.51: Réponse possible à la question :
 "Quel est ce bruit ?"

Enfin, on conçoit aisément que, dans la question: "Est-ce à dix heures que le directeur de l'usine est arrivé ?" (schéma n° 1.52), "heure" ne puisse être considéré comme l'antécédent de "que".

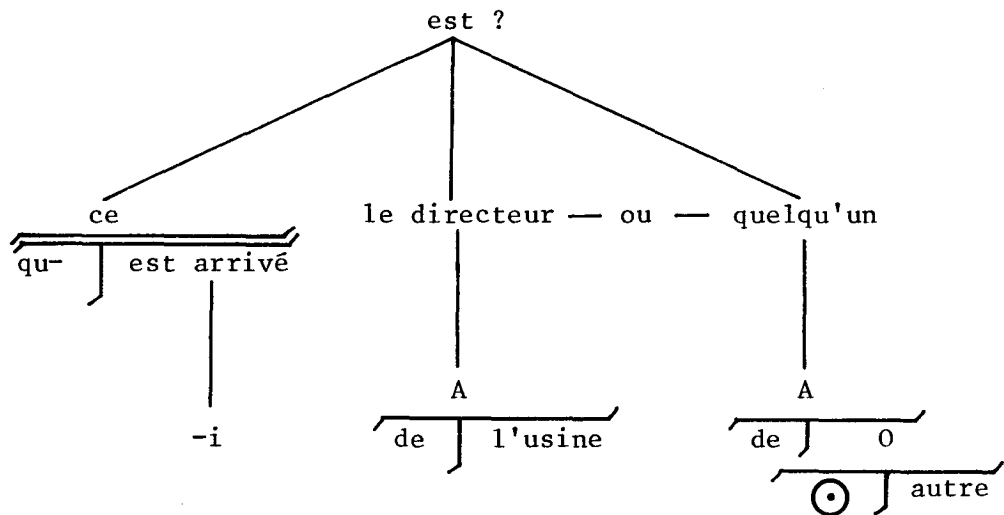


1.52: "Est-ce à dix heures que le directeur de
 l'usine est arrivé ?"

Il nous reste à voir enfin deux cas particuliers : celui où la question est une proposition énonciative ou injonctive, et celui où l'interrogation est présentée sous une forme disjonctive. Aucun de ces cas ne soulève de problème spécifique, comme on peut s'en assurer sur les deux exemples suivants : "Je vous demande de me dire l'heure à laquelle le directeur de l'usine est arrivé" (schéma n° 1.53), et : "Est-ce le directeur de l'usine, ou quelqu'un d'autre, qui est arrivé à dix heures ?" (schéma n° 1.54).



1.53: "Je vous demande de me dire l'heure à laquelle le directeur de l'usine est arrivé"



1.54: "Est-ce le directeur de l'usine, ou quelqu'un d'autre, qui est arrivé ?"

e) Bilan de l'apport de TESNIÈRE à l'analyse de l'interrogation.

Lucien TESNIÈRE n'a malheureusement pas eu le loisir de consacrer une étude d'ensemble à l'interrogation. En outre, il lui arrive parfois de proposer plusieurs représentations stemmatiques pour une même structure, selon le phénomène qu'il désire mettre en évidence. C'est pourquoi nous avons dû à la fois extrapoler certaines de ses analyses pour les appliquer à des cas qu'il n'avait pas envisagés, et les restreindre à une forme standard (moins détaillée par exemple que celle qu'il utilise dans ses stemmas intégraux à la fin de son ouvrage).

Le grand intérêt de cette analyse structurale est de présenter un petit nombre de procédures permettant de dériver l'ensemble des questions possibles, relativement à une situation concrète donnée, à partir d'un énoncé décrivant cette situation. Dans la perspective qui nous intéresse, cela montre que, sous l'extrême variété des questions envisageables, il est possible de découvrir un substrat commun. Il n'est donc pas tout à fait utopique de vouloir réduire un ensemble de questions (dont les caractéristiques restent à préciser) à une forme unique dont celles-ci sont dérivées. De plus, à lire TESNIÈRE, on a le sentiment que les procédures de réduction à cette forme unique pourraient être essentiellement syntaxiques (voire exclusivement), et susceptibles de s'appliquer à l'en-

semble des langues naturelles. Toutefois, du point de vue qui est le nôtre, les procédures décrites par TESNIÈRE présentent l'inconvénient d'être centrées sur le locuteur, comme si le problème posé était de comprendre comment interroger autrui pour obtenir de lui une information. C'est là d'ailleurs une optique que nous avons retrouvée dans la quasi-totalité des grammaires usuelles, du moins en ce qui concerne l'interrogation. Ce que nous aurions souhaité, c'est disposer d'un outil morpho-syntaxique nous permettant, pour un énoncé de question donné, de ré-écrire celui-ci sous une forme préétablie, en vue de comparer son contenu sémantique à celui d'autres énoncés, quelle que soit leur forme originelle.

Tel n'était évidemment pas l'objectif de TESNIÈRE. On peut pourtant se demander s'il n'est pas possible de faire fonctionner ces procédures déductives en sens inverse, afin de découvrir la phrase énonciative qui sous-tend l'interrogation. Cela est probablement réalisable pour l'interrogation connexionnelle, dont tous les nucléus sont pleins. Par contre, le problème se posera de remplir le nucléus vide de l'interrogation nucléaire ; voire, problème plus difficile encore, de reconstituer les nucléus qui, dans la phrase énonciative correspondante, sont subordonnés au nucléus vide (et, par conséquent, généralement effacés dans l'interrogation). Une solution à étudier consisterait à remplacer le mot interrogatif du nucléus vide par un mot plein général ("quelqu'un", "quelque chose", "d'une certaine manière", "à un certain moment"). Nous reviendrons sur ce point dans la conclusion de ce chapitre.

Parmi les autres apports de la syntaxe structurale, il faut signaler la description du mécanisme de l'interrogation connexionnelle opposée à l'interrogation nucléaire. Il n'est pas sans importance en effet de savoir que la notion de portée de l'interrogation n'est pas un concept purement descriptif, mais qu'elle correspond à une différence fondamentale, traduisant la possibilité de deux projets distincts de la part du locuteur. Nous avons vu toutefois que les multiples variantes de l'interrogation connexionnelle introduisaient des nuances assez proches de certaines formes d'interrogation nucléaire ; et ce, non seulement en latin ou en russe, comme le montre TESNIÈRE, mais également en français. Ainsi, à : "Le directeur est-il arrivé ?", on pourra comparer : "Est-ce le directeur qui est arrivé ?" et : "Est-ce qu'il est arrivé, le directeur ?". Cet exemple (parmi beaucoup d'autres) de la relative indépendance du sémantique et du structural nous conduira à ré-examiner ultérieurement le problème de la portée de l'interrogation.

1.2.2. Les grammaires génératives

Les travaux de Noam CHOMSKY sont tellement connus qu'il pourrait paraître superflu d'en rappeler les principaux aspects. Il nous a semblé nécessaire toutefois, avant d'examiner l'application d'un modèle génératif à l'analyse des interrogations, de décrire le cadre théorique dans lequel se situent les grammaires génératives, ne serait-ce qu'en raison des bouleversements que la perspective chomskyenne a provoqué dans les conceptions linguistiques traditionnelles, et du nombre considérable des travaux de recherche qu'elle a inspirés.

La difficulté d'une telle description naît du foisonnement des modèles génératifs présentés dans la littérature scientifique ; mais elle se trouve accrue en raison de leur caractère souvent fragmentaire, et explicitement hypothétique. CHOMSKY lui-même, dans ses nombreuses publications, a remis en cause certaines des conceptions qu'il avait développées à des stades antérieurs de sa réflexion ; conceptions qu'il avait d'ailleurs pris soin de présenter comme des hypothèses à éprouver, et non comme des résultats assurés. En outre, il n'a pas jusqu'ici proposé une vue d'ensemble de la grammaire d'une langue (l'anglais en l'occurrence) ; il a préféré ouvrir à la recherche linguistique des perspectives stimulantes mais partielles, et parfois contradictoires. Au point qu'on a pu écrire que CHOMSKY "est ambigu, définit peu, varie d'un endroit à l'autre", et qu' "il y aura toujours chez lui, comme dans la Bible, une phrase qu'on n'aura pas citée, et qui dit le contraire de ce qu'on fait dire à sa doctrine en général" [Mounin 1972 : 222].

C'est pourquoi nous nous limiterons à une description succincte des principes des grammaires génératives, tels qu'ils apparaissent à travers CHOMSKY, ou à travers ses disciples et commentateurs. Nous présenterons ensuite les grandes lignes de la syntaxe du français de DUBOIS et DUBOIS-CHARLIER, et la manière dont ce modèle génératif rend compte des formes interrogatives. Nous concluons sur l'intérêt que présentent les grammaires génératives pour notre recherche.

a) La perspective générative en linguistique.

Noam CHOMSKY appelle *langue* "un ensemble (fini ou infini) de phrases, chacune d'elles étant de longueur finie et composée d'un ensemble fini d'éléments. Toutes les langues naturelles, sous leur forme écrite ou parlée, répondent à cette définition, puisque toute langue naturelle comporte un nombre fini de phonèmes (ou de lettres dans son alphabet) et que chaque phrase peut être représentée comme une suite finie de ces phonèmes (ou lettres) bien que le nombre de phrases soit infini" [Chomsky 1957 : 15]. Si l'on considère que l'ensemble des phrases (correctes grammaticalement) possibles dans une langue donnée est infini, on ne peut fonder une grammaire sur la seule analyse d'un corpus d'énoncés, quelle que soit sa dimension ; au contraire, "la grammaire d'une langue doit *projeter* le corpus fini et toujours plus ou moins accidentel des énoncés observés sur l'ensemble (présupposé infini) des phrases grammaticales" [Chomsky 1957 : 17]. Pour ce faire, il n'y a d'autre choix que de partir de l'ensemble fini des éléments de la langue.

Si l'on considère l'ensemble (fini) des chaînes (d'une longueur donnée, ou d'une longueur inférieure ou égale à un seuil donné) que l'on peut former à l'aide de ces éléments, on constate qu'une partie seulement des phrases ainsi obtenues est conforme aux règles syntaxiques (indépendamment de leur signification). La linguistique générative s'efforce de dégager les règles permettant d'*engendrer*, à partir des éléments de la langue, les phrases correctes grammaticalement et elles seules. On pourrait, de ce fait, considérer les grammaires génératives comme des modèles simulables visant à reproduire le comportement du locuteur idéal [Mounin 1972 : 221], un tel locuteur étant supposé connaître parfaitement la langue qu'il parle, et ne pas commettre d'erreurs lorsqu'il s'exprime dans cette langue [Chomsky 1965 : 12].

Les grammaires génératives, comme toute théorie syntaxique, ne s'intéressent pas, en principe, au sens des mots. Ainsi, du point de vue de la syntaxe, la phrase : "le silence vertébral indispose la voile licite" est aussi correcte que : "le signal vert indique la voie libre", sur le modèle de laquelle elle a été créée [Tesnière 1959 : 41-42] ; alors que : "signal vert indiquer voie libre", malgré sa signification perceptible immédiatement et sans ambiguïté, n'appartient pas à la langue française. Pour simuler la production de phrases

grammaticalement correctes dans une langue donnée, on aurait pu imaginer un automate capable d'engendrer (aléatoirement ou systématiquement) des chaînes quelconques d'éléments de la langue, et d'éliminer ensuite les phrases non grammaticales ; les règles de sélection des phrases grammaticales constituant alors un modèle de la syntaxe de la langue. Si cette voie n'a pas été suivie par la linguistique générative, c'est probablement en raison du pouvoir explicatif très réduit qu'aurait présenté une telle syntaxe.

En effet, CHOMSKY propose de tester l'adéquation d'une grammaire donnée par son aptitude à rendre compte des cas d'ambiguïté dus à une homonymie de construction [Chomsky 1957 : 93-100]. Depuis le "dico Romanos Carthaginienses vincere" de l'oracle, les exemples d'ambiguïté syntaxique sont nombreux. Citons : "flying planes can be dangerous" [Chomsky 1965 : 38] ; "the police were ordered to stop drinking about midnight" [Crystal 1971 : 211] ; "la belle ferme le voile" [Moreau 1975 : 100] ; "j'achète un livre à mon fils" [Dubois 1969 : 18] ; "Pierre propose à Jean de venir" [Ruwet 1967 : 293] ; etc. Il est probable que les règles de sélection des phrases grammaticales parmi l'ensemble des chaînes possibles ne permettraient pas de rendre compte de ces ambiguïtés de façon simple. Par contre, on peut se fixer comme objectif de concevoir un automate susceptible d'engendrer les seules phrases grammaticalement correctes à partir de règles de combinaison d'éléments de base, et sans faire intervenir de règles de sélection. Un automate de ce type est une grammaire générative.

L'éventualité d'énoncés ambigus par construction conduit à distinguer la structure syntaxique *de surface* (caractéristiques syntaxiques de l'énoncé) et la structure syntaxique *profonde*, qui doit être dépourvue d'ambiguïté. Pour reprendre l'un des exemples ci-dessus, l'ambiguïté de "Pierre propose à Jean de venir" vient de ce que cet énoncé peut provenir de deux structures profondes différentes, l'une signifiant que Pierre dit à Jean : "si tu le désires, je viendrai", et l'autre, que Pierre lui dit : "viens si tu le désires". On comprend sur cet exemple que ce sont les structures profondes qui déterminent la représentation sémantique de la phrase [Chomsky 1972 : 13]. Si l'on suppose définie la structure profonde des énoncés, une grammaire générative doit ensuite énoncer un ensemble de règles permettant de dériver d'une structure profonde donnée la ou les structures de surface correspondantes. Plus l'on s'efforcera de simplifier la grammaire, c'est-à-dire de

réduire le nombre des règles de dérivation, plus la structure profonde devra être abstraite [Ruwet 1967 : 272-273] ; il y a donc plusieurs grammaires génératives possibles pour une langue donnée. D'autre part, pour qu'un automate simulant le fonctionnement du locuteur idéal soit capable d'engendrer des phrases en nombre infini à partir d'un nombre fini d'éléments de base, il est nécessaire d'introduire la récursivité dans les processus de dérivation.

Pour résumer, une grammaire générative est un modèle hypothético-déductif [Mounin 1972 : 205], défini par la donnée d'un vocabulaire terminal fini V_T , d'un vocabulaire auxiliaire fini V_A , d'un élément initial $\Sigma \in V_A$, et d'un nombre fini de règles de dérivation ; ces règles sont de la forme $A \rightarrow \phi$, avec $A \in V_A$ et $\phi \in \{V_A \cup V_T\}^*$ (cf. [Gross 1967 : 79]), le symbole E^* désignant l'ensemble des chaînes formées par concaténation d'éléments de E . Les règles de dérivation sont des règles dites "de ré-écriture", énonçant comment un élément (unique) du vocabulaire auxiliaire se ré-écrit sous la forme d'une chaîne d'éléments empruntés à l'ensemble réunissant le vocabulaire auxiliaire et le vocabulaire terminal. Pour que le langage ainsi engendré ne soit pas vide, il faut qu'il existe au moins une règle terminale dont la partie droite soit une chaîne ne comprenant aucun élément du vocabulaire auxiliaire : règle de la forme $A \rightarrow \tau$, avec $A \in V_A$ et $\tau \in V_T^*$ [Gross 1967 : 80].

Pour illustrer ces notions, nous empruntons à CHOMSKY un très court exemple [Chomsky 1958 : 213-214], qui présente en outre l'intérêt de mettre en évidence la propriété de récursivité des grammaires génératives. Soit la grammaire suivante :

$V_T = \{came, man, that, the, unfortunate, was\}$

$V_A = \{Adj, Art, Cop, N, SN, SV, V, \Sigma\}$

Règle 1 : $\Sigma \rightarrow SN + SV$

Règle 2 : $SN \rightarrow that + \Sigma$

Règle 3 : $SN \rightarrow Art + N$

Règle 4 : $SV \rightarrow Cop + Adj$

Règle 5 : $SV \rightarrow V$

Règle 6 : $Adj \rightarrow unfortunate$

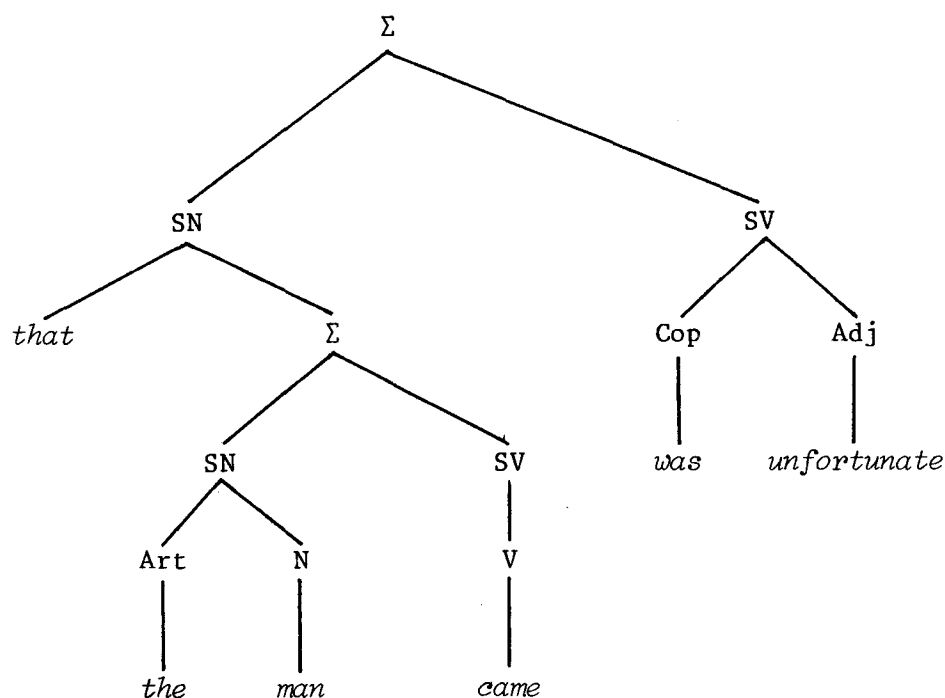
- Règle 7 : Art \rightarrow *the*
 Règle 8 : Cop \rightarrow *was*
 Règle 9 : N \rightarrow *man*
 Règle 10 : V \rightarrow *came*

Les règles 6 à 10 sont les règles terminales de la grammaire. L'opérateur \rightarrow indique la ré-écriture, l'opérateur + la concaténation. Les mots du vocabulaire auxiliaire se lisent : Adj = adjectif, Art = article, Cop = copule, N = nom, SN = syntagme nominal, SV = syntagme verbal, V = verbe, Σ = phrase.

Un premier exemple de réalisation de cette grammaire pourrait être obtenu par l'application successive des règles 1,3,5,7,9,10 :

- (1) Σ \rightarrow SN + SV
 (3) \rightarrow Art + N + SV
 (5) \rightarrow Art + N + V
 (7) \rightarrow *the* + N + V
 (9) \rightarrow *the* + *man* + V
 (10) \rightarrow *the* + *man* + *came*

Soit la phrase : "the man came". On aurait de même, par application (dans l'ordre indiqué) des règles 1,3,4,7,9,8, et 6 : "the man was unfortunate". La récursivité est introduite dans cette grammaire par la règle 2, dont l'application conduit à l'enchâssement d'une phrase à l'intérieur d'une autre phrase, comme élément de celle-ci. Par exemple, par application successive des règles non terminales 1,2, 1,3,5,4, puis de l'ensemble des règles terminales, on engendrera la phrase : "that the man came was unfortunate". La structure syntaxique de cette phrase peut être représentée graphiquement par une arborescence dont la racine est le symbole initial Σ , et les sommets terminaux, des éléments du vocabulaire terminal (cf. schéma n° 1.55).



1.55 : Représentation en arbre correspondant à l'énoncé "that the man came was unfortunate".

L'intérêt pour notre recherche de la perspective ouverte par CHOMSKY réside surtout dans la possibilité de montrer qu'un petit nombre de formes de structures profondes permet de rendre compte de la variété des énoncés à valeur de question tels qu'ils se rencontrent dans les tests et les questionnaires. La difficulté à laquelle nous nous trouvons confronté tient d'abord au fait que CHOMSKY lui-même n'a pas encore présenté une grammaire générative complète, susceptible de rendre compte de la totalité des règles syntaxiques d'une langue naturelle (l'anglais, en l'occurrence) ; ensuite, à ce que les principaux modèles (fragmentaires) proposés par CHOMSKY ne s'appliquent qu'à la langue anglaise ; enfin, à la multiplicité des tentatives pour constituer une grammaire générative du français. Nous avons en définitive opté, afin de donner une vue d'ensemble du problème, pour la grammaire générative de la langue française traitant de la plus grande variété d'énoncés : les *Eléments de linguistique française : syntaxe*, de Jean DUBOIS et Françoise DUBOIS-CHARLIER. En l'état actuel des résultats apportés par la linguistique générative, les auteurs avaient à choisir entre une recension des questions en suspens, ou un exposé synthétique, mais élémentaire, des notions

de base appliquées au français ; c'est cette seconde solution, plus risquée mais plus utile pour nous, qu'ils ont adoptée.

b) La syntaxe du français de DUBOIS et DUBOIS-CHARLIER

Le modèle générateur que nous décrivons comporte : un ensemble de règles définissant les relations grammaticales entre les éléments qui constituent les structures profondes (*composante catégorielle*) ; un ensemble de règles lexicales permettant le remplacement des symboles catégoriels de la structure profonde par des mots du lexique ; un ensemble de règles de transformation régissant le passage de la structure profonde à la structure de surface ; et enfin un ensemble de règles phonologiques ou phonétiques qui déterminent la réalisation de la structure de surface abstraite en une phrase de la langue.

Le tableau n° 1.56 ci-après présente les principales règles de réécriture des constituants catégoriels de la structure profonde de la phrase de base (règles n° 1 à 30), ainsi que quelques règles lexicales (règles n° 31 à 45). Comme dans le paragraphe précédent, le symbole \rightarrow correspond à la ré-écriture, et le symbole + à la concaténation. En outre, l'accolade est utilisée dans les règles syntaxiques pour indiquer une alternative entre plusieurs constituants : par exemple (règle 3), le constituant de phrase est formé d'un premier élément qui est soit Affir, soit Inter, soit Imp. Les parenthèses indiquent les éléments facultatifs : dans la règle de ré-écriture du constituant de phrase, il peut y avoir ou ne pas y avoir le constituant Nég, le constituant Emph, le constituant Passif ; par contre, les constituants Affir, Inter, et Imp étant obligatoires, l'un d'eux est nécessairement présent dans la structure profonde. Le signe + placé en exposant exprime la possibilité de répéter le symbole correspondant : ainsi, dans la ré-écriture du noyau (règle n° 11), il peut y avoir plusieurs SP (syntagmes prépositionnels) concaténés. Enfin, les *mots* écrits en italique sont des mots du langage, c'est-à-dire des éléments du vocabulaire terminal ; en toute rigueur, nous aurions dû, dans la règle de ré-écriture du constituant Nég (règle n° 6), mentionner le symbole Ne comme premier élément, et énoncer ensuite la règle lexicale : Ne \rightarrow *ne* (cette remarque vaut également pour les règles syntaxiques 9, 17, 18, 23, 29, et 30).

- (1) $\Sigma \rightarrow$ Const + P
- (2) Const \rightarrow $\left\{ \begin{array}{l} \text{Affir} \\ \text{Inter} \\ \text{Imp} \end{array} \right\} + (\text{Nég}) + (\text{Emph}) + (\text{Passif})$
- (3) Affir \rightarrow Inton_{affir}
- (4) Inter \rightarrow $\left\{ \begin{array}{l} (\text{SN}_{\text{inter}}) + \text{Inton}_{\text{inter}} \\ (\text{SN}_{\text{inter}}) + \text{SN}_{\text{que}} + (\text{Inton}_{\text{inter}}) \end{array} \right\}$
- (5) Imp \rightarrow Prés_{imp} + (Adv_{imp}) + Inton_{imp}
- (6) Nég \rightarrow ne + Adv_{nég}
- (7) Emph \rightarrow $\left\{ \begin{array}{l} \text{SN} \\ \text{SA} \\ \text{SP} \end{array} \right\} + \text{Accent}_{\text{emph}}$
- (8) Passif \rightarrow Aux_{être} + SP_{pas}
- (9) Aux_{être} \rightarrow être + PP
- (10) SP_{pas} \rightarrow Pré_p + SN_{pas}
- (11) P \rightarrow SN+SV+(SP)⁺
- (12) SN \rightarrow No+GN
- (13) No \rightarrow $\left\{ \begin{array}{l} \text{sing} \\ \text{plur} \end{array} \right\}$
- (14) GN \rightarrow D+N
- (15) D \rightarrow (PréArt)+(Dém)+Art+(PostArt)
- (16) Art \rightarrow $\left\{ \begin{array}{l} \text{Déf} \\ \text{Indéf} \end{array} \right\}$
- (17) PréArt \rightarrow $\left\{ \begin{array}{l} \text{Quanta} \\ \text{QuantR} + \text{de} \end{array} \right\}$
- (18) PostArt \rightarrow $\left(\left\{ \begin{array}{l} \text{Cardinal} \\ \text{Numéral} \end{array} \right\} + (\text{Dénotatif}) \right)$
- (19) SV \rightarrow Aux + GV
- (20) Aux \rightarrow Tps +(Parf)+(M)+(Parf)
- (21) Tps \rightarrow $\left(\left\{ \begin{array}{l} \text{Futur} \\ \text{Subj} \end{array} \right\} \right) + \left\{ \begin{array}{l} \text{Prés} \\ \text{Pas} \end{array} \right\} + \text{Pe} + \text{No}$
- (22) Pas \rightarrow $\left\{ \begin{array}{l} \text{Imparfait} \\ \text{Pas}_{\text{hist}} \end{array} \right\}$
- Σ = phrase de base (symbole initial ; cf : 2, 11)
- Const = constituant de phrase (cf : 3 à 8)
- Affir = affirmation (Inton_{affir} = intonation affirmative)
- Inter = interrogation (cf : 30 ; SN_{inter} = syntagme nominal (cf:12) interrogatif, identique au SN sujet de P ; Inton_{inter} = intonation interrogative)
- Imp = impératif (cf : 44 ; Prés_{imp} = présent impératif ; Inton_{imp} = intonation impérative)
- Nég = négation (cf : 43)
- Emph = emphase (cf : 12,26,28 ; Accent_{emph} = accent emphatique)
- Passif = passif (cf : 9,10)
- Aux_{être} = auxiliaire être (PP = participe passé)
- SP_{pas} = syntagme prépositionnel passif (cf : 29 ; SN_{pas} = syntagme nominal passif, constituant vide dont la place sera occupée par le SN sujet de P)
- P = noyau (cf : 12,19,28)
- SN = syntagme nominal (cf : 13,14)
- No = nombre (sing=singulier ; plur = pluriel)
- GN = groupe nominal (cf : 15 ; N=nom)
- D = déterminant (cf : 16 à 18, 40)
- Art = article (cf : 41,42)
- PréArt = préarticle (cf : 31,32)
- PostArt = postarticle (cf : 33,34,39)
- SV = syntagme verbal (cf : 20,25)
- Aux = auxiliaire (cf : 21,23,24)
- Tps = temps, personne, nombre (cf : 13,22 ; Futur = futur ; Subj = subjonctif ; Prés = présent ; Pe = personne)
- Pas = passé (Imparfait = imparfait ; Pas_{hist} = passé historique ou passé simple)

- (23) Parf → $\left. \begin{array}{l} \text{avoir} \\ \text{être} \end{array} \right\} + \text{PP}$
- (24) M → $\left. \begin{array}{l} \text{Mod} \\ \text{Asp} \end{array} \right\} + \text{Inf}$
- (25) GV → $\left. \begin{array}{l} \text{Cop} + \left\{ \begin{array}{l} \text{SN} \\ \text{SA} \\ \text{SP} \end{array} \right\} \\ \text{V}+(\text{SN})+(\text{SP}) \end{array} \right\} + \text{GA} + (\text{SP})$
- (26) SA → (Adv_{degré}) + GA + (SP)
- (27) GA → Adj + (SP)
- (28) SP → Prép_p + SN
- (29) Prép_p → *de* + (Prép) + (Adv)
- (30) SN_{que} → *que* + Déf + (quel) + N

REGLES DE RÉ-ÉCRITURE LEXICALES

- (31) Quanta → { *tous, tout, toute, toutes* }
- (32) QuantR → { *assez, beaucoup, moins, peu, plus, trop* }
- (33) Cardinal → { *un, deux, trois, ...* }
- (34) Numéral → { *aucun, certain, chaque, différents, divers, maints, nul, plusieurs, quelconque, quelque* }
- (35) Mod → { *devoir, pouvoir* }
- (36) Asp → { *aller, être en train de, être sur le point de, venir de* }
- (37) Cop → { *devenir, être, rester* }
- (38) Prép → { *à, de, dès, en, par, pour, sous, sur* }
- (39) Dénotatif → { *autre, même* }
- (40) Dém → { *ce, que* }
- (41) Déf → { *la, le, les* }
- (42) nDéf → { *des, in, une, ∅* }
- (43) Adv_{neg} → { *guère, jamais, pas, plus, point, ...* }
- (44) Adv_{imp} → { *donc* }
- (45) Adv_{degré} → { *aussi, fort, moins, plus, très, trop, ...* }

REMARQUES

- Quanta = quantitatif absolu
- QuantR = quantitatif relatif
- Cardinal = adjectif cardinal
- Numéral = adjectif indéfini distributif
- Mod = modalité
- Asp = aspectuel
- Cop = copule
- Prép = préposition racine (certaines prépositions, comme *avant, avec, après*, sont constituées de *à* + Adv ; d'autres, comme *depuis, devant*, sont faites du constituant *de* + Adv ; d'où la règle 29)
- Dém = démonstratif (dans la composante morphophonologique, on a la règle Dém+Déf → { *ce, cet, cette, ...* })
- Déf = article défini
- nDéf = article non défini
- Adv_{neg} = adverbe négatif
- Adv_{imp} = adverbe impératif
- Adv_{degré} = adverbe de degré

Pour illustrer le fonctionnement de ce modèle syntaxique, nous avons choisi une phrase courte, ne comportant qu'un verbe principal, et sans adjectif qualificatif épithète : "tous les autres enfants ont dû cueillir ces pommes". Nous ne présentons que le début de la ré-écriture de la structure profonde de la phrase, l'ensemble de ces opérations étant figuré sur le schéma n° 1.57.

<i>Règle appliquée</i>	<i>Dérivation</i>
(1) $\Sigma \rightarrow \text{Const} + \text{P}$	Const + P
(2) Const \rightarrow Affir	Affir + P
(3) Affir \rightarrow Inton _{affir}	Inton _{affir} + P
(11) P \rightarrow SN + SV	Inton _{affir} + SN + SV
(12) SN \rightarrow No + GN	Inton _{affir} + No + GN + SV
(13) No \rightarrow Plur	Inton _{affir} + Plur + GN + SV
(14) GN \rightarrow D + N	Inton _{affir} + Plur + D + N + SV
(15) D \rightarrow PréArt + Art + PostArt	Inton _{affir} + Plur + PréArt + Art + PostArt + N + SV
etc.	

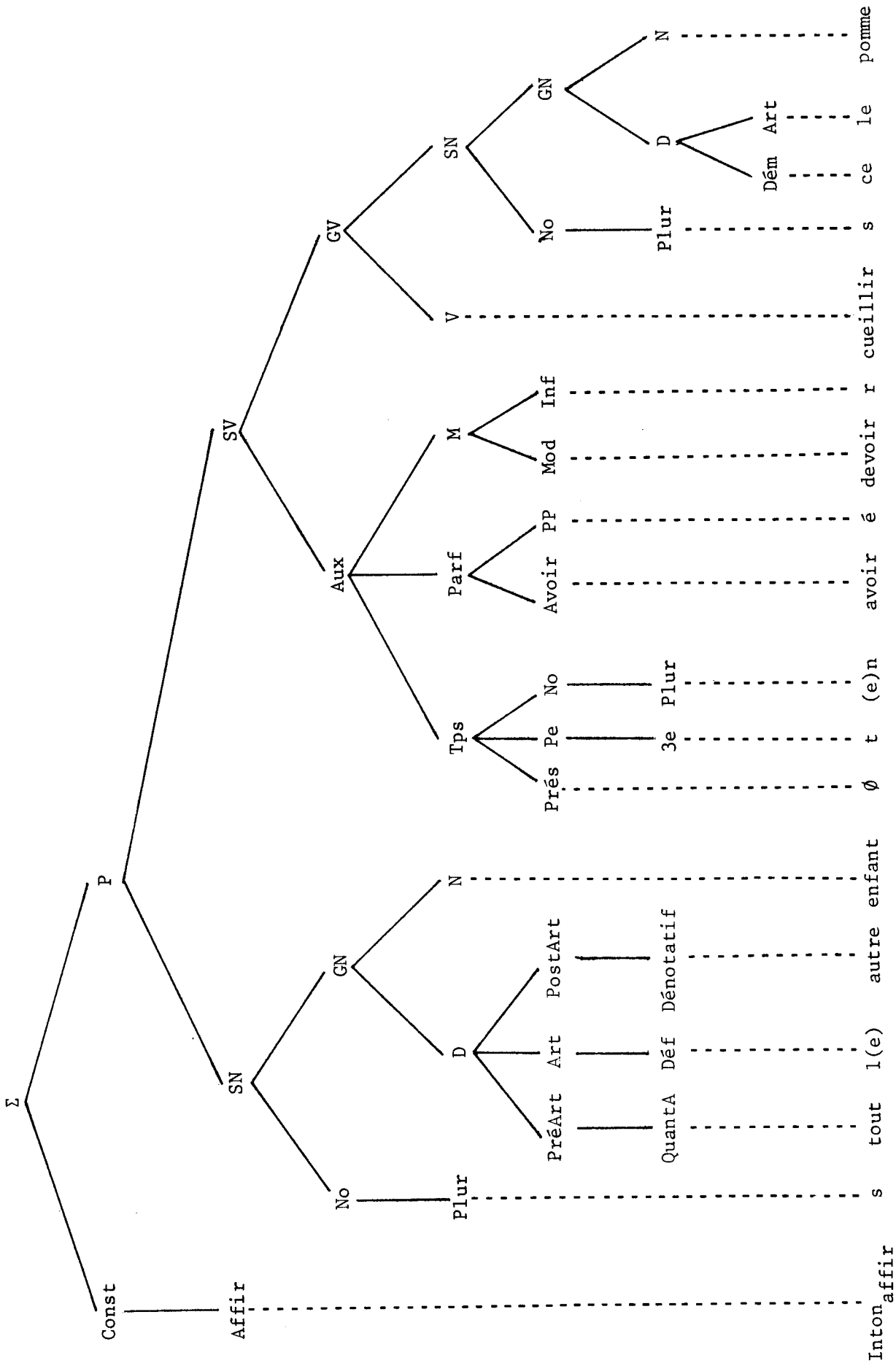
La suite préterminale ainsi obtenue par l'application des règles de ré-écriture syntaxiques est :

Inton_{affir} + Plur + QuantA + Déf + Dénotatif + N + Prés + 3e + Plur + avoir + PP + Mod + Inf + V + Plur + Dém + Art + N.

L'application de règles de ré-écriture lexicales nous donne la structure profonde :

Inton_{affir} + s + tout + l(e) + autre + enfant + \emptyset + t + (e)n + avoir + é + devoir + r + cueillir + s + ce + le + pomme.

Le passage de la structure profonde à la structure de surface fait appel à un autre type de règles de ré-écriture, les *transformations*. En effet, une grammaire purement syntagmatique deviendrait rapidement d'une telle complexité que son maniement, jusqu'à l'obtention de la phrase réalisée, soulèverait des difficultés presque insurmontables (cf. [Chomsky 1957 : 39-50 ; 1958 : 219-223] etc.). Le recours à un modèle transformationnel, greffé en quelque sorte sur le modèle syntagmatique qui engendre la structure profonde, accroît la capacité



1.57 : Structure profonde de la phrase : "tous les autres enfants ont dû cueillir ces pommes".

généralive de la grammaire [Dubois 1969 : 14-16]. Les transformations fondamentales sont le *déplacement* et l'*effacement* de constituants ; ces deux transformations permettent de rendre compte de l'ensemble des modifications de la structure profonde, jusqu'à l'obtention de la structure de surface abstraite à laquelle on appliquera les règles phonologiques ou phonétiques. En pratique, on préfère décrire ces modifications à l'aide de "macro-transformations" constituées d'une suite ordonnée de transformations élémentaires (fondamentales). Par exemple, la *substitution* (déplacement suivi d'un effacement) ; la *permutation* (série de déplacements successifs) ; etc. Les "macro-transformations" les plus importantes sont la transformation affixale (T_{af}) et la transformation d'enchâssement (T_{ench}).

La *transformation affixale* décrit l'opération de déplacement qui affecte les constituants affixes de Aux dans le syntagme verbal, et de No dans le syntagme nominal (pour ce qui est des cas les plus courants ; elle intervient également dans la transformation adjectivale). Pour l'essentiel, les affixes sont des suffixes nominaux (-age, -eur, -té, -tion) ou des désinences verbales.

Dans le syntagme nominal (cf. règles 12 et 13), la transformation affixale régit l'accord en nombre des constituants du groupe nominal. Les constituants (affixes) de No se ré-écrivent : sing $\rightarrow \emptyset$, plur $\rightarrow s$. La transformation affixale entraîne le déplacement de l'affixe après chacun des constituants de GN, et, en descendant, de D. Dans l'exemple précédent, on obtient par application de T_{af} au premier SN :

s+tout+le+autre+enfant $\xrightarrow{T_{af}}$ tout+s+le+s+autre+s+enfant+s

Dans le second SN, on applique au préalable la règle morphophonologique : ce + le \rightarrow ce (cf. règle n° 40, remarque), ce qui nous donne ensuite :

s+ce+pomme $\xrightarrow{T_{af}}$ ce+s+pomme+s.

Dans le syntagme verbal (cf. règles 19 et 20), cette transformation s'applique aux constituants de Tps, qui constituent un affixe unique, et aux constituants PP et Inf. Convenons par exemple de ré-écrire : Prés $\rightarrow \emptyset$, 3e pers $\rightarrow t$, Plur $\rightarrow (e)n$, PP $\rightarrow é$, Inf $\rightarrow r$; on applique ensuite la règle morphophonologique : $\emptyset+t+(e)n \rightarrow (e)nt$. Si l'on remplace par X la partie du syntagme verbal non affectée par la transformation affixale,

l'application de celle-ci nous donne :

(e)nt+avoir+é+devoir+r+cueillir+X $\xrightarrow{T_{af}}$ avoir+(e)nt+devoir+é+cueillir+r+X

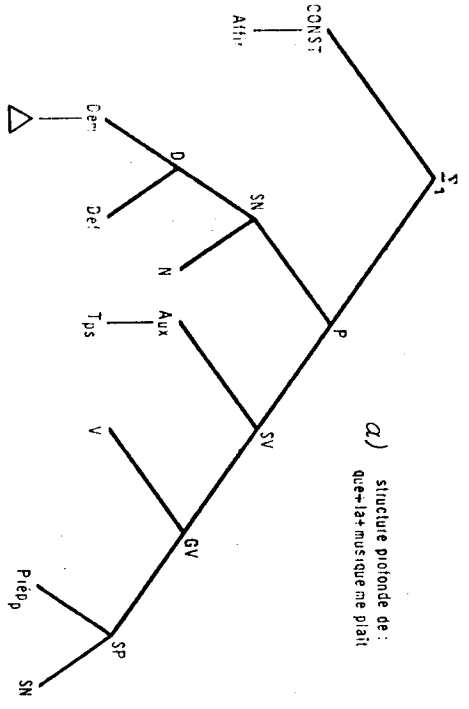
En appliquant la transformation affirmative qui reporte l'intonation en fin d'énoncé, puis ultérieurement des règles phonologiques et phonétiques telles que : avoir+(e)nt \rightarrow ont , devoir+é \rightarrow dû, cueillir+r \rightarrow cueillir, on obtiendrait la réalisation de la phrase sous sa forme définitive.

La *transformation d'enchâssement* introduit la récursivité dans la grammaire, en rendant compte des subordinées complétives, circonstancielles, relatives et interrogatives indirectes. Elle permet également d'ailleurs de rendre compte des compléments du nom (comme l'adjectif qualificatif épithète), des attributs du complément (après un verbe comme *considérer, croire, juger, penser*), de la nominalisation d'une phrase de base, de la construction des phrases avec "c'est" (distinctes des phrases simplement emphatiques). Nous nous limiterons ici à l'exemple de la subordinée relative déterminative, en décrivant une espèce particulière de T_{ench} , la transformation relative T_{rel} .

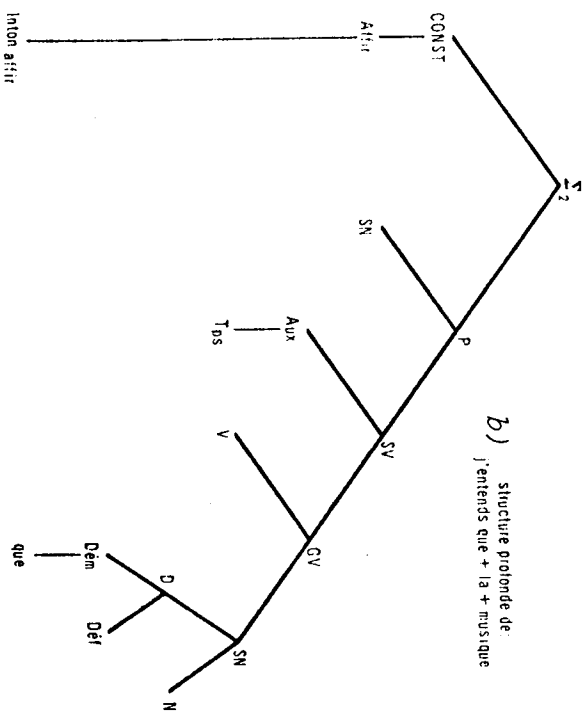
La *transformation relative* a pour effet d'enchâsser une phrase Σ_2 (subordonnée relative) dans une phrase matrice Σ_1 . Pour démontrer le mécanisme de cette transformation, nous allons reconstituer la phrase : "la musique que j'entends me plaît", à partir de l'enchâssement de deux structures profondes signifiant approximativement : "cette musique me plaît", et : "j'entends cette musique" (cf. schéma n° 1.58). Pour que T_{rel} puisse s'appliquer, il faut par conséquent que, dans les deux phrases sources, les deux syntagmes nominaux concernés comportent dans leur ré-écriture un démonstratif et un nom identique ; c'est-à-dire :

- dans Σ_1 , GN \rightarrow Dém+Déf+N, avec Dém \rightarrow Δ ;
- dans Σ_2 , GN \rightarrow Dém+Déf+N, avec Dém \rightarrow *que* ;
- le N de Σ_1 est identique au N de Σ_2 .

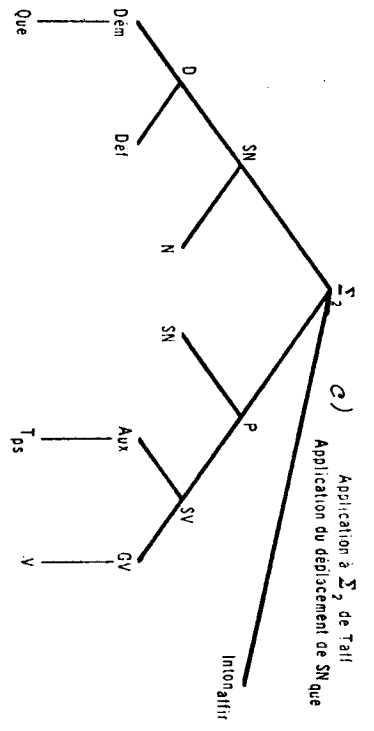
Le symbole Δ est un symbole postiche, utilisé généralement lorsque la ré-écriture syntaxique n'est plus possible [Dubois 1970 : 59,84]. Ici, Δ servira de support à l'enchâssement de Σ_2 .



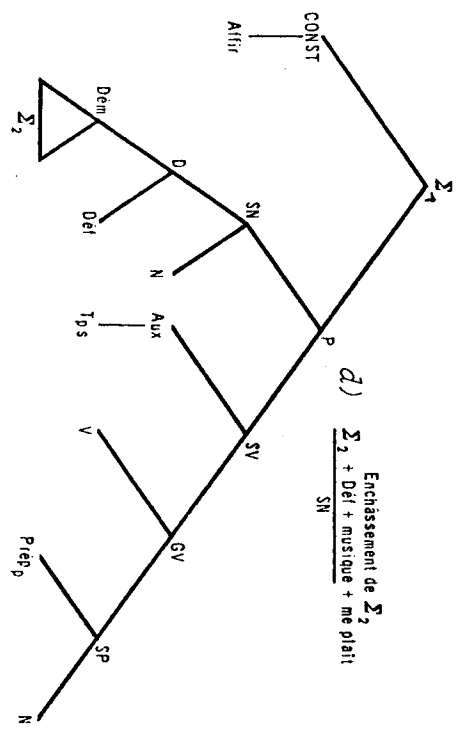
a) structure profonde de :
que+la+musique+me+plait



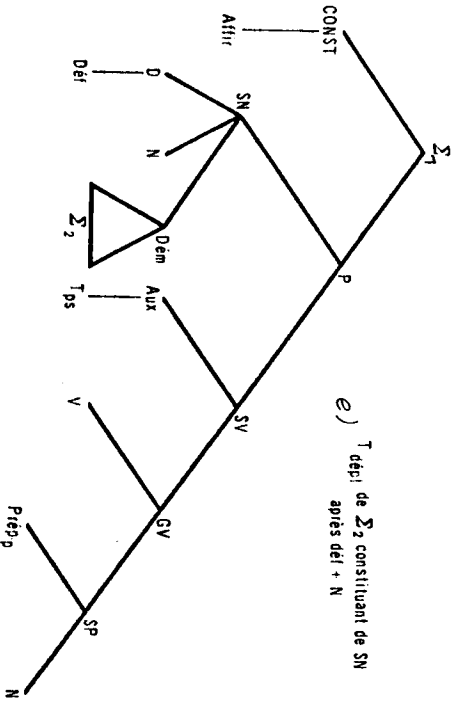
b) structure profonde de :
entendus que+la+musique



c) Application à Σ₂ de Taff
Application du déplacement de SN que



d) Enchâssement de Σ₂
Σ₂ + Déf + musique + me plait



e) T dépli de Σ₂ constituant de SN
après déf + N

Les structures profondes des phrases sources peuvent être représentées schématiquement par :

Σ_1 : Dém+la+musique+me plait (schéma n° 1.58 a)

Σ_2 : j'entends+Dém+la+musique (schéma n° 1.58 b)

La première étape de la transformation est, dans Σ_2 , le déplacement du SN_{que} en tête de phrase (schéma n° 1.58 c) :

Σ_2 : Dém+la+musique+j'entends

La seconde étape est l'enchâssement de Σ_2 dans Σ_1 , à la place du symbole postiche Δ (schéma n° 1.58 d) :

Σ_1 : [Dém+la+musique+j'entends]+la+musique+me plait.

La phrase enchâssée prenant la fonction du symbole qu'elle remplace, Σ_2 se comporte maintenant comme un déterminant du nom. La troisième étape consiste dans le déplacement de Dém(Σ_2 enchâssée) après le nom qu'il détermine (schéma n° 1.58 e) :

Σ_1 : la+musique+[que+la+musique+j'entends]+me plait.

La dernière étape est l'effacement du deuxième nom par le premier :

Σ_1 : la+musique+que+j'entends+me plaît.

Les principales autres transformations d'enchâssement, qui introduisent les subordinées complétives, circonstancielles, et interrogatives (indirectes), ont pour effet le remplacement par Σ_2 d'une proforme nominale (ProN) de Σ_1 . Une *proforme* est "un mot du dictionnaire qui comporte dans sa définition un ensemble de traits lexicaux fondamentaux, à l'exception de tout trait sémantique définissant un sens [...]". Les proformes restent des objets abstraits, comme les mots du lexique, et ce sont les règles morphophonologiques qui peuvent ensuite les réaliser, cette réalisation pouvant être diverse selon les combinaisons dans lesquelles entre une proforme" [Dubois 1970 : 65]. Par exemple, la proforme nominale (ProN) peut se réaliser en *on*, *un* (dans *quelqu'un*), *chose* (dans *quelque chose*), etc. La proforme verbale (ProV) est le verbe *faire*, homonyme de *faire* (produire) et de *faire* (factitif). On voit sur ces exemples que les proformes constituent l'équivalent dans la structure profonde de ce que sont, dans l'énoncé, les mots pleins généraux de TESNIÈRE.

Pour qu'il puisse y avoir enchâssement d'une complétive, d'une circonstancielle, ou d'une interrogative, il faut qu'il y ait dans Σ_1 un groupe nominal de la forme : GN \rightarrow Dém+Déf+ProN, dans lequel ProN est la ré-écriture de N. Si la phrase à enchâsser Σ_2 est une phrase affirmative, lorsque le ProN est un constituant du SN complément du verbe dans le SV de Σ_1 , la proposition ainsi formée est une subordonnée complétive ; lorsque le ProN est un constituant du SN élément de SP dans le SV de Σ_1 , la proposition est une subordonnée circonstancielle. Si la phrase à enchâsser est une phrase interrogative, et si le ProN est un constituant du SN complément du verbe dans le SV de Σ_1 , la subordonnée ainsi formée est une interrogative indirecte. Nous examinerons en détail la transformation d'enchâssement d'une interrogative indirecte au paragraphe suivant. Auparavant, nous donnons à titre d'exemple la structure (simplifiée) de la phrase avec subordonnée complétive : "je vous affirme qu'Albert est tombé" (schéma n° 1.59), que l'on comparera avec le stemma proposé par TESNIÈRE pour le même énoncé (schéma 1.29).

c) L'interrogation selon DUBOIS et DUBOIS-CHARLIER.

La règle de ré-écriture n° 4 (tableau n° 1.56) donne deux manières de ré-écrire le constituant Inter. La première forme : Inter \rightarrow (SN_{inter}) + Inton_{inter}, correspond à l'interrogation directe totale, à réponse *oui/non* ; la seconde : Inter \rightarrow (SN_{inter}) + SN_{que} + (Inton_{inter}), à l'interrogation directe partielle, dont la réponse est un syntagme nominal correspondant à celui qui est formé avec le pronom, l'adjectif, ou l'adverbe interrogatif. Si l'on tient compte du caractère obligatoire ou facultatif des constituants, on arrive par conséquent à deux formes d'interrogation directe totale, et quatre formes d'interrogation directe partielle ; en pratique, ces dernières nous paraissent pouvoir être ramenées à deux formes d'interrogation directe partielle, en écartant les variantes sans intonation interrogative, rares dans le langage courant, et sans utilité dans les tests ou les questionnaires. Les quatre formes d'interrogation directe sont donc :

totale :	}	Inton _{inter}	ex : "Pierre vient ?"
		SN _{inter} + Inton _{inter}	ex : "Pierre vient-il ?"
partielle :	}	SN _{que} + Inton _{inter}	ex : "Comment va Pierre ?"
		SN _{inter} + SN _{que} + Inton _{inter}	ex : "Comment Pierre va-t-il ?"

En outre, les auteurs font une place à part aux interrogatives directes avec *est-ce que*, ainsi qu'aux interrogatives indirectes.

La première formule de transformation interrogative, qui engendre les interrogatives totales, compte trois étapes. Au départ, en appelant X une suite quelconque de constituants, on a par exemple la séquence :

$$\Sigma : \text{Inter} + \text{SN} + \text{Tps} + \text{V} + \text{X}$$

(pouvant avoir comme réalisation : *Inter+Pierre+vient+∅*).

La ré-écriture du constituant interrogatif donne (règle n° 4) :

$$\Sigma : \text{SN}_{\text{inter}} + \text{Inton}_{\text{inter}} + \text{SN} + \text{Tps} + \text{V} + \text{X}$$

avec un SN_{inter} identique au SN sujet de la phrase (réalisation : *Pierre+?+Pierre+vient+∅*).

La première étape de la transformation T_{inter} est le déplacement de l'intonation interrogative en fin de phrase :

$$\Sigma : \text{SN}_{\text{inter}} + \text{SN} + \text{Tps} + \text{V} + \text{X} + \text{Inton}_{\text{inter}}$$

(réalisation : *Pierre+Pierre+vient+∅+?*).

La seconde étape consiste à déplacer SN derrière le premier verbe rencontré :

$$\Sigma : \text{SN}_{\text{inter}} + \text{Tps} + \text{V} + \text{SN} + \text{X} + \text{Inton}_{\text{inter}}$$

(réalisation : *Pierre+vient+Pierre+∅+?*).

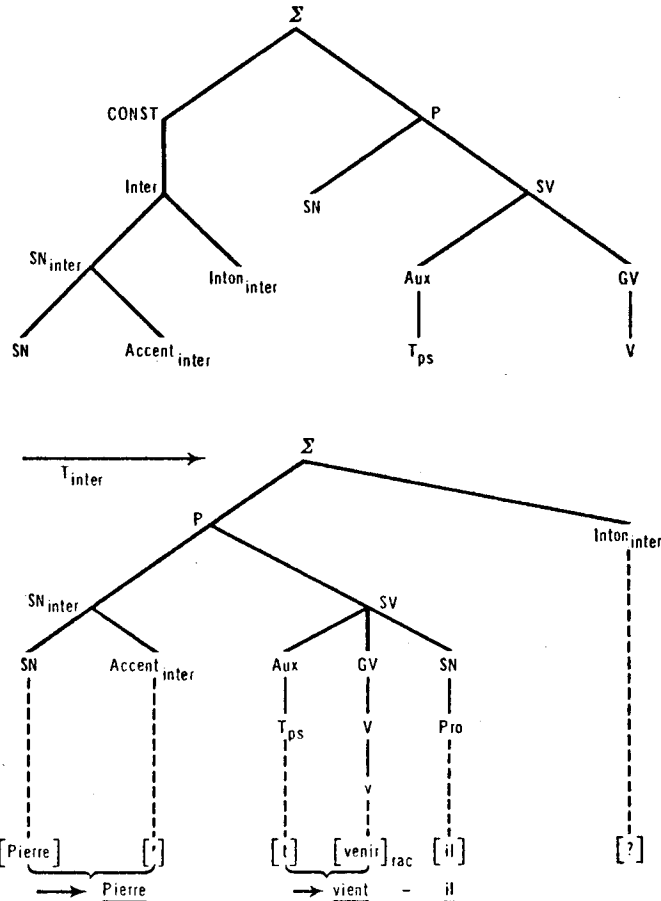
La troisième étape est le remplacement de SN par un pronom :

$$\Sigma : \text{SN}_{\text{inter}} + \text{Tps} + \text{V} + \text{Pro} + \text{X} + \text{Inton}_{\text{inter}}$$

(réalisation : *Pierre+vient+il+∅+?*).

Cette transformation (dont le résultat est représenté par le schéma n° 1.60) rend compte de la mise sous forme interrogative de phrases plus complexes, telles que "le directeur de l'usine est-il arrivé à dix heures ?", comme il est aisé de le vérifier. En particulier, même lorsque le Aux de GV se ré-écrit : $\text{Aux} \rightarrow \text{Tps} + \text{Parf} + \text{M} + \text{Parf}$ ("Pierre a-t-il dû venir ?"), et même lorsque le GV contient la copule Cop ("Pierre a-t-il pu être là ?"), la seconde étape de la transformation rend compte de la place du pronom rappelant le SN sujet, puisque le SN à pronominaliser se place après le premier verbe rencontré (qui peut être un constituant

de M, de Parf, ou la copule) et non nécessairement le V de GV. Si la ré-écriture du constituant est : $\text{Inter} \rightarrow \text{Inton}_{\text{inter}}$, il est clair que seule la première étape de la transformation peut s'appliquer.



1.60 : Transformation interrogative appliquée à l'interrogation totale ("Pierre vient-il ?) [Dubois 1970 : 212].

La seconde formule de transformation interrogative engendre les interrogatives partielles. Elle s'applique lorsque l'on a la ré-écriture : $\text{Inter} \rightarrow (\text{SN}_{\text{inter}}) + \text{SN}_{\text{que}} + \text{Inton}_{\text{inter}}$. En appelant X et Y des suites quelconques de constituants non modifiés par T_{inter} , nous avons par exemple la séquence :

$$\Sigma : \text{Inter} + \text{SN}_1 + T_{\text{ps}} + V + X + \text{SN}_2 + Y$$

dans laquelle SN_1 est sujet du verbe, et SN_2 complément (réalisation possible : $\text{Inter} + \text{Pierre} + \text{verra} + \emptyset + \text{quelqu'un} + \emptyset$).

La ré-écriture du constituant interrogatif donne (règle n° 4) :

$$\Sigma : \text{SN}_{\text{inter}} + \text{SN}_{\text{que}} + \text{Inton}_{\text{inter}} + \text{SN}_1 + \text{Tps} + \text{V} + \text{X} + \text{SN}_2 + \text{Y}$$

La première étape est la même que précédemment (déplacement de $\text{Inton}_{\text{inter}}$) :

$$\Sigma : \text{SN}_{\text{inter}} + \text{SN}_{\text{que}} + \text{SN}_1 + \text{Tps} + \text{V} + \text{X} + \text{SN}_2 + \text{Y} + \text{Inton}_{\text{inter}}$$

La seconde étape également (déplacement du SN sujet après le premier verbe rencontré) :

$$\Sigma : \text{SN}_{\text{inter}} + \text{SN}_{\text{que}} + \text{Tps} + \text{V} + \text{SN}_1 + \text{X} + \text{SN}_2 + \text{Y} + \text{Inton}_{\text{inter}}$$

La troisième étape est propre aux interrogatives partielles. Elle consiste à substituer le SN_{que} au SN sur lequel porte l'interrogation. Ce peut être le SN sujet du verbe (ex : "qui verra quelqu'un ?") ou complément (ex : "qui Pierre verra-t-il ?"), ou bien le SN constituant d'un SP (ex : "de quelle manière Pierre va-t-il ?" → "comment Pierre va-t-il ?"). Nous prenons ci-après l'exemple d'une interrogation partielle portant sur le SN complément (SN_2). Nous avons par conséquent :

$$\Sigma : \text{SN}_{\text{inter}} + \text{Tps} + \text{V} + \text{SN}_1 + \text{X} + \text{SN}_{\text{que}} + \text{Y} + \text{Inton}_{\text{inter}}$$

La quatrième étape correspond à la troisième étape de la formule précédente de T_{inter} (pronominalisation du SN sujet) :

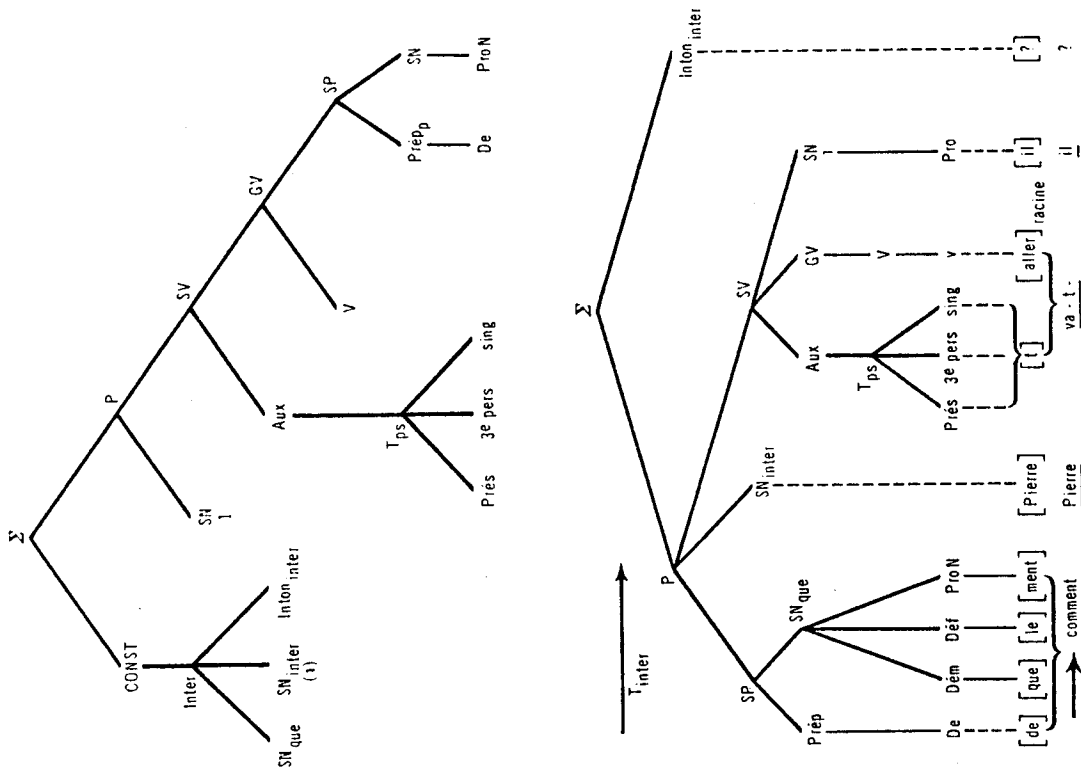
$$\Sigma : \text{SN}_{\text{inter}} + \text{Tps} + \text{V} + \text{Pro} + \text{X} + \text{SN}_{\text{que}} + \text{Y} + \text{Inton}_{\text{inter}}$$

La cinquième étape est le déplacement du SN_{que} en tête de la phrase :

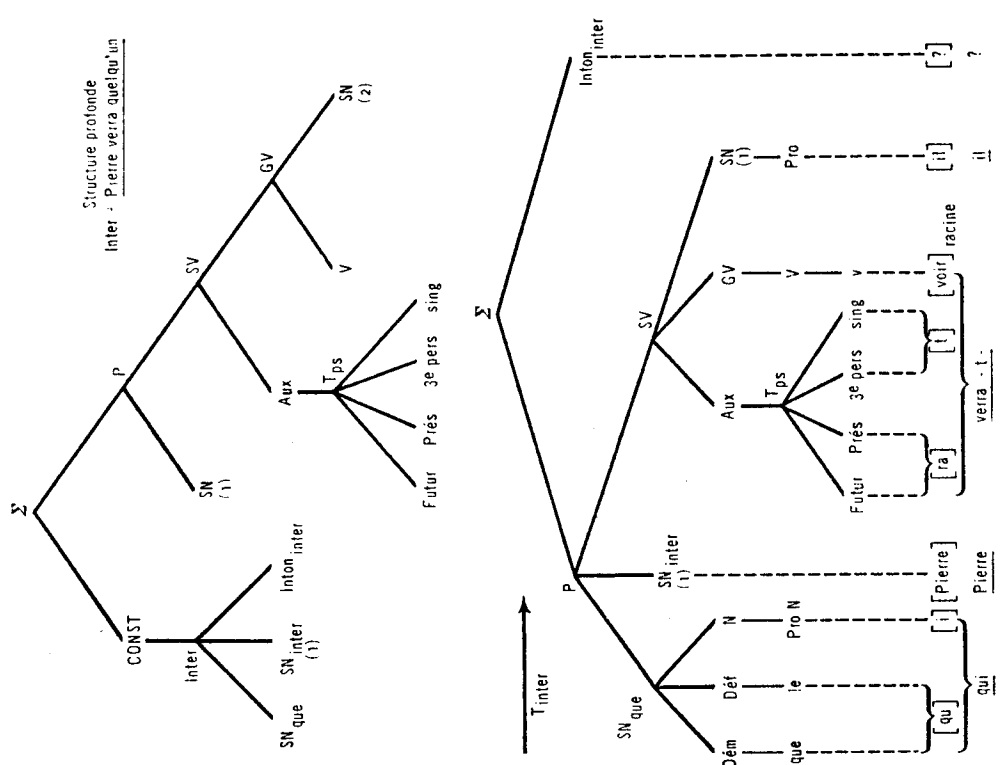
$$\Sigma : \text{SN}_{\text{que}} + \text{SN}_{\text{inter}} + \text{Tps} + \text{V} + \text{Pro} + \text{X} + \text{Y} + \text{Inton}_{\text{inter}}$$

Cette transformation rend compte du passage d'une structure profonde que l'on pourrait représenter par : *Inter+Pierre verra quelqu'un*, à une phrase réalisée en : "qui Pierre verra-t-il ?" (cf. schéma n° 1.61). On montrerait de même qu'elle rend compte du passage de : *Inter+Pierre va de quelque manière*, à : "comment Pierre va-t-il ?" (cf. schéma n° 1.62).

Les diverses formes de la transformation interrogative peuvent être représentées par un organigramme unique (cf. schéma n° 1.63). Les cases de l'organigramme correspondent aux opérations élémentaires (déplacement, substitution,



1.62 : Transformation interrogative appliquée à l'interrogation partielle portant sur le complément circonstanciel ("comment Pierre va-t-il ?") [Dubois 1970 : 219].

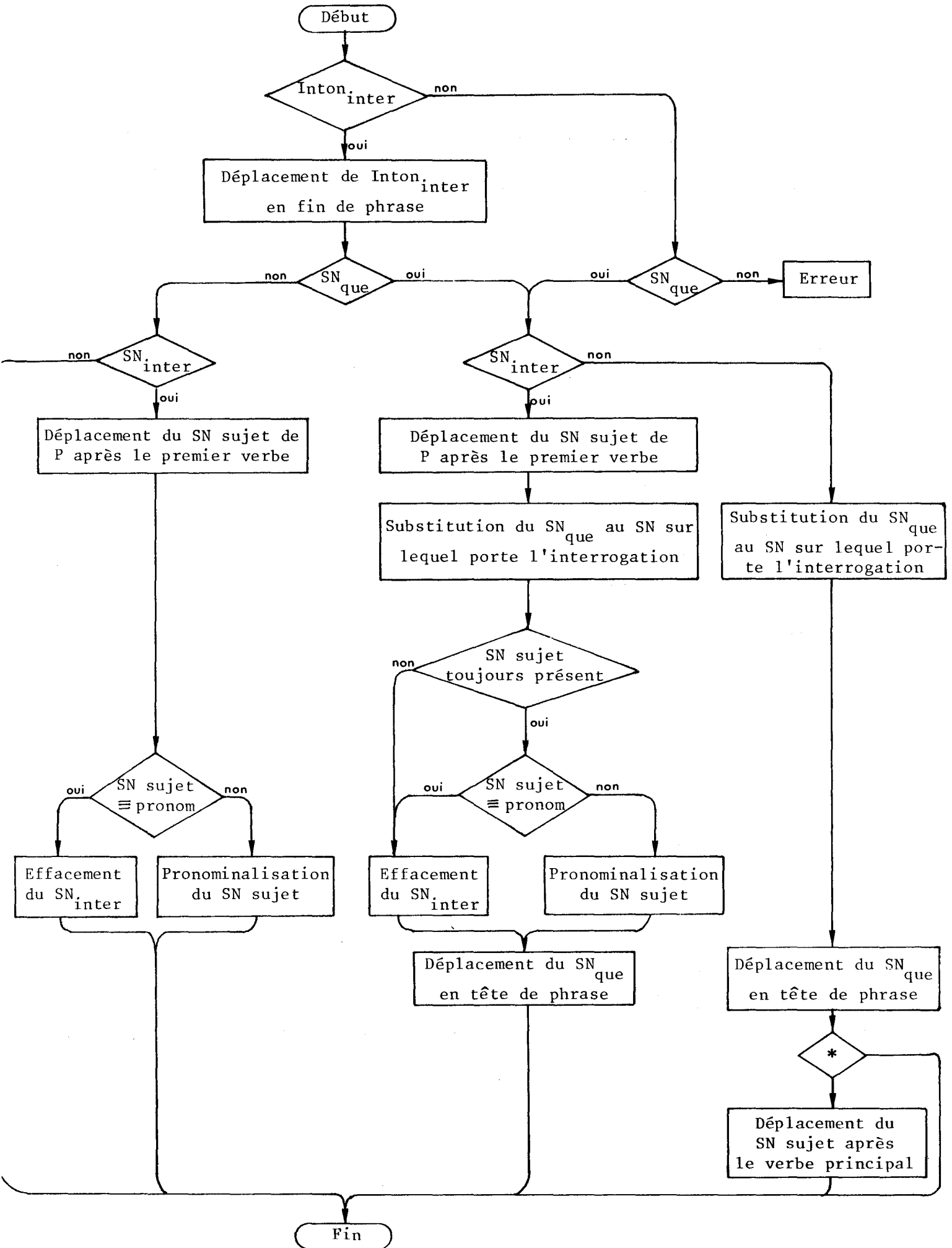


1.61 : Transformation interrogative appliquée à l'interrogation partielle portant sur le complément d'objet ("qui Pierre verra-t-il ?") [Dubois 1970 : 215].

Structure profonde
Inter : Pierre verra qui qu'un

pronominalisation). Les premiers aiguillages dépendent de la manière dont le constituant Inter a été réécrit : présence ou absence de $\text{Inton}_{\text{inter}}$, SN_{que} , SN_{inter} . Les aiguillages suivants correspondent à divers cas particuliers. Ainsi, lorsque l'on a $\text{Inter} \rightarrow \text{SN}_{\text{inter}} + \text{SN}_{\text{que}} + (\text{Inton}_{\text{inter}})$, la pronominalisation du SN sujet ne peut pas s'appliquer lorsque l'interrogation porte sur le SN sujet lui-même ("qui viendra ?") puisque ce SN n'est plus présent dans la phrase (SN_{que} lui a été substitué), ni lorsque le SN sujet est déjà un pronom ("qui verra-t-il ?") ; par contre, dans ces conditions, il y a effacement du SN_{inter} (sans cela, on aurait : * "qui quelqu'un viendra ?" et : * "qui il verra-t-il ?"). Ceci s'applique également à $\text{Inter} \rightarrow \text{SN}_{\text{inter}} + \text{Inton}_{\text{inter}}$ (sinon, on aurait : * "il est-il heureux ?"). D'autre part, dans le cas où $\text{Inter} \rightarrow \text{SN}_{\text{que}} + (\text{Inton}_{\text{inter}})$, il peut y avoir déplacement du SN sujet après le verbe principal (et non après le premier verbe rencontré). Ce déplacement est obligatoire lorsque l'interrogation porte sur un SN complément comportant une proforme nominale ayant le trait non-animé ("qu'a pensé ton père ?") ; il est interdit si la réalisation de la phrase aboutit à deux SN sans préposition concaténés (* "à qui ont donné tes parents ce livre ?"). Dans l'organigramme, nous avons remplacé par un astérisque l'ensemble des conditions assez complexes qui régissent ce déplacement (cf. : [Dubois 1970 : 221-223]).

Contrairement à TESNIÈRE, les auteurs ne considèrent pas *est-ce que* comme une particule interrogative, mais comme la réalisation d'une phrase matrice interrogative $\text{Inter} + \text{SN}_1 + \text{Aux} + \text{Cop} + \text{SN}_2 \rightarrow \text{Inter} + \text{ce}(\text{la}) \text{est que}$ (avec la règle $\text{Aux} \rightarrow \text{Tps}$, les constituants Parf et M étant exclus), dans laquelle est enchâssée une phrase affirmative. Par exemple, la question "est-ce que Paul viendra ?" est issue de l'enchâssement d'une phrase Σ_2 : $\text{Affir} + \text{Paul viendra}$, dans la phrase matrice Σ_1 : $\text{Inter} + \text{ce}(\text{la}) \text{est} + \text{Dém} + \text{Déf} + \text{ProN}$. De même, l'interrogation portant sur le SP : "pourquoi est-ce que Paul viendra ?" provient de Σ_2 : $\text{Affir} + \text{Paul viendra}$, enchâssée dans Σ_1 : $\text{Inter} + \text{ce}(\text{la}) \text{est} + \text{Dém} + \text{Déf} + \text{ProN} + \text{pour une raison}$. Par contre, la question "est-ce Paul qui viendra ?" contient une affirmative Σ_2 : $\text{Affir} + \text{quelqu'un viendra}$ enchâssée par T_{rel} dans une interrogative Σ_1 : $\text{Inter} + \text{ce quelqu'un est Paul}$; si l'on réécrit $\text{Inter} \rightarrow \text{Inton}_{\text{inter}}$, on a la variante : "c'est Paul qui viendra ?".



1.63 : Organigramme de la transformation interrogative T_{inter}

Les interrogatives indirectes sont des phrases interrogatives enchâssées dans le SN complément d'une phrase matrice. Par exemple, l'enchâssement par T_{compl} de Σ_2 : Inter+*Paul viendra* dans Σ_1 : Affir+*je demande ce(la)* donne la phrase : "je demande si Paul viendra", la combinaison Dém+Déf (provenant de la ré-écriture du SN complément : SN \rightarrow Dém+Déf+ProN \rightarrow *ce(la)*) se ré-écrivant *si*. Lorsque l'interrogative est une interrogative partielle (avec SN_{que}), la combinaison Dém+Déf (provenant de SN_{que} \rightarrow Dém+Déf+ProN) est effacée. A titre d'illustration, nous avons repris l'énoncé "je vous demande qui est arrivé à dix heures", déjà analysé dans l'optique de la syntaxe structurale de TESNIÈRE (cf.: schéma n° 1.35). La phrase enchâssée a pour structure profonde (cf.: schéma n° 1.64) :

$$\Sigma_2 : \text{Inter} + \text{SN} + \text{Tps} + \text{Parf} + \text{V} + \text{SP}$$

L'interrogation portant sur le sujet, le SN_{inter} aura la même forme que le SN sujet, et il se confondra avec le SN_{que}. On peut par conséquent choisir pour le constituant de phrase : Inter \rightarrow SN_{que} (pour simplifier, on peut omettre l'intonation, puisqu'elle sera effacée au style indirect). On a donc la ré-écriture :

$$\Sigma_2 : \text{SN}_{\text{que}} + \text{SN} + \text{Tps} + \text{Parf} + \text{V} + \text{SP}$$

Le SN sujet a une proforme nominale comme constituant (le noyau de la phrase est : P \rightarrow *quelqu'un est arrivé à dix heures*) ; le SN_{que} a également comme constituant une proforme nominale : SN_{que} \rightarrow Dém+Déf+ProN. Après ré-écriture de SN_{que}, et substitution de SN_{que} à SN, on obtient :

$$\Sigma_2 : \text{Dém} + \text{Déf} + \text{ProN} + \text{Tps} + \text{Parf} + \text{V} + \text{SP}$$

Le SN_{que} se ré-écrit : *que*+Déf+ProN \rightarrow *qui*. On aboutit ainsi à :

$$\Sigma_2 : \text{qui} + \text{Tps} + \text{Parf} + \text{V} + \text{SP}$$

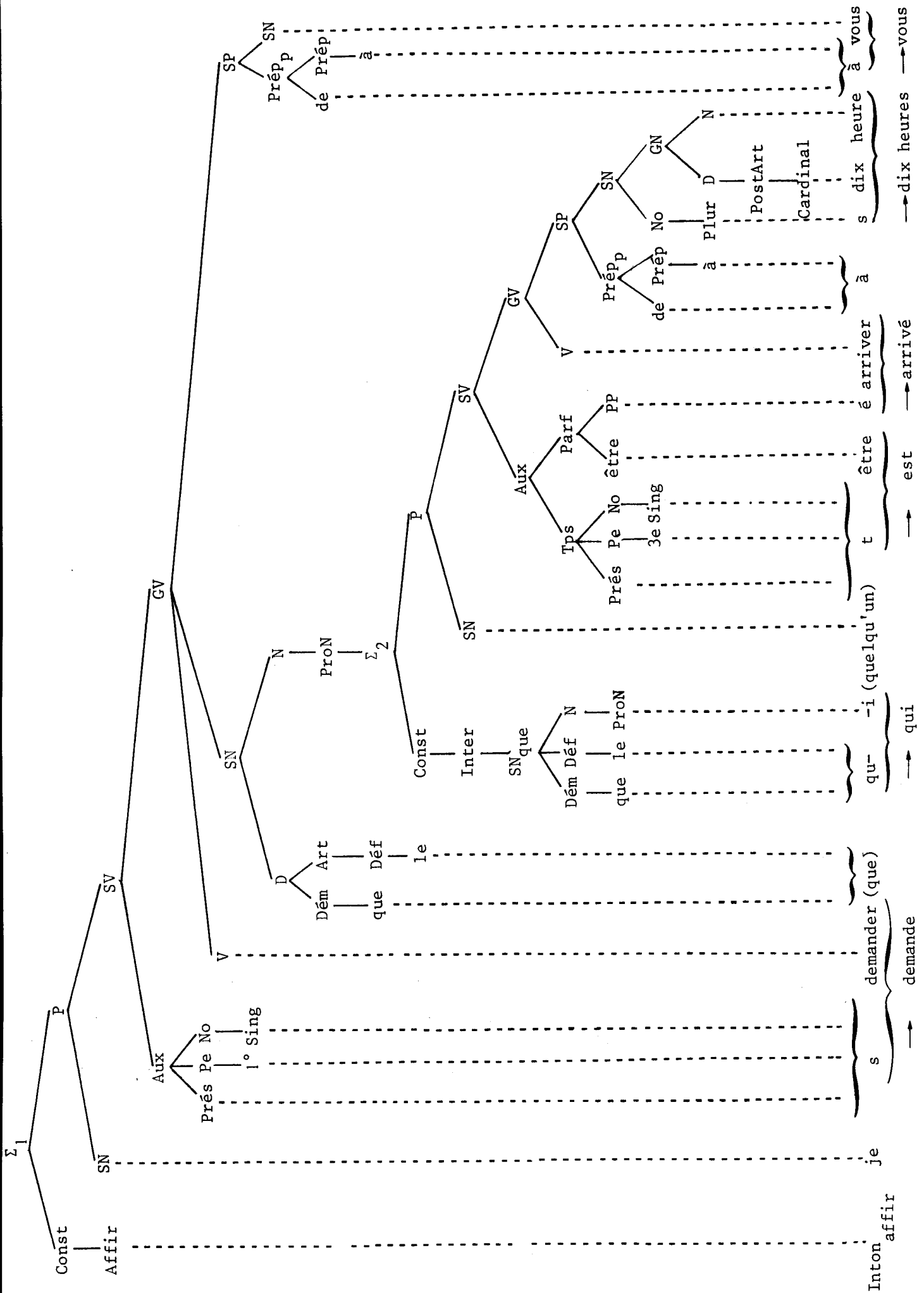
La phrase matrice Σ_1 a pour structure profonde (cf. schéma n° 1.64) :

$$\Sigma_1 : \text{Affir} + \text{SN}_1 + \text{Aux} + \text{V} + \text{SN}_2 + \text{SP}$$

Le SN complément a pour constituant une proforme nominale, et se ré-écrit : SN₂ \rightarrow Dém+Déf+ProN. La phrase Σ_2 est enchâssée à la place de ProN :

$$\Sigma_1 : \text{Affir} + \text{SN}_1 + \text{Aux} + \text{V} + \text{Dém} + \text{Déf} + \Sigma_2 + \text{SP}$$

$$\Sigma_1 : \text{Affir} + \text{SN}_1 + \text{Aux} + \text{V} + \text{Dém} + \text{Déf} + [\text{qui} + \text{Tps} + \text{Parf} + \text{V} + \text{SP}_2] + \text{SP}_1$$



1.64: Structure profonde de l'énoncé : "je vous demande qui est arrivé à dix heures" (comparer avec le schéma n° 1.35).

La phrase interrogative avec SN_{que} ($\rightarrow \text{qui}$) efface la suite Dém+Déf qui précède la proforme supportant l'enchâssement ($\text{ProN} \rightarrow \Sigma_2$). Nous obtenons finalement :

$$\Sigma_1 : \text{Affir} + SN_1 + \text{Aux} + V + \text{qui} + \text{Tps} + \text{Parf} + V + SP_2 + SP_1$$

dont on peut donner la réalisation approximative : "je demande [qui est arrivé à dix heures] à vous".

Il est clair que, pour aboutir à la structure de surface "je vous demande qui est arrivé à dix heures", quelques étapes restent à franchir. Nous n'avons évidemment présenté ici que les principales règles de ré-écriture sans lesquelles il est impossible de comprendre l'application du modèle génératif de DUBOIS et DUBOIS-CHARLIER à l'étude de l'interrogation. Nous avons délibérément omis un point tout à fait primordial, qui est celui de l'ordre dans lequel les transformations doivent être effectuées ; or, dans certains cas, une permutation dans l'ordre des transformations peut conduire à des interprétations sémantiques très différentes (par exemple dans le cas de phrases ayant : $\text{Const} \rightarrow \text{Inter} + \text{Nég} + \text{Passif}$). En outre, nous avons négligé d'autres problèmes importants, tels que la pronominalisation du syntagme nominal, ou la forme d'interrogation partielle avec un adjectif interrogatif (ex. : "quel film Pierre verra-t-il ?"). Présenter l'essentiel de cette grammaire générative aurait nécessité (au minimum) une centaine de pages. Ce bref aperçu devrait néanmoins permettre d'évaluer l'apport de la linguistique générative à l'analyse des énoncés à valeur de question.

d) Apport des modèles génératifs à l'analyse de l'interrogation.

Nous l'avons dit, la linguistique générative est en rapide développement. De nombreux points y sont encore du domaine de la recherche, et les modèles génératifs proposés dans la littérature sont soit très imparfaits, soit limités dans leur extension ; dans tous les cas, ils sont donc sujets à révision. Ainsi, CHOMSKY lui-même a proposé plusieurs versions de l'analyse des phrases interrogatives en anglais (comparer par exemple [Chomsky 1957 : 69-78 ; 1962 : 60-70 ; etc.]).

Pour la langue française, parmi les problèmes non (ou mal) résolus, figure le cas de l'interrogation portant sur le complément d'objet du verbe, lorsque ce complément est un être humain. Soit par exemple [Ruwet 1972 : 276-277] la phrase énonciative : "Pierre a mangé une pomme". On peut en dériver les interrogatives : "qui a mangé une pomme ?", et : "qu'a mangé Pierre ?". En appliquant les mêmes règles de dérivation à la phrase : "Pierre a rencontré Paul", on obtient : "qui a rencontré Paul ?", et * "qui a rencontré Pierre ?" ; cette dernière forme, parallèle à : "qu'a mangé Pierre ?", est inacceptable, en raison de l'impossibilité dans laquelle se trouve l'interlocuteur de distinguer le sujet de l'objet (TESNIÈRE fait justement remarquer l'intérêt des formes renforcées *qui est-ce qui*, *qui est-ce que*, etc. pour lever cette ambiguïté [Tesnière 1959 : 199]). On vérifierait aisément que cette interrogation agrammaticale est régulièrement engendrée par la transformation T_{inter} . Il en serait de même pour : * "qui critiquera cet imbécile ?", ou : * "quel auteur cite ce conférencier ?" ; voire pour la forme plus claire sémantiquement (compte tenu de ce que l'interlocuteur sait) : * "quels soldats commandent ces officiers ?" [Ruwet 1972 : 278-279].

De prime abord, le modèle génératif que nous avons décrit paraît capable d'engendrer l'ensemble des formes observables d'énoncés interrogatifs. Par contre, nous venons de constater qu'il risque d'engendrer également des phrases agrammaticales. Du point de vue de la linguistique générative, c'est une imperfection notable. Dans notre optique, il ne s'agit que d'un problème secondaire : à condition que le modèle soit susceptible d'engendrer *la totalité* des formes observables, il nous importe peu qu'il en produise également d'autres qui ne sont pas dans la langue. En effet, notre problème se limite à rechercher un substrat commun à la plus grande variété possible d'énoncés à valeur de question observables. De ce point de vue, le modèle génératif de DUBOIS et DUBOIS-CHARLIER est assez satisfaisant, puisqu'il permet de dériver d'une même structure profonde : Inter+(Nég)+(Emph)+(Passif)+P plus de quarante structures superficielles d'interrogatives directes distinctes (selon la ré-écriture du constituant de phrase Const et du constituant Inter ; et selon que l'interrogation porte sur P tout entier, ou sur le SN sujet de P, le SN complément dans GV, le SN élément de SP dans P ou dans GV) ; en outre, l'application des règles morphophonologiques augmente cette variété. En outre également, la production

d'interrogatives indirectes par enchâssement dans d'autres phrases rend le nombre d'énoncés réalisables à partir d'un noyau P pratiquement illimité.

Il convient toutefois de faire ici deux réserves. La première porte sur le sens des phrases ainsi engendrées. Il est évident que le sens de l'interrogation ne sera pas le même selon que le constituant de phrase se ré-écrira : $\text{Const} \rightarrow \text{Inter}$, ou : $\text{Const} \rightarrow \text{Inter} + \text{Nég} + \text{Passif}$; ou que l'on aura : $\text{Inter} \rightarrow \text{SN}_{\text{inter}} + \text{Inton}_{\text{inter}}$, ou : $\text{Inter} \rightarrow \text{SN}_{\text{inter}} + \text{SN}_{\text{que}}$; et, dans ce dernier cas, selon le SN sur lequel portera l'interrogation. Certes, à ce premier stade de notre analyse, le sens des énoncés importe moins que leurs caractéristiques morpho-syntaxiques ; et, de ce point de vue, les modèles génératifs sont particulièrement puissants, puisqu'ils rendent compte, à partir d'une même structure profonde, d'énoncés sémantiquement distincts. Toutefois, pour l'analyse logico-sémantique ultérieure, cette puissance risque de se révéler plus gênante qu'utile.

La seconde réserve concerne l'interprétation de la structure profonde des énoncés. Dans un souci de pédagogie (partagé par les auteurs du modèle présenté), nous avons proposé une traduction de la structure profonde correspondant à : "la musique que j'entends me plaît", à : "Pierre vient-il ?", et à "qui Pierre verra-t-il ?". Ce procédé doit être regardé comme une entorse grave aux principes de la linguistique générative, analogue au remplacement des symboles algébriques par des valeurs numériques dans l'exposé d'une démonstration. Les éléments de base qui constituent la structure profonde des énoncés sont des éléments abstraits appartenant au vocabulaire auxiliaire V_A ; la réalisation de la phrase n'est que le stade ultime du processus générateur, par application des règles de ré-écriture terminales, des transformations, et des règles morphologiques. L'application prématurée des règles terminales est par conséquent un contre-sens. Si l'on prend soin de bien distinguer les phrases-noyaux de leur structure profonde (leur *indicateur syntagmatique*), comme le recommande CHOMSKY [Chomsky 1965 : 34 - 35], il n'est plus possible de décrire le contenu sémantique commun à un ensemble d'énoncés ayant la même structure profonde, et des structures de surface (et par conséquent des réalisations) différentes. Or, ce problème ne manquera pas de se poser dans l'analyse des énoncés interrogatifs (cf. chapitre 2). Une solution commode pour appréhender le contenu sémantique du noyau P consistera à s'intéresser à la réalisation de la phrase $\Sigma \rightarrow \text{Affir} + \text{P}$;

la transformation affirmative peut en effet être considérée comme la transformation qui altère le moins la structure profonde, puisqu'elle se borne au déplacement de l'intonation affirmative en fin de phrase :

$$\text{Affir+P} \rightarrow \text{Inton}_{\text{affir}} + \text{P} \xrightarrow{\text{T}_{\text{affir}}} \text{P+Inton}_{\text{affir}}$$

Dans ces conditions, les modèles génératifs nous donnent le moyen de ramener une grande diversité d'énoncés interrogatifs à une structure profonde de la forme Inter+P, dont on étudiera le contenu en réalisant la phrase Affir+P.

Sur le plan morpho-syntaxique, la grammaire générative nous paraît présenter certains avantages par rapport à la syntaxe structurale de TESNIÈRE. Tout d'abord, elle fait moins appel (implicitement) à la signification des énoncés : la notion de nucléus, indispensable pour analyser les formes d'interrogation, est en effet une notion purement sémantique, alors que le modèle génératif que nous avons décrit ne fait intervenir les *traits* sémantiques (des noms, des adjectifs, des verbes) que dans quelques cas particuliers ; encore pourrait-on peut-être ramener ces traits à des propriétés purement distributionnelles. Ensuite, ce modèle propose une vision plus synthétique de l'interrogation particulière (interrogation nucléaire, questions en *qu-*), puisque l'objet sur lequel porte l'interrogation est toujours un syntagme nominal. Enfin, sur la grammaire générative de l'anglais de CHOMSKY se fondent les travaux logico-sémantiques de KATZ, qui propose en particulier une théorie complète de l'interrogation (que nous étudierons au chapitre 2).

Par contre, la grammaire générative présente l'inconvénient (apparent) d'être beaucoup plus abstraite que la syntaxe de TESNIÈRE, et par conséquent de faire appel à des mécanismes plus complexes pour rendre compte d'un énoncé identique (comparer les schémas 1.29 et 1.59 ; 1.35 et 1.64). En outre, c'est également un modèle hypothético-déductif, qui se prête mal à l'analyse d'un corpus d'énoncés. Comme pour la syntaxe de TESNIÈRE, le problème reste posé du passage de l'énoncé observé à sa structure sous-jacente (démarche inductive),

1.2.3. Les grammaires transformationnelles

Il peut paraître surprenant de n'aborder les théories de Zellig Sabbetai HARRIS qu'après avoir parlé des thèses de Noam CHOMSKY, qui fut son élève, puis son disciple dissident. De même, il peut sembler arbitraire d'opposer grammaires génératives et grammaires transformationnelles, alors que la grande majorité des syntaxes génératives sont également transformationnelles. La raison de ce choix est que l'oeuvre de HARRIS, peut-être plus difficile d'accès que celle de CHOMSKY, nous a toutefois paru plus homogène, plus assurée, et d'un intérêt pratique plus grand que les grammaires génératives, tout au moins dans la perspective d'une formalisation des énoncés à valeur de question.

CHOMSKY est probablement, pour le chercheur en linguistique ou pour le professeur de langues, un auteur plus stimulant que HARRIS : il propose des modèles syntaxiques relativement simples (donc faciles à enseigner) qui sont autant d'hypothèses à vérifier (donc de directions de recherches). Mais pour celui qui n'est désireux que d'emprunter à la linguistique des méthodes ou des résultats relativement solides, les modèles hypothético-déductifs de CHOMSKY présentent nous semble-t-il un intérêt moindre que les procédures inductives de HARRIS. De plus, comme nous le verrons au chapitre 2, la formalisation de l'interrogation proposée par HARRIS est en complet accord avec les postulats de la logique érotétique. Aussi, cette formalisation nous apparaît-elle comme le maillon indispensable pour passer de l'analyse morpho-syntaxique de l'interrogation à son étude logico-sémantique.

Compte-tenu de l'unité des travaux de HARRIS, nous avons jugé bon d'en donner un très bref aperçu d'ensemble, avant d'exposer ses vues sur la grammaire transformationnelle et l'analyse syntaxique de l'interrogation.

a) La démarche de Z.S. HARRIS

"Lorsqu'on considère la carrière de ce linguiste dans son ensemble, ce qui frappe d'abord c'est le sérieux dans la réflexion, la persévérance dans une direction bien tracée, la cohérence" [Mounin 1972 : 170-171]. L'oeuvre de HARRIS présente en effet une remarquable continuité : partant de l'analyse phonologique des énoncés, il aboutit à l'étude de la syntaxe et à l'analyse du discours, en

appliquant avec rigueur et constance une méthode simple. Cette méthode se ramène à l'application réitérée de deux procédures successives : la définition d'unités élémentaires, et l'analyse des relations de ces unités les unes par rapport aux autres [Harris 1951 : 6].

Ainsi, au niveau phonologique, la première étape consiste à distinguer des éléments sonores discrets correspondant chacun à un son, et à identifier les répétitions de ces éléments dans des énoncés différents. Outre qu'il est inutile de pousser la segmentation au point de séparer des sons qui se présentent toujours associés dans la langue que l'on étudie [Harris 1951 : 42-44], il faut pouvoir considérer comme équivalents des sons susceptibles de varier légèrement d'une répétition à l'autre (ou selon le contexte ou le locuteur) : HARRIS utilise pour cela le *test d'appariement*, qui permet de regrouper sous un même symbole phonétique un ensemble de sons voisins, et de définir avec précision ce qu'il convient de distinguer à l'intérieur d'une séquence phonétique. Cette procédure [Harris 1951 : 29-41 ; 1968 : 24-27] consiste à faire prononcer plusieurs fois par un locuteur, dans un ordre quelconque, deux courts segments d'énoncé de sa langue maternelle (deux mots par exemple) dont on se demande s'il faut ou non les distinguer, et d'inviter un auditeur de même langue maternelle à identifier les répétitions. Si les diagnostics de l'auditeur sont corrects, il y a bien lieu de distinguer les deux segments (et par conséquent les traits phonologiques par lesquels ils diffèrent). Après avoir identifié tous les éléments phonologiques de la langue (phonèmes), il reste à décrire la structure phonologique de celle-ci, en déterminant, parmi toutes les suites possibles de phonèmes, quelles sont celles que l'on rencontre effectivement [Harris 1951 : 150-155].

Cette même démarche en deux étapes se retrouve au niveau de la morphologie. Il faut tout d'abord déterminer les séquences de phonèmes pouvant constituer les unités morphologiques (racines, préfixes, suffixes), en identifiant les séquences de phonèmes interchangeables dans des énoncés identiques par ailleurs. La règle étant que, "si, dans un environnement global - X, on observe les combinaisons AB, AD, et CD (où A,B,C, et D sont des segments de parole phonologiquement identifiables), alors, que l'on observe ou non CB, on peut considérer A,B,C et D comme des segments morphémiques discrets dans l'environnement - X"

[Harris 1951 : 158-159]. Il est ensuite possible de définir des classes de morphèmes (noms, types de verbes, etc.) en fonction de leurs relations distributionnelles dans les énoncés, c'est-à-dire selon les ressemblances relevées dans leurs environnements. Remarquons que HARRIS a développé également une procédure de détermination des morphèmes à partir du nombre de phonèmes différents que l'on peut observer après une séquence donnée de phonèmes (cf. p.ex. [Harris 1968 : 27-32], ainsi qu'une application au français dans [Harris 1955 : 55-57]).

Au niveau syntaxique, on peut définir un ensemble relativement réduit de structures de phrases (description en constituants immédiats) qui sont les seules combinaisons élémentaires de morphèmes observables dans une langue donnée. On peut également définir des règles de substitution ou d'adjonction susceptibles d'enrichir la structure de base (cf. p. ex. [Harris 1946]). L'étape suivante consiste à déterminer le degré d'acceptabilité (ou de vraisemblance) de phrases formées selon un schéma donné ; par exemple, "l'homme a donné un livre au garçon" est plus acceptable que : "l'homme a donné un livre à la table". Avec le même vocabulaire, mais un schéma de phrase différent, on obtiendrait : "un livre a été donné par l'homme au garçon", et : "un livre a été donné par l'homme à la table" ; la relation entre les degrés d'acceptabilité est la même pour les deux couples de phrases. Sur la base d'observations de cette nature, il est possible de définir des classes d'équivalence entre des schémas de phrase présentant les mêmes inégalités d'acceptabilité pour toutes leurs réalisations (cf. : [Harris 1965]) ; ces classes sont des *classes de transformation* d'une structure de phrase en une autre structure de phrase équivalente. Nous verrons plus loin de quelle nature sont ces transformations.

Les recherches de HARRIS ne s'arrêtent pas à la syntaxe de la phrase. Sous l'étiquette d'*analyse du discours*, il propose en effet une procédure permettant de déterminer, dans un texte quelconque, des classes d'équivalence d'éléments du texte en fonction de l'identité ou de l'équivalence distributionnelle de leur environnement (cf. p.ex. [Harris 1952]). Une telle procédure peut constituer, nous semble-t-il, une préparation à l'analyse sémantique d'un texte (un questionnaire complet par exemple) ; d'autant que la signification d'un

morphème ou d'une suite de morphèmes dépend, dans une certaine mesure, de son environnement, même si cette relation est loin d'être bi-univoque (cf. p. ex. : [Harris 1954 : 31-32]).

L'intérêt de l'oeuvre de HARRIS, pour notre recherche et pour les recherches futures sur le questionnaire, tient à son caractère opératoire. Toutes les procédures décrites ont été automatisées, ou sont présentées avec une rigueur telle qu'elles semblent automatisables ; en outre, ces procédures n'utilisent que les caractéristiques manifestes des matériaux linguistiques (sons ou textes écrits), sans jamais faire appel au sens. Le chercheur dispose par conséquent d'un ensemble d'outils susceptibles d'être appliqués au texte des tests et des questionnaires (voire aux protocoles de tests projectifs, d'observations, ou d'entretien), dans une phase préalable à une analyse sémantique.

Toutefois, dans la perspective plus limitée d'une description formelle des énoncés interrogatifs, les travaux de HARRIS sur la syntaxe de la phrase apportent également une solution au problème posé au départ : la réduction des énoncés interrogatifs à une forme canonique unique.

b) Les éléments de la phrase : opérateurs et arguments

Dans la présentation des principales conceptions de HARRIS en matière de syntaxe, nous nous en tiendrons à la terminologie et à la symbolique utilisées dans son ouvrage le plus récent, *Notes du cours de syntaxe* [Harris 1976]; cela nous conduira à reformuler les indications extraites en particulier de *Structures mathématiques du langage* [Harris 1968] et de *The Elementary Transformations* [Harris 1964].

La distinction entre opérateurs et arguments résulte de l'observation suivante : "il existe certaines séquences de mots $A...DE...G$ telles que les séquences $A...DE...G$ et $A...D$ appartiennent à l'ensemble des phrases, alors que

la séquence $E...G$ n'y appartient pas. Dans de telles situations, nous dirons que $E...G$ est un opérateur qui nécessite $A...D$ comme argument" [Harris 1976 : 14]. Par exemple, dans la phrase : "le livre de cuisine est tombé", la séquence : "le livre est tombé" appartient à l'ensemble des phrases, alors que ni la séquence : "de cuisine", ni : * "le de cuisine est tombé" n'y appartiennent ; par conséquent, "de cuisine" est un opérateur qui nécessite "livre" comme argument [Harris 1976 : 25].

Il serait commode de réduire toute phrase de la langue à une suite d'opérateurs et d'arguments. Cependant, il existe des phrases courtes dans lesquelles chaque partie nécessite les autres. Par exemple : "la neige fond". On remarque toutefois que "fond" a des propriétés qui caractérisent les opérateurs (il porte l'indication du temps et se place après son premier argument), et que "neige" se comporte comme l'argument unique de "fond". Cette constatation permet d'étendre à ce type de phrases les notions d'opérateurs et d'arguments, et de considérer que "toute séquence de mots qui consiste en un opérateur concaténé avec un ou plusieurs arguments ordonnés, est un discours " [Harris 1976 : 26].

Il est donc possible de décrire la structure de la phrase sur le modèle d'une structure algébrique (cf. p. ex. : [Harris 1966]), en termes d'*opérateurs* et d'*arguments*. Un opérateur peut d'ailleurs être l'argument d'un autre opérateur. Ceci nous conduit à distinguer à l'intérieur de l'ensemble des arguments, les *arguments élémentaires*, arguments qui ne sont pas eux-mêmes des opérateurs ; l'ensemble N de ces arguments correspond en gros aux noms (à certaines exceptions près). Pour classer les opérateurs O , qui sont pour la plupart des verbes, on peut se baser sur le nombre d'arguments qu'ils requièrent (comme TESNIÈRE fait pour les verbes à partir du nombre de leurs actants), ou sur la nature de leurs arguments. Ce dernier critère conduit à distinguer les *opérateurs élémentaires*, dont tous les arguments sont des arguments élémentaires, et les *opérateurs s'appliquant à un (ou plusieurs) discours*, dont l'un des arguments au moins est lui-même un opérateur.

Comme exemples d'opérateurs élémentaires à un seul argument O_n , on peut citer les verbes comme *démissionner* ("Jean démissionne") ou *dormir* ("Max dort"), des adjectifs comme *vieux* ("Max est vieux"), des adverbes comme *debout*

("Pierre est debout"). Les opérateurs élémentaires à deux arguments O_{nn} comprennent des verbes comme *manger* ("Max mange du poisson") ou *rayer* ("le diamant raje l'acier"), des adverbes comme *près* ("François est près de la maison"), des noms relationnels comme *père* ("Robert est le père de Gérard"). Parmi les opérateurs élémentaires à trois arguments O_{nnn} , citons des verbes comme *mettre* ("Bernard a mis du sel sur la laitue") et des noms comme *ambassadeur* ("Campbell est ambassadeur de Grande-Bretagne en France"). Sur ces quelques exemples, on note qu'en français (comme en anglais) l'opérateur suit immédiatement son premier argument (nous verrons plus loin que les discours où cette règle de position et de contiguïté n'est pas respectée sont liés par une relation d'équivalence à des discours où elle est appliquée). On note également que certains mots (prépositions en français et en anglais, prépositions ou suffixes de cas dans d'autres langues) jouent le rôle d'indicateurs du rang de l'argument (pour les second et troisième arguments principalement).

Certains opérateurs non élémentaires s'appliquent à la fois à des arguments élémentaires et à un discours (c'est-à-dire à un autre opérateur, accompagné de ses arguments). Exemples de O_{no} : les verbes tels que *savoir* ("Jean sait que Paul est arrivé"), *croire* ("je crois qu'il est parti"), *espérer* ("nous espérons qu'il viendra"), *voir* ("Claude a vu que Denis traversait la rue"), *dire* ("il a dit qu'il viendrait"). Exemples de O_{on} : les verbes comme *surprendre* ("que la neige fonde surprend Jacques") et les prépositions de temps, de lieu, de manière, etc., telles que *pendant* ("il est mort pendant la guerre"), *à* ("je suis né à Paris"), etc. Exemples de O_{nno} : le verbe *informer* ("Henri a informé Louis de ce que Marie avait gagné le prix"), et à nouveau le verbe *dire* ("Jean dit à François que la neige fond").

Les opérateurs ne s'appliquant qu'à des discours (n'ayant que des opérateurs comme arguments) sont, en français (comme en anglais), de trois types. Le premier type, O_o , n'a qu'un argument : c'est le cas par exemple pour le nom *fait* ("qu'il soit français est un fait") ou l'adjectif *possible* ("qu'il soit en faute est possible") ; en anglais, c'est également le cas pour un verbe comme *to continue* ("the child's crying continued"). Les opérateurs s'appliquant à deux discours, O_{oo} , constituent les deux autres types : ils peuvent en effet être soit associatifs, soit non associatifs. Exemples de O_{oo} associatifs :

les conjonctions *et* ("Jean a téléphoné et Paul est revenu"), *mais* ("je le lui ai proposé mais il a refusé"), *ou* ("je lui ai demandé s'il prendrait le même ou s'il changerait"). Le type des O_{oo} non associatifs est illustré par des verbes ou des expressions signifiant *causer*, *impliquer* ("que Jean soit arrivé a fait que Richard est parti").

A ces opérateurs s'appliquant à des discours ou des éléments de discours, HARRIS ajoute deux autres espèces d'opérateurs, les *opérateurs de méta-discours*, dont nous allons voir le rôle au paragraphe suivant, et les *opérateurs méta-linguistiques*, qui permettent de rendre compte de phrases telles que : "Marie est un nom", ou "Je suis venu est une phrase".

c) Les transformations élémentaires

Les exemples de discours que nous avons donnés pour illustrer les divers types d'opérateurs ont tous une structure relativement simple ; en particulier, tous les *discours élémentaires* (les phrases formées autour d'un opérateur élémentaire) sont de la forme : "premier argument-opérateur-(second argument)-(troisième argument)", sans permutation ni adjonction d'aucune sorte. Ils font partie de ce que HARRIS appelle les *discours de stricte concaténation* : ceux-ci "constituent un sous-ensemble important de la langue parce qu'ils ont une structure simple, et, comme on le verra, parce qu'ils peuvent exprimer tout ce qui peut être exprimé dans la langue. Ce ne sont pas les seuls discours. Cependant, tous les autres discours de la langue sont liés à ceux de concaténation par une relation d'équivalence particulière" [Harris 1976 : 37]. Cette relation d'équivalence définit, sur l'ensemble des discours, des classes de transformation dont les éléments constituent des variantes d'un même discours ; les procédures de passage d'une forme de discours à une autre forme appartenant à la même classe d'équivalence sont les opérations de *transformation* des énoncés.

"Une relation transformationnelle entre deux ensembles de propositions A_1 et A_2 peut être décrite comme une application biunivoque ϕ de A_1 sur A_2 , telle que chaque proposition de A_2 est identique à son image inverse dans A_1 , à l'exception de quelques additions, permutations ou omissions qui accompagnent ϕ " [Harris 1968 : 69]. La transformation peut être décrite également comme une

opération sur les séquences de mots. Les transformations paraphrastiques élémentaires (ou : *réductions*) sont de trois sortes : la permutation, l'attachement, l'effacement. L'application, dans un ordre déterminé, d'une série de ces transformations, permet de rendre compte de toutes les variantes d'un discours donné. Selon les langues, il peut s'y ajouter des transformations morphophonémiques obligatoires, telles que certaines règles d'accord (du type : pluriel de *cheval* ou de *knife*) ou les variantes de verbes irréguliers (comme *aller* en français, ou *to be* en anglais). Ces transformations permettent de rendre compte des irrégularités de la langue (cf. [Gross 1977 : 235]). En général, les transformations paraphrastiques peuvent s'effectuer dans les deux sens (contrairement à la présentation usuelle des transformations en grammaire générative) ; toutefois, il est commode pour l'exposé de choisir une direction donnée (cf. [Harris 1968 : 69-71]), tout en considérant que la variante du discours choisie comme point de départ de la suite de transformations est une *forme distinguée* pour représenter une classe de transformations, et non un énoncé "source" (cf. [Harris 1976 : 38]).

Le groupe des *permutations* ϕ_p rassemble toutes les transformations qui modifient simplement l'ordre des opérateurs et des arguments. En anglais, les ϕ_p peuvent servir à mettre l'accent sur certains éléments de l'énoncé, en opérant un effet de contraste avec les autres phrases du discours [Harris 1968 : 75]. Par exemple :

- "I will think about the problem while driving"
- (ϕ_p) → "I will think, while driving, about the problem"
- (ϕ_p) → "I will, while driving, think about the problem"
- (ϕ_p) → "I, while driving, will think about the problem"
- (ϕ_p) → "while driving, I will think about the problem" [Harris 1968 : 88].

En français, la transformation ϕ_p est illustrée par le maître de philosophie de Monsieur Jourdain (MOLIÈRE, *Le Bourgeois Gentilhomme*, II IV), qui l'applique d'ailleurs de manière abusive (les phrases engendrées n'appartiennent pas toutes à la langue). Nous verrons au paragraphe suivant la place de la permu-

tation dans une chaîne de transformations permettant de passer d'un énoncé à un autre énoncé équivalent. A titre d'exemples simples, citons :

- le déplacement de la particule préverbale ("pronom personnel")

[Gross 1968 : 56] ; exemple :

* "j'espère voir vous"

(ϕ_p) → "j'espère vous voir".

- le déplacement du déterminant adverbial [Gross 1977 : 42], qui peut donner lieu à des permutations alternées successives :

"il a été bu beaucoup trop de vin"

(ϕ_p) → * "il a été beaucoup bu trop de vin"

(ϕ_p) → "il a été beaucoup trop bu de vin"

(ϕ_p) → * "il a beaucoup été trop bu de vin"

(ϕ_p) → "il a beaucoup trop été bu de vin" [Gross 1977 : 43].

- le déplacement du prédéterminant du groupe nominal ; exemple :

"Luc tient à à peu près trois livres"

(ϕ_p) → "Luc tient à peu près à trois livres" [Gross 1977 : 56].

- la permutation de longueur [Harris 1964 : 516] "qui a pour effet de permuter certains opérateurs courts (ou certains opérateurs accompagnés de leur second objet) avec certains seconds arguments longs de leur argument" [Harris 1976 : 148]. Par exemple :

* "je crois qu'il vienne à Paris impossible"

(ϕ_p) → "je crois impossible qu'il vienne à Paris" (cf.: [Gross 1968 : 120]).

Les transformations d'*attachement* ϕ_m ont pour résultat la combinaison d'un opérateur avec l'un de ses arguments. Les principales formes d'attachement sont relatives (cf. [Harris 1976 : 125-138]) :

- aux flexions, qui ont pour effet de combiner en un seul morphème un opérateur (sous la forme d'un affixe) et l'un de ses arguments. C'est par exemple le cas, en français comme en anglais, de la marque du temps dans les formes verbales.

- aux mots composés, qui sont le produit d'un mécanisme analogue au précédent. Exemples : "childhood", "anti-nucléaire", "parapluie", etc.

- à l'intonation de phrase (et éventuellement aux signes de ponctuation qui en sont les équivalents graphiques). Exemples (accent contrastif, et intonation de question) :

"I like books and not other things"
 (ϕ_m) → "I like books and not other things" [Harris 1968 : 90].
 * "I ask what will he write"
 (ϕ_m) → "I ask : what will he write ?" [Harris 1968 : 107].

Les transformations d'*effacement* ϕ_z ont pour effet de réduire à zéro certains morphèmes, et, dans certains cas, d'introduire des substituts des morphèmes effacés (pronominalisation) [Harris 1968 : 89 ; 1976 : 46]. Les cas d'effacement ou de pronominalisation les plus notables sont :

- l'effacement de mots répétés, en particulier sous les opérateurs *et* et *ou* ; exemple :

"Jean a vu Marie et Jean a appelé Marie"
 (ϕ_z) → "Jean a vu et appelé Marie" [Harris 1976 : 101].

- la pronominalisation de mots répétés (qui, dans l'exemple précédent, aurait pu donner : "Jean a vu Marie et l'a appelée"), qui porte le plus souvent sur des noms ou des adjectifs ; ainsi [Gross 1977 : 36-37] :

"nous avons gagné un gâteau, la moitié du gâteau me revient"
 (ϕ_z) → "nous avons gagné un gâteau, la moitié m'en revient" ;

"Luc est trop gros"
 (ϕ_z) → "Luc l'est trop".

- l'effacement de constantes, telles que les mots indéfinis ("quelque chose", "certaines personnes", "quelque part", ...), et le pronom relatif *qui* suivi du verbe *être*. Par exemple :

"je regarde des gens jouer aux échecs"
 $(\phi_z) \rightarrow$ "je regarde jouer aux échecs" [Harris 1976 : 109] ;

"Luc apprécie cela"
 $(\phi_z) \rightarrow$ "Luc apprécie" [Gross 1977 : 190] ;

"Les invités, qui étaient déçus, sont repartis"
 $(\phi_z) \rightarrow$ "Les invités, déçus, sont repartis" [Gross 1977 : 32].

- l'effacement de mots appropriés (parmi lesquels les formes performatives du type : "je vous demande : "). Exemples :

"il lit les écrits de Shakespeare"
 $(\phi_z) \rightarrow$ "il lit Shakespeare" [Harris 1976 : 114] ;

"Luc est resté pendant beaucoup de temps dans son lit"
 $(\phi_z) \rightarrow$ "Luc est resté beaucoup de temps dans son lit" [Gross 1977 : 40] ;

"je te demande : vient-il ?"
 $(\phi_z) \rightarrow$ "vient-il ?" [Harris 1976 : 123].

Pour rendre compte de manière satisfaisante des conditions d'application de la transformation d'effacement, il est nécessaire d'introduire la notion d'*opérateur de méta-discours*, signalée au paragraphe précédent. On peut distinguer deux types d'opérateurs de méta-discours : l'opérateur porteur d'adresses, de forme O_o , et l'opérateur *wh-* (ou *qu-*) introduisant une relative, de forme O_{oo} . De tels opérateurs "peuvent paraître particulièrement complexes et n'être qu'une fiction théorique, mais ils sont en fait inévitables pour toute analyse simple du langage" [Harris 1976 : 85].

En effet, considérons par exemple la phrase : "j'ai écrit à Jean et j'ai téléphoné à Jean". Si cette phrase admet pour variante (après effacement) : "j'ai écrit et téléphoné à Jean", c'est à la condition que la personne désignée par le premier "Jean" soit la même que celle désignée par le second "Jean". Or, rien dans la phrase originale n'indique que cette condition soit remplie. D'au-

tre part, cette identité du référent n'a rien d'évident, comme en témoigne par exemple la phrase suivante : "Monsieur André Henry, grand secrétaire aux affaires extérieures du Grand Orient de France, nous demande de préciser qu'il n'est pas Monsieur André Henry, président de la FEN" (*Le Monde*, 16 septembre 1978, page 7). Enfin, on ne peut faire l'hypothèse que l'information sur les identités soit connue du locuteur et du récepteur en dehors de la phrase à laquelle elle se réfère. Par conséquent, il est nécessaire de postuler un opérateur de méta-discours O_o , s'appliquant à la phrase originale, et annonçant que la phrase "a un second argument du second argument désignant la même personne que le second argument du premier argument" (ou en abrégé : "contient en 2.2. le même individu qu'en 1.2"). De même, la phrase : "j'ai acheté un livre et elle en a acheté un" provient par ϕ_z de : "j'ai acheté un livre et elle a acheté un livre", avec application de l'opérateur de méta-discours : "contient 1.2 identique à 2.2" [Harris 1976 : 84-85].

L'opérateur O_{oo} formant la proposition relative contient également une assertion sur l'identité de deux adresses. Les différences avec l'opérateur de méta-discours O_o porteur d'adresses tiennent à ce que l'identité affirmée s'applique non seulement aux mots contenus dans les adresses, mais aussi et toujours aux individus ou objets auxquels renvoient ces mots (les *référents*), comme dans le premier exemple ci-dessus ; et à ce que l'opérateur *wh-* (ou *qu-* en français) relie deux phrases, et non plus seulement deux arguments d'une même phrase. Le pronom relatif peut donc être considéré comme une variante de cet opérateur ; par exemple, la phrase : "l'homme que j'ai vu est parti" peut être considérée comme une variante de : "*l'homme est parti* contient un argument 1 identique à l'argument 2 de *j'ai vu l'homme*" [Harris 1976 : 88-89].

d) L'interrogation selon HARRIS

Toute phrase bien formée peut être réduite, par transformations successives, à un ou plusieurs discours élémentaires. Nous l'avons vu, un discours élémentaire est une forme assertive construite autour d'un opérateur élémentaire unique. Par conséquent, et puisque les transformations paraphrastiques peuvent fonctionner dans les deux sens, toute forme interrogative peut être considérée

comme la transformation d'une ou de plusieurs formes assertives.

Dans *Structures mathématiques du langage*, HARRIS développe trois exemples d'interrogation en anglais ; il n'est malheureusement pas toujours possible de les transposer tels quels en français, sous peine d'introduire certaines transformations morphophonématiques supplémentaires qui les alourdiraient inutilement. Par contre, dans les *Notes du cours de syntaxe*, on trouve quelques exemples en langue française, mais peu développés. Toutefois, à partir de l'ensemble de ce matériel, et en l'adaptant au besoin, il est relativement aisé de présenter la manière dont cet auteur analyse les formes interrogatives. Les exemples choisis couvrent les cas les plus courants d'interrogation : question partielle portant sur le sujet du verbe, l'objet, le verbe lui-même ; puis question portant sur l'ensemble de la phrase (interrogation totale, à réponse *oui/non*).

La question : "qui est ici ?" (réponse : "Jean") est une paraphrase de : "je vous demande si N_1 est ici ou si N_2 est ici ou ... ou si Jean est ici ou si N_j est ici". Les N_i (dont "Jean" fait partie) désignent l'ensemble des arguments possibles de l'opérateur O_n "(est) ici", c'est-à-dire l'ensemble des sujets possibles. Le verbe "demander (si)" est un opérateur O_{nno} (ou O_{no}) qui s'applique à "ou" [Harris 1976 : 83-84], le *ou* pouvant d'ailleurs être répété [Harris 1968 : 148, 211 ; 1976 : 128]. Les étapes de la réduction sont les suivantes [Harris 1976 : 156] :

① *effacement* (ϕ_z) des arguments se correspondant et comportant des adresses. Dans cet exemple, les mots identiques étant des opérateurs, il est nécessaire que leurs relations d'opérateurs à arguments soient les mêmes pour que la variante zéro puisse leur être substituée à la deuxième, troisième, $j^{\text{ième}}$ adresse (cf. [Harris 1976 : 101]). On obtient ainsi : "je vous demande si N_1 est ici ou N_2 ou ... ou Jean ou N_j ".

② *permutation* (ϕ_p) du résidu de la deuxième, troisième, $j^{\text{ième}}$ phrase après effacement : "je vous demande si N_1 ou N_2 ou ... ou Jean ou

N_j est ici".

③ *pronominalisation* (ϕ_z) des N_i répétés. La pronominalisation de la disjonction "si N_1 ou N_2 ou ... ou N_j " se fait automatiquement par combinaison avec "si", sous la forme "qui" ou "que" (cf. [Harris 1968 : 107, n27]). Ce qui donne : "je vous demande qui est ici".

④ *attachement* (ϕ_m) de l'intonation interrogative (marquée par le point d'interrogation) : "je vous demande : qui est ici ?".

⑤ *effacement* (ϕ_z) du performatif. Nous reviendrons à la fin de ce chapitre sur cette notion de performatif, introduite par J.L. AUSTIN. Rappelons simplement pour l'instant que les verbes signifiant *dire*, *demander*, *ordonner*, qui permettent l'attachement d'une intonation assertive, interrogative, ou impérative (respectivement), ont une forme distinguée, où les deux premiers arguments sont le locuteur ("je") et le récepteur ("te", "vous"), et où le temps est le présent de l'indicatif. Cette situation a la propriété unique de rendre effective l'assertion, l'interrogation, ou le commandement. Dans cette situation, et dans celle-ci seulement, le verbe est effaçable [Harris 1976 : 122-123]. On aboutit ainsi à : "qui est ici ?".

Pour résumer, la question portant sur le sujet de la phrase est obtenue par la suite de transformations paraphrastiques :

- "Je vous demande si N_1 est ici ou si N_2 est ici ou ... ou si Jean est ici ou si N_j est ici"
- (ϕ_z) → "Je vous demande si N_1 est ici ou N_2 ou ... ou Jean ou N_j "
- (ϕ_p) → "je vous demande si N_1 ou N_2 ou ... ou Jean ou N_j est ici"
- (ϕ_z) → "je vous demande qui est ici"
- (ϕ_m) → "je vous demande : qui est ici ?"
- (ϕ_z) → "qui est ici ?"

On obtient de même la question portant sur l'objet de la phrase (le deuxième argument de l'opérateur) par une série de transformations compa-

nable. Dans l'adaptation que nous faisons de l'exemple original [(Harris 1968 : 107)], certaines formes intermédiaires n'appartiennent pas à la langue. Soit la question : "que va-t-il écrire ?" (réponse : "une lettre"), on a la chaîne de transformations :

- "je demande s'il va écrire N_1 ou s'il va écrire N_2 ou ... ou s'il va écrire une lettre ou s'il va écrire N_j "
- $(\phi_z) \rightarrow$ "je demande s'il va écrire N_1 ou N_2 ou ... ou une lettre ou N_j "
- $(\phi_p) \rightarrow$ *"je demande si N_1 ou N_2 ou ... ou une lettre ou N_j il va écrire"
- $(\phi_z) \rightarrow$ "je demande (ce) qu'il va écrire"
- $(\phi_p) \rightarrow$ *"je demande que va-t-il écrire"
- $(\phi_m) \rightarrow$ "je demande : que va-t-il écrire ?"
- $(\phi_z) \rightarrow$ "que va-t-il écrire ?"

Evidemment, pour que cet exemple soit complètement développé, il serait nécessaire d'explicitier les transformations rendant compte par exemple des variations de forme du pronom interrogatif ("que", "qu'", "quoi", "ce que") qui n'existent pas en anglais ("what"), ou de l'apparition du *t* euphonique. Cette présentation suffit toutefois à explicitier le mécanisme de dérivation de la question à l'aide des seules transformations paraphrastiques.

L'exemple donné par HARRIS de question portant sur le verbe de la phrase ([Harris 1968 : 107]) nous a paru plus difficile à adapter que l'exemple précédent, le français ne disposant pas, avec le verbe "faire", d'un mot équivalent au verbe anglais "do". C'est pourquoi nous le donnons sous sa forme originale. La question est : "what will he do ?" (réponse : "smoke"). On suppose la possibilité de remplacer : "he will smoke" par : "he will do smoking", et plus généralement : "he will V_i " par : "he will do V_i ing" (cette transformation de la forme verbale est décrite, et discutée, dans [Harris 1964 : 503-504] et dans [Harris 1968 : 108, 190]). On obtient la série :

- "I ask whether he will do V_1 ing or he will do V_2 ing or ... or he will do smoking or he will do V_j ing"
- $(\phi_z) \rightarrow$ "I ask whether he will do V_1 ing or V_2 ing or ... or smoking or V_j ing"

- $(\phi_p) \rightarrow$ *"I ask whether V_1 ing or V_2 ing or ... or smoking or V_j ing he will do"
 $(\phi_z) \rightarrow$ "I ask what he will do"
 $(\phi_p) \rightarrow$ *"I ask what will he do"
 $(\phi_m) \rightarrow$ "I ask : what will he do ?"
 $(\phi_z) \rightarrow$ "what will he do ?"

Lorsque la question porte sur l'ensemble de la phrase (interrogation totale), le mécanisme de dérivation est analogue à celui observé pour les questions partielles. La seule différence est que, dans la forme distinguée, la disjonction porte non plus sur un ensemble de phrases S_i ("je demande si S_1 ou si S_2 ou ... ou si S_j "), mais sur une phrase unique et sa négation ("je demande si S_1 ou ne pas S_1 ") [Harris 1976 : 128]. Dans ces conditions, la négation de S_1 n'apporte pas d'information par rapport à S_1 elle-même, et peut par conséquent prendre la variante zéro (effacement de mots appropriés ϕ_z). Ainsi, la question : "restez-vous ?" s'obtient par les transformations :

- "je vous demande si vous restez ou si vous ne restez pas"
 $(\phi_z) \rightarrow$ "je vous demande si vous restez"
 $(\phi_z) \rightarrow$ *"je vous demande vous restez"
 $(\phi_m) \rightarrow$ "je vous demande : vous restez ?"
 $(\phi_p) \rightarrow$ "je vous demande : restez-vous ?"
 $(\phi_z) \rightarrow$ "restez-vous ?"

Pour résumer, l'essentiel des propositions de HARRIS concernant l'interrogation tient dans les citations suivantes :

- "il est possible de dériver toutes les formes interrogatives ou impératives de formes assertives avec un opérateur performatif W " [Harris 1964 : 528].

- "toutes les questions (à la fois les questions à réponse *oui/non* et les questions *wh-*) peuvent être obtenues de façon simple, sans recourir à des transformations spéciales, à partir d'un opérateur particulier W [performatif] opérant sur une disjonction de S [phrases élémentaires] :

je demande si S_1 ou si S_2 ... ou si S_n

dans laquelle la suite *ou si S_2 ... ou si S_n* peut être remplacée par la forme effaçable *ou non*" [Harris 1965 : 561].

- "*what did he take ?* peut se dériver comme transformée de ses réponses *he took a book, he took a pen, etc.*" [Harris 1968 : 190].

Dans le langage des *Notes du cours de syntaxe*, nous dirions que toute question appartient à la même classe d'équivalence (ou : est une paraphrase transformationnelle) d'une forme distinguée, caractérisée par un opérateur O_{nno} qui est le verbe *demande* au présent de l'indicatif, avec comme premier argument : *je*, comme second argument : *te* ou : *vous*, et comme troisième argument l'opérateur O_{oo} de disjonction : *ou*, éventuellement répété ; ce dernier opérateur a pour arguments un ensemble de phrases qui constituent l'ensemble des réponses possibles à la question. Ainsi, dans l'exemple très simple que nous avons déjà utilisé (cf. schémas n° 1.35 et 1.64) : "je vous demande qui est arrivé à dix heures", l'ensemble des réponses possibles serait : " N_1 est arrivé à dix heures ; N_2 est arrivé à dix heures ; ... ; le directeur est arrivé à dix heures ; N_j est arrivé à dix heures". La forme distinguée de l'ensemble des questions appartenant à la même classe d'équivalence serait obtenue en appliquant l'opérateur de disjonction (*ou*) à l'ensemble des réponses possibles, puis l'opérateur performatif (*je vous demande si*) à la suite ainsi formée.

e) Apport des grammaires transformationnelles à l'analyse de l'interrogation

La réponse apportée par HARRIS à l'analyse des formes interrogatives ne ressemble ni à la solution de TESNIÈRE, ni à celle de CHOMSKY. Sous sa forme la plus récente [Harris 1976], elle conduit à définir des ensembles de phrases synonymes, et des règles de passage d'une forme appartenant au même ensemble, sans qu'il soit nécessaire de postuler l'existence d'une phrase énonciative génératrice des diverses questions possibles (TESNIÈRE), ou d'une structure profonde abstraite source d'un ensemble d'énoncés de même signification (CHOMSKY). Ainsi, l'identité de sens que l'on peut constater empiriquement entre des questions telles que : "pour qui avez-vous l'intention de voter ?"

et : "avez-vous l'intention de voter pour X , pour Y , ou pour Z ?", se trouve fondée en théorie par l'existence d'opérations permettant de passer de l'une à l'autre de ces formulations. Il semble donc que nous disposions, avec la grammaire transformationnelle, de l'outil conceptuel nous permettant de subsumer, sous une forme promue au rang de forme canonique, la diversité des énoncés interrogatifs. D'une manière inattendue pour qui ne s'est nourri que des grammaires traditionnelles, mais avec de solides arguments (cf. [Harris 1976 : 16-17, 38]), HARRIS choisit d'ailleurs comme élément distingué représentatif d'une classe de questions synonymes une forme que nous n'avons jamais observée dans notre corpus : l'interrogative indirecte disjonctive introduite par une énonciative (à valeur performative) : "je vous demande si vous avez l'intention de voter pour X ou si vous avez l'intention de voter pour Y ou si vous avez l'intention de voter pour Z". L'intérêt de cette forme est qu'il est extrêmement aisé d'identifier chacun des discours élémentaires qui la composent, et qui constituent autant de réponses possibles à la question posée.

Toutefois, en l'état actuel des recherches de HARRIS ou de GROSS, il nous faut émettre deux réserves importantes. La première renvoie à nos remarques concernant la polymorphie des énoncés exprimant une question. Il est possible de considérer par exemple que les énoncés : "pour qui allez-vous voter ?" et : "pour quel candidat allez-vous voter ?", sont à peu près équivalents (dans un questionnaire d'enquête) ; et que : "je voudrais que vous me disiez", "voudriez-vous me préciser", et : "indiquez-moi", ont tous trois pour effet d'inciter la personne interrogée à répondre à la question posée. Or, nous n'avons pas trouvé, dans les publications s'inspirant des travaux de HARRIS que nous avons consultées, ni d'ailleurs chez HARRIS lui-même, d'analyse complète des multiples variantes d'énoncés à valeur de question. L'analyse du premier exemple nous conduirait à adopter comme forme distinguée : "je vous demande si vous allez voter pour X ou ... ou si vous allez voter pour Z", la disjonction portant sur des discours élémentaires du type : "vous allez voter pour X". Pour passer à l'énoncé : "vous allez voter pour le candidat X ", il faut postuler l'application d'un opérateur de méta-discours O_{oo} : "*vous allez voter pour X* contient un argument 2 identique à l'argument 1 de *X est candi-*

dat" ; d'où : "vous allez voter pour X qui est candidat", d'où enfin : "vous allez voter pour le candidat X ". Seulement, pour que nous puissions considérer les formes : "pour qui" et : "pour quel candidat" comme équivalentes (équivalence que l'on peut d'ailleurs constater empiriquement à partir d'un corpus de questions), il nous faut postuler l'identité de "voter pour X" et "voter pour le candidat X". La transformation sous-jacente ϕ_z (effacement de mots appropriés) devrait se fonder sur le fait que l'opérateur O_{nn} : "voter" implique que son second argument ait la qualité de candidat (on vote pour un candidat, on boit une boisson, etc.). Cette analyse n'est nullement en contradiction avec celles de HARRIS (cf.p.ex. [Harris 1976 : 114-115]) ou de GROSS (cf. [Gross 1977 : 39-40]) ; mais en l'absence d'un corpus suffisant (ou de résultats théoriques appropriés), elle risquerait de nous entraîner vers des considérations logico-sémantiques que nous n'aborderons qu'au chapitre 2. D'autre part, le second exemple nous amène à placer dans une même classe de transformations une forme assertive, une forme interrogative, et une forme impérative ; la justification de cette classification se trouve dans la notion de forme performative, qui nous sera nécessaire pour réaliser la réduction à une forme unique des énoncés de questions. Remarquons cependant que le problème des transformations permettant de passer de l'un à l'autre de ces énoncés n'a pas, à notre connaissance, été entièrement résolu par les grammaires transformationnelles.

La seconde réserve que nous formulons porte sur le sens de la relation d'équivalence établie par HARRIS entre des discours différents. Nous l'avons vu, cette relation se fonde sur la constatation suivante : "dans une phrase S_i , il peut être possible d'identifier une phrase S_j accompagnée de matériel supplémentaire X [...]. [Ceci est] le résultat d'une analyse, puisqu'il est nécessaire de démontrer que la portion S_j de S_i est en réalité la même phrase que la phrase indépendante S_j et non pas simplement les mots de S_j qui apparaîtraient ensemble pour d'autres raisons. [Une démonstration possible] consiste à prendre pour chaque X un ensemble de S_i , et à voir si leurs inégalités de vraisemblance sont identiques ou liées de façon déterminée *a priori* aux inégalités de vraisemblance des S_j correspondants "[Harris 1976 : 13-14]. L'application de cette procédure conduirait sans doute à ranger dans une même classe de transformations les tournures performatives citées à l'alinéa précé-

dent. Mais elle conduirait également à considérer comme synonymes les énoncés fameux des premières recherches sur la formulation des questions d'enquête (cf. [Cantril 1940], [Rugg 1941]), tels que : "do you think the United States should allow (not forbid) speeches against democracy ?". Or nous savons bien qu'en matière d'enquêtes par sondage, deux formulations "synonymes" ne donnent pas lieu à des distributions de réponses identiques ; doit-on alors toujours parler de synonymie ? Nous reviendrons ultérieurement sur ces problèmes de sémantique des questions. Nous pouvons toutefois affirmer dès maintenant que les grammaires transformationnelles, dans leur état actuel, ne nous permettent pas encore de répondre d'une manière satisfaisante au problème posé au départ, celui de la réduction à une forme canonique des énoncés à valeurs de questions.

L'analyse transformationnelle des énoncés de questions utilisés dans les tests et les questionnaires reste par conséquent à faire. Ce type de recherche, de type inductif au départ, conduit à des résultats plus stables que la démarche hypothético-déductive, mais demande en contrepartie une somme d'efforts considérable. Il n'est donc pas certain qu'une telle recherche puisse être menée à son terme dans un proche avenir. Malgré cela, l'apport de HARRIS et de ses continuateurs à l'analyse des énoncés interrogatifs nous paraît décisif au moins à trois égards.

Tout d'abord, les procédures d'analyse linguistique mises au point à l'université de Pennsylvanie nous paraissent pouvoir être appliquées sans modification à l'étude des tests et des questionnaires d'enquête. Toutefois, si l'on désire recueillir par ce moyen des éléments en vue de réduire à une forme unique la variété des formulations, il serait judicieux, dans la phase initiale de cette analyse, à la fois de limiter à un étroit domaine d'application (par exemples : questions portant sur le vote, ou sur la lecture des quotidiens, ou sur la profession exercée), et de l'étendre à la plus grande variété possible de sources (instituts de sondage). De cette manière, compte tenu de la diversité des styles propres aux auteurs de questionnaires (et du peu d'avancement des tentatives de standardisation des énoncés), la principale source de variations sera d'ordre essentiellement linguistique. De ce point de vue, les questionnaires de personnalité ou les tests de connaissances ou d'aptitudes constitueraient

un matériau moins riche en formes linguistiques, et moins abondant.

Ensuite, HARRIS est parvenu à faire ce que ni TESNIÈRE, ni CHOMSKY n'ont pu réaliser dans le cadre de leurs conceptions linguistiques : réduire l'interrogation totale et l'interrogation partielle à une forme unique. Nous avons vu en effet que, pour la grammaire transformationnelle, la seule différence entre ces deux types d'interrogation réside dans la nature des discours élémentaires concaténés par l'opérateur *ou* dans la forme distinguée (avec l'opérateur *demander si* à la première personne du présent de l'indicatif, et *tu* ou *vous* comme deuxième argument) : ou bien il n'y a qu'un discours élémentaire et sa négation, ou bien il y a plusieurs discours élémentaires différant les uns des autres par au moins un de leurs éléments. Dans le premier cas, on dira que l'interrogation est totale ; dans le second, qu'elle est partielle. Mais dans tous les cas, la réponse est l'un des discours concaténés, et les formulations des questions sont obtenues au moyen des mêmes variantes [Harris 1976 : 128]. Il s'agit bien, par conséquent, d'une seule et même forme linguistique.

Enfin, la conception de HARRIS sur l'interrogation prend pour point de départ l'ensemble des réponses possibles à la question. L'importance de ce point de départ avait déjà été entrevue par certains linguistes avant même l'apparition de la notion de transformation ; ainsi, DAMOURETTE et PICHON insistent sur "la liaison intime qui existe entre l'interrogation et la réponse", faisant remarquer que l'interrogation "reste en suspens tant au point de vue logique qu'au point de vue mélodique", et que la réponse en complète le sens, apportant "la solution mélodique et sémantique de l'interrogation" [Damourette 1931 : IV, 305-306]. On peut dire que TESNIÈRE et CHOMSKY, chacun à leur manière, tiennent compte de la réponse dans leur présentation de l'interrogation : TESNIÈRE, en partant d'une phrase énonciative contenant une réponse à chacune des questions possibles relativement à la situation décrite par cette phrase ; CHOMSKY, en faisant l'hypothèse d'une structure profonde telle que, si l'on remplace le constituant de phrase interrogatif par celui de l'affirmation, la réalisation du noyau de la phrase donne l'une des réponses possibles à la question. L'intérêt de la conception harrissienne, fondée sur la disjonction de l'ensemble des réponses possibles, est qu'elle rejoint l'un des postulats de base de la logique érotétique (postulat de HAMBLIN), selon lequel connaître l'ensemble des propo-

sitions qui constituent une réponse à une question est équivalent à connaître la question elle-même [Hamblin 1958 : 162].

Ainsi, à défaut de nous apporter une théorie achevée applicable à l'ensemble des énoncés à valeur de question et permettant de réduire ceux-ci à une forme unique, HARRIS nous propose une telle théorie pour l'ensemble des formes interrogatives au sens strict, et nous décrit le substrat commun à tous les énoncés à valeur de question : une forme performative opérant sur une disjonction de l'ensemble des réponses possibles à la question. Cette conception nous permettra de faire le passage de l'étude morpho-syntaxique de ces énoncés à leur analyse logico-sémantique.

1.2.4. Le substrat commun aux énoncés à valeur de question

Après avoir passé en revue trois grands courants de la recherche contemporaine en matière de syntaxe, nous constatons que leur caractéristique commune est d'accorder une place centrale à la notion de *transformation*. Certes, une telle notion n'est pas radicalement nouvelle, comme l'a fait remarquer Nicolas RUWET [Ruwet 1967 : 222-228]. Par exemple, Antoine ARNAULT et Claude LANCELOT y ont eu recours dans la *Grammaire générale et raisonnée de Port-Royal*, publiée en 1660 ("Dieu invisible a créé le monde visible" ; cf. [Gross 1967 : 186, 190-191]). Dans la première moitié du vingtième siècle, des auteurs comme Otto JESPERSEN, son disciple Andreas BLINKENBERG, ou Charles BALLY, ont utilisé des notions assez voisines. Plus récemment, Emile BENVENISTE, qui ne se réclame pas des grammaires transformationnelles, a décrit sous le nom de *transposition* un mécanisme analogue à celui des transformations : afin de rendre compte des huit emplois réguliers différents du génitif en latin, il est conduit à analyser divers types de conversion d'une forme verbale en une forme nominale ("fugere laborem" → "fugiens laboris" ; "consul advenit" → "adventus consulis" ; "animus patitur" → "patientia animi" et : "pati dolorem" → "patientia doloris" ; etc.), et à conclure "que tous les emplois du génitif sont engendré par cette relation de base, qui est de nature purement syntaxique" [Benveniste 1962 : 147]. D'autre part, nous avons vu à propos de l'interrogation

(§ 1.1.3.a) que les grammaires traditionnelles utilisent la notion de transformation sans la nommer, pour décrire les règles de passage du style direct au style indirect. Mais en tout état de cause, si certains des auteurs cités font de cette notion un usage assez rigoureux, aucun n'en fait un usage systématique.

Or, celle-ci nous paraît très importante pour notre propos, non seulement en elle-même, mais surtout par ses implications. En l'absence d'une théorie syntaxique complète sur la manière d'interroger en français, nous pouvons cependant avancer l'hypothèse qu'en l'état actuel des recherches morphosyntaxiques, une telle théorie doit nécessairement s'appuyer sur la notion de transformation et ses corrélats (classe de transformation ; phrase-noyau ou discours élémentaire ; forme distinguée ou source). Si cette hypothèse devait se révéler fondée, il resterait toutefois de nombreux problèmes en suspens, dont la solution est indispensable pour édifier une telle théorie ; nous en rappellerons les principaux, en indiquant quelles nous paraissent être les directions de recherche envisageables à court terme en vue de leur apporter des éléments de réponse. Enfin, en anticipant sur les résultats espérés de telles recherches, nous proposerons une définition du substrat commun à l'ensemble des énoncés à valeur de question, dans la perspective de leur réduction à une forme canonique.

a) La notion de transformation et ses implications

Pour qui désire réduire les énoncés à valeur de question à une forme unique, la notion de transformation nous paraît être une notion fondamentale. En effet, à certaines conditions que nous allons examiner, une telle notion peut déboucher sur l'établissement d'une relation d'équivalence syntaxique (voire sémantique) entre des énoncés différents, et sur la décomposition d'énoncés longs ou complexes en segments élémentaires standardisés. Il est possible par conséquent, à partir de relations d'équivalence entre énoncés et de règles de passage établies d'une manière appropriée, de définir des classes d'énoncés distincts ayant tous la même valeur ; de caractériser ensuite chaque classe par un énoncé distingué ; de décomposer enfin cet énoncé distingué en segments élémentaires pouvant être considérés comme des "atomes" syntaxiques (ou sémanti-

ques). Appliquée à l'ensemble des énoncés à valeur de question, une telle démarche nous permettrait d'atteindre l'objectif fixé au début de ce chapitre, et de fonder une analyse sémantique ultérieure de l'ensemble des formes linguistiques utilisées pour l'interrogation. Mais auparavant, il nous faut ré-examiner brièvement les avantages et les inconvénients respectifs (de ce point de vue) des trois conceptions de la notion de transformation que nous avons présentées.

La principale critique formulée à l'égard de la notion de *translation* selon TESNIÈRE est qu'elle est fondée sur une réduction de la syntaxe à la morphologie, c'est-à-dire sur une confusion entre les *fonctions* et les *mots* qui les remplissent ordinairement [Ruwet 1967 : 228-229]. Une illustration plaisante en est donnée (involontairement) par TESNIÈRE lui-même, avec son exemple fameux : "le signal vert indique la voie libre"/"le silence vertébral indispose la voile licite" [Tesnière 1959 : 41-42] ; RUWET fait remarquer que, contrairement à ce qu'en dit TESNIÈRE, ces deux phrases n'ont pas la même structure syntaxique, puisque "libre" est un *attribut* de "voie", alors que "licite" ne peut être considéré que comme *épithète* de "voile" [Ruwet 1967 : 372]. Cette critique il est vrai nous paraît concerner essentiellement la translation du premier degré, qui rend compte de la structure des phrases simples, et ne pas diminuer sensiblement la valeur de la translation du second degré pour expliquer la structure des phrases complexes. En outre, la conception de TESNIÈRE présente l'avantage (pour notre recherche d'équivalences syntaxiques et sémantiques entre interrogations) de faire opérer la translation secondaire sur des énoncés réels, et non sur des structures abstraites, et de ne pas négliger la relation liant les questions à leurs réponses. Il faut simplement regretter que TESNIÈRE ait été un chercheur isolé sans grande postérité, car on trouve également en germe dans son oeuvre des idées que nous retrouvons chez HARRIS ou chez CHOMSKY ; par exemple l'énonciative ("Alfred est venu") comme source possible de l'interrogation connexionnelle ("est-ce qu'Alfred est venu ?") [Tesnière 1959 : 615-616] ; ou encore la disjonction ("est-ce Alfred, ou Albert, ou Antoine qui parle ?") comme source possible de l'interrogation nucléaire ("qui parle ?") [Tesnière 1959 : 193 § 7].

Chez CHOMSKY, les transformations ne s'appliquent qu'à des structures profondes, et elles ne sont pas réversibles ; ces propriétés peuvent apparaître comme un obstacle à l'établissement de relations d'équivalence entre énoncés. En pratique, rien n'interdit de considérer comme équivalents deux énoncés réalisant la même structure profonde. Par exemple : "does John have a chance to live ?" est équivalent à : "has John a chance to live ?" [Chomsky 1957 : 73] ; de même : "Pierre viendra-t-il ?" équivaut à : "Pierre viendra ?" [Dubois 1970 : 210-211]. Mais les relations d'équivalence possibles sont très limitées, puisque ni : "dites-moi si Pierre viendra", ni : "je vous demande si Pierre viendra", ni même : "est-ce que Pierre viendra ?", ne sont équivalents aux énoncés de l'exemple précédent, alors que tous ont pourtant la même signification. A cet égard, les grammaires génératives ne se distinguent guère des grammaires traditionnelles, qui classeraient ces énoncés selon la modalité de la phrase principale plutôt que d'après le contenu de l'élément porteur d'interrogation (pourtant le plus important au niveau sémantique). Toutefois, si l'on décompose la structure profonde des cinq énoncés sémantiquement équivalents, on constate qu'ils ont tous en commun un même élément, qui est la structure sous-jacente à la *phrase-noyau* : "Pierre viendra". Pour les grammaires génératives en effet, les phrases-noyaux (*kernel sentences*) sont celles que l'on obtient par application des seules transformations (ou règles de ré-écriture) obligatoires [Chomsky 1957 : 51], ce qui exclut par exemple les transformations passive, négative, d'enclassement, etc. Par conséquent, les phrases noyaux sont des phrases simples, déclaratives, actives [Chomsky 1957 : 88], obtenues à partir de l'élément P (dans la symbolique de la syntaxe de DUBOIS et DUBOIS-CHARLIER) de l'indicateur syntagmatique. S'il est exact que la décomposition des indicateurs syntagmatiques des énoncés de même signification permet d'identifier l'élément porteur du sens identique, il est également vrai que des énoncés ayant des indicateurs syntagmatiques différents ne peuvent en aucun cas être considérés comme syntaxiquement équivalents. Ce dernier point nous semble être un obstacle très important à l'utilisation de la notion de transformation selon CHOMSKY pour réduire à une forme canonique un ensemble d'énoncés à valeur de question sémantiquement équivalents.

Chez HARRIS, la relation de transformation est symétrique, transitive, et réflexive (si l'on admet la transformation identique), et elle s'applique à des

phrases du discours ; elle conduit par conséquent à définir des classes d'équivalence, dont les éléments sont les uns par rapport aux autres des "transformées synonymes" (ou des "paraphrases transformationnelles") [Harris 1976 : 15]. Appliquée à un corpus de même nature que le nôtre, mais de taille plus importante et de contenu plus homogène (cf. les remarques du § 1.1.5 et du § 1.2.3.e), la procédure inductive préconisée par HARRIS devrait normalement conduire à ranger dans une même classe d'équivalence la totalité des énoncés ayant pour fonction de demander "la même chose" à l'interlocuteur (cf. les cent mille manières de demander son nom à quelqu'un) ; nous définirons plus loin ce qu'il convient d'entendre par : "la même chose" (§ 1.2.4.e). Mais, nous l'avons vu, les transformations sont également (comme chez TESNIÈRE ou CHOMSKY) des *opérations* sur les séquences de mots ; la différence essentielle est que, chez HARRIS, ces opérations sont réversibles (Il est vrai que CHOMSKY présente l'irréversibilité comme une commodité technique plus que comme une nécessité théorique. cf. [Chomsky 1957 : 88]). Dans les publications antérieures aux *Notes du cours de syntaxe*, HARRIS posait, comme condition à la réversibilité de la transformation, que les éléments des phrases (noms, verbes, marques du temps, etc.) soient identiques dans toutes les phrases appartenant à la même classe de transformation [Harris 1957 : 394-397]. En conséquence, les deux énoncés : "j'ai acheté un livre, un livre vient de paraître", et : "j'ai acheté un livre qui vient de paraître" ne sont pas équivalents ; le passage du premier au second impose l'utilisation d'une transformation incrémentielle (non paraphrastique) ϕ_c , qui *ajoute* au premier énoncé le connecteur "qui" [Harris 1968 : 86-87]. La modification apportée sur ce point particulier, par les *Notes du cours de syntaxe*, à la théorie de la transformation selon HARRIS, est le remplacement de la notion de transformation incrémentielle par celle d'opérateur de méta-discours. Sur l'exemple que nous avons utilisé, cela revient à considérer que l'énoncé équivalent au second énoncé est : "j'ai acheté un livre contient un argument 2 identique à l'argument 1 de un livre vient de paraître" (cf. [Harris 1976 : 87-91] et le § 1.2.3.c, ci-dessus). La phrase contient donc deux discours élémentaires connectés par un opérateur de méta-discours.

Avant de revenir sur ce problème, notons que seule la conception de HARRIS peut conduire, à partir de la notion de transformation, à définir des relations d'équivalence entre un grand nombre d'énoncés de même sens. Ceci est

possible parce que HARRIS, contrairement à CHOMSKY, considère la transformation comme une *relation* entre énoncés ; et que, contrairement à TESNIÈRE et CHOMSKY, il pose que toutes les transformations (paraphrastiques) élémentaires sont *réversibles*. Encore faut-il nous assurer, avant d'adopter la problématique harrissienne, que cet avantage n'entraîne pas, en contrepartie, des propositions par trop irréalistes. Il faut bien admettre que la présence, dans le discours, d'opérateurs de méta-discours (comme dans le dernier exemple ci-dessus) n'est pas très naturelle. Certes, pour construire la proposition relative : "qui vient de paraître", il est nécessaire de postuler que le mot : "livre" a le même référent dans les deux discours élémentaires ; c'est ce qu'expriment la liaison anaphorique entre l'élément variable du pronom relatif et son antécédent chez TESNIÈRE [Tesnière 1959 : 561 § 13], et les conditions nécessaires et suffisantes à l'application de la transformation relative dans les grammaires génératives (cf. [Dubois 1970 : 249], [Chomsky 1965 : 189]). Mais chez TESNIÈRE ou CHOMSKY, l'identité du référent est en quelque sorte postulée extérieurement à l'énoncé lui-même ; cela signifie alors, comme le souligne HARRIS à propos des grammaires traditionnelles, que "l'information sur les identités, qui est cruciale pour la compréhension de la phrase, est connue du locuteur et du récepteur (comment ?) en dehors de la phrase à laquelle elle se réfère" [Harris 1976 : 85]. Or, il est clair que si l'on désire décomposer une phrase telle que : "j'ai acheté un livre qui vient de paraître", en phrases simples (TESNIÈRE), en phrases-noyaux (CHOMSKY), ou en discours élémentaires (HARRIS), *sans aucune perte d'information*, il faut postuler l'existence, dans la phrase elle-même, d'une indication sur la relation entre le relatif et son antécédent. La décomposition de la phrase doit par conséquent aboutir à trois éléments : "j'ai acheté un livre" + "un livre vient de paraître" + *identité des référents* (cf. p.ex. : [Delaunay 1974 : 61]. Cette analyse est à rapprocher de l'analyse par TESNIÈRE de la phrase simple : "Alfred parle", en trois éléments : "Alfred" + "parle" + *connexion*. cf. [Tesnière 1959 : 11-12], et § 1.2.1.a ci-dessus).

Malgré l'intérêt des notions de translation ou de transformation de type Chomskyen, il nous semble donc, à l'issue de ce premier bilan, que nous devons opter de préférence pour la conception harrissienne des transformations si nous voulons parvenir à une théorie opératoire de l'interrogation. Cette conception cumule en effet les principaux "avantages" (dans la perspective qui est

la nôtre) des deux autres conceptions : explicitation du lien entre questions et réponses, décomposition des phrases complexes en éléments simples, relations d'équivalence syntaxique (et sémantique) entre énoncés de formes différentes. Cependant, nous allons voir que HARRIS à lui seul ne résoud pas tous les problèmes posés par la réduction à une forme canonique des énoncés à valeur de question.

Pour qu'il soit possible de ramener à une forme unique l'ensemble des formes d'énoncé à valeur de question, en s'appuyant sur la transformation harriessienne et ses corrélats, il faut, sinon découvrir la solution des problèmes auxquels HARRIS n'apporte pas de réponse complète ou explicite, du moins s'assurer que ces problèmes seraient toutefois justiciables d'une solution qui ne soit pas en contradiction avec les principes des grammaires transformationnelles. Il est utile en outre de revenir au préalable sur deux points déjà abordés précédemment, mais non complètement traités, celui de la portée de l'interrogation, et celui des relations entre questions et réponses.

b) Le problème de la portée de l'interrogation

Le problème de la portée de l'interrogation pourrait s'énoncer ainsi : y-a-t-il deux types d'interrogation irréductibles l'une à l'autre, l'interrogation totale (ou : générale) et l'interrogation partielle (ou : particulière), ou bien s'agit-il d'un seul et même type ? Nous avons vu que, pour TESNIÈRE comme pour DUBOIS et DUBOIS-CHARLIER, il s'agit de deux types absolument différents ; mais nous avons également souligné (§ 1.2.1.e) les difficultés que l'on éprouve à vouloir marquer la frontière entre ces deux types dans la grammaire structurale de TESNIÈRE. Poussant plus loin l'analyse, Ryszard ZUBER montre qu' "il n'y a pas une différence catégorielle entre ces deux types de questions", car "la question dite générale n'est pas du tout générale en fait. Elle ne peut pas porter sur toute la proposition et, pour reprendre la terminologie des Grecs, elle n'est jamais « d'ignorance totale » mais, comme la question particulière, « d'ignorance partielle » " [Zuber 1972 : 131]. Sans en reprendre la démonstration (qui s'appuie sur des considérations logiques que nous n'aborderons qu'au chapitre 2), on peut illustrer la thèse de ZUBER à partir d'un exemple simple. Soit la question générale : "est-ce qu'il a plu hier ?" ; parmi les réponses possibles à cette question, il y a : "non, il a plu avant-hier", et :

"non, il a neigé hier". On peut donc, selon le contexte, la comprendre soit comme : "est-ce hier qu'il a plu ?", soit comme : "est-ce que ce qui s'est passé hier c'est de pleuvoir ?" ; or, à chacune de ces interprétations peut être associée une question particulière : "quand a-t-il plu ?", et : "que s'est-il passé hier ?". Remarquons d'ailleurs que l'intonation peut suffire à marquer l'interprétation souhaitée : "est-ce qu'il a plu *hier* ?"/"est-ce qu'il a *plu* hier ?" [Zuber 1972 : 132-133] (Ce dernier point a été très souvent souligné ; par exemple pour l'anglais par CHOMSKY, notamment pour introduire la notion de *focus* [Chomsky 1972 : 38-55], et pour le russe par PANFILOV, pour illustrer les rapports entre intonation et ordre des mots [Panfilov 1963 : 27]).

TESNIÈRE avait été conduit, pour résoudre le problème ainsi posé, à distinguer un cas particulier d'interrogation connexionnelle, celui où l'interrogation porte sur un pôle de la connexion [Tesnière 1959 : 209 § 14] ; il aurait pu également considérer le cas où l'interrogation ne porte que sur une connexion, ou une partie de l'ensemble des connexions de la phrase. La divergence entre TESNIÈRE et ZUBER tient au critère de classification utilisé : mécanisme de la réponse chez ZUBER, présence d'un contenu sémantique dans les nucléus chez TESNIÈRE. Mais nous verrons que l'analyse du mécanisme de la réponse à l'interrogation connexionnelle, seulement esquissée par TESNIÈRE, aurait pu le conduire à dissocier l'antinomie "nucléaire/connexionnelle" de l'opposition "générale/particulière". Par contre, si pour DUBOIS et DUBOIS-CHARLIER cette opposition se retrouve dans la structure profonde de la phrase, et apparaît donc comme irréductible (présence ou absence du SN_{que} dans le constituant interrogatif), CHOMSKY avait proposé antérieurement une analyse de l'interrogation dans laquelle les interrogatives particulières sont dérivées d'interrogatives générales [Chomsky 1957 : 67-78]. Ainsi, la phrase déclarative : "John ate an apple", obtenue par application des transformations obligatoires de la grammaire, donne, si l'on applique la transformation facultative T_{int} : "did John eat an apple ?" ; cette dernière phrase permet d'obtenir, par application de la transformation facultative T_{w} : "what did John eat ?", et : "who ate an apple ?". Mais, dans des publications ultérieures, CHOMSKY paraît avoir avalisé des travaux comme ceux de KATZ et POSTAL, qui ont postulé l'existence, dans la structure profonde des phrases interrogatives, d'un marqueur de l'interrogation Q et d'un ou plusieurs indicateurs du champ de l'interrogation wh (qui, lorsqu'ils sont éléments de syntagmes nominaux, caractérisent l'interrogation particulière)

[Katz 1964 : 125-188], éléments dont la présence déclenche des transformations obligatoires [Chomsky 1965 : 180-182]. Cette conception est celle dont se sont entre autres inspirés DUBOIS et DUBOIS-CHARLIER [Dubois 1970] et Gérard GENOT [Genot 1973]. Quant à HARRIS, il élude en fait le problème, puisqu'il part de l'ensemble des réponses possibles. Dans l'exemple donné par ZUBER, on aurait, selon le contexte, ou bien : "je vous demande s'il a plu hier ou s'il a plu avant-hier ou ... ou s'il a plu samedi dernier" ; ou bien : "je vous demande s'il a plu hier ou s'il a grêlé hier ou ... ou s'il a neigé hier" ; ou bien : "je vous demande s'il a plu hier ou s'il n'a pas plu hier" ; etc. En procédant ainsi, HARRIS fait implicitement l'hypothèse que l'ensemble des réponses possibles dépend du discours en question (cf. [Harris 1968 : 94-95]), c'est-à-dire du contexte (ce qui pose d'autres problèmes, comme nous le verrons).

c) Le problème des relations entre questions et réponses

Les relations entre questions et réponses peuvent être mises en évidence par l'étude du mécanisme de réponse aux questions. Pour le français, la variété des formes de réponses a été bien décrite par DAMOURETTE et PICHON ; ceux-ci distinguent trois grandes catégories de réponses [Damourette 1931 : IV 306-307, 346-349] :

- les réponses *entières*, très rares, qui sont des phrases énonciatives complètes, ayant un sens par elles-mêmes, indépendamment de la question qui les a suscitées (et dont elles reprennent d'ailleurs les termes). Exemples : "Et tu refuses de l'épouser ? - Et je refuse de l'épouser" ; "Qu'y-a-t-il eu de décidé ? - Il y a eu de décidé que tout serait préparé".

- les réponses *anaphorisées*, dans lesquelles certains termes de la question sont remplacés par des éléments anaphoriques (pronoms personnels ou adverbiaux). Exemples : "Jean est arrivé ? - Il est arrivé" ; "combien veux-tu de morceaux de sucre ? - J'en veux trois".

- les réponses *zeugmatiques*, les plus fréquentes dans la littérature. Un *zeugme* est "un membre de phrase où sont impliquées logiquement des idées explicitées dans un autre membre de phrase" [Damourette 1931 : IV 276]. Dans les réponses zeugmatiques, les idées impliquées logiquement se trouvent explicitées dans l'énoncé interrogatif qui les a suscitées. Exemples : "Nous parti-

rons ? - Quand tu voudras" ; "Comment t'appelles-tu ? - Jacques".

Cette classification purement syntaxique est intéressante à plus d'un égard. Tout d'abord, elle illustre parfaitement la relation très forte qui lie questions et réponses. Ensuite, elle est absolument indépendante de la portée de l'interrogation (nous avons à dessein choisi pour chaque catégorie de réponses un exemple d'interrogation générale et un exemple d'interrogation particulière). Enfin, exprimée en termes de grammaire transformationnelle, elle permet de montrer l'équivalence de ces trois formes : la réponse anaphorisée étant obtenue à partir de la réponse entière par *pronominalisation* de mots présents à la fois dans la question et la réponse entière ; et la réponse zeugmatique provenant de l'*effacement*, dans la réponse entière, de certains mots répétés dans la réponse entière. On peut par conséquent considérer la réponse entière comme la forme distinguée représentant un ensemble de réponses équivalentes, et analyser le mécanisme de la réponse à partir de la comparaison de l'énoncé interrogatif et des réponses entières qui lui correspondent. Notons que HARRIS a signalé la nature transformationnelle du lien unissant les réponses anaphorisées et les réponses zeugmatiques à l'énoncé interrogatif qui les précède dans le dialogue, comme d'ailleurs de celui qui existe entre les questions elliptiques et l'assertion à laquelle elles succèdent [Harris 1957 : 425-426] ; mais il ne semble pas en avoir donné une analyse détaillée pour l'anglais (Pour l'allemand, on trouve un exemple de réécriture d'une réponse anaphorisée sous sa forme entière dans : [Philipp 1974 : 105]).

Si l'on prend comme point de départ de l'analyse les réponses entières et elles seules, il est assez facile de construire une typologie des couples question-réponses. Soit par exemple l'assertion simple : "Je vais à Paris". En s'inspirant des énoncés interrogatifs cités dans les grammaires françaises que nous avons consultées, on peut établir une liste de questions usuelles auxquelles cette assertion constitue une réponse. On remarquera que, dans la majeure partie des cas, les différences dans la forme des questions dépend étroitement de l'ensemble des autres réponses possibles que l'on associe à l'énoncé choisi comme point de départ. Les principales formes de couples question-réponses que l'on peut ainsi déterminer empiriquement sont les suivantes (la raison des lacu-

nes dans leur numérotation apparaîtra plus loin) :

- 1 { Q 1 = "Allez-vous à Paris ?"
R 1 = "Je vais à Paris/Je ne vais pas à Paris"
- 2 { Q 2 = "Allez-vous à Paris ou n'allez-vous pas à Paris ?" (avec des variantes possibles comme Q'₂ = "Allez-vous à Paris, (oui) ou non ?")
R 2 = "Je vais à Paris/Je ne vais pas à Paris"
- 12 { Q 12 = "Allez-vous à Paris ou allez-vous à Bordeaux ?" (variante : Q'12 = "Allez-vous à Paris ou à Bordeaux ?")
R 12 = "Je vais à Paris/Je vais à Bordeaux"
- 15 { Q 15 = "Allez-vous à Paris ou restez-vous à la campagne ?"
R 15 = "Je vais à Paris/Je reste à la campagne"
- 19 { Q 19 = "Où allez-vous ?"
R 19 = "Je vais à Paris/Je vais à Bordeaux/Je vais à Marseille/.../Je vais à Lille/Je vais à Clermont-Ferrand".

Il est certain que ces exemples sont assez artificiels, et que l'on rencontrera rarement de tels dialogues en-dehors des cours élémentaires de français pour étrangers. Dans un dialogue réel, certaines redondances sont évidemment supprimées. Mais si l'on prend le parti de ne considérer que la forme entière des réponses, il est rationnel de s'en tenir au même parti pour les questions elles-mêmes ; les comparaisons en sont facilitées. D'autre part, nous n'avons pas pris en considération certaines formes de réponses possibles, comme : "je n'en sais rien", "cela ne vous regarde pas", etc. Des réponses de ce genre se rencontrent assurément dans la conversation, comme d'ailleurs dans les examens psychologiques ou les enquêtes psychosociales ; leur analyse sortirait du cadre purement syntaxique que nous nous sommes fixé dans ce chapitre, et poserait des problèmes logico-sémantiques que nous n'aborderons qu'au chapitre 2. Pour le moment, nous nous limitons aux couples *question entière - réponses entières*, dans lesquels par conséquent la réponse reprend les termes de l'interrogation qui l'a provoquée.

En appliquant à cet ensemble de couples question-réponses les classifications linguistiques déjà décrites (§ 1.1.2.b et c), on ne peut que constater les désaccords (voire les contradictions) entre les auteurs que nous avons cités. Il est toutefois possible de regrouper leurs classifications en trois familles :

1) Celles qui opposent le type Q 19 = "Où allez-vous ?" à tous les autres. C'est le cas pour DAMOURETTE et PICHON, qui appellent la Q 19 : *interrogation partielle*, et considèrent les phrases de même forme que tous nos autres exemples comme des *interrogations totales* (en donnant à ces mots un sens très différent de celui que leur donnent WAGNER et PINCHON par exemple). On s'en convaincra en examinant les interrogations totales suivantes, d'après DAMOURETTE et PICHON : "Tu vas à la banque, ou j'y vais ?" ; "C'est-il la logique et l'idéologie, ou bien c'est-il la comparaison historique qui nous mettent à même d'établir des rapprochements de ce genre ?" [Damourette 1931 : IV 329, 336]. On relève également cette opposition chez Ronald W. LANGACKER, pour qui les *specification questions* (dont la Q 19 fait partie) se caractérisent par un mot interrogatif indéterminé (du groupe des mots en *wh-*, liés au pronom-adjectif indéfini *some* : *who - someone, why - for some reason, etc.*), alors que les *alternative questions* se définissent par la présence de la conjonction *or*. Lorsque l'alternative ne porte que sur une phrase et sa négation (cette dernière pouvant être explicite, comme dans Q 2, ou implicite comme dans Q 1), on a la sous-classe des *yes-no questions* [Langacker 1972 : 188-189]. De même, Viviane ALLETON distingue en chinois l'*interrogation substitutive* (à laquelle correspond Q 19), dont la réponse est obtenue en substituant un mot plein au mot interrogatif indéfini, et l'*interrogation exclusive* (ex. : Q 1, Q 2, Q 12, Q 15), à laquelle on répond en sélectionnant un des membres de l'alternative proposée. Parmi les interrogations exclusives, on peut définir des sous-classes : l'*interrogation oui-non* (ex. : Q 1), et l'*interrogation par répétition* (ex. : Q 2), dans laquelle le verbe est répété sous sa forme négative [Alleton 1973 : 31-35].

2) Celles qui opposent les types Q 1 et Q 2 à tous les autres. Otto JESPERSEN range Q 1 et Q 2 parmi les *questions nexales* (*nexus-questions*), car l'interrogation porte sur la relation (*nexus*) entre le sujet et le prédicat de la phrase ; à l'autre pôle de sa classification, Q 19 constitue un bon exemple des *questions sur x* (*x-questions*), dans lesquelles l'interrogation vise à établir la nature d'une inconnue *x*. Les critères d'identification que propose JESPERSEN sont l'intonation, et le contenu de la réponse : "la réponse à une question nexale est soit *oui* soit *non* ; mais la réponse à une question sur *x* peut être à peu près n'importe quoi en fonction de la situation, sauf *oui* et *non*". Par conséquent, Q 12 et Q 15 sont des questions sur *x*, "bien qu'elles

ressemblent à des questions nexales d'un point de vue formel" [Jespersen 1924 : 432-433]. La catégorisation de Léonard BLOOMFIELD semble tout à fait parallèle à celle de JESPERSEN, à ceci près qu'il appelle le premier type : *question à laquelle on répond par oui ou par non*, et le second type : *question à compléter (supplément question)* [Bloomfield 1961 : 109]. Comme nous l'avons vu, Lucien TESNIÈRE oppose de même l'*interrogation connexionnelle* (ex. : Q 1 et Q 2), qui porte sur la connexion entre nucléus pleins, à l'*interrogation nucléaire* (ex. : Q 12, Q 19), où la réponse consiste à remplir le nucléus vide [Tesnière 1959 : 191-215]. Bien que le cas de la Q 15 ne soit pas traité explicitement, il est clair qu'il ne peut s'agir d'une interrogation connexionnelle.

3) Celles qui n'envisagent que les deux pôles Q 1 - Q 2 d'une part et Q 19 d'autre part, sans que la possibilité des Q 12 et Q 15 soit seulement mentionnée. Nous avons vu que c'est notamment le cas pour WAGNER et PINCHON, qui ne considèrent que l'*interrogation totale* (ex. : Q 1, Q 2), marquée par la position du pronom personnel sujet, l'intonation, ou la particule "est-ce que", et l'*interrogation partielle* (ex. : Q 19), introduite par un pronom, un adjectif, ou un adverbe interrogatif [Wagner 1962 : 525-538] ; aucun des exemples cités par ces auteurs n'est de la forme des Q 12 ou Q 15. C'est à partir des mêmes critères morpho-syntaxiques qu'André TELLIER distingue en anglais les questions à réponse *yes/no* et les questions de type *wh-* [Tellier 1974 : 106-110]. Nous avons vu également que la grammaire générative et transformationnelle du français de DUBOIS et DUBOIS-CHARLIER ne traite explicitement que des formes correspondant d'une part à Q 1 et Q 2 (*interrogation oui/non*), et d'autre part à Q 19 (*interrogation "que"*) [Dubois 1970 : 208-227].

d) Description formelle des grandes catégories de questions

Ce survol montre à l'évidence que l'analyse du mécanisme de l'interrogation et de la réponse n'a en général pas suffisamment retenu l'attention des linguistes. C'est pourquoi il n'est pas inutile de procéder plus systématiquement que nous ne l'avons fait jusqu'ici, et de passer en revue toutes les formes simples d'interrogation pouvant avoir la phrase : "Je vais à Paris" comme élé-

ment de l'ensemble des réponses possibles. En s'inspirant de la décomposition des phrases-noyaux, dans certains travaux de HARRIS, en nom ou pronom sujet N , verbe V (en négligeant la marque du temps), et objet (direct ou indirect) Ω [Harris 1964 : 484 ; 1968 : 76], on peut représenter la structure de la phrase : $S = \text{"Je vais à Paris"}$ par : $S = n_1 v_1 \omega_1$. Si l'on admet l'équivalence du "je" dans la réponse et du "vous" dans la question ("je_R" = "vous_Q" = n_1), car ils ont le même référent, et si l'on néglige l'ordre des mots, la question : $Q 1 = \text{"Allez-vous à Paris ?"}$ peut s'écrire :

$$Q 1 = n_1 v_1 \omega_1 ?$$

(l'ordre des symboles correspond par conséquent ici à l'ordre des mots dans l'interrogation marquée par la seule intonation). Pour rendre compte, selon la même symbolique, de la question $Q 2$, il faut introduire la *négation*, par exemple sous la forme d'un surlignement des éléments sur lesquels elle porte, et la *disjonction*, que l'on écrira : *ou*. Ainsi (en utilisant des parenthèses lorsque la disjonction opère sur des phrases entières de forme : $N V \Omega$, et non sur des mots isolés), on pourra écrire :

$$Q 2 = (n_1 v_1 \omega_1) \text{ ou } (\overline{n_1 v_1 \omega_1}) ?$$

Nous avons implicitement défini : $n_1 = \text{"je}_R = \text{"vous}_Q$, $v_1 = \text{"vais}_R = \text{"allez}_Q$ et $\omega_1 = \text{"à Paris"}$. Nous pouvons poser : $\omega_2 = \text{"à Bordeaux"}$, et écrire :

$$Q 12 = (n_1 v_1 \omega_1) \text{ ou } (n_1 v_1 \omega_2) ?$$

En posant par contre : $v_2 = \text{"reste}_R = \text{"restez}_Q$, et $\omega_2 = \text{"à la campagne"}$, on écrira :

$$Q 15 = (n_1 v_1 \omega_1) \text{ ou } (n_1 v_2 \omega_2) ?$$

Enfin, en adoptant comme symbole pour les mots interrogatifs indéfinis (qui ?, que ?, où ?, comment ?, etc.) le mot "que", nous avons :

$$Q 19 = n_1 v_1 \text{ que } ?$$

Cette écriture symbolique nous permet d'engendrer d'une manière systématique l'ensemble des formes syntaxiques concevables pour représenter des

questions dont l'une des réponses possibles est : "Je vais à Paris" = $n_1 v_1 \omega_1$. En nous bornant à la négation et la disjonction opérant sur des éléments N , V , et Ω (ou sur des phrases formées de ces éléments), et en limitant les effets de la disjonction à des alternatives binaires (comme : S ou \bar{S} , ω_1 ou ω_2), nous obtenons les 23 types de questions du tableau n° 1.65. Avant d'étudier sur cet ensemble les relations entre questions et réponses, il nous faut nous assurer que ces formes existent, qu'elles sont distinctes les unes des autres, et qu'elles sont les seules possibles dans le cadre que nous nous sommes fixé.

Le problème de l'existence de ces types de questions peut être résolu empiriquement en s'assurant qu'elles ont une traduction ayant un sens en langage naturel. Nous avons ainsi constaté que $Q 1$, $Q 2$, $Q 12$, $Q 15$ et $Q 19$ existent en ce sens, même si le langage courant préfère à la forme entière que nous avons retenue une forme de question équivalente, mais moins redondante. Mais comment traduire $Q 3$ en français ? Il est clair que : * "Allez-vous à Paris ou allez-pas-vous à Paris ?" est inacceptable. Pourtant, l'ensemble $R3$ qui lui est associé pourrait être : "Je vais à Paris/Ce n'est pas moi qui vais à Paris", en utilisant la périphrase emphatique "c'est ... qui". On peut alors se demander si l'interrogation : "Est-ce vous, ou non, qui allez à Paris ?" n'est pas la réalisation en français de la formule $Q 3$ (\bar{n}_1 pouvant d'ailleurs aussi se traduire par : "quelqu'un d'autre", plus acceptable du point de vue de l'usage de la langue) ; d'autant que ce serait la seule traduction possible d'une phrase comme : "you are going to Paris ?" (sous-entendu : "or is it not you ?", à la condition toutefois qu'il s'agisse d'une vraie question, et non d'une interrogation rhétorique exprimant l'étonnement). La difficulté est du même ordre pour $Q 4$; alors que l'on pourrait demander : "Are you flying to Paris (or aren't you flying) ?", ou : "Летите ли вы в Париж (или не летите)?", ou encore : "Fliegen Sie nach Paris (oder fliegen Sie nich) ?", le français imposerait : "Vous allez à Paris en avion, ou autrement ?", phrase dont la structure n'est plus morphologiquement celle de $Q 4$. Les difficultés que nous rencontrons ici sont donc d'ordre lexicographique plus que d'ordre syntaxique. Par contre, pour $Q 9$, la difficulté paraît insurmontable dans toutes les langues, si l'on désire que la phrase ait un sens ; à moins que $Q 9$ soit une autre écriture pour $Q 2$, ce que nous examinerons plus loin. Sur le

		<u>N</u>	<u>V</u>	<u>Ω</u>		<u>N'</u>	<u>V'</u>	<u>Ω'</u>
X	{	Q1	=	n_1	v_1	ω_1	?	
		Q2	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(\overline{n_1} \ \overline{v_1} \ \overline{\omega_1})$?	
B	{	Q3	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(\overline{n_1} \ v_1 \ \omega_1)$?	
		Q4	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ \overline{v_1} \ \omega_1)$?	
		Q5	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ v_1 \ \overline{\omega_1})$?	
		Q6	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(\overline{n_1} \ \overline{v_1} \ \omega_1)$?	
		Q7	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(\overline{n_1} \ v_1 \ \overline{\omega_1})$?	
		Q8	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ \overline{v_1} \ \overline{\omega_1})$?	
		Q9	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(\overline{n_1} \ \overline{v_1} \ \overline{\omega_1})$?	
A	{	Q10	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_2 \ v_1 \ \omega_1)$?	
		Q11	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ v_2 \ \omega_1)$?	
		Q12	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ v_1 \ \omega_2)$?	
		Q13	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_2 \ v_2 \ \omega_1)$?	
		Q14	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_2 \ v_1 \ \omega_2)$?	
		Q15	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_1 \ v_2 \ \omega_2)$?	
		Q16	=	$(n_1 \ v_1 \ \omega_1)$	ou	$(n_2 \ v_2 \ \omega_2)$?	
C	{	Q17	=	que	v_1	ω_1	?	
		Q18	=	n_1	que	ω_1	?	
		Q19	=	n_1	v_1	que	?	
		Q20	=	que	que	ω_1	?	
		Q21	=	que	v_1	que	?	
		Q22	=	n_1	que	que	?	
		Q23	=	que	que	que	?	

1.65 : Questions simples pouvant avoir une réponse entière de la forme : N V Ω

modèle de Q 12, il est facile d'imaginer une traduction des types Q 10 à Q 16; Nous avons déjà vu l'exemple cité par DAMOURETTE et PICHON : "Tu vas à la banque, ou j'y vais ?", qui est une réalisation (sur un autre thème) de Q 10. Pour Q 11, on aurait : "Are you flying to Paris, or driving ?", dont la traduction nécessite en français une autre structure syntaxique ("Vous allez à Paris en avion, ou en voiture ?"). Remarquons d'ailleurs que, pour Q 11, nous ne sommes pas tout à fait obligé de recourir aux langues étrangères comme pour Q 4, qui lui est pourtant apparenté ; on peut avoir en français : "Vous allez à Paris, ou vous en revenez ?" (avec, il est vrai, un changement de préposition dans Ω ...). Quant à Q 16, qui est la forme la plus complexe de la série, elle peut être réalisée par exemple en : "Allez-vous à Paris, ou bien votre épouse prend-elle des vacances ?". Les types Q 17 à Q 19 ne posent également aucun problème de réalisation en français. Les types Q 20 à Q 22 soulèvent celui de ce que TESNIÈRE a appelé l'*interrogation binucléaire*. Alors que l'on trouve dans de nombreuses langues des questions équivalentes à Q 21, de la forme : "Quis quem verberat ?", "τίς τίνα τύπτει ; ", "Who strikes whom ?" , "кто кого бьёт ?" , etc., le français est obligé de diviser la phrase en deux interrogations conjointes : "Quel est celui qui frappe et qui frappe-t-il ?" [Tesnière 1959 : 201-203]. Or, il semble que, sous l'influence de la diffusion de l'anglais, la langue française commence à utiliser des tournures telles que : "Qui est qui ?", voire : Q 21 = "Qui va où ?". Le type Q 20 = "Qui fait quoi à Paris ?" est déjà plus douteux ; quant à Q 22 = "Vous faites quoi où ?", une telle phrase est à la limite de l'acceptabilité en français (faute d'une forme verbale indéterminée plus maniable que "faire" cf. [Ruwet 1967 : 230]). Enfin, le type Q 23 semble n'être traduisible dans aucune langue (cf. [Tesnière 1959 : 203 § 12]). On peut en conséquence douter de son existence sur le plan linguistique (comme d'ailleurs de sa signification éventuelle) ; à moins que cette forme ne rende compte de l'interrogation du type : "Comment ?", "Plaît-il ?", ou : "Quoi ?", qui porte sur l'ensemble de la phrase qui précède dans le dialogue, quelle que soit celle-ci.

Le problème de la non-identité des types de questions se pose surtout pour Q 1 et Q 2. LANGACKER opte pour l'identité, en considérant Q 1 comme dérivée de Q 2 : "Is it raining, or isn't it raining ?" → "Is it raining, or

not ?" → "Is it raining ?" [Langacker 1972 : 188]. Toutefois, si l'on part non plus de l'énoncé interrogatif, mais de l'ensemble des réponses possibles, on peut supposer que : R2 = "Je vais à Paris/Je ne vais pas à Paris" = "oui/non" ; alors que R1 serait beaucoup plus ouvert : "Je vais à Paris/Je vais à Bordeaux/C'est ma femme qui y va/J'en reviens/.../Je ne vais pas à Paris" (Nous reviendrons d'ailleurs sur cette analyse à la fin de ce paragraphe, pour conclure que Q 1 est une forme atypique). Par contre, le problème des ressemblances éventuelles entre Q 2 et Q 9 (si ce dernier type existe) est plus délicat à résoudre ; sa solution dépend des rapports syntaxiques qui lient la négation portant sur les éléments de la phrase, et la négation portant sur la phrase entière. Un moyen de le résoudre serait peut-être d'analyser les analogies entre la concaténation de N, V, et Ω , et un connecteur logique (comme la conjonction). En l'absence de données précises dans ce domaine, il nous semble prématuré de trancher sur ce point, qui d'ailleurs n'est pas crucial pour notre recherche. Nous pouvons donc admettre qu'à l'exception de Q 9, dont l'existence et la spécificité ne sont pas démontrées, tous les types de Q 1 à Q 22 existent dans au moins certaines des langues naturelles, et ont une identité propre.

Il reste à voir si notre liste est exhaustive. Etant données les règles que nous avons suivies pour engendrer cette série, il n'est guère possible que de créer des types hybrides, comme par exemple :

$$n_1 v_1 \text{ que : } (n_1 v_1 \omega_1) \text{ ou } (n_1 v_1 \omega_2) ?$$

réalisable en : "Où allez-vous : vous allez à Paris ou vous allez à Bordeaux ?". On voit que ce nouveau type est obtenu par concaténation de Q 19 et de Q 12, et que l'ensemble des réponses possibles à cette interrogation est R 12. L'adjonction d'un énoncé de la forme de Q 19 n'ajoute par conséquent rien à Q 12, et les questions : "Où allez-vous : à Paris ou à Bordeaux ?", et : "Vous allez à Paris ou à Bordeaux ?" sont équivalentes. On vérifierait aisément qu'il en est de même pour les formes hybrides que l'on peut former en concaténant un énoncé de la série Q 17 à Q 22, avec l'énoncé correspondant de la série Q 10 à Q 15. Par ailleurs, toute autre combinaison de deux (ou plus) des types déjà constitués ne paraît pas avoir de sens.

L'analyse des relations entre une question et l'ensemble des réponses possibles nous ramène pour l'essentiel aux remarques de LANGACKER et d'ALLETON. On peut en effet classer la plupart des types que nous avons engendrés en deux grandes catégories ; ceux pour lesquels la réponse consiste à répéter un des membres de l'alternative, et ceux pour lesquels la réponse impose de remplacer les mots interrogatifs par des mots pleins. Cette dernière catégorie comprend toutes les formes de Q 17 à Q 22. La première catégorie se subdivise en alternative portant sur des phrases-noyaux complètes (Q 2), alternatives portant sur un seul élément de phrase (Q 3, Q 4, Q 5, Q 10, Q 11, Q 12), alternatives portant sur deux éléments (Q 6, Q 7, Q 8, Q 13, Q 14, Q 15), et alternatives portant sur plus de deux éléments (Q 16). Il semble qu'il soit nécessaire de créer une troisième catégorie spécialement pour le type Q 1, qui ne propose pas d'alternative, et qui ne contient pas de mot interrogatif indéfini. Nous avons en effet observé que l'ensemble des réponses possibles R1 peut être beaucoup plus étendu que R2 (qu'il contient). En première approximation, il semble que les réponses possibles pour l'ensemble des types de Q 17 à Q 22 puissent également servir de réponses à Q 1.

Une autre approche, peut-être plus intéressante, consiste à adopter une optique ensembliste relativement aux réponses possibles. Commençons par examiner les interrogations qui ne portent que sur un élément de la phrase ; par exemple Ω . Considérons pour simplifier dans les réponses R_i le contenu R'_i portant sur Ω :

$$R' 5 = \omega_1 / \bar{\omega}_1$$

$$R' 12 = \omega_1 / \omega_2$$

$$R' 19 = \omega_1 / \omega_2 / \omega_3 / \dots / \omega_i / \dots / \omega_n$$

Appelons \mathcal{U}_i l'univers sur lequel porte la question Q_i . Si l'on définit Ω_ℓ comme l'ensemble des compléments d'objets possibles $\omega_i \in \Omega_\ell$, de la phrase $S = n_1 v_1 \Omega = \text{"Je vais à ..."}$, dans un contexte donné L , nous avons :

$$\mathcal{U}_5 = \mathcal{U}_{12} = \mathcal{U}_{19} = \Omega_\ell$$

Traduisons l'ensemble des réponses à chaque question en termes d'ensembles de parties de \mathcal{U}_i :

$$R' 5 = \{\{\omega_1\}, \{\omega_2, \omega_3, \dots, \omega_i, \dots, \omega_n\}\}$$

$$R' 12 = \{\{\omega_1\}, \{\omega_2\}\}$$

$$R' 19 = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \dots, \{\omega_i\}, \dots, \{\omega_n\}\}$$

Nous voyons immédiatement que poser la question Q 12 consiste à effectuer une sélection préalable d'un sous-ensemble de Ω_ℓ , $U_{12} = \{\omega_1, \omega_2\}$, au sein duquel l'interlocuteur doit choisir le terme qui lui convient. Alors qu'en posant Q 5 ou Q 19, le locuteur laisse à la personne interrogée la possibilité de répondre, quel que soit l'endroit où elle se rend. Si par exemple l'interlocuteur se rend à Lille, et qu'on lui demande (Q 12) : "Allez-vous à Paris ou à Bordeaux ?", *stricto sensu* l'interlocuteur ne peut pas répondre, sous peine de se placer hors des conditions de l'interrogation. Par contre, il peut répondre à Q 19 : "Je vais à Lille", et à Q 5 : "Non, ce n'est pas à Paris que je vais".

Il est clair toutefois que le locuteur qui pose Q 5 exprime par ce moyen une attente différente de celle qu'il aurait eue s'il avait posé Q 19. En demandant : "Allez-vous à Paris ?", plutôt que : "Où allez-vous ?", il privilégie dans Ω_ℓ l'élément $\omega_1 = \text{"à Paris"}$; soit qu'il pense que cette réponse a plus de chances de lui être donnée que n'importe quelle autre, soit qu'il ait son intérêt particulièrement éveillé si, et seulement si, la personne interrogée reconnaît aller à Paris (nous écartons volontairement ici l'interrogation rhétorique, encore que l'on puisse probablement la réduire à ce deuxième cas). Appelons *énumération* $K_i \in \mathcal{P}(\mathcal{U}_i)$ l'ensemble des éléments ω_i explicitement cités dans l'énoncé de la question Q_i . Nous avons :

$$K 5 = \{\omega_1\}$$

$$K 12 = \{\omega_1, \omega_2\}$$

$$K 19 = \emptyset$$

Appelons d'autre part *sélection* $U_i \in \mathcal{P}(\mathcal{U}_i)$ l'ensemble des éléments ω_i préalablement sélectionnés par le fait de poser la question Q_i .

Nous avons :

$$U_5 = \{\omega_1\} \cup \{\omega_2, \omega_3, \dots, \omega_i, \dots, \omega_n\} = \mathcal{U}_5 = \Omega_\ell$$

$$U_{12} = \{\omega_1\} \cup \{\omega_2\} = \{\omega_1, \omega_2\}$$

$$U_{19} = \{\omega_1\} \cup \{\omega_2\} \cup \{\omega_3\} \cup \dots \cup \{\omega_i\} \cup \dots \cup \{\omega_n\} = \mathcal{U}_{19} = \Omega_\ell$$

Si nous considérons que ce que nous avons dit pour Ω peut l'être également pour N ou pour V , nous pouvons caractériser les trois types de questions ne portant que sur un seul élément de la phrase en déterminant l'étendue de la sélection préalable qu'elles opèrent, et celle de l'énumération qui est faite dans l'énoncé de la question.

Examinons maintenant les questions portant sur deux éléments de la phrase. Afin de pouvoir disposer dans tous les cas d'une traduction en français (au moins approximative), considérons les questions qui portent à la fois sur N et sur Ω . Ce sont : $Q_7 =$ "Vous allez à Paris, ou c'est quelqu'un d'autre qui va ailleurs ?" ; $Q_{14} =$ "Vous allez à Paris, ou c'est votre ami qui va à Bordeaux ?" ; et $Q_{21} =$ "Qui va où ?". Appelons comme précédemment Ω_ℓ l'ensemble des objets possibles $\omega_i \in \Omega_\ell$ de la phrase $S = n_1 v_1 \Omega =$ "Je vais à ...", et N_ℓ l'ensemble des sujets possibles $n_j \in N_\ell$ de la phrase $S' = N v_1 \omega_1 =$ "... va à Paris", dans un contexte donné L identique pour S et S' .

L'univers sur lequel porte l'interrogation est le même pour les trois questions:

$$\mathcal{U}_7 = \mathcal{U}_{14} = \mathcal{U}_{21} = N_\ell \times \Omega_\ell$$

La sélection et l'énumération pour chacune des trois questions sont donc des parties du produit cartésien de l'ensemble N_ℓ des sujets possibles et de l'ensemble Ω_ℓ des objets possibles. L'énumération $K_i \in \mathcal{P}(N_\ell \times \Omega_\ell)$ est aisée à définir :

$$K_7 = \{(n_1, \omega_1)\}$$

$$K_{14} = \{(n_1, \omega_1), (n_2, \omega_2)\}$$

$$K_{21} = \emptyset$$

Pour la sélection par contre, il nous faut commencer par analyser les contenus R^i des réponses possibles à chacune des questions. La question Q 7 donne le choix entre : "Je vais à Paris", et : "C'est quelqu'un d'autre qui va ailleurs". On ne peut donc répondre ni : $\bar{n}_1 v_1 \omega_1 =$ "Ce n'est pas moi qui vais à Paris", ni : $n_1 v_1 \bar{\omega}_1 =$ "Ce n'est pas à Paris que je vais". Sont donc écartés de la sélection tous les couples de la forme : (n_i, ω_1) , et : (n_1, ω_i) , avec $i \neq 1$. L'ensemble des réponses à Q 7 est un ensemble à deux éléments non complémentaires dans \mathcal{U}_7 :

$$R^7 = \{ \{ (n_1, \omega_1) \}, \{ (n_2, \omega_2), (n_2, \omega_3), \dots, (n_2, \omega_n), (n_3, \omega_2), (n_3, \omega_3), \dots, (n_3, \omega_n), \dots, (n_m, \omega_2), (n_m, \omega_3), \dots, (n_m, \omega_n) \} \}$$

En procédant de même pour Q14, on remarque que, s'il est vrai que cette question sélectionne deux éléments de N_ℓ et deux éléments de Ω_ℓ , on ne peut avoir dans les réponses possibles les quatre couples d'éléments correspondant au produit cartésien de ces sous-ensembles. En effet, on ne peut répondre ni : $n_2 v_1 \omega_1 =$ "C'est mon ami qui va à Paris", ni : $n_1 v_1 \omega_2 =$ "C'est à Bordeaux que je vais". De telles réponses, valables respectivement pour Q 10 et pour Q 12, se situent ici hors des conditions de l'interrogation. Nous avons par conséquent :

$$R^{14} = \{ \{ (n_1, \omega_1) \}, \{ (n_2, \omega_2) \} \}$$

Enfin, Q 21 accepte comme réponses tous les couples (n_i, ω_j) possibles :

$$R^{21} = \{ \{ (n_1, \omega_1) \}, \{ (n_1, \omega_2) \}, \{ (n_1, \omega_3) \}, \dots, \{ (n_1, \omega_n) \}, \\ \{ (n_2, \omega_1) \}, \{ (n_2, \omega_2) \}, \dots, \{ (n_2, \omega_n) \}, \dots, \{ (n_m, \omega_1) \}, \\ \{ (n_m, \omega_2) \}, \{ (n_m, \omega_n) \} \}$$

Nous pouvons sur ces bases déterminer la sélection opérée par ces trois types de questions :

$$U_7 = \{\{n_1\} \times \{\omega_1\}\} \cup \{\{\bar{n}_1\} \times \{\bar{\omega}_1\}\}$$

$$U_{14} = \{(n_1, \omega_1)\} \cup \{(n_2, \omega_2)\}$$

$$U_{21} = \{(n_1, \omega_1)\} \cup \{(n_1, \omega_2)\} \cup \dots \cup \{(n_m, \omega_n)\} = \mathcal{U}_{21}$$

L'ensemble de schémas n° 1.66 illustre ces développements. On y a représenté la zone de la sélection préalable par un semis de points, et la zone de l'énumération par des hachures.

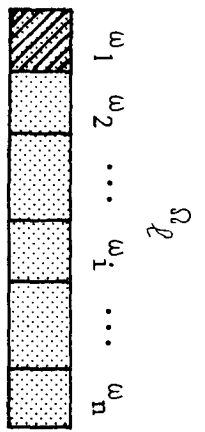
Tout ceci peut assez facilement être généralisé. Ainsi, nous avons vu que Q 7 (où l'interrogation porte sur deux éléments de la phrase) opère une sélection préalable dans \mathcal{U}_7 , alors que ce n'est pas le cas pour la question apparemment de la même catégorie Q 5 (où l'interrogation ne porte que sur un seul élément). Appelons $I = \alpha_i, \beta_i, \dots, \omega_i$ les éléments de la phrase (quelles que soient leur nature et leur fonction grammaticales) sur lesquels porte l'interrogation. En généralisant R'5 et R'7 à l'ensemble B des types Q 3 à Q 9, on obtient :

$$\begin{aligned} R'B = \{ & \{(\alpha_1, \beta_1, \dots, \omega_1)\}, \{(\alpha_2, \beta_2, \dots, \omega_2)\}, (\alpha_2, \beta_2, \dots, \omega_3), \\ & (\alpha_2, \beta_2, \dots, \omega_n), \dots, (\alpha_2, \beta_p, \dots, \omega_2), \dots, (\alpha_2, \beta_p, \dots, \omega_n), \\ & \dots, (\alpha_m, \beta_p, \dots, \omega_n)\} \end{aligned}$$

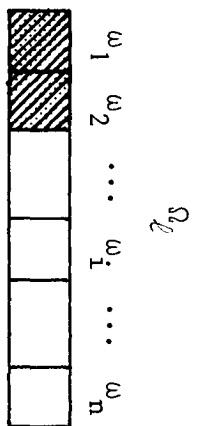
D'où la formule générale de sélection :

$$U B = \{\{\alpha_1\} \times \{\beta_1\} \times \dots \times \{\omega_1\}\} \cup \{\{\bar{\alpha}_1\} \times \{\bar{\beta}_1\} \times \dots \times \{\bar{\omega}_1\}\}$$

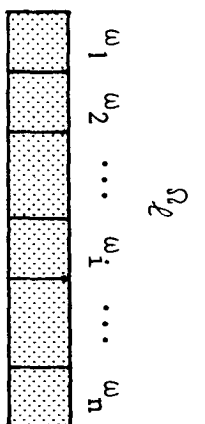
Pour le cas particulier où $I = \{\alpha_i\}$, on retrouve la formule de Q 5, qui s'applique à toutes les questions de cette catégorie ne portant que sur un élément (Q 3, Q 4, Q 5). On peut donc dire que les questions Q 3 à Q 9 opèrent toujours une sélection préalable, mais que, dans le cas particulier (le plus fréquent en pratique) où l'interrogation ne porte que sur un seul élément, cette sélection reste sans effet, et $U = \mathcal{U}$.



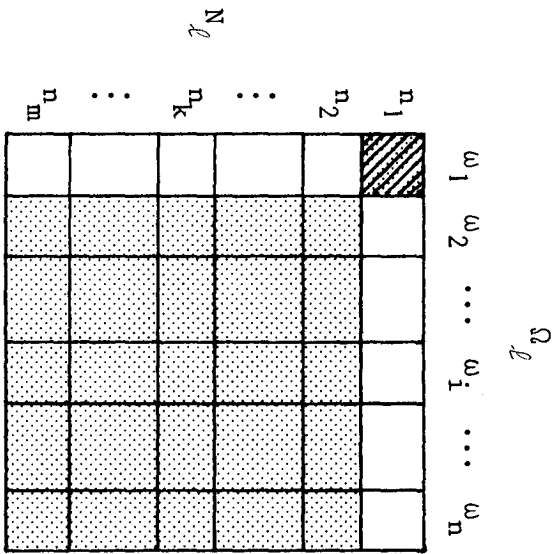
Q5 = $(n_1 v_1 \omega_1)$ ou $(n_1 v_1 \bar{\omega}_1)$?



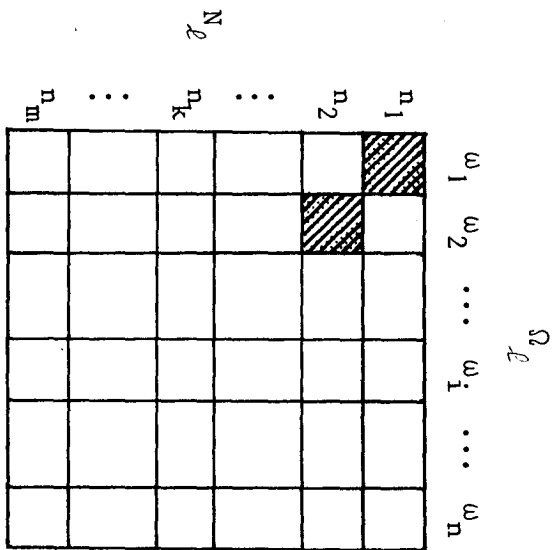
Q12 = $(n_1 v_1 \omega_1)$ ou $(n_1 v_1 \omega_2)$?



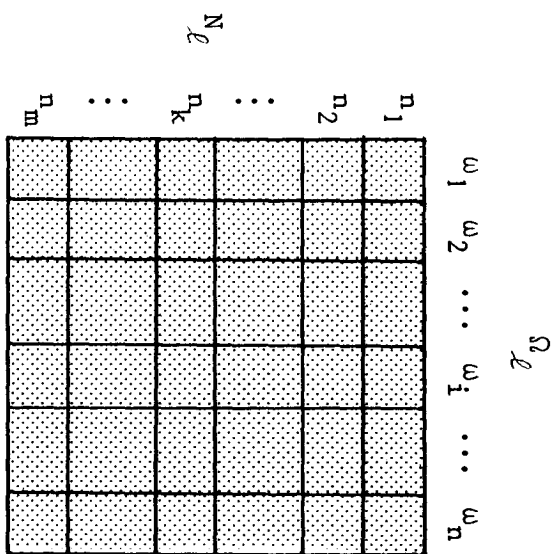
Q19 = $n_1 v_1$ que ?



Q7 = $(n_1 v_1 \omega_1)$ ou $(\bar{n}_1 v_1 \bar{\omega}_1)$?



Q14 = $(n_1 v_1 \omega_1)$ ou $(n_2 v_1 \omega_2)$?



Q21 = que v_1 que ?

1.66 Sélection et énumération pour quelques types de questions de la forme : $N V \Omega$.

On peut également généraliser sans trop de difficultés ce que nous avons constaté pour les questions de la catégorie $A = Q 10$ à $Q 16$, en considérant que nous avons $k > 2$ modalités dans l'alternative de l'énoncé. Si aucun des éléments sur lesquels porte l'interrogation n'est répété, nous avons :

$$R'A = \{ \{(\alpha_1, \beta_1, \dots, \omega_1)\}, \{(\alpha_2, \beta_2, \dots, \omega_2)\}, \dots, \{(\alpha_k, \beta_k, \dots, \omega_k)\} \}$$

et :

$$U A = \bigcup_{i=1}^k \{(\alpha_i, \beta_i, \dots, \omega_i)\}$$

Mais il est vraisemblable que l'on rencontrera des cas où certains termes se trouvent répétés. Par exemple : $(n_1 v_1 \omega_1)$ ou $(n_1 v_1 \omega_2)$ ou $(n_2 v_1 \omega_2)$? = "Vous allez à Paris, vous allez à Bordeaux, ou c'est votre femme qui va à Bordeaux ?". Pour rendre compte de ce cas plus général, on posera :

$$R'A = \{ \{(\alpha_1, \beta_1, \dots, \omega_1)\}, \{(\alpha_2, \beta_1, \dots, \omega_2)\}, \dots, \{(\alpha_i, \beta_j, \dots, \omega_1)\} \}$$

et, en définissant $P_i = \{(\alpha_j, \beta_1, \dots, \omega_p)\}$ comme une partie de \mathcal{U} ne contenant qu'un représentant de chacune des catégories de termes sur lesquelles porte l'interrogation :

$$U A = \bigcup_{i=1}^k P_i$$

De même, pour la catégorie $C = Q 17$ à $Q 23$, nous avons :

$$R'C = \{ \{(\alpha_1, \beta_1, \dots, \omega_1)\}, \{(\alpha_1, \beta_1, \dots, \omega_2)\}, \{(\alpha_1, \beta_1, \dots, \omega_n)\}, \\ \{(\alpha_1, \beta_2, \dots, \omega_1)\}, \dots, \{(\alpha_1, \beta_2, \dots, \omega_n)\}, \dots, \{(\alpha_m, \beta_p, \dots, \omega_n)\} \}$$

et la sélection :

$$U C = \bigcup_{i=1}^m \bigcup_{j=1}^p \dots \bigcup_{\ell=1}^n \{(\alpha_i, \beta_j, \dots, \omega_\ell)\} = \mathcal{U}.$$

L'ensemble de ces résultats est résumé dans le tableau n° 1.67, ainsi que quelques propriétés moins importantes.

Catégorie	B	A	C
Forme syntaxique générale	$(\alpha_1, \beta_1, \dots, \omega_1)$ ou $(\bar{\alpha}_1, \bar{\beta}_1, \dots, \bar{\omega}_1)$?	$(\alpha_1, \beta_1, \dots, \omega_1)$ ou $(\alpha_2, \beta_1, \dots, \omega_2)$ ou $(\alpha_2, \beta_2, \dots, \omega_3)$ ou... ou $(\alpha_i, \beta_j, \dots, \omega_1)$?	que α , que β, \dots , que ω ?
Exemples (cf. tableau n° 1.65)	Q5 = $(n_1 v_1 \omega_1)$ ou $(n_1 v_1 \bar{\omega}_1)$? Q7 = $(n_1 v_1 \omega_1)$ ou $(\bar{n}_1 v_1 \bar{\omega}_1)$?	Q12 = $(n_1 v_1 \omega_1)$ ou $(n_1 v_1 \omega_2)$? Q14 = $(n_1 v_1 \omega_1)$ ou $(n_2 v_1 \omega_2)$?	Q19 = $n_1 v_1$ que ? Q21 = que v_1 que ?
Sélection U dans $\mathcal{P}(\mathcal{U})$	$\{\{\alpha_1\} \times \{\beta_1\} \times \dots \times \{\omega_1\}\}$ U $\{\{\bar{\alpha}_1\} \times \{\bar{\beta}_1\} \times \dots \times \{\bar{\omega}_1\}\}$	$\bigcup_{i=1}^k P_i, P_i = \{(\alpha_j, \beta_1, \dots, \omega_p)\}$	$\bigcup_{i=1}^m \bigcup_{j=1}^p \dots \bigcup_{\ell=1}^n \{(\alpha_i, \beta_j, \dots, \omega_\ell)\}$
Nombre de réponses possibles	$ R' = 2$	$ R' = k$	$ R' = \mathcal{U} $
Cardinal de l'énumération	$ K = 1$	$ K = k$	$ K = 0$
Propriétés remarquables	$ I = 1 \rightarrow U = \mathcal{U}$ $ I > 1 \rightarrow U \subset \mathcal{U}$	$K = U$ $ K = R' $	$K = \emptyset$ $U = \mathcal{U}$

Il est temps maintenant d'examiner les propriétés des types Q 1 et Q 2 à la lumière des considérations qui précèdent. Rappelons que :

$$Q 2 = (n_1 v_1 \omega_1) \text{ ou } (\overline{n_1 v_1 \omega_1}) ? = S \text{ ou } \bar{S} ?$$

avec comme réalisation possible : "Vous allez à Paris, ou non ?". On peut considérer que des réponses de la forme : $\overline{n_1 v_1 \omega_1} =$ "Ce n'est pas moi qui y vais", $n_2 v_1 \omega_1 =$ "C'est mon ami qui y va", $n_3 v_1 \omega_2 =$ "C'est ma femme qui (part, et elle) va à Bordeaux", etc., si elles ne sont pas des réponses explicites à Q 2 *stricto sensu*, représentant néanmoins des situations recouvertes par le : "Non, je ne vais pas à Paris". On a alors :

$$R'2 = \{ \{ (n_1, v_1, \omega_1) \} , \{ (n_1, v_1, \omega_2), \dots, (n_1, v_1, \omega_n), (n_2, v_1, \omega_1), \dots, (n_2 v_1 \omega_n), \dots, (n_2 v_2 \omega_1), \dots, (n_m, v_p, \omega_n) \} \}$$

et par conséquent :

$$U 2 = \bigcup_{i=1}^m \bigcup_{j=1}^p \bigcup_{l=1}^n \{ (n_i v_j \omega_l) \} = \mathcal{U}.$$

C'est pourquoi, si l'on opte pour l'existence d'une interrogation générale, contrairement au point de vue développé par ZUBER (§ 1.2.4.b), on peut ajouter une quatrième colonne au tableau n° 1.61. Cette dernière catégorie X se distingue de la catégorie B en ce que, bien que le nombre des réponses possibles et le cardinal de l'énumération soient comparables ($|R'2| = 2$; $|K 2| = 1$), la sélection n'opère jamais pour Q 2, alors qu'elle est effective pour la catégorie B lorsque $|I| > 1$. En pratique cependant, compte-tenu de la relative rareté en français des interrogations portant sur plus d'un élément, il n'est peut être pas indispensable de retenir la catégorie de l'interrogation générale.

Le cas de Q 1 est plus complexe. Nous avons vu que LANGACKER considère cette forme comme une transformée de Q 2 ; il nous semble que, ou bien ce point de vue peut être défendu pour au moins tous les types de Q 2 à Q 16, ou bien il ne peut l'être pour aucun type. En effet, si nous considérons (comme nous l'avons fait dans le cours de ce paragraphe) que $R 1 =$ "Je vais à Paris/Je vais à Bordeaux/C'est ma femme qui y va/J'en reviens/.../Je ne vais

pas à Paris", on peut montrer que : $\forall i, R_i \subseteq R_1$. Il serait aisé ensuite de rendre compte de l'énoncé Q 1 à partir des autres types de Q 2 à Q 16 par la transformation d'effacement des autres termes de la disjonction. Si par contre nous nous en tenons à la forme syntaxique de Q 1, nous constatons que, bien qu'interrogatif, cet énoncé ne propose pas de choix. Il est vrai que les habitudes culturelles nous conduisent à considérer Q 1 comme une question de la forme Q 2 (ou Q 9 ? ou d'un autre type encore ? cf. ZUBER). Ce qui nous ramène au problème précédent. C'est pourquoi nous avons préféré écarter Q 1 de notre classification, en considérant cette forme ambiguë comme atypique.

Il y aurait certainement de nombreux développements à ajouter à propos de cette tentative de formalisation. Par exemple :

- présenter certaines propriétés remarquables, comme la relation : $\forall Q_i : K_i \subseteq U_i \subseteq \mathcal{U}_i$, dans un contexte donné (quelle que soit la question, tous les éléments cités dans l'énoncé sont des éléments de la sélection ; tous les éléments de la sélection sont des éléments de l'univers de l'interrogation) ;
- analyser les rapports entre les diverses formes de réponses (entières, anaphoriques, zeugmatiques), les "factifs strumentaux" *oui, non, si* [Damourette 1931 : IV 350-352], et les catégories de questions (ainsi, la réponse *oui/non* n'est possible qu'avec les questions de la catégorie B) ;
- établir un parallèle entre l'opérateur linguistique de disjonction *ou* et l'opérateur logique d'exclusion, ainsi qu'entre la concaténation et l'opérateur logique de conjonction ;
- prendre en considération les réponses de la forme : n_1 et $n_2 \vee_1 \omega_1 =$ "Pierre et Jacques vont à Paris", telles que la partie de l'univers de l'interrogation qui lui correspond contienne plus d'un élément.

Les bases morpho-syntaxiques sur lesquelles se fonde cette analyse sont toutefois à la frontière de la logique des questions, qui occupe une place importante dans le chapitre suivant. C'est pourquoi nous concluons ce chapitre 1 sur la définition du substrat morpho-syntaxique commun à toutes les questions.

e) La forme canonique des énoncés à valeur de question.

Au terme de cette analyse morpho-syntaxique, nous disposons des éléments nécessaires pour décrire la forme canonique des énoncés à valeur de question. Une telle forme devrait être aisée à analyser sur le plan sémantique, tout en étant réductible à toutes les autres formes syntaxiques selon des règles bien définies. Pour cela, la forme canonique devrait satisfaire à trois conditions précises, que nous allons énumérer. Considérons pour cela un ensemble d'énoncés de structures syntaxiques différentes, mais de même signification. On admettra comme sémantiquement équivalents tous les énoncés susceptibles d'être substitués les uns aux autres (dans un questionnaire ou dans une conversation), sans que le comportement induit par ces énoncés chez l'interlocuteur en soit modifié de façon perceptible. En adoptant une terminologie harrissienne, il nous faudra élire, dans cet ensemble d'énoncés équivalents, une *forme distinguée* qui présente les propriétés :

- d'être totalement explicite, et dépourvue d'ambiguïtés d'origine non lexicale ;
- de pouvoir jouer un rôle de "source", dont tous les énoncés équivalents pourraient être dérivés à l'aide de transformations simples ;
- d'être elle-même dérivable par transformation de n'importe lequel des énoncés équivalents.

La première condition est indispensable pour une analyse logico-sémantique ultérieure. La forme canonique devra par conséquent restreindre l'emploi des mots pleins généraux ("pro-éléments" [Ruwet 1968 : 263], ou mots indéfinis), comme "quelqu'un", "quelque chose", "quelque part", etc., qui n'ont pas de contenu proprement sémantique [Tesnière 1959 : 60 § 7]. En particulier, si de tels mots figurent dans la forme canonique, ils ne doivent pas être un des termes sur lesquels porte l'interrogation. On proscriera donc, dans cette forme distinguée, ce que TESNIÈRE appelle les mots généraux interrogatifs, comme "qui", "où", "pourquoi" ; et, à plus forte raison, les mots généraux non interrogatifs comme "quelqu'un", "quelque part", "pour quelque raison", si ces locutions correspondent à un mot interrogatif dans l'une des transformées para-

phrastiques. D'autre part, pour faciliter l'analyse de contenu de l'énoncé distingué, celui-ci devrait se présenter comme un ensemble de phrases simples (phrases-noyaux, ou discours élémentaires), sur lesquelles agiraient des opérateurs de méta-discours ; chaque phrase-noyau et chaque opérateur de méta-discours pouvant être assimilé à une unité sémantique (ce que fait par exemple Charles E. OSGOOD dans son *evaluative assertion analysis* [Osgood 1959 : 45-47]). Enfin, pour qu'il soit totalement explicite, un énoncé à valeur de question doit énumérer l'ensemble des réponses proposées à l'interlocuteur. Dans les termes de la description formelle des catégories de question au paragraphe précédent, la forme canonique doit appartenir à la catégorie A, pour laquelle l'énumération coïncide avec la sélection ($K = U$).

La seconde condition constitue une sorte de justification, au niveau de la théorie syntaxique, des équivalences de sens constatées empiriquement. Elle revient à dire que, si une forme d'énoncé donnée est considérée comme forme canonique, on doit pouvoir en déduire toutes les formes (syntaxiques) d'énoncés équivalents. Les règles de transformation dépendent étroitement de la langue considérée. Nous allons donc examiner les relations transformationnelles existant en français dans un ensemble d'énoncés équivalents. Nous nous limiterons pour cela à un cas très simple, dans lequel la phrase de base est de la forme $S = N V \Omega$, et où l'interrogation porte sur l'objet Ω . En outre, et bien que dans notre exemple le verbe et l'objet soient décomposables en éléments ($V = t V'$ = auxiliaire marquant le futur proche, suivi du verbe ; et $\Omega = PN$ = préposition suivie d'un nom), nous ne tiendrons pas compte des transformations possibles mettant en jeu leur décomposition. Soient les six énoncés de questions suivants :

- ① "Je vous demande si vous allez voter pour X ou si vous n'allez pas voter pour lui"
- ② "Allez-vous voter pour X, ou allez-vous voter pour quelqu'un d'autre ?"
- ③ "Vous allez voter pour X, pour Y, pour Z, ou pour T ?"
- ④ "Pour qui est-ce que vous allez voter ?"
- ⑤ "Je vous demande si c'est pour X, pour Y, pour Z, ou pour T que vous allez voter"
- ⑥ "Est-ce pour X que vous allez voter ?"

Il est facile de montrer qu'ils peuvent être dérivés, par des transformations paraphrastiques, de la forme distinguée :

Ⓡ "Je vous demande si vous allez voter pour X , ou si vous allez voter pour Y , ou si vous allez voter pour Z , ou si vous allez voter pour T .

Le passage de la forme distinguée Ⓡ à l'énoncé ⓐ demande trois transformations (les symboles renvoient au tableau n° 1.68) :

1) Effacement, par ϕ_{z6} , des autres possibilités que "voter pour X", et remplacement de ces possibilités par : "voter pour non X", ce qui donne la phrase agrammaticale :

* "Je vous demande si vous allez voter pour X ou si vous allez voter pour non X ".

2) Déplacement ϕ_{p3} sur le verbe de la négation attachée à l'objet (avec la transformation morphophonématique obligatoire de *non* en *ne ... pas*) :

"Je vous demande si vous allez voter pour X ou si vous n'allez pas voter pour X ".

3) Pronominalisation ϕ_{z7} de l'objet répété :

ⓐ "Je vous demande si vous allez voter pour X ou si vous n'allez pas voter pour lui".

La dérivation de ⓑ demande six étapes :

1) Effacement ϕ_{z6} comme ci-dessus.

2) Remplacement ϕ_{m2} de "non X" par son équivalent dans la langue : "quelqu'un d'autre" :

"Je vous demande si vous allez voter pour X ou si vous allez voter pour quelqu'un d'autre".

Effacements :

- z 1 : effacement des NV répétés, sauf la première apparition
- z 2 : effacement des NV répétés, sauf la dernière apparition
- z 3 : effacement des *si* répétés, sauf la première apparition
- z 4 : effacement des *ou* répétés, sauf la dernière apparition
- z 5 : effacement des *que* non suivis d'une subordonnée
- z 6 : remplacement de : " ω_2 *ou* ... *ou* ω_n " par : " $\bar{\omega}_1$ ", avec effacement des membres de l'alternative portant sur les ω_i , $i \neq 1$
- z 7 : pronominalisation de ω_1 répété
- z 8 : remplacement de : "*si* N *non* V Ω " (après : "*si* N V Ω *ou*") par : "*non*"
- z 9 : effacement de : "*ou non*"
- z 10 : pronominalisation de : "*si* ω_1 *ou si* ... *ou si* ω_n " en : *qu-* (avec effacement des NV répétés)
- z 11 : effacement de *si* (après le performatif W, et éventuellement répété)
- z 12 : effacement du performatif W.

Permutations :

- p 1 : déplacement de Ω en tête de phrase : N V Ω \rightarrow Ω N V
- p 2 : déplacement du sujet après le verbe : N V \rightarrow V N
- p 3 : déplacement sur le verbe de la négation portant sur l'objet :

$$V \text{ non } \Omega \rightarrow \text{non } V \Omega$$

Attachements :

- m 1 : attachement de la périphrase emphatique : "*c'est* ... *que*".
- m 2 : traduction de "*non* Ω " par : "quelqu'un d'autre" (ou : "quelque chose d'autre").
- m 3 : attachement de l'intonation interrogative

1.68 : Principales transformations permettant de dériver de leur forme canonique des énoncés de questions simples, de base N V Ω , avec interrogation portant sur Ω

3) Effacement ϕ_{z11} de la conjonction de subordination *si* placée après le performatif :

*"Je vous demande vous allez voter pour X ou vous allez voter pour quelqu'un d'autre".

4) Permutation ϕ_{p2} du sujet et du verbe :

*"Je vous demande allez-vous voter pour X ou allez-vous voter pour quelqu'un d'autre".

5) Attachement ϕ_{m3} de l'intonation interrogative :

"Je vous demande : allez-vous voter pour X , ou allez-vous voter pour quelqu'un d'autre ?"

6) Effacement ϕ_{z12} du performatif : "je vous demande" :

ⓑ "Allez-vous voter pour X , ou allez-vous voter pour quelqu'un d'autre ?".

Pour passer de ⓐ à ⓑ, il faut également six transformations élémentaires :

1) Effacement ϕ_{z1} de : "vous allez voter" répété :

*"Je vous demande si vous allez voter pour X ou si pour Y ou si pour Z ou si pour T".

2) Effacement ϕ_{z3} des *si* répétés :

"Je vous demande si vous allez voter pour X ou pour Y ou pour Z ou pour T".

3) Effacement ϕ_{z4} des *ou* répétés :

"Je vous demande si vous allez voter pour X , pour Y , pour Z ou pour T".

4) Effacement ϕ_{z11} du *si* suivant le performatif :

*"Je vous demande vous allez voter pour X , pour Y , pour Z , ou pour T ".

5) Attachement ϕ_{m3} de l'intonation interrogative :

"Je vous demande : vous allez voter pour X , pour Y , pour Z , ou pour T ?".

6) Effacement ϕ_{z12} du performatif :

ⓐ "Vous allez voter pour X , pour Y , pour Z , ou pour T ?".

La dérivation de ⓐ fait intervenir sept transformations élémentaires:

1) Déplacement ϕ_{p1} de l'objet en tête de phrase :

*"Je vous demande si pour X vous allez voter, ou si pour Y vous allez voter, ou si pour Z vous allez voter, ou si pour T vous allez voter".

2) Pronominalisation ϕ_{z10} des objets X, Y, Z, et T précédés de *si* en *qu-*, avec effacement corrélatif des NV répétés devenus littéralement sans objet (et transformation morphophonématique obligatoire *qu-* → *qui*) :

"Je vous demande pour qui vous allez voter".

3) Attachement ϕ_{m1} de la périphrase emphatique *c'est ... que* :

*"Je vous demande c'est pour qui que vous allez voter".

4) Permutation ϕ_{p1} appliquée à : "c'est pour qui" :

* "Je vous demande pour qui c'est que vous allez voter".

5) Permutation ϕ_{p2} appliquée au même membre de phrase :

*"Je vous demande pour qui est-ce que vous allez voter".

6) Attachement ϕ_{m3} de l'intonation interrogative :

"Je vous demande : pour qui est-ce que vous allez voter ?"

7) Effacement du performatif ϕ_{z12} :

ⓓ "Pour qui est-ce que vous allez voter ?"

La dérivation de ⓔ permet d'introduire les transformations ϕ_{z2} et ϕ_{z5} :

1) Déplacement ϕ_{p1} comme pour l'énoncé précédent.

2) Attachement ϕ_{m1} de la périphrase emphatique :

"Je vous demande si c'est pour X que vous allez voter, ou si c'est pour Y que vous allez voter, ou si c'est pour Z que vous allez voter, ou si c'est pour T que vous allez voter".

3) Effacement ϕ_{z2} des membres de phrase répétés, sauf le dernier :

* "Je vous demande si c'est pour X que ou si c'est pour Y que ou si c'est pour Z que ou si c'est pour T que vous allez voter".

4) Effacement ϕ_{z5} des *que* non suivis de subordonnée :

"Je vous demande si c'est pour X ou si c'est pour Y ou si c'est pour Z ou si c'est pour T que vous allez voter".

5) Effacement ϕ_{z1} appliqué à $NV = "c'est"$:

* "Je vous demande si c'est pour X ou si pour Y ou si pour Z ou si pour T que vous allez voter".

6) Effacement ϕ_{z3} des *si* répétés :

"Je vous demande si c'est pour X ou pour Y ou pour Z ou pour T que vous allez voter".

7) Effacement ϕ_{z4} des *ou* répétés :

ⓔ "Je vous demande si c'est pour X , pour Y , pour Z , ou pour T que vous allez voter".

Enfin, la dérivation de \textcircled{f} illustre l'application des transformations ϕ_{z8} et ϕ_{z9} :

1 et 2) Application de ϕ_{p1} et ϕ_{m1} comme ci-dessus.

3) Application de ϕ_{z6} , comme pour \textcircled{a} et \textcircled{b} :

*"Je vous demande si c'est pour X que vous allez voter ou si c'est pour non X que vous allez voter".

4) Report ϕ_{p3} de la négation de l'objet sur le verbe :

"Je vous demande si c'est pour X que vous allez voter ou si ce n'est pas pour X que vous allez voter".

5 et 6) Effacement ϕ_{z1} et ϕ_{z5} de mots répétés :

"Je vous demande si c'est pour X que vous allez voter, ou si ce n'est pas pour X".

7) Remplacement ϕ_{z8} de la phrase négative : "si ce n'est pas pour X" par "non" :

"Je vous demande si c'est pour X que vous allez voter, ou non".

8) Effacement ϕ_{z9} de : "ou non" :

"Je vous demande si c'est pour X que vous allez voter".

9 à 12) Application à : "Je vous demande si c'est pour X" des transformations ϕ_{z11} , ϕ_{p2} , ϕ_{m3} , et ϕ_{z12} (comme dans les étapes 3 à 6 de la dérivation de \textcircled{b}):

\textcircled{f} "Est-ce pour X que vous allez voter ?".

Les six exemples qui précèdent étaient nécessaires pour présenter l'ensemble des dix-huit transformations paraphrastiques qui permettent de passer de la forme canonique \textcircled{r} aux énoncés de même signification (cf. tableau n° 1.68). On vérifierait sans peine que ces transformations rendent compte de

la dérivation d'autres formes d'interrogation, comme par exemple :

- ⓐ "Allez-vous voter pour X ou pour quelqu'un d'autre ?"
- ⓑ "Je vous demande si c'est pour X que vous allez voter, ou pour Y ,
ou pour Z , ou pour T "
- ⓒ "Est-ce que vous allez voter pour X ?"
- ⓓ "Est-ce que c'est pour X que vous allez voter ?"

Comme le montre la figure n° 1.69, il est possible de dériver de ⓐ plus de 180 énoncés appartenant à la langue, à l'aide des seules transformations du tableau n° 1.68. Dans ce graphe, la forme canonique ⓐ est, par construction, la racine de l'arborescence.

Pour vérifier si la seconde condition est satisfaite par la forme distinguée ⓐ, il faut s'assurer que tous les énoncés équivalents à ⓐ peuvent en être dérivés, quelle que soit leur structure syntaxique. Par exemple, la forme suivante ne figure pas dans le graphe de la figure n° 1.69 :

"Allez-vous voter pour X , ou est-ce que vous allez voter pour
quelqu'un d'autre ?".

Il s'agit en effet d'une forme hybride, mêlant deux modes d'interrogation (inversion sujet/verbe, et particule *est-ce que*). Pourtant, on pourrait en rendre compte en supposant qu'après ϕ_{z6} et ϕ_{m2} (cf. les étapes 1 et 2 de la dérivation de ⓑ), on applique ϕ_{m1} au second membre seulement de l'alternative, puis ϕ_{z11} , ϕ_{p2} , ϕ_{m3} , et ϕ_{z12} . Par contre, on vérifierait sans peine que :

"Allez-vous voter pour X , Y , Z , ou T ?"

ne peut être dérivé de ⓐ à l'aide des seules transformations que nous avons décrites. Cela tient à ce que nous n'avons pas tenu compte de la possibilité de décomposer $\Omega = PN$. Il suffit, pour que la dérivation soit possible, d'introduire la transformation d'effacement (sauf la première apparition) de la préposition *pour* répétée.

Les deux exemples suivants posent par contre des problèmes plus complexes. Ainsi :

"Pour qui allez-vous voter : pour X , pour Y , pour Z , ou pour T ?"

suppose ou bien que l'on introduise une transformation de répétition de l'objet du verbe ($N V \Omega \rightarrow N V \Omega \Omega$), ou bien que l'on décompose les transformations de pronominalisation comme ϕ_{z10} en pronominalisation proprement dite (sans effacement), et effacement du mot pronominalisé. De même, l'énoncé :

"Pour quel candidat allez-vous voter ?"

ne peut en aucun cas être dérivé de (r) . Mais on peut le dériver de :

(r') "Je vous demande si vous allez voter pour le candidat X ou si vous allez voter pour le candidat Y ou si vous allez voter pour le candidat Z ou si vous allez voter pour le candidat T".

En utilisant une transformation d'effacement du mot approprié (d'ailleurs répété) : "le candidat", on obtient (r) . En tenant compte des remarques qui précèdent, on pourrait par conséquent dériver de (r') environ quatre cents énoncés interrogatifs appartenant à la même classe de transformation. La phrase de base de (r') , sur laquelle opère la disjonction, est, comme nous l'avons vu (§ 1.2.3.e), de la forme : "*Vous allez voter pour X* contient un argument 2 identique à l'argument 1 de *X est candidat*", d'où : "*Vous allez voter pour X qui est candidat*", d'où enfin : "*Vous allez voter pour le candidat X*".

Ce rapide examen semble confirmer la possibilité de dériver d'une forme canonique unique un ensemble d'énoncés interrogatifs équivalents, à l'aide des transformations paraphrastiques. Toutefois, il est clair que la preuve ne pourra en être apportée qu'à l'issue de recherches ultérieures sur les paraphrases des énoncés à valeur de question, dans le cadre d'une syntaxe transformationnelle du français.

La troisième condition nous impose de vérifier que l'on peut remonter, par induction, de n'importe quel énoncé interrogatif à la source dont il est dérivé. Il est facile de s'assurer que les transformations du tableau n° 1.68 sont réversibles, sauf en ce qui concerne ϕ_{z6} , ϕ_{z10} , et ϕ_{p3} . On notera d'ailleurs que ces trois transformations sont celles qui déterminent le passage d'une forme

de question à une autre (comme on s'en assurera sur le graphe n° 1.69) :

$A \xrightarrow{\phi_{z6}} B$, $A \xrightarrow{\phi_{z10}} C$, $B \xrightarrow{\phi_{p3}} X$. Nous allons donc devoir examiner en détail la possibilité d'inverser chacune de ces trois transformations.

La transformation ϕ_{z6} , appliquée à \textcircled{r} , donne (cf. exemples \textcircled{a} et \textcircled{b}) :

*"Je vous demande si vous allez voter pour X ou si vous allez voter pour non X".

Comment, à partir de cet énoncé, reconstituer la source \textcircled{r} ? Cela suppose la connaissance de l'univers \mathcal{U} dont X est un élément, de telle sorte que l'on puisse énumérer tous les éléments du complément à X dans \mathcal{U} . C'est-à-dire que la connaissance compréhensive (ou intensive) de $\mathcal{U} = \{a : a \text{ est un candidat aux élections en question}\}$ ne suffit pas : sa connaissance extensive est indispensable (et elle seule). La difficulté est la même pour la transformation ϕ_{z10} qui, dans l'exemple \textcircled{d} , donne l'énoncé :

"Je vous demande pour qui vous allez voter".

Pour retrouver l'étape précédente de la dérivation, il faut savoir que :

$\mathcal{U} = \{X, Y, Z, T\}$. Le problème ne se pose pas en fait pour le linguiste qui dispose d'un corpus suffisamment étendu de phrases équivalentes, et qui connaît ainsi l'ensemble des réponses possibles à la question dans le discours dont il s'agit (cf. sur ce point : [Harris 1968 : 94-95]). Il est évident qu'il ne se pose pas non plus pour le psychologue ou le sociologue analysant un questionnaire, puisque ou bien figure dans le contexte immédiat de l'énoncé interrogatif la liste des réponses possibles (si la question est fermée ou précodée), ou bien la catégorisation des réponses libres se fonde sur une grille de codification (ou d'analyse de contenu) qui constitue une énumération de \mathcal{U} .

Par contre, l'inversion de la transformation ϕ_{p3} peut être une source d'ambiguïté. En effet, on a dans l'exemple \textcircled{a} la séquence : "vous n'allez pas voter pour X" qui provient par ϕ_{p3} de : *"vous allez voter pour non X". Soit, compte-tenu de l'ambiguïté de la langue française,

$n_1 v_1 \bar{\omega}_1 \xrightarrow{\phi_{p3}} n_1 \bar{v}_1 \omega_1$ ou : $n_1 v_1 \bar{\omega}_1 \xrightarrow{\phi_{p3}} \overline{(n_1 v_1 \omega_1)}$. Pour retrouver

l'énoncé de l'étape précédente, il faut savoir que l'interrogation porte sur Ω . Dans notre exemple, il ne semble pas y avoir d'ambiguïté, pour des raisons essentiellement lexicales. Mais si l'on pose $n_1 = \text{"vous"}$, $v_1 = \text{"soutenir"}$, $\omega_1 = \text{" X "}$, de : "vous ne soutenez pas X" on peut passer à : "vous soutenez Y" (avec $\omega_2 = \text{"Y"}$) si l'interrogation porte sur Ω , ou à : "vous combattez X" (avec $v_2 = \text{"combattre"}$), si l'interrogation porte sur V. En pratique toutefois, comme pour les exemples précédents, le contexte devrait permettre de trancher.

En conclusion, malgré de sérieuses réserves sur le plan théorique (réserves que des recherches empiriques ultérieures permettront peut être de lever), nous pouvons considérer les trois conditions nécessaires à l'existence d'une forme canonique comme satisfaites, du moins dans la perspective essentiellement pratique qui est la nôtre. Aussi pouvons-nous définir la forme canonique des énoncés à valeur de question comme *un énoncé performatif du type : "je vous demande si ..."*, opérant sur une disjonction de phrases assertives dont l'ensemble constitue l'ensemble des réponses possibles à l'interrogation.

Pour répondre à la question posée au début de ce chapitre, nous ajouterons qu'une question est un énoncé dont la structure syntaxique est soit identique à celle de la forme canonique, soit dérivable de celle-ci uniquement à l'aide de transformations paraphrastiques (aux transformations morphophonématiques obligatoires près).

L'intérêt des définitions qui précèdent est de mettre en évidence les trois points suivants :

- la connaissance extensive de l'ensemble des réponses possibles est essentielle pour l'analyse d'une question ;
- il existe une classe de transformation des énoncés performatifs servant à introduire une question, dont la variété doit permettre de rendre compte de la *modalité* de l'énoncé (interrogative, énonciative, injonctive) ;
- les transformations paraphrastiques entre énoncés permettent de rendre compte des relations de synonymie entre questions de structures morphosyntaxiques différentes.

Cependant, en l'état actuel des recherches linguistiques, il est difficile de définir la classe des performatifs, ou d'analyser les différences de sens entre des formulations apparemment "synonymes", en des termes "purement morpho-syntaxiques". C'est pourquoi nous aborderons ces problèmes au chapitre 2, dans le cadre d'une analyse logico-sémantique des énoncés à valeur de question, et dans la première partie du chapitre 3, en présentant l'ensemble des traits sémantiques qu'il est ou non possible de prendre en compte dans une formalisation fondée sur des relations simples entre énoncés ou éléments d'énoncés.



2. QUESTIONS ET RÉPONSES

La réduction à une forme canonique des énoncés à valeur de question met en évidence l'étroite interdépendance qui lie la question à l'ensemble des réponses possibles à celle-ci. Cette constatation nous conduit, contrairement à la majorité des grammaires traditionnelles, à introduire dans notre analyse des questions la notion de réponse possible, et même à accorder à cette notion une place prépondérante. L'objet de ce second chapitre est l'étude des liens existant entre une question et l'ensemble de ses réponses possibles. Corrélativement, nous traiterons également des différents types de réponse, et des moyens permettant de les identifier. Nous serons ainsi amené à poser le problème de la validité d'une réponse, c'est-à-dire de sa pertinence et de sa valeur informative relativement à une question donnée.

Dans les situations courantes, la plupart des locuteurs n'éprouvent guère de difficultés pour décider si un énoncé déclaratif répond ou ne répond pas à une question, voire s'il y répond complètement ou seulement partiellement. Cela ne signifie évidemment pas que le problème technique de l'identification des réponses valides soit aisé à résoudre, mais seulement qu'il existe sur ce point un large consensus entre personnes parlant une même langue ; consensus dont les fondements restent à expliciter. Toutefois, ce consensus semble faire défaut à propos d'un cas particulier, illustré par exemple par les mésaventures de Candide, déserteur involontaire de l'armée du roi des Bulgares : "On lui demanda juridiquement ce qu'il aimait le mieux d'être fustigé trente-six fois par tout le régiment, ou de recevoir à la fois douze balles de plomb dans la cervelle. Il eut beau dire que les volontés sont libres, et qu'il ne voulait ni l'un ni l'autre, il fallut faire un choix" (VOLTAIRE, *Candide ou l'optimiste*, ch. 2). Dans cet exemple, celui qui pose la question enferme son interlocuteur dans une alternative : " A ou B " ; dans cette optique, ni : "les deux", ni : "ni l'un ni l'autre", ne constituent une réponse acceptable

aux yeux du locuteur (encore que l'interlocuteur soit ici d'un avis contraire). De même, on connaît la célèbre caricature qui représente Charles Alexandre CALONNE (sous la figure d'un singe) s'adressant aux membres de l'Assemblée des Notables (figurés par des animaux de basse-cour) pour leur demander à quelle sauce ils désirent être mangés. A la réponse : "Mais nous ne voulons pas être mangés du tout !!!", CALONNE réplique : "Vous sortez de la question" (Bibliothèque Nationale, Estampes : QB 1 22/2/1787). Comme la réponse de Candide, celle des Notables présente la particularité de nier la présupposition de la question. Le problème est de savoir si la négation d'une présupposition est ou n'est pas une réponse acceptable à la question.

La notion de présupposition est, plus qu'il n'apparaît sur ces deux exemples, une notion tout à fait fondamentale pour nous, au même titre que la notion de réponse possible (à laquelle elle est d'ailleurs liée). Son importance théorique apparaîtra tout au cours de ce second chapitre. Son intérêt pratique peut être illustré de plusieurs manières. Tout d'abord, il est clair qu'une réponse qui nie une présupposition apporte plus d'information sur le problème en cause qu'une réponse complètement hors du sujet. Par exemple, à la question : "Combien avez-vous d'enfants ?", il est possible de répliquer : "Plus que je n'aurais souhaité en avoir" ; ou : "Une centaine, au moins !" ; ou : "Pourquoi me demandez-vous cela ?" ; ou : "Quel beau temps pour la saison !" ; ou encore : "Mon dieu, j'ai oublié mes papiers !". Pour des raisons diverses, aucune de ces répliques ne peut constituer une réponse acceptable à la question, du moins dans les situations courantes (cf. : "Comment ne pas répondre à une question" [Weiser 1975]). Par contre, "Je n'ai pas d'enfants", qui nie la présupposition de la question, doit cependant être considéré comme une information utile, au même titre que la réponse : "J'en ai quatre". D'où l'intérêt de prévoir, dans la codification des déclarations de l'interlocuteur, une rubrique particulière pour la négation des présuppositions, distincte à la fois des réponses proprement dites et de l'aveu d'ignorance ou du refus de répondre. Ensuite, le type de critique que l'on adresse le plus fréquemment à la formulation d'une question est qu'elle enferme la personne interrogée dans une problématique qui peut n'être pas la sienne, et que celle-ci devrait par conséquent avoir la latitude de rejeter. Supposons par exemple que l'on demande, dans une

enquête par questionnaire : "Pour lutter contre les facteurs génétiques de la délinquance, quelle est à votre avis, parmi les mesures suivantes, celle que le gouvernement devrait adopter en priorité ?" ; en choisissant l'une des mesures figurant sur la liste qui lui est soumise, le répondant se sentirait cautionner la thèse selon laquelle il existe des facteurs biologiques héréditaires prédisposant à la délinquance. Il est clair que cette conception ne peut, en l'état actuel de nos connaissances, être présupposée (c'est-à-dire être implicitement présentée comme sinon certaine, du moins communément admise), et imposée de ce fait à l'interlocuteur. On remarquera que la plupart des critiques adressées par les politologues aux questions utilisées dans les sondages d'opinion traduisent en réalité le rejet des présupposés de ces questions.

Il nous paraît donc nécessaire, en traitant du problème des réponses possibles à une question, d'aborder parallèlement celui des présupposés de la question. Dans ce dernier domaine, l'apport de la linguistique contemporaine est considérable, et nous ne pourrions en donner qu'un bref aperçu. Par contre, le problème des réponses possibles à une question a été peu étudié par les linguistes. Il est vrai qu'au premier abord, l'identification des réponses acceptables à une question donnée semble ne pas présenter de difficulté particulière, comme le montre apparemment l'exemple des questions à réponse par *oui* ou *non*, voire celui des questions de choix entre un nombre fini de modalités unies par une relation de disjonction (c'est-à-dire la forme canonique définie au chapitre 1). Or, ceci peut être vrai dans le cas de réponse entière à une question générale, à condition toutefois que les termes utilisés dans la réponse ne s'écartent pas trop de ceux qui figurent dans l'énoncé de la question. Car sinon se posent des problèmes de synonymie des termes ou d'équivalence sémantique des phrases, donc de transformations paraphrastiques fondées éventuellement sur des équivalences de traits lexicaux. Dans le cas de questions particulières, nous verrons (§ 2.1.2.a) qu'aucune des grammaires présentées dans le premier chapitre ne propose, en l'état actuel de leur développement, de solution vraiment satisfaisante au problème de l'identification des réponses acceptables, malgré la tentative intéressante de Jerrold J. KATZ (§ 2.1.2.b), inspirée des travaux de CHOMSKY.

C'est pourquoi nous avons, dans ce chapitre, accordé une place prépondérante à la logique érotétique, qui propose un ensemble de modèles formels permettant de définir clairement les relations entre questions, réponses, et présupposés, et de distinguer divers types de questions, de réponses, et de relations entre questions. Comme les grammaires génératives et les grammaires transformationnelles, la logique érotétique est actuellement en plein développement ; mais contrairement à celles-ci, elle ne semble pas avoir suscité beaucoup d'intérêt chez les auteurs de langue française. C'est pourquoi il nous a paru nécessaire de lui consacrer de plus amples développements que nous ne l'avons fait au premier chapitre pour les thèses de HARRIS ou de CHOMSKY, et les travaux qu'elles ont inspirés.

La première partie de ce chapitre constitue une transition entre l'analyse linguistique et l'approche logique du problème des relations entre questions et réponses. Dans la seconde partie, nous présentons d'une manière relativement détaillée l'un des trois systèmes majeurs de logique érotétique, celui de Nuel D. BELNAP Jr. Cette présentation est complétée, dans la troisième partie, par un énoncé plus bref et plus sélectif d'autres systèmes, parmi lesquels ceux de Lennart ÅQVIST et de Tadeusz KUBIŃSKI. Enfin, dans la quatrième partie, nous dressons le bilan des apports de ces modèles formels à l'analyse des relations entre questions, réponses, et présupposés.

2.1. DE LA LINGUISTIQUE A LA LOGIQUE DES QUESTIONS

Cette première partie du chapitre 2 est donc, d'une certaine manière, une introduction à la logique des questions, qui fait l'objet des seconde et troisième parties. Avant le bref survol historique de la logique érotétique qui la clôt (§ 2.1.3), nous présentons la contribution des linguistes à l'analyse des relations entre les questions et leurs réponses et présupposés (§ 2.1.2). Nous avons jugé utile de faire précéder cette présentation par un exposé succinct des résultats principaux (selon les objectifs de ce travail) des recherches des linguistes sur le problème général des relations entre énoncés, en insistant tout particulièrement sur la relation de présupposition (§ 2.1.1).

2.1.1. Les relations sémantiques entre énoncés déclaratifs

Les relations entre énoncés déclaratifs ont été particulièrement étudiées par les linguistes, en partie probablement à cause du parallèle possible entre l'analyse sémantique de ces énoncés et la logique des propositions. Nous verrons en effet (§ 2.1.1.d) que certaines des relations usuelles entre énoncés ont des propriétés semblables à celles de relations logiques, qui peuvent ainsi leur servir de modèles formels ; tel est le cas par exemple des couples synonymie/équivalence, et inconsistance/exclusion réciproque. Mais les relations qui, de notre point de vue, sont les plus importantes, à savoir les relations d'implication et surtout de présupposition, ne trouvent guère d'équivalents formels satisfaisants. Aussi développerons-nous particulièrement les aspects proprement linguistiques de la relation de présupposition, en indiquant les critères permettant d'identifier les présupposés, ainsi que les manières de les classer, avant d'aborder les problèmes de formalisation de cette relation.

a) *Le problème linguistique des présupposés d'un énoncé.*

"Chaque fois que quelqu'un traduit une pensée en mots, il y a dans son esprit beaucoup plus d'idées que son énoncé n'en exprime. Parmi ces idées, certaines entretiennent une relation particulière avec la pensée exprimée : [...]"

ce sont ses présuppositions" [Collingwood 1940 : 21]. A la suite de COLLINGWOOD (voire de FREGE), les linguistes considèrent que, dans le contenu de tout énoncé déclaratif, il est possible de distinguer (au moins) deux composantes : le *posé*, qui est en principe l'information que le locuteur veut apporter explicitement à son interlocuteur, et le *présupposé*, information qui "est présentée comme une donnée à partir de laquelle on parle, mais qui n'est pas directement en jeu dans la parole" [Ducrot 1972 : 23] (cf. également [Strawson 1964 : 80] pour le développement de ces notions). Par exemple [Leech 1974 : 293], dans l'énoncé : "Le livre que vous avez volé à la bibliothèque est intéressant", le *posé* est : "Le livre est intéressant", et le *présupposé* : "Vous avez volé le livre à la bibliothèque" (on trouvera un échantillon de la variété des présupposés dans [Keenan 1971 : 46-47]). Pour identifier et dégager les présupposés d'un énoncé, il est possible d'utiliser certaines transformations non paraphrastiques ; ces transformations ont la propriété de ne pas avoir la même action sur le *posé* et sur le *présupposé*. Tel est en particulier le cas de la négation.

Le *test de la négation* pour identifier la présupposition d'un énoncé semble dû à Gottlob FREGE (cf. [Frege 1892 : 116]) ; mais on peut en découvrir une forme embryonnaire dans la *Logique de Port-Royal* [Arnault 1683 : 190]. Le principe de ce test est que, lorsque l'on remplace un énoncé par sa négation, seul le *posé* est nié ; le *présupposé* reste affirmé (dans le cas d'une présupposition affirmative). Ainsi, nous avons le couple d'énoncés suivant [Leech 1974 : 294]:

{ "La blonde qu'il a épousé est une riche héritière"
 { "La blonde qu'il a épousé n'est pas une riche héritière".

Le contenu implicite commun à ces deux énoncés est : "Il a épousé une blonde". De même, l'énoncé : "Sam ne se rend pas compte que Irv est un martien" présuppose : "Irv est un martien" [Lakoff 1970 : 33]. La présupposition de ces énoncés est un élément qu'il est facile d'identifier et d'isoler dans les énoncés eux-mêmes. Tel n'est pas toujours le cas. Ainsi, dans le couple [Lakoff 1970 : 34] :

{ "Irv fait semblant d'être malade"
 { "Irv ne fait pas semblant d'être malade"

le présupposé : "Irv n'est pas malade" ne parait pas susceptible d'être extrait des énoncés par une procédure purement syntaxique ; même dans l'optique de la syntaxe générative, on ne trouve pas dans la structure profonde des énoncés cités d'indicateur syntagmatique à constituant négatif pouvant correspondre à la présupposition.

Cette constatation nous conduit à introduire une distinction entre deux types de présuppositions [Kiefer 1973 : 92] : celles qui dépendent du sens d'un élément du lexique (ici, l'expression : "faire semblant de"), que l'on appellera *présuppositions lexicales* ; et celles qui ne font pas partie du sens d'un élément particulier, mais qui dépendent plutôt du sens de l'énoncé, et que l'on appellera *présuppositions non lexicales*. Le test de la négation peut être appliqué également aux présuppositions lexicales. Par exemple :

- "Marie a (n'a pas) nettoyé la pièce" présuppose : "la pièce était sale" [Kiparski 1970 : 351] ;

- "Le chien a (n'a pas) couru après le chat" présuppose que le chat s'est enfui [Fillmore 1971 : 382] ;

- "Léna consentit (ne consentit pas) à venir" présuppose que l'on a demandé à Léna de venir [Kiefer 1973 : 90].

Le test de la négation a fait l'objet de nombreuses critiques. La plus fréquemment relevée concerne la *portée* de la négation. Bertrand RUSSEL a fait remarquer que la négation de : "L'actuel roi de France est chauve" peut signifier soit : "Il existe une entité qui est actuellement roi de France, et cette entité n'est pas chauve" ; soit : "Il est faux qu'il existe une entité qui est actuellement roi de France et qui est chauve" [Russel 1905b : 53]. Nous reviendrons ultérieurement sur les conséquences logiques de cette ambiguïté sémantique de la négation dans certaines langues naturelles. Signalons seulement que cet exemple introduit naturellement une seconde distinction relative aux présuppositions : la division en *présuppositions existentielles* (de la forme : "Il existe actuellement un roi de France"), et *présuppositions non existentielles* (cf.p.ex. [Kiefer 1973 : 88-89]). La critique formulée par RUSSEL a été souvent reprise. Il est aisé de montrer que la négation peut porter soit sur la phrase entière, soit sur un élément de la phrase. Par exemple,

l'énoncé : "Beaucoup de flèches ont atteint la cible" peut être transformé par la négation en : "Pas beaucoup de flèches ont atteint la cible", "Beaucoup de flèches n'ont pas atteint la cible", voire (sous réserve que cette phrase soit grammaticale en français) : "Pas beaucoup de flèches n'ont pas atteint la cible" (cf. [Chomsky 1972 : 56]) ; sans omettre évidemment : "Il est faux que beaucoup de flèches aient atteint la cible". De même : "Arthur aboie" peut avoir pour négation : "Il n'est pas vrai qu'Arthur aboie", ou : "Ce n'est pas Arthur qui aboie" [Zuber 1972b : 18-19]. Dans le premier cas, on nie la faculté qu'aurait Arthur d'aboyer, alors que, dans le second, on exprime que ce quelqu'un (cet animal ?) qui aboie actuellement n'est pas Arthur.

Cette critique de l'ambiguïté de la négation a été reformulée par Oswald DUCROT de la manière suivante [Ducrot 1972 : 37-39]. La négation remplit, dans les langues naturelles, au moins deux fonctions différentes : une fonction surtout *descriptive*, comme dans l'énoncé : "Il n'y a pas un nuage dans le ciel" ; et une fonction *métalinguistique* (ou : *polémique* [Ducrot 1973 : 123-124]), que l'on emploie non pour décrire, mais pour marquer son opposition à une affirmation antérieure (qui peut être implicite), comme dans l'énoncé : "Le mur n'est pas blanc". La négation descriptive ("ne ... pas" portant en général sur le syntagme verbal) n'affecte pas la présupposition, alors que la négation métalinguistique ("il est faux que ...") peut affecter à la fois le posé et le présupposé. En conséquence, c'est à la négation descriptive qu'il faut recourir pour identifier les présuppositions d'un énoncé. Toutefois, la négation descriptive elle-même n'est pas exempte de défauts. Considérons par exemple le couple d'énoncés :

- | | |
|---|--|
| } | - "L'effondrement du pont a été provoqué par le séisme" |
| } | - "L'effondrement du pont n'a pas été provoqué par le séisme". |

On peut en extraire la présupposition : "Le pont s'est effondré". Mais si l'on considère les mêmes énoncés sous forme active :

- | | |
|---|---|
| } | - "Le séisme a provoqué l'effondrement du pont" |
| } | - "Le séisme n'a pas provoqué l'effondrement du pont" |

il est clair que seul le premier énoncé présuppose : "Le pont s'est effondré", tandis que le second suggère au contraire l'idée que le pont a tenu bon [Leech 1974 : 314]. L'échec du test de la négation est imputable ici aux effets, sur la portée de la négation, de l'ordre des mots (et en conséquence de la place de l'accent tonique ; cf. [Bickerton 1975 : 52-53]).

De nombreuses autres transformations ont été proposées pour identifier les présuppositions d'un énoncé. Nous ne citerons ici que les principales, en les illustrant brièvement :

- L'*adverbialisation* est une généralisation du test de la négation, fondée sur l'hypothèse que l'adverbe ajouté à l'énoncé n'affecte pas le présupposé [Langendoen 1971a : 341-342]. Ainsi, dans les trois énoncés suivants :

- "Rocky n'a pas accusé Max d'avoir dilapidé le butin"
- "Rocky a avec raison accusé Max d'avoir dilapidé le butin"
- "Rocky a violemment accusé Max d'avoir dilapidé le butin"

ni la négation, ni les adverbes ne modifient les présuppositions lexicales (liées au sens du verbe : "accuser") : "Quelqu'un a dilapidé le butin" et : "Dilapider le butin est mal".

- La *mise sous forme interrogative* fait porter le questionnement sur le posé, et n'affecte pas le présupposé. Ainsi, dans le couple [Ducrot 1972 : 22-23] :

- "Pierre s' imagine que Jacques va venir"
- "Est-ce que Pierre s' imagine que Jacques va venir ?"

le présupposition : "Jacques ne viendra pas" (due au verbe : "s'imaginer") n'est pas touchée par l'interrogation (On trouverait d'autres exemples dans [Kiparsky 1970 : 351]).

- La *mise sous forme impérative* a les mêmes effets. Par exemple : "Poursuivez le voleur !" ou : "Je vous demande de poursuivre le voleur" a pour présupposition : "Le voleur s'enfuit" [Kiparsky 1970 : 353].

- L'*enchaînement d'un autre énoncé*, lié au précédent par une conjonction autre que *et* et *si*, n'affecte que le posé. Par exemple, l'énoncé : "Jean

ne prend plus de caviar au petit déjeuner", pose : "Jean ne prend pas actuellement de caviar au petit déjeuner", et présuppose : "Jean prenait autrefois du caviar au petit déjeuner". Si à l'énoncé primitif, on enchaîne un énoncé explicatif, comme : "Jean ne prend plus de caviar au petit déjeuner parce qu'il a peur de grossir", on voit que l'explication ne concerne pas le présupposé [Ducrot 1972 : 81].

Après avoir utilisé les transformations négative, interrogative, et d'enchaînement pour distinguer posé et présupposé dans les énoncés, Oswald DUCROT fait remarquer que ces procédures sont utiles pour prouver l'existence de ces deux formes de contenu, mais insuffisantes pour identifier et extraire à coup sûr d'un énoncé donné ce qu'il pose et ce qu'il présuppose [Ducrot 1972 : 103]. En effet, leur application conduit à considérer comme présupposé "une foule d'indications sémantiques si hétéroclites qu'il deviendrait impossible d'élaborer une théorie tant soit peu systématique pour la prévision des présupposés" ; par exemple, "Seul Pierre est venu" présuppose non seulement : "Pierre est venu", mais aussi : "D'autres avaient des chances de venir", "La venue de Pierre n'était pas impossible" (ce qui présuppose : "Pierre n'était pas alité, en voyage, etc."), "Le destinataire de la phrase comprend le français", etc. D'autre part, dans certains cas, le critère d'identification des présupposés est défini de façon circulaire ; par exemple, si l'on ne peut distinguer syntaxiquement la négation descriptive de la négation métalinguistique d'un énoncé donné, on utilise pour cela le fait que le présupposé est ou n'est pas affecté par la négation. Enfin, les règles d'identification laissent souvent subsister, pour un énoncé donné, la possibilité d'analyses différentes [Ducrot 1972 : 103-105]. C'est pourquoi nous allons, dans les paragraphes suivants, décrire tout d'abord les différents contenus des présuppositions, pour ensuite examiner les propriétés logico-sémantiques des présupposés et leurs relations avec d'autres formes d'énoncés.

b) Les différents types de présupposés

Nous avons déjà signalé incidemment deux dichotomies permettant de classer les présuppositions : l'alternative *existentielle/non existentielle*, et l'alternative *lexicale/non lexicale*. Nous allons introduire également deux autres dichotomies : *idiosyncrasique/universelle*, et *du premier ordre/d'un autre ordre*. Dans chacune de ces quatre dichotomies, le premier terme est le plus important ; le second étant défini par rapport à lui, et servant à désigner l'ensemble des présuppositions n'appartenant pas à la première catégorie. C'est pourquoi nous nous sommes attaché principalement à définir le premier terme de chaque dichotomie, et à décrire les règles d'identification qui lui correspondent.

Les *présuppositions existentielles* au sens large correspondent à deux sous-catégories : celle des présuppositions concernant l'*existence* d'individus (personnes ou objets), et celle des présuppositions relatives à l'*occurrence* d'un événement (cf. [Bickerton 1975 : 48]). Soient par exemple les trois énoncés :

- (a) "Le loup est carnivore"
- (b) "J'ai blessé un loup"
- (c) "Le loup que j'ai blessé m'a mordu"

Le statut existentiel associé au mot "loup" n'est pas le même dans les trois cas. Dans l'énoncé (a), il s'agit d'un terme générique : l'existence d'un individu correspondant à la classe ainsi désignée n'est ni posée, ni présupposée (Il s'agit d'une existence *hypothétique*. On aurait, sur le même modèle : "la licorne a une corne au milieu du front") ; seule est présupposée l'existence du nom de la classe. Cette existence du nom générique est d'ailleurs également présupposée dans les deux autres exemples. En outre, dans (b), l'existence d'un individu élément de la classe est *posée* ; dans (c), elle est *présupposée* [Bickerton 1975 : 49-50].

En règle générale, on peut considérer qu'en français l'usage d'un nom propre est toujours assorti de la présupposition d'existence. En effet, un nom propre désigne toujours en principe un être défini, singulier. Par contre, le problème se pose d'identifier, parmi les noms communs (comme le mot "loup" dans l'exemple ci-dessus), ceux auxquels est associée une présupposition d'existence. Zeno VENDLER a proposé le critère d'*extraction existentielle*, fondé sur la constatation suivante. Pour qu'un nom commun puisse désigner un être particulier (personne, animal, ou objet) dont l'existence est présupposée, il est nécessaire que ce nom soit précédé de l'article défini. En outre, il doit être ou bien immédiatement suivi d'une proposition (relative) qui l'identifie ; ou bien précédé (dans l'ensemble du discours) d'une autre apparition, sans l'article défini, où il se trouve identifié ; ou bien son identification doit être impliquée par le contexte du discours (cf. [Vendler 1967 : 129-130]). Le test d'extraction existentielle consiste alors à construire une phrase avec le nom commun N en question, de la forme : "Il y a un N qui (que, dont, avec quoi, etc.) ..." ; si cette phrase peut être acceptée en tant que paraphrase du discours, la présupposition d'existence est avérée [Vendler 1967 : 131-132]. Soit le texte suivant :

"J'ai tiré un ours et un élan. L'*ours* que j'ai tiré s'est enfui, mais l'*élan* est tombé raide mort. Le *fusil* a marché à la perfection, mais la *carte* que vous m'avez donnée est inexacte".

On déduit facilement la présupposition d'existence des individus ou des objets désignés par les mots en italique en appliquant le test d'extraction existentielle ; seul, le mot *fusil* peut poser un problème, mais son existence est impliquée par le contexte ("Il y a un fusil que j'avais emporté, et avec lequel j'ai tiré") [Vendler 1967 : 130].

La présupposition d'*occurrence* d'un événement est une présupposition d'existence appliquée à une phrase entière. Cette présupposition peut être mise en évidence par la transformation suivante. On ajoute à la phrase l'expression : "le fait que" ; si la phrase résultante est toujours acceptable, la présupposition d'occurrence est avérée [Kiparsky 1970 : 348-349]. Soient par exemple les énoncés :

- { (a) "Je déplore qu'il pleuve"
 (b) "Il se peut qu'il pleuve"

La transformation appliquée à (a) donne la phrase acceptable : "Je déplore le fait qu'il pleuve" ; donc la présupposition : "Il pleut" est avérée. Par contre, appliquée à (b), la transformation n'aboutit à aucune phrase acceptable, même après application de transformations paraphrastiques : *"Le fait qu'il pleuve se peut", *"Le fait qu'il pleuve est possible" (cette dernière phrase étant, en français, au seuil de l'acceptabilité, ce qui n'est pas le cas en anglais). La présupposition d'occurrence est étroitement liée au sens des mots qui introduisent l'énoncé, ou sont introduits par lui. Lorsque la présupposition d'occurrence est vraie, l'énoncé est sujet de syntagmes verbaux tels que : "suffit", "est bizarre", etc. ; ou il est complément de verbes *factifs* comme : "regretter", "savoir", "ignorer", etc. [Kiparsky 1970 : 345, 347] (cf. également [Karttunen 1973 : 57-59]). Ce sont toujours des verbes qui "préservent la réalité" [Vendler 1967 : 133]. Lorsque la présupposition est fausse, on trouve par contre : "est possible", "semble", etc. ; ou encore : "supposer", "croire", "s'imaginer", etc.

Les *présuppositions non existentielles* peuvent être des présuppositions liées au sens des mots qui les introduisent. Par exemple, dans les énoncés suivants [Kiefer 1973 : 90] :

- { - "Jean craint que Léna ne vienne"
 - "Jean espère que Léna viendra"

les verbes "craindre" et "espérer" présupposent que le fait décrit est ressenti comme (respectivement) malheureux ou heureux. Ces énoncés constituent également des exemples de présuppositions lexicales.

Les *présuppositions lexicales* dépendent, nous l'avons dit, de traits sémantiques des mots qui les introduisent. Charles J. FILLMORE en donne les

exemples suivants [Fillmore 1971 : 381] :

- { (a) "Harry a accusé Mary d'avoir écrit l'éditorial"
 (b) "Harry a critiqué Mary pour avoir écrit l'éditorial"

Le premier énoncé présuppose que Harry considère le fait d'écrire l'éditorial comme "mal", et pose que Harry a affirmé que Mary en est l'auteur ; l'énoncé (b) au contraire présuppose que Harry considère Mary comme l'auteur de l'éditorial, et pose que Harry juge le résultat de l'activité de Mary comme "mauvais". Ces présuppositions sont contenues dans le sens des verbes "accuser" et "critiquer". Ainsi, la définition de ces verbes devrait comporter les indications [Fillmore 1971 : 391-392] :

accuser : - "x accuse y de z"

- présuppose : x et y sont humains
z est une activité
x juge z comme "mauvais"
- pose : x indique que y est cause de z

critiquer : - "x critique y pour z"

- présuppose : x et y sont humains
z est une activité ou son résultat
x juge que y est cause de z
- pose : x indique que z est "mauvais"

(on trouvera une analyse des présuppositions liées aux verbes de jugement dans [Fillmore 1969]).

Pour identifier les présuppositions lexicales, Ferenc KIEFER propose de remplacer les sujets et les compléments des verbes par des variables. Dans les pseudo-énoncés :

- { "X craint Y"
 "X espère Y"

les présuppositions lexicales dépendant des verbes sont (seules) conservées

[Kiefer 1973 : 92]. Ce test met en évidence les traits sémantiques qui distinguent des verbes ayant des significations voisines, lorsque ces traits affectent les arguments des verbes (en particulier les compléments). Par exemple, les verbes "se douter", "se demander", "s'imaginer", ont des présuppositions différentes quant à la réalité de leur objet :

- "Marie se doute que sa soeur est une sorcière" présuppose : "La soeur de Marie est une sorcière" ;

- "Marie se demande si sa soeur est une sorcière" ne présuppose ni : "La soeur de Marie est une sorcière", ni : "La soeur de Marie n'est pas une sorcière" ;

- "Marie s' imagine que sa soeur est une sorcière" présuppose : "La soeur de Marie n'est pas une sorcière".

On peut associer à ces complétives respectivement les traits : *factuel/non factuel/irréel* (ou : *antifactuel*). Cette triade se retrouve par exemple pour les verbes de volition :

- "John insiste pour lire les lettres de Mary" (*factuel*)
- "John désire lire les lettres de Mary" (*non-factuel*)
- "John souhaite pouvoir lire les lettres de Mary" (*irréel*)

(cf. [Leech 1974 : 302-303]). Cette distinction, très importante sur le plan sémantique, permet de rendre compte du caractère inacceptable de phrases comme : "Je ne me rends pas compte qu'il est parti" [Kiparsky 1970 : 349], ou : "Je ne sais pas que Brown est mort" [Barker 1974 : 33]. Elle est en partie parallèle à la distinction entre verbes factifs et non factifs.

Comme exemple de *présuppositions non lexicales*, on peut citer des présuppositions existentielles, comme dans les énoncés : "Le roi de France est en bonne santé", "Jean écrit de la poésie dans son bureau" [Kiefer 1973 : 87, 107-108].

La distinction entre *présuppositions idiosyncrasiques* et *présuppositions universelles* (ou *générales*) est moins tranchée que les dichotomies précédentes. En principe, dans les cas de présuppositions lexicales, les présuppositions universelles sont celles qui appartiennent "aux définitions et aux axiomes de la sémantique du lexique d'une langue donnée", alors que les présuppositions idiosyncrasiques sont particulières à un item lexical [Kiefer 1973 : 99]. Par exemple, les verbes "craindre" et "espérer" ont tous deux comme présupposition que leur sujet est humain (présupposition universelle, qu'ils partagent avec de nombreux autres verbes) ; mais "espérer" présuppose que le complément d'objet est ressenti comme "bon", alors que "craindre" présuppose qu'il est ressenti comme "mauvais" (présuppositions idiosyncrasiques). Ces présupposés lexicaux peuvent évidemment entraîner des restrictions de sélection (par exemple dans le choix des sujets et des compléments des verbes). Dans le cas de présuppositions non lexicales, les présuppositions idiosyncrasiques sont celles qui varient selon la langue dans laquelle l'énoncé est exprimé. Par exemple, en anglais : "John has lived in Princeton" présuppose que John est vivant, alors que : "John lived in Princeton" présuppose qu'il est mort [Chomsky 1972 : 65] (en français, l'usage du passé simple ou du passé composé n'entraîne pas de différences aussi marquées dans les présupposés). D'autre part, reprenant une phrase déjà utilisée (à d'autres fins) dans [Chomsky 1972 : 67] : "Jean est grand pour un Watusi", Ferenc KIEFER montre que sa traduction en suédois ou en hongrois *pose* la nationalité de Jean, alors que les versions anglaise ou française la *présupposent* [Kiefer 1973 : 106-107]. De même, Ryszard ZUBER fait remarquer que : "Il est trop jeune pour un ministre" présuppose : "Il est ministre", alors que : "Il est trop jeune pour être ministre" ne présuppose rien de tel. La présupposition citée semble due en français à la présence du verbe "être" ; dans d'autres langues, elle serait rendue par l'emploi de deux formes adverbiales différentes pour traduire "trop" [Zuber 1972b : 65-66]. C'est pourquoi cette distinction, due à Manfred BIERWISH pour les présuppositions lexicales, semble pouvoir s'appliquer également aux présuppositions non lexicales.

Dans le cas de présuppositions lexicales, le test pour distinguer entre présupposition générale et présupposition idiosyncrasique consiste à

enfeindre la règle de sélection liée à la présupposition en question. Si la phrase ainsi obtenue est inacceptable, c'est qu'il s'agit d'une présupposition universelle ; si elle est seulement bizarre, la présupposition enfreinte est idiosyncrasique. Par exemple [Kiefer 1973 : 100] :

- (a) *"Le rocher espère que Léna va venir"
- (b) *"Jean espère que Léna va venir et sent qu'il lui sera néfaste que Léna vienne"

Dans le premier énoncé, la présupposition enfreinte est universelle ; la phrase n'a pas de signification décelable (Un autre exemple serait : "Le vieux garçon est un protozoaire" [Kiefer 1973 : 114], où la violation d'une présupposition universelle conduit également à une phrase dépourvue de sens). L'énoncé (b), par contre, paraît seulement bizarre, et aurait pu être le fait d'un étranger connaissant mal le français ; la présupposition enfreinte est idiosyncrasique. Nous n'avons pas trouvé de test comparable qui soit applicable aux présuppositions non lexicales (hormis la traduction dans d'autres langues, à condition de bien choisir celles-ci ...). D'autre part, le test pour présuppositions lexicales que nous avons décrit n'est malheureusement pas toujours efficace, ce qui "suggère qu'il y a une différence de degré plutôt qu'une différence de nature entre les deux présuppositions [universelle et idiosyncrasique]" [Kiefer 1973 : 101].

La définition des *présuppositions du premier ordre* (et d'ordres supérieurs) se fonde sur la notion d'*enchaînement* des présuppositions. Soient par exemple les trois énoncés [Lakoff 1970 : 35] :

- (a) "Peu d'hommes ont cessé de battre leur femme"
- (b) "Il y a des hommes qui ont cessé de battre leur femme"
- (c) "Il y a des hommes qui ont battu leur femme"

En première analyse, on peut dire que (a) présuppose (b), et que (b) présuppose (c). En outre, dans cet exemple, (a) présuppose également (c) (nous aborderons ultérieurement le problème de la transitivité dans les enchaînements de présuppositions). Nous pouvons dire par conséquent que (a) *présuppose immédiatement* (b), alors que (a) présuppose, mais pas immédiatement, (c). On appellera (b) : *présupposition du premier ordre* de (a), et (c) : *présupposition du deuxième ordre* de (a). Plus formellement, si x , y , et z sont des phrases quelconques, on dira que x présuppose immédiatement y si et seulement si x présuppose y , et qu'il n'existe pas de z tel que x présuppose z et z présuppose y [Lakoff 1970 : 37]. Alors, z est une présupposition du premier ordre de x .

Deux tests sont proposés par George LAKOFF pour identifier les présuppositions du premier ordre. Tous deux consistent à ajouter à l'énoncé un membre de phrase, qui affecte différemment la présupposition du premier ordre et celles des ordres plus élevés. Le premier test est l'*adjonction d'un syntagme "qualificateur"*, qui peut être en anglais de la forme : "if any at all have", "or maybe none at all are", etc. [Lakoff 1970 : 35-37]. Soit l'énoncé (a) ci-dessus ; l'adjonction d'un syntagme "qualificateur" donne : "Peu d'hommes ont cessé de battre leur femme, si même quelques-uns ont cessé". Le nouvel énoncé ainsi obtenu est acceptable ; mais, contrairement à (a), il ne présuppose plus qu'il y ait des hommes qui aient cessé de battre leur femme ((b)), tout en continuant à présupposer qu'il y a des hommes qui ont battu leur femme ((c)). Le second test est l'*adjonction d'un commentaire infirmant* (comme : "il est bizarre, surprenant, curieux, etc., que ...") pouvant porter soit sur la phrase entière, soit sur l'une de ses présuppositions ; lorsque le commentaire porte sur une présupposition d'ordre deux, ou d'un ordre plus élevé, l'ensemble apparaît comme illogique. Par exemple [Lakoff 1970 : 38] :

- (d) "Peu d'hommes ont cessé de battre leur femme, et il est curieux que quelques-uns même aient cessé"
- (e) *"Peu d'hommes ont cessé de battre leur femme, et il est curieux que quelques-uns même l'aient jamais battue"

Dans l'énoncé (d), le commentaire porte sur la présupposition du premier ordre (b) de l'énoncé (a); la phrase est acceptable. Par contre, dans l'énoncé (e), qui porte sur une présupposition du deuxième ordre (c), le commentaire semble ne pas découler logiquement de la première partie de l'énoncé. Notons qu'un commentaire confirmant (comme : "il est normal que ...", "on peut s'attendre à ce que ...") n'est acceptable que s'il porte sur l'ensemble de l'énoncé ; même appliqué à une présupposition du premier ordre, il donne lieu à un énoncé illogique (cf. [Lakoff 1970 : 39 n3]). Nous reviendrons sur cette analyse à la fin du § 2.1.1.d.

c) Les essais de formalisation de la relation de présupposition

Pour formaliser les relations sémantiques entre les énoncés déclaratifs, les linguistes prennent en général pour modèle les relations logiques entre propositions. D'où les essais pour appliquer aux langues naturelles des concepts logiques tels que : équivalent, contradictoire, négation, conjonction, disjonction, implication. Nous verrons au paragraphe suivant l'usage que les linguistes font de ces notions. En ce qui concerne la présupposition, la logique classique ne fournit pas de relation qui lui corresponde de façon relativement satisfaisante. C'est pourquoi il est nécessaire tout d'abord d'étudier les rapports qui existent entre les valeurs de vérité d'un énoncé, et celles de ses présupposés (ou de sa présupposition du premier ordre).

Dès l'antiquité, les philosophes ont été embarrassés par le problème du lien entre la valeur logique d'une proposition, et celle de sa présupposition du premier ordre. Ainsi, EUBULIDE de Milet proposait le raisonnement suivant (dit : *sophisme du cornu* : ὁ κερᾶτικός, parfois attribué également à DIODORE CRONOS, et à CHRYSIPPE) : "Tu n'as pas perdu tes cornes ; donc, tu as des cornes" (cf. [Diogène Laërce : I, 142, 289 n234 ; II 117], [Mallet 1845 : 51-52], et [Makovevski 1967 : 79]), présupposant que : "ce que l'on n'a pas perdu, on le possède toujours". La notion de présupposé n'ayant pas été dégagée avant la fin du dix-neuvième siècle, il n'était pas possible de poser le problème en termes de relation de présupposition ; c'est pourquoi ce raisonnement fallacieux était considéré comme un syllogisme dont la majeure (ici : le

présupposé) est fautive. Dans une optique voisine, la *Logique de Port-Royal* considère le posé et le présupposé comme deux énoncés conjoints (formant une proposition composée), dont la composition n'est pas expressément marquée ; il est donc nécessaire, pour en extraire le sens, d'explicitier ces deux énoncés (cf. [Arnault 1683 : 177, 184-190]). De même, KANT aborde le problème des jugements exponibles, "qui impliquent simultanément, bien que de façon cachée, une affirmation et une négation" (ex.: "Peu d'hommes sont savants" = "Beaucoup d'hommes ne sont pas savants" & "Quelques hommes sont savants") ; mais c'est pour conclure que ce problème "relève non de la logique, mais de la grammaire" [Kant 1800 : 120 § 31].

FREGE a été conduit à dégager la notion de présupposition à partir de celle de *dénotation*. La dénotation d'un mot ou d'une expression, à distinguer de sa signification, est l'objet que ce mot ou cette expression sert à désigner (son *réfèrent*). Un même objet peut être désigné par plusieurs mots ou expressions (ex. : "Vénus" = "L'étoile du matin" = "L'étoile du soir"), et un mot ou une expression peut avoir un sens sans avoir de dénotation (ex. : "la suite qui converge le moins rapidement") [Frege 1892 : 102-108]. Si l'on applique cette distinction à une proposition affirmative prise comme un tout, la signification de la proposition est la pensée qu'elle exprime, et sa dénotation est sa valeur de vérité [Frege 1892 : 108-112]. Lorsque la proposition comporte une relative n'ayant pas pour antécédent un nom propre (ou un terme équivalent), cette relative n'a pas pour dénotation une valeur de vérité, mais un objet ; elle est interchangeable avec un nom propre sans altération de la valeur de vérité de la proposition, et elle peut par conséquent être considérée comme un nom propre composé. Sur l'exemple célèbre : "Celui qui a découvert la forme elliptique des orbites planétaires est mort dans la misère" (où la relative désigne KÉPLER), FREGE montre que "quand on énonce une affirmation, on suppose toujours sans le dire que les noms propres y figurant, qu'ils soient simples ou composés, ont une dénotation" [Frege 1892 : 115] (le texte original porte : "Wenn man etwas behauptet, so ist immer die Voraussetzung selbstverständlich, daß [...]"; certains auteurs préfèrent traduire : "Voraussetzung" par : "présupposition" plutôt que par : "supposition" cf. [Garner 1971 : 27], et [Zuber 1972b : 9]). Cette présupposition, selon laquelle le mot KÉPLER désigne effecti-

vement quelqu'un, vaut tout autant pour la négation : "KÉPLER ne mourut pas dans la misère", que pour l'affirmation. Lorsque la présupposition de l'existence d'une dénotation est fautive (lorsqu'au lieu de désigner effectivement un objet, elle "en éveille seulement l'illusion sans rien dénoter en fait" [Frege 1892 : 116]), la proposition n'est ni vraie ni fautive. Soit par exemple l'énoncé : "Après la prise du Schleswig-Holstein au Danemark, la Prusse et l'Autriche se brouillèrent". Pour un étranger ayant une connaissance médiocre de l'histoire européenne, et qui tiendrait pour faux que le Schleswig-Holstein ait été pris au Danemark, l'élément de phrase : "après la prise du Schleswig-Holstein" n'aurait pas de dénotation ; en conséquence, la proposition toute entière ne serait pour lui ni vraie ni fautive [Frege 1892 : 118 n 1]. Autrement dit, et en généralisant le raisonnement de FREGE, si les présupposés sont faux, la proposition n'a pas de valeur de vérité : elle est *indéterminée* (certains auteurs considèrent même que lorsque la présupposition est fautive, FREGE estime qu'il n'y a pas assertion, mais seulement énonciation d'une forme verbale ; cf. [Garner 1971 : 28]. En pratique, cette thèse conduit également à l'absence de valeur de vérité pour la proposition considérée).

Bertrand RUSSEL a défendu un point de vue différent, sur son exemple également célèbre : "Le roi de France est chauve" ; le fait qu'au moment où RUSSEL écrit, il n'existe pas de roi de France, rend automatiquement la proposition fautive [Russel 1905b : 46-47]. L'argumentation de RUSSEL contre FREGE est bien connue : les expressions ayant un même référent ne sont pas toujours interchangeables ("Scott est-il l'auteur de *Waverley* ?" \neq "Scott est-il Scott ?") ; la loi du tiers exclu impose que soit vraie l'une des affirmations contradictoires : "Le roi est chauve", "Le roi n'est pas chauve" ; enfin affirmer ou nier l'existence d'un individu ou d'un objet peut conduire à des contradictions [Russel 1905b : 47-48]. Pour notre propos, il paraît plus important d'examiner plus en détail quelles valeurs de vérité RUSSEL attribue à une proposition selon que ses présupposés sont vrais ou faux. L'exemple ci-dessus peut être paraphrasé en : "Une et une seule entité a la propriété d'être actuellement roi de France, et cette entité a la propriété d'être chauve" (cf. [Russel 1905b : 53]). KIEFER propose la formalisation de cette proposition en [Kiefer 1973 : 86] :

$$\exists x (Rx \ \& \ \forall z, \forall y ((Rz \ \& \ Ry) \ \rightarrow \ (z \leftrightarrow y)) \ \& \ Cx)$$

avec :

Rx = "x est actuellement roi de France"

Cx = "x est chauve"

Dans cette représentation de l'énoncé, il y a par conséquent trois termes conjoints : l'affirmation qu'un individu est roi de France, l'affirmation que le roi de France est unique, et l'affirmation que ce même individu est chauve. Pour que la proposition soit fausse, il suffit que l'un des termes conjoints soit faux (c'est-à-dire : qu'il n'y ait pas actuellement de roi de France, ou qu'il y en ait plusieurs, ou que le roi de France ne soit pas chauve). Ainsi que nous l'avons déjà noté à propos de la portée de la négation, l'énoncé : "Le roi de France n'est pas chauve" peut recevoir deux interprétations :

$$\left\{ \begin{array}{l} A = \overline{\exists x (Rx \ \& \ \forall z, \forall y ((Rz \ \& \ Ry) \ \rightarrow \ (z = y)) \ \& \ Cx)} \\ B = \exists x (Rx \ \& \ \forall z, \forall y ((Rz \ \& \ Ry) \ \rightarrow \ (z = y)) \ \& \ \overline{Cx}) \end{array} \right.$$

La première se trouve être vraie, la seconde fausse [Russel 1905b : 53].

S'il est assez facile de formuler des objections théoriques contre ces deux thèses (cf.p.ex. : [Strawson 1964 : 86-87]), il n'est pas aisé de déterminer laquelle présente le moins d'inconvénients pratiques. La conception de FREGE laisse subsister des lacunes dans la table de vérité, ce qui limite les possibilités de calcul ; en outre, en limitant la signification de l'énoncé au seul posé [Frege 1892 : 116], elle chasse le présupposé du contenu de la proposition, pour en réduire la fonction à celle d'une simple condition d'énonciation. D'autre part, la conception de RUSSEL rejoint celle de la *Logique de Port Royal*, qui met sur le même plan posé et présupposé (puisque la conjonction est symétrique). Or, les analyses qui précèdent montrent qu'il est souhaitable qu'un modèle logique de la présupposition n'accorde pas un statut identique à : "Le roi de France est unique s'il existe", "Untel est roi de France", et "Untel est chauve". De plus, RUSSEL attribue un poids excessif au critère d'existence d'un référent à ce dont on parle, par rapport à ses effets dans le langage courant. Ainsi, tout énoncé général de la forme : "le x qui satisfait ϕx satisfait ψx " (ex. : "le x qui est roi de France est

chauve") présuppose l'existence de phrases de la forme " ϕa " et " ψa " dans lesquelles a est un nom (ex. : "Louis IX est roi de France", "Louis IX est chauve") [Russel 1940 : 110]. Si ce nom ne désigne pas un individu réel, il peut avoir un sens, mais il ne dénote rien ; tel est le cas par exemple de noms comme Apollon ou Priam, ou d'expressions comme : "l'actuel roi de France". Toute phrase dans laquelle entre un nom d'individu (ou de classe d'individus) sans dénotation est fautive. On ne peut donc rien énoncer de vrai concernant un personnage mythologique, et la proposition " x est un Centaure" est fautive quel que soit x [Russel 1905a : 100]. RUSSEL défend sa théorie des descriptions en affirmant à juste titre que, pour des besoins techniques, il est indispensable de disposer de langages techniques, différant de ceux de la vie quotidienne [Russel 1957 : 124]. Mais dans un champ d'application qui concerne justement l'analyse des langages de la vie quotidienne, trop de différences constitue un inconvénient. Ainsi, la présupposition qu'une personne est vivante diffère de la présupposition qu'elle existe [Kiefer 1973 : 105] ; *il y a* des choses qui n'*existent* pas réellement, comme Zeus, ou le bateau dont on a rêvé [Vendler 1967 : 133]. On prend, dans la vie courante, le risque de "parler pour ne rien dire", en laissant la vérité ou la fausseté de ce que l'on dit subordonnée en quelque sorte à la vérification ultérieure de ce qui est présupposé.

Ces considérations pratiques nous amènent à conclure que le modèle logique de la présupposition présenté par RUSSEL ne correspond pas aux besoins de notre analyse. D'ailleurs, à la suite de STRAWSON (cf. [Strawson 1950 ; 1964] sur ses deux descriptions successives de la présupposition), la grande majorité des linguistes a adopté la perspective de FREGE (parfois, il est vrai, avec certains accommodements). Nous l'avons vu, cette thèse laisse indéterminée la valeur logique d'un énoncé dont l'un des présupposés (au moins) est faux. Il est intéressant toutefois de définir les inconvénients d'un modèle inspiré de FREGE, mais ne présentant pas de "trous logiques". Soient deux énoncés A et B ; on peut exprimer la thèse de FREGE en disant que : *A présuppose B* signifie que si A est vrai, B est vrai, et si A est faux, B est toujours vrai [Leech 1974 : 293]. Il est possible d'étendre ce point de vue en disant que : "A présuppose B" signifie que l'on a à la fois : "A implique B" et : "non A implique B". Soit

(si l'on pose : A psp B = "A présuppose B") :

$$(A \text{ psp } B) \stackrel{\text{df}}{=} (A \rightarrow B) \ \& \ (\bar{A} \rightarrow B)$$

qui vérifie par conséquent :

$$(A \text{ psp } B) \rightarrow B.$$

La relation ainsi définie est :

- irreflexive (A psp A n'est ni une tautologie, ni une contradiction) ;

- asymétrique (on n'a ni : (A psp B) \rightarrow (B psp A) , ni :
(A psp B) \rightarrow $\overline{(B \text{ psp } A)}$ quels que soient A et B) ;

- transitive (on a : ((A psp B) & (B psp C)) \rightarrow (A psp C)
quels que soient A , B , et C).

Ces propriétés du modèle correspondent-elles aux propriétés de la relation de présupposition telle qu'elle se présente dans les langues naturelles ? L'irreflexivité paraît fondée, puisque, si la présupposition est en général anti-réflexive (puisque la plupart des énoncés ont pour contenu à la fois un posé et des présupposés), il lui arrive d'être réflexive (dans le cas des présupposés existentiels, et aussi dans les énoncés redondants du genre : "John aime la fille qu'il aime" cf. [Keenan 1971 : 49]). De prime abord, l'antisymétrie aurait mieux convenu que l'asymétrie ; mais l'antisymétrie aurait entraîné l'antiréflexivité. Par contre, la transitivité est en désaccord avec certaines observations de LAKOFF, qui a relevé des exemples de phrases complexes dans lesquelles la transitivité de la présupposition semblait ne pas jouer. Ainsi [Lakoff 1970 : 41] : "Si Harry était parti, Bill s'en serait rendu compte" présuppose que Harry n'est pas parti, alors que le verbe "se rendre compte" entraîne la vérité de son complément : "Harry est parti". LAKOFF cite également : "Max faisait semblant de se rendre compte qu'il était malade", "Si Sue avait eu des ennuis, Irv aurait fait semblant d'être malade", "J'ai demandé à Sam de faire semblant d'être malade", etc. [Lakoff 1970 : 40-53]. A l'examen, ces exemples ne sont pas sans analogie avec le *paradoxe du menteur*

(ὁ ψευδόμενος : "Je ne dis jamais la vérité", ou : "Epiménide le Crétois dit que les Crétois mentent toujours", attribué à EUBULIDE de Milet) ; en effet, tous les exemples de Lakoff contiennent un élément non factuel, introduit soit par une interrogative indirecte, soit par une conditionnelle irréaliste, soit par un verbe antifactif ("faire semblant de"). A la limite, le raisonnement de LAKOFF pourrait être étendu à la négation métalinguistique des énoncés, cas dans lequel on montrerait sans difficulté que la relation de présupposition est toujours anti-transitive. On peut au contraire admettre avec DUCROT que "les propositions conditionnelles - qu'elles soient potentielles ou irréelles - ont cette propriété générale d'annuler le présupposé de la principale lorsqu'il leur est identique" [Ducrot 1972 : 187]. Ainsi, dans l'énoncé : "Si Pierre avait été à Paris, il serait resté", la proposition principale introduit le présupposé : "Pierre était à Paris", qui est contredit par la subordonnée, et de ce fait annulé. On peut également faire appel au principe dit de *projection des présupposés*, selon lequel les présupposés des subordonnées ne se combinent pas avec ceux de la principale, mais constituent plutôt des présupposés de l'ensemble de l'énoncé (cf. [Langendoen 1971b : 57]). Il nous paraît plus conforme à l'usage de la langue de rendre compte, de l'une ou l'autre de ces manières, des exceptions apparentes à la règle de transitivité de la présupposition ("exceptions" d'ailleurs rares ; cf. [Lakoff 1970 : 56]).

Si ce modèle logique de la présupposition apparaît, somme toute, comme acceptable sur le plan pratique, il présente cependant un inconvénient majeur sur le plan théorique. Il est en effet facile de démontrer que, quelles que soient les propositions A et B, si B est vrai, alors A présuppose B :

$$B \rightarrow (A \text{ psp } B)$$

(plus généralement, la valeur logique de la relation est équivalente à celle du présupposé : $(A \text{ psp } B) \leftrightarrow B$). Cette propriété conduit évidemment à des absurdités sur le plan pratique : toute proposition doit compter parmi ses présupposés non seulement l'ensemble des tautologies, mais également n'importe quelle proposition se trouvant être vraie. Ce point est une conséquence des "paradoxes" de la relation d'implication, à partir de laquelle nous avons

défini la relation de présupposition (Nous reviendrons au paragraphe suivant sur ces "paradoxes"). Il nous paraît justifier le rejet de cette formalisation.

La formalisation adoptée par de nombreux auteurs à la suite de FREGE n'est guère plus satisfaisante nous semble-t-il. Elle repose sur la thèse qu'un énoncé dont la présupposition est fautive ne peut avoir de valeur de vérité (la question ne se pose pas [Strawson 1950 : 12] ; l'énoncé est hors de propos [Geach 1950 : 34] ; l'énonciation est nulle et non avenue [Austin 1962 : 51]), ou que cette valeur est indéterminée ([Vendler 1967 : 132], [Katz 1972 : 143]). LEECH résume cette conception en disant que A présuppose B si les deux conditions suivantes sont satisfaites [Leech 1974 : 292] :

- si l'énonciation de A est valide (pourvue de sens), alors B est nécessairement vrai ;

- si B est faux, alors l'énonciation de A est invalide ou vide (nulle, dépourvue de sens).

(En termes de logique bivalente, on pourrait dire que la vérité de la validité de A implique la vérité de B). Pour que l'on puisse dire si A est vrai ou faux, il faut au préalable que son énonciation soit valide. Ce corollaire interdit une formalisation du type :

$$\left\{ \begin{array}{l} (A \text{ psp } B) \rightarrow ((A \ \& \ B) \rightarrow A) \\ (A \text{ psp } B) \rightarrow ((\bar{A} \ \& \ B) \rightarrow \bar{A}) \end{array} \right.$$

car l'expression à droite du premier signe d'implication est une tautologie ; elle ne dépend pas par conséquent de la relation de présupposition. Il nous paraît nécessaire, pour que cette conception ait un sens, d'introduire une distinction entre l'énoncé tout entier et le posé, et d'écrire : A p C pour : A pose C. On aurait alors :

$$\left\{ \begin{array}{l} ((A \text{ psp } B) \ \& \ (A \text{ p } C)) \rightarrow ((B \ \& \ C) \rightarrow A) \\ ((A \text{ psp } B) \ \& \ (A \text{ p } C)) \rightarrow ((B \ \& \ \bar{C}) \rightarrow \bar{A}) \end{array} \right.$$

Ainsi, la valeur de A reste indéterminée quand B est faux ; quand B est vrai, elle dépend également de la valeur de C . Il est facile de constater que les formalisations inspirées de FREGE conduisent, comme la conception précédente, à affirmer que tout énoncé présuppose toutes les tautologies. En outre, ces formalisations présentent un inconvénient supplémentaire, qui est d'autoriser des lacunes dans la table de vérité.

Cet examen rapide des représentations logiques possibles de la relation sémantique de présupposition aboutit au rejet de modèles trop simples fondés sur le calcul classique des propositions. D'autres directions de recherche restent ouvertes, dont certaines commencent à être explorées. Une tentative intéressante pour concilier la thèse fregéenne avec la loi du tiers exclu de la logique classique, est la formalisation de Bas C. VAN FRAASSEN. Cet auteur définit un langage formel L comportant les opérateurs de négation, de disjonction, et d'implication, une valuation binaire vrai/faux, et un opérateur de "nécessitation" : $A \text{ n } B$ tel que : A nécessite B si et seulement si chaque fois que A est vrai, B est vrai également [Van Fraassen 1968 : 138, 140]. La relation de nécessitation englobe les relations d'implication et de présupposition, mais ne s'y limite pas. Considérons un ensemble E de phrases vraies dans une situation donnée ; toutes ces phrases ont par définition des présuppositions vraies, car : $(A \ \& \ (A \ \text{n} \ B)) \ \rightarrow \ B$. Donc, si A est un élément de E , A est vrai ; si \bar{A} est un élément de E , A est faux ; dans les autres cas, A n'a pas de valeur définie. Ce modèle préserve la loi du tiers exclu ($A \vee \bar{A}$ est toujours vrai), mais non la bivalence [Van Fraassen 1968 : 142]. Par contre, le métalangage M qui permet de parler de L est lui strictement bivalent, et permet d'appliquer la logique classique à l'analyse des relations entre présupposition et valeur logique (cf. [Van Fraassen 1968 : 145-146]).

Parmi les autres directions de recherche pouvant conduire à des modèles logiques plus satisfaisants de notre point de vue, signalons pour conclure ce paragraphe :

- le recours aux opérateurs modaux (\square = "il est nécessaire que" ; \diamond = "il est possible que") pour rendre compte des relations de présupposition et d'implication. Cette solution a été envisagée par LAKOFF, qui suggère que

certaines formules comme $\Box (A/B) = \text{"A est nécessaire étant donné B"}$, soient interprétées comme signifiant : "A est nécessaire lorsque A présuppose B" [Lakoff 1970 : 53] (cf. également : [Lakoff 1970 : 93-106 ; 1971 a : 234-236]). De même, ZUBER, à partir de l'analyse de la portée de la négation, distingue entre fausseté restreinte (impossibilité) et fausseté contingente (non nécessité), et pose [Zuber 1972 b : 21-22] :

$$\left\{ \begin{array}{l} (A \text{ psp } B) \rightarrow (\bar{B} \rightarrow \Box \bar{A}) \\ (A \text{ psp } B) \rightarrow (\Diamond A \rightarrow B) \end{array} \right.$$

Mais ces recherches ne sont pas encore assez avancées pour que l'on puisse juger de leur valeur pratique.

- l'utilisation d'une logique plurivalente, permettant de conserver la conception de FREGE sans avoir de "trous logiques". Par exemple, dans une logique trivalente analogue à celle de ŁUKASIEWICZ, prenant les valeurs *vrai/indéterminé/faux*, on pourrait définir la relation de présupposition de telle sorte que l'on ait : $A \text{ psp } B = \text{vrai}$, si et seulement si : $B = \text{vrai}$, ou : $B = \text{faux}$ et $A = \text{indéterminé}$; et : $A \text{ psp } B = \text{faux}$, si et seulement si : $B = \text{faux}$ et : $A = \text{vrai}$ ou $A = \text{faux}$ (Signalons que, dans une telle hypothèse, on ne peut avoir : $(A \text{ psp } B) = (A \rightarrow B) \& (\bar{A} \rightarrow B)$, au sens donné par ŁUKASIEWICZ à l'implication).

- la définition d'un modèle formel de la présupposition de manière proprement inductive, à partir des propriétés observées sur les couples d'énoncés. Par exemple, si l'on écarte les énoncés tautologiques et les couples d'énoncés équivalents, on peut chercher à définir une relation de présupposition qui soit antiréflexive, antisymétrique, et transitive (relation d'ordre strict). Une telle relation sort évidemment du cadre du calcul des propositions.

d) Les autres types de relations sémantiques entre énoncés.

Geoffrey LEECH énumère huit types de propriétés ou relations sémantiques de base [Leech 1974 : 84-87] :

- *Synonymie*. X et Y sont synonymes = X a la même valeur de vérité que Y . Ex. : "Je suis un orphelin" est synonyme de : "Je suis un enfant et je n'ai ni père ni mère".

- *Implication*. X implique Y = si X est vrai, Y est vrai ; si Y est faux, X est faux. Ex. : "Je suis un orphelin" implique : "Je n'ai pas de père".

- *Inconsistance* (exclusion réciproque). X est inconsistant avec Y (X et Y s'excluent mutuellement) = si X est vrai, Y est faux, et si Y est vrai, X est faux. Ex. : "Je suis un orphelin" est inconsistant avec : "J'ai un père".

- *Présupposition positive*. X présuppose positivement Y = tout locuteur qui énonce X considère que la vérité de Y est admise par l'interlocuteur. Ex. : "Votre père est-il chez lui ?" présuppose positivement : "Vous avez un père".

- *Présupposition négative*. X présuppose négativement Y = tout locuteur qui énonce X considère que la fausseté de Y est admise par l'interlocuteur. Ex. : "S'il avait un père, cela ne se passerait pas comme cela" présuppose négativement : "Il a un père".

- *Tautologie*. X est une tautologie = X est toujours vrai. Ex. : "Cet orphelin n'a pas de père".

- *Contradiction*. X est une contradiction = X est toujours faux. Ex.: "Cet orphelin a un père".

- *Irrégularité*. X est sémantiquement irrégulier = X est absurde (c'est-à-dire : n'a pas de valeur de vérité), car X présuppose une contradiction. Ex. : "Le père de l'orphelin boit énormément".

Dans la perspective qui est la nôtre, ces relations ne présentent pas toutes le même intérêt. Les plus importantes nous paraissent être la relation

de présupposition (déjà analysée) et la relation d'*implication*. Comme la présupposition, l'implication joue deux rôles très différents : un rôle *linguistique*, lié à l'utilisation dans le langage courant d'expressions comme : "si ..., alors ..." et ses synonymes ; et un rôle *métalinguistique*, dans le langage technique des logiciens et des sémanticiens. Mais tandis que le rôle proprement linguistique de la notion de présupposition est assez réduit, l'usage de l'implication est fréquent dans la langue quotidienne. C'est pourquoi il nous faut décrire brièvement tout d'abord les formes variées d'utilisation dans le langage courant d'expressions conditionnelles ; ceci nous conduira à distinguer de l'implication deux formes plus faibles de relations entre énoncés : la *supposition*, et la *suggestion*. Nous pourrons ensuite présenter les tentatives qui ont été faites pour formaliser la relation d'implication entre énoncés.

Les relations conditionnelles ont probablement suscité autant de recherches que la relation de présupposition. Il est vrai qu'ARISTOTE n'a pas, en dépit de ses intentions, développé l'analyse du "syllogisme hypothétique" (cf. *Premiers Analytiques*, I, 44 [Aristote : 50 a 15 - 50 b 4]). Par contre, la logique stoïcienne a accordé une place centrale au jugement hypothétique (τὸ συνημμένον) : "si A, alors B". Et il mérite d'être souligné que PHILON de Mégare proposait du jugement hypothétique une table de validité qui est exactement la table de vérité de l'implication. Dans la phrase : "S'il fait jour, je parle", le jugement exprimé n'est incorrect que s'il fait jour et que je ne parle pas, c'est-à-dire si le premier terme est vrai et le second faux ; il est correct par contre dans les trois autres cas possibles, y compris s'il fait nuit et que je parle ("le faux implique le vrai"). SEXTUS EMPIRICUS, à qui l'on doit la relation de la thèse de PHILON, lui oppose les objections de DIODORE CRONOS, pour conclure que le problème de la validité de cette forme de raisonnement "échappe à la compréhension" [Sextus Empiricus HP : II 114] (cf. également [Sextus Empiricus AM : VIII 109-115], [Mallet 1845 : 127-137], et [Makovelski 1967 : 221-247]). La difficulté de l'analyse de la relation d'implication entre énoncés dépend beaucoup, comme nous allons le voir, de l'ambiguïté des langues naturelles.

Ainsi, étudiant la relation linguistique liée à l'expression : "Si ... alors ...", et qu'il appelle relation de *supposition*, Oswald DUCROT décrit cinq formes différentes de *si* [Ducrot 1972 : 169-178] :

- *standard*, correspondant à l'expression des relations de causalité, de condition nécessaire, de condition suffisante, ou de condition nécessaire et suffisante, selon les cas ; ex. : "Si Pierre vient, Jacques partira".

- *oppositif*, où "la supposition permet de situer l'affirmation qui la suit dans le mouvement général du discours" ; ex. : "S'il a de l'esprit, il a aussi du coeur" (paraphrasable en : "Il est bon de signaler ses qualités morales, parce qu'on risquerait de lui attribuer seulement des qualités intellectuelles").

- *contrastif*, qui "semble également destiné à justifier, toute question de vérité mise à part, l'acte d'affirmation accompli dans la [proposition] principale", mais en mettant l'accent sur la forme de l'assertion plus que sur son contenu ; ex. : "Si la Cité est le coeur de Paris, le quartier latin en est l'âme" (paraphrasable en : "Si tu admets une façon de parler comme *La Cité est le coeur de Paris*, alors tu dois admettre aussi qu'on dise *Le quartier latin en est l'âme*").

- *prépositionnel*, qui "permet d'annuler les présupposés, de la principale en les introduisant, sous forme d'hypothèse, dans la subordonnée" ; ex. : "Si Pierre est à Paris, il y restera certainement".

- *préventif* (non dénommé par DUCROT), qui prémunit le locuteur contre l'objection éventuelle de l'interlocuteur que cela ne le concerne ou ne l'intéresse pas ; ex. : "Si tu as soif, il y a de la bière dans le frigidaire".

L'auteur subsume ces divers usages de la supposition en considérant que l'énonciation de : "Si A, alors B" est toujours un acte qui demande à l'interlocuteur d'accepter provisoirement la proposition A comme cadre du discours (et plus spécifiquement de B). On perçoit mieux, à partir de cette description, la nature des rapports entre le supposé et le présupposé. Le présupposé est imposé à l'interlocuteur : celui-ci se voit contraint à l'accepter par les conventions du

discours, et en outre cette contrainte est implicite. Le supposé, par contre, est toujours explicitement proposé à l'interlocuteur, et présenté comme une hypothèse récusable [Ducrot 1972 : 179-185].

La relation de *suggestion* a toutes les apparences d'une relation d'implication entre énoncés, comme il apparaît dans les exemples suivants :

- Ⓐ "Personne d'autre que Pierre n'est venu" suggère : "Pierre est venu" [Leech 1974 : 319] ;
- Ⓑ "Si nous étions mariés, nous serions heureux" suggère : "Nous ne sommes pas heureux" [Leech 1974 : 323] ;
- Ⓒ "Sa jambe de bois n'a pas empêché John de danser avec Marie" suggère : "John a dansé avec Marie" [Karttunen 1973 : 75] ;
- Ⓓ "John fut en mesure de partir de bonne heure" suggère : "John est parti de bonne heure" [Karttunen 1973 : 76].

En effet, on serait tenté de dire que l'énoncé Ⓐ implique que Pierre est venu. Toutefois, LEECH souligne que, s'il est vrai qu'un énoncé comme : "Seul Pierre est venu" présuppose bien la venue de Pierre, il n'en est pas de même pour l'énoncé Ⓐ, qui pourrait être suivi d'une remarque telle que : "... et d'ailleurs, Pierre n'est pas venu non plus". De même, KARTTUNEN note que, si on lit l'énoncé Ⓒ sans autre contexte, "on garde probablement l'impression que John a dansé avec Marie en dépit de sa jambe de bois". Or, les verbes tels que : "empêcher", "dissuader", "décourager", "s'abstenir" n'ont de valeur implicative que sous leur forme affirmative ; la preuve en est justement fournie par le fait que Ⓒ "peut, sans aucune contradiction, être enchâssé dans un contexte où il devient clair que John n'a pas dansé avec Marie [...]. Néanmoins, en l'absence de toute preuve contraire, Ⓒ semble bien suggérer que John a dansé avec Marie" [Karttunen 1973 : 69-70, 75-76].

La relation d'*implication* entre énoncés apparaît, à la suite de cette analyse, comme un cas particulier du *si standard* de DUCROT. Au sens, encore relativement imprécis, de la définition de LEECH citée au début de ce paragraphe, l'usage de la relation d'implication dans le langage courant prend des formes variées. Par exemple :

- Ⓐ "Si vous tondez mon gazon, je vous donnerai cinq dollars" [Karttunen 1973 : 76-77] ;
- Ⓑ "S'il pleut, je resterai" [Ducrot 1973 : 136] ;
- Ⓒ "Je ne viendrai pas sans que tu en sois averti" [Ducrot 1973 : 139] ;
- Ⓓ "Encore une bière et je pars" [Lakoff 1970 : 10] ;
- Ⓔ "Si les animaux y sont interdits, les chiens y sont interdits".

Les quatre premiers exemples concernent des relations d'implication entre phrases. Une analyse un peu poussée des exemples Ⓐ et Ⓑ montrerait que la relation d'implication entre énoncés tend, en français comme en anglais, à être interprétée comme une condition non seulement suffisante, mais également nécessaire [Ducrot 1973 : 136], comme s'il existait une tendance naturelle de l'esprit humain à parachever les conditionnelles en biconditionnelles [Karttunen 1973 : 76]. L'énoncé Ⓒ est intéressant en tant que variante de : "si je viens, tu en seras averti" ; l'énoncé Ⓓ, en tant que dépendant fortement du contexte pour son interprétation ("Si je bois encore une bière, alors je partirai" ; "Si tu bois encore une bière, alors je pars" ; "Si tu me verses encore une bière dans le cou, je m'en vais" ; etc.). On aura un aperçu de la variété de l'expression de la relation d'implication entre énoncés dans le langage courant en français en consultant : [Blum 1971 : 85] et [Coyaud 1972 : 46, 92]. Le dernier exemple exprime une relation d'implication simplement lexicale, fondée sur la relation sémantique d'*hyponymie* (cf. [Leech 1974 : 100-102] ; [Lyons 1978 : 236-239]), puisque "chien" est hyponyme d'"animal" ; mais on trouverait également dans le langage courant l'expression d'implications fondées sur des relations d'*homonymie* (ex. : "Il doit collectionner les timbres, puisque c'est un philatéliste"). Bien que la relation d'implication présente certains traits

communs avec la relation de présupposition, il est possible de les distinguer. Par exemple, dans le couple d'énoncés :

- { (a) "Nixon a refusé d'essayer de faire taire Agnew"
 (b) "Nixon n'a pas essayé de faire taire Agnew"

(a) implique (b) , mais ne le présuppose pas [Lakoff 1970 : 45]. D'ailleurs, dans le dialogue, le désaccord de l'interlocuteur avec ce qui est impliqué par l'énoncé ne s'exprime pas de la même manière que son désaccord avec ce que l'énoncé présuppose, ce qui peut constituer un critère permettant de les différencier (cf. [Ducrot 1973 : 213-217]).

Au niveau métalinguistique, nous citerons tout d'abord l'exemple bien connu d'AUSTIN : "Si « Le chat est sur le paillason » entraîne « Le paillason est sous le chat » , alors « Le paillason n'est pas sous le chat » entraîne « Le chat n'est pas sur le paillason » " [Austin 1962 : 49], qui rappelle la règle de contraposition de la relation logique d'implication. Pour l'usage métalinguistique de cette relation entre énoncés, on pourrait encore citer BARKER, disant qu'en affirmant : "Marie ne sait pas que Brown est mort", le locuteur implique que Brown est mort [Barker 1974 : 34] ; ou DUCROT analysant le raisonnement de PASCAL sur la grâce suffisante dans la *Seconde Provinciale*, ou celui de MONTESQUIEU sur l'esclavage dans l'*Esprit des Lois* [Ducrot 1973 : 179-208]. Pour l'usage métalinguistique de cette relation entre lexèmes (*hyponymie*), on citera seulement l'exemple de CARNAP pour introduire la notion de *postulat de sens* : "Si Jack est célibataire, alors il n'est pas marié" [Carnap 1952 : 222].

La formalisation de la relation d'implication ne pose pas de problème, en principe, au niveau du métalangage ; on peut penser en effet que la relation sémantique se confond alors avec la relation logique, comme c'est le cas au moins pour les exemples de AUSTIN, DUCROT et CARNAP cités ci-dessus. Au niveau du langage courant, on se heurte à deux types de difficultés. Le premier tient à certaines propriétés du modèle, qui ont d'ailleurs reçu le nom de "paradoxes de l'implication" :

- $\bar{A} \rightarrow (A \rightarrow B)$ = "le faux implique n'importe quoi" (*ex falso sequitur quodlibet*) ;

- $B \rightarrow (A \rightarrow B)$ = "le vrai est impliqué par n'importe quoi" (*verum sequitur ad quodlibet*).

Il est évident que ces propriétés s'écartent trop de celles observables sur les énoncés dans l'usage courant pour que ce modèle soit satisfaisant ; toutefois, on peut à la rigueur les considérer comme des paradoxes formels, sans incidence pratique sur le langage de la vie quotidienne. Par contre, le second type de difficultés est dû à un écart entre l'usage effectif et le modèle, sous certaines transformations. En particulier, la transformation négative n'a pas toujours le même effet sur la signification des énoncés. Par exemple [Ducrot 1972 : 181], alors que le premier énoncé ci-dessous vérifie la loi de contraposition (propriété du modèle logique : $(A \rightarrow B) \rightarrow (\bar{B} \rightarrow \bar{A})$), l'application de cette loi au second conduit à une absurdité :

- "Si Pierre vient, je le recevrai" implique logiquement : "Si je ne reçois pas Pierre, c'est qu'il ne sera pas venu" ;

- "Si Pierre vient, je ne le recevrai pas" implique logiquement : "Si je reçois Pierre, c'est qu'il ne sera pas venu".

Une explication possible de cet écart entre le modèle et les faits linguistiques est que la condition exprime ici une supposition, et non une implication (cf. [Ducrot 1972 : 182-183]).

L'adoption d'un modèle logique pour la relation d'implication nécessite deux démarches parallèles (et complémentaires) : l'analyse plus poussée des emplois de cette relation dans les langues naturelles, et des effets des diverses transformations sur le sens des énoncés correspondants ; et la recherche de modèles plus complexes, empruntés aux logiques modales ou plurivalentes. Dans la première perspective se situent des travaux comme ceux de DUCROT, sur les effets des transformations négative et interrogative appliquées aux énoncés exprimant une relation d'implication [Ducrot 1972 : 172-175 ; 1973 : 263-268]. Dans la seconde, on rencontre des essais de formalisation comme ceux

de LAKOFF, montrant qu'un énoncé de la forme "Si A , alors B" ne peut pas toujours être traduit par : $A \rightarrow B$, mais doit dans certains cas l'être par : $\Diamond (A \rightarrow B)$ [Lakoff 1970 : 6-9]. Les ressources de la logique modale ont attiré l'attention de certains linguistes ; toutefois, un modèle comme l'implication stricte de LEWIS (A implique strictement B = $A \times B = \Diamond (A \& \bar{B})$) présente l'inconvénient de laisser des "trous" dans la table de vérité. Peut être cet inconvénient spécifique des logiques modales pourrait-il être pallié en recourant à une logique trivalente comme celle de ŁUKASIEWICZ (en donnant aux valeurs la signification : *vrai/indéterminé/faux*). Toutefois, malgré leur grande richesse potentielle (si l'on a n valeurs de vérité, on a n^2 relations binaires distinctes possibles ; cf. [Blanché 1957 : 48]), les logiques plurivalentes ne semblent pas avoir suscité beaucoup d'intérêt chez les linguistes en quête de représentations formelles.

L'analyse des relations sémantiques entre expressions et entre énoncés est rendue difficile par les ambiguïtés des langues naturelles, comparées aux langages formels (cf. p.ex. : [Russel 1940 : 91-107], [Quine 1965 : *passim*], [Zuber 1972 b : 18-19], [Ducrot 1973 : 117-131], etc.). D'où certains désaccords entre les auteurs que nous avons cités. Soit par exemple les énoncés :

- (a) "Peu de filles vont venir"
- (b) "Il y a des filles qui vont venir"

Pour LAKOFF, (b) est la présupposition du premier ordre de (a) [Lakoff 1970 : 36], alors que pour LEECH, (b) est seulement suggéré par (a) [Leech 1974 : 294-295]. Le raisonnement de LEECH fait d'ailleurs appel aux mêmes arguments pour montrer sur cet exemple que la relation de suggestion est plus faible que la relation de présupposition (effets des commentaires infirmants ; cf. également [Leech 1974 : 318]) que celui de LAKOFF pour établir qu'il s'agit d'une présupposition du premier ordre. Toutefois, si les critères linguistiques d'identification des diverses relations sémantiques entre énoncés ne sont pas

encore assez précis pour lever ces ambiguïtés dans tous les cas observables, les analyses présentées dans les pages qui précèdent suffisent à montrer que, dans le contenu de tout énoncé, il convient de distinguer ce que cet énoncé pose, suppose, présuppose, suggère, et implique. Cette règle s'applique évidemment à l'analyse du contenu des énoncés déclaratifs que sont les réponses aux questions ; compte-tenu de la définition que nous avons donnée, au chapitre 1, de la forme canonique des énoncés à valeur de question, elle doit pouvoir s'appliquer également à l'analyse du contenu des énoncés interrogatifs eux-mêmes.

2.1.2. Les relations sémantiques entre questions, réponses, et présupposés.

C'est surtout à l'occasion d'analyses d'énoncés déclaratifs que le cadre théorique relatif à la variété des contenus possibles d'un énoncé (et, partant, aux diverses relations sémantiques entre énoncés) a été élaboré. Nous allons voir cependant que l'impact de ces théories sémantiques, comme d'ailleurs celui de certaines des théories syntaxiques présentées au chapitre 1, sur l'analyse des relations entre questions, réponses, et présupposés, est loin d'être nul. Il manque toutefois aux résultats que nous exposons au § 2.1.2.a, de s'intégrer dans une théorie linguistique globale. C'est pour pallier cet inconvénient que Jerrold J. KATZ a proposé un cadre théorique général (improprement intitulé : "logique des questions") d'analyse des énoncés interrogatifs. L'apport de la théorie de KATZ est important ; mais il paraît sans commune mesure avec celui des théories logiques des questions.

a) Les apports de la linguistique

Les travaux dont nous avons donné un aperçu dans les pages qui précèdent, et dans la seconde partie du premier chapitre, débouchent principalement sur des solutions partielles à deux problèmes importants, concernant les relations entre questions et réponses, et les composantes du contenu des énoncés interrogatifs (et en particulier le rôle des présupposés).

Le problème des relations entre questions et réponses est rarement abordé par les grammairiens, et, de ce point de vue, DAMOURETTE et PICHON d'une part, TESNIÈRE d'autre part, constituent des exceptions remarquables. La syntaxe structurale de TESNIÈRE précise en effet que l'on répond à une interrogation nucléaire en remplissant le nucléus vide sur lequel porte l'interrogation [Tesnière 1959 : 197 § 17] ; et à une interrogation connexionnelle en affirmant ou en niant l'existence de la connexion [Tesnière 1959 : 211 § 2], ce qui peut être également réalisé par la répétition des termes de la question sous forme énonciative, éventuellement assortie d'une négation [Tesnière 1959 : 215 § 26]. Si DUBOIS et DUBOIS-CHARLIER ne traitent pas explicitement du problème des réponses aux questions, il est toutefois assez facile de se représenter un mécanisme de réponse aux questions parallèle à celui décrit par TESNIÈRE. En effet, on peut obtenir une réponse entière à une question de structure profonde Σ par transformation de cette structure en une forme Σ' , dans laquelle le constituant Affir a été substitué au constituant Inter de Σ . En outre, s'il s'agit d'une interrogation totale, on peut avoir Const \rightarrow Inter + Nég ; s'il s'agit d'une interrogation partielle, au syntagme nominal avec démonstratif *que* (SN_{que}) doit être substitué un syntagme nominal standard, ne comportant pas de proforme nominale (ProN = "quelque chose", "quelqu'un", "de quelque manière", etc.). Ces déductions sont corroborées par les solutions proposées par d'autres formes de grammaires génératives inspirées de CHOMSKY (cf. pour l'anglais : [Katz 1964 : 143-144]). Enfin, il est clair que pour HARRIS, toute question contient, sous sa forme distinguée, l'énumération de l'ensemble de ses réponses entières possibles.

Toutefois, ces théories ne permettent pas, dans le cas général, de résoudre de façon satisfaisante le problème suivant : étant donné un énoncé interrogatif A, comment établir, pour tout énoncé déclaratif B, si celui-ci est ou n'est pas une réponse acceptable à la question A ? (Par : "acceptable", nous entendons provisoirement "en accord avec la conception intuitive qu'en ont les locuteurs de la langue considérée"). Remarquons tout d'abord que ni la grammaire structurale de TESNIÈRE, ni les formes de syntaxe générative que nous avons passées en revue, ne permettent de déterminer les *réponses entières* acceptables pour une question particulière donnée. Soit par exemple l'interrogation : "Qui

est venu ?" ; en appliquant les règles syntaxiques de construction des réponses entières, on obtiendrait des phrases comme : "Pierre est venu", "Jules César est venu", "Le désespoir est venu", "Un rocher est venu". Si l'on soumettait ces réponses à diverses personnes de langue maternelle française, selon la procédure des enquêtes ou des expérimentations linguistiques, il est probable que seule la première réponse serait acceptée comme telle par la quasi-totalité des locuteurs ; la seconde pourrait éventuellement l'être à certaines conditions (par exemple, que "Jules César" soit le surnom ou le nom d'une personne ou d'un animal vivant actuellement, ou que le thème général du discours soit l'histoire romaine) ; la troisième serait probablement rejetée par tous ceux (la majorité?) qui n'y verraient pas une métaphore autorisée par la licence poétique ; quant à la quatrième, elle serait vraisemblablement jugée absurde par tous les locuteurs. En dehors de la linguistique distributionnelle (qui constate plus qu'elle n'explique), aucune théorie purement morpho-syntaxique ne permet de rendre compte de l'inacceptabilité des réponses citées. C'est pourquoi les théories sémantiques inspirées des grammaires génératives font en général l'hypothèse d'existence de *traits* sémantiques caractérisant la signification d'un élément lexical, et de *restrictions de sélection*, qui déterminent les règles d'association entre lexèmes à partir de ces traits sémantiques (cf. p.ex. : [Chomsky 1965 : 95-96], [Katz 1966 : 224-239], [Leech 1974 : 141-146]). On rencontre d'ailleurs le même genre de difficulté dans les grammaires transformationnelles de type HARRIS. Il est vrai que l'on peut considérer comme réponse acceptable tout élément de la disjonction de la forme distinguée correspondant à l'interrogation ou toute paraphrase de cet élément. Mais la difficulté se trouve repoussée, ainsi que nous l'avons noté, au niveau des transformations paraphrastiques aboutissant à la forme distinguée (cf. la transformation φ_{Z6} dans l'exemple figuré sur le schéma n° 1.69, au § 1.2.4.e).

A plus forte raison est-il difficile de définir des règles syntaxiques pour l'identification, parmi les énoncés suivants, de ceux que l'on peut raisonnablement considérer comme des réponses à la question : "Est-ce que Pierre est venu ce matin ?". Soient les énoncés déclaratifs :

- Ⓐ "Si quelqu'un était venu, je l'aurais su"
- Ⓑ "Je ne sais pas qui est Pierre"
- Ⓒ "Le facteur est passé ce matin"
- Ⓓ "Personne n'est venu depuis hier"
- Ⓔ "Je n'ai pas été là de la matinée"
- Ⓕ "Il est venu un grand monsieur brun d'une trentaine d'années, portant moustaches, mais il n'a pas dit son nom"

Il est certainement clair pour toute personne connaissant suffisamment la langue française que les énoncés Ⓐ et Ⓓ équivalent à des réponses négatives à la question posée. De ce dernier, on peut en effet déduire que "Pierre n'est pas venu depuis hier". De l'énoncé Ⓐ, on peut parvenir à la même conclusion à condition de faire en outre l'hypothèse que, si l'interlocuteur *savait* que Pierre est venu, il le *dirait* ; or, ce point n'est pas lié à l'énoncé assertif lui-même, mais à des règles sociales que l'on suppose admises par tous. Les énoncés Ⓒ et Ⓕ sont peut-être des réponses à la question, si Pierre est le facteur dans le premier cas, et si la description concorde et est suffisante dans le second cas. Les énoncés Ⓑ et Ⓔ sont-ils des réponses ou des aveux d'ignorance ? Dans quelle mesure un aveu d'ignorance ne répond-il pas d'une certaine manière à la question posée ?

A notre connaissance, la syntaxe transformationnelle n'est pas armée pour résoudre le problème de l'identification de toutes les formes de réponses non entières à une interrogation donnée. Il est nécessaire, pour cela, de faire appel au *sens* des énoncés. Intuitivement, pour qui connaît bien la langue, il est presque toujours possible de dire si un énoncé répond totalement à une question, ou seulement partiellement, ou s'il n'y répond pas du tout ; voire, éventuellement, s'il suggère une réponse sans toutefois la donner directement. Mais ceci ne se peut que parce que l'on comprend le sens des énoncés. Il est alors possible de définir diverses catégories sémantiques de réponses à une question : réponses directes ou indirectes, totales ou partielles. A notre connaissance, peu de linguistes se sont préoccupés de définir les différentes sortes de réponses à une question ; à plus forte raison, peu de travaux proposent des procédures concrètes d'identification (ou de reconnaissance) de ces différentes sortes.

Cependant, une voie de recherche possible serait l'extension aux relations entre questions et réponses des résultats de travaux sur les "propos rapportés". En effet, de telles recherches ont pour objet les relations entre un discours source, et des énoncés présentés comme le compte-rendu de ce discours (cf. p.ex. : [Zwicky 1971]) ; relations comparables à celles que l'on peut définir entre une réponse entière à une question, et d'autres formes de réponses plus couramment utilisées dans l'usage quotidien. Nous verrons, au cours de ce chapitre, comment une analyse logico-sémantique des énoncés interrogatifs conduit à classer et identifier ces formes particulières de réponses. Par contre, il existe des cas, analogues aux exemples (b) et (e) ci-dessus, pour l'interprétation desquels il est nécessaire de faire appel soit au contenu implicite des énoncés, soit même au contexte de l'énonciation. Ce type d'information est, nous l'avons vu, subsumé sous le concept de présupposition.

Considérons tout d'abord l'énoncé assertif conditionnel : "Si Pierre vient, Jacques restera à Paris". Il se décompose en trois éléments :

- *Présupposé* : "Jacques est actuellement à Paris"
- *Supposé* : "Pierre viendra"
- *Posé* : "La venue de Pierre entraînera la présence ultérieure de Jacques à Paris"

Si nous considérons maintenant l'interrogation générale qui lui correspond : "Si Pierre vient, est-ce que Jacques restera à Paris ?", on voit qu'elle conserve les mêmes présupposé et supposé que l'assertion ci-dessus, mais que le posé et lui seul n'a plus le même statut : "dans l'interrogation, en effet, le posé est mis en doute, et le présupposé est maintenu" [Ducrot 1972 : 276]. Si l'on veut, on peut exprimer le posé de l'interrogation sous la forme :

Posé : "La venue de Pierre entraînera la présence ultérieure de Jacques à Paris, ou la venue de Pierre n'entraînera pas la présence ultérieure de Jacques à Paris".

Ce qui, du point de vue logique, est une tautologie : $A \vee \bar{A}$ (Ceci n'est toujours le cas que pour les questions à réponse par *oui* ou *non*).

Il est possible de décomposer plus avant l'énoncé interrogatif ci-dessus, en distinguant par exemple les présupposés existentiels du présupposé ("Jacques existe et est un homme vivant", "Paris existe et est une ville"), les présupposés du supposé ("Pierre existe et est un homme vivant", "Pierre peut se déplacer", "Pierre n'est pas ici"), etc. Henry S. LEONARD a proposé diverses distinctions dans les présupposés d'une question, qui d'ailleurs recourent partiellement certaines classifications déjà exposées à propos des énoncés assertifs. En particulier, il oppose présuppositions primaires et présuppositions secondaires, et également présuppositions factuelles et présuppositions formelles. D'autre part, KEENAN (et d'autres auteurs) proposent de distinguer présuppositions logiques (ou sémantiques) et présuppositions pragmatiques.

La distinction la plus importante en pratique est entre *la présupposition primaire* et *les présuppositions secondaires* d'une interrogation. Soit par exemple la question : "A quelle vitesse John Sprague conduisait-il hier soir dans Main Street ?". Elle comporte une présupposition primaire : "John Sprague conduisait hier soir dans Main Street", et un ensemble de présuppositions secondaires parmi lesquelles les présuppositions existentielles concernant John Sprague et Main Street. Toute présupposition telle que sa vérité est impliquée (au sens logique) par la vérité d'une ou plusieurs autres présuppositions de la même question est une présupposition secondaire ; sinon, c'est une présupposition primaire [Leonard 1967 : 39]. Toute question a une présupposition primaire, et une seule (Nous retrouvons ici une notion analogue à la notion de présupposition du premier ordre, de LAKOFF). La conception de LEONARD se fonde évidemment sur la transitivité de la relation de présupposition, et il est aisé de montrer que ses définitions peuvent être mises en échec. Par exemple, la question : "Jacques s'imaginait-il que la réunion finirait aussi tard ?" a en principe deux présuppositions secondaires contradictoires : "la réunion a fini tard", et :

"la réunion n'a pas fini tard", présupposée (présupposition lexicale) par le verbe : "s'imaginer" (cf. : "Jacques s'imaginait que la réunion finirait tard") [Ducrot 1972 : 276-277]. En conséquence, selon la définition de LEONARD, il ne pourrait pas y avoir de présupposition primaire à cette question. De prime abord, il paraît possible d'adopter le point de vue de DUCROT, déjà mentionné à propos des énoncés déclaratifs, selon lequel certains termes peuvent annuler certains présupposés : à l'effet d'annulation de "s'imaginer" se combinerait dans le cas cité un effet dû à l'adverbe "aussi", effet qu'il importerait d'analyser. En fait, DUCROT considère plutôt "que l'on touche ici les limites de l'analyse en présupposés", et que cela "montre la nécessité de construire une théorie plus compréhensive" pour rendre compte de ces observations [Ducrot 1972 : 277].

La distinction entre *présupposition factuelle* et *présupposition formelle* est illustrée par les deux questions suivantes :

- (a) "A quelle vitesse John conduisait-il hier ?"
- (b) "Quelle est la largeur de ce bureau ?"

qui ont pour présupposition primaire respectivement :

- (a') "Jones conduisait hier"
- (b') "Ce bureau a une largeur"

La présupposition (a') est dite *factuelle* car elle requiert une observation, un témoignage sur un événement qui s'est produit. La présupposition (b') est dite *formelle*, parce que l'existence d'une largeur est contenue dans la description ou la définition de l'objet appelé : "bureau" [Leonard 1967 : 36-37]. Un critère de distinction est que, si une présupposition formelle (au moins) est fausse, la question n'a pas de sens (comme par exemple : "Quelle est la hauteur de l'honnêteté ?"), alors qu'elle conserve un sens si une présupposition factuelle est fausse. On peut rencontrer certaines présuppositions formelles particulièrement subtiles, comme dans la question : "George Washington était-il un plus grand homme qu'Abraham Lincoln ?", qui présuppose entre autres choses que la grandeur (au sens de : "supériorité morale et intellectuelle") est

susceptible d'être sinon mesurée, du moins ordonnée [Leonard 1967 : 37-38].

La distinction faite par KEENAN entre présuppositions *logiques* et présuppositions *pragmatiques* [Keenan 1971] se retrouve chez d'autres auteurs (cf. p.ex. [Barker 1974 : 38]), et peut d'ailleurs s'appliquer également aux énoncés déclaratifs. Les présuppositions pragmatiques dépendent du contexte de leur énonciation, c'est-à-dire du locuteur, de ses interlocuteurs, des autres personnes présentes, de l'environnement physique, des significations sociales et culturelles attachées à l'acte d'énonciation, etc. [Keenan 1971 : 49]. Ces caractéristiques (que LEONARD ne considère d'ailleurs pas comme d'authentiques présuppositions [Leonard 1967 : 35-36]) sont bien des présuppositions en ce sens que leur ignorance peut rendre l'énoncé soit totalement incompréhensible, soit très difficile à interpréter. Comme exemples de présuppositions pragmatiques, on peut citer les relations entre locuteur et interlocuteur liées au sexe, à l'âge, et au statut social, qui déterminent le choix des expressions utilisées dans la plupart des langues naturelles. Quant aux présuppositions logiques des énoncés interrogatifs, KEENAN les définit comme "les phrases qui sont des conséquences logiques de chacune de leurs réponses" [Keenan 1971 : 48]. Nous retrouverons cette propriété fondamentale des présuppositions des questions dans les thèses essentielles de la logique érotétique.

Que se passe-t-il lorsque la présupposition primaire d'une question est fautive ? Le philosophe grec MÉNÉDEME d'Érétrie avait déjà été confronté au problème par ALEXINOS, qui "lui demanda un jour s'il avait cessé de battre son père. Notre philosophe lui répondit que n'ayant jamais commencé il ne pouvait avoir cessé", et il refusa d'entrer dans le jeu d'ALEXINOS et de répondre à cette question par *oui* ou *non* [Diogène Laërce : I 156] (cf. également [Mallet 1845 : 189]). La *Logique de Port Royal* adopte la même solution : "Si, connaissant la probité d'un juge, on me demandait, *s'il ne vend plus la justice*, je ne pourrais pas répondre simplement par *non*, parce que le *non* signifierait, qu'il ne la vend plus ; mais laisserait croire en même temps que je reconnais qu'il l'a autrefois

vendue" [Arnault 1683 : 190]. Cette forme d'interrogation est appelée *réunion de plusieurs questions en une* (πλείω ἐρωτήματα ἐν ποιεῖν [Aristote : 167 b 38]), *sophisme des questions multiples* (*sophisma plurium quaestionum* pour les scolastiques ; *fallacia plurium interrogationum* [Whately 1848 : 180]), *question complexe* [Creighton 1898 : 181-182], *confusion de plusieurs questions* [Makovelski 1967 : 204], etc. Un exemple historique souvent cité dans la littérature anglo-saxonne est la question posée à la Royal Society par le roi Charles II, question que l'on peut résumer en : "Quelles sont les raisons qui expliquent le fait qu'un poisson mort pèse plus lourd qu'un poisson vivant ?" ; ce qui implique en réalité deux questions distinctes : "Est-ce un fait avéré ?", "Si oui, quel en est la cause ?" (cf. [Whately 1848 : 181-182, 207]). Un autre exemple souvent cité (que nous retrouverons très fréquemment sous la plume des logiciens traitant des questions) est une adaptation de la question d'ALEXINOS, sous la forme : "Avez-vous cessé de battre votre femme ?". R.G. COLLINGWOOD montre que cette interrogation se décompose en (au moins) quatre questions successives [Collingwood 1940 : 38] :

- "Avez-vous une femme ?"
- "Aviez-vous l'habitude de la battre ?"
- "Avez-vous l'intention de ne plus le faire ?"
- "Avez-vous commencé à mettre cette intention à exécution ?".

Sur ce modèle, P.T. GEACH a proposé une critique de la thèse de RUSSEL sur les présuppositions, en montrant qu'une question comme : "Le roi de France est-il chauve ?" se décompose en trois questions successives, dont les deux premières mettent en cause les présupposés de la troisième [Geach 1950 : 33] :

- "Quelqu'un est-il actuellement roi de France ?"
- (Si oui) "Y-a-t-il actuellement plusieurs personnes distinctes qui sont chacune roi de France ?"
- (Si non) "L'actuel roi de France est-il chauve ?"

Ce qui revient à dire que RUSSEL tombe dans le sophisme des questions multiples lorsque la présupposition d'existence porte sur le sujet de la présupposition [Geach 1950 : 34].

Lorsque la présupposition primaire d'une question est fautive, LEONARD considère que cette question est *invalidée*, c'est-à-dire qu'elle n'a pas de réponse [Leonard 1967 : 41]. Dans l'exemple ci-dessus, si l'on n'a pas répondu *oui* à la première question et *non* à la seconde, "la réponse affirmative *oui il est chauve* n'est pas fautive mais simplement hors de propos" [Geach 1950 : 33].

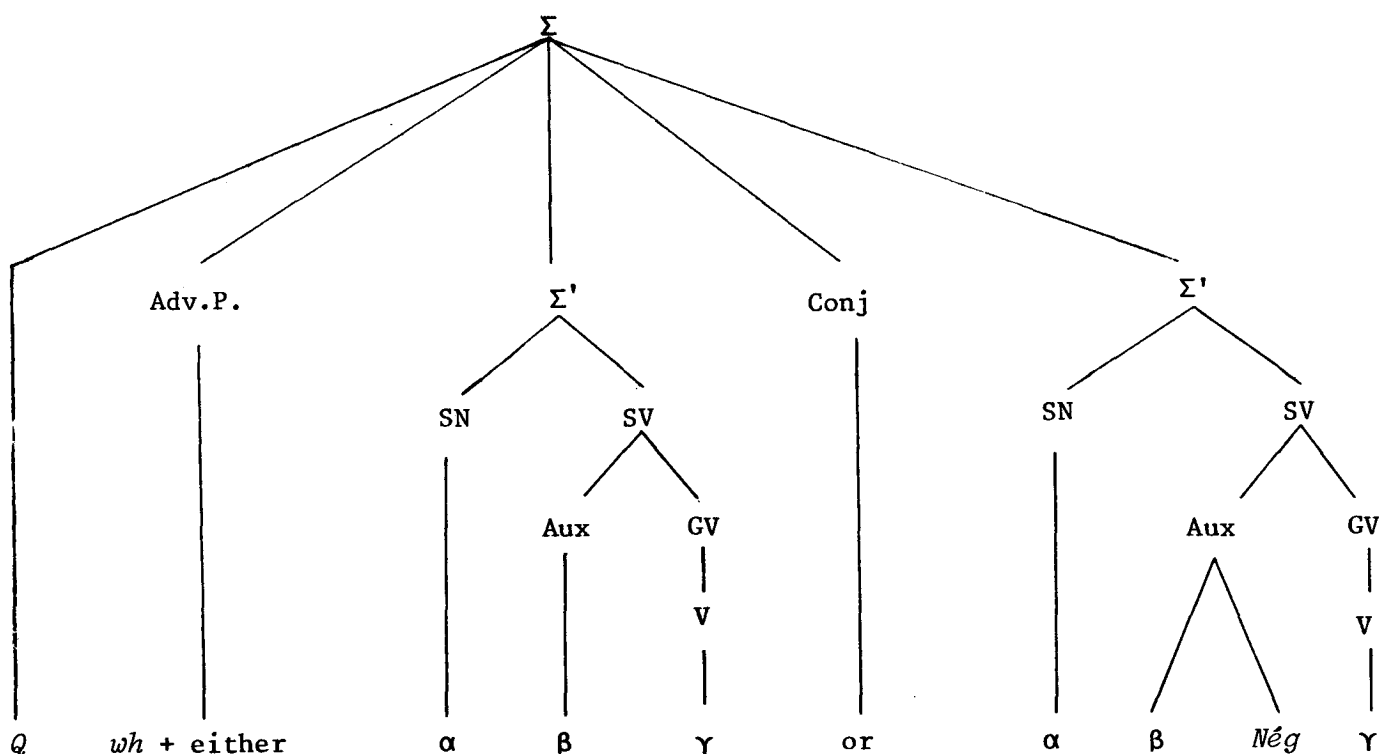
b) L'analyse logico-sémantique des questions de Jerrold J. KATZ

A partir d'une analyse linguistique des énoncés interrogatifs fondée sur la grammaire générative de CHOMSKY, Jerrold J. KATZ a formulé une théorie logico-sémantique des questions qui vise à "l'explication de certaines propriétés et certaines relations logiques des interrogatives" [Katz 1972 : 201]. La démarche suivie consiste à dresser un parallèle entre certaines formes énonciatives et des formes interrogatives qui leur sont apparentées, et à examiner si les propriétés des premières peuvent être attribuées aux secondes, moyennant éventuellement quelques modifications. Les propriétés les plus importantes selon KATZ sont relatives aux expressions analytiques, aux expressions contradictoires, et aux relations d'implication entre expressions.

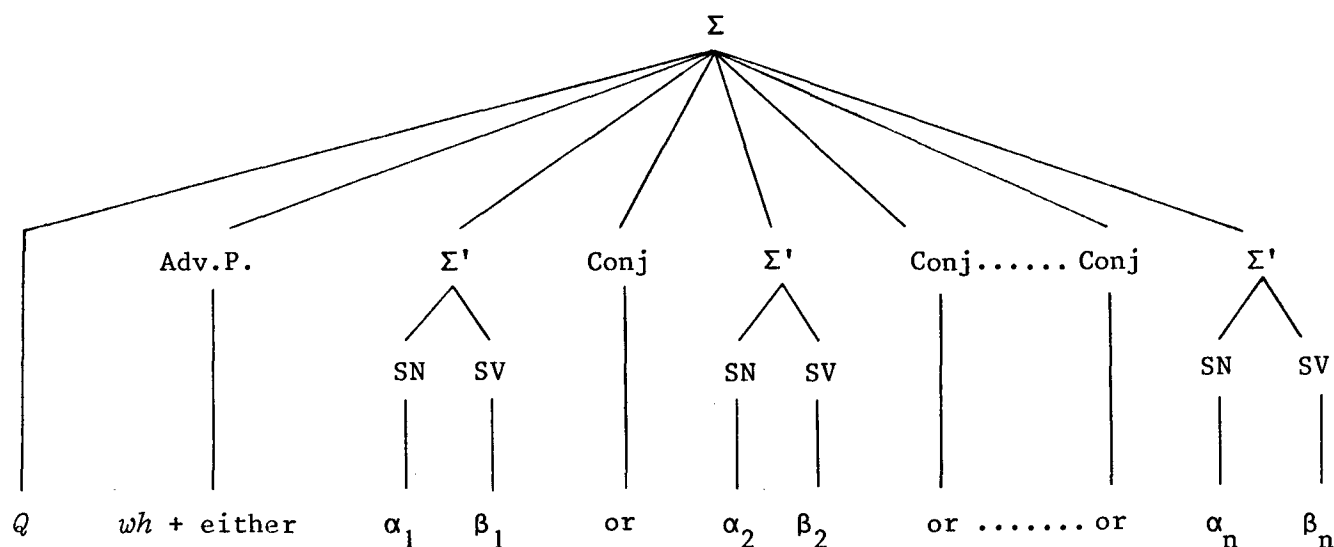
La typologie des interrogatives de Jerrold KATZ emprunte sa terminologie à Otto JESPERSEN. Il distingue deux grandes catégories de questions : les *questions nexales*, qui correspondent aux questions générales, et proposent le choix entre plusieurs modalités (ex. : "Jean demande si Marie est partie, ou non") ; et les *questions sur x*, qui correspondent aux questions particulières (ex. : "Jean demande quand Marie est partie"). Il ajoute toutefois que "sur le plan sémantique, il s'agit d'une distinction fondamentale, mais sur le plan syntaxique, les questions nexales sont simplement une variante de plus dans la structure révélée par les différents spécimens de questions sur *x* que nous avons examinés" [Katz 1972 : 207]. En effet, selon la théorie de KATZ et POSTAL, en anglais, toute question recèle dans sa structure profonde à la fois un marqueur de l'interrogation *Q* et un élément *wh* (correspondant à l'élément

qu en français), non spécifique de l'interrogation (et pouvant être également présent dans les relatives, certaines complétives, etc.) [Katz 1964 : 157], et indiquant ici l'étendue du champ de Q [Katz 1972 : 205]. Mais en pratique (nous l'avons signalé au § 1.2.4.b) la distinction entre les deux formes de questions reste importante ; elle est aisée à faire dans la structure profonde elle-même, comme nous allons le voir.

Les *questions nexales* se subdivisent en deux types : les questions nexales *simples*, correspondant à l'alternative entre une expression et sa négation (ex. : "Marie est-elle partie ?"), et auxquelles on peut répondre par *oui* ou *non* ; et les questions nexales *complexes*, qui proposent le choix entre plusieurs modalités exclusives mais non contradictoires (ex. : "Marie est-elle partie à pied, en autobus, ou en taxi ?"). La structure profonde correspondant à ces deux types de question nexale est représentée dans les figures n° 2.1. et 2.2. On voit que l'auteur postule l'existence, dans la structure profonde



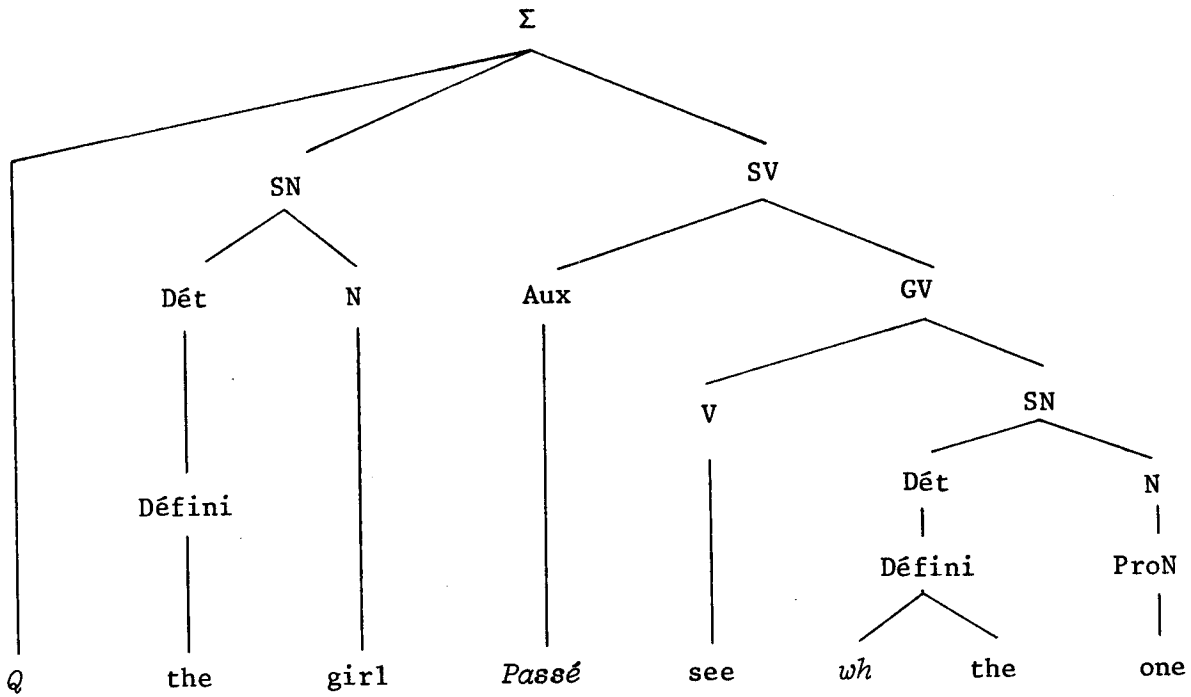
2.1. : Structure profonde d'une question nexale simple [Katz 1972 : 217] (ex.: "Did Mary go or didn't Mary go ?" [Katz 1972 : 209]).



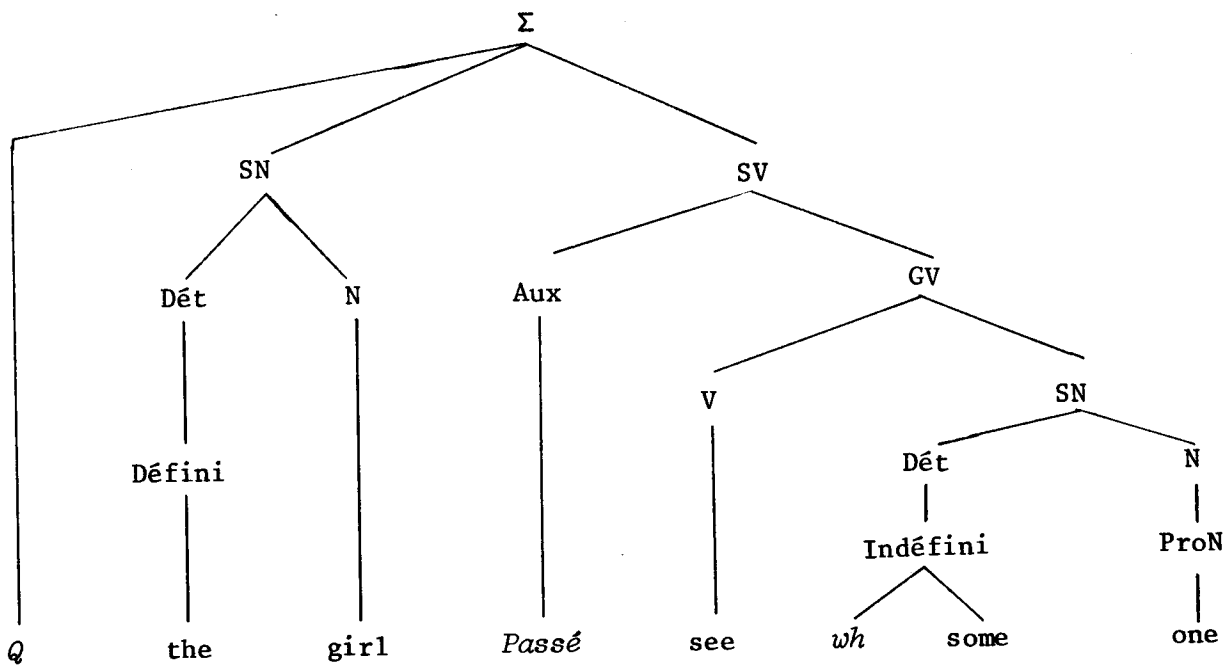
2.2. : Structure profonde d'une question nexale complexe [Katz 1972 : 218].

d'une question nexale, d'un *adverbe de phrase* (Adv. P.) dominant *wh* [Katz 1964 : 155]. Dans la réalisation de la phrase, le constituant adverbial est effacé lorsqu'il est précédé de *Q* en position initiale [Katz 1964 : 157], ce qui est toujours le cas dans les questions nexales directes. D'autre part, la structure profonde postulée par KATZ est très sensiblement différente de celle que postulent DUBOIS et DUBOIS-CHARLIER (cf. § 1.2.2.c et figure n° 1.60), puisque toute question nexale contient au moins deux phrases Σ' enchassées, reliées par la conjonction *ou*. Cette analyse est en fait assez proche de la manière dont HARRIS rend compte de l'interrogation totale ; d'ailleurs, KATZ et POSTAL soulignent eux-mêmes qu'une question comme : "Vas-tu rentrer chez toi ?" peut être paraphrasée en : "J'exige que tu me répondes « oui, je vais rentrer chez moi » ou « non, je ne vais pas rentrer chez moi »" [Katz 1964 : 136-137].

Les *questions sur x* sont traitées par KATZ et POSTAL d'une manière relativement semblable à l'analyse de DUBOIS et DUBOIS-CHARLIER, à ceci près que les auteurs proposent une distinction importante sur le plan sémantique entre les questions *portant sur un domaine indéfini* (ex. : "Qui avez-vous vu ?") et celles *portant sur un domaine défini* (ex. : "Lequel avez-vous vu ?") [Katz 1964 : 146-152]. Au niveau morpho-syntaxique, cette différence se limite à la



2.3. : Structure profonde d'une question sur x portant sur un domaine défini ("which one did the girl see ?") [Katz 1964 : 167].



2.4. : Structure profonde d'une question sur x portant sur un domaine indéfini ("Who did the girl see ?")

nature du déterminant du syntagme nominal sur lequel porte l'interrogation (cf. figures n° 2.3. et 2.4.). On notera que l'élément *wh* n'est qu'un constituant du déterminant de ce syntagme nominal, alors que chez DUBOIS et DUBOIS-CHARLIER, la marque de l'interrogation partielle est elle-même un syntagme nominal : SN_{que}. Une question sur *x* de la forme : "Qui quelqu'un a-t-il vu ?" peut être paraphrasée en : "J'exige que vous me répondiez « quelqu'un a vu *x* » " [Katz 1964 : 138].

Sur ces bases morpho-syntaxiques, KATZ définit les deux notions fondamentales de *présupposition* et de *réponse possible* à une question. Ces deux notions correspondent aux deux types d'information que l'interlocuteur peut tirer de la structure linguistique d'une question [Katz 1964 : 188].

La *présupposition* d'une question est une proposition (ou une conjonction de propositions) dont la vérité est la condition pour que la question exprime une demande d'information [Katz 1972 : 210]. Par exemple, pour que la phrase interrogative : "Qui a vu Marie ?" ait valeur de question, il faut que l'assertion : "Quelqu'un a vu Marie" soit vraie (ou soit acceptée comme telle par l'interlocuteur) [Katz 1964 : 187]. D'où les deux définitions suivantes : [Katz 1972 : 210-211] :

- une proposition *p* est la *présupposition* de la question *q* si et seulement si *p* est représentée par la lecture d'une phrase qui a la même structure profonde que *q*, sauf que le marqueur de l'interrogation *Q* y est remplacé par celui de l'énonciation, et que les éléments *wh*, devenus sans objet, sont effacés.

- une proposition *p'* est une *composante de la présupposition* de la question *q* si et seulement si *p'* est représentée par la lecture d'une phrase pouvant être déduite de la présupposition de *q*.

A titre d'illustration, la question nexale : "Marie est-elle rentrée chez elle ou Marie est-elle partie pour l'école ?" a pour présupposition : "Ou bien Marie

est rentrée chez elle ou bien Marie est partie pour l'école". De même, la question sur x : "Qu'est-ce que Jean a mangé ?" a pour présupposition l'une des deux prépositions équivalentes : "Jean a mangé quelque chose", ou : "Quelque chose a été mangée par Jean". Dans ce dernier cas, "Qu'est-ce que" provient d'une structure profonde du syntagme nominal SN \rightarrow Indéfini + ProN correspondant à : "*wh* + quelque + chose" ("*wh* + some + thing").

La notion de *réponse directe possible* à une question varie selon la catégorie à laquelle appartient la question. Nous avons par conséquent les deux définitions suivantes [Katz 1972 : 212] :

- la proposition exprimée par une phrase p est une réponse directe possible à la *question nexale* q si et seulement si la structure profonde de p est identique à la structure profonde d'un des éléments Σ' de la disjonction présentée par q (cf. figures n° 2.1. et 2.2.) ; ou bien p est une paraphrase d'une phrase satisfaisant à la condition ci-dessus.

- la proposition exprimée par une phrase p est une réponse directe possible à la *question sur x* q si et seulement si la structure profonde de p est identique à celle de la présupposition de q , sauf que tout syntagme nominal SN assorti d'un élément *wh* est remplacé par un syntagme nominal SN' contenant plus d'information que SN (ayant plus de marqueurs sémantiques que SN) ; ou bien p est une paraphrase d'une phrase satisfaisant à la condition ci-dessus.

Ces définitions ne peuvent s'appliquer qu'à condition que la question soit *décidable*, c'est-à-dire ne contienne pas d'expression contradictoire. On peut en outre définir une réponse *indirecte possible* à une question q comme une proposition décidable qui n'est pas une réponse directe à q , et dont découle une réponse directe possible à q [Katz 1972 : 216].

Les réponses possibles sont à distinguer d'autres formes d'énoncés pouvant faire suite à une question, sans toutefois constituer une réponse à celle-ci [Katz 1972 : 213-215] :

- l'*élusion de la question*, dont le contenu informationnel est identique ou inférieur à celui de la question elle-même ; une telle réplique à une question q est par conséquent soit la présupposition de q , soit un énoncé identique à l'un des éléments conjoints de la présupposition de q . Par exemple, soit la question : "Quand John a-t-il cessé de battre sa femme ?" ; sont des élusions de la question les répliques : "John a cessé de battre sa femme à un certain moment", et : "John a une femme".

- le *rejet de la question*, dont le contenu informationnel est en contradiction avec celui de la question ; une telle réplique à une question q est une proposition qui entraîne la négation de la présupposition de q . Sur le même exemple que précédemment, sont des rejets de la question les répliques : "John n'a jamais cessé de battre sa femme", "John n'a jamais battu sa femme à quelque moment que ce soit", et : "John n'a pas de femme".

- l'*aveu d'ignorance*, qui répond non au contenu de la question, mais à une présomption du locuteur relativement à ce que sait l'interlocuteur. Par exemple : "Je ne sais pas quand John a cessé de battre sa femme".

Enfin, KATZ définit une *question soluble linguistiquement* comme une question dont la présupposition *peut* être vraie, c'est-à-dire dont la présupposition n'est ni indécidable, ni contradictoire [Katz 1972 : 215, 222]. Une phrase est *indécidable* si son sujet n'a aucun référent ; tel est le cas par exemple de : "L'homme reine est malheureux", ou de : "Le cadavre vivant a attaqué quelqu'un", en raison du caractère contradictoire des termes désignant le sujet de la phrase. C'est pourquoi sont insolubles linguistiquement les questions : "L'homme reine est-il malheureux ?", et : "Qui le cadavre vivant a-t-il attaqué ?" [Katz 1972 : 221-222]. Une phrase est *contradictoire* si son sujet et son prédicat contiennent des éléments antonymes, c'est-à-dire soit des éléments contradictoires, comme : "parfait/imparfait", "vivant/mort" ; soit des éléments contraires, comme : "riche/pauvre", "supérieur/inférieur" ; soit des éléments inverses, comme : "acheter/vendre", "employeur/employé" [Katz 1972 : 159]. Par exemple : "Jean accepte les offres qu'il refuse", "Jean est courageux et timoré", "Jean a un crâne chauve couvert de cheveux". Aussi, la question : "Qui tuera le cadavre ?" est-elle insoluble linguistiquement parce que

sa présupposition : "Quelqu'un tuera le cadavre" est une phrase contradictoire [Katz 1972 : 222].

Pour aborder le problème des propriétés logiques des questions, KATZ part de certaines propriétés des phrases énonciatives ; il considère en particulier les phrases *analytiques* (toujours vraies, comme p.ex. : "Jean achète à ceux qui lui vendent") et les phrases *contradictaires* (toujours fausses, comme p.ex. : "Jean vend des choses à ceux qui ne les lui achètent pas"), et recherche parmi les énoncés interrogatifs ceux qui peuvent être considérés comme équivalents à ces types de phrases énonciatives. Cela le conduit à la notion de question *résolue linguistiquement* [Katz 1972 : 217]. La définition de cette notion varie selon le type de la question.

Une *question nexale simple* est dite :

- *résolue linguistiquement par l'affirmative* si le premier terme de la disjonction est une expression analytique. Exemple : "Une vieille fille est-elle du sexe féminin, ou non ?"

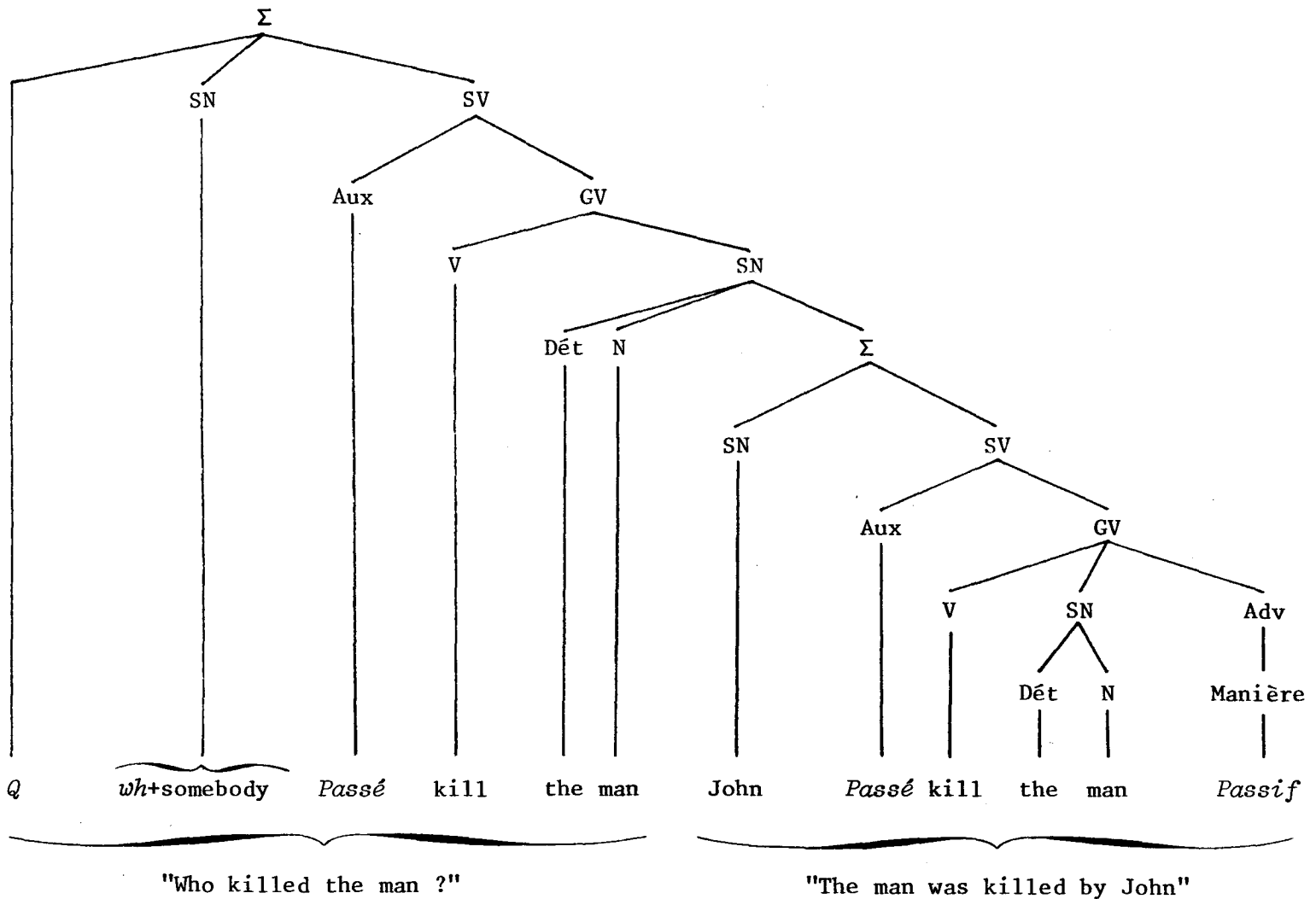
- *résolue linguistiquement par la négative* si le premier terme de la disjonction est une expression contradictoire. Exemple : "Une vieille fille est-elle du sexe masculin, ou non ?".

Une *question nexale complexe* est dite :

- *résolue linguistiquement par l'affirmative* si l'un des termes de la disjonction est une expression analytique. Exemple : "Une vieille fille est-elle du sexe masculin ou du sexe féminin ?".

- *résolue linguistiquement par la négative* si tous les termes de la disjonction sont des expressions contradictoires. Exemple : "Une vieille fille est-elle mariée, ou du sexe masculin ?".

Une question sur x est dite résolue linguistiquement lorsque sa structure profonde contient à la fois un élément qui a la forme de la structure profonde d'une question q' et un élément qui a la forme de la structure profonde d'une réponse possible à la question q' (ce dernier élément se présentant comme une relative enchâssée). Par exemple, la question : "Qui a tué l'homme qui a été tué par Jean ?" contient, dans sa structure profonde, la structure profonde de la question : "Qui a tué l'homme ?", conjointement à celle de : "L'homme a été tué par Jean", qui est une réponse possible à cette dernière question (cf. figure n° 2.5.). Comme exemples de question nexale résolue linguistiquement, on peut également citer : "Quelle est la couleur du chariot rouge ?", "Où se trouve le chapeau que j'ai sur la tête ?", "Quelle heure est-il à minuit juste ?", etc. [Katz 1972 : 203].



2.5. : Structure profonde d'une question nexale résolue linguistiquement ("Who killed the man who was killed by John ?") [Katz 1972 : 219].

Considérant ensuite les relations d'*implication* entre phrases énonciatives (comme : "Marie a eu un cauchemar" → "Marie a rêvé"), KATZ définit entre les questions une forme de relation analogue. Soit par exemple le couple de questions nexales simples :

$$\left\{ \begin{array}{l} Q_1 = \text{"Jean est-il roi ?"} \\ Q_2 = \text{"Jean est-il un monarque ?"} \end{array} \right.$$

Il est clair qu'une réponse affirmative à Q_1 implique nécessairement une réponse affirmative à Q_2 . De même, dans le couple de questions sur x :

$$\left\{ \begin{array}{l} Q_1 = \text{"Qui a volé un chat ?"} \\ Q_2 = \text{"Qui a volé un animal ?"} \end{array} \right.$$

toute réponse à Q_1 est par cela même une réponse à Q_2 .

c) *Insuffisance des analyses de type linguistique.*

Nous l'avons dit, les travaux linguistiques sur l'interrogation et sur les réponses aux interrogations sont actuellement en plein développement. Les résultats que nous avons présentés ne sont qu'un échantillon des acquis des recherches en cours (encore que nous ayons sélectionné ceux qui nous ont paru les plus utiles pour l'étude des tests et questionnaires). Il serait par conséquent prématuré de porter un jugement sur la valeur de ces travaux, et surtout de préjuger leur fécondité et leur intérêt pratique. Malgré tout, nous pouvons raisonnablement penser que ces recherches n'ont pas encore atteint un niveau qui les rendent directement utilisables, tant sur le plan technique que sur celui de la théorie. En effet, nous n'avons pas trouvé de moyens techniques permettant par exemple, dans tous les cas et sans risque d'erreur, d'identifier les présupposés d'un énoncé interrogatif ou déclaratif, ou de décider si un énoncé donné constitue une réponse totale ou une réponse partielle, ou ne constitue pas une réponse du tout à une question donnée. Sur le plan théorique, les essais de formalisation des relations d'implication ou de

présupposition ne paraissent pas avoir abouti à des modèles ajustables aux données observées. En outre, si l'on excepte la tentative de KATZ, nous n'avons pas trouvé de théorie linguistique cohérente qui traite de tous les aspects logico-sémantiques de l'interrogation. Enfin, l'essai de KATZ lui-même ne supporte pas la comparaison avec les théories que propose la logique érotétique, comme le montre Frank FAIR.

Dans un article très critique relatif à la théorie de KATZ, Frank FAIR fait tout d'abord remarquer qu'il est possible de compléter les définitions et théorèmes proposés par KATZ, en ajoutant que lorsque, dans un couple de questions nexales simples, *oui* à Q_1 implique *oui* à Q_2 , alors de même *non* à Q_2 implique *non* à Q_1 [Fair 1975 : 284]. Par analogie avec les fractions, il appelle de tels couples de questions *réciproques* [Fair 1975 : 285], et montre que cette relation peut être définie d'une manière plus rigoureuse et plus complète dans le cadre des théories de logiciens tels que HARRAH, BELNAP, ou ÅQVIST. Nous verrons en effet (§ 2.2.5.e) que la relation d'*inclusion* entre questions, qui correspond à la relation décrite par KATZ, a en logique érotétique à la fois un fondement théorique plus assuré, et un champ d'application plus étendu. En outre, la logique érotétique a le mérite de faire apparaître, dans la relation d'implication entre questions, certaines assertions implicites qui justifient cette relation. Par exemple, une réponse positive à Q_1 = "Jean est-il un roi ?" entraîne une réponse positive à Q_2 = "Jean est-il un monarque ?" *relativement à un ensemble H d'expressions* parmi lesquelles : "Tout roi est un monarque" (cf. [Fair 1975 : 286-287], et § 2.2.5.a). Enfin, d'autres relations logico-sémantiques peuvent être définies sans difficulté (nous en verrons des exemples dans les seconde et troisième parties de ce chapitre). C'est pourquoi FAIR conclut : "Bien que je pense que l'article de KATZ est intéressant pour ses efforts en vue d'appliquer la grammaire moderne aux questions, je pense également que sa tentative pour fonder une logique des questions repose sur des analogies avec la logique des propositions qui sont fallacieuses et obscurcissantes [...]. A mon avis, les relations sur lesquelles KATZ attire l'attention peuvent être traitées d'une meilleure manière en utilisant les concepts dégagés par David HARRAH et les autres

pionniers de ce champ de recherches logiques. L'approche de KATZ est une voie sans issue ; la leur est un nouveau point de départ pour la logique" [Fair 1975 : 289].

2.1.3. La logique érotétique : brève présentation historique.

Le terme de : "logique érotétique", dans son acception moderne de : "logique des questions", semble dû à Mary et Arthur PRIOR, dans [Prior 1955]. On trouve toutefois l'adjectif ἐρωτητικός chez PLATON, servant par exemple à SOCRATE pour qualifier les héros d'Homère de "questionneurs habiles" (ἐρωτητικοὶ ; *Cratyle*, 398 e). Mais c'est chez ARISTOTE qu'il prend son sens le plus technique, puisque celui-ci définit la dialectique comme un art qui *procède par interrogation* (ἐρωτητική τέχνη), par opposition aux arts qui *démontrent* la nature de quelque chose [Aristote : 172 a 16]. Il est vrai que cet adjectif, d'ailleurs rarement usité, a un sens plus large chez ARISTOTE que chez les logiciens contemporains.

La logique érotétique s'est développée au vingtième siècle. Apparue vers 1925, elle a véritablement pris son essor à partir de 1958, et est encore actuellement en pleine expansion. Cependant, sans vouloir tomber dans le travers qui consiste à faire remonter toute nouveauté à l'antiquité gréco-latine, il est possible de trouver chez ARISTOTE, et également chez les Stoïciens, de véritables précurseurs des travaux actuels. Nous présentons ci-après les quelques jalons historiques antérieurs au vingtième siècle qui ont été redécouverts, et qui constituent en quelque sorte la préhistoire (ou plus exactement la proto-histoire) de la logique des questions ; nous décrivons ensuite les grandes étapes du développement de la logique érotétique depuis 1925 environ jusqu'en 1976.

a) Les précurseurs (de l'antiquité au dix-neuvième siècle)

En relisant l'*Organon* à la lumière des recherches récentes en logique des questions, il est facile de trouver chez ARISTOTE des éléments d'une théorie de l'interrogation ; mais, en raison du peu d'intérêt qu'il porte aux formes

d'énoncés n'exprimant pas un jugement, l'auteur n'a pas éprouvé le besoin de les rassembler et de les systématiser. Ces éléments peuvent être présentés de manière synthétique sous quatre rubriques :

① *Questions réelles et pseudo-questions.* On peut, avec Robert SOKOLOWSKI, distinguer chez ARISTOTE deux types de pseudo-questions : celui qui repose sur une tautologie, dans lequel la question est purement verbale, et celui qui porte sur un fait non avéré [Sokolowski 1971 : 243-244]. Pour illustrer le premier type, SOKOLOWSKI propose l'énoncé : "Pourquoi une symphonie est-elle une composition musicale pour orchestre en forme de sonate ?". Dans cet exemple, "symphonie" a pour prédicat sa propre définition, et la question se ramène à : "Pourquoi A est-il A ?". Par conséquent, toute réponse serait tautologique, et n'affirmerait rien (cf. [Aristote : 90 b 33-36]). Le second type de pseudo-question porte sur l'essence d'une chose dont nous ne savons pas véritablement qu'elle existe, "et chercher ce qu'est une chose sans savoir qu'elle existe, c'est assurément ne rien chercher du tout" [Aristote 93 a 27]. Aussi, une question réelle ne doit-elle porter que sur un phénomène réel, qui ne soit pas une simple coïncidence, un simple accident.

② *Présupposés d'une question.* De la citation qui précède, on peut conclure que toute question présuppose l'existence de ce sur quoi on s'interroge. De plus, dans le cas d'une question à réponse par *oui* ou *non*, on pose "que quelque chose appartient ou n'appartient pas à une autre chose" [Aristote : 24 a 25].

③ *Types de questions.* "Nous nous posons quatre sortes de questions : le *fait*, le *pourquoi*, *si* la chose existe, et enfin *ce* qu'elle est" [Aristote : 89 b 23-25] ; ce que HIŽ paraphrase en : "1. Est-ce ainsi ? / 2. Pourquoi ? / 3. Est-ce là ? Cela existe-t-il ? / 4. Qu'est-ce ?" [Hiz 1962 : 253]. SOKOLOWSKI oppose, dans ces quatre types fondamentaux, les questions d'existence (εἰ ἐστὶ) aux autres formes d'interrogation (ὅτι) ; il qualifie les premières d'*herméneutiques* car "elles essayent d'interpréter des significations acceptées", et les secondes de *scientifiques* car "elles essayent de comprendre

des faits enregistrés" [Sokolowski 1971 : 250]. Cette opposition est d'ailleurs suggérée par ARISTOTE lui-même ("Quand nous connaissons le fait, nous cherchons le pourquoi" [Aristote : 89 b 29]). Par ailleurs, on connaît les dix questions fondamentales qui définissent les catégories [Aristote : 1 b 25 - 2 a 4], ce qui constitue une sorte de classification des contenus de toutes les questions possibles. Au niveau de la forme de la question, ARISTOTE privilégie deux types particuliers : la question *dialectique* (c'est-à-dire la question de controverse par excellence, même si l'on peut l'utiliser également dans l'interrogation scientifique ; cf. sur ce dernier point [Aristote : 77 a 36-39]), "à laquelle il est possible de répondre par *oui* ou par *non*" [Aristote : 158 a 15] ; et la question *à sens multiples*, dont les termes sont obscurs [Aristote : 160 a 16-34], ou qui réunit plusieurs questions en une seule [Aristote : 175 b 39 - 176 a 18 ; 181 a 35 - 181 b 18], et à laquelle il est impossible de répondre directement. En fait, la question dialectique est, pour ARISTOTE, "la source commune dont la proposition affirmative et la proposition négative sont toutes deux issues" [Prior 1955 : 56]. Ce qui explique que la question "portant sur la nature même d'une chose", qui n'est pas une interrogation dialectique [Aristote : 20 b 27], ne soit guère prise en compte dans une logique centrée sur l'analyse de la proposition.

④ *Théorie de l'interrogation.* L'importance d'une théorie de l'interrogation est due au fait que tout notre savoir consiste dans les réponses que nous obtenons aux quatre types fondamentaux de questions [Aristote : 89 b 35-36]. La base de toute question non ambiguë est une proposition de forme *sujet-prédicat*, décrivant le phénomène qui fait problème. Répondre à l'interrogation, c'est découvrir le moyen terme qui explique cette proposition ; c'est par là même identifier la cause du phénomène. Ce mécanisme s'applique aux quatre types fondamentaux d'interrogation. En effet, "il est clair qu'il y a identité entre la nature de la chose et pourquoi elle est. La question : *qu'est-ce que l'éclipse ?* et sa réponse : *la privation de la lumière de la Lune par l'interposition de la Terre*, sont identiques à la question : *pourquoi y-a-t-il éclipse ?* [...] et sa réponse : *en raison du manque de lumière quand la Terre vient à s'interposer* [...] Connaître ce qu'est une chose revient à connaître pourquoi elle est" [Aristote : 90 a 14-18,31]. En d'autres termes,

la proposition de départ fait problème parce que l'on ne peut expliquer le lien qui unit sujet et prédicat ; le moyen terme recherché est ce qui s'identifie à la fois au prédicat de la proposition (pour exprimer ce qu'il est réellement) et au sujet (pour exprimer son substrat, c'est-à-dire ce qui sert de support à son existence). Par exemple, à la question : "Pourquoi a-t-il des taches rouges sur le corps ?", la réponse est contenue dans le moyen terme : "Rougeole" ; ce mot exprime l'idée que ces taches rouges "sont réellement" la rougeole, et identifie d'autre part la rougeole au corps qui en est le support [Sokolowski 1971 : 248-249]. Comme le fait remarquer LOHMANN, le syllogisme aristotélicien n'apporte aucune information nouvelle, mais permet simplement le passage de la connaissance vulgaire (δόξα) à la connaissance scientifique (ἐπιστήμη) ; c'est par conséquent dans le cadre de cette connaissance vulgaire, ou dans le cadre de l'expérience (ἐμπειρία), que les questions d'explication se posent (cité par [Sokolowski 1971 : 259 n 8]).

La logique des Stoïciens apparaît, au travers des relations qui en sont faites (en particulier par DIOGÈNE LAËRCE et SEXTUS EMPIRICUS), comme une oeuvre d'importance comparable à la logique aristotélicienne ; malheureusement, aucun des innombrables traités mentionnés par les historiens et les critiques ne nous est parvenu. Ceci est d'autant plus regrettable que, depuis les travaux de Victor BROCHARD (réédités à titre posthume dans [Brochard 1912 : 220-251]), on s'accorde généralement à reconnaître à la logique stoïcienne une profonde originalité par rapport à celle d'ARISTOTE, et même un certain modernisme (cf. par exemple la distinction entre le signe : σημάιον, la signification du signe : πρᾶγμα, et le signifié : τυγχάνον ; l'analyse du raisonnement hypothétique et de la relation d'implication ; la place donnée, dans la connaissance, à l'hypothèse et à sa vérification par l'expérience).

Il semble qu'un auteur fécond tel que CHRYSIPPE de Soles (ou de Tarse), à qui l'on attribue plus de 700 ouvrages, en ait consacré plusieurs à la logique érotétique. Sur les 119 titres d'écrits logiques mentionnés par

DIOGÈNE LAËRCE, 12 au moins traitent des questions ou des réponses aux questions (cf. [Diogène Laërce : II 118]). Si l'on se fonde sur l'analyse que fait cet auteur de la logique des Stoïciens (dans son chapitre sur ZÉNON de Cittium), on note tout d'abord que le problème de l'interrogation occupe dans ces théories une place entière, au même titre que l'énonciation d'un jugement, ou que d'autres actes de parole, tels que l'ordre, ou le serment [Diogène Laërce : II 73] ; alors que chez Aristote, l'état de celui qui interroge est un état essentiellement privatif et transitoire [Sokolowski 1971 : 245], l'interrogation n'existant que par rapport au jugement qui en est la réponse. L'importance de la logique érotétique pour les Stoïciens peut être entrevue à travers cet éloge de la dialectique : "[...] Sans elle, on ne peut ni interroger, ni répondre comme il faut, mais l'on donne dans le défaut qui consiste à tout nier, même le réel, en sorte que l'on tombe dans l'absurdité et la sottise [...]. La dialectique apprend à bien parler et à penser juste, à discuter sur les sujets proposés, à répondre aux questions posées, toutes choses que seul peut faire correctement l'homme très versé dans la dialectique" [Diogène Laërce : II 67]. Les deux points de la logique érotétique des Stoïciens rapportés par DIOGÈNE LAËRCE sont la thèse selon laquelle les questions n'ont pas de valeur de vérité, et la classification des questions selon la portée de l'interrogation : τὸ ἐρώτημα est une question bien déterminée, comme le jugement, mais elle demande une réponse par *oui* ou *non* (exemple : "Fait-il jour ?") ; τὸ πύσμα est une chose à laquelle il faut répondre par une phrase complète (exemple : "Où habitez-il ?") [Diogène Laërce : II 73-74]. Peut-être cette distinction est-elle contenue en germe chez ARISTOTE (cf. [Aristote : 20 b 27-31]), comme le pensent A. et M. PRIOR (cf. [Prior 1955 : 45]) ; mais en tout cas le mérite revient aux Stoïciens de l'avoir explicitée et développée.

On ne connaît guère mieux les apports ultérieurs à la logique érotétique. Dans une étude concernant les thèses d'AVICENNE (IBN - SINÂ) sur la logique des questions, Nicholas RESCHER a montré que celui-ci s'est inspiré des dix catégories d'ARISTOTE et des cinq prédicables de PORPHYRE pour développer une

classification des questions selon leur contenu. Cette classification se fonde sur deux dichotomies : questions de base, concernant la substance et les causes d'une chose, opposées à questions subsidiaires, concernant les traits accidentels ; et questions de fait, portant sur la réalité, opposées à questions de discours, portant sur les opinions et le sens des mots [Resher 1967 : 1-3].

D'autre part, on peut trouver d'autres jalons historiques, par exemple dans la *Logica Hamburgensis* de Joachim JUNGE. Cet auteur reprend la distinction des Stoïciens, en opposant l'*interrogatio* (ἑρώτημα ; ex. : "Y a-t-il quatre éléments ?") et la *siscitatio* (ou *percontatio*, πύσμα ; ex. : "Combien y-a-t-il d'éléments des corps naturels ?") [Junge 1638 : 358-359] ; en outre, il propose une classification des *interrogationes* en cinq sous-catégories, ce qui conduit à la classification des *quaestiones* représentée dans le tableau n° 2.6. D'autre part, il signale la possibilité d'autres distinctions ; par exemple, selon le contenu de la question [Junge 1638 : 365], ou selon que celle-ci est absolue ou comparative [Junge 1638 : 367, 502-503].

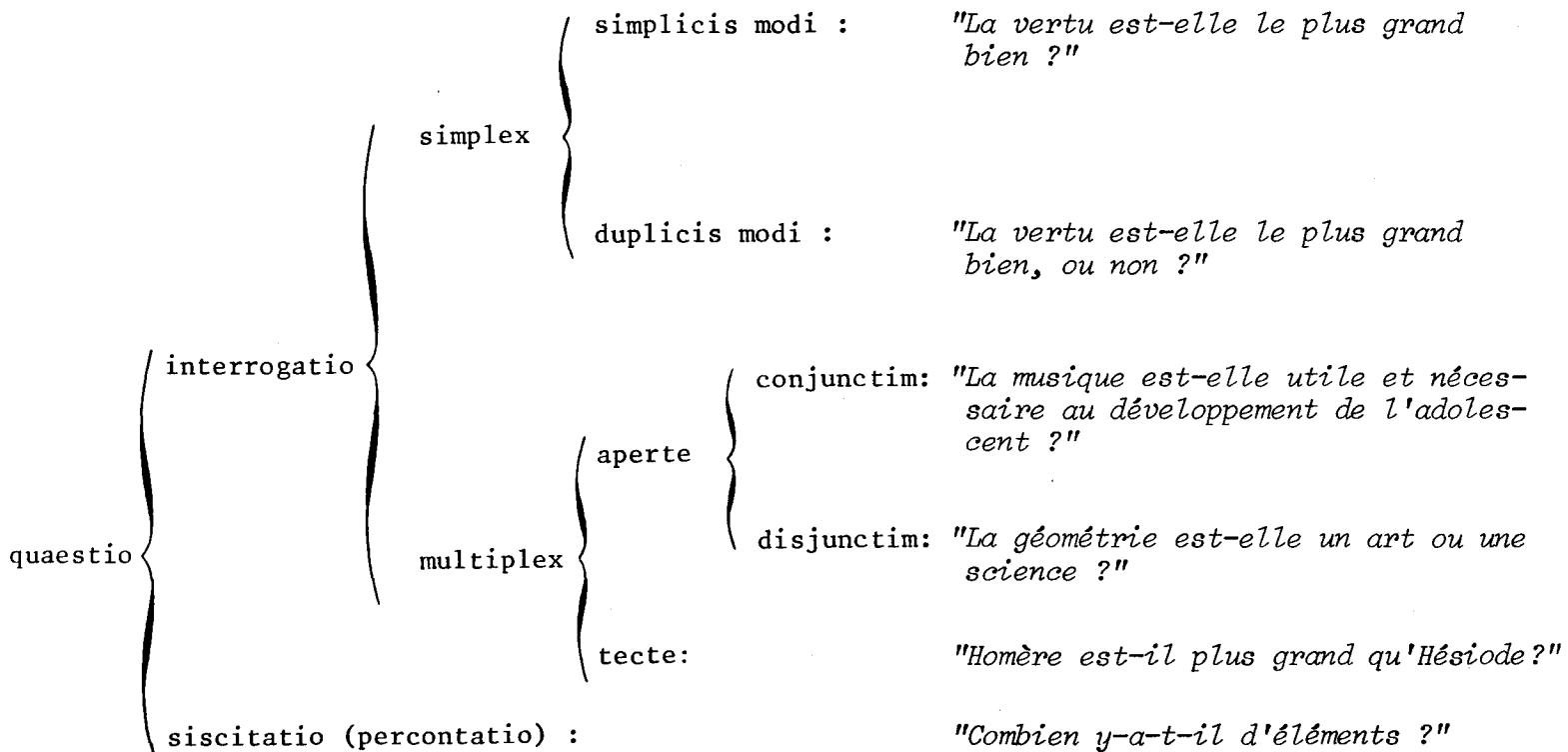


Tableau n° 2.6 : Classification des questions selon JUNGE (d'après [Junge 1638 : 358-363]).

Un dernier jalon important avant le vingtième siècle est constitué par les contributions de Richard WHATELY, archevêque de Dublin et polygraphe. Dans ses *Eléments de Logique*, WHATELY reprend la distinction stoïcienne des questions selon la portée de l'interrogation : "on peut considérer que toutes les questions appartiennent à l'une ou l'autre des deux catégories suivantes : « Quel sera le prédicat d'un sujet donné » ; et : « Quelle copule, affirmative ou négative, unira un sujet donné à un prédicat donné ». En bref, nous demandons ou bien : « Qu'est-ce que A ? », ou bien : « A est-il B , ou non ? ». La première catégorie de questions dépend du Philosophe ; la seconde, de l'Avocat" [Whately 1848 : 265]. En effet, certaines personnes ont besoin d'être persuadées, d'autres, d'être informées. "Les premières sont supposées avoir présents à l'esprit les *termes* de la proposition considérée, et on leur demande d'examiner si cette proposition particulière est vraie ou fausse ; les autres ne sont pas supposées connaître les termes de la conclusion, mais rechercher quelle proposition doit être reçue pour vraie. On peut décrire les premières [...] comme ayant un doute au sujet de la *copule* ; les autres, au sujet du *prédicat* "[Whately 1857 : 24]. En outre, on peut distinguer les questions *verbales*, qui portent sur la signification et l'extension des termes employés, et les questions *réelles*, qui concernent l'objet représenté par les termes (ce qui suppose un accord préalable sur leur signification) [Whately 1848 : 270-271]. Cette dichotomie est parallèle à la division des sophismes en ceux qui portent sur les mots (*in dictione*) et ceux qui portent sur le fond (*extra dictionem*) [Whately 1848 : 158]. Enfin, WHATELY mentionne également les questions conditionnelles [Whately 1848 : 64], et le sophisme de la question multiple [Whately 1848 : 180-182].

L'originalité de WHATELY, par rapport à ses prédécesseurs, consiste tout d'abord en l'intérêt porté à ce que CHOMSKY appellera le *focus* d'un énoncé (cf. p.ex. : [Chomsky 1972 : 38-55]), c'est-à-dire au centre d'intonation de la phrase, dont la place varie selon la nature de la question à laquelle elle répond. Ainsi, l'énoncé : "Vous trouverez cette thèse chez Bacon" peut être opposé à : "Vous trouverez dans Bacon une thèse *différente*", ou bien à : "Vous trouverez cette thèse chez un *auteur différent*" ; et de même : "Cet homme est un *meurtrier*" ou : "Tu ne *voleras* pas" ne répondent pas aux mêmes

questions que : "Cet homme *est* un meurtrier", ou : "Tu ne voleras *pas*" [Whately 1848 : 100 ; 1857 : 313].

Si l'on s'en tient plus strictement à la logique érotétique (indépendamment des problèmes linguistiques qui lui sont liés), l'apport de WHATELY porte sur trois points, qui n'apparaissent d'ailleurs que tout à fait incidemment dans les *Eléments de Logique* et les *Eléments de Rhétorique* :

① Réduction des deux types de question à une forme unique. Partant de l'analyse des questions ayant pour réponse possible : "Tu ne voleras pas", WHATELY fait remarquer que la question de type *whether* : "Est-ce que l'on peut voler, ou non ?" peut être considérée comme une question de type *what* : "quelle valeur de vérité a la proposition : il est permis de voler ?" (cf. [Whately 1857 : 313, 313 n 20], et [Prior 1955 : 45]). A partir de cette remarque, M. et A. PRIOR montrent qu'il est possible de définir deux procédures de réduction des questions *whether* aux questions *what* [Prior 1955 : 46-51].

② Description de la présupposition d'une question. Dans la question posée par Charles II à la *Royal Society* (≈ "Pourquoi un poisson mort pèse-t-il plus lourd qu'un poisson vivant ?" ; cf. [Whately 1848 : 181]), le roi abusait les membres de celle-ci en leur demandant implicitement de lui accorder la vérité du fait rapporté ; l'affirmation selon laquelle un poisson mort pèse plus lourd qu'un poisson vivant est ce que WHATELY appelle une *assomption indirecte* [Whately 1848 : 207], c'est-à-dire un présupposé.

③ Procédure d'assurance d'une question. Toujours à propos du même exemple, WHATELY suggère une procédure (qui sera reprise en particulier par ÅQVIST ; cf. § 2.3.1. a et b) pour rendre la question toujours valide : "Le moyen de réfuter [le sophisme] est, bien entendu, de répondre *séparément* à chacune des questions", c'est-à-dire à : "premièrement : est-ce un *fait* ? deuxièmement : si c'est un fait, quelle en est la cause ?" [Whately 1848 : 180, 182].

Ce survol des précurseurs de la logique érotétique est nécessairement lacunaire. Il semble en effet que "l'art d'interroger et de répondre", c'est-à-dire la dialectique (cf. PLATON, *Cratyle*, 390 c), ait tenu une place importante dans les préoccupations des philosophes de l'antiquité, mais aussi que la renommée de la logique d'ARISTOTE ait eu pour conséquence l'effacement progressif (et conséquemment la disparition matérielle) des travaux des autres logiciens. Il semble également que, jusqu'au premier quart du vingtième siècle, les auteurs qui ont traité du problème logique de l'interrogation ne l'aient fait que de manière incidente, et sans lui accorder beaucoup d'intérêt. On pourrait cependant relever d'autres travaux que ceux que nous avons développés, en parcourant par exemple la *Logique de Port Royal* (cf. [Arnault 1683 : 233-234]), les *Règles pour la direction de l'esprit* (cf. [Descartes : X, 428-429]), ou les premiers ouvrages de philosophie du langage comme le *Traité* de Alexander Bryan JOHNSON (cf. [Johnson 1836 : 241-253]). Mais de telles recherches se révèlent rapidement décevantes, en raison de la modicité des matériaux recueillis.

Pour clore ce paragraphe, signalons deux points historiques exhumés par Nuel D. BELNAP, qui sont à l'origine de dénominations devenues usuelles en logique des questions. Le premier est le "théorème du cinquième gymnosophe", qui correspond au dicton : "A question stupide, réponse stupide" ; on en trouvera l'énoncé et la démonstration dans [Belnap 1976 : 131-133]. La première mention connue de ce théorème se trouve dans la *Vie d'Alexandre* (PLUTARQUE, *Vies*, IX, 64 6-7). Au cours de sa conquête de l'Inde, Alexandre captura dix sages et décida de les exécuter, à moins qu'ils ne résolvent correctement le problème qu'il leur poserait. Au cinquième, il demanda quel était le plus ancien, du jour ou de la nuit ; celui-ci répondit : "le jour est le plus ancien, d'un jour seulement". Alexandre s'étant étonné de la réponse, le cinquième gymnosophe ajouta que "les questions dépourvues de sens appellent nécessairement des réponses dépourvues de sens" (τῶν ἀπόρων ἐρωτήσεων ἀνάγχη καὶ τὰς ἀποκρίσεις ἀπόρους εἶναι). Le second point évoqué par BELNAP concerne l'expression anglaise : "un choix à la Hobson" (*HOBSON'S choice*) [Belnap 1976 : 104]. Thomas HOBSON (1544 ? - 1631) est un loueur de chevaux de Cambridge qui possédait une quarantaine de bons chevaux ; quand un client venait, il le conduisait à l'écurie, où le choix était grand, et l'obligeait

à prendre le premier cheval rencontré dans les stalles en partant de la porte, considérant qu'ainsi aucun client n'était mieux ou plus mal servi que les autres. D'où l'expression citée, pouvant se traduire par : "C'est cela, ou rien" (cf. Thomas WARD, *England's Reformation*, IV : "Where to elect there is but one / 'Tis Hobson's choice/ Take that or none"). A la suite de BELNAP, nous avons appelé *question de HOBSON* une question ne présentant qu'une seule réponse possible ; cette notion, sans utilité pratique immédiate, présente un grand intérêt théorique (comme nous le verrons au § 2.2.4.d).

b) *Naissance et développement de la logique des questions.*

Au cours des cinquante années qui s'étendent approximativement de 1875 à 1925, les problèmes du statut logique de la question, et de son rôle par rapport au jugement, ont été abordés par un assez grand nombre d'auteurs ; on en trouvera un aperçu dans [Löw 1928 : 369-374, 428-434], et dans [Moritz 1940 : 124-132]. Il ne semble pas, au travers des mentions qui en sont faites, que ces auteurs aient beaucoup apporté à la logique érotétique ; il y a là toutefois un champ de recherches historiques intéressant à deux égards : les auteurs cités sont parfois des logiciens de grand renom (BOLZANO, HUSSERL, MEINONG), et il est possible que la relecture de certains textes, à la lumière des développements récents de la logique érotétique, aboutisse à des découvertes non négligeables. Si l'on s'en tient aux auteurs dont l'influence, sur l'essor qu'a connu depuis la logique des questions, est attestée, il semble que les premiers textes importants soient : deux articles d'AJDUKIEWICZ, publiés dans les tomes VII (1922-1923) et X (1926-1927) de la revue *Ruch Filozoficzny*, et portant, le premier sur la question de la forme : "Qu'est-ce que P ?", le second sur l'analyse sémantique de la proposition interrogative ; un texte de R. INGARDEN intitulé : *Essentielle Fragen*, publié en 1925 dans le *Jahrbuch für Philosophie und phänomenologische Forschung* ; et enfin un article de Friedrich LÖW sur "la logique de la question", publié en 1928. Nous n'avons pas pu avoir accès aux trois premiers de ces textes ; c'est pourquoi, dans cette présentation de la logique érotétique, le premier texte analysé sera celui de LÖW, ceux d'AJDUKIEWICZ et d'INGARDEN n'étant connus de nous que par les mentions qui en ont été faites ultérieurement.

Ces publications marquent le début d'une première période, dans l'histoire de la logique érotétique, qui s'étend jusqu'à 1940 environ. Cette période est caractérisée par un renouveau de l'intérêt porté à la logique des questions, mais sans que la manière d'aborder les problèmes se soit dégagée de préoccupations proprement philosophiques. Si l'on considère les deux principales bibliographies concernant la logique des questions ([Egli 1976] et [Ficht 1978]), il semble qu'il n'y ait eu que fort peu de publications sur le sujet entre 1940 et 1955, et en tous cas aucun article ayant été repris dans les publications ultérieures. Une seconde période, qui dure encore actuellement, commence avec l'article de WHEATLEY sur les questions de délibération, et celui de Mary et Arthur PRIOR, intitulé : *Erotetic Logic*, publiés tous deux en 1955 ; mais c'est surtout en 1958 que la logique des questions commence à s'affranchir de ses arrière-pensées philosophiques (avec la publication d'un article de C.L. HAMBLIN qui aura une influence importante sur des auteurs tels que BELNAP et HARRAH), et à s'affirmer en tant que branche autonome de la logique moderne (avec une communication de KUBIŃSKI au XIIe congrès international de philosophie).

La première période commence donc peu après 1920. A cette date apparaît un auteur dont l'influence se fera peu à peu sentir sur tous ceux qui s'intéressent aux rapports entre le langage et la pensée, entre la linguistique et la logique. Il s'agit de Ludwig WITTGENSTEIN, dont le *Tractatus Logico-philosophicus*, achevé en 1918, est publié en 1921. L'influence directe de WITTGENSTEIN sur la logique érotétique est assez limitée. Lui-même s'est peu préoccupé du "jeu de langage" qu'est l'interrogation (cf. p.ex. [Wittgenstein 1921 : § 6.5, 6.51 ; 1953 : § 21-24, 47, 516-517]), et il ne semble pas que les auteurs qui se réclament de lui aient beaucoup apporté à la logique des questions (cf. p.ex. : [Stenius 1969 : 251-252, 270]). Par contre, les effets de la problématique de WITTGENSTEIN sont sensibles actuellement chez la plupart des auteurs (en particulier anglo-saxons) qui analysent les relations sémantiques entre les énoncés interrogatifs et la formalisation logique des questions.

Si l'influence immédiate de WITTGENSTEIN sur la logique érotétique semble avoir été inexistante, il n'en est pas de même pour INGARDEN et LÖW. On doit à INGARDEN la définition d'une question bien posée ("Une question est *correcte* si elle a une réponse vraie" ; cité par [Kubiński 1966 : 22]), et l'analyse des conditions de la *solubilité* d'une question (cf. [Löw 1928 : 425-428]) ; les thèses d'INGARDEN ont été ultérieurement reprises (et critiquées) par LÖW, MORITZ, et KUBIŃSKI. La "Logique de la question" de Friedrich LÖW aborde le problème de la "nature" logique et psychologique de l'interrogation, et conclut à la spécificité du "questionnement" (*das Fragen*) en tant que comportement de recherche d'un état de fait existant [Löw 1928 : 435 § 1-3, 436 § 15-17]. Mais ce qui fait l'intérêt actuel de ce texte est moins son apport à la psychologie de l'interrogation que l'analyse logique qu'il développe, et qui concerne la structure de la question, ses présupposés, le mécanisme de la réponse, et les conditions de solubilité de la question. C'est pour quoi nous présentons, au § 2.3.4.b, un résumé de l'apport technique de LÖW à la logique érotétique. En ce qui concerne les écrits d'AJDUKIEWICZ au cours de cette période, nous ne les connaissons que par l'influence qu'ils ont exercée, trente années plus tard, sur Tadeusz KUBIŃSKI. On trouvera toutefois, au § 2.3.4.a, une brève présentation des conceptions d'AJDUKIEWICZ à partir de son ouvrage le plus récent (publié à titre posthume en 1965).

Félix COHEN, qui publie en 1929 un article intitulé "Qu'est-ce qu'une question ?", semble avoir ignoré les travaux que nous venons de mentionner. La thèse centrale de COHEN est qu'une question peut être identifiée à une fonction propositionnelle. En effet, un énoncé interrogatif se présente le plus souvent comme une proposition dans laquelle une constante a été remplacée par un mot interrogatif tel que : "qui ?", "quand ?", "où ?", "pourquoi ?", etc. qui joue le rôle d'une variable. Un énoncé tel que : "César est-il mort ?" constitue une exception apparente à cette description ; mais c'est en raison, dit COHEN, de l'absence dans la langue d'un mot représentant une variable dont les valeurs seraient : "est/n'est pas" [Cohen 1929 : 353-356]. Si l'on prend soin de distinguer entre fonction propositionnelle, et proposition existentielle ou universelle (c'est-à-dire entre une question telle que : " $\sin^2 x = 1$ ", et une assertion comme : " $\forall x : \sin^2 x + \cos^2 x = 1$ " [Cohen 1929 : 358 n 7]), alors

"toute fonction propositionnelle est une question, quoiqu'elle puisse être indéterminée ou dépourvue de signification, et toute valeur d'une telle fonction est une réponse, quoiqu'elle puisse être fausse" [Cohen 1929 : 359]. COHEN distingue les questions *valides* (qui ont une réponse vraie et une seule), les questions *invalides* (qui n'ont aucune réponse vraie), et les questions *indéterminées* (qui ont plusieurs réponses vraies, abstraction faite de toute ambiguïté due au langage) [Cohen 1929 : 359-363]. LEWIS et LANGFORD font remarquer que, si la conception de COHEN est correcte, toute question doit avoir une réponse vraie, car "une fonction propositionnelle peut toujours être transformée en une proposition vraie par spécification ou par généralisation" [Lewis 1932 : 333]. Ainsi, la question : "Qui est l'actuel roi de France ?" donne par généralisation la proposition vraie :
 " $(\exists x : x \text{ est l'actuel roi de France})$ ".

Les articles concernant la logique des questions publiés entre 1930 et 1940 sont de moindre intérêt que les précédents. Toutefois, Moritz SCHLICK, traitant des questions insolubles, distingue entre l'insolubilité logique et l'insolubilité empirique ; une question *insoluble empiriquement* est une question à laquelle il est pratiquement ou techniquement impossible d'apporter une réponse, en raison de circonstances accidentelles (ex. : "Que fit Platon à 8 heures du matin, le jour de ses cinquante ans ?" ; "Combien pesait Homère ?") [Schlick 1935 : 99-101]. Mais il est toujours possible alors de décrire les procédures permettant de découvrir la réponse, même si ces procédures ne peuvent être mises en oeuvre pour des raisons techniques. En dernière analyse, la *signification* d'une question n'est rien d'autre que la description de la méthode à appliquer pour découvrir la réponse ; corrélativement, une question insoluble logiquement est une question sans signification [Schlick 1935 : 102-103]. Par contre, la contribution d'Eugeniu SPERANTIA au Congrès international de philosophie scientifique tenu la même année, apporte peu d'éléments nouveaux, en-dehors de la notion d'*exhaustivité* de la réponse [Sperantia 1935 : 24-27] ; cependant, SPERANTIA peut être considéré un peu comme un précurseur de COLLINGWOOD dans son article sur les présupposés métaphysiques de toute question [Sperantia 1937]. Enfin, un texte très documenté de Manfred MORITZ développe des conceptions assez proches de celles de COHEN, puisqu'il identifie

les questions à des formes propositionnelles pouvant prendre une valeur, et une seulement [Moritz 1940 : 132-137] ; il considère d'autre part que les questions à compléter sont réductibles aux questions de décision, sous la forme d'une disjonction des réponses possibles ("Quelle est la couleur de cette table ?" équivaut à : "Cette table est-elle marron ?", "Cette table est-elle noire ?", etc. [Moritz 1940 : 144]).

La fin de cette première période est marquée par un ouvrage qui, de prime abord, semble assez extérieur à la logique érotétique : l'*Essai sur la métaphysique*, de R.G. COLLINGWOOD, que nous avons déjà cité à propos de la notion de présupposition (§ 2.1.1. a) et des questions multiples (§ 2.1.2. a). Cet auteur développe une théorie très intéressante sur les présupposés de toute question, théorie que l'on peut résumer en quelques propositions [Collingwood 1940 : 21-33] :

- Toute assertion, faite par qui que ce soit en quelque circonstance que ce soit, est faite en réponse à une question ; celle-ci lui est logiquement antérieure.

- Toute question contient des présupposés ; les uns sont contenus directement (présupposés immédiats), les autres indirectement (présupposés des présupposés).

- Le présupposé immédiat doit être assumé par le locuteur pour que celui-ci puisse poser la question ; sinon, la question est une question absurde, elle "ne se pose pas".

- "Assumer" signifie : "supposer par une libre décision" ; il n'est pas nécessaire de croire en la vérité d'un présupposé pour l'assumer.

- Il existe des présupposés relatifs et des présupposés absolus. Les premiers jouant le rôle de présupposition d'une question, et de réponse à une autre question (ex. : "Vous avez l'habitude de battre votre femme"). Les présupposés absolus ne sont réponse à aucune question (ex. : "Tout ce qui arrive a une cause") ; ils sont invérifiables.

- Toute question traitant un présupposé absolu comme une proposition (ex. : "Est-il vrai que tout ce qui arrive a une cause ? Quelle preuve en

a-t-on ? Comment cela peut-il être démontré ?") est une question absurde. D'autre part, selon COLLINGWOOD, les présupposés absolus ne sont pas des propositions, mais des faits historiques [Collingwood 1940 : 66] ; et l'analytique transcendantale (cf. [Kant 1781 : 2ème partie, 1ère division]) correspond à l'étude des présuppositions [Collingwood 1940 : 237-242].

La seconde période dans l'histoire de la logique érotétique commence en 1955, avec deux publications. La première en date, de J.M.O. WHEATLEY, attire l'attention sur un type particulier de question, la question de délibération. L'auteur prend pour point de départ la distinction, due à Renford BAMBROUGH, entre les *questions d'investigation*, dont la réponse est une information qui peut être vraie ou fausse, et les *questions de délibération*, dont la réponse est une décision ou une résolution [Wheatley 1955 : 49-50] (Cette distinction recoupe une distinction faite par CARNAP entre questions *théoriques* et questions *pratiques* [Wheatley 1955 : 55]). Exemples de questions de délibération : "Est-ce que je vais voter Conservateur ou Travilliste aux prochaines élections ?" ; "Quand nous réunirons-nous de nouveau ?" ; "Partirons-nous pour la lune la semaine prochaine ?" [Wheatley 1955 : 50, 55, 59]. Entre également dans cette catégorie la question, pour des jurés, de savoir si l'accusé est coupable ou non coupable ; pour des naufragés, de déterminer si l'on est ou non jeudi aujourd'hui ; pour des personnes interrogées au cours d'un sondage, d'être d'accord ou pas d'accord avec un énoncé exprimant une opinion [Wheatley 1955 : 51-52]. Convenons qu'une expression E est dite avoir un contenu (sémantique) délibératif "si elle peut être utilisée pour exprimer la décision d'un individu X d'exécuter une action possible (ou une série d'actions possibles) Y" [Wheatley 1955 : 58]. Une question de délibération est une question admettant comme réponse ce type d'expression. Elle a un sens même si l'action envisagée n'est pas réalisable (ex. : "Nous partirons pour la lune la semaine prochaine") ; mais elle n'en a pas si l'expression E est absurde (ex. : "Nous partirons pour la lune hier") ou contradictoire (ex. : "Je voterai Conservateur quand je voterai Libéral"). La réponse à une question de

délibération n'a pas de valeur de vérité ; par contre, la résolution exprimée par la réponse peut être *assumée* ou non (p.ex. : si l'on a décidé de voter Conservateur, la réponse : "Je voterai Conservateur" est assumée, la réponse : "Je voterai Libéral" ne l'est pas) [Wheatley 1955 : 58-59]. Ce texte ne semble pas avoir eu de postérité, à part un article de Bernard MAYO paru l'année suivante, dans lequel l'auteur considère que la réponse à une question de délibération est en réalité toujours une injonction ; prendre une décision, c'est en dernière analyse accepter de se soumettre à cette injonction [Mayo 1956 : 63].

La seconde publication de l'année 1955 est l'article de Mary et Arthur PRIOR, que nous avons déjà mentionné. Il rappelle un certain nombre de notions propres à la logique des questions, ou applicables à celle-ci. En particulier, la distinction de WHATELY entre questions *whether* et questions *what* [Prior 1955 : 44-45] ; la description de deux procédures de réduction des questions *whether* aux questions *what*, inspirées de WHATELY [Prior 1955 : 46-49] ; et l'extension aux questions des classifications des jugements selon le rapport de pensée (la relation) qu'ils impliquent [Prior 1955 : 51-55] et selon leur modalité [Prior 1955 : 55-58] (cf. [Kant 1781 : 2e partie, 1ère div., 1er livre, I 2 § 9]). Les auteurs ignorent manifestement les travaux de la plupart de leurs devanciers, et en particulier ceux de LÖW. Par rapport à ceux-ci, le principal intérêt de cet article réside dans l'analyse qui est faite de l'interrogation authentiquement conditionnelle (ex. : "Si vous mettez un chapeau ce soir, de quelle couleur sera-t-il ?"), distinguée de l'interrogation simple contenant une condition hypothétique (ex. : "Si vous étiez poursuivi par un taureau, que feriez-vous ?") ; dans l'interrogation conditionnelle, "si l'antécédent est faux, il n'y a tout simplement pas de question", alors que dans une question hypothétique, refuser l'hypothèse ne dispense pas de répondre à la question [Prior 1955 : 52-53]. Par contre, la conclusion de l'article, selon laquelle la logique érotétique ne pourra pas être développée dans le cadre de la logique symbolique [Prior 1955 : 59], ne s'est pas trouvée confirmée par les événements.

Au cours des années suivantes, il faut mentionner le premier article de Gerold STAHL, intitulé : "La logique des questions", paru en 1956 dans les *Anales de la Universidad de Chile*, et les *Principles of Right Reason*, de Henry

S. LEONARD (1957). La théorie de STAHL a fait ultérieurement l'objet d'autres articles, et nous en donnons une vue d'ensemble au § 2.3.4.c ; sa caractéristique principale est d'identifier une question à l'ensemble de ses réponses possibles. L'apport de LEONARD à la logique érotétique est surtout de type classificatoire ; nous avons présenté, au § 2.1.2.a, ses divisions des présuppositions en primaires/secondaires d'une part, factuelles/formelles d'autre part. En ce qui concerne les questions, LEONARD les classe en questions *pratiques* (Que faire dans une situation concrète donnée ?) et questions *théoriques* (concernant la nature ou les caractéristiques d'un élément du monde qui nous entoure) [Leonard 1967 : 3-5, 31-32], ainsi qu'en questions *directes* (à réponse *oui/non*), *semi-directes* (choix entre plusieurs modalités) et *indirectes* (analogues aux questions *what*, ou questions à compléter) [Leonard 1967 : 33-34]. La première distinction est identique à celle de BAMBROUGH et de CARNAP rapportée par WHEATLEY. D'autre part, LEONARD considère qu'une question est *valide* si et seulement si elle est précise, sans ambiguïté, et n'a pas de présuppositions fausses [Leonard 1967 : 41-42]. Une contribution plus originale de LEONARD à la logique érotétique est sa thèse exposée ultérieurement, selon laquelle un énoncé interrogatif a une valeur de vérité (indépendamment du caractère "honnête" ou "malhonnête" de la question), identique à celle de son contenu propositionnel (cf.: [Leonard 1959 ; 1961], [Wheatley 1961], [Hurrel 1964]).

La publication, en 1958, de l'article de C.L. HAMBLIN intitulé : *Questions*, marque un tournant dans l'histoire de la logique érotétique, et le début de l'essor véritable de cette branche nouvelle de la logique. HAMBLIN aborde en effet le problème de l'analyse logique des questions sous un angle radicalement nouveau : au lieu de présenter la définition précise de notions telles que "question", "réponse", etc., il propose un système axiomatique relatif au statut logique de ces notions, et à leurs interrelations. Les principaux éléments de cette axiomatique sont [Hamblin 1958 : 162-165] :

- *Postulat 1* : Toute réponse à une question est une proposition. Il importe par conséquent de distinguer entre la réponse elle-même, et son codage (un mot, un geste, une croix devant une phrase écrite, etc.).

- *Postulat 2* : Connaître ce qui a valeur de réponse à une question donnée est équivalent à connaître la question. Ceci n'a de sens que si l'on accepte le postulat 1, c'est-à-dire si les réponses considérées sont toujours des propositions.

- *Corrolaire* : Deux questions qui ont exactement les mêmes réponses sont équivalentes.

- *Postulat 3* : Les réponses possibles à une question constituent un ensemble exhaustif d'éventualités mutuellement exclusives. L'exhaustivité de l'ensemble assure la solubilité de la question, puisqu'ainsi aucune des possibilités logiquement envisageables n'est exclue (sinon, la question reste acceptable *relativement* à certains présupposés). Le caractère mutuellement exclusif des éventualités élimine les réponses incomplètes.

- *Théorème 1* : Toute question n'ayant qu'une réponse possible est une tautologie ; réciproquement, si la réponse à une question est une tautologie, elle est l'unique réponse possible à cette question.

- *Théorème 2* : Toute question a une réponse.

La publication de ce système axiomatique a eu un grand retentissement auprès des logiciens s'intéressant aux questions ; et en particulier, le deuxième postulat a été discuté et finalement adopté par la plupart d'entre eux. Signalons en outre que, dans ce texte relativement court (dix pages), HAMBLIN :

- énonce quelques propriétés de moindre importance (la disjonction de toutes les réponses possibles à une question est une tautologie ; une proposition contradictoire ne peut être une réponse à une question ; une question contradictoire est valide, et sa réponse est une tautologie) [Hamblin 1958 : 165] ;

- suggère la définition de certaines relations entre questions (inclusion, équivalence, conjonction), et d'une caractéristique particulière, la *force* de la question (qui est nulle lorsque la question n'a qu'une réponse possible) [Hamblin 1958 : 166] ;

- met en évidence les analogies entre une logique des questions et la théorie de l'information (une question équivaut à une partition des univers

possibles ; le contenu informationnel d'un message dépend non seulement des caractéristiques du message, mais aussi du nombre et des caractéristiques des autres messages possibles) [Hamblin 1958 : 166-168].

Par ailleurs, ainsi que nous l'avons signalé, l'année 1958 est marquée par la communication de Tadeusz KUBIŃSKI au XIIe congrès international de philosophie (Venise, septembre 1958), qui présente en quelques pages l'embryon d'une théorie formelle complète de l'interrogation, fondée sur une extension du calcul des propositions. Mais, en dépit de son importance, cette communication ne semble pas avoir eu la même influence immédiate que l'article de HAMBLIN. A partir de 1958, le nombre et la variété des travaux relatifs à la logique des questions commencent à croître rapidement. L'importance numérique et technique de ces travaux se traduira même par la citation de la logique érotétique dans des ouvrages destinés à un large public, comme l'*Encyclopaedia Britannica* [Resher 1974]. Peut être y-a-t-il eu à cette floraison d'articles, et même de livres entiers consacrés aux questions, deux causes essentielles : la découverte d'une possibilité d'extension de la logique formelle encore peu explorée, et l'émergence de nouveaux champs d'application pour une logique de l'interrogation. La réalité de ce dernier facteur est d'ailleurs attestée par le développement de recherches appliquées à divers domaines extra-linguistiques, tels que les systèmes de communications (cf. p.ex. : [Harrah 1963 : 25-53] et ci-dessous le § 2.3.4.d, ainsi que [Mac Kay 1969 : 31-38, 94-104]), le dialogue avec des civilisations extra-terrestres (cf. p.ex. : [Freudenthal 1960 : 52, 90-114, 140-141]), la philosophie des sciences (cf. [Bromberger 1966 a], et ci-dessous le § 2.3.3.), voire le questionnaire psycho-social (cf. [Pawłowski 1969]).

Au niveau de la conception de systèmes formels, il y a lieu de signaler quelques travaux isolés, comme les systèmes dus à Martin STEINMAN Jr. (qui formule en particulier une règle d'interprétation sémantique valable à la fois pour une assertion et une question [Steinman 1959]), et à M.J. CRESSWELL (qui formalise la relation "être une réponse à une question" à partir de la logique modale [Cresswell 1963]). Toutefois, il est à peu près certain que les conceptions récentes, mais relativement mineures, de la logique des questions n'ont

pas encore été toutes recensées, et que des recherches devraient être entreprises dans cette direction (ainsi, par exemple, il est fait mention, dans [Fatkin 1970], d'un travail sur "la logique des questions et des réponses" dû à E.K. VOÏCHVILLO, publié en 1966 ; ce texte ne figure dans aucune bibliographie consacrée à la logique érotétique). Par contre, si l'on s'en tient aux grandes théories qui ont été publiées, il semble bien que nous soyons parvenu à faire une recension exhaustive. Aussi décrivons-nous, dans les pages qui suivent, les trois conceptions majeures de la logique érotétique contemporaine : celles de Nuel D. BELNAP Jr. (§ 2.2.), de Lennart ÅQVIST (§ 2.3.1), et de Tadeusz KUBIŃSKI (§ 2.3.2), qui représentent trois directions de recherches bien différenciées. Nous y présentons en outre un résumé de quelques autres théories qui sont soit plus anciennes (LÖW , § 2.3.4.b ; AJDUKIEWICZ, § 2.3.4.a), soit moins originales (HARRAH, § 2.3.4.d) ou moins fécondes (STAHL, § 2.3.4.c), soit encore moins générales (BROMBERGER, § 2.3.3) que les trois grands systèmes évoqués ci-dessus.

2.2. LA LOGIQUE DES QUESTIONS ET DES RÉPONSES DE Nuel D. BELNAP

L'analyse des questions de Nuel D. BELNAP Jr. constitue l'une des tentatives les plus complètes (avec celles de Lennart ÅQVIST et Tadeusz KUBIŃSKI) pour édifier une logique érotétique. Le but visé est la réalisation d'un langage formel qui facilite la communication scientifique entre un questionneur et un répondant. "Au niveau du langage-objet, nous voulons créer un système formel soigneusement construit, qui puisse servir à poser des questions et à y répondre. Au niveau du méta-langage, nous voulons élaborer un ensemble de concepts permettant de classer les questions et les réponses, de les évaluer, et de définir entre elles des relations" [Belnap 1976 : 1].

Pour BELNAP, un point fondamental de la logique érotétique est la relation question-réponse [Belnap 1966 : 609]. Aussi prend-il comme point de départ le postulat de HAMBLIN, qu'il reformule ainsi : "Ne pas comprendre une question, c'est précisément ne pas savoir ce qui compte pour une réponse à cette question" [Belnap 1963 : 58]. Par conséquent, la connaissance de ce qui compte pour une réponse à une question q est une condition nécessaire à la connaissance de cette question q . Dire qu'une question q est imprécise, ambiguë, ou dépourvue de sens, équivaut à dire que, quel que soit A , la proposition : " A est une réponse à la question q " est elle-même imprécise, ambiguë, ou dépourvue de sens. D'où l'importance de la notion de *réponse directe* à une question, et du *critère de décidabilité* ainsi formulé : "A partir du seul examen des caractéristiques syntaxiques d'une question, il doit être possible de décider effectivement si une expression donnée est ou non une réponse directe à la question" [Belnap 1963 : 58].

Nous ne présentons ci-après que l'essentiel (dans notre perspective) des résultats auxquels BELNAP est parvenu : principaux concepts de base de la logique érotétique, description morphologique des types de questions et des combinaisons de questions, et analyse sémantique des questions et des réponses. Dans la mesure où la terminologie et la symbolique utilisées par l'auteur ont évolué depuis sa première publication sur le sujet [Belnap 1963], nous avons dans la quasi-totalité des cas donné la préférence sur ces points aux travaux les plus récemment publiés, en particulier à *The Logic of Questions and Answers*

[Belnap 1976], écrit en collaboration avec Thomas B. STEEL.

2.2.1. Les concepts de base

a) Les fondements assertoriques

Pour construire sa théorie, BELNAP utilise un langage formel L , qui n'est autre que le calcul appliqué des prédicats du premier ordre muni de la relation d'identité. Ce langage a pour éléments :

- des variables d'individus : w, x, y, z, \dots ;
- des constantes d'individu ;
- des constantes de fonction n -aires : f, g, \dots , servant à désigner des fonctions ayant des individus pour valeurs, et des n -uples d'individus comme arguments ;
- des constantes de prédicat n -aires : F, G, \dots , servant à désigner des classes de n -uples d'individus ;
- la relation d'identité ($=$) ;
- un connecteur propositionnel à une place, la négation (\neg) ;
- des connecteurs propositionnels à deux places : conjonction ($\&$), disjonction (\vee), implication (\rightarrow), équivalence (\leftrightarrow) ;
- les quantificateurs universel (\forall) et existentiel (\exists) ;
- des signes de ponctuation : parenthèses, virgule, barre oblique.

Termes et expressions sont définis par récurrence selon les règles habituelles. Ainsi pour les *termes* :

- ① Les constantes d'individu et les variables d'individu sont des termes ;
- ② Si a_1, \dots, a_n sont des termes, et si f est une constante de fonction n -aire, alors $f(a_1, \dots, a_n)$ est un terme.

La formation des *expressions* est régie par les règles suivantes :

- ① Si a_1 et a_2 sont des termes, alors $(a_1 = a_2)$ est une expression bien formée ;
- ② Si F est une constante de prédicat n -aire, et si a_1, \dots, a_n sont des termes, alors $F(a_1, \dots, a_n)$ est une expression bien formée ;
- ③ Si A est une expression bien formée, alors $(\forall x)A$ et $(\exists x)A$ sont des expressions bien formées ;
- ④ Si A est une expression bien formée, alors \bar{A} est une expression bien formée ;
- ⑤ Si A et B sont des expressions bien formées, alors $(A \& B)$, $(A \vee B)$, $(A \rightarrow B)$, $(A \leftrightarrow B)$ sont des expressions bien formées.

L'utilisation des quantificateurs conduit à introduire, par commodité, les définitions suivantes. On appellera :

- *occurrence liée* de la variable x dans l'expression A , toute occurrence de x figurant dans une partie bien formée de A , de la forme $(\forall x)B$ ou $(\exists x)B$;
- *occurrence libre* de x dans A , toute occurrence de x qui n'y est pas liée ;
- *variable liée* dans l'expression A , toute variable x ayant dans A des occurrences liées ;
- *variable libre* dans A , toute variable x ayant des occurrences libres dans A ;
- *nom*, tout terme sans variable libre (servant par conséquent à désigner un individu) ;
- *référent* d'un nom, l'individu unique désigné par ce nom ;
- *proposition*, toute expression bien formée ne contenant aucune variable libre ;

- *forme propositionnelle n-aire*, toute expression bien formée contenant exactement n variables libres.

Pour les formes propositionnelles, on adoptera la convention d'écriture suivante : $Ax_1 \dots x_n$, désignant une forme propositionnelle n -aire quelconque A , ayant n variables libres x_1, \dots, x_n . L'intérêt de cette écriture est de faire figurer explicitement le nom des variables libres de A . Le résultat de l'opération de substitution, dans la forme propositionnelle $Ax_1 \dots x_n$, d'une occurrence de chacun des termes b_i à chaque occurrence de chacune des variables libres x_i , s'écrira par conséquent : $Ab_1 \dots b_n$.

Un sous-ensemble distingué des formes propositionnelles unaires joue un rôle particulier en logique des questions. Il s'agit des *conditions catégorielles élémentaires*. A chaque condition catégorielle élémentaire Cx est associée une classe de substitution, appelée *catégorie nominale* ; c'est un ensemble de noms, représentant l'ensemble des valeurs que peut prendre la variable x . Par exemple, soit la condition catégorielle $Cx = "x \text{ est un océan}"$, la catégorie nominale associée à Cx est l'ensemble des noms d'océans : {Atlantique, Arctique, Indien, ...}. Un même référent pouvant être désigné sous plusieurs noms différents, il nous faut distinguer la catégorie nominale (ensemble de noms) de la catégorie réelle (ensemble de référents). C'est pourquoi à chaque condition catégorielle est également associé un ensemble d'entités, appelé *catégorie réelle* ; entre la catégorie nominale et la catégorie réelle associées à une même condition catégorielle, il existe une relation sémantique qui, à chaque nom, fait correspondre une entité et une seule. La liste des *prédicats catégoriels* entrant dans les conditions catégorielles élémentaires est supposée donnée dans le langage L . Les conditions catégorielles non élémentaires sont formées par récurrence : seules sont des conditions catégorielles non élémentaires les expressions formées par la conjonction ou la disjonction de conditions catégorielles, élémentaires ou non.

Les *règles d'interprétation* de ce langage formel sont celles du calcul des prédicats. A chaque espèce de symboles désignant des individus, des fonctions, ou des prédicats, est associé un ensemble non-vide appelé *domaine*,

dans lequel les symboles de l'espèce correspondante prendront leurs valeurs. Soit D le domaine des individus ; le domaine d'un prédicat n -aire F est le sous-ensemble de D^n , dont les éléments sont les n -uples d'individus pour lesquels la relation définie par F est vraie. Une *interprétation* M est définie comme l'ensemble des domaines, et l'ensemble des valeurs prises par les constantes (d'individu, de fonction, de prédicat). Si l'on appelle D le domaine des individus, on peut également définir une interprétation comme un couple ordonné formé de D , et d'une *fonction d'interprétation*, ayant pour arguments les différentes constantes et variables ; cette fonction a pour valeurs des individus de D (pour les constantes et les variables d'individus), des fonctions sur D (pour les constantes de fonctions), et des relations sur D (pour les constantes de prédicat) [Belnap 1976 : 9]. Une *assignation de valeurs* α dans l'interprétation M attribue à chacune des variables une valeur prise dans son domaine ; c'est par conséquent une application de l'ensemble des variables d'individu dans D .

Les *règles d'évaluation* de la valeur d'une expression bien formée doivent permettre de décider si l'expression est "vraie (ou fausse) dans l'interprétation M pour l'assignation de valeurs α à ses variables libres" [Belnap 1963 : 55]. Les principales règles d'interprétation sont (en adoptant les abréviations : "ssi" pour : "si et seulement si", et : "vrai dans $M\alpha$ " pour : "vrai dans l'interprétation M pour l'assignation de valeurs α ") :

① $(a_1 = a_2)$ est vraie dans $M\alpha$ ssi a_1 et a_2 ont le même référent dans $M\alpha$;

② $F(a_1, \dots, a_n)$ est vraie dans $M\alpha$ ssi le n -uple formé par les référents de a_1, \dots, a_n dans $M\alpha$ est dans la classe des n -uples associés à F dans $M\alpha$;

③a $(\forall x)A$ est vraie dans $M\alpha$ ssi A est vraie dans $M\alpha$ et dans toute assignation $M\beta$ ne différant de $M\alpha$ que par la valeur donnée à x ;

③b $(\exists x)A$ est vraie dans $M\alpha$ ssi A est vraie dans $M\alpha$ ou dans au moins une assignation $M\beta$ ne différant de $M\alpha$ que par la valeur donnée à x ;

④ \bar{A} est vraie dans $M\alpha$ ssi A n'est pas vraie dans $M\alpha$;

⑤a) $(A \& B)$ est vraie dans $M\alpha$ ssi A est vraie dans $M\alpha$ et B est vraie dans $M\alpha$;

⑤b) $(A \vee B)$ est vraie dans $M\alpha$ ssi, de A et de B , l'une au moins des deux expressions est vraie dans $M\alpha$;

⑤c) $(A \rightarrow B)$ est vraie dans $M\alpha$ ssi l'une au moins des expressions \bar{A} et B est vraie dans $M\alpha$;

⑤d) $(A \leftrightarrow B)$ est vraie dans $M\alpha$ ssi les deux expressions $(A \rightarrow B)$ et $(B \rightarrow A)$ sont toutes deux vraies dans $M\alpha$.

C'est à partir de ces bases "assertoriques" que l'auteur édifie sa logique des questions et des réponses. Il a besoin pour cela d'un certain nombre de notions et de symboles complémentaires, que nous introduirons au fur et à mesure de leur nécessité dans le cours de l'exposé.

b) La structure des questions élémentaires

Soit une question (abstraite) q . On appelle *forme interrogative* I (en abrégé : *interrogative*) un énoncé (lexical) correspondant à la question q ; cet énoncé est dit *poser* la question q . Deux interrogatives distinctes peuvent se trouver poser la même question q .

Questions et interrogatives élémentaires se composent de deux parties distinctes : le *sujet* σ , qui présente un ensemble de modalités de réponses, et la *requête* ρ , qui indique au répondant la forme que devra prendre sa réponse. Lorsqu'il sera nécessaire de distinguer le sujet d'une interrogative du sujet de la question à laquelle elle correspond, on appellera le premier *sujet lexical* et le second *sujet abstrait*.

Toute question élémentaire est de la forme :

$$q = ? \rho \sigma$$

En d'autres termes, poser une question revient à proposer à l'interlocuteur un ensemble de matériaux (*modalités de réponse*) qui lui serviront à construire sa

réponse, et à lui imposer une manière de sélectionner et de présenter ces matériaux [Belnap 1963 : 18]. Cette définition est particulièrement bien adaptée aux questions posées dans les tests de connaissances ou d'aptitudes, et dans les questionnaires d'opinion, d'attitudes ou de personnalité, dans lesquelles la *consigne* (requête) est toujours explicite et distincte des modalités soumises au jugement de la personne interrogée.

La forme du sujet σ détermine le type de la question posée. Parmi l'ensemble des formes possibles, BELNAP distingue deux cas particuliers, dont il développe l'analyse :

- présentation des modalités en extension, qui détermine les questions *whether* (exemple : "Portait-elle son chapeau rouge ou son chapeau vert ?") ;

- présentation des modalités en compréhension, qui détermine les questions *which* (exemple : "Qui est l'actuel roi de France ?").

Dans le premier cas, les modalités de la réponse figurent explicitement dans l'énoncé et sont en nombre fini ("elle portait son chapeau rouge"/"elle portait son chapeau vert") ; dans le second, elles doivent être construites à partir du contenu de l'énoncé et peuvent être en nombre infini ("Untel est l'actuel roi de France", avec "Untel" = "Louis XIX"/"Charles XI"/"François II"/etc.). Nous verrons aux paragraphes suivants comment les différentes formes de sujet peuvent être réécrites dans le langage formel proposé par BELNAP.

La forme de la requête est, à quelques réserves près, la même pour tous les types de questions. Elle comporte trois éléments : la dimension de sélection s , l'exigence d'exhaustivité c , et l'exigence de distinction d :

$$\rho = (s \ c \ d)$$

Et par conséquent :

$$q = ? (s \ c \ d) \sigma$$

On appelle *sélection* S l'ensemble des modalités qu'une réponse sélectionne parmi l'ensemble des modalités du sujet σ .

La *dimension de sélection* s indique le nombre minimum v et le nombre maximum u de modalités que la réponse demandée doit sélectionner. Elle s'écrit $s = \frac{u}{v}$ avec $0 < v \leq u$. Il peut arriver que le nombre maximum u ne soit pas précisé ; on écrira alors $s = \frac{-}{v}$.

L'*exigence d'exhaustivité* c demande au répondant de préciser qu'il a bien sélectionné toutes les modalités qu'il estime vraies. Par exemple, à l'interrogation : "Portait-elle son chapeau rouge ou son chapeau vert ?", une réponse possible est la phrase : "Elle portait son chapeau rouge". Si l'exigence d'exhaustivité est formulée dans la question sous-jacente, la réponse sera : "Elle portait son chapeau rouge, et non son chapeau vert". L'utilité de cette exigence est plus évidente dans la question : "Quels sont *tous* les nombres premiers compris entre 10 et 20 ?", qui a pour réponse possible (d'ailleurs fausse) : "Il n'y a pas de nombres premiers compris entre 10 et 20, autres que 11, 13, et 17" (cf. [Belnap 1963 : 27-34]). Dans l'énoncé de la question, l'exigence d'exhaustivité est souvent exprimée par les mots : "tous", "la plupart", "au moins un tiers", "5 %", etc. Nous nous limiterons dans cet exposé aux cas extrêmes que sont l'exigence d'exhaustivité (maximum), et l'absence d'exigence d'exhaustivité. Par conséquent, lorsque l'exigence d'exhaustivité figure dans la question, elle est maximum, et s'écrit : $c = \forall$; dans le cas contraire, on écrit $c = -$.

L'*exigence de distinction* est nécessaire lorsque certains des noms cités dans la sélection risquent de désigner un même référent. Par exemple : "7" et "VII" ; "l'auteur de Macbeth" et "Shakespeare". Lorsque cette exigence est spécifiée, on écrit : $d = \neq$; dans le cas contraire, on écrit : $d = -$.

En combinant ces divers éléments, on peut distinguer un grand nombre de types de requêtes observables dans les formes d'interrogation. BELNAP en privilégie six [Belnap 1976 : 75] :

- l'exemple unique : $\rho = \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right)$;
- les exemples multiples : $\rho = \left(\begin{array}{c} - \\ 1 \end{array} - - \right)$;
- les exemples multiples et distincts : $\rho = \left(\begin{array}{c} - \\ 1 \end{array} - \neq \right)$;
- la réponse vraie unique : $\rho = \left(\begin{array}{c} 1 \\ 1 \end{array} \forall - \right)$;

- l'énumération exhaustive : $\rho = (\bar{1} \vee -)$;
- l'énumération exhaustive et distincte : $\rho = (\bar{1} \vee \neq)$.

c) Le concept de réponse directe à une question.

Une *réponse directe* $d(I)$ à une interrogation élémentaire I est une réponse qui répond entièrement à la question, sans imposer au locuteur de tâche complémentaire, et sans lui apporter plus d'information qu'il n'en demande.

"Pour présenter les choses sous un angle psychologique, nous dirons qu'une réponse directe est précisément le type de réaction que le locuteur entend provoquer par sa question" [Belnap 1976 : 13]. Soit par exemple l'interrogation :

- "Quelle est, en degrés centigrades, la température de congélation de l'eau dans des conditions normales ?"

Sont considérées comme des réponses directes à cette question les phrases :

- "La température de congélation de l'eau dans des conditions normales est zéro degrés centigrade" ;
- "La température de congélation de l'eau dans des conditions normales est 37 degrés centigrades".

Nous voyons que la valeur de vérité de la proposition n'intervient pas dans son appréciation en tant que réponse directe. Par contre, la forme de la réponse est déterminante. N'est pas une réponse directe la proposition suivante :

- "La température de congélation de l'eau dans des conditions normales est indiquée dans tous les manuels de chimie élémentaire".

En effet, bien qu'elle fournisse au locuteur le moyen de découvrir l'information qu'il cherche, elle lui impose une tâche complémentaire pour obtenir une réponse complète et précise. Il en est de même pour les réponses :

- "La température de congélation de l'eau dans des conditions normales est 32 degrés Fahrenheit" ;
- "La température de congélation de l'eau dans des conditions normales est $\sqrt{\pi}$ degrés centigrades".

D'autre part, ne peut être considérée comme une réponse directe la phrase :

- "La température de congélation de l'eau dans des conditions normales est zéro degré centigrade, et sa température d'ébullition est cent degrés centigrades".

Bien que cette phrase apporte toute l'information demandée, elle apporte également une information superflue, hors du sujet. D'un point de vue pratique, ses inconvénients (qui la font écarter des réponses directes) sont nombreux : risque d'égarer le locuteur, coût plus élevé en termes de recherche d'information et d'analyse du contenu de la réponse, etc.

Enfin, on dira que la réponse :

- "zéro degré centigrade"

est une *réponse codée*, c'est à dire une abréviation, interprétable grâce au contexte, d'une réponse directe (cette terminologie est empruntée à HAMBLIN).

Ajoutons que l'on donne le nom de *réplique* à "la multitude des bruits plus ou moins chargés de sens qui peuvent faire suite à une question" [Belnap 1963 : 20-21], et qui ne sont pas des réponses directes, tels que :

- "Je ne sais pas",
- "C'est une bonne question",
- "Demandez à Untel".

D'un point de vue formel, une réponse directe $d(I)$ à une interrogative élémentaire I est une expression A , composée en principe de trois éléments : la sélection S , l'affirmation d'exhaustivité C , et l'affirmation de distinction D . Ces trois éléments sont unis par le connecteur $\&$; pour que la réponse soit vraie, il faut (et il suffit) que ses trois éléments soient vrais simultanément. En première approximation, sa forme générale est donc :

$$d(I) = S \& C \& D.$$

Nous avons déjà vu que la *sélection* est l'ensemble des modalités présentées par le sujet σ et retenues par le répondant ; son contenu dépend de la nature de σ . La sélection est le seul élément toujours obligatoire d'une réponse directe.

L'affirmation d'exhaustivité "est toujours la déclaration du degré d'exhaustivité de la sélection, relativement au nombre total des modalités vraies présentées par la question" [Belnap 1976 : 50]. Lorsque l'exigence d'exhaustivité est absente de l'interrogative, l'affirmation d'exhaustivité est absente de la réponse (on pourra dire également que cette affirmation est *vide*, ou *toujours vraie* ; cf. [Belnap 1963 : 96]) ; lorsqu'elle est maximum, on l'écrit : $c = \text{Max}(\sigma, S)$. L'affirmation de distinction (ou de *non redondance* [Belnap 1963 : 117]) est la déclaration que chaque élément de la sélection S désigne un individu (réel) différent. Lorsque l'exigence de distinction est absente de la question, l'affirmation de distinction est absente de la réponse ; lorsqu'elle est présente, elle s'écrit : $D = \text{Dist}(\sigma, S)$.

A partir de ces définitions, il est possible de décrire quatre types de réponses directes, selon que la requête de la question comporte ou non les exigences d'exhaustivité et de distinction :

$\rho = (s - -)$	$d(I) = S$
$\rho = (s \vee -)$	$d(I) = S \ \& \ \text{Max}(\sigma, S)$
$\rho = (s - \neq)$	$d(I) = S \ \& \ \text{Dist}(\sigma, S)$
$\rho = (s \vee \neq)$	$d(I) = S \ \& \ \text{Max}(\sigma, S) \ \& \ \text{Dist}(\sigma, S)$.

La dimension de sélection s influe sur le cardinal de S , mais non sur la forme de la réponse directe. Nous allons voir maintenant de manière plus détaillée la forme et le contenu des réponses directes aux questions *whether* et aux questions *which*.

2.2.2. Types de questions fondamentaux

a) Les quatre formes distinguées de questions *whether*

Les questions *whether* sont considérées par l'auteur comme le type de question élémentaire le plus simple, c'est-à-dire le plus aisé à formaliser. Dans les questions de ce type, en effet, le sujet énumère explicitement les éléments possibles de réponse (modalités) ; il y a une quasi-identité entre le sujet lexical et le sujet abstrait ; et l'exigence de distinction n'est jamais posée.

Le sujet abstrait d'une question *whether* est un ensemble de propositions : $\{A_1, \dots, A_n\}$. En principe, un ensemble unique de propositions peut être représenté par divers sujets lexicaux, principalement en faisant varier l'ordre des propositions, ou en répétant certaines d'entre elles [Belnap 1963 : 71]. Toutefois, pour simplifier la présentation formelle, on admettra que le sujet lexical énumère les propositions du sujet abstrait dans le même ordre, et sans répétitions (ni omission, évidemment). On écrira le sujet lexical d'une question *whether* :

$$\sigma = (A_1, \dots, A_n) ;$$

et par conséquent, la forme générale de ce type d'interrogative s'écrit :

$$I = ? \rho(A_1, \dots, A_n).$$

On admettra en outre, à titre provisoire, qu'aucune des modalités de réponse A_i présentées dans le sujet n'est une conjonction dont tous les membres sont eux-mêmes des modalités de réponse. Par exemple, on écartera une séquence telle que : $\sigma = (B, A, C, A\&C, B\&A\&C)$; mais on acceptera par contre la séquence : $\sigma = (B, A, C, A\&D, B\&A\&E)$.

Les interrogatives de forme *whether* sont aisées à identifier, et BELNAP en donne de nombreux exemples. Ainsi :

$$I = \text{"La prochaine crise concernera-t-elle Cuba ou Berlin ?"}$$

dont le sujet lexical peut s'écrire :

$$\sigma = (\text{"La prochaine crise concernera Cuba"}, \text{"La prochaine crise concernera Berlin"}).$$

Ou encore :

$$I = \text{"Parmi les lecteurs de bande 1,2 et 3, lequel est disponible ?"}$$

avec : $\sigma = (\text{"Le lecteur 1 est disponible"}, \text{"Le lecteur 2 est disponible"}, \text{"Le lecteur 3 est disponible"}).$

Ou enfin :

$I = \text{"Du boeuf, du veau, ou du mouton, lequel est en promotion aujourd'hui ?"}$,

avec : $\sigma =$ ("Le boeuf est en promotion aujourd'hui", "Le veau est en promotion aujourd'hui", "Le mouton est en promotion aujourd'hui").

L'exigence de distinction ne s'applique pas aux questions *whether*, dans la mesure où, contrairement aux questions *which*, la sélection qu'elles requièrent ne porte pas sur des noms d'individus, mais sur des propositions. Si redondance il y a, celle-ci est déjà présente dans le sujet abstrait de la question, et la formalisation adoptée par l'auteur ne permet pas d'exiger du répondant qu'il examine les relations d'implication possibles entre les modalités de réponse qui lui sont proposées (cf. [Belnap 1976 : 61]). En conséquence, à partir des six types de requête qu'il a distingués, BELNAP définit quatre formes de questions *whether* : exemple unique, exemples multiples, réponse vraie unique, énumération exhaustive.

La forme générale d'une question *whether* demandant un exemple unique est :

$$I = ? \left(\begin{matrix} | \\ | \end{matrix} - - \right) (A_1, \dots, A_n)$$

que BELNAP illustre par l'interrogative :

- "*De la cuisine, de l'office, et de la cave, quelle pièce vous paraît aussi valable qu'une autre pour commencer à chercher l'épingle à chapeau manquante ?*".

Si l'on symbolise le sujet abstrait de cette interrogative par : $\sigma = \{Cu, O, Ca\}$, il y a trois réponses directes possibles : Cu, O, Ca. On dira qu'une proposition B est une réponse directe à cette forme de question si et seulement si B est une des modalités A_i présentées par σ [Belnap 1963 : 95]. Soit : $d(I) = S$, avec : $S = \{A_i\}$ et : $A_i \in \sigma$

Une interrogative *whether* demandant des exemples multiples est de la forme :

$$I = ? \left(\begin{matrix} \bar{ } \\ | \end{matrix} - - \right) (A_1, \dots, A_n).$$

BELNAP propose comme représentant de cette catégorie l'interrogative :

- "Parmi les phrases suivantes, citez en au moins une qui soit vraie :
le maître d'hôtel cache quelque chose ; la femme de chambre en
sait plus qu'elle ne dit ; cela vaudrait la peine de réinterro-
ger le jardinier".

Soit $\sigma = \{M, F, J\}$. Il y a cette fois-ci sept réponses directes possibles :

M , F , J , M&F , M&J , F&J , et M&F&J . On dira qu'une conjonction de propo-
sitions $B = A_j \& \dots \& A_k$ est une réponse directe à cette forme de question si et
seulement si toute proposition A_i figurant dans l'expression B est un élé-
ment de σ . Soit : $d(I) = S$, avec : $S = \{A_j, \dots, A_k\}$ et : $\{A_j, \dots, A_k\} \in \mathcal{P}(\sigma)$.

La forme d'une question *whether* à réponse vraie unique est :

$$I = ? \left(\bigvee_{i=1}^n \neg (A_i, \dots, A_n) \right).$$

Contrairement aux catégories précédentes (dont les illustrations sont plus pro-
ches de la problématique d'Agatha CHRISTIE que de celle des auteurs de tests ou
de questionnaires, voire de la vie quotidienne), cette forme de question se ren-
contre très fréquemment. C'est pourquoi les interrogatives citées désormais
seront, dans la mesure du possible, extraites de notre corpus. Par exemple :

- "Depuis le début de votre carrière, pensez-vous avoir plutôt évo-
lué en tant que : homme d'étude, homme de prévision, homme
d'organisation ?" (veuillez cocher une case, une seule pour
chaque question) (BNIST : 39)

En symbolisant le sujet abstrait par : $\sigma = \{E, P, O\}$, les trois réponses directes
possibles à cette interrogative sont :

$$d(I)_1 = \{E\} \& \text{Max}(\sigma, S) = E \& \bar{P} \& \bar{O}$$

$$d(I)_2 = \{P\} \& \text{Max}(\sigma, S) = P \& \bar{E} \& \bar{O}$$

$$d(I)_3 = \{O\} \& \text{Max}(\sigma, S) = O \& \bar{E} \& \bar{P}$$

On peut illustrer $d(I)_1$ par l'énoncé : "Depuis le début de ma carrière, je
pense avoir plutôt évolué en tant qu'homme d'étude, et non en tant qu'homme
de prévision ou qu'homme d'organisation". Il ressort de cet exemple que l'af-
firmation d'exhaustivité (maximum) est exprimée par la conjonction des négat-
ions des modalités non sélectionnées (ou par une expression équivalente :
dans l'énoncé ci-dessus, $(\overline{P \vee O})$ est mis pour : $\bar{P} \& \bar{O}$).

Pour généraliser ceci, nous dirons qu'à une interrogative de la forme :

$$I = ? \left(\bigvee_{i=1}^n - \right) (A_1, \dots, A_n),$$

correspond une réponse directe de la forme :

$$d(I) = \bar{A}_1 \& \bar{A}_2 \& \dots \& \bar{A}_{i-1} \& A_i \& \bar{A}_{i+1} \& \dots \& \bar{A}_n$$

La réponse directe sélectionne l'ensemble à un élément $\{A_i\}$, et affirme que sont ainsi sélectionnées toutes les modalités vraies présentées par le σ de I (ou si l'on préfère, d(I) affirme que la modalité sélectionnée A_i est la seule vraie parmi toutes les modalités présentées) [Belnap 1963 : 98-99]. On remarque que si l'on pose : $\{B_1, \dots, B_r\}$ = ensemble de tous les membres de σ n'appartenant pas à la sélection = $\sigma - S$, l'affirmation d'exhaustivité (maximum) s'écrit ici [Belnap 1976 : 53-54] :

$$\text{Max}(\sigma, \{A_i\}) = \bar{B}_1 \& \dots \& \bar{B}_r$$

La forme générale pour une interrogative *whether* demandant une liste exhaustive est :

$$I = ? \left(\bar{\bigvee}_{i=1}^n - \right) (A_1, \dots, A_n)$$

Soit par exemple la question d'enquête :

"[...] Dans quels domaines parmi les suivants souhaitez-vous que l'Etat dépense moins :

- Education Nationale
- Subventions aux entreprises nationalisées déficitaires
- Défense Nationale
- Santé publique
- Aide sociale
- Subventions à l'agriculture
- Aide aux pays sous-développés ?" (IFOP 69-5 : 5b)

dont on symbolisera le sujet abstrait par : $\sigma = \{Ed, En, D, Sp, As, Ag, P\}$. Une telle interrogation a : $2^{|\sigma|-1} = 127$ réponses directes possibles. Par exemple :

$$d(I)_j = Ed \& En \& \bar{D} \& \bar{Sp} \& \bar{As} \& \bar{Ag} \& \bar{P}$$

dont l'équivalent linguistique pourrait être :

$d(I)_j =$ "Je souhaite que l'Etat dépense moins pour l'Education Nationale et pour les subventions aux entreprises nationalisées déficitaires, mais je ne souhaite pas qu'il dépense moins pour la Défense Nationale, ni pour la santé publique, ni pour l'aide sociale, ni pour les subventions à l'agriculture, ni pour l'aide aux pays sous-développés".

Par conséquent, à une interrogative de la forme : $I = ? (\bar{\quad} \vee -)$
 (A_1, \dots, A_n) correspond une réponse directe de la forme :

$$d(I) = \bar{A}_1 \& \dots \& \bar{A}_{i-1} \& A_i \& \bar{A}_{i+1} \& \dots \& \bar{A}_{j-1} \& A_j \& \dots \& \bar{A}_n$$

Comme pour la catégorie précédente, la réponse directe est une conjonction des modalités sélectionnées, conjointe à une conjonction des négations des modalités non sélectionnées.

b) Variété et réductibilité des questions whether

A partir de l'ensemble des développements qui précèdent, il est assez aisé d'identifier dans notre corpus les questions whether, de les formaliser, et de décrire l'ensemble des réponses directes possibles qui leur correspondent. Par exemple :

- "Dans les prochaines années, quelles sont parmi les suivantes les trois mesures auxquelles le Gouvernement doit donner la priorité :
- Limiter les impôts, même s'il faut sacrifier certains équipements comme les hôpitaux, les autoroutes, les écoles
- Aider l'agriculture à se moderniser
- Protéger les petits commerçants
- Augmenter les avantages sociaux des salariés
- Soutenir la valeur du Franc
- Développer les équipements publics : téléphone, autoroutes, hôpitaux, écoles
- Aider les industries à devenir compétitives ?"

(IFOP 69-5 : 4)

pourra être formalisée en :

$$I = ? \left(\begin{matrix} 3 \\ 3 \end{matrix} \vee - \right) (I, A, P, S, F, E, C)$$

et aura : $\frac{|\sigma|!}{u!(|\sigma| - u)!} = 35$ réponses directes possibles. Si l'on accepte d'interpréter l'énoncé de l'interrogative comme une invitation à citer au moins une mesure et au plus trois, la formalisation devient :

$$I = ? \binom{3}{1} \vee -) (I, A, P, S, F, E, C)$$

avec 63 réponses possibles.

Dans de nombreux tests d'aptitudes ou inventaires de personnalité, on rencontre des séries de questions *whether* ne différant que par leur sujet, et ayant par conséquent la même requête. Par exemple, pour une requête $\rho = \binom{1}{1} \vee -) :$

- "Dans les phrases suivantes il manque un mot. Trouvez le mot qui convient le mieux, parmi ceux qui sont en dessous de la phrase"
(ECO 5 : CP)
- "Il faut trouver parmi les 5 dessins placés sur la même ligne, à droite, celui qui manque dans le coin inférieur droit du grand carré"
(EAI 5/2 : Ma)
- "On vous donne deux membres de phrases, précédés des lettres (a) et (b) : on vous demande de choisir, en faisant une croix sur la lettre qui le précède, celui qui correspond à ce que vous êtes, à ce que vous aimez, à ce que vous avez l'habitude de faire, de penser ou de sentir"
(E-I/6)

Exemples de requête $\rho = \binom{2}{2} \vee -) :$

- "Vous trouverez ci-dessous des pensées accompagnées de quatre ou six phrases. Vous ferez une croix dans les cercles correspondants aux deux phrases dont le sens vous paraît se rapprocher le plus (ou s'éloigner le moins) de la pensée. Rappelez-vous pour chaque pensée que vous devez indiquer deux phrases et deux phrases seulement"
(BV 50)
- "Dans chaque groupe de mots, soulignez les deux d'entre eux qui indiquent ce que les objets possèdent toujours. Soulignez deux mots et rien que deux dans chaque ligne" (NB : 4)

Certaines interrogations *whether* élémentaires posent des problèmes particuliers de formalisation. C'est le cas par exemple des questions qui posent une alternative simple, comme les questions à réponse par *oui* ou *non*. Ainsi, l'interrogation :

- "Estimez-vous qu'il faut faire adopter à ses enfants ses propres convictions en matière de religion ?" (SOFRES 686 : 15)

peut être formalisée également en :

$$\begin{aligned} I_1 &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A, \bar{A}) \\ I_2 &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} \vee \\ - \end{array} \right) (A, \bar{A}) \\ I_3 &= ? \left(\begin{array}{c} - \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A, \bar{A}) \\ I_4 &= ? \left(\begin{array}{c} - \\ 1 \end{array} \begin{array}{c} \vee \\ - \end{array} \right) (A, \bar{A}) \end{aligned}$$

La forme I_1 (exemple unique) et la forme I_2 (réponse vraie unique) apparaissent comme pratiquement équivalentes. En effet, les deux réponses directes possibles à la première sont :

$$\left\{ \begin{array}{l} d(I_1)_1 = A \\ d(I_1)_2 = \bar{A} \end{array} \right.$$

Pour la seconde, nous avons :

$$\left\{ \begin{array}{l} d(I_2)_1 = A\bar{A} = A \\ d(I_2)_2 = \bar{A}\bar{A} = \bar{A} \end{array} \right.$$

(cf. [Belnap 1963 : 100]). Pour des raisons de simplicité toutefois, il est permis de préférer la première forme. De même, les formes I_3 (exemples multiples) et I_4 (énumération exhaustive) sont pratiquement équivalentes, comme nous allons le voir ; mais elles se distinguent des formes précédentes en ce qu'elles acceptent la réponse (en principe contradictoire) "oui et non". Les trois réponses directes possibles à I_3 sont :

$$\left\{ \begin{array}{l} d(I_3)_1 = A \\ d(I_3)_2 = \bar{A} \\ d(I_3)_3 = A\bar{A} \end{array} \right.$$

Les trois réponses directes possibles à I_4 sont :

$$\left\{ \begin{array}{l} d(I_4)_1 = A\bar{A} = A \\ d(I_4)_2 = \bar{A}\bar{A} = \bar{A} \\ d(I_4)_3 = A\bar{A} \end{array} \right.$$

Le choix de la forme de la question (abstraite) correspondant à l'interrogative dépend par conséquent ici de la signification attachée à la conjonction des deux membres de l'alternative (cf. [Belnap 1976 : 59]).

Parmi les interrogations élémentaires dont la formalisation fait problème, nous citerons :

- "Utilisez-vous du rouge à lèvres, du maquillage pour les yeux, du fond de teint, du vernis à ongles, du démaquillant, de la laque pour cheveux ?" (SOFRES 686 : 8)

Il est possible d'écrire la question correspondante sous la forme :

$$I = ? \left(\bar{1} \vee - \right) (R, M, F, V, D, L)$$

dans laquelle : R = "J'utilise du rouge à lèvres", M = "J'utilise du maquillage pour les yeux", etc. La difficulté provient ici de ce que, parmi les réponses possibles prévues sur le questionnaire, figure l'énoncé : "Je n'utilise aucun de ces produits". Il faudrait donc réécrire la requête, en posant : $\rho = \left(\bar{0} \vee - \right)$. Or BELNAP interdit de poser $\vee = 0$, "pour nous épargner les difficultés techniques qui surgiraient dans le cas de sélections vides" [Belnap 1976 : 43]. Nous sommes donc contraint ici, sous peine de renoncer aux facilités offertes par la formalisation de l'auteur, de considérer : "Je n'utilise aucun de ces produits" non comme une réponse directe, mais bien comme une réplique, ou mieux comme une *réponse corrective* (cf. § 2.2.5.c).

Avant d'en terminer avec les questions *whether* élémentaires, signalons qu'il est possible de réduire toutes les formes de questions *whether* que nous avons décrites à une même forme, celle de l'exemple unique. Il nous faut pour cela lever au préalable l'interdiction (provisoire) présentée au début de ce paragraphe, d'avoir des modalités de réponse qui soient des conjonctions d'autres modalités. Il est alors aisé de montrer que toute autre forme de question *whether* est équivalente à une question demandant un exemple unique, dont le sujet est l'ensemble des réponses directes possibles à la question originale [Belnap 1963 : 104]. On dira que deux interrogatives I_1 et I_2 sont *érotétiquement équivalentes* si à chaque réponse directe $d(I_1)$ correspond une réponse directe $d(I_2)$ qui lui est équivalente, et si, réciproquement, à chaque $d(I_2)$

correspond une $d(I_1)$ qui lui est équivalente ([Belnap 1976 : 137] ; cf. § 2.2.5 e).

Soit par exemple l'interrogative demandant plusieurs exemples :

$$I = ? \left(\begin{array}{c} \bar{} \\ 1 \end{array} - - \right) (A, B, C)$$

Lui est érotétiquement équivalente l'interrogation demandant un exemple unique :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (A, B, C, A \& B, A \& C, B \& C, A \& B \& C)$$

De même, à l'interrogation demandant une réponse unique :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (A, B, C)$$

correspond :

$$I = ? \left(\begin{array}{c} \bar{} \\ 1 \end{array} - - \right) (A \& \bar{B} \& \bar{C}, \bar{A} \& B \& \bar{C}, \bar{A} \& \bar{B} \& C)$$

Enfin, à l'interrogation demandant une énumération exhaustive :

$$I = ? \left(\begin{array}{c} \bar{} \\ 1 \end{array} \vee - \right) (A, B, C)$$

correspond :

$$I = ? \left(\begin{array}{c} \bar{} \\ 1 \end{array} - - \right) (A \& \bar{B} \& \bar{C}, \bar{A} \& B \& \bar{C}, \bar{A} \& \bar{B} \& C, A \& B \& \bar{C}, \\ A \& \bar{B} \& C, \bar{A} \& B \& C, A \& B \& C)$$

A partir de ces exemples, on étendrait sans peine la procédure de réduction à d'autres interrogatives *whether* présentant un sujet qui comporterait un plus grand nombre de modalités de réponse, ou une requête de nature différente. Toutefois, lorsque u est élevé, ou n'est pas défini (exemples multiples ou énumération exhaustive), le nombre de réponses directes possibles croît rapidement avec la taille de σ , et la forme réduite n'est guère maniable [Belnap 1963 : 106] ; ceci est peut être la raison qui a poussé l'auteur à ne pas aborder de nouveau le problème de la réduction des interrogatives à une forme unique dans ses publications postérieures à 1963.

De notre point de vue cependant, ces possibilités de réduction des interrogatives *whether* à une forme unique nous paraissent avoir un grand intérêt théorique, dans une perspective unificatrice ; d'autant plus que, comme

nous le verrons, certaines interrogatives *which* peuvent également être réduites à cette même forme unique. En outre, on remarquera la quasi-identité de forme entre l'interrogative *whether* demandant un exemple unique, et la forme canonique des énoncés à valeur de question définie au chapitre 1.

c) Les six formes distinguées de questions *which*

Ce type de question pose des problèmes de formalisation beaucoup plus complexes que les questions *whether*. C'est pourquoi, dans ce paragraphe, nous ne traiterons en détail que du cas le plus simple des questions *which*, celui des questions à une seule variable. Ce cas est d'ailleurs le seul qui se rencontre effectivement dans les tests et questionnaires en langue française, puisque l'interrogation binucléaire (et à plus forte raison l'interrogation *n*-nucléaire) y est inusitée. Nous nous bornerons à signaler ensuite le cas général, en ne l'illustrant que par un seul exemple, d'ailleurs encore relativement simple.

Comme exemple simple de question *which*, on peut choisir : "Citez un entier positif qui soit un nombre premier compris entre 10 et 20" [Belnap 1972 : 332]. La requête est ici du type exemple unique : $\rho = (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -)$. Quant au sujet lexical, on peut considérer qu'il se compose de deux éléments distincts :

- une *condition de catégorie* C_x , définissant la classe de substitution dans laquelle la variable x prend ses valeurs (ici : $C_x = "x$ est un entier positif") ;

- une *forme propositionnelle* P_x , par référence à laquelle les modalités de réponse sont présentées (ici : $P_x = "x$ est un nombre premier compris entre 10 et 20").

La forme générale d'une interrogative *which* à une seule variable peut donc s'écrire :

$$I = ? \rho (C_x // P_x)$$

On voit sur cet exemple que les modalités de réponse nominales P_a sont obtenues en substituant à x , dans P_x , les noms a éléments de la classe de substitution définie par C_x ; ici : "1 est un nombre premier compris entre 10 et 20",

"2 est un nombre premier compris entre 10 et 20", etc. Contrairement aux questions *whether*, les questions *which* peuvent présenter des modalités de réponse en nombre très élevé, voire (comme ici) infini.

La formalisation du sujet lexical d'une interrogative *which* donnée n'a pas nécessairement une solution unique. Ainsi, dans l'exemple cité, on peut également poser :

$Cx = "x \text{ est un nombre premier}"$

$Px = "x \text{ est compris entre 10 et 20}"$.

Le seul problème est que, par définition, les conditions de catégorie sont en nombre fini, et font partie intégrante du langage formel L . Créer une condition catégorielle entièrement nouvelle (c'est-à-dire qui ne soit ni une conjonction ou une disjonction de conditions catégorielles, ni le résultat d'un changement de variable d'une condition catégorielle), c'est en fait modifier le langage (cf. [Belnap 1976 : 32]). D'ailleurs, pour qu'il soit toujours possible de dire si un énoncé est ou non une réponse directe à une question *which*, il est nécessaire de pouvoir dire si un nom donné appartient ou non à la classe de substitution définie par Cx [Belnap 1963 : 61, 91]. Or, s'il est facile de dire si un nombre donné est ou n'est pas un entier positif, il l'est moins de dire s'il est ou non premier. Toujours dans l'exemple cité, on pourrait également affranchir la variable de sa condition de catégorie, en reportant le prédicat de la condition catégorielle ci-dessus dans la forme propositionnelle :

$\sigma = (x//x \text{ est un nombre premier} \ \& \ x \text{ est compris entre 10 et 20})$.

La variable x peut alors prendre ses valeurs dans l'ensemble des noms associés aux individus quels qu'ils soient. Le locuteur peut éventuellement recourir à cette forme "s'il ne sait pas que la réponse doit être cherchée parmi les nombres entiers, ou s'il ne veut pas fournir cette information au répondant" [Belnap 1963 : 120].

En outre, le choix de la condition catégorielle peut déterminer le nombre, ou même l'existence des réponses vraies à la question. Soit par exemple l'interrogative :

$I = "Quelles \text{ sont les racines de l'équation } x^2 = 4 ?"$

Si l'on pose :

$$\sigma = (x \text{ est un nombre rationnel positif} // x^2 = 4)$$

il n'y a qu'une réponse vraie : $x = 2$.

Si l'on pose :

$$\sigma = (x \text{ est un nombre rationnel} // x^2 = 4)$$

il y a deux réponses vraies : $x = 2$ et $x = -2$.

D'autre part, si l'on a :

$$I = \text{"Quelles sont les racines de l'équation } x^2 = -4 \text{ ?"}$$

avec :

$$\sigma = (x \text{ est un nombre rationnel} // x^2 = -4)$$

il n'y a aucune réponse vraie ; ce qui n'est évidemment plus le cas avec :

$$\sigma = (x \text{ est un nombre imaginaire} // x^2 = -4)$$

[Belnap 1963 : 90-92]. On dira que "poser une question dont le sujet ne détermine aucune modalité vraie en raison d'un conflit entre la condition catégorielle et la forme propositionnelle" constitue une *erreur catégorielle* [Belnap 1976 : 33].

La distinction que nous avons faite entre catégorie nominale (classe des noms) et catégorie réelle (classe des référents) nous conduit à distinguer deux types de modalités de réponse aux questions *which* :

- les modalités *nominales*, obtenues par substitution d'un nom α_i à la variable libre x de la forme propositionnelle Px ,

- les modalités *réelles*, qui sont définies par rapport à une interprétation. On dira que, pour une interprétation donnée M , les modalités réelles présentées dans M par le sujet lexical $\sigma = (Cx // Px)$ sont tous les couples $\langle f, Px \rangle$, où f est une fonction qui associe à x une entité du sous-ensemble du domaine des individus de M dans lequel la propriété Cx est vérifiée. Pour une interprétation donnée M , relative à une inconnue x et une forme propositionnelle Px , on dira que la modalité nominale Pa signifie dans M

la modalité réelle $\langle f, Px \rangle$, si le nom α sert à désigner dans M l'individu $f(x)$. Par exemple, soit une interprétation M relative à une inconnue x et une équation : $32 - x = 5$, les noms : "XXVII", " 27_{10} ", " 33_8 ", " 11011_2 ", etc. servent à désigner dans M la même entité : "vingt-sept".

Nous avons déjà vu une interrogative *which* demandant un exemple unique :

$I =$ "Citez un entier positif qui soit un nombre premier compris entre 10 et 20"

$= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (x \text{ est un entier positif} // x \text{ est un nombre premier compris entre 10 et 20})$

$d(I) =$ "9 est un nombre premier compris entre 10 et 20".

Sur cette réponse directe (fausse), on voit que le prédicat catégoriel C peut ne pas apparaître dans l'énoncé de la réponse à une question *which* [Belnap 1972 : 332]. Une proposition A est une réponse directe à une interrogative de la forme : $I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (Cx // Px)$ si et seulement si A est de la forme Pa , où le nom α est un élément de la classe de substitution associée à Cx .

Une interrogative *which* demandant des exemples multiples s'écrit :

$I = ? \left(\begin{array}{c} \bar{1} \\ 1 \end{array} - - \right) (Cx // Px)$.

Comme illustration de cette forme d'interrogative, on peut avoir :

$I =$ "Citez des nombres premiers compris entre 10 et 20"

[Belnap 1976 : 38], avec comme réponse possible la forme codée :

"13 et 17"

représentant la réponse directe :

$d(I) =$ "13 est un nombre premier compris entre 10 et 20 et 17 est un nombre premier compris entre 10 et 20".

Une proposition A est une réponse directe à une interrogative

$I = ? \left(\begin{array}{c} \bar{1} \\ 1 \end{array} - - \right) (Cx // Px)$, si et seulement si elle est de la forme

$A = Pa_1 \& \dots \& Pa_i \& \dots \& Pa_k$, où les noms a_i sont des éléments de la classe de substitution associée à Cx .

Une interrogative *which* demandant des exemples multiples et distincts a pour représentation symbolique :

$$I = ? \left(\begin{array}{c} \bar{ } \\ 1 \end{array} - \neq \right) (Cx//Px).$$

En posant :

$I =$ "Citez plusieurs villes françaises de plus de 1 million d'habitants"

soit : $\left\{ \begin{array}{l} Cx = "x \text{ est une ville française}" \\ Px = "x \text{ a plus de 1 million d'habitants}" \end{array} \right.$

On peut obtenir comme réponse codée correspondante :

"Marseille, la porte de l'Orient, et la moderne Massilia".

L'exigence de distinction qui est exprimée par la question doit se traduire dans la réponse directe par l'affirmation qu'il n'existe aucun couple des noms cités désignant un même référent ; ou si l'on préfère, on a ici :

$\text{Dist}(\sigma, S) = (\text{Marseille} \neq \text{la porte de l'Orient}) \& (\text{Marseille} \neq \text{la moderne Massilia}) \& (\text{la porte de l'Orient} \neq \text{la moderne Massilia}).$

Pour que l'exigence de distinction ne soit pas satisfaite, il suffit qu'une seule de ces propositions élémentaires soit fausse. Une réponse directe à une interrogative $I = ? \left(\begin{array}{c} \bar{ } \\ 1 \end{array} - \neq \right) (Cx//Px)$ est une expression A de la forme $A = (Pa_1 \& \dots \& Pa_i \& \dots \& Pa_k) \& (a_1 \neq a_2) \& (a_1 \neq a_3) \& \dots \& (a_1 \neq a_k) \& (a_2 \neq a_3) \& \dots \& (a_{k-1} \neq a_k)$. Notons que l'affirmation de distinction peut s'écrire de manière abrégée en :

$$\text{Dist}(\sigma, S) = \forall i, \forall j > i (a_i \neq a_j)$$

Une interrogative *which* demandant une réponse vraie unique a pour forme générale :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (Cx//Px)$$

Soit par exemple :

$I =$ "Qui a écrit le *Cid* ?"

Le pronom interrogatif "qui" définit la condition catégorielle comme le fait d'être un être animé, ou même un être humain. On a donc :

$$I : ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (x \text{ est un être humain} // x \text{ a écrit le } Cid)$$

Une réponse directe à cette question doit non seulement sélectionner un nom d'individu a supposé avoir écrit le *Cid*, mais encore affirmer qu'aucun autre individu ayant la propriété d'être humain n'a également écrit le *Cid*.

Pour affirmer l'exhaustivité maximum, BELNAP pose :

$$\text{Max} (\sigma, S) = \forall x (Cx \rightarrow (Px \rightarrow x = a)),$$

où a est le nom de l'individu cité dans la sélection (cf. [Belnap 1976 : 55]). Il nous paraît plus clair d'appliquer à cette expression l'équivalence formelle : $p \rightarrow q \leftrightarrow (\overline{p} \ \& \ \overline{q})$, et d'écrire :

$$\text{Max} (\sigma, S) = \forall x ((Cx \ \& \ Px) \rightarrow (x = a)).$$

Ce qui peut être paraphrasé en disant que, si l'on a à la fois Cx et Px , alors on ne peut avoir que $x = a$. Dans notre exemple, si x est un être humain et si x a écrit le *Cid*, alors x ne peut être que l'individu nommé "Pierre Corneille". On dira donc qu'une expression A est une réponse directe à une interrogative $I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (Cx // Px)$ si et seulement si $A = Pa \ \& \ \forall x (Cx \rightarrow (Px \rightarrow x = a))$, avec $a =$ nom appartenant à la classe de substitution associée à Cx .

Les interrogatives *which* qui demandent une énumération exhaustive sont du type :

$$I = ? \left(\begin{array}{c} - \\ 1 \end{array} \vee - \right) (Cx // Px).$$

On en trouve de très nombreux exemples dans les tests et les questionnaires, comme :

- "Maintenant dites-moi tous les mots que vous pouvez qui riment avec : jour" (TMI : 13)
- "Quelles sont toutes les marques de voitures que vous avez déjà possédées, y compris, bien sûr, votre (ou vos) voiture(s) actuelle(s) ?" (SOFRES 967 : 15)

- "Qu'est-ce que le général de Gaulle a fait de mieux depuis qu'il est au pouvoir ?" (IFOP 58-11 : 37)
- "A votre avis, quelles sont les rues ou places de M. les plus agréables pour faire du lèche-vitrine ?" (SOFRES 659 : 1)
- "D'après ce que vous avez pu remarquer, quelles sont les marques qui ont les stations les mieux équipées pour la voiture ?" (SOFRES 340 : 25 c)

L'affirmation d'exhaustivité des réponses directes doit, ici aussi, exprimer qu'en dehors des noms cités dans l'énoncé de la réponse, aucun autre nom du domaine de Cx ne satisfait à la condition de la forme propositionnelle Px :

$$\text{Max}(\sigma, S) = \forall x (Cx \rightarrow (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k))).$$

Par conséquent, une expression A est une réponse directe à une interrogative $I = ? (\bar{1} \vee -) (Cx//Px)$ si et seulement si $A = (Pa_1 \& \dots \& Pa_i \& \dots \& Pa_k) \& \forall x (Cx \rightarrow (Px \rightarrow (x = a_1 \vee \dots \vee x = a_i \vee \dots \vee x = a_k)))$, avec a_i = noms appartenant à la classe de substitution associée à Cx .

Enfin, les interrogatives *which* qui demandent une énumération exhaustive et distincte ont pour formule :

$$I = ? (\bar{1} \vee \neq) (Cx//Px)$$

BELNAP en donne comme exemple l'interrogative :

$I =$ "Qui étaient les dénonciateurs de Catilina ?"

avec la réponse codée (fausse) :

"Cicéron et Marcus Tullius".

La réponse directe complète doit comporter les trois éléments conjoints :

- $S =$ "Cicéron était un dénonciateur de Catilina & Marcus Tullius était un dénonciateur de Catilina"
- $\text{Max}(\sigma, S) =$ " x est une personne" \rightarrow (" x est un dénonciateur de Catilina" \rightarrow ($x =$ "Cicéron" \vee $x =$ "Marcus Tullius"))
- $\text{Dist}(\sigma, S) =$ "Cicéron" \neq "Marcus Tullius"

C'est évidemment la fausseté de l'affirmation de distinction qui entraîne ici la fausseté de la réponse directe. En règle générale, on considèrera qu'une expression A est une réponse directe à une interrogative $I = ? (\bar{1} \vee \neq)$ ($Cx//Px$) si et seulement si $A = (Pa_1 \& \dots \& Pa_i \& \dots \& Pa_k) \& \forall x (Cx \rightarrow (Px \rightarrow (x = a_1 \vee \dots \vee x = a_i \vee \dots \vee x = a_k))) \& (\forall i, \forall j > i (a_i \neq a_j))$, en posant que les a_i sont des noms appartenant à la classe de substitution associée à Cx .

Il n'est pas toujours aisé de formaliser une interrogative *which* donnée. Nous avons déjà remarqué avec l'auteur que le sujet lexical pouvait dans certains cas être défini de plusieurs manières. Nous allons voir que le choix de la forme de la requête pose également des problèmes si l'on veut retrouver à travers le langage les intentions réelles du locuteur. Soit par exemple :

$I =$ "Quels sont tous les écrivains contemporains que vous connaissez ?"

qui est une question de notoriété spontanée (les questions de notoriété suggérée étant des questions *whether*). Le sujet lexical peut être :

$\sigma = (x \text{ est un écrivain contemporain} // \text{je connais } x)$

Dans ce cas, on peut difficilement accepter de poser :

$\rho = (\bar{1} \vee -)$

car l'on aurait :

$d(I) = (\text{"je connais } a_1 \text{"} \& \dots \& \text{"je connais } a_k \text{"}) \& \forall x (\text{"}x \text{ est un écrivain contemporain" } \rightarrow (\text{"je connais } x \text{" } \rightarrow (x = a_1 \vee \dots \vee x = a_k)))$

ce qui est vraisemblablement faux, puisqu'en notoriété suggérée le locuteur s'apercevrait qu'il connaît bien d'autres écrivains contemporains. En se plaçant au point de vue de l'analyse des réponses directes, l'interrogative ci-dessus n'est susceptible que des deux formalisations suivantes (compte non tenu de l'éventuelle exigence de distinction) :

$$I_1 = ? \left(\begin{array}{c} \bar{ } \\ 1 \end{array} - - \right) (x \text{ est un écrivain contemporain//je connais } x)$$

$$I_2 = ? \left(\begin{array}{c} \bar{ } \\ 1 \end{array} \forall - \right) (x \text{ est un écrivain contemporain//je me souviens de } x)$$

Soit ou bien une interrogative demandant des exemples multiples (malgré l'apparente exigence d'exhaustivité exprimée par le : "tous") et conservant dans son sujet lexical les termes de l'énoncé, ou bien une interrogative demandant une liste exhaustive, mais dont la forme propositionnelle ne reprend pas les termes de l'énoncé.

De même, il paraît naturel de formaliser la requête correspondant à la question :

- "Dis-moi les noms des mois de l'année"

(TERMAN : 9 ans, supplément 1)

par : $\rho = \left(\begin{array}{c} \bar{ } \\ 1 \end{array} \forall - \right)$, ou bien : $\rho = \left(\begin{array}{c} 12 \\ 12 \end{array} \forall - \right)$. Or, si l'on se réfère au système de notation correspondant, on constate que cette partie de l'item est réussie si les mois sont nommés en moins de 20 secondes, avec une erreur ou omission au maximum. Le langage formel conçu par BELNAP ne permet pas de rendre compte des contraintes de temps, ni même (en principe du moins) de la possibilité d'accepter une erreur. Par contre, on peut symboliser l'acceptation d'une éventuelle omission par : $\rho = \left(\begin{array}{c} 12 \\ 11 \end{array} - - \right)$.

d) Caractères généraux des questions which

Avant de passer à d'autres types de questions, il nous reste à présenter pour mémoire la forme générale des questions *which* et de leurs réponses directes, et à examiner s'il est possible de réduire certaines questions *which* à une forme canonique.

La forme générale des questions *which* correspond à une interrogative portant simultanément sur plusieurs inconnues (question dite *relationnelle* [Belnap 1976 : 24]). Les exemples d'interrogatives à plusieurs variables sont courants en algèbre (systèmes d'équations à n inconnues) ; on peut fabriquer en français des questions assez artificielles, dans le style : "Qui fait quoi à qui ?", "Qui va où ?", etc. Mais ce type d'énoncé interrogatif ne semble pas

usité dans les tests et les questionnaires de notre corpus. Dans sa forme générale, le sujet d'une question *which* à n variables comporte une forme propositionnelle n -aire, que l'on écrira : $Px_1 \dots x_n$. Chaque variable x_i peut être soit liée à une condition catégorielle $C_i x_i$, soit libre. En supposant que l'on ait r variables liées à une condition catégorielle, et $n-r$ variables libres, une question *which* s'écrit par conséquent :

$$I = ? \rho (C_1 x_1, \dots, C_{r-1} x_{r-1}, C_r x_r, x_{r+1}, \dots, x_n // Px_1 \dots x_n)$$

Toute réponse directe à une question *which* comporte obligatoirement une sélection des modalités de réponse présentées par le sujet ; cette sélection s'écrit :

$$S = (Pa_{i_1} \dots a_{i_n} \ \& \ \dots \ \& \ Pa_{p_1} \dots a_{p_n})$$

avec a_{i_k} = nom appartenant à la classe de substitution associée à $C_k x_k$.

La réponse directe peut comporter en outre, selon la forme de la requête, une affirmation d'exhaustivité (maximum) et une affirmation de distinction. L'affirmation d'exhaustivité (maximum) est de la forme :

$$\begin{aligned} \text{Max}(\sigma, S) &= \forall x_1, \dots, \forall x_n (C_1 x_1 \ \& \ \dots \ \& \ C_r x_r \\ &\rightarrow (Px_1 \dots x_n \rightarrow ((x_1 = a_{i_1} \ \& \ \dots \ \& \ x_n = a_{i_n}) \\ &\vee \dots \vee (x_1 = a_{p_1} \ \& \ \dots \ \& \ x_n = a_{p_n})))) \end{aligned}$$

L'affirmation de distinction exprime l'idée que deux modalités de réponse figurant dans la sélection doivent différer par au moins une des valeurs attribuées aux variables sur lesquelles porte l'interrogation. Ainsi, pour deux modalités $Pa_{i_1} \dots a_{i_n}$ et $Pa_{j_1} \dots a_{j_n}$, affirmer qu'elles sont distinctes, c'est affirmer que :

$$(a_{i_1} \neq a_{j_1}) \vee \dots \vee (a_{i_k} \neq a_{j_k}) \vee \dots \vee (a_{i_n} \neq a_{j_n})$$

Pour l'ensemble des modalités sélectionnées, l'affirmation de distinction s'écrira, selon les mêmes conventions de notation :

$$\begin{aligned}
\text{Dist}(\sigma, S) = & ((a_{1_1} \neq a_{2_1}) \vee \dots \vee (a_{1_k} \neq a_{2_k}) \vee \dots \vee \\
& (a_{1_n} \neq a_{2_n})) \ \& \ ((a_{1_1} \neq a_{3_1}) \vee \dots \vee (a_{1_k} \neq a_{3_k}) \vee \dots \vee \\
& (a_{1_n} \neq a_{3_n})) \ \& \ \dots \ \& \ ((a_{1_1} \neq a_{p_1}) \vee \dots \vee (a_{1_n} \neq a_{p_n})) \ \& \\
& ((a_{2_1} \neq a_{3_1}) \vee \dots \vee (a_{2_n} \neq a_{3_n})) \ \& \ \dots \ \& \ ((a_{2_1} \neq a_{p_1}) \\
& \vee \dots \vee (a_{2_n} \neq a_{p_n})) \ \& \ \dots \ \& \ ((a_{p-1_1} \neq a_{p_1}) \vee \dots \vee \\
& (a_{p-1_n} \neq a_{p_n}))
\end{aligned}$$

Notation que l'on pourra abréger [Belnap 1976 : 65] en :

$$\text{Dist}(\sigma, S) = \bigwedge_{(1 \leq i < j \leq p)} \bigvee_{(1 \leq k \leq n)} (a_{i_k} \neq a_{j_k})$$

Pour illustrer sur un exemple simple cette présentation des questions *which*, soit l'interrogation bi-nucléaire [Belnap 1976 : 76-77] :

I = "Quels garçons sont les frères de quelles filles ?"

que l'on peut interpréter comme une question demandant une énumération exhaustive et distincte :

$$I = ? \left(\bigvee_1 \neq \right) (Gx, Fy // Pxy)$$

dans laquelle :

$$\left\{ \begin{array}{l} Gx = "x \text{ est un garçon}" \\ Fy = "y \text{ est une fille}" \\ Pxy = "x \text{ est frère de } y" \end{array} \right.$$

Soit une interprétation M, telle que, à Gx soit associé la classe d'individus : {Emile, François, Georges, Henri} en abrégé : {e, f, g, h} ; et à Fy, la classe d'individus : {Annie, Béatrice, Colette, Denise}, en abrégé : {a, b, c, d}. Une réponse codée à cette interrogative pourrait être :

"Emile et Béatrice, François et Denise, Georges et Annie, Georges et Colette, et Henri et Béatrice".

Cet énoncé correspond à la réponse directe :

$$\begin{aligned}
 d(I) = & (Peb \ \& \ Pfd \ \& \ Pga \ \& \ Pgc \ \& \ Phb) \ \& \\
 & \forall x, \forall y (Gx \ \& \ Fy \rightarrow (Pxy \rightarrow (x = e \ \& \ y = b) \vee (x = f \ \& \ y = d) \\
 & \vee (x = g \ \& \ y = a) \vee (x = g \ \& \ y = c) \vee (x = h \ \& \ y = b) \ \& \\
 & ((e \neq f \vee b \neq d) \ \& \ (e \neq g \vee b \neq a) \ \& \ (e \neq g \vee b \neq c \ \& \\
 & (e \neq h \vee b \neq b) \ \& \ (f \neq g \vee d \neq a) \ \& \ (f \neq g \vee d \neq c) \ \& \\
 & (f \neq h \vee d \neq b) \ \& \ (g \neq g \vee a \neq c) \ \& \ (g \neq h \vee a \neq b) \ \& \\
 & (g \neq h \vee c \neq b))
 \end{aligned}$$

Au paragraphe 2.2.2.b, nous avons vu que l'on pouvait réduire les principales formes de questions *whether* à une forme unique, celle de l'interrogative demandant un exemple unique. Dans une perspective unificatrice, il serait souhaitable de parvenir à réduire également les questions *which* à une question *whether* demandant un exemple unique. Le principal obstacle à cette réduction tient à l'impossibilité pratique d'énumérer un ensemble de modalités de réponse infini. Toutefois, même dans le cas fini, la réduction des questions *which* à une forme équivalente de question *whether* n'est réalisable qu'à la condition de formaliser, dans le cadre des questions *whether*, l'énumération d'un ensemble fini de modalités définies en compréhension. Ceci nécessite l'introduction des notions d'interrogative quantifiée et d'interrogative relativisée, que nous aborderons aux paragraphes 2.2.4.b, c, et d.

2.2.3. Autres formes de questions simples

La division des interrogatives simples en questions *whether* et en questions *which*, selon la manière dont le sujet présente les modalités de réponse possibles, peut paraître recouvrir l'ensemble des formes simples de questions observables. Telle est un peu l'impression que donnent en particulier les premiers travaux de BELNAP ("il est en général facile de dire dans laquelle des deux catégories ranger une question de la langue courante" [Belnap 1963 : 45]), même s'il affirme par ailleurs qu' "il n'y a pas de raison de supposer cette classification exhaustive" [Belnap 1963 : 37 n.17].

Dans *The Logic of Questions and Answers*, l'auteur propose une autre présentation des questions simples, qui élargit sensiblement le champ des formes de questions possibles [Belnap 1976 : 78-79]. Le calcul des prédicats du premier ordre peut être considéré comme présentant six éléments fondamentaux : les propositions, les formes propositionnelles, les noms, les termes avec variables libres, les connecteurs fonctionnels, et les quantificateurs. BELNAP propose de classer les questions simples selon les éléments qu'elles *posent* et ceux qu'elles *demandent*. En nous limitant aux questions qui posent un seul élément et en demandant un seul également, nous pouvons en principe définir 36 catégories. L'auteur en décrit quelques unes, en commençant par situer les questions *which* et les questions *whether* dans ce schéma. Ainsi, les questions *which* posent une *forme propositionnelle*, et demandent des *noms* à substituer aux variables pour engendrer les modalités de réponse possibles. Les questions *whether* posent une séquence de *propositions*, et demandent une fonction de vérité (*connecteur*) ayant ces propositions comme arguments ; cette présentation (due à [Stahl 1969 : 213 |]) impliquant que l'on considère une réponse directe à une question *whether* comme une conjonction des modalités de réponse ou de leurs négations. BELNAP décrit très brièvement quatre autres formes de questions ayant leur place dans ce schéma : questions de description, d'identité, de quantité, et questions *what*. S'y ajoutent les questions de causalité, qui se situent un peu en marge de cette classification, mais qui méritent plus qu'une allusion.

a) Les questions de description

Les questions de description posent un nom b , et demandent une forme propositionnelle $H_i x$, appelée *descripteur*. Soit par exemple :

I = "De quelle couleur est Dick ?"

Cette interrogative pose le nom $b = \text{"Dick"}$, et demande des descripteurs de la forme : $H_i x = \text{"x est blanc"}$, $H_j x = \text{"x est roux"}$, $H_k x = \text{"x est noir"}$, etc. En appelant Hx la condition : " x est < une couleur >", à laquelle sont associés les descripteurs $H_i x$ (pour tout i), on écrira le sujet de I :

$\sigma = \text{Des} (Hx // b)$

Comme exemple de réponse directe à cette interrogative, en supposant que la

requête exprime la demande d'exemples multiples et distincts, on peut avoir :

$$d(I) = (H_i b \ \& \ H_k b) \ \& \ \overline{\forall x (H_i x \leftrightarrow H_k x)}$$

qui formalise la réponse codée : "Dick est blanc et noir".

Appartiennent à ce type de questions des interrogatives de la forme : "Qui est Victor Hugo ?", ou : "Comment allez-vous ?". Dans les tests et les questionnaires, on en trouve des illustrations du genre :

- "Quel âge as-tu ?" (TERMAN : 5 ans, supplément)
- "D'après vous, quelles sont les principales qualités de F. ?" (SOFRES 967 : 40)

b) Les questions d'identité

Les questions d'identité posent un nom et demandent un nom. Soit par exemple :

I : "Quel était le vrai nom de Barbe-Bleue ?"

dont le sujet peut être formalisé en :

$$\sigma = \text{Ident} (Cx // b)$$

où b est le nom posé ("Barbe-Bleue"), et où Cx est une condition catégorielle ($Cx = "x$ est un personnage ayant réellement existé"), à laquelle est associé un ensemble de noms a_1, \dots, a_i . Les modalités de réponse sont de la forme : $b = a_i$, soit ici : "Le vrai nom de Barbe-Bleue était Gilles de Rais". Selon la nature de la requête, les réponses directes peuvent associer à la sélection les affirmations d'exhaustivité et de distinction.

c) Les questions de quantité

Les questions de quantité posent une forme propositionnelle Px et demandent un quantificateur. On peut écrire leur sujet sous la forme :

$$\sigma = \text{Combien} (Px)$$

avec par exemple $Px = "x \text{ est un académicien}"$, ce qui donne l'interrogation :

$I = "Combien y-a-t-il d'académiciens ?"$.

On trouve de nombreux exemples de ce type de question dans les tests et les questionnaires :

- "Combien as-tu de doigts à une main ?" (TERMAN : 7 ans, 1°)
- "Combien de spectateurs l'ont admiré ?" (TMI : 10c)
- "Vous-même, avant votre mariage, combien de personnes estimez-vous avoir rencontrées avec qui un mariage aurait été possible ?" (INED 081 : opinions 9)
- "En tout, combien de voitures a-t-on achetées depuis 5 ans [...] ?" (SOFRES 769 : 33)
- "Combien d'heures faites-vous par semaine ?" (GEMAS 76 : 106)

Toutefois, la formalisation des questions de quantité reste à faire, et le mécanisme de la réponse n'a pas été analysé.

d) Les questions what

Les questions *what* posent une forme propositionnelle Ax et demandent une forme propositionnelle $H_i x$. Comme pour les questions de description, les $H_i x$ sont des descripteurs ; l'ensemble des $H_i x$ est associé à une condition Hx . L'auteur distingue quatre variétés de questions *what*, selon la nature de la relation unissant les deux formes propositionnelles (équivalence, implication, conjonction) : questions d'équivalence, de condition nécessaire, de condition suffisante, et d'intersection.

Une question d'équivalence peut être illustrée par exemple par l'interrogative :

$I = "Qu'est-ce qu'un nombre premier ?"$
 $= ? \left(\begin{matrix} 1 \\ 1 \end{matrix} - - \right) \text{Equiv} (Hx // Ax)$

avec : $Hx = "x \text{ est un nombre}"$, et : $Ax = "x \text{ est premier}"$. A la condition Hx sont associés les descripteurs : $H_i x = "x \text{ est un multiple de } 2"$, $H_j x = "x \text{ n'est divisible sans reste que par } 1 \text{ et par } x"$, etc. Une réponse directe à I est de la forme :

$$d(I) = \forall x (Ax \leftrightarrow H_j x)$$

Peuvent s'y ajouter selon la requête les affirmations d'exhaustivité (dans le cas fini) et de distinction.

Une *question de condition nécessaire*, du type :

I = "Que sont les mammifères ?"

est formalisable en :

$$I = ? (\begin{matrix} 1 \\ 1 \end{matrix} - -) Nec (Hx // Ax)$$

avec : Hx = "x est < un élément d'une classe d'êtres vivants >", et : Ax = "x est un mammifère". Si $H_1 x =$ "x est un vertébré", on aura :

$$\begin{aligned} d(I) &= \forall x (Ax \rightarrow H_1 x) \\ &= \text{"les mammifères sont des vertébrés"}. \end{aligned}$$

Une *question de condition suffisante* peut être :

I = "Quels animaux font partie des mammifères ?" formalisable (dans l'hypothèse de la demande d'un exemple unique) en :

$$I = ? (\begin{matrix} 1 \\ 1 \end{matrix} - -) Suf (Hx // Ax)$$

avec : Hx = "x est < un animal >" et : Ax = "x est un mammifère". Soit $H_1 x =$ "x est un cheval", on aura :

$$\begin{aligned} d(I) &= \forall x (H_1 x \rightarrow Ax) \\ &= \text{"les chevaux sont des mammifères"}. \end{aligned}$$

Remarquons que la question : "What sort of things are mammals ?" est ambiguë (contrairement à ses traductions françaises), puisqu'elle peut aussi bien demander une condition nécessaire qu'une condition suffisante.

Une *question d'intersection* enfin, comme :

I = "Quelles sortes de choses abstraites peuvent-elles être premières ?"

est formalisable en :

$$I = ? \left(\begin{matrix} 1 \\ 1 \end{matrix} - - \right) \text{Inter} (Hx // Ax)$$

et avoir pour réponse directe :

$$d(I) = \exists x (H_1 x \ \& \ Ax)$$

$$= \text{"les nombres peuvent être premiers"}$$

avec : $H_1 x = "x \text{ est un nombre}"$, et $Ax = "x \text{ est premier}"$.

Les *questions what* se rencontrent dans les tests d'aptitudes ou de connaissances, sans qu'il soit toujours facile de dire de quelle variété il s'agit. Ainsi, les questions de vocabulaire du BINET-SIMON, du Terman ou du TMI (pour nous limiter à notre corpus) sont de la forme :

- "Qu'est-ce que c'est qu'un cheval ?" (BS Z : 13)

Si l'on se réfère aux règles d'appréciation de la réponse, lorsque les sujets sont des enfants, les réponses acceptables se classent en définitions par l'usage et définitions supérieures à l'usage [Binet 1911 : 56-57]. Les définitions par l'usage jugées acceptables sont par exemple : "c'est pour traîner les voitures", "c'est pour courir", "c'est pour s'asseoir dessus". Si l'on considère qu'il s'agit, non d'un usage possible, mais d'un usage effectif (cheval que l'enfant a vu, ou dont il a entendu parler), les réponses ci-dessus sont de la forme :

$$d(I) = \exists x (H_1 x \ \& \ Ax)$$

avec : $Ax = "x \text{ est un cheval}"$

$$H_1 x = "x \text{ traîne les voitures}"$$

La question est interprétée comme une question d'intersection. Les réponses supérieures à l'usage peuvent être classées (grossièrement) en réponses scolaires (ex. : "c'est un quadrupède"), en descriptions des caractères visibles (ex. : "ça a quatre jambes"), en réponses sur la nature ou la composition (ex. : "c'est de la viande"), en définitions de type "aristotélicien" (genre prochain et différence spécifique, cette dernière souvent centrée sur l'usage ; ex. : "c'est une bête qui traîne les voitures"). Les trois premières catégories nous paraissent être de la forme :

$$d(I) = \forall x (Ax \rightarrow H_1 x)$$

La question est par conséquent perçue comme une question de condition nécessaire. Quant à la définition "aristotélicienne", elle nous semble associer les deux types précédents :

$$d(I) = \forall x (Ax \rightarrow H_i x) \quad \& \quad \exists x (H_j x \& Ax)$$

avec :

$$Ax = "x \text{ est un cheval}"$$

$$H_i x = "x \text{ est une bête}"$$

$$H_j x = "x \text{ traîne les voitures}."$$

Cette forme n'entre évidemment pas dans la classification des questions simples proposée par BELNAP. Notons toutefois qu'une définition "aristotélicienne" rigoureuse ne citerait pas, comme différence spécifique, un caractère accidentel ; ne s'agissant pas d'un test pour enfants, une réponse de ce type ne pourrait être que la conjonction :

$$d(I) = \forall x (Ax \rightarrow H_i x) \quad \& \quad \forall x (H_k x \rightarrow Ax).$$

e) Les questions de causalité

Les questions de causalité diffèrent des formes de question que nous avons traitées jusqu'ici en ce qu'elles posent deux éléments à la fois : une forme propositionnelle Ax , et un nom b . Si l'on admet avec l'auteur qu'une explication ne saurait raisonnablement être considérée comme la seule possible, on considérera qu'une question de causalité n'exige ni exhaustivité, ni même distinction. La forme générale des questions de causalité s'écrira :

$$I = ? \left(\begin{array}{l} | \\ 1 \end{array} - - \right) \text{ Pourquoi } (x // Ax, b)$$

Pour illustrer l'analyse de ce type de question, BELNAP prend l'exemple suivant :

$$I = "Pourquoi *Grimbaum* s'écrit-il avec un umlaut ?"$$

La réponse codée : "Parce que ce mot est d'origine allemande", est, selon [Bromberger 1966a:77-78], mise pour la réponse directe :

$$d(I) = Bb \quad \& \quad Cb \quad \& \quad \forall x (Bx \rightarrow (Ax \leftrightarrow (Cx \vee Dx)))$$

avec :

$b = \text{"Grünbaum"}$

$Ax = \text{"x s'écrit avec un umlaut"}$

$Bx = \text{"x est un mot anglais"}$

$Cx = \text{"x est un mot d'origine allemande"}$

$Dx = \text{"x est un mot qui provient d'une autre langue que l'allemand utilisant l'umlaut"}$.

La réponse $d(I)$ peut être paraphrasée par l'énoncé : "*Grünbaum* est un mot anglais et c'est un mot d'origine allemande, et aucun mot anglais ne s'écrit avec un umlaut à moins qu'il ne soit d'origine allemande ou qu'il provienne d'une autre langue que l'allemand utilisant l'umlaut".

Nous reviendrons ultérieurement sur certaines questions de causalité relevées dans les tests ou les questionnaires, auxquelles ce schéma s'applique difficilement. L'exemple suivant par contre illustre bien l'analyse de BROMBERGER-BELNAP :

- "*Pourquoi n'écoutez-vous pas davantage l'ORTF ?*" (ORTF 65 : 9)

En affectant aux éléments de la question et aux réponses précodées les symboles suivants :

$b = \text{"je"}$

$Ax = \text{"x n'écoute pas davantage l'ORTF"}$

$Bx = \text{"x est un auditeur de la radio"}$

$Cx = \text{"x a la télévision"}$

$Dx = \text{"x trouve difficilement l'ORTF sur le cadran"}$

$Ex = \text{"x reçoit mal l'ORTF sur son poste"}$

$Fx = \text{"x ne connaît pas bien les émissions de l'ORTF"}$

$Gx = \text{"x trouve les émissions de l'ORTF moins bonnes qu'avant"}$

$Hx = \text{"x préfère les émissions des autres stations"}$,

comme la question comporte la consigne : "une seule réponse", on a :

$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) \text{ Pourquoi } (x // Ax, b)$

Si l'interviewé a répondu : "parce que je trouve les émissions de l'ORTF moins bonnes qu'avant", cette réponse codée correspond à :

$$d(I) = Bb \ \& \ Gb \ \& \ \forall x (Bx \rightarrow (Ax \leftrightarrow (Cx \vee Dx \vee Ex \vee Fx \vee Gx \vee Hx)))$$

f) Irréductibilité à une forme unique.

Nous le voyons, l'étude des formes simples de questions, autres que les interrogatives *whether* et *which*, est seulement esquissée par l'auteur. Il semble d'ailleurs que BELNAP ait considéré primitivement comme des variantes des questions *which* des énoncés de la forme :

- "Qui est le Président de l'Assemblée Nationale ?"
- "Qui était Christophe Colomb ?"
- "Pourquoi les corps lourds tombent-ils ?"

ainsi que les questions introduites par : "où", "quand", "comment", etc.

[Belnap 1963 : 41-43]. Cette manière de voir très réductionniste permettait par exemple d'écrire le sujet de la dernière interrogative ci-dessus sous la forme :

$$\sigma = (x \text{ est une phrase // } x \text{ est une explication du fait que les corps lourds tombent}).$$

Il est clair cependant qu'une telle représentation est purement verbale, et n'apporte rien à l'analyse du mécanisme de la réponse aux questions de causalité [Belnap 1976 : 84].

Il semble donc qu'il faille renoncer à réduire les questions de causalité aux questions *which*, l'intérêt de l'analyse de BROMBERGER étant au contraire de permettre l'identification d'une réponse à une question de causalité à partir de sa seule forme syntaxique [Belnap 1966 : 610]. Mais il est d'autres obstacles plus dirimants aux visées réductionnistes. Soit par exemple la question de description déjà citée :

"De quelle couleur est Dick ?"

Malgré l'apparence, il n'est pas possible de la symboliser par :

$$I = ? \rho (x \text{ est une couleur // } x \text{ est la couleur de Dick}).$$

En effet, la condition catégorielle ainsi posée implique que l'on dispose, dans le langage formel, de noms pour désigner les propriétés exprimées par les

descripteurs. Ceci sort évidemment du cadre d'un calcul des prédicats du premier ordre [Belnap 1976 : 80].

2.2.4. Les questions complexes

Toutes les formes de questions présentées dans les paragraphes précédents étaient des formes simples de questions absolues sans quantificateur. Nous allons maintenant montrer brièvement comment les questions simples peuvent être combinées pour former des questions *composées* ; comment les questions absolues peuvent être relativisées, pour former en particulier des questions *conditionnelles* ; comment enfin les quantificateurs universel et existentiel sont utilisés pour former des questions *quantifiées*. Nous terminerons ce paragraphe sur l'intérêt de combiner ces diverses propriétés pour rendre compte de formes particulières d'interrogatives, et également pour ré-écrire les principales variétés de questions *which* sous la forme d'une question demandant un exemple unique.

a) Composition de questions simples

Les questions composées étudiées par BELNAP sont de quatre types : conjonction de questions, union de questions, union de sujets, union de requêtes. En règle générale (et ceci est particulièrement vrai pour les opérateurs booléens), l'opérateur utilisé s'applique aux réponses directes, et non aux questions elles-mêmes. Pour simplifier l'écriture, on utilisera cependant entre les questions élémentaires le même opérateur qu'entre les réponses directes aux questions élémentaires.

Une *conjonction de questions* est une interrogative de la forme :

$$I = I_1 \ \& \ \dots \ \& \ I_i \ \& \ \dots \ \& \ I_n$$

dans laquelle les I_i sont des interrogatives simples. Une réponse directe à une conjonction d'interrogatives est la conjonction d'une réponse directe à chacune des interrogatives élémentaires qui la composent :

$$d(I) = d(I_1) \ \& \ \dots \ \& \ d(I_i) \ \& \ \dots \ \& \ d(I_n)$$

L'ensemble des réponses possibles à une conjonction de questions peut être représenté par le produit cartésien des ensembles des réponses directes possibles à chacune des questions élémentaires de la conjonction. Soit par exemple :

$$\begin{aligned} I &= \text{"Etes-vous déjà allé en Allemagne, et êtes-vous déjà allé en} \\ &\quad \text{Suède ?"} \\ &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (A, \bar{A}) \ \& \ ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (S, \bar{S}) \end{aligned}$$

La liste des réponses directes possibles à I est : $A \ \& \ S$, $A \ \& \ \bar{S}$, $\bar{A} \ \& \ S$, $\bar{A} \ \& \ \bar{S}$; cette liste correspond à l'ensemble des couples de modalités de réponse : $\{A, \bar{A}\} \times \{S, \bar{S}\} = \{(A, S), (A, \bar{S}), (\bar{A}, S), (\bar{A}, \bar{S})\}$.

Une *union de questions* est une interrogative de la forme :

$$I = I_1 \ \cup \dots \cup \ I_i \ \cup \dots \cup \ I_n$$

Par exemple :

$$\begin{aligned} I &= \text{"Etes-vous déjà allé en Allemagne, ou êtes-vous déjà allé en} \\ &\quad \text{Suède ?"} \\ &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (A, \bar{A}) \ \cup \ ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (S, \bar{S}) \end{aligned}$$

La raison du choix de l'opérateur booléen, de préférence à l'opérateur logique de disjonction, apparaît lorsque l'on considère qu'une expression A est une réponse directe à I si et seulement si A est une réponse directe à l'une au moins des I_i : par conséquent, l'ensemble des réponses possibles à une union de questions est l'union des ensembles des réponses possibles à chacune des questions élémentaires qui la constituent. Dans l'exemple cité, l'ensemble des réponses possibles à I est : $\{A, \bar{A}\} \cup \{S, \bar{S}\} = \{A, \bar{A}, S, \bar{S}\}$. Nous verrons également, à propos de la quantification des questions, que l'opérateur d'union peut être utilisé, en logique érotétique, comme une sorte de quantificateur existentiel.

Une *union de sujets* s'écrit :

$$\sigma = \sigma_1 \ \cup \dots \cup \ \sigma_i \ \cup \dots \cup \ \sigma_n$$

et l'ensemble des modalités de réponse présentées par σ est l'union des ensembles des modalités de réponse présentées par les σ_i . L'intérêt de cette notion est qu'elle ne se confond pas avec l'union de questions, et qu'elle permet de rendre compte de formes de questions mixtes. Le premier point est facile à montrer sur un exemple simple. Soit :

$$I_1 = ? \left(\begin{array}{c} 2 \\ 2 \end{array} - - \right) \sigma_1 \quad , \quad \text{avec } \sigma_1 = (A, B, C)$$

$$I_2 = ? \left(\begin{array}{c} 2 \\ 2 \end{array} - - \right) \sigma_2 \quad , \quad \text{avec } \sigma_2 = (D, E, F)$$

Soit la réponse :

$$d(I) = A \ \& \ D$$

Il est aisé de constater que cette réponse est bien une réponse directe à : $I_3 = ? \left(\begin{array}{c} 2 \\ 2 \end{array} - - \right) \sigma_1 \cup \sigma_2$, mais qu'elle ne saurait en aucun cas être une réponse directe à $I_4 = I_1 \cup I_2$ (En règle générale, l'union des interrogatives et l'union des sujets correspondants admettent les mêmes réponses directes si et seulement si toutes les requêtes demandent un exemple unique). Les unions de sujets constituent une formalisation commode pour des *questions mixtes*, comme l'interrogation *which-whether* ci-dessous, extraite par BELNAP d'un formulaire administratif :

I = "Je n'ai jamais été marié ; j'ai été marié pour la première fois en _____".

La requête est ici une demande de réponse vraie unique, commune à l'union d'un sujet *whether* et d'un sujet *which*. Cet exemple est formalisable en :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) ((\text{je n'ai jamais été marié}) \cup (\text{x est une date // j'ai été marié pour la première fois en x}))$$

L'interrogative *whether* : $I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (\text{je n'ai jamais été marié})$ est ce que l'on appelle une *question de HOBSON*, de la forme : $? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (A) = \text{"Dites-moi que A"}$. Sans utilité considéré isolément, ce type de question *whether*, demandant une réponse vraie unique (ou un exemple unique) et n'ayant qu'une réponse directe possible : $d(I) = A$, joue un rôle dans les unions de sujets et dans les interrogatives conditionnelles.

Une union de requêtes s'écrit :

$$\rho = \rho_1 \cup \dots \cup \rho_i \cup \dots \cup \rho_n$$

L'intérêt de cette notion est illustrée par l'auteur par l'exemple suivant :

I = "S'il n'y a pas plus de cinq paires de nombres premiers, quelles sont-elles ? S'il y en a plus de cinq, quels en sont au moins six exemples ?"

formalisable en :

$$I = ? \left(\left(\begin{matrix} 5 \\ 1 \end{matrix} \forall \neq \right) \cup \left(\begin{matrix} - \\ 6 \end{matrix} - \neq \right) \right) (x, y // P_x \& P_y \& x + 2 = y)$$

avec $P_x = "x \text{ est un nombre premier}"$.

b) Relativisation des questions absolues

Pour aborder le problème des *questions relativisées*, il nous faut au préalable définir le caractère opératoire d'une question. On dira qu'une question est *opératoire relativement à une interprétation M* si et seulement si elle demande une réponse dans M. Par définition, les questions absolues ont les deux propriétés suivantes : elles sont opératoires quel que soit M, et elles ont le même ensemble de réponses directes possibles quel que soit M (les questions relativisées qui possèdent également ces deux propriétés sont appelées : questions *catégoriques*).

Une *question conditionnelle* I est une question J relativisée, dont le caractère opératoire dépend de la réalisation d'une condition P. On l'écrira :

$$I = (P/J)$$

et l'on dira que I demande une réponse dans une interprétation M si et seulement si : la condition P est vraie dans M et la question J demande une réponse dans M. Si la condition P n'est pas vraie, la notion de réponse directe à I n'est pas définie. Si la question conditionnelle I est opératoire dans M, une expression A est une réponse directe à I dans M si et seulement si A est une réponse directe à J dans M. Soit par exemple [Belnap 1969 : 126] :

$$I = \text{"Si la marée est haute, le rocher est-il recouvert par l'eau ?"} \\ = (H / ? (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -) (R, \bar{R}))$$

Cette question n'a de réponse possible que si la marée est haute ; et dans ce cas, l'ensemble des réponses directes possibles est : $\{R, \bar{R}\}$.

Une variété intéressante de questions conditionnelles est celle des *questions assurées*. Lorsque l'on ignore si une question a une réponse directe vraie (si elle est *sûre* : cf. § 2.2.5.d), on peut prendre la précaution de demander : "si une telle réponse existe, donnez-la moi s'il vous plait". Soit par exemple :

$$J = \text{"Était-ce un suicide ou un meurtre ?"} \\ = ? (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -) (S, M)$$

Si le décès est dû à une mort naturelle ou à un accident, la question J n'a pas de réponse directe vraie ; en toute rigueur, elle ne devrait pas alors être posée. On peut *assurer* J en demandant :

$$I = \text{"S'il s'agit bien d'un suicide ou d'un meurtre, duquel des deux s'agit-il ?"} \\ = (S \vee M / ? (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -) (S, M))$$

Cette formulation évite de solliciter une réponse vraie dans les cas où il n'en existe pas : la condition non satisfaite rend simplement la question J non opératoire. Appliquée aux *questions de HOBSON*, cette forme de relativisation aboutit à la forme d'interrogative usuelle :

$$I = (P / ? (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -) (P)) \\ = \text{"Si P est vrai, dites-le moi"}$$

observable fréquemment dans la vie courante (ex. : "Si vous partez, prévenez-moi").

Les exemples de questions assurées se rencontrent surtout dans les tests et les questionnaires remplis directement par le sujet (épreuves collectives, questionnaires postaux, etc.). Par exemple :

- "Quand deux mots signifient la même chose, soulignez le mot même"
(NB : 3)
- "Lisez chaque phrase et barrez : V si c'est toujours vrai [...] "
(ECO 5 : 8)
- "Si vous n'appartenez pas à l'une des 14 catégories ci-contre écrivez dans le cadre profession (n° 8) les renseignements qui vous concernent"
(REL 55)
- "Si vous avez interrompu votre activité professionnelle au cours de votre carrière, décrivez avec précision dans quelles circonstances"
(F.ING 69 : III 15)
- "Si vous avez hésité, qu'est-ce qui a finalement déterminé votre choix ?"
(F.ING 69 : VI 9)
- "Si vous envisagez de travailler professionnellement dès la fin de vos études à l'I.U.T., savez-vous déjà dans quelle branche d'activité vous exerceriez ?"
(OCDE IUT : 50a)

Par contre, les autres variétés de questions conditionnelles ne figurent pratiquement jamais dans notre corpus, et semblent d'ailleurs d'une utilité très restreinte dans les tests et les questionnaires (à part les consignes du type : "Quand vous aurez fini, dites-moi combien vous avez trouvé"). Il faut se garder de confondre en effet les questions conditionnelles avec les questions hypothétiques (très fréquentes dans notre corpus) et les questions "puisque" (beaucoup plus rares, mais cependant observables).

Les questions hypothétiques correspondent à des énoncés interrogatifs assez voisins de ceux des questions conditionnelles. Soit par exemple l'interrogative conditionnelle [Belnap 1972 : 335] :

$$\begin{aligned}
 I &= \text{"Si vous sortez tout de suite, est-ce que vous prenez un parapluie ?"} \\
 &= (S / ? (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - -) (P, \bar{P}))
 \end{aligned}$$

devant être comprise comme ne demandant une réponse que si l'interlocuteur est effectivement sur le point de sortir (l'anglais dit sans ambiguïté : "If you are going, are you taking an umbrella ?"). Sa signification est très différente de celle de l'interrogative hypothétique :

J = "Si vous sortiez, est-ce que vous prendriez un parapluie ?".

Ceci est aisé à constater en partant des réponses directes à chacune des questions. La première a pour réponses possibles :

$$\begin{cases} d(I)_1 = \text{"Oui, je prends un parapluie"} \\ d(I)_2 = \text{"Non, je ne prends pas de parapluie"}. \end{cases}$$

"Mais si le répondant ne sort pas, [...] alors tout ce que nous pouvons dire est qu'aucune réponse n'est requise "[Belnap 1976 : 104]. La seconde question a pour réponses possibles [Belnap 1976 : 96] :

$$\begin{cases} d(J)_1 = \text{"Oui, si je sortais, je prendrais un parapluie"} \\ d(J)_2 = \text{"Non, si je sortais, je ne prendrais pas de parapluie"}. \end{cases}$$

Malgré certaines réserves sur la correspondance entre le "si" et l'implication telle qu'elle a été définie, on peut symboliser les réponses directes à J par :

$$\begin{cases} d(J)_1 = S \rightarrow P \\ d(J)_2 = S \rightarrow \bar{P} \end{cases}$$

La question J devient en conséquence :

$$J = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (S / \rightarrow / (P, \bar{P}))$$

(une autre écriture est également possible : cf [Belnap 1976 : 96-97]). Ainsi, les questions hypothétiques se distinguent des questions conditionnelles en ce que la condition exprimée dans l'énoncé interrogatif ne détermine pas le caractère opératoire de la question, mais au contraire est un élément du sujet de l'interrogative, et figure dans chaque modalité de réponse possible.

A titre d'exemples d'interrogatives hypothétiques tirées de notre corpus, nous citerons :

- "Si vous dînez au restaurant, préférez-vous qu'un autre commande le menu pour vous ?" (BERN 35 : 63)
- "Si vous travailliez dans une usine de mécanique, préféreriez-vous travailler à l'établissement de dessins soignés, ou discuter avec des acheteurs éventuels du produit ?" (CAT 16 PF : 154)
- "Lorsque vous êtes seul à la maison, vous préférez : a) lire un livre b) jouer à un jeu qui vous permette de bouger" (E/I 6 : 23)

- "S'il arrivait que le parti ou la tendance que vous préférez ne puisse présenter un candidat, quelle solution choisiriez-vous ? Vous abstenir, ou bien voter pour le candidat d'un autre parti ?" (IFOP 58-11 : 30a)
- "Si vous aviez un fils ou une fille en âge de se marier, quel conseil lui donneriez-vous à ce sujet ?" (INED 081 : opinions 2)
- "S'il vous arrive de manquer un numéro de ce magazine, avez-vous l'impression qu'il vous manque quelque chose ?" (SOFRES 686 : 27)

On remarquera que, même si la condition exprimée dans l'énoncé est au présent de l'indicatif, il ne peut s'agir de questions conditionnelles ; en effet, l'interlocuteur n'a pas à s'interroger sur la valeur de vérité de la condition pour décider s'il répondra ou non, et l'énoncé de la condition figure nécessairement dans les réponses directes.

Pour illustrer les questions "puisque", nous prendrons un exemple proche des précédents :

I = "Puisque vous sortez, est-ce que vous prenez un parapluie ?"

Les réponses directes à la question sont |Belnap 1976 : 98| :

$$\begin{cases} d(I)_1 = \text{"Oui, je sors et je prends un parapluie"} \\ d(I)_2 = \text{"Non, je sors mais je ne prends pas de parapluie"} \end{cases}$$

formalisables en :

$$\begin{cases} d(I)_1 = S \ \& \ P \\ d(I)_2 = S \ \& \ \bar{P} \end{cases}$$

La question elle-même s'écrit :

$$I = S \ / \ \& \ / \ ? \ \left(\begin{matrix} 1 \\ 1 \end{matrix} \ - \ - \right) \ (\ P, \ \bar{P} \)$$

Cette forme de question, comme la question hypothétique, est toujours opératoire. En français, les questions "puisque" sont introduites également par des locutions comme : "maintenant que ...", "étant donné que ...", "compte-tenu de ce que ...", etc. Dans les tests et les questionnaires, son utilité paraît très réduite, si ce n'est pour rappeler à la personne interrogée une de ses réponses antérieures, ou bien une situation supposée connue de tous :

- "Maintenant, ça fait déjà quelque temps que vous avez votre ...
[modèle de voiture cité à Q 1] ; à l'usage, et dans l'ensemble, en êtes-vous très satisfait, assez satisfait, pas très satisfait, ou plutôt mécontent ?" (SOFRES 967 : 7)
- "Comme vous le savez, l'Algérie et les territoires d'Outre-Mer vont aussi voter pour le référendum. Comment croyez-vous que va voter l'Algérie ? Et l'Afrique noire ?" (IFOP 58-9 : 19)

c) Quantification des questions simples

La quantification des questions se réduit à l'usage du quantificateur universel \forall . En effet, BELNAP déclare n'avoir pas découvert d'utilisation du quantificateur existentiel \exists , et "ne pas s'attendre à en découvrir, tant que l'on n'a pas trouvé auparavant un usage pour l'opération (finie) de disjonction sur des interrogatives, dont elle serait une généralisation" [Belnap 1976 : 100]. Une question universalisée $\forall x Ix$ comporte une variable libre x . Soit par exemple :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (P_x , \overline{P_x})$$

la question $\forall x Ix$ peut être paraphrasée en "Pour tout x , est-ce que x est P ou \overline{P} ?". Si le domaine des x est fini, il est possible d'apporter une réponse directe à la question : on dira que $\forall x Ix$ demande une réponse dans M si et seulement si une réalisation Ia de Ix demande une réponse dans M ; et que, si $\forall x Ix$ est opératoire dans M , une expression A est une réponse directe à $\forall x Ix$ si et seulement si elle est de la forme :

$A = A_1 a_1 \& \dots \& A_i a_i \& \dots \& A_n a_n$, où les $A_i a_i$ sont les réponses directes à l'ensemble des n réalisations Ia_i de I [Belnap 1972 : 338].

Si le domaine des x est infini, il est toutefois possible de répondre effectivement à $\forall x Ix$ à la condition que l'ensemble des x correspondant à l'une des deux réponses directes soit fini. Si, dans l'exemple ci-dessus, on a $\overline{P_x}$ pour $x = a$, $x = b$, et $x = c$, et P_x pour tous les autres x , on a alors :

$$d(I) = \overline{P_a} \& \overline{P_b} \& \overline{P_c} \& \forall x (x \neq a \& x \neq b \& x \neq c \rightarrow P_x)$$

La règle générale est alors qu'une interrogative de la forme $\forall x Ix$ a pour réponse directe :

$$A_1 a_1 \ \& \ \dots \ \& \ A_n a_n \ \& \ \forall x (x \neq a_1 \ \& \ \dots \ \& \ x \neq a_n \ \rightarrow Bx)$$

où les expressions : $A_1 x, \dots, A_n x, Bx$ sont toutes des réponses directes à l'interrogative I [Belnap 1976 : 99].

Il est possible d'introduire des conditions catégorielles au niveau du quantificateur lui-même, et de poser par exemple :

$$\begin{aligned} I &= \text{"Pour tout } x \text{ qui est un nombre entier compris entre 10 et 20,} \\ &\quad x \text{ est-il premier ou non ?"} \\ &= \forall_{[Cx]} ? \left(\begin{array}{l} 1 \\ 1 \end{array} - - \right) (Px, \overline{Px}) \end{aligned}$$

avec : $Cx = \text{"}x \text{ est un nombre entier compris entre 10 et 20"}$.

$Px = \text{"}x \text{ est premier"}$.

Les réponses directes seront de la forme :

$$Pa_1 \ \& \ \dots \ \& \ Pa_n \ \& \ \forall x (Cx \rightarrow (x \neq a_1 \ \& \ \dots \ \& \ x \neq a_n \rightarrow \overline{Px}))$$

où les a_i sont des nombres entiers compris entre 10 et 20, et réputés premiers par le répondant.

A défaut de questions existentialisées, il est possible de trouver des *questions unionisées* qui leur ressemblent beaucoup, et qui utilisent le symbole de l'union comme un quantificateur existentiel. Ainsi :

$$Ux ? \left(\begin{array}{l} 1 \\ 1 \end{array} - - \right) (Px, \overline{Px})$$

signifie : "Pour un x quelconque, dites-moi s'il est P ou \overline{P} ". Une question unionisée $Ux Ix$ demande une réponse dans M si et seulement si il existe au moins une réalisation Ia de Ix qui demande une réponse dans M ; et une expression A est une réponse directe dans M à $Ux Ix$ si et seulement si elle est une réponse directe dans M à une réalisation par substitution Ia de Ix [Belnap 1972 : 340].

d) Autres formes de questions complexes

Le langage défini par BELNAP permet évidemment d'analyser des formes de questions à la fois quantifiées, conditionnelles, et composées. Les formes les plus intéressantes sont les conjonctions et les unions d'interrogatives conditionnelles, et les conditionnelles universalisées et unionisées.

Une *conjonction de conditionnelles* est de la forme :

$$I = I_1 \ \& \ \dots \ \& \ I_j \ \& \ \dots \ \& \ I_n$$

où chaque I_j est une interrogative conditionnelle. Une conjonction de conditionnelles I demande une réponse dans M si et seulement si une I_j au moins demande une réponse dans M . Appelons $\{I_\alpha, \dots, I_\mu\}$ l'ensemble de toutes les interrogatives conditionnelles parmi les I_j qui demandent une réponse dans M ; une expression A est une réponse directe dans M à I si et seulement si $A = A_\alpha \ \& \ \dots \ \& \ A_\mu$, où les A_α sont des réponses aux I_α correspondantes. Comme illustration d'une conjonction de conditionnelles, l'auteur cite [Belnap 1969 : 127] :

$I =$ "Si la marée est haute, le rocher est-il recouvert par l'eau, et s'il s'agit bien d'un suicide ou d'un meurtre, duquel des deux s'agit-il ?

$$= (H / ? \binom{1}{1} - -) (R, \bar{R}) \ \& \ (S \vee M / ? \binom{1}{1} - -) (S, M)$$

Notons le parallèle entre une conjonction de *questions de HOBSON* conditionnelles et une question *whether* demandant une énumération exhaustive. Par exemple :

$$\left\{ \begin{array}{l} I = (A / ? \binom{1}{1} - -) (A) \ \& \ (B / ? \binom{1}{1} - -) (B) \\ J = ? \binom{-}{1} \vee -) (A, B) \end{array} \right.$$

Les réponses directes $d(I)$ et $d(J)$ ont, dans une même interprétation M , la même sélection S ; les différences entre $d(I)$ et $d(J)$ tiennent à ce que $d(J)$ exprime en outre l'exhaustivité (maximum), et à ce que J peut avoir des réponses fausses [Belnap 1976 : 106].

Une union de conditionnelles s'écrit :

$$I = I_1 \cup \dots \cup I_j \cup \dots \cup I_n$$

où chaque I_j est une interrogative conditionnelle. Une union de conditionnelles I demande une réponse dans M si et seulement si au moins une des I_j demande une réponse dans M . Si I est opératoire dans M , une expression A est une réponse directe à I dans M si et seulement si A est une réponse directe dans M à l'une des I_j . En remplaçant, dans l'énoncé de l'exemple utilisé à l'alinéa précédent, la conjonction *et* par : *ou*, on obtient l'union de conditionnelles suivante :

$$I = (H / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (R, \bar{R})) \cup (S \vee M / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (S, M))$$

[Belnap 1969 : 128]. On remarquera à ce propos le parallélisme entre une union de questions de HOBSON conditionnelles et une question *whether* demandant un exemple unique. Ainsi :

$$\left\{ \begin{array}{l} I = (A / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A)) \cup (B / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (B)) \\ J = ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A, B) \end{array} \right.$$

Toute réponse directe vraie dans M à la question J est également une réponse directe vraie dans M à la question I . Par contre, seule l'interrogative J peut avoir des réponses fausses [Belnap 1976 : 106].

Une forme possible de conditionnelle universalisée est :

$$I = \forall x (Cx / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (Px, \bar{Px}))$$

que l'on peut réaliser en : "Pour tout x , si x est un nombre compris entre 10 et 20, x est-il premier?". Cette question a évidemment 2^9 réponses directes possibles, dont une seule est vraie. Si l'on remplace la question *whether* à deux modalités de réponse par une question de HOBSON conditionnelle, on obtient :

$$I = \forall x (Cx / (Px / ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (Px)))$$

qui a exactement les mêmes réponses directes que la question *which* demandant une énumération exhaustive :

$$J = ? \left(\begin{smallmatrix} - \\ 1 \end{smallmatrix} \vee - \right) (Cx // Px)$$

Nous avons donc là un moyen de réduire une question *which* à une forme de question *whether*. Cette réduction ne présente d'intérêt que lorsque le nombre de C qui sont P est fini (cf. [Belnap 1972 : 339]).

De même, une *conditionnelle unionisée* peut avoir la forme :

$$I = \cup x (Cx / ? (\frac{1}{1} - -) (Px, \overline{Px}))$$

réalisable en : "Pour un x quelconque compris entre 10 et 20, x est-il premier ?". En remplaçant la question à deux modalités de réponse par une *question de HOBSON*, on obtient :

$$I = \cup x (Cx / ? (\frac{1}{1} - -) (Px))$$

qui équivaut à une question *which* demandant un exemple unique :

$$J = ? (\frac{1}{1} - -) (Cx // Px)$$

On peut de même réécrire une question *which* demandant une réponse unique, en utilisant une *question de HOBSON* conditionnelle, à laquelle on ne répond que si la réponse vraie est unique. Soit alors le couple de questions équivalentes :

$$\left\{ \begin{array}{l} I = \cup x (Cx / (Px \& \forall y (Cy \rightarrow (Py \rightarrow y = x))) / ? (\frac{1}{1} - -) (Px)) \\ J = ? (\frac{1}{1} \forall -) (Cx // Px) \end{array} \right.$$

Les autres variantes étudiées par BELNAP ne semblent pas être exactement équivalentes à une question *which* à réponse unique (cf. [Belnap 1972 : 340-341]).

Pour ré-écrire les questions *which* demandant des exemples multiples, l'auteur introduit une variable F qui prend ses valeurs dans des ensembles finis $\{a_1, \dots, a_n\}$. On écrira $y \in F$ si et seulement si $y = a_1 \vee \dots \vee a_n$.

On a alors le couple de questions équivalentes :

$$\left\{ \begin{array}{l} I = \cup F \forall x (x \in F / (Cx / ? (\frac{1}{1} - -) (Px))) \\ J = ? (\frac{1}{1} - -) (Cx // Px) \end{array} \right.$$

L'interrogative I a des réponses directes de la forme Pa_i , $Pa_i \& Pa_j$, $Pa_i \& \dots \& Pa_k$, telles que les a_i soient tous des C [Belnap 1972 : 341].

2.2.5. Analyse sémantique des questions et des réponses

Pour analyser la signification des questions et de leurs réponses, il est nécessaire de décrire les relations sémantiques existant entre questions et réponses, c'est-à-dire entre des questions et des assertions. Il serait commode de disposer pour cela du même outillage logique que celui qui est applicable aux propositions. Or, comme l'a vigoureusement rappelé HAMBLIN dans une discussion des thèses de LEONARD et de HARRAH, "les questions ne sont pas des propositions" [Hamblin 1963]. Ce qui implique que des concepts tels que valeur de vérité, implication, équivalence, etc. ne puissent être appliqués tels quels aux interrogatives. Cependant, comme l'avait souligné HARRAH, toute question a un *contenu propositionnel* : en d'autres termes, poser une question, ce n'est pas seulement requérir de l'information, c'est aussi en apporter à l'interlocuteur. Ainsi, demander le nom du Monsieur que l'on a aperçu avec Madame Untel apporte de l'information au répondant [Belnap 1976 : 115] ; et si l'on considère que l'affirmation : "Jean avait l'habitude de battre sa femme" est un mensonge, la question : "Jean a-t-il cessé de battre sa femme ?" en est un également [Belnap 1963 : 130]. D'où la tentation d'étendre aux questions la notion de valeur de vérité, pourtant applicable en principe aux seules expressions.

BELNAP sort élégamment de ce dilemme en créant la catégorie des *quasi-expressions*, auxquelles appartiennent à la fois les expressions et les questions [Belnap 1976 : 121] (cf. dans les travaux antérieurs les *quasi-propositions* [Belnap 1963 : 134]). Les notions de valeur de vérité, d'implication logique, et d'équivalence logique, s'appliquent à toutes les quasi-expressions sans distinction, permettant ainsi de définir certaines relations entre les questions et certains types particuliers d'expressions, comme les réponses directes et les pré-suppositions.

a) Valeur de vérité d'une question

Nous allons commencer par définir ce qu'est une *question vraie* (ou fausse). Rappelons tout d'abord comment se définit la valeur de vérité d'une réponse à une question [Belnap 1976 : 71-72]. Une *réponse abstraite* à une interrogative I se définit comme un couple ordonné $\langle I, S \rangle$, où S est

la sélection abstraite contenue dans la réponse. Soit

$S = \{A_1, \dots, A_i, \dots, A_n\}$ la sélection abstraite d'une réponse directe

à une question *whether*. Une réponse abstraite à une question *whether* est dite vraie dans une interprétation M si et seulement si toutes les A_i éléments de S sont vraies dans M (et, éventuellement, si l'affirmation de complétude est également vraie dans M). Soit maintenant $S = \{Pa_1, \dots, Pa_i, \dots, Pa_n\}$

la sélection abstraite d'une réponse directe à une question *which*. Une réponse abstraite à une question *which* est dite vraie dans une interprétation M si et seulement si tout élément Pa_i est vrai dans M (et, s'il y a exigence de complétude, si et seulement si toute modalité réelle présentée par I et qui est vraie dans M est un élément de la sélection S ; et, s'il y a exigence de distinction, si et seulement si aucune modalité de réponse vraie dans M n'apparaît plus d'une fois dans la sélection S).

On dira qu'une question est vraie dans une interprétation M si et seulement si au moins une de ses réponses directes est vraie dans une interprétation M [Belnap 1963 : 133]. Plus spécifiquement [Belnap 1976 : 119] :

- une question *whether* est vraie dans M si et seulement si au moins une de ses réponses directes est vraie dans M ;
- une question *whether* est fausse dans M si et seulement si toutes ses réponses directes sont fausses dans M ;
- une question *which* est réellement vraie dans M si et seulement si au moins une de ses réponses réelles est vraie dans M ;
- une question *which* est nominalement vraie dans M si et seulement si au moins une de ses réponses nominales directes est vraie dans M ;
- une question *which* est réellement fausse dans M si et seulement si toutes ses réponses réelles sont fausses dans M ;
- une question *which* est nominalement fausse dans M si et seulement si toutes ses réponses nominales directes sont fausses dans M .

Considérons maintenant des quasi-expressions X et X' , et des ensembles de quasi-expressions H et H' . Pour fixer les idées, on peut se

représenter un ensemble H comme l'ensemble des propositions et des interrogatives que le locuteur (ou l'interlocuteur) sait ou croit être vraies [Belnap 1976 : 124]. En pratique, il est utile de relativiser la valeur de vérité d'une quasi-expression non seulement à une interprétation M , mais aussi à un ensemble de quasi-expressions H . En effet, soit par exemple la question : "Jean a-t-il cessé de battre sa femme ?". Une telle question est vraie si l'ensemble H auquel elle appartient a également pour élément la proposition : "Jean a battu sa femme" ; elle est fausse si H , au contraire, a parmi ses éléments la proposition : "Jean n'a jamais battu sa femme" [Belnap 1976 : 131]. On dira qu'une interprétation M est une *H-interprétation* si et seulement si toute expression élément de H est vraie dans M . Ceci posé, les principales définitions relatives aux quasi-expressions sont les suivantes [Belnap 1976 : 122] :

- X est (logiquement) *H-vrai* s'il est vrai dans toutes les H -interprétations ;
- X est *H-consistant* s'il est vrai dans au moins une H -interprétation ;
- X est *H-inconsistant* s'il n'est vrai dans aucune H -interprétation ;
- X est *H-consistant* avec X' si X et X' sont vrais simultanément dans au moins une H -interprétation ;
- X est *H-inconsistant* avec X' si X et X' ne sont jamais vrais simultanément dans une même H -interprétation ;
- X' *H-implique* (propositionnellement) X si et seulement si X est vrai dans toute H -interprétation dans laquelle X' est vrai ;
- H' *H-implique* (propositionnellement) X si et seulement si X est vrai dans toute H -interprétation dans laquelle toute quasi-expression élément de H' est vraie ;
- X et X' sont (propositionnellement) *H-équivalents* si et seulement si X H -implique X' et X' H -implique X (c'est-à-dire si et seulement si X et X' sont vrais dans exactement les mêmes H -interprétations).

Un corollaire de la définition de l'implication entre quasi-expressions, et de la définition de la vérité d'une question, est que toute réponse directe $d(I)$ H -implique I .

b) Présupposition d'une question

Nous disposons maintenant des concepts nécessaires pour aborder le problème de la présupposition d'une question. BELNAP commence par affirmer que : "toute question présuppose *précisément* qu'une au moins de ses réponses directes est vraie", et qu' "une phrase *exprime la présupposition* d'une question si sa vérité est une condition nécessaire et suffisante pour que la question ait au moins une réponse vraie" [Belnap 1966 : 610]. C'est affirmer, contrairement à LEONARD par exemple, qu'une question n'a pas plusieurs présuppositions mais (au plus) une seule. Ce que l'on entend généralement par présuppositions d'une question recouvre, en effet, outre *la* présupposition au sens de BELNAP, diverses assertions liées au contexte de la question (présuppositions pragmatiques), et diverses assertions impliquées par la présupposition elle-même (présuppositions secondaires au sens de LEONARD) (cf. § 2.1.2.a). Soit par exemple la question fameuse : "Jean a-t-il cessé de battre sa femme ?". Les *présuppositions pragmatiques* usuelles sont que le locuteur ne connaît pas la réponse vraie à sa question, qu'il désire la connaître, qu'il pense que son interlocuteur la connaît ou peut en prendre connaissance, qu'il suppose que ce dernier acceptera de la lui donner, etc. Parmi les *présuppositions secondaires*, il y a le fait qu'il existe un personnage appelé Jean, qu'il est marié, qu'il avait l'habitude de battre quelqu'un, etc. ; ces propositions sont impliquées logiquement par *la* présupposition de la question, qui est : "Jean avait l'habitude de battre sa femme".

Pour formaliser ceci, nous dirons qu'une interrogative *I* présuppose une expression *A* si et seulement si *A* est vraie dans toutes les interprétations dans lesquelles *I* est vraie. Cette définition entraîne les propriétés suivantes :

- *I* H-présuppose *A* si et seulement si *I* H-implique *A* ;
- *A* exprime la présupposition *Prés(I)* de *I* relativement à *H* si et seulement si *A* est (propositionnellement) H-équivalent à *I* ;
- quand *Prés(I)* est définie, toute réponse directe *d(I)* H-implique *Prés(I)*.

La conséquence la plus importante sur le plan sémantique est que *le contenu propositionnel d'une interrogative est équivalent au contenu propositionnel de sa présupposition.*

Il est facile d'appliquer ces notions à l'analyse des questions *whether*. En effet, soit I une interrogative *whether*; une expression A exprime la présupposition $Prés(I)$ relativement à H si et seulement si A est H -vraie dans toutes les interprétations dans lesquelles I est H -vraie. $Prés(I)$ est alors la disjonction de l'ensemble de ses réponses directes $d(I)$. On a par exemple [Belnap 1963 : 137] :

$$\begin{aligned} I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (A, B) & \quad Prés(I) = (A \& \bar{B}) \vee (\bar{A} \& B) \\ I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (A, B) & \quad Prés(I) = A \vee B \\ I = ? \left(\begin{array}{c} \bar{1} \\ 1 \end{array} \vee - \right) (A, B) & \quad Prés(I) = A \vee B \vee (A \& B) = A \vee B \end{aligned}$$

En reprenant la question citée plus haut, si l'on pose [Belnap 1976 : 22, 114] :

$$\begin{aligned} I &= \text{"Jean a-t-il cessé de battre sa femme ?"} \\ &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (B \& C, B \& \bar{C}) \end{aligned}$$

avec : $B = \text{"Jean avait l'habitude de battre sa femme"}$

$C = \text{"Jean a cessé de battre sa femme"}$

on retrouve bien la présupposition :

$$Prés(I) = (B \& C) \vee (B \& \bar{C}) = B$$

Remarquons à ce propos que l'on peut avoir des *questions sans présupposition*, c'est-à-dire des questions qui restent vraies quelle que soit l'interprétation choisie. Ce sont des interrogatives *whether* posant une alternative entre une modalité de réponse et sa négation. Par exemple [Belnap 1976 : 17] :

$$\begin{aligned} I &= \text{"Le verre est-il liquide à 20° centigrades ?"} \\ &= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (L, \bar{L}) \\ Prés(I) &= L \vee \bar{L} \end{aligned}$$

Pour les questions *which*, le problème de la définition de la présupposition est plus compliqué, en raison de la distinction entre modalités nominales et modalités réelles. BELNAP considère "qu'il n'est en général pas possible de découvrir pour chaque interrogative *which* I une expression exprimant la présupposition *nominale* de I" [Belnap 1976 : 121]. Par contre, il est possible de formuler la présupposition *réelle* d'une question *which*. Si I est une interrogative *which*, une expression A exprime la présupposition réelle *Prés(I)* relativement à H si et seulement si A est H-vraie dans toutes les interprétations dans lesquelles I est réellement H-vraie. Dans le cas le plus complexe, celui d'une question demandant une énumération exhaustive et distincte, on a :

$$I = ? \left(\begin{matrix} u \\ v \end{matrix} \vee \neq \right) (Cx // Px)$$

$$Prés(I) = P_1 \ \& \ P_2 \ \& \ P_3$$

Dans l'expression de la présupposition, chaque élément exprime qu'une exigence de la requête est satisfaite :

$$P_1 = \text{"au moins un C est P"}$$

$$P_2 = \text{"au plus u C sont P"}$$

$$P_3 = \text{"au moins v C sont P"}$$

Dans ce cas précis, P_3 rend P_1 redondant. Dans le cas général, P_1 est le seul élément obligatoire de l'expression ; P_2 correspond à l'affirmation que l'exigence d'exhaustivité est satisfaite lorsque u est défini ; P_3 correspond à l'affirmation que l'exigence de distinction est satisfaite.

c) Classifications sémantiques des réponses

L'analyse des relations sémantiques entre certaines quasi-expressions conduit à définir une typologie des réponses à une question, et une nouvelle typologie des questions. Cette analyse permet en particulier d'étudier des formes de réponses qui, tout en n'étant pas des réponses directes à la question posée, apportent cependant une certaine information à celui qui la pose.

Soit $r(I)$ une réponse quelconque à une interrogative I , et H un ensemble de quasi-expressions. On dira que la réponse $r(I)$ est, relativement à H [Belnap 1976 : 125] :

- *non informative* si I H-implique $r(I)$; par exemple, si $I =$ "Portait-elle son chapeau rouge ou son chapeau vert ?", et $r(I) =$ "L'un ou l'autre", on a : $? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (R, V) \rightarrow R \vee V$.
- *informative* sinon (c'est-à-dire si I , ou ce qui revient au même $Prés(I)$ si elle est définie, ne H-implique pas $r(I)$).
- *stupide* si $r(I)$ est H-inconsistante (c'est-à-dire si $r(I)$ n'est vraie dans aucune H-interprétation).
- *possible* sinon (c'est-à-dire si $r(I)$ est vraie dans au moins une H-interprétation).
- *relativement stupide* si $r(I)$ est H-inconsistante avec I (c'est-à-dire si $r(I)$ et I ne sont jamais simultanément vraies dans une H-interprétation).
- *relativement possible* sinon (c'est-à-dire si $r(I)$ et I sont simultanément vraies dans au moins une H-interprétation).

Si l'on examine les relations d'implication entre une réponse quelconque $r(I)$ à une interrogative I , et une réponse directe $d(I)$ à cette même interrogative, on est amené à distinguer quatre grands types de réponses :

- réponse complète : $r(I) \rightarrow d(I)$
- réponse partielle : $r(I) \leftarrow d(I)$
- réponse éliminatoire : $r(I) \rightarrow \overline{d(I)}$
- réponse quasi-éliminatoire : $r(I) \leftarrow \overline{d(I)}$

A ces quatre types s'ajoutent les deux cas particuliers suivants :

- réponse exactement complète : $(r(I) \rightarrow d(I)) \& (d(I) \rightarrow r(I))$
- réponse exactement éliminatoire : $(r(I) \rightarrow \overline{d(I)}) \& (\overline{d(I)} \rightarrow r(I))$.

Nous allons décrire et illustrer ci-après chacun de ces types de réponses, ainsi que certaines de leurs variantes présentant un intérêt pour l'analyse sémantique des tests ou des questionnaires. Auparavant, nous préciserons que ces relations définissent l'ensemble des réponses pertinentes à une interrogative donnée. BELNAP pose en effet qu'une expression $r(I)$ est *érotétiquement pertinente* à une interrogative I relativement à un ensemble de quasi-expressions H , si et seulement si $r(I)$ est une réponse à I qui soit (au moins) ou complète, ou partielle, ou éliminatoire, ou quasi-éliminatoire, relativement à H [Belnap 1963 : 148].

Une expression $r(I)$ est une *réponse complète* à une interrogative I relativement à H , si et seulement si $r(I)$ H -implique au moins une réponse directe à I . Il s'ensuit que toute modalité de réponse d'une question *whether* est une réponse complète à la question, même pour une question demandant une réponse vraie unique. Par exemple [Belnap 1963 : 138] :

$$I = \text{"Portait-elle son chapeau rouge ou son chapeau vert ?"} \\ = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (R, V)$$

la réponse codée : $r(I) = \text{"son chapeau rouge"} = R$, est une réponse complète à I , puisqu'elle implique la réponse directe $d(I) = R \ \& \ \bar{V}$. Bien plus, si l'une des modalités de réponse est elle-même une disjonction, tout élément de la disjonction peut constituer à lui seul une réponse complète. Soit l'exemple (inspiré de [Belnap 1976 : 126]) :

$I = \text{"Est-ce que vous vous décidez pour l'achat de la chaise et du fauteuil, ou est-ce que vous hésitez toujours entre l'armoire et le buffet ?"}$

$$= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (C \ \& \ F, A \ \vee \ B)$$

la réponse $r(I) = \text{"J'achète l'armoire"} = A$ est une réponse complète, puisqu'elle implique $d(I) = A \ \vee \ B$.

Une expression $r(I)$ est une *réponse partielle* à une interrogative I relativement à H si et seulement si $r(I)$ est H -impliquée par au moins une réponse directe à I . Bien que n'apportant pas toute l'information exigée d'une

réponse directe, une réponse partielle apporte néanmoins de l'information au locuteur (dans le cas général du moins). Dans le dernier exemple ci-dessus, la réponse $r(I) = \text{"J'achète la chaise"} = C$ est une réponse partielle, puisque la réponse directe $d(I) = C \ \& \ F$ implique C . Toutefois, il existe deux cas particuliers dans lesquels une réponse partielle n'apporte aucune information. Il y a tout d'abord ce que HARRAH appelle une "question test" [Harrah 1963 : 37]. Soit :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (A \ \& \ \bar{A}, B)$$

la réponse $d(I) = A \ \& \ \bar{A}$ implique n'importe quelle proposition, et ainsi toute proposition quelle qu'elle soit compte pour une réponse partielle à I . D'où la notion de réponse partielle régulière, correspondant ici aux réponses partielles impliquées par $d(I) = B$. Une expression $r(I)$ est une *réponse partielle régulière* à une interrogative I relativement à H si et seulement si $r(I)$ est impliquée par au moins une réponse directe à I qui soit H -consistante (c'est-à-dire vraie dans au moins une H -interprétation, ce qui n'est évidemment pas le cas pour $A \ \& \ \bar{A}$). Le second cas particulier est plus couramment observable : il correspond à une réponse qui est une disjonction de toutes les modalités de réponses possibles. Ainsi :

$$I = \text{"Était-ce un suicide ou un meurtre ?"}$$

$$= ? \left(\begin{array}{c} 1 \\ 1 \end{array} - - \right) (S, M)$$

$$r(I) = \text{"C'était l'un ou l'autre"}$$

$$= S \vee M.$$

La réponse $r(I)$ est une réponse partielle régulière. Néanmoins, elle n'apporte aucune information, car elle est impliquée par toutes les réponses directes à I . C'est pourquoi la variété de réponse partielle la plus utile en pratique est la réponse partielle fortement régulière, définie comme suit : une expression $r(I)$ est une *réponse partielle fortement régulière* à l'interrogative I relativement à H , si et seulement si $r(I)$ est une réponse partielle régulière à I et n'est pas H -impliquée par au moins une réponse directe à I . En d'autres termes, une réponse partielle fortement régulière est une réponse qui est à la fois partielle régulière et informative.

Une expression $r(I)$ est une *réponse exactement complète* à une interrogative I relativement à H , si et seulement si $r(I)$ est (propositionnellement) H -équivalente à une réponse directe à I . Par exemple, dans l'alternative :

$$I = ? \left(\begin{array}{l} 1 \\ 1 \end{array} \vee - \right) (A, \bar{A})$$

la réponse : $r(I) = A$ équivaut à la réponse directe $d(I) = A \& \bar{\bar{A}}$, et est une réponse exactement complète à I . Dans notre exemple précédent, $r(I) =$ "J'hésite entre le buffet et l'armoire" = $B \vee A$ est équivalente à $d(I) = A \vee B$, sans toutefois être identique à $d(I)$. L'intérêt de la distinction entre une réponse exactement complète et une réponse directe tient à ce qu'une réponse directe doit toujours pouvoir être effectivement évaluée, alors que l'équivalence propositionnelle n'est pas toujours décidable. Par exemple, $r(I) =$ "Le verre est liquide à $(2^4 + 2^2)^\circ$ centigrades" n'est propositionnellement équivalent à $d(I) =$ "Le verre est liquide à 20° centigrades" que relativement aux règles usuelles de l'arithmétique. Remarquons qu'une réponse exactement complète est une réponse à la fois partielle et complète. Remarquons en outre que (bien que la réciproque ne soit pas vraie) toute réponse directe est une réponse exactement complète.

Une expression $r(I)$ est une *réponse éliminatoire* à une interrogative I relativement à H si et seulement si $r(I)$ H -implique la négation d'au moins une réponse directe à I . On dira que $r(I)$ *élimine* $d(I)_j$ de I si et seulement si : $r(I) \rightarrow \overline{d(I)_j}$. Dans l'exemple déjà cité, $r(I) =$ "Je ne veux pas acheter la chaise" = \bar{C} élimine la réponse directe $d(I) =$ "J'achète la chaise et le fauteuil" = $C \& D$. La négation d'une réponse partielle est toujours une réponse éliminatoire puisque : $A \rightarrow B$ équivaut à : $\bar{B} \rightarrow \bar{A}$. On peut par conséquent faire correspondre aux réponses éliminatoires les variétés déjà définies pour les réponses partielles. On dira qu'une expression $r(I)$ est une *réponse éliminatoire régulière* à une interrogative I relativement à H , si et seulement si $r(I)$ H -implique la négation d'au moins une réponse directe à I qui soit H -consistante. En effet, si la réponse directe est déjà une contradiction (par exemple : $d(I) = A \& \bar{A}$), sa négation n'apporte aucune information. De même, on dira qu'une expression $r(I)$ est une *réponse éliminatoire fortement*

régulière à une interrogative I relativement à H , si et seulement si $r(I)$ est une réponse éliminatoire régulière à I et il existe au moins une réponse directe à I dont la négation n'est pas impliquée par $r(I)$.

Si une réponse éliminatoire régulière élimine toutes les réponses directes à une interrogative, il s'agit d'une variété de réponse éliminatoire particulièrement importante en pratique, appelée réponse correctrice. Par définition, une expression $r(I)$ est une *réponse correctrice* à une interrogative I relativement à H , si et seulement si $r(I)$ H -implique la négation de toutes les réponses directes à I . Cette définition suffit pour les questions *whether*.
A la question :

$$I = \text{"Portait-elle son chapeau rouge ou son chapeau vert ?"} \\ = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (R, V)$$

la réponse $r(I) = \text{"Ni l'un ni l'autre"} = \bar{R} \ \bar{V}$ est une réponse correctrice $Corr(I)$, qui nie la présupposition $Prés(I) = R \vee V$. Si l'on considère I comme une question demandant une réponse vraie unique : $I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \vee - \right) (R, V)$, il y a deux réponses correctrices possibles : $Corr(I)_1 = \text{"Ni l'un ni l'autre"} = \bar{R} \ \bar{V}$, et $Corr(I)_2 = \text{"Les deux"} = R \ \& \ V$, toutes deux équivalant à la négation de $Prés(I) = (A \ \& \ \bar{B}) \vee (\bar{A} \ \& \ B)$. Pour les questions *which*, il faut tenir compte de la distinction entre présupposition nominale et présupposition réelle. On dira par conséquent qu'une expression $r(I)$ est une *réponse correctrice nominale* à une interrogative I relativement à H , si et seulement si $r(I)$ H -implique la négation de toutes les $d(I)$. Au niveau réel, on aura les deux définitions suivantes (qui peuvent d'ailleurs s'appliquer également aux questions *whether*) :

- une expression $r(I)$ est une *réponse correctrice* à l'interrogative I relativement à H , si et seulement si I est (réellement) fausse dans toutes les H -interprétations dans lesquelles $r(I)$ est vraie ;

- une expression $r(I)$ est une *réponse exactement correctrice* à une interrogative I relativement à H , si et seulement si $r(I)$ est vraie dans toutes les H -interprétations dans lesquelles I est fausse, et dans elles seules.

Une expression $r(I)$ est une *réponse quasi-éliminatoire* à une interrogative I relativement à H , si et seulement si $r(I)$ est H -impliquée par la négation d'au moins une réponse directe à I (dans ses premiers travaux, BELNAP appelle ce type de réponse : *réponse de la quatrième espèce*). Dans l'exemple de l'achat de meubles ci-dessus : $I = ? \left(\begin{matrix} 1 \\ 1 \end{matrix} - - \right) (C \ \& \ F, A \ \vee \ B)$, la réponse : $r(I) = \text{"Je n'achète pas l'armoire"} = \bar{A}$ est une réponse quasi-éliminatoire, car impliquée par la négation de : $d(I) = A \ \vee \ B$. Remarquons que la négation d'une réponse complète est toujours une réponse quasi-éliminatoire.

Enfin, une expression $r(I)$ est une *réponse exactement éliminatoire* à une interrogative I relativement à H , si et seulement si $r(I)$ est (propositionnellement) H -équivalente à la négation d'au moins une réponse directe à I . En d'autres termes, une réponse exactement éliminatoire est à la fois éliminatoire et quasi-éliminatoire. Dans l'exemple rappelé à l'alinéa précédent, la réponse : $r(I) = \text{"Ce ne sera pas le buffet, ni d'ailleurs l'armoire"} = \bar{B} \ \& \ \bar{A}$ équivaut à la négation de la réponse directe $d(I) = A \ \vee \ B$. Notons à ce propos que la réponse exactement éliminatoire : $r(I) = \bar{B} \ \& \ \bar{A}$ est la négation de la réponse exactement complète : $r(I)' = B \ \vee \ A$.

d) Classifications sémantiques des questions

De même que l'analyse des relations sémantiques entre quasi-expressions conduit à une typologie des réponses, elle conduit également à une typologie sémantique des questions, indépendante de la typologie syntaxique déjà décrite.

Soit I une interrogative quelconque, et H un ensemble de quasi-expressions. On dira que l'interrogative I est, relativement à H :

- *sûre*, si et seulement si I est (logiquement) H -vraie, c'est-à-dire propositionnellement impliquée par H . Si $Prés(I)$ est définie, I est sûre relativement à H si et seulement si $Prés(I)$ est H -vraie. En d'autres termes, I est sûre relativement à H si et seulement si H a pour élément, ou implique propositionnellement $Prés(I)$.

- *risquée* sinon, c'est-à-dire si et seulement si I (ou $Prés(I)$ si elle est définie) n'est pas vraie dans toutes les H-interprétations.

- *stupide*, si et seulement si I est H-inconsistante, c'est-à-dire si et seulement si I n'est vraie dans aucune H-interprétation. Par conséquent, \bar{I} (ou : $Corr(I)$ si elle est définie) est propositionnellement impliquée par H . En d'autres termes, I est stupide relativement à H si et seulement si H a pour élément, ou implique propositionnellement, $Corr(I)$.

- *possible* sinon, c'est-à-dire si et seulement si I (ou $Prés(I)$ si elle est définie) n'est pas fausse dans toutes les H-interprétations.

Pour illustrer ces définitions, reprenons l'exemple :

$$I = \text{"Jean a-t-il cessé de battre sa femme ?"} \\ = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (B \ \& \ C , B \ \& \ \bar{C})$$

Cette question est à la fois *risquée* et *possible*. Mais si H a pour élément B , elle devient *sûre* relativement à H , alors que si H a pour élément \bar{B} , elle devient *stupide* relativement à H . De même, l'interrogative :

$$I = \text{"Quel est le plus grand nombre ?"} \\ = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} \forall \\ - \end{array} \right) (N_x // \forall y : x > y)$$

est *possible* absolument. Mais si H est l'ensemble des postulats usuels de l'arithmétique, I est à la fois *risquée* et *stupide* relativement à H .

Il existe une variété absolue des questions sûres et des questions stupides. Une interrogative I est dite *absolument sûre* si et seulement si elle est vraie dans toute interprétation. Tel est le cas en particulier des questions à réponse par oui ou non régulières ; ainsi :

$$I = \text{"Le verre est-il liquide à 20° centigrades ?"} \\ = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (L , \bar{L})$$

a une réponse vraie quel que soit l'état de l'univers [Belnap 1976 : 131].

La variété absolue des questions stupides se définit comme suit :
une interrogative I est une *interrogative muette* si et seulement si elle est fausse (c'est-à-dire : n'a pas de réponse vraie) dans toute interprétation. Par exemple :

$$I = \text{"Citez au moins trois propositions vraies parmi les suivantes :} \\ A, B \text{"} \\ = ? \left(\frac{1}{3} - - \right) (A, B)$$

est une interrogative muette en raison de la contradiction entre son sujet et sa requête. On peut d'ailleurs considérer que la question : "Citez un exemple de nombre dont le nom n'a jamais été, n'est pas, et ne sera jamais cité par qui que ce soit" est muette en ce sens qu'il existe certainement une réponse vraie, mais qu'il est impossible de l'énoncer. Enfin, une *question de HOBSON* conditionnelle universalisée de la forme :

$$I = \text{"Pour tout } C, \text{ si c'est un } P, \text{ dites-le moi"} \\ = \forall x (Cx // (Px / ? \left(\frac{1}{1} - - \right) (Px)))$$

est une question muette si la liste des C qui sont P est infinie [Belnap 1972 : 339].

L'examen des relations sémantiques entre un ensemble donné de quasi-expressions H, et les réponses directes d(I) à une interrogative I, conduit à distinguer d'autres types de questions. En particulier, on dira qu'une interrogative I est :

- *soluble par H*, si H implique propositionnellement au moins une réponse directe à I.

- *insoluble par H*, si H n'implique propositionnellement aucune réponse directe à I.

- *rhétorique* relativement à H, si $H \cup \{I\}$ implique propositionnellement au moins une réponse directe à I.

- *semi-rhétorique* relativement à H , si $H \cup \{I\}$ implique propositionnellement la négation d'au moins une réponse directe à I (ce type n'est pas défini explicitement par l'auteur : comparer [Belnap 1963 : 151] et [Belnap 1976 : 134-135] ; mais il comble une lacune et présente un intérêt pour l'analyse de la dernière catégorie ci-dessous).

- *discutable* relativement à H , si $H \cup \{I\}$ n'implique propositionnellement aucune réponse directe à I .

- *fortement discutable* relativement à H , si $H \cup \{I\}$ n'implique propositionnellement ni au moins une réponse directe à I , ni la négation d'au moins une réponse directe à I .

On voit par conséquent qu'une question est *discutable* si elle n'est pas rhétorique, c'est-à-dire si aucune réponse n'est garantie vraie a priori par $H \cup \{I\}$; et qu'une question est *fortement discutable* si elle n'est ni rhétorique, ni semi-rhétorique, c'est-à-dire si toutes les réponses directes possibles $d(I)$ et leur négation sont H -consistantes avec I .

Pour illustrer ces définitions, appelons H_q et H_r les ensembles de quasi-expressions correspondant à l'information dont disposent respectivement celui qui pose la question, et celui dont on attend une réponse. Dans une situation où l'on demande (par exemple) un renseignement, on peut supposer que le questionneur estime sa question *discutable* relativement à H_q , et *soluble* relativement à H_r . S'il s'est trompé sur ce dernier point, la personne interrogée pourra répondre un : "Je ne sais pas", signifiant : "Votre interrogative n'est pas soluble par l'ensemble des expressions représentant l'information dont je dispose actuellement" [Belnap 1976 : 135]. Dans une situation d'examen, la question posée par l'examineur est *rhétorique* relativement à H_q , et *soluble* ou *insoluble* par rapport à H_r selon que le candidat connaît ou ne connaît pas la réponse à la question posée (cf. : [Belnap 1963 : 151]).

En règle générale, il n'est pas possible de dire, à partir de la seule forme d'une interrogative, si celle-ci est rhétorique ou non. Il est vrai que nous pouvons dire par exemple que, d'après le théorème de GÖDEL, il ne peut exister de formalisation de l'arithmétique relativement à laquelle toute question arithmétique soit rhétorique. Toutefois, en pratique, le seul type de

question qui soit rhétorique de par sa forme est la *question de HOBSON*. Par exemple [Belnap 1963 : 152 ; 1976 : 135], les questions *whether* :

$$\left\{ \begin{array}{l} I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A) \\ I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} \vee \\ - \end{array} \right) (A) \\ I = ? \left(\begin{array}{c} - \\ 1 \end{array} \begin{array}{c} \vee \\ - \end{array} \right) (A) \end{array} \right.$$

sont rhétoriques ; de même que la question *which* :

$$I = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (Cx // Px)$$

dans le cas particulier où la classe de substitution associée à la condition catégorielle Cx ne contient qu'un seul nom.

Enfin, si l'on s'intéresse aux relations sémantiques entre les réponses directes à une même interrogation I , on dira qu'une interrogation I est, relativement à un ensemble H de quasi-expressions :

- *exclusive*, si et seulement si, dans toute H -interprétation, au plus une réponse directe à I est vraie. Par conséquent, la vérité d'une réponse directe donnée entraîne la fausseté de toutes les autres réponses directes possibles.

- *faiblement exclusive* (nominalement), si et seulement si, pour tout couple de réponses directes à I , ou bien elles sont logiquement H -équivalentes, ou bien elles sont H -inconsistantes.

- *indépendante* si et seulement si aucune réponse directe à I ne H -implique une autre réponse directe à I distincte.

- *minimale*, si et seulement si, pour toute expression A qui est une réponse directe à I , il existe une H -interprétation M dans laquelle A est la seule $d(I)$ qui soit vraie (toutes les autres réponses directes possibles étant fausses dans M relativement à H).

Le concept d'*interrogative exclusive* demande à être précisé. En particulier, une question *whether* sera dite exclusive si au plus une réponse

abstraite est vraie dans chaque H-interprétation ; elle sera dite *nominalement exclusive* si dans chaque H-interprétation au plus une réponse directe (nominale) est vraie. Ainsi, une interrogative *whether* demandant une réponse unique : $\rho = (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \vee -)$, est nominalement exclusive relativement à H ; si sa requête demande une énumération exhaustive : $\rho = (\begin{smallmatrix} \bar{1} \\ 1 \end{smallmatrix} \vee -)$, elle est *faiblement nominalement exclusive* relativement à H, puisque deux réponses directes peuvent être logiquement H-équivalentes.

Pour les questions *which*, le problème est plus complexe, et nécessite la définition de deux formes d'équivalence entre deux réponses réelles. Soient deux réponses réelles à une même interrogative *which* ; elles seront dites :

- *équivalentes par leurs éléments* si les ensembles de modalités apparaissant dans la sélection de chacune de ces réponses sont identiques.

- *équivalentes à une permutation près* si la sélection de l'une est une permutation de la sélection de l'autre.

La seconde définition est évidemment plus restrictive que la première. On dispose ainsi des moyens de qualifier plus précisément certaines formes d'interrogatives *which*. Une interrogative *which* demandant une réponse unique : $\rho = (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} \vee -)$ est exclusive relativement à H. Si elle demande une énumération exhaustive : $\rho = (\begin{smallmatrix} \bar{1} \\ 1 \end{smallmatrix} \vee -)$, elle est *exclusive au niveau des éléments* relativement à H ; en effet, dans aucune H-interprétation elle ne peut avoir deux réponses réelles équivalentes par leurs éléments. Si elle demande une énumération exhaustive et distincte : $\rho = (\begin{smallmatrix} \bar{1} \\ 1 \end{smallmatrix} \vee \neq)$, elle est *exclusive au niveau des permutations* relativement à H, puisque dans aucune H-interprétation elle ne peut avoir deux réponses réelles équivalentes à une permutation près.

e) Relations sémantiques entre questions

BELNAP étudie quatre formes principales de relations entre interrogatives. Soient I_1 et I_2 deux interrogatives quelconques, et H un ensemble de quasi-expressions ; on dira que, relativement à H :

- I_1 contient I_2 , si toute réponse directe H-consistante à I_1 est une réponse complète, relativement à H, à I_2 .

- I_1 est érotétiquement équivalent à I_2 si, pour toute réponse directe à I_1 , il existe une réponse directe à I_2 qui lui est H-équivalente, et si, pour toute réponse directe à I_2 , il existe une réponse directe à I_1 qui lui est H-équivalente ([Belnap 1976 : 137], à comparer avec : [Belnap 1963 : 154]).

- I_1 est érotétiquement pertinent à I_2 si et seulement si au moins une réponse directe à I_1 est pertinente à I_2 , relativement à H.

- I_1 annule I_2 si toute réponse directe à I_1 est, relativement à H, soit une réponse corrective, soit une réponse complète à I_2 .

La relation d'inclusion entre interrogatives est très différente de la relation ensembliste d'inclusion. Tout d'abord, remarquons que si l'ensemble des réponses à I_1 contient l'ensemble des réponses à I_2 , alors I_2 contient I_1 . Soit par exemple :

$$\begin{cases} I_1 = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A, B, C) \\ I_2 = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A, B) \end{cases}$$

On voit que : $\{A, B, C\}$ contient : $\{A, B\}$, ce qui a pour conséquence que I_2 contient I_1 . Remarquons également que la réciproque n'est pas nécessairement vraie. Soit par exemple :

$$\begin{cases} I_1 = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A \& B, C, D) \\ I_2 = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \begin{array}{c} - \\ - \end{array} \right) (A, C, D) \end{cases}$$

On voit que I_1 contient I_2 , puisque en particulier $d(I_1) = A \& B$ est une réponse complète à I_2 ($A \& B$ implique A). Par contre, l'ensemble des réponses à I_2 ne contient pas l'ensemble des réponses à I_1 .

La relation d'inclusion entre interrogatives est réflexive, non symétrique, et transitive. La réflexivité découle de la définition même : toute réponse

directe à I est également une réponse complète à I . La non symétrie est bien mise en évidence dans les deux exemples ci-dessus ; en effet :

$? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A, B, C)$ ne contient pas : $? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A, B)$, et :

$? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A, C, D)$ ne contient pas : $? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A \& B, C, D)$. La transitivité de l'inclusion entre interrogatives est une conséquence de la transitivité de l'implication entre réponses (qui sert de base à la définition des réponses complètes).

La *relation d'équivalence érotétique* est réflexive, symétrique, et transitive. Elle ne doit pas être confondue avec l'équivalence propositionnelle entre interrogatives. Par exemple, les deux interrogatives ci-dessous ont évidemment des contenus propositionnels équivalents :

$$\begin{cases} I_1 = ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (A, \bar{A}) \\ I_2 = ? \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix} - - \right) (B, \bar{B}) \end{cases}$$

puisque : $A \vee \bar{A} = B \vee \bar{B}$. Il est non moins évident qu'elles ne sont pas érotétiquement équivalentes, puisqu'aucune $d(I_1)$ n'est équivalente à une $d(I_2)$, et réciproquement. L'intérêt pratique de cette relation (par exemple dans la construction d'un questionnaire) est la possibilité de remplacer une interrogative donnée par une autre interrogative érotétiquement équivalente.

Bien entendu, il serait d'un plus grand intérêt pratique encore d'étendre cette relation d'équivalence érotétique aux ensembles d'interrogatives. BELNAP amorce cette extension avec la notion de partition directe. Une série ordonnée d'interrogatives (I_1, \dots, I_n) est une *partition directe* d'une interrogative I si la conjonction $I_1 \& \dots \& I_n$ est érotétiquement équivalente à I . La série (I_1, \dots, I_n) est une *partition quasi-directe* si la conjonction $I_1 \& \dots \& I_n$ est érotétiquement équivalente à $I \& \dots \& I$ (n fois).

La *relation de pertinence* entre interrogatives est réflexive, symétrique, et non transitive. La réflexivité est aisée à démontrer : toute $d(I)$ est toujours une réponse complète à I , et donc érotétiquement pertinente à I . La symétrie est une conséquence de la relation de pertinence entre une réponse et une question : est pertinente à I toute réponse $r(I)$ qui soit implique,

soit est impliquée par une $d(I)$ ou sa négation. Toute relation de ce type entre une $r(I_1)$ et une $d(I_2)$ entraîne une relation symétrique entre I_1 et I_2 . La non transitivité peut être illustrée par l'exemple suivant ; soit [Belnap 1963 : 156] :

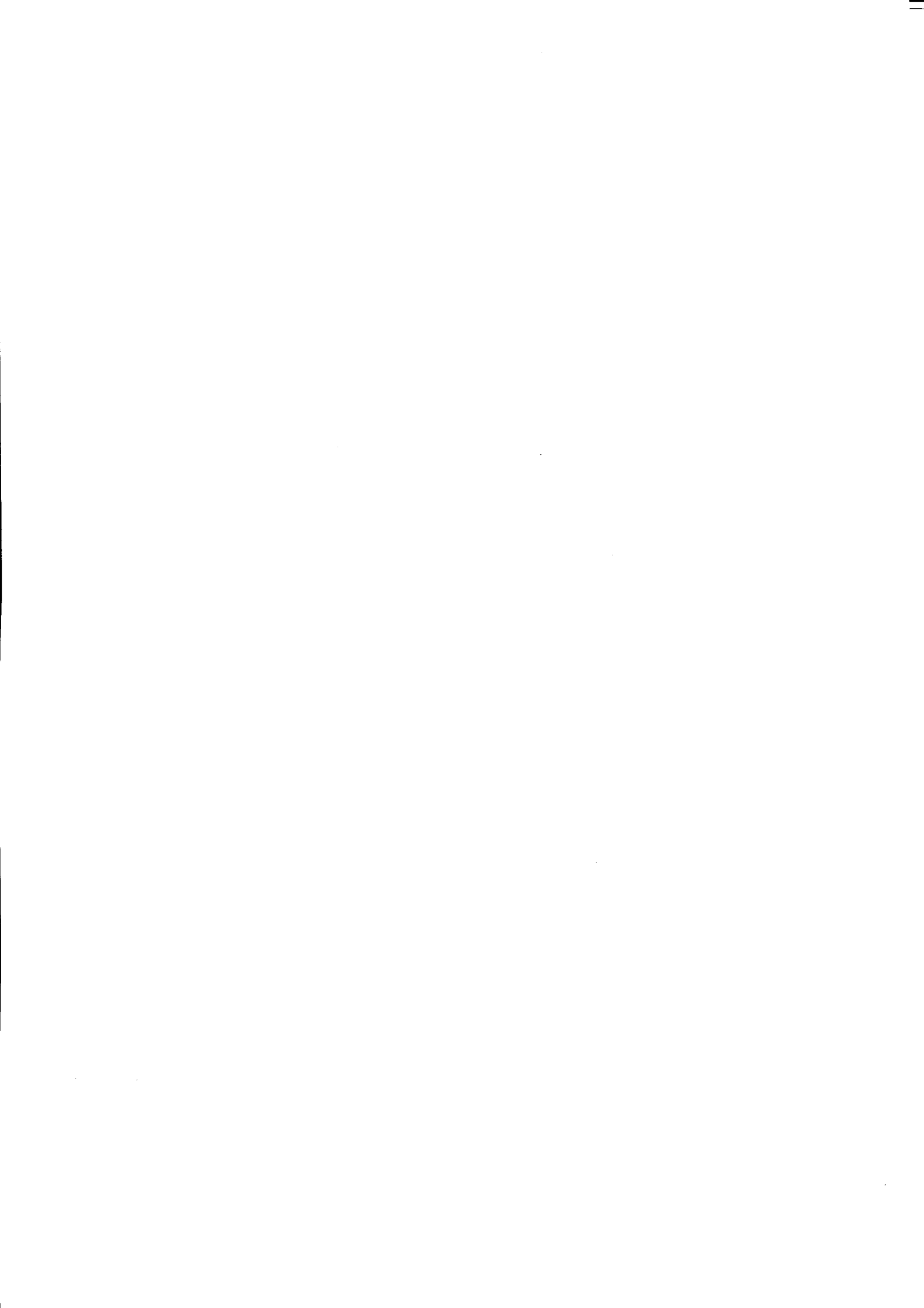
$$\left\{ \begin{array}{l} I_1 = ? \left(\begin{array}{l} 1 \\ 1 \end{array} \vee - \right) (A, B) \\ I_2 = ? \left(\begin{array}{l} 1 \\ 1 \end{array} \vee - \right) (B, C) \\ I_3 = ? \left(\begin{array}{l} 1 \\ 1 \end{array} \vee - \right) (C, D) \end{array} \right.$$

I_1 est pertinente à I_2 , I_2 à I_3 , mais I_1 n'est pas pertinente à I_3 .

La relation d'*annulation* d'une interrogative I_2 par une autre interrogative I_1 traduit le fait qu'obtenir une réponse à I_1 rend inutile le fait de poser I_2 . En effet, si $d(I_1)$ est une réponse complète à I_2 , elle implique une $d(I_2)$, et apporte par conséquent une réponse pertinente à I_2 ; si $d(I_1) = Corr(I_2)$, elle apprend au locuteur que $Prés(I_2)$ est fausse, et que I_2 n'a pas de réponse vraie. Dans les deux cas, il n'y a pas lieu de poser I_2 . Soit par exemple [Belnap 1976 : 138] :

$$\left\{ \begin{array}{l} I_1 = \text{"Est-ce la fin du paragraphe, ou doit-il y avoir encore un} \\ \quad \text{dernier alinéa ?"} \\ I_2 = \text{"Combien doit-il y avoir d'autres alinéas ensuite ?"} \end{array} \right.$$

I_1 annule I_2 . Remarquons sur cet exemple qu'il n'est pas nécessaire, pour que I_1 annule I_2 , que I_1 contienne I_2 (cf. [Belnap 1963 : 157]).



2.3. BRÈVE PRÉSENTATION D'AUTRES SYSTÈMES DE LOGIQUE ÉROTÉTIQUE

Il aurait été nécessaire de présenter les théories de Lennart ÅQVIST et de Tadeusz KUBIŃSKI avec au moins autant de détails que celle de Nuel D. BELNAP. Cela aurait sans doute entraîné quelques redites, mais les apports de ces auteurs à la logique des questions méritaient mieux qu'une trentaine de pages chacun. Le souci de brièveté nous a conduit à faire une sélection dans leurs travaux en fonction de nos propres objectifs ; la sélection d'un logicien aurait certainement été toute autre. La présentation que nous faisons du résultat des recherches d'ÅQVIST (§ 2.3.1) et de KUBIŃSKI (§ 2.3.2) ne saurait par conséquent donner une idée véritable de l'ampleur et de la rigueur de leurs théories. Par contre, elle devrait suffire à montrer l'intérêt qu'il y a, pour le psychologue ou le sociologue, à se familiariser avec elles.

Dans la perspective d'une application de la logique des questions aux questionnaires sociologiques et aux tests psychologiques, nous avons accordé ensuite une place relativement importante (un peu plus d'une dizaine de pages) à la théorie de Sylvain BROMBERGER sur les questions de causalité, problème encore mal résolu mais d'une grande importance théorique et pratique (§ 2.3.3). Enfin, nous avons mentionné certains travaux qui, dans l'optique de notre recherche, nous ont paru d'une importance moindre, mais non négligeable (§ 2.3.4).

2.3.1. La théorie logique des interrogatives de Lennart ÅQVIST

La manière dont ÅQVIST pose le problème d'une logique érotétique est très originale par rapport aux approches de HARRAH ou de BELNAP, qui lui sont antérieures, et dont il a eu connaissance. En effet, sans méconnaître l'importance de l'analyse des relations entre une question et l'ensemble de ses réponses possibles, ÅQVIST estime que "les relations logiques entre questions peuvent être étudiées sans que l'on ait besoin au préalable de définir la nature des relations question-réponse" [Åqvist 1965 : 11]. L'auteur considère qu'une question équivaut à une *impérative épistémique*, c'est-à-dire à une phrase composée de deux éléments dont l'un (épistémique) exprime une ignorance du locuteur, et l'autre (impératif) traduit la demande du locuteur que cette ignorance soit levée. Remarquons la parenté entre cette conception et celles par exemple du linguiste Otto JESPERSEN ("une

interrogation est aussi une sorte d'ordre, c'est-à-dire l'ordre de dire quelque chose à celui qui pose la question, de lui fournir l'information qu'il réclame" [Jespersen 1924 : 431]), du cybernéticien Donald M. MACKAY ("On peut considérer les questions comme des requêtes (ou des ordres selon leur forme ou leur intonation) demandant l'élaboration d'un instrument linguistique ayant pour fonction la mise à jour [du système d'information du locuteur]" [Mackay 1969 : 130]), ou même de logiciens tels que Richard Mervyn HARE (pour qui les questions sont des "crypto-impératives", qui "peuvent être traduites en ordres sans aucune déperdition de signification" [Hare 1949 : 24]). Toutefois, la théorie d'ÅQVIST est à notre connaissance le seul développement systématique de cette conception des interrogatives.

Le langage formel utilisé par l'auteur s'inspire à la fois de la logique déontique (cf. par exemple : [Kalinowski 1972]), domaine au développement duquel il a lui-même apporté une contribution, et de la logique épistémique développée par Jaakko HINTIKKA. Ce langage est le calcul appliqué des prédicats du premier ordre avec identité, muni en outre d'opérateurs épistémiques et d'opérateurs déontiques. Parmi ceux-ci, les plus importants sont :

- l'opérateur déontique ! , que l'on pourrait traduire par : "Faites en sorte (maintenant) que". Si P est une expression (purement assertive), ! P est une expression du langage formel considéré, et se lit : "Faites en sorte (maintenant) que P".

- l'opérateur épistémique K , que l'on peut paraphraser en "Je sais (maintenant) que". Si P est une expression (assertive), KP est une expression du langage, et se lit : "Je sais (maintenant) que P".

- l'opérateur épistémique K# , qui se lit : "Vous savez (maintenant) que" (K#P = "Vous savez (maintenant) que P").

La théorie d'ÅQVIST est développée à deux niveaux. Le niveau le plus élémentaire correspond au calcul des propositions : c'est le système PIE (*propositional imperative-epistemic logic*), qui permet de rendre compte des questions *whether*. Le niveau le plus élaboré est celui du système QIE (*quantified imperative-epistemic logic*), qui correspond au calcul des prédicats avec variables libres et quantificateurs. Ce système sert à formaliser les questions *what*

(qui, dans la terminologie d'ÅQVIST, correspondent aux questions *which* de BELNAP), auxquelles d'ailleurs peuvent être réduites les questions *whether* (cf. § 2.3.1.c). Tel qu'il est décrit par l'auteur, ce langage ne permet pas de rendre compte de certaines formes de questions, en particulier de celles qui correspondent à des énoncés introduits par *pourquoi*, *comment*, *qu'est-ce que* (définition abstraite), etc. Cette limitation correspond (comme chez BELNAP) aux limites du calcul des prédicats du premier ordre. C'est pourquoi l'auteur envisage une extension ultérieure du système QIE en vue de traiter également des questions de causalité ou de définition [Åqvist 1965 : 113-117]. Par contre, la théorie d'ÅQVIST pourrait s'appliquer à certaines formes d'énoncés interrogatifs, comme les pseudo-questions à valeur d'interdiction du genre : "Vous n'allez pas faire cela ?" (cf. [Åqvist 1965 : 11]).

Comparée à la logique des questions et des réponses de BELNAP, la théorie des interrogatives d'ÅQVIST apparaît comme un système d'une importance comparable, et de même ampleur. Dans les deux cas, les champs d'application semblent très étendus, et les fondements théoriques solidement assurés. Par contre, la différence de style entre les auteurs est particulièrement sensible. L'oeuvre de BELNAP apparaît plutôt comme une construction *ad hoc*, élaborée pour rendre compte d'un ensemble de questions aussi variées que possible dont l'auteur serait parti pour édifier sa théorie. La théorie d'ÅQVIST donne l'impression au contraire d'être avant tout une construction abstraite, basée sur l'intuition qu'une logique érotétique pouvait naître d'une combinaison de la logique déontique et de la logique épistémique. D'où une attention particulière apportée ensuite au codage, dans le langage formel, d'un énoncé interrogatif donné, et aux mérites (et inconvénients) respectifs des diverses formalisations possibles d'une même question. La présentation que donne ÅQVIST de sa logique érotétique est certainement plus complexe et plus difficile d'accès que l'oeuvre de BELNAP ; mais la contrepartie de cette complexité est une plus grande latitude dans l'interprétation sémantique des énoncés, offrant une gamme de nuances beaucoup plus étendue que celles que peuvent offrir non seulement les langues naturelles, mais aussi les autres systèmes formels. Pour conclure ce bref parallèle entre BELNAP et ÅQVIST, nous reprendrons à notre compte le jugement de HARRAH, qui "suggère que toute personne intéressée par les questions étudie l'oeuvre d'ÅQVIST et également l'oeuvre de BELNAP à laquelle il fait référence, et qu'elle attende pour se

prononcer sur les mérites respectifs de ces deux approches que celles-ci aient été développées plus avant" [Harrah 1967 : 404]. Il ne semble pas malheureusement qu'ÅQVIST ait mené à bien la rédaction de la seconde partie de sa recherche, exposant en détail la logique des interrogatives dont [Åqvist 1965] n'est que la présentation analytique. En outre, l'auteur a esquissé (dans [Åqvist 1971]) une révision des fondements de son système de logique déontique-épistémique, afin d'éliminer quelques imprécisions (dans l'usage des quantificateurs) et de corriger certains paradoxes liés à l'opérateur épistémique et aux questions conditionnelles ; mais il n'a pas (à notre connaissance) présenté un exposé complet de son nouveau système appliqué aux interrogatives.

Dans cette brève présentation de l'oeuvre d'ÅQVIST, nous nous fondons essentiellement sur l'exposé général de [Åqvist 1965]. Nous traitons d'abord successivement des questions *whether*, et des questions *what* demandant une énumération exhaustive. Cela nous donne l'occasion d'étudier diverses formalisations possibles d'une même question ; de présenter les notions de *noyau* d'une question, et de présupposition ; de décrire diverses procédures d'*assurance* d'une question ; et enfin d'introduire les questions conditionnelles et les questions mixtes *whether-what*. Nous abordons ensuite plus succinctement le problème des autres formes de questions *what*. Nous terminons cette présentation par l'analyse des propriétés sémantiques des réponses aux questions.

a) Les questions *whether*

En pratique, la manière dont ÅQVIST traite les questions *whether* ne s'écarte pas beaucoup de celle de BELNAP. Toutefois, la manière de poser le problème est évidemment très différente. Soit par exemple l'interrogative :

I = "Est-ce Brutus ou Cassius qui a tué César ?"

Elle équivaut à l'affirmation d'une ignorance limitée, que l'on peut exprimer par : "Je sais que c'est Brutus ou Cassius qui a tué César, mais je ne sais pas lequel des deux", affirmation assortie d'une demande signifiant : "Faites en sorte que je le sache". Cette forme d'interrogative est justiciable du système PIE. Soient en effet les propositions :

B = "C'est Brutus qui a tué César"

C = "C'est Cassius qui a tué César"

elles constituent les deux modalités de réponse présentées par la question. Cependant, l'ambiguïté du langage rend possibles trois interprétations, selon que l'on considère l'interrogative comme demandant un exemple unique, une réponse vraie unique, ou une énumération exhaustive.

Dans le premier cas (exemple unique), on peut paraphraser la question par : "Quelle est *au moins* une modalité de réponse vraie parmi les suivantes : B, C ?". La formalisation dans le système PIE devient :

$$I = ! (KB \vee KC)$$

que l'on peut lire : "Faites en sorte que je sache que c'est Brutus qui a tué César ou que je sache que c'est Cassius qui a tué César" (même si en réalité ce sont les deux). On abrège cette formalisation en écrivant :

$$I = ?_2 (B,C)$$

On appelle *noyau* de la question son seul contenu épistémique, c'est-à-dire l'expression obtenue en supprimant de sa forme symbolique développée l'opérateur déontique :

$$\text{Noyau}(I) = KB \vee KC$$

On appelle *présupposition* d'une question I une proposition purement assertive P telle que, de toutes les propositions assertives impliquées par I (les "présuppositions" au sens large), P soit la plus forte, c'est-à-dire celle qui implique toutes les autres sans être impliquée par aucune d'entre elles [Åqvist 1965 : 132]. On obtient la présupposition *Prés*(I) en supprimant de la forme symbolique développée de I tout opérateur épistémique ou déontique :

$$\text{Prés}(I) = B \vee C$$

On retrouve ainsi le contenu propositionnel de la partie positive de l'affirmation d'ignorance limitée citée plus haut (symbolisable en : $K(B \vee C) \& (\overline{KB \vee KC})$).

La seconde interprétation (réponse vraie unique) équivaut à la paraphrase : "Quelle est l'*unique* modalité vraie parmi les suivantes : B,C ?", symbolisable en :

$$\begin{aligned} I &= ?_2^u (B,C) \\ &= ! (K(B \& \overline{C}) \vee K(\overline{B} \& C)) \end{aligned}$$

Remarquons l'identité :

$$?_2^u (B, C) = ?_2 (B \& \bar{C}, \bar{B} \& C)$$

Le noyau de la question est ici :

$$\text{Noyau}(I) = K(B \& \bar{C}) \vee K(\bar{B} \& C)$$

La présupposition :

$$\text{Prés}(I) = (B \& \bar{C}) \vee (\bar{B} \& C)$$

exprime l'idée que Brutus et Cassius ne peuvent être simultanément meurtriers de César.

La troisième interprétation (énumération exhaustive) correspond à : "Quelles sont *toutes* les modalités de réponse vraies, *et elles seules*, parmi les suivantes : B, C ?", formalisable en :

$$\begin{aligned} I &= ?_2^c (B, C) \\ &= ! (K(B \& \bar{C}) \vee K(\bar{B} \& C) \vee K(B \& C)) \\ &= ?_3 (B \& \bar{C}, \bar{B} \& C, B \& C) \end{aligned}$$

Au noyau :

$$\text{Noyau}(I) = K(B \& \bar{C}) \vee K(\bar{B} \& C) \vee K(B \& C)$$

correspond la présupposition :

$$\begin{aligned} \text{Prés}(I) &= (B \& \bar{C}) \vee (\bar{B} \& C) \vee (B \& C) \\ &= B \vee C \end{aligned}$$

Dans ses préoccupations, ÅQVIST accorde une place relativement importante à l'*assurance* des questions. Il présente trois méthodes permettant d'assurer une question *whether* risquée, de sorte qu'elle ait toujours au moins une réponse vraie. Soit par exemple la question risquée déjà analysée (§ 2.2.5.b) :

$$\begin{aligned} I &= \text{"Jean a-t-il cessé de battre sa femme ?"} \\ &= ?_2 (B \& C, B \& \bar{C}) \\ &= ! (K(B \& C) \vee K(B \& \bar{C})) \end{aligned}$$

Le piège de cette question est que, quoique risquée, elle a l'apparence d'une question sûre, puisque toute question dichotomique (à réponse par *oui* ou *non*) régulière est sûre. En réalité, l'alternative n'est pas entre deux contradictoires ($B \& C$, $\overline{B \& C}$), mais entre deux contraires ($B \& C$, $B \& \overline{C}$). On pourrait rendre cette question sûre en modifiant l'énoncé :

$$\begin{aligned} I' &= \text{"Est-il vrai que Jean a une femme, qu'il l'a battue, et qu'il a} \\ &\quad \text{cessé de la battre ?"} \\ &= ?_2 (B \& C, \overline{B \& C}) \\ &= ?_1 (B \& C) \end{aligned}$$

(notation abrégée pour les questions à réponse *oui-non*). Mais en agissant ainsi, on modifie le sens de la question, puisque l'on altère la seconde modalité de réponse ($B \& \overline{C}$ devient : $\overline{B \& C}$). Une bonne méthode d'assurance d'une question doit avant tout ne modifier aucune des modalités de réponse de la question d'origine.

La première méthode d'assurance, dite *méthode whether-whether*, consiste à ajouter aux modalités existantes une modalité supplémentaire, qui est la négation de la disjonction des modalités d'origine. Dans notre exemple, on ajoutera : $((\overline{B \& C}) \vee (B \& \overline{C})) = \overline{B}$, ce qui donnera :

$$\begin{aligned} I' &= \text{"Jean a-t-il cessé de battre sa femme, ou bien est-ce qu'il n'a pas} \\ &\quad \text{de femme qu'il ait battue ?"} \\ &= ?_3 (\overline{B}, B \& \overline{C}, B \& C) \\ &= ! (K\overline{B} \vee K(B \& \overline{C}) \vee K(B \& C)) \end{aligned}$$

La procédure générale, étant donnée une question *whether* risquée I , de la forme :

$$I = ?_n (P_1, \dots, P_n)$$

est de la remplacer par une question I' , de la forme :

$$I' = ?_{n+1} ((\overline{P_1 \vee \dots \vee P_n}), P_1, \dots, P_n)$$

qui est toujours sûre.

La seconde méthode, dite *méthode de WHATELY-PRIOR*, consiste à décomposer la question en deux questions successives, dont la seconde est conditionnelle :

I_1 = "Est-il vrai que Jean a une femme et qu'il avait l'habitude de la battre ?"

$$= ?_2 (B, \bar{B})$$

$$= ?_1 (B)$$

I_2 = Si oui : "A-t-il cessé de la battre ?"

La question conditionnelle I_2 s'écrit :

$$I_2 = ?_{2/1} (B \& C, B \& \bar{C} / B)$$

et la question I d'origine est remplacée par la conjonction :

$$I' = ?_1 (B) \quad \& \quad ?_{2/1} (B \& C, B \& \bar{C} / B)$$

La règle générale étant de remplacer la question I par :

$$I' = ?_1 (P_1 \vee \dots \vee P_n) \quad \& \quad ?_{n/1} (P_1, \dots, P_n / P_1 \vee \dots \vee P_n)$$

La troisième méthode, dite *méthode whether-if*, se limite à rendre conditionnelle la question originale, en posant :

I' = "S'il est vrai que Jean a une femme et qu'il avait l'habitude de la battre, a-t-il cessé de le faire ?"

$$= ?_{2/1} (B \& C, B \& \bar{C} / B)$$

Soit dans le cas général :

$$I' = ?_{n/1} (P_1, \dots, P_n / P_1 \vee \dots \vee P_n)$$

Cette forme est sûre, et n'altère aucune des modalités de réponses. La réponse : "Il n'est pas marié et n'a donc pas l'habitude de battre sa femme" = \bar{B} est une pseudo-réponse directe à I' (cf. § 2.3.1.d).

Cette formalisation des questions *whether* conduit à certains "paradoxes des questions conditionnelles" ; par exemple [Åqvist 1971 : 53], les expressions suivantes sont valides :

- $\bar{Q} \rightarrow ?_{n/1} (P_1, \dots, P_n / Q)$
- $!\bar{Q} \rightarrow ?_{n/1} (P_1, \dots, P_n / Q)$
- $?_n (P_1, \dots, P_n) \rightarrow ?_{n/1} (P_1, \dots, P_n / Q)$

ce qui constitue plutôt un inconvénient pour une logique des questions. C'est pourquoi dans sa révision de 1971, l'auteur propose de considérer la question conditionnelle comme la forme de base de toute question, et d'en dériver la question inconditionnelle en posant :

$$?_n (P_1, \dots, P_n) = ! (KP_1 \vee \dots \vee KP_n / P_1 \vee \dots \vee P_n) \ \& \ (P_1 \vee \dots \vee P_n)$$

En d'autres termes, l'opérateur déontique devient un opérateur à deux arguments : $!(P / Q)$ = "Faites en sorte que P dans la conjoncture que Q", et poser une question consiste *aussi* à affirmer la vérité de sa présupposition.

b) Les questions *what* demandant une énumération exhaustive

Bien que les questions *what* d'ÅQVIST correspondent aux questions *which* de BELNAP (en ce sens que ce sont en gros les mêmes énoncés qui sont formalisables sous ces formes abstraites), les principes de leur formalisation et, partant, leurs variétés, sont très différents de ceux des questions *which*. Trois dimensions permettent de classer les questions *what* : leur nombre de variables libres distinctes, leur quantificateur, et le fait que les variables libres soient ou non soumises à une condition de catégorie [Åqvist 1965 : 83-84, 108]. Ces dimensions se retrouvent également chez BELNAP. Dans ce paragraphe, nous ne traiterons que des questions *what* à une seule variable libre, demandant une énumération exhaustive, et dont la variable est indépendante de toute condition catégorielle. Au paragraphe suivant, nous aborderons les questions *what* demandant au moins n réponses vraies, et celles qui demandent exactement n réponses vraies ; nous traiterons également des questions *what* soumises à des conditions catégorielles, mais en nous limitant toujours aux questions à une seule variable libre.

Soit la question : "Qui sont les membres de l'Orchestre Philharmonique de Vienne ?". Cette question est susceptible, dans le système QIE, de quatre

traductions distinctes, selon que celui qui pose la question requiert ou non l'exhaustivité (K), et selon qu'il présume ou non l'existence d'au moins un membre de l'Orchestre (*présomption de non-vacuité* E). A partir d'une forme de base (B) pour laquelle ni l'exhaustivité, ni la non-vacuité ne sont requises, on peut engendrer les trois autres formes (KB, EB, EKB) de questions *what* demandant une énumération exhaustive. Ces quatre formes sont liées entre elles par des relations d'implication (dans le système QIE) représentées sur la figure n° 2.7. ci-après. Nous allons analyser chacune de ces formes, en posant : $Ox = "x \text{ est un membre de l'Orchestre Philharmonique de Vienne}"$.

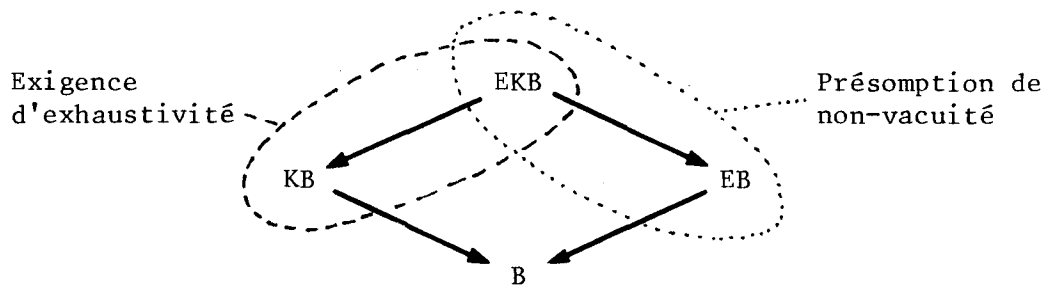


Figure n° 2.7 : Relations d'implication entre quantificateurs d'une question *what* demandant une énumération exhaustive [Åqvist 1965 : 93].

La *forme B* est évidemment la plus simple. L'énoncé de la question exprime au minimum la requête suivante : "Que pour tout membre de l'Orchestre, je sache qui il est et aussi qu'il est membre de l'Orchestre". Dans le système QIE, "Je sais qui est x " s'écrit : $\exists y, K(y = x)$ (cf. [Åqvist 1965 : 27] à ne pas confondre avec : $K(\exists y(y = x)) = "je \text{ sais que } x \text{ existe}"$ [Åqvist 1965 : 149]). Mais comme les règles s'appliquant à l'opérateur épistémique K imposent que toute variable liée à cet opérateur (ici : x et y) soit connue du sujet implicite de l'opérateur (ici, le locuteur), nous devons utiliser une formalisation un peu plus complexe ; la question devient donc [Åqvist 1965 : 86-87] :

$$I = ! (\forall x (Ox \rightarrow \exists y ((y = x) \& KOy)))$$

$$= (? \underset{B}{I} x) Ox$$

On vérifierait aisément que le noyau de la question ainsi formalisée n'implique aucunement l'existence d'au moins un x membre de l'Orchestre. D'autre part, lorsque le locuteur dispose de l'énumération exhaustive des membres de l'Orchestre, rien ne lui indique en principe qu'il possède la réponse attendue. Il peut

se trouver par conséquent que, de tout membre de l'Orchestre, le locuteur sache à la fois qu'il est et qu'il est membre de l'Orchestre, sans savoir toutefois que la totalité des membres de l'Orchestre a été passée en revue.

La forme *KB* permet d'échapper à cette dernière objection. Elle peut être paraphrasée par : "Faites en sorte que je sache que : pour tout membre de l'Orchestre, je sais qu'il est et aussi qu'il est membre de l'Orchestre"

[Aqvist 1965 : 90]. Nous avons alors :

$$\begin{aligned} I &= ! K (\forall x (Ox \rightarrow \exists y ((y = x) \& KOy))) \\ &= (? \overset{I}{KB} x) Ox \end{aligned}$$

Comme pour la forme *B*, le noyau de la forme *KB* n'implique nullement l'existence d'au moins un x membre de l'Orchestre. Par contre, elle impose, en plus des informations requises par la forme *B*, que le locuteur sache qu'il dispose effectivement de l'énumération exhaustive recherchée.

La forme *EB* permet de présumer l'existence d'au moins un x membre de l'Orchestre, mais sans exiger de savoir que l'on dispose de l'information demandée. Elle ajoute à la forme *B* la condition : $\exists x, KOx$ = "Qu'il existe au moins un individu dont je sache qu'il est membre de l'Orchestre" ; cette condition entraîne la *présomption de non-vacuité* : $\exists x, Ox$. Soit :

$$\begin{aligned} I &= ! (\exists x(KOx) \& \forall x(Ox \rightarrow \exists y ((y = x) \& KOy))) \\ &= (? \overset{I}{EB} x) Ox \end{aligned}$$

Remarquons d'ailleurs que l'on obtiendrait des résultats équivalents en pratique à la formalisation ci-dessus en exprimant la présomption de non-vacuité de façon moins contraignante, c'est-à-dire en remplaçant $\exists x, KOx$ par : $K(\exists x, Ox)$, ou même simplement par : $\exists x, Ox$ (cf. [Aqvist 1965 : 93]). On obtient alors :

$$I' = (\exists x, Ox) \& (? \overset{I}{B} x) Ox .$$

La forme *EKB*, la plus complète, se déduit aisément des formes précédentes. On a par conséquent :

$$\begin{aligned} I &= ! (\exists x(KOx) \& K(\forall x (Ox \rightarrow \exists y((y = x) \& KOy)))) \\ &= (? \overset{I}{EKB} x) Ox \end{aligned}$$

On peut également constituer la forme plus faible, mais pratiquement équivalente quant aux résultats :

$$I' = (\exists x, O_x) \ \& \ (\overset{1}{KB} x) O_x.$$

Il est intéressant de noter que la forme B peut être remplacée par une forme EB conditionnelle qui lui est équivalente (dans le système QIE). En effet, la forme :

$$I = (\overset{1}{B} x) O_x$$

équivalait à la forme conditionnelle [Åqvist 1965 : 125-126] :

$$\begin{aligned} I' &= (\overset{1}{EB} x / 1) (O_x / \exists x, O_x) \\ &= ! ((\exists x, O_x) \rightarrow \text{Noyau} (\overset{1}{EB} x) O_x) \\ &= ! ((\exists x, O_x) \rightarrow (\exists x(KO_x) \ \& \ \forall x(O_x \rightarrow \exists y((y = x) \ \& \ KO_y)))) \end{aligned}$$

que l'on peut paraphraser en disant : "Faites en sorte que, s'il existe au moins un individu qui soit membre de l'Orchestre, alors : il existe au moins un individu dont je sache qu'il est membre de l'Orchestre, et pour tout membre de l'Orchestre je sache qui il est et aussi qu'il est membre de l'Orchestre". A cette forme, on peut préférer une forme plus forte puisque demandant en outre l'assurance d'exhaustivité :

$$I'' = (\overset{1}{EKB} x / 1) (O_x / \exists x, O_x)$$

De même, la forme KB peut être remplacée par une forme mixte *whether-what* qui lui est équivalente (dans le système QIE). Soit comme précédemment :

$$I = (\overset{1}{KB} x) O_x$$

A cette question *what* on peut associer (par une relation de disjonction) une question *whether* signifiant : "Faites en sorte que je sache s'il existe ou non au moins un individu qui soit membre de l'Orchestre". On obtient alors :

$$\begin{aligned}
I' &= (? \frac{1}{KB} x ; +1) (Ox ; (\overline{\exists x, Ox})) \\
&= ! (Noyau (? \frac{1}{KB} x) Ox \vee K (\overline{\exists x, Ox})) \\
&= ! (K(\forall x(Ox \rightarrow \exists y((y = x) \& KOy))) \vee K (\overline{\exists x, Ox}))
\end{aligned}$$

Cette forme mixte signifie à peu près : "Faites en sorte que je sache que : pour tout membre de l'Orchestre, je sais qui il est et aussi qu'il est membre de l'Orchestre ; ou bien que je sache que personne n'est membre de l'Orchestre". On peut à cette formalisation préférer la forme plus forte :

$$I'' = (? \frac{1}{EKB} x ; +1) (Ox ; (\overline{\exists x, Ox}))$$

qui demande l'affirmation d'exhaustivité.

Selon l'auteur, ces formes équivalentes constituent en quelque sorte la forme fondamentale des questions *what* de type B ou KB : "Dans notre discussion des questions *what* demandant une énumération exhaustive, nous avons affirmé que le quantificateur $(? \frac{1}{KB} a)$ est en fait un quantificateur formant des questions mixtes *whether-what*, et que $(? \frac{1}{B} a)$ est en fait un quantificateur formant des questions *what* conditionnelles ("assurées conditionnellement)" [Åqvist 1965 : 125]. Toutefois, il existe une différence importante entre la forme originale et la forme jugée équivalente (et fondamentale) : la première est risquée, alors que la seconde est sûre.

L'assurance des questions *what* suit des voies parallèles à celles des questions *whether* [Åqvist 1965 : 140]. Soit une question risquée $I = (? \frac{1}{EKB} x) Ox$. Sa présupposition $Prés(I)$ s'obtient, comme pour les questions *whether*, en supprimant de sa forme développée tout opérateur épistémique ou déontique. Nous avons par conséquent :

$$\begin{aligned}
I &= ! (\exists x(KOx) \& K(\forall x(Ox \rightarrow \exists y((y = x) \& KOy)))) \\
Prés(I) &= (\exists x, Ox) \& \forall x(Ox \rightarrow \exists y((y = x) \& Oy))
\end{aligned}$$

Les trois méthodes d'assurance des questions *what* correspondant à celles décrites au paragraphe précédent pour les questions *whether* sont :

- la méthode *whether-what*, qui consiste à remplacer la question *what* par une question mixte de la forme :

$$I' = (? \overset{1}{\text{EKB}} x ; +1) (Ox ; \overline{\text{Prés}(I)})$$

- la *méthode de WHATELY-PRIOR*, qui a pour résultat la conjonction de deux questions, la première, inconditionnelle, portant sur la présupposition de I , et la seconde étant la question I rendue conditionnelle :

$$I' = ?_2 (\text{Prés}(I) , \overline{\text{Prés}(I)}) \& (? \overset{1}{\text{EKB}} x / 1) (Ox / \text{Prés}(I))$$

- la *méthode what-if* enfin, qui rend la question I conditionnelle :

$$I' = (? \overset{1}{\text{EKB}} x / 1) (Ox / \text{Prés}(I))$$

Ces trois méthodes s'appliquent à toutes les formes de questions *what*, y compris celles qui dépendent de conditions catégorielles (cf. § 2.3.1.c).

c) Autres types de questions what

Parmi les questions *what* non soumises à des conditions catégorielles, ÅQVIST traite également des questions demandant *au moins n réponses vraies*, et de celles qui demandent *exactement n réponses vraies*. La classification des questions *what* est par conséquent très différente de celle des questions *whether*; aussi, les types de questions *what* d'ÅQVIST ne recourent-ils pas les formes distinguées de questions *which* de BELNAP.

Comme exemple de question demandant *au moins n réponses vraies*, l'auteur propose [Åqvist 1965 : 95-96] : "Citez au moins trois nombres premiers compris entre 10 et 20". La formalisation de cet énoncé dans le système QIE donne :

$$\begin{aligned} I &= ! (\exists x, \exists y, \exists z, K(Px \& Py \& Pz \& x \neq y \& x \neq z \& y \neq z)) \\ &= (? \overset{1}{C_3} x) Px \end{aligned}$$

avec : $Px = "x$ est un nombre premier compris entre 10 et 20". La forme développée de la question montre qu'elle présuppose l'existence d'au moins trois entités distinctes satisfaisant P :

$$\text{Prés}(I) = \exists x, \exists y, \exists z (Px \& Py \& Pz \& x \neq y \& x \neq z \& y \neq z)$$

La forme générale de ce type de question est :

$$\begin{aligned}
 (? \underset{C}{C}_n^1 x) Px = & ! (\exists x_1, \dots, \exists x_n, K(Px_1 \& \dots \& Px_n \& x_1 \neq x_2 \& \dots \\
 & \& x_1 \neq x_n \& x_2 \neq x_3 \& \dots \& x_{n-1} \neq x_n))
 \end{aligned}$$

Lorsque $n = 1$, on retrouve la question *which* demandant un exemple unique ; lorsque $n > 1$, on peut assimiler ce type de question *what* à une question *which* demandant des exemples multiples et distincts (avec $v = n$, et : $u = -$).

Un énoncé assez proche demandant *exactement n réponses vraies* serait : "Citez exactement trois nombres premiers compris entre 10 et 20", ou mieux : "Quels sont les trois nombres premiers compris entre 10 et 20 ?". Soit :

$$\begin{aligned}
 I = & ! (\exists x, \exists y, \exists z, K(Px \& Py \& Pz \& x \neq y \& x \neq z \\
 & \& y \neq z \& \forall w(Pw \rightarrow (w = x \vee w = y \vee w = z)))) \\
 = & (? \underset{D}{D}_3^1 x) Px
 \end{aligned}$$

La présupposition de la question :

$$\begin{aligned}
 Prés(I) = & \exists x, \exists y, \exists z (Px \& Py \& Pz \& x \neq y \& x \neq z \& y \neq z \\
 & \& \forall w(Pw \rightarrow (w = x \vee w = y \vee w = z)))
 \end{aligned}$$

exprime comme précédemment qu'il existe au moins trois entités distinctes qui satisfont P , mais elle affirme également (à tort dans ce cas précis) qu'il y en a au plus trois. La forme générale est évidemment :

$$\begin{aligned}
 (? \underset{D}{D}_n^1 x) Px = & ! (\exists x_1, \dots, \exists x_n, K(Px_1 \& \dots \& Px_n \& x_1 \neq x_2 \& \dots \\
 & \& x_1 \neq x_n \& \dots \& x_{n-1} \neq x_n \& \forall y(Py \\
 & \rightarrow (y = x_1 \vee \dots \vee y = x_n))))
 \end{aligned}$$

Lorsque $n = 1$, nous avons l'équivalent d'une question *which* demandant une réponse vraie unique ; lorsque $n > 1$, nous avons une forme de question *which* demandant une énumération exhaustive et distincte, avec : $v = u = n$ (Notons qu'une question demandant exactement n réponses vraies et ayant m variables libres peut être réduite à une question demandant au moins une réponse vraie avec $m.n$ variables libres [Åqvist 1965 : 123-124]).

Bien que l'auteur n'ait explicitement distingué comme principales formes de questions *what*, que les six types que nous avons décrits, cette classification est loin d'épuiser les possibilités du système QIE. Nous l'avons vu déjà à propos des variantes qui, dans l'interprétation d'un même type de question, permettent de rendre compte de nuances dans les intentions ou les attentes du locuteur. ÅQVIST mentionne d'autres formes simples de symbolisation, qu'il n'a pas jugé nécessaire de retenir comme types de questions *what*. La première est une forme (A) plus faible que la forme B demandant une énumération exhaustive ; en reprenant l'exemple des membres de l'Orchestre Philharmonique de Vienne, elle devient [Åqvist 1965 : 87-88] :

$$\begin{aligned} I &= (? \frac{1}{A} x) O_x \\ &= ! (\forall x, O_x \rightarrow \exists y((y = x) \& \exists z, K(z = y))) \end{aligned}$$

Cette forme est paraphrasable en : "Pour chaque individu membre de l'Orchestre, faites-moi savoir qui il est". On démontre qu'elle est impliquée par la forme : $(? \frac{1}{B} x) O_x$, qui demande en outre que l'on fasse savoir que l'individu cité est membre de l'Orchestre. La seconde forme simple mentionnée (E) est intermédiaire entre le type C demandant au moins n réponses vraies, et le type D demandant exactement n réponses vraies. Dans l'exemple des trois nombres premiers compris entre 10 et 20, elle devient [Åqvist 1965 : 121-122] :

$$\begin{aligned} I &= (? \frac{1}{E_3} x) P_x \\ &= ! (\exists x, \exists y, \exists z (K (P_x \& P_y \& P_z \& x \neq y \& x \neq z \& y \neq z)) \\ &\quad \& \forall w (K P_w \rightarrow (w = x \vee w = y \vee w = z))) \end{aligned}$$

On peut rendre cette forme par l'énoncé : "Faites en sorte qu'il y ait exactement trois entités dont je sache que ce sont des nombres premiers compris entre 10 et 20". Elle est plus forte que la forme C en ce qu'elle limite la connaissance requise par le locuteur aux seuls trois exemples demandés, et moins forte que la forme D en ce qu'elle ne requiert pas que ces trois exemples soient les seuls possibles. Cette forme E n'a pas été retenue par l'auteur en dernière analyse, pour des raisons de consistance avec des formes voisines.

En outre, ÅQVIST mentionne une extension ensembliste du système QIE qui permet de rendre compte, en particulier, de questions mettant en jeu des quantités [Åqvist 1965 : 118-119]. Par exemple, soit la question : "Combien y-a-t-il

d'abonnés à votre revue ?" ; si l'on pose : $Ax = "x \text{ est un abonné de la revue }"$, on peut la formaliser en :

$$\begin{aligned} I &= ! \exists y, K(y = | \{x : Ax\} |) \\ &= (? \underset{C_1}{C}^1 y) (y = | \{x : Ax\} |) \end{aligned}$$

où : $| \{x : Ax\} |$ est le cardinal de l'ensemble des abonnés. A titre d'exemple plus complexe, soit la question : "Qui sont $n\%$ des abonnés de votre revue ?". Le pronom interrogatif *qui* implique que les entités recherchées sont des êtres humains, ce que l'on exprimera par $Hx = "x \text{ est un être humain}"$. En utilisant la même notation que ci-dessus, on obtient l'expression QIE :

$$\begin{aligned} I &= ! \exists y, K((y \subseteq \{x : Ax\} \ \& \ |y| = \frac{n \cdot |\{x : Ax\}|}{100}) \ \& \ \forall z(Hz \ \& \ z \in y \\ &\rightarrow \exists w((w = z) \ \& \ KAw))) \end{aligned}$$

Enfin, avant d'abandonner le sujet des questions *what* non soumises à des conditions catégorielles, remarquons que l'on peut réduire les trois formes de questions *whether* à une question *what*. En reprenant l'exemple déjà utilisé :

$$I = "Est-ce Brutus ou Cassius qui a tué César ?"$$

et en utilisant le prédicat métalinguistique Y pour signifier :

$$Yx = "x \text{ est identique à l'une des modalités } B \text{ ou } C, \text{ et } x \text{ est vrai}"$$

on peut traduire : $?_2(B,C)$ par : "Quelle est au moins une modalité de réponse vraie parmi les suivantes : c'est Brutus qui a tué César ; c'est Cassius qui a tué César ?" [Åqvist 1965 : 65] formalisable en [Åqvist 1965 : 117] :

$$\begin{aligned} I &= (? \underset{C_1}{C}^1 x) Yx \\ &= ! (\exists x, KYx) \end{aligned}$$

De même, la forme : $?_2^u(B,C)$ peut également s'écrire :

$$\begin{aligned} I &= (? \underset{D_1}{D}^1 x) Yx \\ &= ! (\exists x, K(Yx \ \& \ \forall w(Yw \rightarrow w = x))) \end{aligned}$$

Enfin, la forme : $?_2^C (B,C)$ équivaut pratiquement à :

$$I = (?_{EKB}^1 x) Yx$$

$$= ! (\exists x(KYx) \ \& \ K(\forall x(Yx \rightarrow \exists w((w = x) \ \& \ KYw))))$$

Sur ces bases, on peut donc établir la table de correspondance suivante :

Question <i>whether</i> demandant :	Question <i>what</i> demandant :
- un exemple unique	- au moins une réponse vraie (C)
- une réponse vraie unique	- exactement une réponse vraie (D)
- une énumération exhaustive	- une énumération exhaustive (EKB)

Tableau n° 2.8. Réduction des questions *whether* aux questions *what*.

Dans tous les exemples de questions *what* analysés jusqu'ici, la variable libre était indépendante de conditions catégorielles. Or, soit la question : "Quels sont les Autrichiens qui sont membres de l'Orchestre Philharmonique de Vienne ?" [Åqvist 1965 : 109] ; en posant : $Ax = "x \text{ est un Autrichien}"$, et : $Ox = "x \text{ est un membre de l'Orchestre Philharmonique de Vienne}"$, on pourrait exprimer cette question par :

$$I = (?_B^1 x) (Ax \ \& \ Ox)$$

$$= ! (\forall x(Ax \ \& \ Ox \rightarrow \exists y((y = x) \ \& \ K(Ay \ \& \ Oy))))$$

ce qui signifierait que le locuteur désire savoir de certaines entités que ce sont des Autrichiens, et que ce sont des membres de l'Orchestre. La formulation de l'énoncé est moins exigeante, et on en rendrait mieux compte en plaçant l'expression Ay hors du champ de l'opérateur épistémique K :

$$I = ! (\forall x(Ax \rightarrow (Ox \rightarrow \exists y(Ay \ \& \ (y = x) \ \& \ KOy))))$$

Mais on peut également aborder le problème sous un autre angle, en posant simplement :

$$\begin{aligned} I &= (? \frac{1}{B} x) O_x \\ &= ! (\forall x(O_x \rightarrow \exists y((y = x) \& KO_y))) \end{aligned}$$

mais en ajoutant la condition que les variables x et y prennent leurs valeurs dans un sous-ensemble de l'ensemble des individus, la classe des Autrichiens. On écrira alors :

$$\begin{aligned} I &= (? \frac{1}{B} \overset{A}{x}) O_x \\ &= ! (\forall x(A_x \rightarrow (O_x \rightarrow \exists y ((y = x) \& KO_y)))) \end{aligned}$$

ce qui constitue une question *what* (demandant une énumération exhaustive) *soumise à une condition catégorielle A*.

Bien qu'ÅQVIST n'ait pas développé ce point, il est évidemment possible de concevoir des exemples de questions qui soient à la fois mixtes, conditionnelles, et soumises à des conditions catégorielles. On en trouvera une illustration dans l'analyse faite par HARRAH de l'ouvrage d'ÅQVIST [Harrah 1967 : 404].

Comme les questions *whether*, les questions *what* sont affectées par la révision du système PIE-QIE. Non seulement les formes inconditionnelles sont dérivées des formes conditionnelles, mais toutes les formes de questions *what* doivent être réinterprétées en raison de l'existence de quatre quantificateurs existentiels et de quatre quantificateurs universels (cf. [Åqvist 1971 : 43-44, 54-55]). Toutefois, nous limitant aux seuls points effectivement développés par l'auteur, nous ne présentons que la dérivation, à partir de la forme conditionnelle de base des questions *demandant au moins n réponses vraies* :

$$\begin{aligned} I &= (? \frac{1}{C_n} x / 1) (P_x / Q) \\ &= ! (Q \rightarrow \text{Noyau} ((? \frac{1}{C_n}) P_x)) \\ &= ! (\text{Noyau} ((? \frac{1}{C_n} x) P_x) / Q) \end{aligned}$$

[Åqvist 1965 : 125 ; 1971 : 54], de la forme inconditionnelle correspondante :

$$\begin{aligned} I' &= (? \underset{C_n}{\overset{1}{C}} x) Px \\ &= ! (Noyau((? \underset{C_n}{\overset{1}{C}} x) Px) / Prés((? \underset{C_n}{\overset{1}{C}} x) Px)) \& Prés((? \underset{C_n}{\overset{1}{C}} x) Px) \\ &= ! (Noyau(I') / Prés(I')) \& Prés(I') \end{aligned}$$

[Åqvist 1971 : 54].

d) Les réponses aux questions

Bien que ÅQVIST ait édifié une *logique des questions*, et non une logique des questions et des réponses, l'analyse syntaxique et sémantique des réponses aux questions reste pour lui un problème essentiel, auquel il consacre une place importante. Deux notions déjà définies sont nécessaires pour cette analyse : celle de *noyau* d'une question (forme développée de la question moins l'opérateur déontique), et celle de *présupposition* d'une question (forme développée, privée de tout opérateur épistémique ou déontique). La notion de présupposition conduit à celle de *correction* d'une question I :

$$Corr(I) = \overline{Prés(I)}$$

et de là, en fait, à celle de *réponse corrective* $r(I) = Corr(I)$.

Une application intéressante de la notion de réponse corrective est développée par l'auteur sous le nom de *chaîne de questions avec accumulation de corrections* [Åqvist 1969 : 114-118]. Soient par exemple les trois questions :

I_1 = "Quels sont les deux nombres premiers compris entre 13 et 17 (c'est-à-dire dans l'intervalle :]13,17[) ?"

$$\begin{aligned} &= (? \underset{D_1}{\overset{1}{D}} x) Px \\ &= ! (\exists x, \exists y, K(Px \& Py \& x \neq y \& \forall z(Pz \rightarrow (z = x \vee z = y)))) \end{aligned}$$

I_2 = "Citez au moins deux des nombres premiers compris entre 13 et 17"

$$\begin{aligned} &= (? \underset{C_1}{\overset{1}{C}} x) Px \\ &= ! (\exists x, \exists y, K(Px \& Py \& x \neq y)) \end{aligned}$$

$$\begin{aligned}
I_3 &= \text{"Y-a-t-il (au moins) un nombre premier compris entre 13 et 17 ?"} \\
&= ?_1 (\exists x, Px) \\
&= ! (K(\exists x, Px) \vee K(\overline{\exists x, Px}))
\end{aligned}$$

Il existe évidemment $3! = 6$ manières différentes d'ordonner ces trois questions ; il existe par conséquent 6 séquences différentes de présentation de ces questions à un interlocuteur. Toutefois, une seule de ces séquences peut être articulée selon la relation : "Est-ce que [...] ? *Sinon* : est-ce que [...] ?" ; c'est la séquence : " I_1 ; *sinon* I_2 ; *sinon* I_3 ". Plus formellement, on aurait la chaîne de questions :

$$\langle I_1 ; \text{Correc}(I_1) \ \& \ I_2 ; \text{Correc}(I_1) \ \& \ \text{Correc}(I_2) \ \& \ I_3 \rangle$$

paraphrasable en : "Quels sont les deux nombres premiers compris entre 13 et 17 ? / Etant donné qu'il n'y a pas *exactement* deux nombres premiers compris entre 13 et 17, citez-en au moins deux exemples / Etant donné qu'il n'y a ni *exactement* deux nombres premiers compris entre 13 et 17, ni plus de deux, y-en-a-t-il au moins un ?". Soit :

$$\begin{aligned}
I = ! & (Noyau(I_1) \vee (\text{Correc}(I_1) \ \& \ Noyau(I_2)) \vee \\
& (\text{Correc}(I_1) \ \& \ \text{Correc}(I_2) \ \& \ Noyau(I_3)))
\end{aligned}$$

Il est aisé de vérifier que toute autre séquence conduirait à des inconsistances, comme par exemple : $\text{Correc}(I_3) \ \& \ Noyau(I_2)$.

Cela conduit l'auteur à définir deux propriétés d'une chaîne de questions. Soit une chaîne de n questions : $S = \langle I_1, \dots, I_i, \dots, I_j, \dots, I_n \rangle$. On dira que la chaîne S est *susceptible d'accumulation de corrections* si et seulement si, pour tout couple de questions $\langle I_i, I_j \rangle$ tel que $i < j$, la présupposition de I_i n'est pas impliquée (dans le système PIE-QIE) par la présupposition de I_j : $(\text{Prés}(I_j) \rightarrow \text{Prés}(I_i))$. On dira d'autre part que la chaîne S est *susceptible d'inclusion de présuppositions* si et seulement si, pour tout couple $\langle I_i, I_j \rangle$ tel que $i < j$, la présupposition de I_i implique (dans le système PIE-QIE) celle de I_j : $\text{Prés}(I_i) \rightarrow \text{Prés}(I_j)$. Soit un ensemble fini de

questions $\{I_1, \dots, I_n\}$. On démontre qu'il existe au plus une chaîne S des n questions qui soit susceptible à la fois d'accumulation de corrections et d'inclusion de présuppositions.

L'analyse sémantique des réponses selon ÅQVIST repose sur la logique épistémique. L'auteur considère en effet que toute réponse à une question est liée de quelque manière à la satisfaction de la requête exprimée par la question ; et qu'en particulier : "si une proposition P compte pour une réponse à une question I , alors, si le locuteur parvient à *connaître* la vérité de P , alors la requête qu'il exprime en posant I est satisfaite" [Åqvist 1965 : 144]. Pour donner un contenu opératoire à ce postulat, on pose que la requête est satisfaite par la proposition P si celle-ci est une condition suffisante pour la vérité du noyau de la question ; en d'autres termes, si :

$P \rightarrow \text{Noyau}(I)$. En ne considérant pour le moment que le cas, plus facile à exposer, du système PIE, nous allons commencer par définir ce que l'auteur appelle l'*épistémisation* d'une proposition. Soit par exemple la proposition $P =$ "Jean est marié, et il avait l'habitude de battre sa femme" = $M \& B$; on appellera [Åqvist 1965 : 146] :

- *épistémisation maximale* de P l'expression :

$$K_{\max} P = KM \& KB .$$

("Je sais que Jean est marié, et je sais qu'il avait l'habitude de battre sa femme").

- *épistémisation partielle* de P (à condition que $K_{\max} P$ soit définie), ou bien l'épistémisation maximale d'une proposition Q plus faible que P , c'est-à-dire telle que : $(P \rightarrow Q) \& (\overline{Q} \rightarrow \overline{P})$, comme par exemple : KM ("Je sais que Jean est marié") ; ou bien le résultat de la suppression d'au moins une occurrence de K dans $K_{\max} P$, comme par exemple : $M \& KB$ ("Jean est marié, et je sais qu'il avait l'habitude de battre sa femme").

- *épistémisation minimale* de P l'épistémisation partielle obtenue par suppression de toutes les occurrences de K dans $K_{\max} P$, c'est-à-dire la proposition P elle-même.

Reprenons l'exemple de la question *whether* assurée selon la méthode *whether-if* (déjà utilisé au § 2.3.1.a), en modifiant légèrement l'énoncé et la formalisation correspondante :

I = "S'il est vrai que Jean est marié et qu'il avait l'habitude de battre sa femme, a-t-il cessé de le faire ?"

$$= ?_{2/1} (M \& B \& C, M \& B \& \bar{C} / M \& B)$$

qui a pour noyau :

$$\text{Noyau}(I) = (M \& B) \rightarrow (K(M \& B \& C) \vee K(M \& B \& \bar{C}))$$

Il est aisé d'illustrer sur cet exemple les quatre formes de réponses que définit l'auteur [Åqvist 1965 : 146-148, 154]. En effet, une proposition P est :

- une *réponse directe* à I, si P fournit une condition épistémique exactement suffisante à la vérité de *Noyau*(I), c'est-à-dire si et seulement si : $K_{\max} P$ implique *Noyau*(I) et il n'existe pas d'épistémisation partielle Q de P telle que : $(Q \rightarrow \text{Noyau}(I)) \& (\overline{\text{Noyau}(I)} \rightarrow Q)$. Exemples :

$$\begin{cases} d(I) = M \& B \& C \\ d'(I) = M \& B \& \bar{C} \end{cases}$$

- une *réponse suffisante* (ou : *complète*) à I, si P fournit une condition épistémique suffisante à la vérité de *Noyau*(I), c'est-à-dire si et seulement si P implique une réponse directe à I. Par exemple, la réponse :

$$\begin{aligned} r(I) &= M \& B \& C \& D \\ &= \text{"Jean a cessé de battre sa femme et il a déménagé"} \end{aligned}$$

apporte au locuteur une information non requise ; et, en conséquence, la réponse implique une proposition plus faible qui est, elle, une réponse directe à la question :

$$(M \& B \& C \& D) \rightarrow M \& B \& C \rightarrow \text{Noyau}(I)$$

- une *pseudo-réponse directe* à I, si P fournit une condition exactement suffisante à la vérité de *Noyau*(I), c'est-à-dire si et seulement si : la proposition P elle-même implique *Noyau*(I), et il n'existe pas de proposition Q plus faible que P et telle que $Q \rightarrow \text{Noyau}(I)$. Par exemple, la réponse :

$$r(I) = \overline{(M \& B)}$$

= "Jean n'avait pas l'habitude de battre sa femme car il n'a jamais été marié".

implique bien $Noyau(I) = \overline{(M \& B)} \vee (K(M \& B \& C) \vee K(M \& B \& \overline{C}))$, et ainsi constitue bien une condition exactement suffisante à la vérité de $Noyau(I)$, car la requête est satisfaite "automatiquement" par le fait brut (non épistémique) que l'hypothèse $M \& B$ est fautive (cf. [Åqvist 1965 : 148]).

- une *pseudo-réponse suffisante* à I , si P fournit une condition suffisante à la vérité de $Noyau(I)$, c'est-à-dire si et seulement si P implique une pseudo-réponse directe à I . Ainsi :

$$r(I) = \overline{M}$$

= "Jean n'est pas marié"

implique $\overline{(M \& B)}$ (Cette quatrième forme de réponse, imposée par des raisons de symétrie, n'est pas développée, ni même explicitement nommée, par l'auteur).

Les considérations qui précèdent peuvent être étendues au système QIE, et par conséquent aux réponses aux questions *what*, au prix toutefois d'une certaine complexification, spécialement en ce qui concerne la forme EB. La proposition P considérée comme réponse possible est de la forme :

$$P = P_1 \& \dots \& P_m \quad \& \quad \exists x(x = a_1) \& \dots \& \exists x(x = a_n)$$

(à condition qu'il n'y ait pas de P_i de la forme : $\exists x(x = a_j)$), et son épistémisation maximale se définit comme :

$$K_{\max} P = KP_1 \& \dots \& KP_m \quad \& \quad \exists x(K(x = a_1)) \& \dots \& \exists x(K(x = a_n))$$

Mais surtout, il est nécessaire de relativiser la notion de condition épistémique exactement suffisante, en la reformulant comme suit. Soient deux propositions P et R ; on dira que P est une condition épistémique exactement suffisante à la vérité de $Noyau(I)$ *relativement* à R si et seulement si : $(K_{\max} P) \& R$ implique $Noyau(I)$, et il n'existe pas d'épistémisation partielle Q de P telle que $(Q \& R)$ implique $Noyau(I)$ sans être impliquée par lui. La raison de cette modification peut être illustrée à partir de l'exemple suivant :

$$I = (? \underset{EB}{1} x) O_x$$

$$= ! (\exists x(KOx) \ \& \ \forall x(Ox \rightarrow \exists y((y = x) \ \& \ KOy)))$$

avec comme précédemment : $Ox = "x \text{ est un membre de l'OPV}"$. Soit la proposition :

$$P = Oa \ \& \ \forall y(Oy \rightarrow (y = a)) \ \& \ \exists x(x = a)$$

$$= "a \text{ est le seul membre de l'Orchestre, et } a \text{ existe}"$$

On démontre que P *n'est pas* une condition épistémique *exactement* suffisante à la vérité de $Noyau(I)$, car il existe une épistémisation partielle Q de P qui implique $Noyau(I)$:

$$Q = KOa \ \& \ \forall y(Oy \rightarrow (y = a)) \ \& \ \exists x(K(x = a))$$

D'autre part, la proposition :

$$S = KOa \ \& \ \exists x(K(x = a))$$

n'est pas une condition épistémique *exactement* suffisante à la vérité de $Noyau(I)$. D'où une situation de prime abord surprenante, dans laquelle la vérité de : $\forall y(Oy \rightarrow (y = a))$ est nécessaire, mais non celle de : $K(\forall y(Oy \rightarrow (y = a)))$, pour constituer une condition épistémique *exactement* suffisante à la vérité de $Noyau(I)$; en d'autres termes, il est nécessaire que cette proposition soit vraie, mais il n'est pas nécessaire que le locuteur sache qu'elle est vraie [Åqvist 1965 : 149]. C'est pourquoi on dira que : $Oa \ \& \ \exists x(x = a)$ est une condition épistémique *exactement* suffisante à la vérité de $Noyau(I)$ *relativement* à : $\forall y(Oy \rightarrow (y = a))$. Cette relativisation est également nécessaire pour la description des réponses directes à une question soumise à des conditions catégorielles.

Le tableau n° 2.9. donne la formule générale des réponses directes à quelques-unes des formes de questions décrites par ÅQVIST dans la première version (1965) de sa logique des questions.

Dans son ouvrage principal sur la logique des questions et des réponses [Åqvist 1965], l'auteur ne traite que de la situation "typique" (ou la plus

Questions <i>whether</i> :	$d(I) = \text{une des } P_i$
$I = ?_n (P_1, \dots, P_i, \dots, P_n)$	
Questions <i>whether</i> conditionnelles :	$d(I) = \text{une des } P_i$
$I = ?_{n/m} (P_1, \dots, P_i, \dots, P_n / Q_1, \dots, Q_m)$	pseudo-réponse directe : $(\overline{Q_1} \& \dots \& \overline{Q_m})$
Questions <i>what</i> :	$d(I) = Pa_1 \& \dots \& Pa_k \& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$
$I = (?_B^1 x) Px$	relativement à : $\forall x (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k))$
$= (?_{EB}^1 x / 1) (Px / \exists x, Px)$	pseudo-réponse directe = $(\overline{\exists x, Px})$
$I = (?_{KB}^1 x) Px$	problème non entièrement résolu [Åqvist 1965 : 158].
	$d(I) = \text{au minimum toute réponse directe à } I' = (?_{EKB}^1 x; + 1)(Px; (\overline{\exists x, Px}))$
	$= \left\{ \begin{array}{l} \text{soit : } Pa_1 \& \dots \& Pa_k \& \forall x (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k)) \\ \& \exists y (y = a_1) \& \dots \& \exists y (y = a_k) \\ \text{soit : } (\overline{\exists x, Px}) \end{array} \right.$
$I = (?_{EB}^1 x) Px$	$d(I) = Pa_1 \& \dots \& Pa_k \& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$
	relativement à : $\forall x (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k))$
$I = (?_{EKB}^1 x) Px$	$d(I) = Pa_1 \& \dots \& Pa_k \& \forall x (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k))$
	$\& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$
$I = (?_C^n^1 x) Px$	$d(I) = Pa_1 \& \dots \& Pa_n \& a_1 \neq a_2 \& a_1 \neq a_3 \& \dots \& a_1 \neq a_n \& a_2 \neq a_3 \& \dots$
	$\& a_{n-1} \neq a_n \& \exists y (y = a_1) \& \dots \& \exists y (y = a_n)$
$I = (?_D^n^1 x) Px$	$d(I) = Pa_1 \& \dots \& Pa_n \& a_1 \neq a_2 \& a_1 \neq a_3 \& \dots \& a_{n-1} \neq a_n \& \forall y (Py \rightarrow$
	$(y = a_1 \vee \dots \vee y = a_n)) \& \exists z (z = a_1) \& \dots \& \exists z (z = a_n)$
Questions <i>what</i> conditionnelles :	$d(I) = \text{toute réponse directe à } I' = (?_{EKB}^1 x) Px$
$I = (?_{EKB}^1 x/k) (Px/Q_1, \dots, Q_k)$	pseudo-réponse directe = $(\overline{Q_1} \& \dots \& \overline{Q_k})$

<p>Questions mixtes :</p> <p>$I = (?_{EB}^I x; +k) (Px; Q_1, \dots, Q_i, \dots, Q_k)$</p>	<p>soit toute réponse directe à $I' = (?_{EB}^I x) Px$</p> <p>$= d(I') = Pa_1 \& \dots \& Pa_k \& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$</p> <p>relativement à : $\forall x (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k))$</p> <p>soit une des Q_i</p>
<p>$I = (?_{EKB}^I x; +k) (Px; Q_1, \dots, Q_i, \dots, Q_k)$</p>	<p>soit toute réponse directe à $I' = (?_{EKB}^I x) Px$</p> <p>soit une des Q_i</p>
<p>$I = (?_{C_n}^I x; +k) (Px; Q_1, \dots, Q_i, \dots, Q_k)$</p>	<p>soit toute réponse directe à $I' = (?_{C_n}^I x) Px$</p> <p>soit une des Q_i</p>
<p>Questions <i>what</i> soumises à une condition catégorielle :</p> <p>$I = (?_{EB}^I x) Px$</p>	<p>$d(I) = Pa_1 \& \dots \& Pa_k \& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$</p> <p>relativement à :</p> <p>$Ja_1 \& \dots \& Ja_k \& \forall x (Jx \rightarrow (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k)))$</p>
<p>$I = (?_{EKB}^I x) Px$</p>	<p>$d(I) = Pa_1 \& \dots \& Pa_k \& \forall x (Jx \rightarrow (Px \rightarrow (x = a_1 \vee \dots \vee x = a_k)))$</p> <p>$\& \exists y (y = a_1) \& \dots \& \exists y (y = a_k)$</p> <p>relativement à : $Ja_1 \& \dots \& Ja_k$</p>
<p>$I = (?_{C_n}^I x) Px$</p>	<p>$d(I) = Pa_1 \& \dots \& Pa_n \& a_1 \neq a_2 \& a_1 \neq a_3 \& \dots \& a_{n-1} \neq a_n$</p> <p>$\& \exists y (y = a_1) \& \dots \& \exists y (y = a_n)$</p> <p>relativement à : $Ja_1 \& \dots \& Ja_n$</p>
<p>$I = (?_{D_n}^I x) Px$</p>	<p>$d(I) = Pa_1 \& \dots \& Pa_n \& a_1 \neq a_2 \& a_1 \neq a_3 \& \dots \& a_{n-1} \neq a_n$</p> <p>$\& \forall y (Jy \rightarrow (Py \rightarrow (y = a_1 \vee \dots \vee y = a_n)))$</p> <p>$\& \exists z (z = a_1) \& \dots \& \exists z (z = a_n)$</p> <p>relativement à : $Ja_1 \& \dots \& Ja_n$</p>

Tableau n° 2.9 : Exemples de réponses directes aux questions [Åqvist 1965 : 156-159].

courante) caractérisée par le fait que celui qui pose par exemple la question : $?_3^u(A, B, C)$ ne sait pas laquelle de ces trois propositions est vraie, mais sait qu'une d'entre elles est vraie ; et par le fait qu' "il désire accroître son savoir en levant son ignorance à cet égard" [Åqvist 1965 : 8]. Cette situation se caractérise formellement par le fait que le noyau de la question est toujours faux dans le contexte ainsi défini [Åqvist 1969 : 119]. Elle écarte par conséquent une variété de situations, dont l'archétype est la situation d'examen, dans laquelle le locuteur connaît la réponse à sa question, mais désire savoir en fait si son interlocuteur la connaît également. Nous retrouvons cette forme de question dans les tests de connaissance, dans certains items de tests d'aptitudes ou de niveau, et également dans certaines questions d'information ou de notoriété utilisées dans les questionnaires d'enquête. Cette forme est caractérisée par le fait que le noyau de la question est toujours vrai relativement au contexte.

Pour rendre compte de cette nouvelle famille de questions, ÅQVIST a recours à l'opérateur épistémique déjà mentionné en présentant le langage formel utilisé :

$K \# =$ "Vous savez que".

Prenons l'exemple d'un examinateur qui demande :

$I =$ "Qui a composé la Symphonie Héroïque ?"

S'il ne s'agissait pas d'une question d'examen, on pourrait avoir la transcription (cf. : [Åqvist 1965 : 114]) :

$$I = (? \underset{C_1}{C}^1 x) Sx$$

$$= ! \exists x(KSx)$$

(pour simplifier, nous ne tenons pas compte de la soumission de x à la condition catégorielle impliquée par *qui* : $Hx =$ "x est un être humain"). Mais cette formalisation ne convient évidemment pas à une situation d'examen. L'auteur examine trois traductions formelles possibles [Åqvist 1969 : 120], sans toutefois mésestimer la difficulté de cette transcription.

La première traduction pourrait être :

$$\begin{aligned}
 I &= ?_1 (\exists x(K \# Sx)) \\
 &= ! (K(\exists x, K \# Sx) \vee K(\overline{\exists x, K \# Sx})) \\
 &= \text{"Dites-moi si vous savez qui a composé la Symphonie Héroïque"}
 \end{aligned}$$

L'inconvénient de cette solution est d'admettre comme réponse directe :

$$\begin{aligned}
 d(I) &= \exists x, K \# Sx \\
 &= \text{"Je sais qui a composé la Symphonie Héroïque"}
 \end{aligned}$$

réponse qui ne peut évidemment satisfaire l'examineur. Cette formalisation est donc trop faible.

Une traduction plus forte serait :

$$\begin{aligned}
 I &= (? \underset{C_1}{1} x) (K \# Sx) \\
 &= ! \exists x(K \# Sx) \\
 &= \text{"Dites-moi qui est celui dont vous savez qu'il a composé la"} \\
 &\quad \text{Symphonie Héroïque"}
 \end{aligned}$$

Cette formalisation se révèle trop forte pour rendre compte de l'énoncé original. En effet :

$$Prés(I) = \exists x, K \# Sx$$

La réponse : "Je ne sais pas qui a composé la Symphonie Héroïque" n'est donc pas une réponse directe, mais une réponse corrective à I .

Une solution intermédiaire plus satisfaisante réside dans la question mixte *whether-what* :

$$\begin{aligned}
 I &= (? \underset{C_1}{1} x ; +1) (K \# Sx ; (\overline{\exists x, K \# Sx})) \\
 &= ! (\exists x(K \# Sx) \vee K(\overline{\exists x, K \# Sx})) \\
 &= \text{"Dites-moi qui est celui dont vous savez qu'il a composé la"} \\
 &\quad \text{Symphonie Héroïque, ou que vous ne connaissez personne dans} \\
 &\quad \text{ce cas"}.
 \end{aligned}$$

Cette formalisation admet en effet comme réponse directe aussi bien un énoncé citant une personne susceptible d'être le compositeur recherché, qu'un aveu d'ignorance.

2.3.2. La théorie logique des questions de Tadeusz KUBIŃSKI

Les premiers travaux publiés par KUBIŃSKI sur la logique des questions remontent à 1958 (selon la bibliographie de [Kubiński 1970]), c'est-à-dire à l'année même de la publication du célèbre article de C.L. HAMBLIN sur les questions. Depuis cette date, l'auteur a publié au moins une douzaine d'autres articles (dont quelques-uns en anglais), ainsi qu'un ouvrage fondamental, l'*Introduction à la théorie logique des questions* [Kubiński 1970], comparable en importance à [Belnap 1976]. Or, si KUBIŃSKI, au travers de son oeuvre, se révèle parfaitement au courant de l'ensemble des publications relatives à la logique érotétique, il ne semble pas que les logiciens des questions lui aient accordé en retour toute l'attention qu'il mérite. Par exemple, [Kubiński 1970] n'est pas utilisé dans [Belnap 1976] (bien que cité en bibliographie), et n'est analysé dans le *Journal of Symbolic Logic* que 6 ans après sa publication (cf. [Materna 1977]). Peut-être faut-il en voir la raison, non seulement dans les barrières linguistiques, mais aussi dans l'extrême concision de certains textes de KUBIŃSKI (cf. par exemple [Kubiński 1958]) ; ce qui fait dire à son commentateur : "KUBIŃSKI a écrit un livre important et utile qui traite d'un grand nombre de problèmes en relation avec la théorie des questions. Le style de l'ouvrage est peut-être trop concis, ce qui peut parfois nuire à la compréhension" [Materna 1977 : 428]. D'autant que, dans [Kubiński 1970] en particulier, l'auteur laisse au lecteur le soin de procéder lui-même à certains développements purement déductifs.

Nous ne présentons ici que les éléments qui nous ont paru les plus intéressants pour notre recherche parmi les nombreux résultats auxquels aboutit l'auteur. Nous avons délibérément adopté (comme pour l'exposition des travaux d'ÅQVIST) l'optique de l'utilisateur. Ceci nous a conduit à négliger l'arrière-plan théorique qui est particulièrement important chez KUBIŃSKI ; et à donner une place privilégiée à la description des principales catégories de questions,

des types de réponses et de leurs rapports avec les questions de diverses catégories, et enfin des relations entre les présuppositions, les questions et les réponses.

Cette présentation s'appuie principalement sur [Kubiński 1970], dont elle adopte pour l'essentiel la formalisation (en évitant cependant des formes d'écriture trop condensées, et en conservant dans la mesure du possible les symboles déjà utilisés pour BELNAP et ÅQVIST). Toutefois, il nous a paru nécessaire de sacrifier parfois la rigueur à la simplicité de l'exposé, en particulier en ne distinguant pas entre la *lexis* et l'*assertion* de celle-ci, ni entre *opérateur* interrogatif et *foncteur* interrogatif, et en considérant comme équivalentes des expressions qui ne le sont qu'à une permutation près de leurs éléments unis par des opérateurs de conjonction ou de disjonction (Le contexte permet d'ailleurs toujours de lever les ambiguïtés). Dans notre exposition, nous commençons par décrire les variétés de questions numériques, et leur syntaxe ; nous traitons ensuite de toutes les autres formes de questions, dont l'ensemble paraît moins bien structuré que celui des questions numériques ; nous abordons enfin les problèmes sémantiques, avec l'analyse des formes de réponse à une question, et des présuppositions.

a) Les questions numériques

Les *questions numériques* (ou : *questions à compléter* [Kubiński 1958 : 317], ou encore : *questions du deuxième type* [Kubiński 1966 : 109]) correspondent aux questions *which* de BELNAP. Mais l'analyse syntaxique, et la classification interne, qu'en donne KUBIŃSKI diffèrent totalement de celles que propose BELNAP. Une question numérique est formée d'opérateurs interrogatifs numériques, suivis d'une expression comportant des variables dont certaines sont liées aux opérateurs numériques (Ceux-ci jouent un rôle analogue, sur le plan syntaxique, à celui dévolu aux quantificateurs ; d'où le nom de "quantificateurs interrogatifs", utilisé dans les premiers textes sur ce sujet. Cf. par exemple [Kubiński 1958 : 317 ; 1966 : 119]). On peut distinguer trois classes de questions numériques, appelées respectivement de première catégorie, de deuxième catégorie, et de type C.

Les questions numériques de première catégorie requièrent un nombre de réponses déterminé par rapport à un seuil fixé k , mais n'expriment pas d'exigence d'exhaustivité (contrairement aux deux autres classes de questions numériques). La première forme de questions de première catégorie est illustrée par l'énoncé suivant [Kubiński 1970 : 16] :

I = "Citez trois maréchaux de Napoléon qui se sont ralliés au trône des Bourbons"

avec, comme réponse directe :

d(I) = "Lefèvre est un maréchal de Napoléon qui s'est rallié au trône des Bourbons, et Macdonald est un maréchal de Napoléon qui s'est rallié au trône des Bourbons, et Oudinot est un maréchal de Napoléon qui s'est rallié au trône des Bourbons"

Formellement, cette question est de la forme : "Quels sont les k x pour lesquels on a Ax ?", soit :

$I = kx \quad Ax$

$d(I) = Aa_1 \ \& \ \dots \ \& \ Aa_k$

Traduite dans le système de BELNAP, cette question serait une interrogative *which* demandant exactement k exemples, soit : $I = ?\binom{k}{k} - -) (Cx//Ax)$, forme à laquelle BELNAP n'accorde pas d'attention particulière (si l'on excepte le cas où : $k = 1$).

La correspondance entre les deux systèmes devient moins évidente lorsque l'on considère les questions à plusieurs arguments. Soit par exemple l'énoncé [Kubiński 1970 : 18] :

I = "Citez deux écrivains français et deux de leurs romans".

Cet énoncé est susceptible de trois interprétations distinctes, qui déterminent trois variétés de la première forme de questions numériques de première catégorie. A la première variété correspondent les formalisations :

$$I = {}^1_{2/2} xy \quad Axy$$

$$d(I) = Aa_i b_k \quad \& \quad Aa_j b_m, \text{ avec : } i \neq j \text{ et : } k \neq m .$$

= "Balzac a écrit *Le père Goriot* et Stendhal a écrit *La Chartreuse de Parme*"

Cette interprétation considère I comme équivalent à : "Citez deux écrivains français et, pour chacun d'eux, un des romans qu'il a écrits". La seconde variété s'écrit :

$$I = {}^2_{2/2} xy \quad Axy$$

$$d(I) = Aa_i b_k \quad \& \quad Aa_i b_m \quad \& \quad Aa_j b_k \quad \& \quad Aa_j b_m \text{ avec : } i \neq j \text{ et : } k \neq m$$

= "Edmond Goncourt et Jules Goncourt ont écrit ensemble *Charles Demailly et Soeur Philomène*".

Cette seconde interprétation correspond à : I = "Citez deux écrivains français ayant écrit des romans en commun, et citez deux de ces romans". Enfin, la troisième variété a pour forme :

$$I = {}^3_{2/2} xy \quad Axy$$

$$d(I) = Aa_i b_k \quad \& \quad Aa_i b_m \quad \& \quad Aa_j b_n \quad \& \quad Aa_j b_p \text{ avec : } i \neq j, k \neq m, n \neq p .$$

= "Balzac a écrit *Le Père Goriot* et *Eugénie Grandet*, et Stendhal a écrit *La Chartreuse de Parme* et *Le rouge et le noir*".

Soit I = "Citez deux écrivains français et, pour chacun d'eux, deux de leurs romans".

En généralisant ceci, on définit les *questions numériques complexes de première catégorie de première forme à deux arguments*, et leurs réponses directes [Kubiński 1970 : 19-20].

Première variété :

$$I = {}^1_{m/n} xy \quad Axy \quad \text{avec : } m \geq 1 \text{ et : } n \geq 1$$

$$d(I) = (Aa_1 b_1 \quad \& \quad \dots \quad \& \quad Aa_1 b_j) \quad \& \quad (Aa_2 b_{j+1} \quad \& \quad \dots \quad \& \quad Aa_2 b_k) \quad \& \quad \dots \\ \& \quad (Aa_m b_{p+1} \quad \& \quad \dots \quad \& \quad Aa_m b_n)$$

Deuxième variété :

$$I = {}^2_{m/n} \text{ xy } \text{ Axy} \quad \text{avec : } m \geq 1 \text{ et } n \geq 1$$

$$d(I) = (Aa_1 b_1 \& \dots \& Aa_1 b_n) \quad \& \quad (Aa_2 b_1 \& \dots \& Aa_2 b_n) \quad \& \dots$$

$$\quad \& \quad (Aa_m b_1 \& \dots \& Aa_m b_n)$$

Troisième variété :

$$I = {}^3_{m/n} \text{ xy } \text{ Axy} \quad \text{avec : } m \geq 1 \text{ et } n \geq 1$$

$$d(I) = (Aa_1 b_{1_1} \& \dots \& Aa_1 b_{1_n}) \quad \& \quad (Aa_2 b_{2_1} \& \dots$$

$$Aa_2 b_{2_n}) \quad \& \dots \& \quad (Aa_m b_{m_1} \& \dots \& Aa_m b_{m_n})$$

Dans toutes ces formules, les termes dont les symboles sont différents sont supposés être distincts les uns des autres.

Il est évidemment possible de généraliser ces formules en les étendant aux questions complexes à plus de deux arguments. Soit par exemple la question suivante, portant sur trois types d'inconnues [Kubiński 1970 : 20] (et intraduisible en français sous la forme d'une question unique) :

$$I = \text{"Quatre députés se sont prononcés contre trois projets au cours de deux séances de la commission de l'Assemblée Parlementaire. Qui sont ces députés ? De quels projets et de quelles séances s'agit-il?"}$$

$$= {}^{1/1}_{4/3/2} \text{ xyz } \text{ Axyz}$$

avec :

$$\text{Axyz} = \text{"le député } x \text{ s'est prononcé contre le projet } y \text{ à la séance } z \text{ de la commission"}$$

La généralisation à k arguments permet d'ailleurs de mêler dans une même question des variétés différentes de questions numériques de la première forme de la première catégorie, puisque la formule de la question s'écrit [Kubiński 1970 : 21] :

$$I = {}^{i_1/\dots/i_{n-1}}_{m_1/\dots/m_n} \text{ } x_1 \dots x_n \quad \text{Ax}_1 \dots x_n$$

avec :

$$n > 1$$

$$i_\lambda = 1, 2, \text{ ou } 3, \quad \text{pour : } 1 \leq \lambda \leq n-1$$

$x_1 \dots x_n$ = suite de n variables d'individu distinctes

$$m_\mu \geq 1, \quad \text{pour } 1 \leq \mu \leq n$$

La seconde forme de questions numériques de première catégorie est illustrée par l'énoncé [Kubiński 1970 : 17] :

I = "Citez au moins trois pianistes polonais qui ont été lauréats du concours international de piano Frédéric Chopin"

Sa forme générale est, dans le cas le plus simple (un seul argument) :

$$I = k \leq x \quad Ax$$

$$= \text{"Citez au moins } k \text{ } x \text{ tels que } Ax\text{"}$$

$$d(I) = Aa_1 \& \dots \& Aa_m \quad \text{avec } m \geq k .$$

Dans le système de BELNAP, cette forme s'écrirait :

$$I = ? \left(\bar{C}_k - - \right) (Cx // Ax)$$

Cette seconde forme est également généralisable à plus d'un argument, selon les mêmes règles que la première forme.

La troisième forme de questions numériques de première catégorie a pour formule, dans le cas le plus simple :

$$I = k < x \quad Ax$$

$$= \text{"Citez plus de } k \text{ } x \text{ tels que } Ax\text{"}$$

On peut la considérer comme une écriture abrégée de :

$$I = k + 1 \leq x \quad Ax$$

$$= ? \left(\bar{C}_{k+1} - - \right) (Cx // Ax) \quad (\text{chez BELNAP})$$

Elle est donc généralisable à k arguments comme les formes précédentes.

Les questions numériques de deuxième catégorie sont comparables aux questions de première catégorie, à ceci près qu'elles exigent en outre l'exhaustivité de la réponse. Tel est par exemple [Kubiński 1970 : 24] l'énoncé : "Quels sont les deux villes polonaises qui, en 1968, ont plus de 700.000 habitants ?", formalisable en : $I = (k)x \quad Ax$. Une réponse directe à la question serait : "Varsovie et Cracovie sont les deux seules villes polonaises de plus de 700.000 habitants en 1968" ; l'affirmation d'exhaustivité est exprimée comme chez BELNAP. Nous avons par conséquent, dans le cas le plus simple (un argument) :

Première forme :

$$\begin{aligned} I &= \text{"Citez tous les } x, \text{ au nombre d'exactement } k, \text{ qui sont } Ax\text{"} \\ &= (k)x \quad Ax \\ d(I) &= Aa_1 \ \& \ \dots \ \& \ Aa_k \quad \& \quad \forall x (Ax \rightarrow ((x=a_1) \vee \dots \vee (x=a_k))) \end{aligned}$$

Deuxième forme :

$$\begin{aligned} I &= \text{"Citez tous les } x, \text{ au nombre d'au moins } k, \text{ qui sont } Ax\text{"} \\ &= (k\leq)x \quad Ax \\ d(I) &= Aa_1 \ \& \ \dots \ \& \ Aa_m \quad \& \quad \forall x (Ax \rightarrow ((x=a_1) \\ &\quad \vee \dots \vee (x = a_m))) \quad \text{avec } m \geq k \end{aligned}$$

Troisième forme :

$$\begin{aligned} I &= \text{"Citez tous les } x, \text{ au nombre de plus de } k, \text{ qui sont } Ax\text{"} \\ &= (k<)x \quad Ax \\ d(I) &= Aa_1 \ \& \ \dots \ \& \ Aa_m \quad \& \quad \forall x (Ax \rightarrow ((x=a_1) \\ &\quad \vee \dots \vee (x = a_m))) \quad \text{avec } m > k. \end{aligned}$$

Ces trois formes correspondent chez BELNAP respectivement à :

$$\begin{aligned} I_1 &= ? \binom{k}{k} \forall - (Cx // Ax) \\ I_2 &= ? \binom{-}{k} \forall - (Cx // Ax) \\ I_3 &= ? \binom{-}{k+1} \forall - (Cx // Ax) \end{aligned}$$

Les questions numériques de deuxième catégorie sont généralisables à k arguments selon les mêmes principes que les questions de première catégorie.

Les questions de type C demandent une énumération exhaustive, sans toutefois que le nombre d'individus demandé soit précisé dans la question. Par exemple [Kubiński 1970 : 21] : "Quelles sont les villes de plus de 200.000 habitants qui sont sur l'Oder ?". Si l'on interprète la question comme étant de la première catégorie (opérateur : $1 \leq x$), une réponse directe serait : "Breslau et Stettin sont deux villes sur l'Oder de plus de 200.000 habitants". Mais si la question est de type C, la réponse directe doit préciser en outre : "et il n'y en a pas d'autre" ; ou, ce qui est équivalent : "et s'il y a une ville de plus de 200.000 habitants qui est sur l'Oder, c'est Breslau ou Stettin". On a par conséquent la formalisation suivante dans le cas simple :

$$\begin{aligned} I &= \text{"Quels sont tous les } x \text{ tels que } Ax\text{"} \\ &= Cx \quad Ax \\ d(I) &= Aa_1 \ \& \ \dots \ \& \ Aa_m \quad \& \quad \forall x (Ax \rightarrow ((x=a_1 \vee \dots \vee x=a_m))) \end{aligned}$$

Les questions de type C à deux arguments s'interprètent de la même manière que les autres questions numériques. Par exemple [Kubiński 1970 : 22], l'énoncé : "Citez tous les compositeurs de musique symphonique allemands romantiques, et toutes les symphonies qu'ils ont écrites", peut avoir comme opérateurs interrogatifs : ${}^1C/C$, ${}^2C/C$, ou ${}^3C/C$. Si l'on choisit la seconde interprétation (deuxième variété), on a :

$$\begin{aligned} I &= {}^2C/C \ xy \quad Axy \\ d(I) &= Aa_1 \ b_1 \ \& \ \dots \ \& \ Aa_1 \ b_n \ \& \ Aa_2 \ b_1 \ \& \ \dots \ \& \ Aa_2 \ b_n \ \& \\ &\quad \dots \ \& \ Aa_m \ b_1 \ \& \ \dots \ \& \ Aa_m \ b_n \quad \& \quad \forall x (Axb_1 \\ &\quad \rightarrow ((x = a_1) \vee \dots \vee (x = a_m))) \ \& \ \dots \ \& \ \forall x (Axb_n \\ &\quad \rightarrow ((x = a_1) \vee \dots \vee (x = a_m))) \quad \& \quad \forall y \\ &\quad (Aa_1 y \rightarrow ((y = b_1) \vee \dots \vee (y = b_n))) \ \& \ \dots \ \& \ \forall y \\ &\quad (Aa_m y \rightarrow ((y = b_1) \vee \dots \vee (y = b_n))) \end{aligned}$$

Les questions de type C peuvent être généralisées à k arguments.

Nous avons vu jusqu'ici des formes de questions numériques complexes (à plus d'un argument) n'utilisant qu'un même type d'opérateur interrogatif. Il

est possible de généraliser ce qui précède en posant, comme formule générale des questions numériques :

$$I = i_1 / \dots / i_{n-1} \quad \Omega_1 / \dots / \Omega_n \quad x_1 \dots x_n \quad Ax_1 \dots x_n$$

avec :

$$n > 1$$

$$i_\lambda = 1, 2, \text{ ou } 3, \quad \text{pour : } 1 \leq \lambda \leq n-1$$

$$x_1 \dots x_n = \text{ suite de } n \text{ variables d'individu distinctes}$$

$$\Omega_\mu = k, \quad k \leq, \quad k <, \quad (k), \quad (k \leq), \quad (k <), \quad \text{ou } C, \\ \text{pour : } 1 \leq \mu \leq n, \quad \text{avec : } k \geq 1.$$

Ainsi, la *question mixte* [Kubiński 1970 : 23] :

$$I = {}^2C/2 \quad xy \quad Axy$$

a pour réponse directe :

$$d(I) = Aa_1 b_1 \quad \& \quad Aa_1 b_2 \quad \& \quad Aa_2 b_1 \quad \& \quad Aa_2 b_2 \quad \& \quad \dots \quad \& \\ Aa_m b_1 \quad \& \quad Aa_m b_2 \quad \& \quad \forall x (Axb_1 \rightarrow ((x = a_1) \\ \vee \dots \vee (x = a_m))) \quad \& \quad \forall x (Axb_2 \rightarrow ((x = a_1) \vee \dots \vee (x = a_m)))$$

D'autre part, bien que nous n'ayons examiné que des questions portant sur des variables d'individu, l'ensemble des considérations qui précèdent s'applique également à des questions portant sur des variables de prédicat. Par exemple, soit l'énoncé [Kubiński 1970 : 17] : "Qui est le chef militaire français qui a gagné la bataille d'Austerlitz ?". Si cet énoncé signifie : "Quel est le nom" ou : "Quelles sont les coordonnées du chef qui ... ?", il correspond à : $I = \lambda x \quad Ax$, et ses réponses directes sont de la forme : $d(I) = Aa =$ "Napoléon Bonaparte", ou : "L'empereur des Français né à Ajaccio a gagné la bataille d'Austerlitz". Si par contre il exprime la question : "Quel genre d'homme est le chef qui ... ?", l'interrogation porte sur une variable de prédicat (à une place) X^1 . On a dans ce cas la formalisation :

$$I = \lambda X^1 \quad AX^1 \\ d(I) = AP^1$$

la réponse directe d(I) pouvant se traduire par : "Celui qui a gagné la bataille d'Austerlitz est un chef militaire décédé dans la première moitié du 19e siècle", ou : "... est un tacticien de génie". Les questions numériques portant sur des variables de prédicat unaires sont appelées *questions sur des attributs d'individus* [Kubiński 1970 : 33]. Leur forme générale est :

$$I = i_1 / \dots / i_{n-1} \quad \Omega_1 / \dots / \Omega_n \quad X_1^1 \dots X_n^1 \quad A \quad X_1^1 \dots X_n^1$$

avec : $X_1^1 \dots X_n^1$ = suite de n variables de prédicats unaires distinctes (les autres symboles ayant ici la même signification que dans la formule générale de question numérique portant sur des variables d'individus). Lorsqu'il s'agit de prédicats à plus d'une place (prédicats n -aires, avec : $n > 1$), on appelle les questions de ce type : *questions sur des relations entre individus* [Kubiński 1970 : 33]. Sur le modèle de ces questions *homogènes* (dont toutes les variables appartiennent à la même catégorie), il est possible de concevoir des questions numériques complexes *hétérogènes* [Kubiński 1970 : 39-40], c'est-à-dire des questions numériques à plusieurs arguments dans lesquelles l'interrogation porte à la fois sur des variables d'individu et sur des variables de prédicat.

Aux questions numériques simples se rattachent les *questions de définition* [Kubiński 1970 : 42]. Soit par exemple l'énoncé interrogatif : "Qu'est-ce que Buenos-Aires ?", qui a pour réponse réelle directe : "Buenos-Aires est la capitale de l'Argentine". On peut le formaliser en :

$$I = (1) x \quad (x = a)$$

$$d(I) = (b = a) \quad \& \quad \forall x ((x = a) \rightarrow (x = b))$$

avec :

x = variable d'individu

a = constante d'individu = "Buenos Aires"

b = constante d'individu = "la capitale de l'Argentine"

Lorsque la question porte sur des attributs d'individus, les principes de formalisation restent les mêmes. Par exemple :

I = "Qui étaient les huguenots ?"

= (1) X^1 ($X^1 = P^1$)

d(I) = "Les huguenots étaient des Français de confession calviniste vivant aux 16e et 17e siècles"

= ($Q^1 = P^1$) & $\forall X^1 ((X^1 = P^1) \rightarrow (X^1 = Q^1))$

L'affirmation d'unicité (d'exhaustivité) dans les réponses directes indique qu'il s'agit ici de réponses *réelles*. Lorsque la définition demandée est une définition *nominale*, l'opérateur d'identité est remplacé par un prédicat binaire A^2 dont la signification est (approximativement) : "désigne (ou : exprime) dans un système linguistique" (c'est-à-dire : "a pour référent") :

I = 1 x A^2_{ax}

= "Quel est l'individu x désigné par le nom a ?"

Les questions de définition (réelle ou nominale) peuvent avoir pour opérateur interrogatif l'un quelconque des sept opérateurs numériques que nous avons décrits.

b) Les autres types de questions

En dehors des questions numériques, KUBIŃSKI analyse diverses formes de questions que nous passons rapidement en revue. Les trois premières : questions de décision, questions conjonctives, et questions disjonctives, correspondent à peu près aux questions *whether* de BELNAP ; la quatrième : questions conditionnelles, recouvre en gros les questions *whether* relativisées (les questions appartenant à l'une de ces quatre premières catégories peuvent être appelées : *questions du premier type* [Kubiński 1966 : 109], ou : *questions de type "est-ce que"* [Kubiński 1970 : 39]) ; la cinquième : questions sur la valeur logique, est sans équivalent direct chez BELNAP (à moins qu'on les considère comme une variété de questions *whether*), mais présente quelques analogies avec les questions de vérité chez STAHL (cf. § 2.3.4.c) ; enfin, la sixième est celle des questions de causalité.

Les *questions de décision* proposent à l'interlocuteur le choix entre une expression quelconque A , et sa négation, et se lisent : "Est-il vrai que A ?" [Kubiński 1958 : 316]. La formalisation correspondante est [Kubiński 1970 : 27]:

$$I = [\alpha] \quad \alpha A$$

L'expression A est appelée la *racine* de la question de décision ; les réponses directes à une question de décision sont la racine de la question et sa négation [Kubiński 1958 : 318] ; soit ici :

$$\begin{cases} d(I) = A \\ d(I)' = \bar{A} \end{cases}$$

Les *questions disjonctives* sont illustrées par l'énoncé suivant [Kubiński 1970 : 26] : "L'homme que l'on a aperçu à tel endroit et à tel moment portait-il un manteau noir, un veston à carreaux, ou seulement une chemise bleu clair ?", formalisable en :

$$I = [\beta^3] \quad \beta^3 A_1, A_2, A_3$$

Ce type de question correspond chez BELNAP à la question *whether* à réponse vraie unique, de forme :

$$I = ? \left(\begin{matrix} 1 \\ \vee \end{matrix} \right) (A_1, A_2, A_3)$$

Les trois réponses directes possibles à cette question sont :

$$\begin{cases} d(I)_1 = A_1 \ \& \ \bar{A}_2 \ \& \ \bar{A}_3 \\ d(I)_2 = \bar{A}_1 \ \& \ A_2 \ \& \ \bar{A}_3 \\ d(I)_3 = \bar{A}_1 \ \& \ \bar{A}_2 \ \& \ A_3 \end{cases}$$

La forme générale des questions disjonctives est :

$$I = [\beta^n] \quad \beta^n A_1, \dots, A_n$$

$$d(I) = \bar{A}_1 \ \& \ \dots \ \& \ \bar{A}_{i-1} \ \& \ A_i \ \& \ \bar{A}_{i+1} \ \& \ \dots \ \& \ \bar{A}_n$$

Toute réponse directe à une question disjonctive portant sur n expressions est une conjonction dont l'un des membres est une des expressions A_i , et dont les

$n-1$ autres membres sont la négation des autres expressions. Il y a par conséquent n réponses directes possibles. Dans le cas particulier où l'on a :

$$I = [\beta^2] \quad \beta^2 A, \bar{A}$$

la question disjunctive I est strictement équivalente à une question de décision de forme : $I = [\alpha] \quad \alpha A$.

Les *questions conjonctives* proposent une conjonction d'expressions dont chacune peut être vraie ou fausse ; par exemple [Kubiński 1970 : 25] : "Jan Kowalski travaille-t-il dans une banque, habite-t-il dans le centre de l'agglomération, et collectionne-t-il les timbres-poste ?", que l'on écrit :

$$I = [\delta^3] \quad \delta^3 A_1, A_2, A_3$$

Il y a 2^3 réponses directes possibles, parmi lesquelles :

$$\left\{ \begin{array}{l} d(I)_1 = A_1 \ \& \ A_2 \ \& \ A_3 \\ \dots \\ d(I)_3 = \bar{A}_1 \ \& \ A_2 \ \& \ \bar{A}_3 \\ \dots \\ d(I)_8 = \bar{A}_1 \ \& \ \bar{A}_2 \ \& \ \bar{A}_3 \end{array} \right.$$

La forme générale des questions conjonctives est :

$$I = [\delta^n] \quad \delta^n A_1, \dots, A_n$$

Les réponses directes à une question conjonctive sont une conjonction à n membres, qui sont chacun soit une expression, soit sa négation. Pour une question conjonctive donnée, il y a par conséquent 2^n réponses directes possibles. Dans le cas particulier où l'on a : $I = [\delta^1] \quad \delta^1 A$, la question I est strictement équivalente à : $I' = [\alpha] \quad \alpha A$, et à $I'' = [\beta^2] \quad \beta^2 A, \bar{A}$ (cf. [Kubiński 1970 : 27]).

Les *questions conditionnelles*, dans la terminologie de KUBIŃSKI, recouvrent en fait deux types distincts de questions. Les questions du premier type, *questions à antécédents dirimants* [Kubiński 1970 : 29], correspondent approximativement aux questions conditionnelles de BELNAP, à ceci près qu'elles admettent comme réponse directe la négation de la condition. Soit par exemple

Les questions conditionnelles à *antécédent non dirimant* sont assez proche des questions "puisque" (cf. § 2.2.4.b). L'auteur en donne comme exemple [Kubiński 1970 : 28] : "Si les droites d_1 et d_2 sont correctement représentées par les équations e_1 et e_2 , est-ce que ces droites se coupent ?" Cette forme d'interrogation suppose que les interlocuteurs admettent tous deux la vérité de l'antécédent, ce qui réduit à deux le nombre de réponses directes possibles. On a la formalisation :

$$I = [\eta] \quad \eta \quad AB$$

$$\begin{cases} d(I)_1 = A \ \& \ B = \text{"Elles sont représentées correctement et elles se coupent"} \\ d(I)_2 = A \ \& \ \bar{B} = \text{"Elles sont représentées correctement et elles ne se coupent pas"} \end{cases}$$

On ne peut avoir comme réponse directe ni : $\bar{A} = \text{"Elles ne sont pas correctement représentées"}$, ni simplement : $B = \text{"Elles se coupent"}$ ou : $\bar{B} = \text{"Elles ne se coupent pas"}$.

Les questions conditionnelles doivent être distinguées d'énoncés interrogatifs de forme voisine. Nous avons vu l'énoncé :

$$I = \text{"Si J.K. est allé à Cracovie, est-ce qu'il a rendu visite à W. ?"} \\ = [\varepsilon] \quad \varepsilon \quad AB$$

L'énoncé, pourtant très proche :

$$I' = \text{"Est-ce que, si J.K. est allé à Cracovie, il a rendu visite à W. ?"} \\ = [\alpha] \quad \alpha \quad (A \rightarrow B)$$

n'est pas susceptible de la même formalisation, en particulier parce qu'il n'admet pas pour réponse directe : $\bar{A} = \text{"J.K. n'est pas allé à Cracovie"}$. Il n'est pas susceptible non plus de représentation par une question à antécédent non dirimant ; en effet, une traduction convenable paraît être :

$$I' = [\alpha] \quad \alpha \quad (A \rightarrow B)$$

$$\begin{cases} d(I')_1 = A \rightarrow B = \overline{(A \ \& \ \bar{B})} = \bar{A} \vee B \\ d(I')_2 = \overline{(A \rightarrow \bar{B})} = A \ \& \ \bar{B} \end{cases}$$

qui admet comme réponses directes :

Ainsi, la réponse directe positive ne correspond pas à ce qu'elle aurait été si l'on avait utilisé l'opérateur interrogatif [η].

De même, un énoncé tel que : "Est-ce que J.K. a cessé de fumer ?" peut être interprété de deux manières différentes. Si l'on pose :

A = "J.K. a fumé auparavant"

B = "J.K. fume actuellement"

une première interprétation serait :

$I = [\alpha] \quad \alpha \quad (A \ \& \ \bar{B})$

= "Est-ce que : J.K. a fumé auparavant et il ne fume pas actuellement?"

qui admet comme réponses directes :

$$\left\{ \begin{array}{l} d(I)_1 = A \ \& \ \bar{B} \\ d(I)_2 = (A \ \& \ \bar{B}) \end{array} \right. = \bar{A} \vee B = A \rightarrow B$$

Une seconde interprétation aurait la forme :

$I' = [\eta] \quad \eta \quad A\bar{B}$

= "Etant donné que J.K. a fumé auparavant, est-ce qu'il fume actuellement ?"

avec comme réponses possibles :

$$\left\{ \begin{array}{l} d(I')_1 = A \ \& \ \bar{B} \\ d(I')_2 = A \ \& \ B \end{array} \right.$$

Cette dernière interprétation est celle retenue par BELNAP (cf. § 2.2.5.b). La première interprétation s'en distingue en acceptant pour réponse directe : $d(I)_2 =$ "Il n'a jamais fumé auparavant, ou il fume toujours actuellement", ce qui permet de nier indirectement la présupposition en affirmant : "Si J.K. a fumé une fois, il fumera toujours".

Les questions sur la valeur logique ont pour forme [Kubiński 1970 : 32]:

$I = [p_1 \dots p_k] \quad A p_1 \dots p_k$

avec :

p_i = variable pouvant prendre l'une des valeurs *vrai* (V) ou *faux* (F)

$A p_1 \dots p_i$ = expression dans laquelle il n'y a pas d'autres variables libres que les p_i .

Une réponse directe à une question sur la valeur logique substituée à chaque p_i sa valeur (*vrai* ou *faux*), et peut être par exemple, pour :

$$I = [p_1 p_2 p_3 p_4 p_5] \quad A \quad p_1 p_2 p_3 p_4 p_5$$

$$d(I) = A \text{ VFFVF}$$

Si la question porte sur une question, deux cas peuvent se présenter : ou bien il s'agit d'une question numérique, et l'interrogation sur la valeur logique porte sur les individus représentés par les x_i . Par exemple :

$$I = [p] \quad k \quad x \quad A \quad p \quad x$$

a pour réponses directes possibles :

$$\begin{cases} d(I)_1 = A \text{ V } a_1 & \& \dots \& \quad A \text{ V } a_k \\ d(I)_2 = A \text{ F } a_1 & \& \dots \& \quad A \text{ F } a_k \end{cases}$$

Ou bien il s'agit d'une question dont l'opérateur est par exemple $[\delta^n]$, $[\varepsilon]$, ou $[\eta]$, et l'interrogation porte sur la valeur logique de l'opérateur.

A ce type de question sont apparentées les deux formes suivantes. La première est illustrée par l'énoncé [Kubiński 1970 : 33-34] : "Parmi les quatre phrases ci-après, quelles sont les deux seules phrases vraies : Mickiewicz est un poète, Słowacki est un poète, Chopin est un poète, Moniuszko est un poète ?". Sa forme générale est :

$$I = [\gamma_k^m] \quad \gamma_k^m \quad A_1, \dots, A_m \quad \text{avec : } 1 \leq k \leq m$$

= "Citez les k seules phrases vraies parmi les m suivantes : ..."

Dans le système de BELNAP, cette forme de question correspondrait à une question *whether* demandant une liste exhaustive de k éléments, qui s'écrirait :

$$I = ? \left(\binom{k}{k} \vee - \right) (A_1, \dots, A_m)$$

Les réponses directes correspondantes sont une conjonction de m expressions A_i , dont k sont affirmées, et $m-k$ sont niées. Par exemple :

$$I = [\gamma_1^2] \quad \gamma_1^2 \quad A, B$$

$$\begin{cases} d(I)_1 = A \ \& \ \bar{B} \\ d(I)_2 = \bar{A} \ \& \ B \end{cases}$$

Toute réponse directe à : $I = [\gamma_k^m] \quad \gamma_k^m \quad A_1, \dots, A_m$ est une réponse directe à : $I' = [\delta^m] \quad \delta^m \quad A_1, \dots, A_m$; mais la réciproque n'est évidemment pas vraie.

La deuxième forme apparentée aux questions sur la valeur logique est analogue à la forme précédente, moins l'exigence d'exhaustivité :

$$I = [\gamma_k^m] \quad \gamma_k^m \quad A_1, \dots, A_m \quad \text{avec : } 1 \leq k \leq m$$

= "Citez k phrases vraies parmi les m suivantes : ..."

Chez BELNAP, cette forme de question serait une question *whether* demandant exactement k exemples. Les réponses directes correspondantes sont une conjonction de k expressions A_i qui sont ainsi affirmées. Nous retrouverons cette forme de question dans la consigne : "Choisir k/m ", de COOMBS (cf. § 3.1.2.a).

Les *questions de causalité* peuvent, d'une certaine manière, être rattachées aux questions numériques. En effet, elles peuvent avoir pour opérateur interrogatif l'un quelconque des sept opérateurs numériques. Leur particularité réside dans l'emploi d'un foncteur à deux arguments, e , signifiant : "le premier argument est la cause (ou : la raison) du second". KUBIŃSKI distingue trois variétés de ce foncteur selon la nature des arguments. Soit par exemple l'interrogation [Kubiński 1970 : 37] :

$$I = \text{"Quelle est la cause du déclenchement de la première guerre mondiale ?"}$$

$$= 1 \ \Phi \quad e_1 \ \Phi \ a$$

avec :

Φ = variable d'expression

a = constante d'individu

= "le déclanchement de la première guerre mondiale"

e_1 = foncteur ayant pour ler argument une phrase, et pour second argument un nom.

La cause recherchée est une explication exprimée par une phrase A qui sera substituée à Φ pour donner la réponse directe :

$$d(I) = e_1 A a$$

= "Le fait que A est la cause de a "

La forme générale de la première variété de questions de causalité est :

$$I = \Omega \Phi e_1 \Phi a$$

dans laquelle :

$$\Omega = k, k \leq, k <, (k), (k \leq), (k <), \text{ ou } C, \quad \text{avec : } k \geq 1$$

La seconde variété de questions de causalité demande une ou plusieurs constantes d'individu pour expliquer un événement décrit par une proposition A . Par exemple :

$$I = C x e_2 x A$$

$$d(I) = e_2 a_1 A \& e_2 a_2 A \& \dots \& e_2 a_k A \quad \& \quad \forall x (e_2 x A \\ \rightarrow ((x = a_1) \vee \dots \vee (x = a_k)))$$

La troisième variété demande une ou plusieurs propositions pour expliquer un événement décrit par une proposition A . Par exemple :

$$I = l \leq \Phi e_3 \Phi A$$

$$d(I) = e_3 B_1 A \& \dots \& e_3 B_j A .$$

Cette analyse morphologique permet de décrire $7 \times 3 = 27$ formes différentes de questions de causalité. Mais, en l'absence de toute analyse sémantique, son intérêt nous paraît moindre que celui de l'analyse de BROMBERGER (cf. § 2.3.3.), partiellement reprise par BELNAP (cf. § 2.2.3.e).

c) Relations entre questions, réponses, et présuppositions.

Dans [Kubiński 1970], l'auteur développe sept systèmes théoriques qui constituent autant de "logiques des questions" : systèmes S, OR, E, T¹, T², T³, et T⁴ ; il examine en outre les relations existant entre ces sept formes distinctes de logique érotétique. Dans la perspective de l'application des concepts de la logique érotétique à la théorie des tests et des questionnaires, nous nous bornerons ici à présenter quelques définitions et quelques résultats relatifs aux types de réponse (système OR), aux présuppositions (systèmes T¹ à T⁴), et aux relations entre questions (systèmes S et E), sans aborder les problèmes théoriques qui leur sont liés.

La notion de *réponse directe* à une question est une notion fondamentale, à partir de laquelle peuvent être définies d'autres notions telles que celles de réponse indirecte, réponse partielle, etc. [Kubiński 1968 : 186]. Or, à partir de la formalisation de toute question, il est en principe toujours possible de dériver, en un nombre fini de transformations bien déterminées, les réponses directes à cette question [Kubiński 1968 : 185 ; 1970 : 12]. Par exemple [Kubiński 1968 : 186], dans :

I = "Citez une ville capitale d'un état européen"
 = lx Ax
 d(I) = "Varsovie est capitale d'un état européen"
 = Aa

la réponse directe citée est le résultat de deux transformations successives : effacement de l'opérateur interrogatif : lx , et substitution d'une constante : a = "Varsovie" à la variable x.

Les réponses directes aux questions numériques peuvent en principe prendre trois formes. Soit par exemple la question [Kubiński 1970 : 64-65] :

I = "Citez trois poètes polonais morts à l'étranger"
 = 3x Ax

On peut avoir :

$$d(I) = Aa_j \ \& \ Aa_k \ \& \ Aa_m \quad , \text{ avec : } j \neq k \neq m \neq j$$

= "Adam Mickiewicz est un poète polonais mort à l'étranger, et l'auteur de *Messire Thadée* est un poète polonais mort à l'étranger, et l'auteur de *Conrad Wallenrod* est un poète polonais mort à l'étranger".

Si l'on désire que deux des noms a_i ne puissent pas avoir le même référent (contrairement à l'exemple cité où les trois noms désignent la même personne), il faut adjoindre à la forme ci-dessus la *conjonction des négations d'identité* de tous les couples de constantes d'individu. On a par conséquent :

$$d(I) = Aa_j \ \& \ Aa_k \ \& \ Aa_m \quad \& \ \overline{(a_j = a_k)} \ \& \ \overline{(a_j = a_m)} \\ \& \ \overline{(a_k = a_m)}$$

Nous retrouvons donc ici l'affirmation de distinction de BELNAP. De même, pour les questions exigeant l'exhaustivité, l'affirmation de l'exhaustivité de la réponse doit figurer dans la réponse directe. Dans le cas d'une question de type C, nous avons par conséquent :

$$I = Cx \quad Ax$$

$$d(I) = Aa_1 \ \& \ \dots \ \& \ Aa_j \quad \& \ \overline{(a_1 = a_2)} \ \& \ \dots \ \& \ \overline{(a_{j-1} = a_j)} \\ \& \quad \forall x(Ax \rightarrow ((x = a_1) \vee \dots \vee (x = a_j)))$$

Les autres formes de réponse définies par KUBIŃSKI dans son système OR sont les réponses indirectes, les réponses presque directes, les réponses indirectes propres, et les réponses partielles (cf. [Kubinski 1970 : 66]). Une expression A est :

- une *réponse indirecte* à une question I si et seulement si : elle n'est pas une réponse directe, mais au moins une réponse directe peut être déduite de A (en d'autres termes : ssi A n'est pas une réponse directe, mais a pour conséquence au moins une réponse directe à I).

- une *réponse presque directe* à une question I si et seulement si : A est une réponse indirecte, et il existe une expression B qui est une

réponse directe et qui est telle que A et B s'impliquent réciproquement (en d'autres termes : ssi A est équivalente à une réponse directe à I).

- une *réponse indirecte propre* à une question I si et seulement si : A est une réponse indirecte mais n'est pas une réponse presque directe à I.

- une *réponse partielle* à une question I si et seulement si : A n'est ni une réponse directe ni une réponse indirecte à I, et il existe une expression B qui n'est elle aussi ni une réponse directe ni une réponse indirecte à I, et qui est telle que la conjonction A & B est soit une réponse directe, soit une réponse indirecte à I.

La notion de réponse en général est définie comme suit : *une expression A est une réponse à une question I relativement à une expression B* si de la conjonction A & B on peut déduire au moins une réponse directe à la question I (si A & B a pour conséquence au moins une réponse directe à I). L'expression B est alors une *présupposition* de la question I

Les expressions : "A peut être *déduit* de B", ou : "A est une *conséquence* de B", ou encore : "B *implique* A", doivent être comprises ici dans un sens restrictif. En effet, si l'on s'en tient aux règles classiques de la déduction, on aboutit à la conclusion que toute expression contradictoire est une réponse à une question de décision quelle qu'elle soit [Kubiński 1966 : 120]. De même, la question numérique : "Qui a découvert l'Amérique ?" admettrait comme réponse : "Il pleut, et en même temps il ne pleut pas" [Kubiński 1966 : 121], puisque de toute expression contradictoire peut être déduite (par exemple) la réponse directe : "(c'est) Christophe Colomb (qui) a découvert l'Amérique". C'est pourquoi on dira que A est une conséquence de B (ou que A peut être déduit de B, ou que B implique A) si et seulement si : A est une conséquence logique de B et B n'est pas une contradiction (cf. également [Kubiński 1970 : 48-49]).

Par ailleurs, on peut classer les réponses en réponses *positives* et réponses *negatives* [Kubiński 1958 : 319-320 ; 1966 : 121 ; 1970 : 67]. Ainsi, par exemple, A sera une réponse positive, et \bar{A} une réponse négative, à une question de la forme : $I = [\alpha] \quad \alpha A$. Plus généralement, si l'on a une question de décision $I = [\alpha] \quad \alpha A$, on dira que, relativement à une présupposition B, l'expression C est :

- une *réponse positive* à I si, de $B \& C$ on peut déduire A ,
- une *réponse négative* à I si, de $B \& C$ on peut déduire \bar{A} .

Il est aisé d'appliquer cette dichotomie aux réponses directes. Dans l'exemple cité, A est évidemment une réponse directe positive, et \bar{A} une réponse directe négative. Si nous considérons une question numérique telle que :

$I = \text{I} \leq x \quad Ax$, elle a pour réponses directes positives des expressions de la forme Aa_j , et pour unique réponse directe négative l'expression : $\forall x \quad \bar{Ax}$. La généralisation de cette dichotomie à l'ensemble des formes de réponses aux questions n'est qu'esquissée par l'auteur (dans [Kubiński 1958] et [Kubinski 1970]).

Il est possible de distinguer trois catégories de présuppositions des questions [Kubiński 1966 : 122-123] : les présuppositions pragmatiques, les présuppositions sémantiques, et les présuppositions syntaxiques. La *présupposition pragmatique interne* d'une question I , relativement à une personne P au moment T , est le fait que la personne P est persuadée, au moment T , qu'au moins une réponse à la question I est vraie [Kubiński 1966 : 123]. Par exemple, l'expression $\exists x \quad Ax$ est la présupposition pragmatique interne de $I = \text{I} \leq x \quad Ax$ relativement à la personne P au moment T si et seulement si, au moment T , P croit à la vérité de $\exists x \quad Ax$. Celui qui demande : "Qui a perdu la bataille de Waterloo ?" croit à la vérité de la phrase : "Quelqu'un a perdu la bataille de Waterloo" [Kubiński 1970 : 78-79]. La *présupposition pragmatique externe* d'une question I , posée au moment T_1 par une personne P_1 à une personne P_2 devant répondre au moment T_2 , est qu'au moment T_1 , la personne P_1 a la conviction qu'au moment T_2 , la personne P_2 sera à même de répondre à la question I [Kubiński 1966 : 123]. La *présupposition sémantique* d'une question I est qu'il existe au moins une réponse vraie à la question I [Kubiński 1966 : 123]. La *présupposition syntaxique primaire* d'une question I est analogue (mais non strictement équivalente) à la présupposition *Prés(I)* au sens de BELNAP (cf. § 2.2.5.b.). Sa définition varie selon le type auquel appartient la question. Les présuppositions syntaxiques *secondaires* [Kubiński 1970 : 93-94] sont toutes les propositions impliquées par la présupposition syntaxique primaire (Ces définitions sont inspirées de [Leonard 1957 : 34-40] et de [Belnap 1963 : 125]).

Question	Présupposition syntaxique		
<i>Questions simples :</i>			
$k \ x \quad Ax$	$\exists k \leq x \quad Ax$		
$k \leq x \quad Ax$	$\exists k \leq x \quad Ax$		
$k < x \quad Ax$	$\exists k < x \quad Ax$		
$(k) \ x \quad Ax$	$\exists k \ x \quad Ax$		
$(k \leq) \ x \quad Ax$	$\exists m \ x \quad Ax \ , \quad m \geq k$		
$(k <) \ x \quad Ax$	$\exists m \ x \quad Ax \ , \quad m > k$		
$C \ x \quad Ax$	$\exists 1 \leq x \quad Ax$		
<i>Questions complexes (avec n = 1, 2, ou 3) :</i>			
${}^n k / (m) \ xy \quad Axy$	$\exists k \leq x$	$\exists m \ y$	Axy
${}^n k \leq / C \ xy \quad Axy$	$\exists k \leq x$	$\exists 1 \leq y$	Axy
${}^n k < / (m) \ xy \quad Axy$	$\exists k < x$	$\exists m \ y$	Axy
${}^n (k) / (m) \ xy \quad Axy$	$\exists k \ x$	$\exists m \ y$	Axy
${}^n (k \leq) / m \leq \ xy \quad Axy$	$\exists p \ x$	$\exists m \leq y$	$Axy \ , \quad p \geq k$
${}^n (k <) / C \ xy \quad Axy$	$\exists p \ x$	$\exists 1 \leq y$	$Axy \ , \quad p > k$
${}^n C / (k) \ xy \quad Axy$	$\exists 1 \leq x$	$\exists k \ y$	Axy
${}^n (k) / C \ xy \quad Axy$	$\exists k \ x$	$\exists 1 \leq y$	Axy
${}^n (k \leq) / (m \leq) \ xy \quad Axy$	$\exists p \ x$	$\exists q \ y$	$Axy \ , \quad p \geq k, q \leq m$
${}^n (k <) / (m <) \ xy \quad Axy$	$\exists p \ x$	$\exists q \ y$	$Axy \ , \quad p > k, q < m$
${}^n (k \leq) / (m <) \ xy \quad Axy$	$\exists p \ x$	$\exists q \ y$	$Axy \ , \quad p \geq k, q < m$
${}^n (k \leq) / C \ xy \quad Axy$	$\exists p \ x$	$\exists 1 \leq y$	$Axy \ , \quad p \geq k$
${}^n C / (m \leq) \ xy \quad Axy$	$\exists 1 \leq x$	$\exists q \ y$	$Axy \ , \quad q \geq m$

Tableau n° 2.10 : Exemples de présuppositions syntaxiques primaires de questions numériques [Kubiński 1970 : 81-84].

La définition de la présupposition syntaxique primaire d'une question numérique fait appel à une extension du quantificateur existentiel, de la forme :

$\exists k x$ = "il existe k x tels que ..."

$\exists k \leq x$ = "il existe au moins k x tels que ..."

$\exists k < x$ = "il existe au moins $k + 1$ x tels que ..."

Soit I une question numérique ayant m variables libres sur lesquelles porte l'interrogation ; en règle générale, on dira qu'une expression A est la présupposition syntaxique primaire $Prés(I)$ de la question I si et seulement si A est dérivée de I par effacement des m opérateurs interrogatifs, et remplacement de chaque opérateur par un quantificateur existentiel (au sens large défini précédemment) [Kubiński 1970 : 80]. Le tableau n°2.10 présente quelques exemples de présuppositions syntaxiques primaires de questions numériques. On remarquera que des questions de formes différentes peuvent avoir la même présupposition ; tel est le cas par exemple de $I = \exists x Ax$, $I' = \exists x \leq Ax$, et $I'' = \exists x < Ax$, dont la présupposition commune est $Prés(I) = \exists x Ax$ [Kubiński 1970 : 83]. Pour les questions de type "est-ce que", la présupposition syntaxique primaire se présente comme une disjonction de l'ensemble des réponses directes possibles à la question (cf. tableau n°2.11).

Question	Présupposition
$[\alpha] \quad \alpha A$	$A \vee \bar{A}$
$[\beta^n] \quad \beta^n A_1, \dots, A_n$	$(A_1 \& \bar{A}_2 \& \dots \& \bar{A}_n) \vee (\bar{A}_1 \& A_2 \& \dots$ $\& \bar{A}_n) \vee \dots \vee (\bar{A}_1 \& \bar{A}_2 \& \dots \& \bar{A}_{n-1} \& A_n)$
$[\delta^n] \quad \delta^n A_1, \dots, A_n$	$(\bar{A}_1 \& \bar{A}_2 \& \dots \& \bar{A}_n) \vee (A_1 \& \bar{A}_2 \& \dots \& \bar{A}_n)$ $\vee (\bar{A}_1 \& A_2 \& \dots \& \bar{A}_n) \vee \dots \vee (\bar{A}_1 \& \bar{A}_2$ $\& \dots \& A_n) \vee (A_1 \& A_2 \& \dots \& \bar{A}_n) \vee \dots$ $\vee (A_1 \& A_2 \& \dots \& A_n)$
$[\varepsilon] \quad \varepsilon AB$	$(A\&B) \vee (A\&\bar{B}) \vee \bar{A}$
$[\eta] \quad \eta AB$	$(A\&B) \vee (A\&\bar{B})$

Tableau n°2.11 : Présupposition syntaxique primaire des questions de type "est-ce que" [Kubiński 1970 : 85-86].

Si l'on considère l'ensemble des questions constitué par les questions numériques et les questions de type "est-ce que", il faut noter que toute réponse directe $d(I)$ à une question I implique $Prés(I)$ [Kubiński 1970 : 80-85] ; et d'autre part que toute question I appartenant à cet ensemble a ou bien une présupposition syntaxique primaire unique, ou bien une infinité de présuppositions syntaxiques primaires [Kubiński 1970 : 82, 87]. Ce dernier cas se présente lorsqu'un opérateur interrogatif au moins d'une question numérique est de la forme : $(k \leq)$ ou : $(k <)$; en effet, toutes les expressions obtenues en remplaçant l'opérateur par : $\exists m$, avec : $m \geq k$ ou : $m > k$ selon le cas, sont des présuppositions syntaxiques primaires de la question. Cette propriété est un point important de divergence entre la conception de KUBIŃSKI et celle de BELNAP.

La notion de présupposition syntaxique d'une question est importante à plusieurs égards [Kubiński 1970 : 78] :

- cette notion peut servir de base à l'élaboration de systèmes déductifs de logique érotétique ;
- elle apporte une réponse au problème de l'existence (et de la définition) de la valeur de vérité des questions ;
- la présupposition d'une question peut toujours découler soit de l'ensemble des réponses directes à la question, soit au moins de certains sous-ensembles non vides de l'ensemble des réponses directes ;
- elle peut toujours être dérivée de la question à l'aide d'une transformation simple.

Dans certains cas, il peut être utile d'étendre la définition de la présupposition syntaxique primaire d'une question, en disant qu'une expression A est une *présupposition syntaxique primaire au sens large* d'une question I (numérique ou de type "est-ce que") s'il existe une expression B qui est une présupposition syntaxique (au sens strict) de I , et à laquelle A est logiquement équivalente [Kubiński 1970 : 86].

Les systèmes déductifs de logique érotétique T^1, T^2, T^3 , et T^4 sont fondés sur la notion de présupposition syntaxique primaire. Soit $p(I)$ l'ensemble des présuppositions syntaxiques primaires au sens strict de la question I ; $p(I)$ contient ou bien un élément unique $Prés(I)$, ou bien une infinité d'éléments. Le système T^1 définit, entre deux questions quelconques I et J , dont les ensembles de présuppositions syntaxiques primaires au sens strict sont respectivement $p(I)$ et $p(J)$, cinq types de relations [Kubiński 1970 : 87] :

① I est *présuppositionnellement équivalent* à J au sens de T^1 si et seulement si $p(I) = p(J)$.

② I est *contenu présuppositionnellement* dans J au sens de T^1 si et seulement si $p(I) \neq p(J)$ et $p(I) \subset p(J)$.

③ I *contient présuppositionnellement* J au sens de T^1 si et seulement si $p(I) \neq p(J)$ et $p(J) \subset p(I)$.

④ I et J sont *présuppositionnellement disjoints* au sens de T^1 si et seulement si $p(I) \cap p(J) = \emptyset$.

⑤ I et J sont *présuppositionnellement conjoints* au sens de T^1 si aucune des quatre relations précédentes entre I et J n'est vraie.

En remplaçant, dans les définitions qui précèdent, les ensembles de présuppositions syntaxiques primaires au sens strict $p(I)$ et $p(J)$ par les ensembles de présuppositions syntaxiques primaires au sens large $p'(I)$ et $p'(J)$, on obtient les définitions correspondantes du système T^2 .

Le système T^3 définit les cinq relations suivantes [Kubiński 1970 : 88] :

① I est *présuppositionnellement équipotent* à J au sens de T^3 s'il existe une application bijective de $p(I)$ sur $p(J)$ correspondant à une relation d'équivalence logique entre les éléments des ensembles $p(I)$ et $p(J)$.

② I est *présuppositionnellement plus faible* que J au sens de T^3 si et seulement si : I et J ne sont pas présuppositionnellement équipotents,

et il existe une application bijective de $p(J)$ sur $p(I)$ telle qu'à tout élément A de $p(J)$ corresponde un élément de $p(I)$ qui soit une conséquence de A .

③ I est *présuppositionnellement plus puissant* que J au sens de T^3 si et seulement si : J est *présuppositionnellement plus faible* que I et I n'est pas *présuppositionnellement plus faible* que J .

④ I est *présuppositionnellement complètement indépendant* de J au sens de T^3 si et seulement si : il n'existe pas d'application d'un sous-ensemble quelconque non vide de $p(I)$ sur $p(J)$ telle qu'à tout élément A de ce sous-ensemble corresponde dans $p(J)$ une conséquence de A , et il n'existe pas d'application d'un sous-ensemble quelconque non vide de $p(J)$ sur $p(I)$ telle qu'à tout élément B de ce sous-ensemble corresponde dans $p(I)$ une conséquence de B .

⑤ I est *présuppositionnellement partiellement indépendant* de J au sens de T^3 si et seulement si aucune des quatre relations précédentes entre I et J n'est vraie.

On obtient les définitions du système T^4 en remplaçant dans les définitions ci-dessus $p(I)$ et $p(J)$ par $p'(I)$ et $p'(J)$.

KUBIŃSKI définit deux autres systèmes de logique érotétique, à partir cette fois-ci des ensembles de réponses directes $(I)_0$ et $(J)_0$ aux questions I et J . Le système E [Kubiński 1970 : 69-71] présente certaines analogies formelles avec les systèmes T_1 et T_2 . Il définit les relations suivantes :

- ① I est *équivalent* à J si et seulement si $(I)_0 = (J)_0$.
- ② I est *contenu* dans J si et seulement si $(I)_0 \subset (J)_0$.
- ③ I *contient* J si et seulement si $(I)_0 \supset (J)_0$.

④ I et J sont *disjoints* si et seulement si $(I)_o \cap (J)_o = \emptyset$

⑤ I et J sont *conjointes* si et seulement si $(I)_o \cap (J)_o \neq \emptyset$

Le système S [Kubiński 1970 : 50-51] présente de même certaines analogies avec les systèmes T^3 et T^4 . Il définit les cinq relations :

① I est *équipotent* à J si et seulement si : de toute expression élément de $(P)_o$ peut être déduite une expression élément de $(Q)_o$, et réciproquement. Par exemple [Kubiński 1970 : 55-56], sont équipotentes les questions :

$$\begin{cases} I = [\alpha] & \alpha A \\ J = [\beta^2] & \beta^2 A, \bar{A} \\ K = [\delta^1] & \delta^1 A \end{cases}$$

② I est *plus faible* que J si et seulement si, de tout élément de $(J)_o$, peut être déduit un élément de $(I)_o$. Par exemple, si l'on a : $k < m$, la question I ci-dessous est plus faible que la question J [Kubiński 1970 : 54] :

$$\begin{cases} I = k \times Ax \\ J = m \times Ax \end{cases}$$

③ I est *plus puissant* que J si et seulement si J est plus faible que I et I n'est pas plus faible que J.

④ I est *complètement indépendant* de J si I n'est ni plus faible que J, ni plus puissant que J.

⑤ I est *partiellement indépendant* de J si et seulement si aucune des relations précédentes entre I et J n'est vraie. Tel est par exemple le cas, si l'on pose : $k+1 < m$, du couple de questions [Kubiński 1970 : 54]:

$$\begin{cases} I = k < x \quad Ax \\ J = m \leq x \quad Ax \end{cases}$$

Ces différents systèmes formels entretiennent des relations dont certaines font l'objet de théorèmes. Remarquons tout d'abord le parallélisme entre les cinq définitions dans les systèmes T^1, T^2, T^3, T^4, E , et S : la définition ① porte sur une relation réflexive, symétrique, et transitive ; les définitions ② et ③, sur une relation transitive ; les définitions ④ et ⑤, sur une relation symétrique (cf. [Kubiński 1970 : 51, 70, 91]). Les principaux théorèmes reliant certains de ces systèmes nous paraissent être [Kubiński 1970 : 70, 75, 91] :

- Si I est équivalent à J , alors I est équipotent à J .

- Si I est complètement indépendant de J , alors I et J sont disjoints.

- Si I est équivalent ou équipotent à J , alors A est une réponse directe (ou : indirecte, presque directe, indirecte propre, partielle) à I si et seulement si A est une réponse directe (ou : indirecte, presque directe, indirecte propre, partielle) à J .

- Si I est équivalent ou équipotent à J , alors A est une réponse à I relativement à B si et seulement si A est une réponse à J relativement à B .

- Si I et J sont présuppositionnellement équivalents au sens de T^1 (présupposition syntaxique primaire au sens strict), ils le sont également au sens de T^2 (présupposition syntaxique primaire au sens large).

- Si I et J sont présuppositionnellement équipotents au sens de T^3 (présupposition syntaxique primaire au sens strict), ils sont présuppositionnellement équivalents au sens de T^2 (présupposition syntaxique primaire au sens large).

Parmi les implications de ces systèmes formels, nous n'avons retenu que celles relatives à la valeur de vérité d'une question, et celles qui définissent la notion de négation d'une question. KUBIŃSKI traite le problème de la *valeur de vérité d'une question* d'une manière relativement différente de celle de BELNAP (cf. § 2.2.5.a), mais qui lui est logiquement équivalente, en posant qu'une question est vraie dans une interprétation M si et seulement si sa

présupposition syntaxique primaire est vraie dans M [Kubiński 1970 : 94]. Toutefois, cette règle ne peut s'appliquer sans modification aux questions numériques dont l'un au moins des opérateurs interrogatifs est de la forme : $(k \leq)$ ou : $(k <)$, puisque ces dernières ont une infinité de présuppositions syntaxiques primaires. D'où les définitions générales suivantes [Kubiński 1970 : 95] :

- une question est *vraie dans M* si et seulement si au moins une de ses présuppositions syntaxiques primaires est vraie dans M ;

- une question est *fausse dans M* si et seulement si toutes ses présuppositions syntaxiques primaires sont fausses dans M (c'est-à-dire : si et seulement si aucune de ses présuppositions syntaxiques primaires n'est vraie dans M).

La notion de *négation d'une question* est une extension possible du système S , que l'auteur définit de la manière suivante [Kubiński 1970 : 72]. Soit Ω un opérateur interrogatif quelconque ; on dira que :

- une question J est la négation d'une question I de forme : $I = \Omega A$ si et seulement si : J est de la forme : $J = \Omega \bar{A}$ et I est complètement indépendant de J ;

- une question J est la négation d'une question I de forme : $I = \Omega' A_1, \dots, A_n$ si et seulement si : J est de la forme : $J = \Omega' \bar{A}_1, \dots, \bar{A}_n$ et I est complètement indépendant de J .

A titre d'exemple, soit les couples formés d'une question I et sa négation \bar{I} [Kubiński 1970 : 73] :

$$\left\{ \begin{array}{l} I = I x \quad Ax \\ \quad = \text{"Quelle est l'équipe de football qui a gagné le match dimanche ?"} \\ \bar{I} = I x \quad \bar{Ax} \\ \quad = \text{"Quelle est l'équipe de football qui n'a pas gagné le match dimanche ?"} \end{array} \right.$$

$$\left\{ \begin{array}{l} I = C x \quad Ax \\ \quad = \text{"Parmi les proches collaborateurs du Général de Gaulle, quels sont ceux qui ont été Premier Ministre sous la cinquième République ?"} \end{array} \right.$$

$$\left. \begin{aligned} \bar{I} &= C \times \overline{Ax} \\ &= \text{"Parmi les proches collaborateurs du Général de Gaulle, quels sont} \\ &\quad \text{ceux qui n'ont pas été Premier Ministre sous la cinquième Répu-} \\ &\quad \text{blique ?"} \end{aligned} \right\}$$

Le problème de la détermination de la forme de la négation de la réponse n'est pas toujours résolu pour la plupart des variétés possibles de questions numériques (en particulier, lorsque : $k < 1$) ; tout au moins, l'existence d'une négation est-elle assurée pour les formes : $I = k \times Ax$, $I' = k \leq x Ax$, et $I'' = k < x Ax$, si A est une expression normale. Par contre, il est aisé de démontrer que les questions de type "est-ce que" ci-dessous *n'ont pas de négation* :

$$\left\{ \begin{aligned} I &= [\alpha] && \alpha \\ I' &= [\beta^2] && \beta^2 A, \bar{A} \\ I'' &= [\delta^n] && \delta^n A_1, \dots, A_n \end{aligned} \right.$$

En effet, pour toutes les questions de décision, toutes les questions conjonctives, et pour la forme particulière de question disjonctive équivalant à une question de décision, si l'on tente de construire la négation de la question en remplaçant tous les A_i par \bar{A}_i , on obtient une question équipotente à la question primitive [Kubiński 1970 : 72-73]. D'autre part, la négation d'une question disjonctive non équivalente à une question de décision existe toujours, et se formule sans difficulté.

2.3.3. L'analyse des questions de causalité de S. BROMBERGER

Nous avons vu (§ 2.2.3.d) une formalisation des questions de causalité s'inspirant des conceptions de Sylvain BROMBERGER. L'analyse faite par cet auteur (dans : [Bromberger 1966 a]) est en réalité plus complexe, puisqu'elle vise à déterminer les caractéristiques d'une explication de type scientifique.

La forme régulière d'une question de causalité est un énoncé interrogatif, introduit par l'adverbe *pourquoi* ; dont le verbe principal ne se réfère pas à une opinion, mais à un fait (ce qui élimine les questions du genre :

"Pourquoi estimez-vous que ... ?") ; et tel que, si l'on efface l'adverbe interrogatif *pourquoi*, l'énoncé résultant exprime une question *whether* à réponse par *oui* ou *non*. Cette dernière question est appelée la *question interne* de l'énoncé original, et la réponse affirmative à la question interne est la présupposition de la question de causalité. Par exemple :

- *question de causalité* : "Pourquoi le cuivre exposé à l'air verdit-il?"
- *question interne* : "Le cuivre exposé à l'air verdit-il ?"
- *présupposition* : "Le cuivre exposé à l'air verdit".

Une réponse satisfaisante à une question de causalité est un énoncé déclaratif P, introduit par la conjonction *parce que* ("Parce que P"). Le problème traité par BROMBERGER est le suivant. Soient deux propositions vraies, P et R ; quelles sont les conditions nécessaires et suffisantes pour que R soit une réponse satisfaisante à une question de causalité dont la présupposition est P ?

a) Insuffisances de la théorie de HEMPEL

Le modèle décrit par Carl G. HEMPEL sous le nom d'*explication déductive nomologique* peut constituer une solution au problème posé. Considérons par exemple l'expérience conduite par Florin PÉRIER à la demande de Blaise PASCAL, au Puy-de-Dôme, le 19 septembre 1648. Elle avait pour but de vérifier que, plus l'altitude d'un lieu est élevée, plus la hauteur de la colonne de mercure d'un baromètre de TORICELLI placé en ce lieu est faible. On peut décomposer les idées de PASCAL à ce sujet de la manière suivante [Hempel 1966 : 77] :

- Ⓐ En tout lieu, la pression exercée par la colonne de mercure de la branche fermée du baromètre est égale à la pression exercée sur la branche ouverte par la colonne d'air située au-dessus.
- Ⓑ Les pressions exercées par les colonnes d'air et de mercure sont directement proportionnelles au poids de ces colonnes ; le poids des colonnes d'air et de mercure est directement proportionnel à leur hauteur.
- Ⓒ Au cours de l'ascension du Puy-de-Dôme, la hauteur de la colonne d'air située au-dessus du baromètre diminuait progressivement.

Ⓓ (Donc :) Au cours de l'ascension du Puy-de-Dôme, la hauteur de la colonne de mercure dans la branche fermée du baromètre diminuait progressivement.

L'explication déductive nomologique apparaît par conséquent comme un raisonnement déductif, dont la conclusion est le phénomène à expliquer Ⓓ, et dont les prémisses sont soit des lois générales exprimant des concomitances régulières (de la forme : "Dans tous les cas où des conditions du genre F sont réalisées, des conditions du genre G sont également réalisées" [Hempel 1966 : 85]), comme Ⓐ et Ⓑ, soit des énoncés décrivant certaines circonstances particulières, comme Ⓒ. HEMPEL fait toutefois remarquer qu' "un problème qui suscite la recherche d'une explication ne détermine pas pour autant par lui-même le genre de découverte requis pour sa résolution" [Hempel 1966 : 83], cette découverte pouvant concerner par exemple les relations entre des théories existantes et des faits particuliers déjà connus, l'existence de faits particuliers entre inconnus, l'élaboration de théories explicatives nouvelles, ou la remise en cause de théories explicatives existantes.

Comme le souligne BROMBERGER, le modèle de l'explication déductive nomologique n'est pas applicable à l'ensemble des questions de causalité, car il n'énumère pas les conditions nécessaires et suffisantes à la vérité de toute proposition de la forme : "B explique A" [Bromberger 1966 a : 71-72]. Soit par exemple un pylône vertical supportant l'extrémité d'une antenne, dont l'autre extrémité est fixée au sol, en un point situé à 30 mètres du pied du pylône. La longueur de l'antenne est de 50 mètres. A partir du théorème de PYTHAGORE, nous pouvons en déduire que la hauteur du pylône est de 40 mètres. Cependant, le raisonnement correspondant n'explique par *pourquoi* sa hauteur est de 40 mètres. Pour l'auteur, la recherche d'explications scientifiques est essentiellement la recherche de réponses à des questions qui sont insolubles à partir des connaissances, des croyances, et des concepts généralement admis (c'est-à-dire des questions telles que toute personne qui possède ces connaissances, souscrit à ces croyances, et se limite à ces concepts, ne peut y apporter de réponse qui ne soit exposée à des objections décisives, ou qui ne fasse appel à des connaissances ou des moyens dont cette personne ne dispose pas [Bromberger 1966 a : 70]). Dans cette perspective, l'intérêt du modèle de HEMPEL est de montrer qu'une réponse

satisfaisante à une question de causalité doit avoir comme élément la conjonction des prémisses d'une explication nomologique déductive dont la conclusion est la présupposition de la question (Cette condition est nécessaire, mais non suffisante).

b) Règles générales et lois de déviance

La solution proposée par BROMBERGER au problème des questions de causalité se fonde naturellement sur une théorie de l'explication scientifique. Celle-ci fait appel aux concepts de règle générale, de loi de déviance générale ou particulière, de prédicat antonymique et de complément d'une règle générale. La définition précise de ces concepts et leur illustration est donnée ci-après.

Une *règle générale* est une expression vraie ou fausse de la forme :

$$\forall x \quad ((F_1x \ \& \ \dots \ \& \ F_jx) \ \rightarrow \ (S_1x \ \& \ \dots \ \& \ S_kx))$$

Soit par exemple :

$Fy =$ "y est un nom (commun) de la langue française"

$Ey =$ "y prend un s au pluriel"

$\forall y (Fy \rightarrow Ey) =$ "En français, tous les noms prennent un s au pluriel"

Une *loi de déviance générale* énumère les exceptions à la règle générale. C'est une expression vraie de la forme :

$$\begin{aligned} \forall x \quad ((F_1x \ \& \ \dots \ \& \ F_jx) \ \rightarrow \ ((\overline{E}x \ \leftrightarrow \ (A_1x \ \vee \ \dots \ \vee \ A_nx \ \vee \ B_1x \ \vee \ \dots \\ \vee \ B_mx \ \vee \ \dots \ \vee \ R_ex)) \ \& \ ((A_1x \ \vee \ \dots \ \vee \ A_nx) \ \leftrightarrow \ S_Ax) \\ \& \ ((B_1x \ \vee \ \dots \ \vee \ B_mx) \ \leftrightarrow \ S_Bx) \ \& \ ((R_1x \ \vee \ \dots \ \vee \ R_ex) \ \leftrightarrow \ S_Rx))) \end{aligned}$$

Soit, pour l'exemple précédent :

$A_1y =$ "y se termine en *al*, et y n'est pas un élément de la liste : *aval, bal, cal, carnaval, etc.*"

$A_2y =$ "y se termine en *ail*, et y est un élément de la liste : *bail, corail, émail, soupirail, etc.*"

$B_1y = "y \text{ se termine en } eu \text{ et } y \text{ n'est pas le mot } pneu"$.

$B_2y = "y \text{ se termine en } au"$

$B_3y = "y \text{ se termine en } ou, \text{ et } y \text{ est un élément de la liste : } bijoux, \text{ caillou, chou, genou, etc.}"$

$C_1y = "y \text{ se termine en } x"$

$C_2y = "y \text{ se termine en } z"$

$C_3y = "y \text{ se termine en } s"$

$S_Ay = "y \text{ au pluriel, substitue la finale } aux \text{ à } al \text{ ou à } ail"$

$S_By = "y \text{ au pluriel, prend un } x"$

$S_Cy = "y \text{ au pluriel, ne change pas}"$

Ce qui permet d'écrire la loi de déviance générale correspondante :

$$\forall y (Fy \rightarrow (\overline{Ey} \leftrightarrow (A_1y \vee A_2y \vee B_1y \vee B_2y \vee B_3y \vee C_1y \vee C_2y \vee C_3y)) \& ((A_1y \vee A_2y) \leftrightarrow S_Ay) \& ((B_1y \vee B_2y \vee B_3y) \leftrightarrow S_By) \& ((C_1y \vee C_2y \vee C_3y) \leftrightarrow S_Cy))$$

qui régit l'ensemble des exceptions à la règle générale du *s* final comme marque du pluriel des noms en français.

En outre, pour toute loi de déviance générale, les expressions suivantes doivent être vraies :

$$\textcircled{a} \quad \forall x ((F_1x \& \dots \& F_jx) \rightarrow (Ex \vee S_Ax \vee \dots \vee S_Rx))$$

(dans notre exemple : tout nom français au pluriel ou bien prend un *s*, ou bien substitue *aux* à *al* ou à *ail*, ou bien prend un *x*, ou bien ne change pas).

$$\textcircled{b} \quad \forall x ((A_1x \rightarrow (\overline{A_2x} \& \dots \& \overline{A_nx} \& \overline{B_1x} \& \dots \& \overline{R_ex})) \& (A_2x \rightarrow (\overline{A_1x} \& \overline{A_3x} \& \dots \& \overline{A_nx} \& \dots \& \overline{R_ex})) \& \dots \& (R_ex \rightarrow (\overline{A_1x} \& \dots \& \overline{A_nx} \& \dots \& \overline{R_{e-1}x})))$$

(disjonction exclusive ; dans notre exemple : aucun nom se terminant en *al* ne se

termine en *ail*, ni en *eu*, ni en *au*, ni en *ou*, ni en *x*, ni en *z*, ni en *s* ;...; aucun nom se terminant en *s* ne se termine en *al*, ni en *ail*, ni en *eu*, ni ..., ni en *z*).

Ⓒ La loi de déviance générale ne reste pas une proposition vraie à forme de loi si l'on supprime un membre (au moins) de la conjonction initiale (c'est-à-dire un $F_i x$), ou un membre (au moins) d'une des disjonctions intérieures.

Ⓓ La clôture de $F_1 x \& \dots \& F_j x$ n'est ni une tautologie, ni une contradiction.

Ⓔ Aucune des disjonctions intérieures n'a pour clôture une tautologie ou une contradiction.

Une *loi de déviance particulière* est une expression vraie à forme de loi, de formule générale :

$$\forall x (F_1 x \& \dots \& F_j x) \rightarrow (\overline{E} x \leftrightarrow (A_1 x \vee \dots \vee A_n x))$$

et satisfaisant à l'ensemble des conditions énoncées pour les lois générales de déviance. Toute loi de déviance générale est équivalente à une conjonction de lois de déviance particulières (et d'ailleurs toute loi de déviance particulière est également une loi de déviance générale, la distinction faite ici n'ayant pour but que de faciliter l'exposé ; cf. [Bromberger 1966 a : 85, n 12]). Dans notre exemple, une loi de déviance particulière pourrait être :

$$\forall y (F_y \rightarrow (S_B y \leftrightarrow (B_1 y \vee B_2 y \vee B_3 y))).$$

Un *prédicat antonymique* est soit le prédicat substitué à E dans une loi de déviance, soit l'un des prédicats S_A, \dots, S_R dans une loi de déviance générale, soit \overline{E} dans une loi de déviance particulière. Dans notre exemple de loi de déviance générale, l'ensemble des prédicats antonymiques est :

$\{E, S_A, S_B, S_C\}$; l'ensemble des prédicats antonymiques d'une loi de déviance particulière relative au même exemple serait : $\{E, \overline{E}\}$.

Une loi de déviance est le *complément d'une règle générale* si et seulement si cette règle est fautive, et peut être obtenue à partir de la loi de déviance

générale en supprimant les clauses "à moins que" (Pour cela, il suffit de remplacer par sa négation le prédicat qui précède immédiatement le premier signe \leftrightarrow ; puis d'effacer ce signe et tout ce qui suit ; et enfin d'ajuster les parenthèses pour obtenir une expression bien formée). On vérifie aisément que, dans notre exemple, les lois de déviance (générales ou particulières) sont bien des compléments de la règle : $\forall y (Fy \rightarrow Ey)$

c) Les réponses aux questions de causalité

A l'aide de ces définitions, il est maintenant possible de décrire la relation entre une réponse satisfaisante R à une question de causalité I, et la présupposition de cette question $Prés(I) = P$. Soit la question : I = "Pourquoi le pluriel de *cheveu* est-il *cheveux* (et non **cheveus*) ?". La présupposition $Prés(I)$ décrivant un fait d'observation est P = "le pluriel de *cheveu* est *cheveux*". Une réponse satisfaisante est l'assertion R = "*cheveu* se termine en *eu* au singulier".

Avec BROMBERGER, nous dirons que R est une réponse satisfaisante à la question de causalité I si et seulement si les quatre conditions ci-dessous sont vraies (cf. [Bromberger 1966 a : 77-78 ; 85-87 (n14)]) :

① Il existe une loi de déviance (générale ou particulière) telle que P soit une instantiation de l'un des prédicats antonymiques de L. Il est nécessaire que cette loi de déviance soit une loi empirique, afin d'éliminer les "explications" purement verbales, du type de la *virtus dormitiva* de l'opium. Dans notre exemple, L = "En français, tous les noms qui se terminent au singulier soit par *eu* (autres que *pneu*), soit par *au*, soit par *ou* à condition de faire partie de la liste : *bijou*, *caillou*, etc., prennent un *x* au pluriel" = $\forall y (Fy \rightarrow (S_B y \leftrightarrow (B_1 y \vee B_2 y \vee B_3 y)))$. Si l'on pose : $c = \text{"cheveu"}$, la présupposition s'écrit : $P = S_B c$, et la réponse : $R = B_1 c$.

② Il existe une explication déductive nomologique D, qui compte L et R parmi ses prémisses, et dont la conclusion est P. En abrégé, cette explication serait de la forme : "En français, tous les noms communs qui, au singulier, se terminent en *eu*, sauf *pneu*, prennent un *x* au pluriel ; *cheveu* se

termine en *eu* et est distinct de *pneu* ; d'où : *cheveu* fait au pluriel *cheveux*".

③ Il existe un raisonnement ayant l'apparence d'une explication déductive nomologique D' dont les prémisses sont l'ensemble des prémisses de D , plus la loi G dont L est un complément, moins L et R ; raisonnement dont la conclusion est l'un des contraires de P . La différence entre D' et une explication déductive nomologique tient à ce que l'une des prémisses à forme de loi et la conclusion de D' sont fausses. Dans notre exemple, la règle générale : $G =$ "En français, tous les noms prennent un *s* au pluriel" = $\forall y (Fy \rightarrow Ey)$, conduit à la conclusion fautive : $Ec =$ "*cheveu* fait au pluriel *cheveus*".

④ La règle générale G dont L est un complément est telle que, si l'on supprime l'un des éléments conjoints de l'antécédent, la nouvelle règle ainsi obtenue ne peut plus être complétée par une loi de déviance (condition de minimalité).

Il est possible de caractériser une explication déductive nomologique à l'aide des mêmes concepts (cf. [Bromberger 1966 a : 87 (n14)]). En reprenant l'exemple de HEMPEL déjà utilisé, soit la question $I =$ "Pourquoi le mercure descend-il dans le baromètre de TORICELLI lorsque l'on s'élève en altitude ?". Une explication déductive nomologique D est une réponse à la question I si et seulement si D a pour prémisses une règle générale vraie G , et un ou plusieurs énoncés O_i décrivant des faits, et pour conclusion la présupposition $Prés(I)$. Ici, la règle générale est la conjonction des lois (a) et (b) : $G =$ "En tous lieux, il y a égalité des pressions d'air et de mercure dans le baromètre et la pression est directement proportionnelle à la hauteur des colonnes d'air et de mercure". L'énoncé relatif à un fait correspond à (c) : $O =$ "En un lieu donné, la hauteur de la colonne d'air qui est au-dessus est inversement proportionnelle à l'altitude du lieu". Enfin, la conclusion de D , qui est aussi la présupposition de la question, est la description des observations de Florin PÉRIER : $Prés(I) =$ "La hauteur de la colonne de mercure dans le baromètre est inversement proportionnelle à l'altitude du lieu où il est placé".

Nous l'avons vu, BROMBERGER ne considère toutefois pas une telle explication comme une réponse à une question de causalité proprement dite. En

effet : "En général, une question se pose chaque fois qu'il y a une raison de croire qu'elle possède une réponse, bien que cette réponse ne soit pas connue. Dans le cas des questions de causalité, ceci se produit lorsque quelqu'un estime que la présupposition est vraie, la considère comme un écart à une règle générale, et pense que l'on peut généraliser par une loi de déviance les conditions dans lesquelles les écarts à la règle générale se produisent" [Bromberger 1966 a : 80]. Une question de causalité ne peut avoir de réponse satisfaisante lorsque :

- la présupposition est fautive (tout au plus peut-il y avoir une réponse corrective niant la présupposition) ;

- la présupposition est vraie, mais l'observation qu'elle décrit ne constitue pas un écart à une règle générale (une réplique possible est de la forme : "Mais c'est toujours le cas !") ;

- les conditions dans lesquelles se produisent les écarts à la règle générale ne peuvent être généralisées (les répliques que l'on obtient alors sont du type : "Cela arrive parfois").

D'autre part, il semble bien que les réponses aux questions de causalité qui se révèlent les plus fécondes sont justement celles qui énoncent de nouvelles règles générales ; c'est-à-dire celles qui ne constituent pas des réponses satisfaisantes au sens défini ci-dessus. Toutefois, cette formalisation des réponses aux questions de causalité correspond à une étape du développement des théories, se traduisant par une remise en cause des règles générales afin de les amender par des lois de déviance, en attendant de les remplacer par de nouvelles règles.

d) Insuffisances de la théorie de BROMBERGER

Paul TELLER a signalé deux difficultés dans l'analyse de BROMBERGER, sans d'ailleurs leur apporter de solution satisfaisante [Teller 1974]. La première difficulté est qu'il est possible d'énoncer des lois de déviance telles qu'on ne puisse en tirer des réponses satisfaisantes aux questions de causalité. Soit par exemple la loi de déviance particulière déjà citée, relative au pluriel

des noms en *eu*, *au* et *ou* :

$$L = \forall y (Fy \rightarrow (S_B y \leftrightarrow (B_1 y \vee B_2 y \vee B_3 y)))$$

Nous pouvons, pour alléger l'écriture, poser :

$$By = B_1 y \vee B_2 y \vee B_3 y$$

$$L = \forall y (Fy \rightarrow (S_B y \leftrightarrow By))$$

Selon l'analyse de BROMBERGER, $B_c = "$ *cheveu* est un nom qui au singulier se termine soit par *eu* (et il est alors différent de *pneu*), soit par *au*, soit par *ou* (et il fait alors partie de la liste : *bijou*, etc.)" est une réponse satisfaisante à toute question de causalité I dont la présupposition est $Prés(I) = S_B c = "$ *cheveu* prend un x au pluriel". Compte-tenu de la propriété de symétrie de la relation d'équivalence, on peut en déduire une nouvelle loi de déviance :

$$L' = \forall y (Fy \rightarrow (By \leftrightarrow S_B y))$$

et dire alors que $S_B c$ est une réponse satisfaisante à toute question de causalité ayant pour présupposition B_c ("Pourquoi *cheveu* se termine-t-il par *eu* au singulier ?" – "Parce qu'il prend un x au pluriel"). Cela suppose évidemment une règle générale G' portant non plus sur la formation du pluriel, mais sur la finale des noms français au singulier.

TELLER cite un grand nombre de contre-exemples construits sur ce modèle, dont certains reprenant des exemples développés par BROMBERGER. Ainsi, des lois du pendule, on peut déterminer la période d'oscillation p d'un corps pesant suspendu à une corde, si l'on connaît la longueur l de la corde. La loi de déviance est de la forme : $L = \forall x ("x$ est un pendule" $\rightarrow ("x$ a une période $p" \leftrightarrow "x$ a une longueur $l"))$. Mais si l'on considère la période d'oscillation comme une propriété constitutive du pendule, on peut également poser : $L' = \forall x (x$ est un pendule" $\rightarrow ("x$ a une longueur $l" \leftrightarrow "x$ a une période $p"))$. BROMBERGER écarte *a priori* cette deuxième interprétation, en déclarant que la période d'oscillation ne peut être une réponse à la question : "Pourquoi la corde du pendule a-t-elle l mètres de long ?" parce que la période n'existerait pas sans la longueur, alors que la longueur existe, que le pendule oscille ou non. Cet

argument tombe lorsque l'on considère la période d'oscillation comme une propriété inhérente au pendule, qu'il se trouve effectivement en train d'osciller ou non.

Cette difficulté semble liée à l'impossibilité de fournir un équivalent formel à la notion de *loi naturelle*. Selon TELLER, pour définir une *expression à forme de loi*, il faudrait faire appel à certains éléments contextuels comme par exemple la distinction entre variables dépendantes et variables indépendantes, ou entre ce qui est manipulable par l'homme et ce qui ne l'est pas. "Comme chacun sait, les développements naïfs de ces suggestions échouent lamentablement. Néanmoins, de tels exemples renforcent l'idée qu'avant que nous puissions espérer rendre compte de façon adéquate de l'explication [scientifique] ou des réponses aux questions de causalité, nous devons prendre en compte des éléments contextuels qui n'ont pas encore été clairement identifiés" [Teller 1974 : 379].

La seconde difficulté signalée par TELLER est qu'il est possible de transformer la plupart des contre-exemples utilisés par BROMBERGER pour montrer l'insuffisance de la théorie de HEMPEL, en contre-exemples révélant l'insuffisance de la théorie de BROMBERGER. En simplifiant sensiblement, la procédure décrite par TELLER se ramène à :

- choisir une explication déductive nomologique simple ayant pour conclusion une proposition E_d , et pour prémisses une règle générale G et une proposition décrivant les circonstances particulières C_d ;

- élaborer une loi de déviance particulière ayant E parmi ses prédicats antonymiques, de la forme :

$$L = \forall x ((F_1x \ \& \ \dots \ \& \ F_jx) \ \rightarrow \ (Ex \ \leftrightarrow \ (A_1x \ \vee \ \dots \ \vee \ A_nx)))$$

(il est toujours possible d'élaborer une telle loi satisfaisant aux conditions de BROMBERGER, en augmentant le nombre des prédicats conjoints dans le premier terme de l'implication [Teller 1974 : 376]) ;

- transformer l'expression à droite du signe équivalence en y ajoutant la condition C_x , en unissant chaque A_i avec \bar{C}_x , et en éliminant les

$(\bar{C}_x \ \& \ A_j)$ qui sont des contradictions, ce qui donne :

$$L' = \forall x ((F_1x \ \& \ \dots \ \& \ F_jx) \ \rightarrow \ (Ex \ \leftrightarrow \ (Cx \ \vee \ (\bar{C}_x \ \& \ A_1x) \ \vee \ \dots \ \vee \ (\bar{C}_x \ \& \ A_kx))))$$

Si G ne contient pas de constante et est vraie, et si L est vraie, alors L' est vraie également.

Reprenons l'exemple de la hauteur du pylône, et supposons avec BROMBERGER qu'il existe une loi de déviance :

$$L = \forall x ("x \text{ est un pylône}" \ \rightarrow \ ("x \text{ a 40 mètres}" \ \leftrightarrow \ (A_1x \ \vee \ \dots \ \vee \ A_nx)))$$

en posant :

$$Cx = "Au sommet de x est fixée une antenne de 50 mètres dont l'autre extrémité est fixée au sol en un point situé à 30 mètres du pied de x "$$

et en appliquant la procédure décrite ci-dessus, on obtient :

$$L' = \forall x ("x \text{ est un pylône}" \ \rightarrow \ ("x \text{ a 40 mètres}" \ \leftrightarrow \ (Cx \ \vee \ (\bar{C}_x \ \& \ A_1x) \ \vee \ \dots \ \vee \ (\bar{C}_x \ \& \ A_kx))))$$

qui satisfait aux conditions énoncées par BROMBERGER relativement aux réponses aux questions de causalité. Ainsi, selon la théorie de BROMBERGER, Cx est une réponse satisfaisante à la question : "Pourquoi x a-t-il 40 mètres ?".

En conclusion, malgré les progrès qu'il représente par rapport à la théorie de HEMPEL, le modèle de BROMBERGER présente encore de sérieuses insuffisances. Cependant, son intérêt réside dans le fait qu'il constitue la meilleure formalisation des réponses aux questions de causalité qui ait été faite, à notre connaissance du moins. Sur le plan pratique, il est possible de ne pas suivre BROMBERGER dans sa définition de la réponse suffisante à une question de causalité, et de lui préférer la forme plus complète que lui donne BELNAP en s'inspirant de BROMBERGER [Belnap 1976 : 85-86]. Ainsi, dans l'exemple relatif au pluriel du mot *cheveu*, on considèrera $R = B_1c$ comme une réponse

codée, mise pour :

$$R' = F_c \ \& \ B_{1c} \ \& \ \forall y (F_y \rightarrow (S_{By} \leftrightarrow (B_{1y} \vee B_{2y} \vee B_{3y})))$$

= "cheveu est un nom de la langue française" & "cheveu se termine au singulier en *eu* et n'est pas le mot *pneu*" & "Aucun nom de la langue française ne prend un *x* au pluriel, à moins qu'il ne se termine en *eu*, etc."

D'autre part, en posant qu'une loi de déviance est toujours vraie, BROMBERGER pose *ipso facto* qu'une réponse satisfaisante à une question de causalité est toujours vraie. Cependant, il définit ailleurs l'ensemble des modalités de réponse présentées par ce type de question comme l'ensemble des énoncés déclaratifs qui satisfont aux conditions d'une réponse satisfaisante, mais qui peuvent être vrais ou faux [Bromberger 1966 b : 604]. C'est pourquoi nous considérerons que toute réponse directe à une question de causalité I, ayant pour présupposition $Prés(I) = S_c$, est de la forme :

$$d(I) = (F_{1c} \ \& \ \dots \ \& \ F_{jc}) \ \& \ A_{ic} \ \& \ \forall x ((F_{1x} \ \& \ \dots \ \& \ F_{jx}) \rightarrow (S_x \leftrightarrow (A_{1x} \ \vee \ \dots \ \vee \ A_{ix} \ \vee \ \dots \ \vee \ A_{nx})))$$

dans laquelle les termes conjoints peuvent être vrais ou faux.

2.3.4. Survol de quelques autres formes de logique érotétique.

Les quatre auteurs qui font ci-après l'objet d'une présentation de quelques pages, nous ont paru mériter mieux qu'une brève mention dans l'historique que nous avons dressé de la logique érotétique (§ 2.1.3.b), mais moins qu'un exposé technique détaillé. Rappelons toutefois que le mode de présentation que nous avons adopté ne permet pas de conclure qu'il s'agit ici d'auteurs

mineurs. En effet, il nous a paru plus commode de centrer la description détaillée des acquis actuels de la logique érotétique sur un auteur particulier (BELNAP en l'occurrence), et de présenter les autres théories par rapport à la première théorie décrite ; cette manière de procéder introduit par conséquent, si l'on s'efforce d'éviter la redondance, un effet de perspective dont sont victimes les auteurs cités en dernier. En utilisant métaphoriquement le langage de l'analyse factorielle en composantes principales, nous pourrions dire que, pour résumer l'ensemble des apports possibles de la logique érotétique à une théorie du questionnaire, nous avons choisi de faire coïncider le premier axe factoriel avec le système développé par BELNAP. Ce qui nous paraît remarquable, c'est qu'après avoir présenté, comme deuxième et troisième axes, les conceptions d'ÅQVIST et de KUBIŃSKI (en considérant BROMBERGER comme l'auteur d'une théorie locale plus que d'un véritable système), il nous reste encore à décrire autant d'apports techniques ou notionnels originaux tirés des travaux d'AJDUKIEWICZ, de LÖW, de STAHL, et de HARRAH.

a) Kazimierz AJDUKIEWICZ

Nous avons signalé (§ 2.1.3.b) que Kazimierz AJDUKIEWICZ a publié, dès 1923, un article sur la question de type : "Qu'est-ce que P ?", et qu'il a proposé, dès 1926, une analyse sémantique de l'expression interrogative. Toutefois, nous nous basons, pour décrire les grandes lignes de sa conception, sur son dernier traité de logique pragmatique, d'ailleurs inachevé et publié à titre posthume [Ajdukiewicz 1965].

La formalisation d'AJDUKIEWICZ s'applique essentiellement aux questions *which* ; s'il est vrai que l'on pourrait l'adapter à d'autres formes de questions (à réponse *oui/non* par exemple), l'auteur convient que cette extension paraîtrait assez artificielle [Ajdukiewicz 1965 : 87]. Le schéma de base d'une question *which* est une forme propositionnelle unaire (appelée *datum quaestionis*), précédée d'un marqueur interrogatif qui indique quelle est l'in-

connue de l'interrogation, et, éventuellement, quel est son domaine (nous l'avons vu, cette formalisation a été adoptée dans ses principes par KUBIŃSKI pour les questions numériques. cf. § 2.3.2.a, et [Kubinski 1958 : 320]). Par exemple [Ajdukiewicz 1965 : 85-87], dans l'énoncé : "Qui a découvert l'Amérique ?", on a la forme propositionnelle : " x a découvert l'Amérique", et le marqueur "Qui", indiquant que l'interrogation porte sur x , et que cette variable prend ses valeurs dans le domaine des noms de personnes humaines. La forme propositionnelle d'une question *which* véhicule de l'information [Ajdukiewicz 1965 : 88], et détermine dans une certaine mesure la forme de la réponse désirée [Ajdukiewicz 1965 : 85]. Cette analyse, proche du langage naturel, débouche toutefois sur diverses notions importantes que nous avons présentées à propos d'autres auteurs : contenu propositionnel de l'interrogation, condition catégorielle régissant l'inconnue sur laquelle porte la question, forme de la réponse, etc.

L'auteur est conduit à distinguer divers types de réponses aux questions *which* [Ajdukiewicz 1965 : 86-91] :

- *réponse propre*, obtenue en remplaçant, dans la forme propositionnelle de la question, la variable par une constante appartenant au domaine indiqué par le marqueur interrogatif.

- *réponse complète*, qui soit est une réponse propre (*réponse directe complète*), soit a pour conséquence une réponse propre (*réponse indirecte complète*).

- *réponse partielle*, qui, n'étant pas une réponse complète, permet cependant d'éliminer certaines des réponses propres envisageables (exemple de réponse partielle : $I = \text{"Qui a découvert l'Amérique ?"} / r(I) = \text{"Un italien"}$). On peut assimiler une réponse partielle à une réponse indiquant l'intervalle dans lequel la valeur cherchée d'un paramètre doit se trouver [Ajdukiewicz 1965 : 378-379] ; la valeur exacte de ce paramètre n'étant donnée que par une réponse complète.

- réponse *exhaustive*, réponse vraie ayant pour conséquence toutes les réponses propres vraies.

- réponse *annulant la présupposition positive* de la question, réponse non propre niant l'existence d'au moins une réponse propre à la question (exemple : I = "Qui est le fils de Copernic ?" r(I) = "Copernic n'avait pas de fils").

- réponse *annulant la présupposition négative* de la question, réponse indirecte complète exprimant qu'il n'existe pas de réponse fausse à la question.

En outre, l'auteur propose sans les développer les définitions suivantes, relatives aux questions [Ajdukiewicz 1965 : 88-89, 92-94] :

- question *posée improprement*, question dont la présupposition positive (il existe au moins une réponse vraie) ou la présupposition négative (il existe au moins une réponse fausse) n'est pas vraie. Les questions de décision à réponse *oui/non* sont toujours posées proprement.

- question *suggestive*, question posée non pour obtenir de l'information, mais pour en apporter à l'interlocuteur (ex. : I = "Qui Jean a-t-il épousé ? ", posée à quelqu'un qui ignore que Jean est marié).

- question *de mauvaise foi (mala fide)*, question suggestive dans laquelle l'information apportée par le contenu propositionnel est fausse.

- question *insidieuse*, question destinée soit à provoquer chez l'interlocuteur une réponse en contradiction avec ses réponses antérieures, soit à lui faire révéler une information qu'il entendait cacher.

- question *didactique*, question dont le locuteur connaît la réponse, utilisée à des fins pédagogiques (méthode érotématique), et pour laquelle la découverte de la réponse correcte est soit un moyen relativement secondaire de la formation (*question didactique incitative*), soit l'objectif même de la formation (*question didactique primaire*).

b) Friedrich LÖW

L'article de LÖW sur la logique de la question, publié en 1928, compte quatre-vingt pages. A notre connaissance, et sous réserve d'une analyse ultérieure des textes d'AJDUKIEWICZ et d'INGARDEN signalés au § 2.1.3.b, cet article constitue la première présentation complète et détaillée d'une théorie sur la logique des questions. Son intérêt toutefois n'est pas seulement historique : la conception que l'auteur développe, les propriétés qu'il décrit, et les théorèmes qu'il énonce, ne semblent pas avoir été repris, du moins dans leur ensemble, par les autres spécialistes de la logique érotétique. Aussi, cinquante ans après sa publication, cette théorie paraît-elle encore profondément originale. Nous résumons dans ce paragraphe les points les plus importants concernant la nature des questions, les divers types de questions, les présupposés, les formes de réponses, le problème de la solubilité des questions, et les relations entre questions.

KUBIŃSKI a qualifié la conception de LÖW de "psychologique" [Kubiński 1970 : 97]. En effet, celui-ci considère l'acte d'interroger avant tout comme une quête de la réalité, qui ne se réduit ni à la logique du jugement (fût-il hypothétique), ni au seul désir de connaître. L'interrogation est une forme de pensée tendue vers la recherche d'un fait matériel existant, fait matériel qui est déterminé de façon univoque par la signification même de la question [Löw 1928 : 435 §3]. En elle-même, la question pose que l'objet qu'elle recherche existe réellement, non sous la forme d'une supposition ou d'une hypothèse, mais en tant qu'affirmation implicite contenue en elle [Löw 1928 : 376]. Toutes les formes de questions recherchent un fait matériel G , et posent implicitement l'assertion : " G existe" [Löw 1928 : 384]. Mais alors qu'à une question donnée correspond un fait matériel unique, à un fait matériel donné correspond une infinité de questions possibles [Löw 1928 : 401].

Il y a en tout et pour tout deux catégories de questions : les questions de décision, et les questions à compléter. Cette distinction logique (elle ne peut être réduite à une distinction purement grammaticale [Löw 1928 : 415]) est fondée sur la manière dont la question détermine le fait matériel recherché [Löw 1928 : 435 § 6]. Une *question de décision* propose le choix entre

plusieurs propositions décrivant un état de fait, dont une et une seulement est vraie ; le nombre des propositions présentées est fini, et au moins égal à deux [Löw 1928 : 405]. Bien que les différentes espèces de questions de décision soient en nombre illimité (en raison du nombre illimité des conceptions possibles d'états de fait, cf. [Löw 1928 : 435 § 7]), on pourrait, à partir des exemples cités par l'auteur, distinguer celles qui proposent un choix entre une proposition et sa négation (ex. : "Le bleu est-il une couleur ?" ; "Napoléon est-il le vainqueur d'Iéna, ou non ?" [Löw 1928 : 380, 393]), et celles qui proposent un choix entre plusieurs propositions incompatibles, mais non nécessairement contradictoires (ex. : "L'or est-il jaune, ou vert ?" ; "Est-ce Mercure, Vénus, ou Jupiter, qui est la plus grande planète ?" [Löw 1928 : 404, 414]). Une *question à compléter* présente une ébauche de description d'un état de fait existant, à laquelle il manque un élément pour que la description soit complète [Löw 1928 : 384] ; la question apporte en outre la définition de l'ensemble (du groupe, de la classe, de la catégorie, etc.) dont l'élément recherché fait partie [Löw 1928 : 389-390]. Par exemple, dans la question : "Quand l'Amérique a-t-elle été découverte ?", le mot interrogatif : "quand" indique que l'élément recherché est un repère temporel [Löw 1928 : 378]. Dans : "Quelle est la couleur de l'or ?", le substantif : "couleur" détermine l'univers sémantique auquel appartient le mot recherché [Löw 1928 : 389-390]. On peut distinguer deux espèces de questions à compléter, selon le rôle joué, dans la description du fait matériel recherché, par l'élément qui complète cette description [Löw 1928 : 406-408] ; ces deux espèces n'ont pas été dénommées par l'auteur. La première est celle dans laquelle la question porte soit sur le sujet, soit sur le prédicat ; exemples : "Qui est omniscient ?" ; "Qu'est-ce que le rouge ?". La seconde espèce regroupe les questions portant sur un caractère accessoire de la description ; exemple : "Quand l'Amérique a-t-elle été découverte ?". La première espèce correspond approximativement aux mots interrogatifs : "qui ?", "que ?" ; "quoi ?" ; la seconde, aux mots : "quand ?", "comment ?", "combien ?", "où ?", etc. Nous allons voir que l'étude des présupposés des questions permet de mieux comprendre cette distinction au sein des questions à compléter.

Toute question pose au moins une assertion implicite, c'est-à-dire un présupposé ; ce présupposé est lié à la question par une relation d'implication,

analogue à celle qui lie deux propositions [Löw 1928 : 377-378]. L'existence d'au moins un présupposé est une nécessité logique : c'est par ses présupposés que la question s'ancre dans la réalité objective [Löw 1928 : 384]. La nature des propositions impliquées par les questions dépend du type de la question. Les *questions de décision* ont pour unique présupposition qu'un et un seulement des faits matériels décrits dans la question est vrai [Löw 1928 : 381]. Les *questions à compléter* peuvent avoir des présupposés de quatre types différents et hiérarchisés. Soit par exemple la question : "Quand le continent Amérique a-t-il été découvert ?". Parmi les propositions impliquées par cette interrogation, on peut citer, dans l'ordre hiérarchique croissant [Löw 1928 : 376-379] :

- $$\left\{ \begin{array}{l} J_1 = \text{"L'Amérique est un continent"} \\ J_2 = \text{"Le continent Amérique a été découvert"} \\ J_3 = \text{"Le continent Amérique a été découvert à une certaine date"} \\ J_4 = \text{"Le fait matériel recherché par la question existe bien"} \end{array} \right.$$

Toute question à compléter implique nécessairement une proposition de type J_3 et une de type J_4 ; en outre, les questions à compléter de la deuxième espèce, et elles seules, impliquent également une proposition de type J_2 [Löw 1928 : 406-407], comme dans l'exemple ci-dessus. A titre de contre-exemple, la question à compléter de première espèce : "Qui est omniscient ?" implique les propositions : $J_3 = \text{"Quelqu'un est omniscient"}$, et : $J_4 = \text{"Ce quelqu'un existe réellement"}$, mais n'implique pas de proposition de type J_2 .

Lorsque l'état de fait recherché par la question existe, la proposition qui exprime cet état de fait est vraie, et constitue *la réponse* à la question. En tout état de cause, seule une proposition vraie peut être considérée comme une réponse à une question [Löw 1928 : 387]. Une proposition fausse ne peut être qu'un *simulacre de réponse* ; mais il peut évidemment arriver qu'une proposition vraie donnée ne réponde pas à une question donnée : cette proposition est dite alors *étrangère à la question* [Löw 1928 : 391]. Une proposition constitue une réponse à une question de décision si l'état de fait dont elle affirme l'existence est l'un des états de fait décrits par la question. Elle

constitue une réponse à une question à compléter si l'état de fait dont elle affirme l'existence non seulement est conforme à la description incomplète qu'en donne la proposition J_2 ou J_3 (selon que la question est de la seconde ou de la première espèce), mais en outre comporte un élément appartenant à la catégorie définie par la question comme objet de l'interrogation [Löw 1928 : 389-390].

Toute proposition vraie (affirmant l'existence d'un état de fait donné) est une réponse à au moins une question (celle qui vise l'état de fait décrit par la proposition) ; puisqu'à un fait donné correspond une infinité de questions possibles distinctes, toute proposition vraie est une réponse à une infinité de questions différentes. Par exemple, la proposition : "Le continent Amérique a été découvert en 1492" est une réponse à : "Quel continent a été découvert en 1492 ?", à : "Est-ce en 1492 ou en 1592 que le continent Amérique a été découvert ?", etc. [Löw 1928 : 392-393]. Par contre, *toute question a au plus une réponse*. Il peut évidemment y avoir des questions sans réponse : c'est le cas lorsque l'état de fait recherché n'existe pas (ex. : "Quel est le plus grand nombre premier ?" [Löw 1928 : 392]). Mais il ne peut, par définition, y avoir de question ayant plusieurs réponses, bien que les imprécisions du langage aient pu faire naître l'illusion du contraire [Löw 1928 : 393-399]. Soit par exemple la question de décision : "Napoléon est-il le vainqueur d'Iéna, ou non ?" ; on peut répondre : "Napoléon est le vainqueur d'Iéna", ou : "L'exilé de Sainte-Hélène est le vainqueur d'Iéna". Les deux énoncés décrivent un même état de fait, et le second est une réponse correcte, à condition que l'on sache préalablement que "Napoléon" et "l'exilé de Sainte-Hélène" désignent le même individu. Le problème est un peu plus complexe pour les questions à compléter. Soit l'interrogation : "Que peut-on extraire de la houille ?" ; si cet énoncé signifie : "Quels produits peut-on extraire ... ?", l'unique réponse est l'énumération exhaustive des dérivés de la houille (une "réponse partielle" n'est pas une réponse) ; si l'énoncé correspond à : "Quel produit ... ?", la question n'a pas de réponse (puisque'il n'y a pas qu'un seul dérivé de la houille). De même, soit l'énoncé : "Qu'est-ce que l'or ?" ; la réponse à cet énoncé variera selon qu'il signifie : "A quel genre appartient l'or ?" (réponse : "C'est un corps simple") ou : "A quelle espèce appartient l'or ?" (réponse :

"C'est un métal"). Cette forme d'ambiguïté se retrouve dans les mots interrogatifs tels que : "où ?" (mis pour : "dans quelle maison ?", "dans quelle ville ?", "dans quel tiroir ?", etc.), "quand ?" (signifiant : "en quelle année ?", "à quelle heure ?", "à quelle occasion ?", etc.). Le cas des questions demandant un exemple est plus embarrassant. Soit l'énoncé : "Quel est l'un (quelconque) des nombres transcendants ?". L'assertion : " π est l'un des nombres transcendants" peut constituer une réponse dans le cas où le sens de l'énoncé est en fait : "Citez le nombre transcendant dont il a été question dans le cours" ; en dehors de ce cas particulier, la question n'a pas de réponse [Löw 1928 : 397-398].

Une question est dite *soluble* s'il existe une proposition vraie, et une seulement, affirmant l'existence de l'état de fait recherché. Aussi, la condition nécessaire et suffisante à la solubilité d'une question est-elle l'existence de l'objet de sa recherche. De même que toute proposition (vraie ou fausse) pose implicitement qu'elle est vraie, de même toute question pose implicitement que l'état de fait qu'elle recherche existe ; si cet état de fait n'existe pas, la question est *insoluble*. Appelons D^F la présupposition d'une question F . S'il s'agit d'une question de décision, D^F est l'affirmation que, parmi les états de fait décrits par la question, un et un seulement existe s'il s'agit d'une question à compléter, D^F est la proposition J_3 . Si D^F est vrai, la question est soluble ; si D^F est faux, elle est insoluble (cf. [Löw 1928 : 417-421, 424]). Appliquons par exemple cette règle aux questions de décision :

- "Le cuivre est-il un corps simple ou un corps composé ?" est soluble, car le cuivre est nécessairement l'un ou l'autre, et ne peut être à la fois l'un et l'autre.

- "Le cuivre est-il un sulfate ou un nitrate ?" est insoluble, car le cuivre n'est ni l'un ni l'autre.

- "Des nombres 5 et 7, lequel est premier ?" est insoluble, car les deux états de faits décrits sont vrais (5 et 7 sont tous deux premiers).

- "Le nombre 2 est-il un nombre impair, ou un nombre premier ?" est soluble, sauf si l'on pose que tous les nombres premiers sont impairs, auquel cas la question est insoluble.

Pour les questions à compléter, on a les exemples suivants :

- "De quelle couleur est l'or ?" est soluble, car : $D^F \stackrel{3}{=} J_3 =$ "L'or a une couleur déterminée" est vrai.

- "Quel est le plus grand nombre premier ?" est insoluble, car : $J_3 =$ "Il existe un nombre qui est le plus grand nombre premier" est faux.

- "Quelle est la loi qui régit la contraction des gaz sous l'effet de la chaleur ?" est insoluble, car : $J_2 =$ "Les gaz se contractent sous l'effet de la chaleur", étant faux, entraîne la fausseté de $J_3 =$ "Les gaz se contractent sous l'effet de la chaleur selon une loi déterminée".

La solubilité d'une question ne dépend pas de la nature de l'objet évoqué dans la question [Löw 1928 : 422, 436 § 13]. On peut par conséquent énoncer quelques propriétés concernant la solubilité *a priori* de certaines formes de questions :

- Toute question à compléter dont la présupposition J_3 est contradictoire est insoluble (ex. : "Quelle fonction est à la fois continue et discontinue ?").

- Toute question de décision qui pose deux états de fait contradictoires est soluble (ex. : " P est-il p ou non - p ?").

- Toute question de décision qui pose soit des contradictions, soit des tautologies, est insoluble (ex. "Est-ce que S est à la fois p et non- p ou bien à la fois q et non- q ?").

D'autre part, la solubilité d'une question donnée F peut entraîner la solubilité d'une autre question :

- Si F est soluble, alors toute question F' telle que la vérité de D^F équivaut à la vérité de $D^{F'}$ est également soluble.

- Si F est soluble, alors toute question F' telle que la vérité de D^F équivaut à la fausseté de $D^{F'}$ est insoluble.

- Si F est insoluble, alors toute question F' telle que la fausseté de D^F équivaut à la vérité de $D^{F'}$ est soluble.

- Si F est insoluble, alors toute question F' telle que la fausseté de D^F équivaut à la fausseté de $D^{F'}$ est également insoluble.

(On trouvera la discussion et l'illustration de certains de ces théorèmes dans : [Löw 1928 : 422-425]).

Nous avons mentionné la thèse de LÖW selon laquelle toute proposition vraie est une réponse à une infinité de questions. Elle a pour corollaire la possibilité que des questions différentes aient la même réponse. Cette propriété d'*identité de la réponse* constitue une relation fondamentale entre des questions différentes. Exemples de questions ayant une réponse identique [Löw 1928 : 413-415] :

- *questions de décision* : "*S* est-il *p* ou *q* ?", "*S* est-il *p* ou *r* ?", "*S* est-il *p* , ou *q* , ou *r* ?" ont la même réponse si (et seulement si) la description de l'état de fait existant est : "*S* est *p*".

- *questions à compléter* : "En quelle année l'Amérique a-t-elle été découverte ?" et : "Quel continent a été découvert en 1492 ?", ont des présuppositions J_3 différentes ; mais, complétées par l'élément recherché par chaque question, celles-ci donnent naissance à une proposition identique.

- *questions de types différents* : "Quelle est la plus grande des planètes ?" et : "Des planètes Mercure, Vénus, et Jupiter, laquelle est la plus grande ?" visent toutes deux le même état de fait, décrit par la proposition : "Jupiter est la plus grande des planètes". Cette relation d'identité de la réponse intervient dans le théorème suivant : si une question est soluble, alors toutes les questions ayant la même réponse sont solubles [Löw 1928 : 423].

c) Gerold STAHL

La logique des questions de Gerold STAHL est caractérisé par KUBIŃSKI comme étant une "logique des questions sans question" [Kubiński 1966 : 117]. En effet, l'auteur substitue en fait à la notion de question celle de "classe des réponses qui sont suffisantes pour l'obtention d'une réponse vraie à la question posée" [Stahl 1956 : 74], et considère toute question comme une classe de propositions [Stahl 1956 : 74n], ou mieux d'expressions [Stahl 1960 : 293], constituée de l'ensemble des réponses suffisantes à la question [Stahl 1969 : 212].

La catégorie de réponses la plus générale, parmi celles que l'auteur analyse, est celle de réponse propre. Une *réponse propre* doit satisfaire à quatre conditions : être vraie ou fausse, ne pas être sous une forme abrégée, ne pas apporter d'information excédentaire, et ne pas être évasive ; par conséquent, une réponse propre est "une proposition qui contient tous les éléments de la question et rien de plus, et qui apporte une solution au problème exprimé par la question" [Stahl 1956 : 71]. Une réponse propre peut toutefois être *inattendue* ; par exemple, si l'on demande : "Où allez-vous ?", ou bien : "Est-ce Pierre ou Paul l'assassin ?", on ne désire pas en principe obtenir une réponse propre du genre : "Je ne vais nulle part", ou bien : "Ce n'est ni l'un ni l'autre". Une réponse *attendue*, ou *réponse directe*, est une proposition simple qui n'est pas une négation de théorème. Une *réponse parfaite* est soit une réponse directe, soit une conjonction de réponses directes, soit une conjonction des négations des réponses directes. Enfin, une *réponse suffisante* est une expression bien formée (proposition ou forme propositionnelle) qui n'est pas la négation d'un théorème et qui, ou bien implique au moins une réponse parfaite qui n'est pas un théorème, ou bien est un théorème à condition qu'il y ait au moins une réponse parfaite qui soit un théorème [Stahl 1960 : 294].

Toute question étant une classe de réponses suffisantes, celle-ci :

- possède au moins un élément ;
- n'a pour éléments que des expressions bien formées ;
- ne peut pas avoir pour élément la négation d'un théorème ;
- peut avoir des théorèmes pour éléments ;
- peut avoir des propositions fausses pour éléments ;
- a pour sous-classes les réponses parfaites et les réponses directes.

D'autre part, il existe des réponses non suffisantes, telles que les réponses abrégées (réponses codées), ou certaines réponses non évatives qui ne sont pas des expressions bien formées ; il est possible de les transformer de manière à obtenir des réponses suffisantes.

STAHL distingue trois types fondamentaux de questions, dont il pense qu'ils épuisent l'ensemble des formes de questions observables dans la vie courante [Stahl 1960 : 294, 295-297] :

- les *questions individuelles* $[Hx?]$, définies par une constante de prédicat H , et une variable d'individu x sur laquelle porte l'interrogation (dans l'hypothèse d'un prédicat unaire). Ces questions se lisent : $[Hx?] =$ "Quel est l'individu qui satisfait à la condition H ?" (exemples : "Qui est l'assassin ?" ; "Quel est le lieu où Socrate est mort ?"). Une question de ce type peut avoir comme éléments des réponses directes : " Ha ", " Hb ", " Hc " (a , b , et c étant des constantes d'individus) ; d'autres formes de réponses parfaites, comme : " $Ha \ \& \ Hb \ \& \ Hc$ ", ou : " $\forall x (Gx \rightarrow Hx)$ " ; d'autres formes de réponses suffisantes, comme : " $A \ \& \ Ha$ ".

- les *questions fonctionnelles* $[X?a]$, définies par une variable de prédicat X , sur laquelle porte l'interrogation, et une constante d'individu a (s'il s'agit d'un prédicat unaire). On a : $[X?a] =$ "A quelle classe appartient l'individu a ?", avec comme exemples d'éléments : " Ha ", " $Ha \ \& \ Ga \ \& \ Fa$ ", " $\forall X (kX \rightarrow Xa)$ " (avec $k =$ fonction unaire constante du second ordre), etc.

- les *questions de vérité unaires* : $[f?A]$, binaires : $[Af?B]$, etc. définies par une fonction de vérité n -aire, sur laquelle porte l'interrogation, et un ensemble de n propositions A, B , etc. (exemple de question de vérité unaire : $[f?A] =$ "Pleut-il ?"). On a : $[Af?B] =$ "Quelle fonction de vérité binaire lie A et B ?", avec comme éléments : " $A \vee B$ ", " $A \rightarrow B$ ", " $A \leftrightarrow B$ ", etc.

En appliquant aux questions ainsi définies les règles du calcul des classes, il est possible d'énoncer un certain nombre de définitions et de théorèmes [Stahl 1960 : 298], dont les plus importants nous paraissent être :

- Deux questions sont *identiques* si et seulement si elles ont les mêmes réponses suffisantes.

- Une question I est une *sous-question* d'une question $J \neq I$ si toutes les réponses suffisantes à I sont des réponses suffisantes à J .

- L'*union* de deux questions est la classe des réponses qui sont suffisantes pour au moins une de ces questions.

- L'*intersection* de deux questions est la classe des réponses qui sont suffisantes simultanément pour ces deux questions.

- Une question de vérité et son contraire sont identiques : $[f?A] = [f?\bar{A}]$ (elles ont les mêmes réponses suffisantes, même si en pratique une formulation donnée suggère certaines réponses plutôt que d'autres).

- L'intersection d'une question individuelle et de son contraire n'est pas vide : $[Hx?] \cap [\bar{H}x?] \neq \emptyset$ (Il existe en effet des réponses suffisantes pour les deux, telles que : " $\forall x, Hx$ ", ou : " $\forall x, \bar{H}x$ ").

- L'intersection d'une question fonctionnelle et de son contraire n'est pas vide : $[X?a] \cap [\bar{X}?a] \neq \emptyset$ (Sont des réponses suffisantes pour les deux toutes les expressions de la forme : " Ha ", " Ga ", qui ne sont ni des théorèmes, ni des négations de théorèmes).

(D'autres théorèmes peuvent être formulés sans difficulté, concernant en particulier l'existence de réponses suffisantes à une question ou à une combinaison de questions [Stahl 1956 : 74]. Mais leur intérêt pour notre recherche reste limité).

L'ensemble de cette théorie permet de rendre compte de formes de questions plus complexes que les types fondamentaux, comme par exemple :

- les *questions de sélection*, dans lesquelles sont présentées diverses réponses possibles, liées par la conjonction *ou*, comme : "Marie a-t-elle épousé Charles, ou Pierre ?" [Stahl 1956 : 73]. Malgré la disjonction exprimée dans l'énoncé, il s'agit d'une *intersection* de questions de vérité unaires : $[f?A] \cap [g?B] = \text{"Marie a-t-elle épousé Charles ?"} \cap \text{"Marie a-t-elle épousé Pierre"}$ [Stahl 1960 : 299]. En effet, une réponse comme : $B = \text{"Marie a épousé Pierre"}$ est une réponse suffisante non seulement à $[g?B]$, mais aussi à $[f?A]$, puisque : $B \rightarrow \bar{A}$. De même, $C = \text{"Marie a épousé Jean"}$ est une réponse suffisante pour les deux questions élémentaires, puisque : $C \rightarrow \bar{A}$, et : $C \rightarrow \bar{B}$.

- les *questions composées*, qui peuvent être considérées comme la forme abrégée de plusieurs questions élémentaires, ou comme une question élémentaire à plusieurs variables [Stahl 1956 : 73, 73n]. Par exemple : "Où ont-ils pris l'assassin et quand ?" peut être traité comme une question comportant une fonction propositionnelle binaire : $[Hx?y?]$ (cf. : [Stahl 1960 : 297]).

- les *questions conditionnelles*. Celles-ci admettent pour réponses directes à la fois les réponses à la question proprement dite, et la négation des propositions exprimant la condition [Stahl 1956 : 73]. Lorsque la question proprement dite est une question de vérité unaire, on peut avoir selon l'énoncé les formes : $[f?(A \rightarrow B)] = \text{"Si Jeanne est malade, travaille-t-elle ?"}$; $[f?(A \rightarrow \bar{B})] = \text{"Est-ce que, si Jeanne est malade, elle ne travaille pas ?"}$; ou encore : $[f?(A \rightarrow B) \cap [f?(A \rightarrow \bar{B})]] = \text{"Si le ministre arrive aujourd'hui, le député restera-t-il, ou (si le ministre arrive aujourd'hui, le député ne restera-t-il pas ?"}$ [Stahl 1960 : 299-301].

d) David HARRAH

La logique des questions développée par David HARRAH est un élément d'une théorie générale de la communication [Harrah 1963], sur laquelle nous reviendrons au chapitre 3. L'objectif visé par l'auteur est de "construire le modèle d'une procédure rationnelle pour poser des questions et fournir des réponses, applicables aux situations dans lesquelles le récepteur désire acquérir de l'information, analyser des significations, et estimer la valeur de l'émetteur en tant que source d'information" [Harrah 1963 : 25]. Ce modèle doit rendre compte des questions de telle sorte que :

- l'utilisateur puisse effectivement, pour chacune de ses interrogations, élaborer une question pertinente, non ambiguë, et ayant un sens ;
- toute question puisse être reconnue comme telle ;
- à toute question corresponde un énoncé qui lui réponde d'une manière raisonnablement adéquate.

Ce modèle doit en outre :

- permettre de distinguer entre les réponses qui sont complètes et suffisantes, et celles qui ne le sont pas ;

- rendre compte de certaines relations entre questions, comme les relations d'inclusion et d'équivalence ;

- rendre compte des réponses de telle sorte que celles-ci soient des expressions bien formées, susceptibles par conséquent d'être impliquées par des propositions [Harrah 1963 : 25-27].

Les hypothèses de base du modèle proposé par HARRAH sont les suivantes :

- un type particulier de question est identifié à un type particulier d'expression. Ainsi, une question *whether* est assimilée à une disjonction exclusive vraie de propositions, dont chacune constitue une réponse possible à la question ; une question *which*, à la quantification existentielle vraie d'une proposition [Harrah 1961 : 40 ; 1963 : 28]. Toutefois, "quand nous disons que les questions sont des propositions, nous n'affirmons pas que poser une question est la même chose, du point de vue psychologique, que faire une assertion. La psychologie de ces deux comportements peut être différente, mais cela ne constitue pas une objection pour notre propos" [Harrah 1963 : 30].

- "dans cette analyse, la situation question-réponse est traduite non en une séquence de requêtes et de satisfaction de ces requêtes, mais en un jeu joué avec des assertions - plus précisément un jeu d'échange d'informations" [Harrah 1961 : 41]. Dans un tel jeu, le locuteur exprime ce qu'il sait sur un sujet donné ; l'interlocuteur a alors la possibilité de riposter en énonçant une proposition qui répond à l'affirmation du locuteur en lui apportant plus d'information que celui-ci ne déclarait en posséder, diminuant ainsi l'incertitude du locuteur. Dans un tel modèle, les réponses du type : "Je ne sais pas", "cela est possible", "je n'ai pas encore décidé", peuvent être considérées non comme répondant à l'affirmation du locuteur, mais comme constituant plutôt des commentaires sur le jeu lui-même, et sur les coups qui viennent d'être joués [Harrah 1961 : 41 ; 1963 : 29-30].

HARRAH distingue deux types fondamentaux de questions. Le premier type, *question disjonctive à n arguments*, correspond aux questions *whether*. Sa formalisation obéit aux règles suivantes. Soit une suite finie de n propositions : A_1, \dots, A_n . On définit à partir de cette suite la conjonction de propositions :

$$A_i^* = \bar{A}_1 \ \& \ \bar{A}_2 \ \& \ \dots \ \& \ \bar{A}_{i-1} \ \& \ A_i \ \& \ \bar{A}_{i+1} \ \& \ \dots \ \& \ \bar{A}_{n-1} \ \& \ \bar{A}_n$$

dans laquelle une proposition et une seule : A_i , est affirmée, toutes les autres propositions conjointes étant niées. On appelle *disjonction primitive* de la suite de n propositions l'expression :

$$D = A_1^* \vee A_2^* \vee \dots \vee A_i^* \vee \dots \vee A_n^*$$

Toute suite finie d'au moins deux propositions engendre une disjonction primitive et une seule. On dira que I est une question disjonctive à n arguments si et seulement si : $I = D ?$. Une telle question a exactement n réponses directes de la forme : $d(I) = A_i^*$. Remarquons que l'opérateur de disjonction (\vee) se comporte dans D comme un opérateur d'exclusion réciproque (\wedge). Ce type de question correspond chez BELNAP à la question *whether* à réponse vraie unique :

$$I = ? \left(\bigvee_{i=1}^n A_i \right)$$

Le deuxième type de question analysé par HARRAH est celui des questions *which*. Soit une expression contenant n variables libres distinctes : $Ax_1 \dots x_n$, et dont la généralisation existentielle : $\exists x_1, \dots, \exists x_n \ Ax_1 \dots x_n$ est vraie. On appelle *liste exhaustive relativement à* $Ax_1 \dots x_n$ une expression C de la forme :

$$C = \forall x_1, \dots, \forall x_n \left(Ax_1 \dots x_n \longleftrightarrow \left((x_1 = a_1 \ \& \ \dots \ \& \ x_n = a_n) \vee \dots \vee (x_1 = b_1 \ \& \ \dots \ \& \ x_n = b_n) \right) \right)$$

dans laquelle $a_1, \dots, a_n, \dots, b_1, \dots, b_n$ sont des constantes d'individu. Cette expression est équivalente à :

$$\begin{aligned} & Aa_1 \dots a_n \ \& \ \dots \ \& \ Ab_1 \dots b_n \quad \& \quad \forall x_1, \dots, \forall x_n \\ & \left(\overline{(x_1 = a_1 \ \& \ \dots \ \& \ x_n = a_n)} \ \& \ \dots \ \& \ \overline{(x_1 = b_1 \ \& \ \dots \ \& \ x_n = b_n)} \right) \\ & \rightarrow \overline{Ax_1 \dots x_n} \end{aligned}$$

On dira que I est une question *which* si et seulement si : $I = \exists x_1, \dots, \exists x_n Ax_1 \dots x_n$, que l'on abrègera en : $I = ? Ax_1 \dots x_n$. Une question *which* est par conséquent la généralisation existentielle d'une expression comportant au moins une variable libre. Toute réponse directe à une question *which* est une liste exhaustive : $d(I) = C$. Ces définitions apparentent les questions *which* de HARRAH aux questions *which* demandant une énumération exhaustive (et distincte) de BELNAP.

A part les définitions qui précèdent, BELNAP a repris (parfois en les modifiant légèrement) bon nombre des définitions et des théorèmes proposés par HARRAH. C'est pourquoi nous ne mentionnons ici qu'une faible partie des résultats auxquels conduit la théorie de HARRAH. L'essentiel des définitions relatives aux formes de question et de réponses est :

- une *question test* est une question dont le locuteur connaît au moins une réponse directe [Harrah 1963 : 37].

- une *question piège* est une question dont toutes les réponses directes sont inconsistantes [Harrah 1963 : 51, 79].

- toute *réponse directe* à une question est soit de la forme : A_1^* , soit une liste exhaustive. Ceci est une conséquence des définitions déjà mentionnées, et du théorème selon lequel I est une question si et seulement si I est ou bien une question disjonctive, ou bien une question *which* [Harrah 1963 : 33].

- une expression A est une *réponse partielle* à une question I si et seulement si il existe une réponse directe $d(I)$ consistante qui implique A [Harrah 1963 : 39].

- enfin, un *éclaircissement* relativement à une question I est une assertion A telle que $A \& I$ implique l'existence d'une relation d'équivalence entre I et une autre expression (cf. [Harrah 1961 : 43 ; 1963 : 51 n1]). Pour illustrer cette notion intéressante, l'auteur cite le dialogue suivant, dans lequel toutes les répliques sont des éclaircissements, mais non des réponses à la question posée [Harrah 1963 : 51] :

Q : "Est-ce que Throop est Démocrate ?"

R : "Throop n'est pas le genre de personne à s'engager politiquement"

Q : "Mais est-ce que d'habitude il vote Démocrate ?"

R : "Il ne vote jamais pour un parti : il vote pour un candidat".

Q : "Mais a-t-il déjà voté pour des candidats du parti Démocrate ?"

R : "Ordinairement, ce n'est pas *pour* un candidat qu'il vote, mais *contre* un autre candidat".

En conséquence du postulat selon lequel une question *est* une expression, HARRAH énonce un certain nombre de théorèmes exprimant une équivalence entre des expressions interrogatives et des expressions assertives. Par exemple [Harrah 1963 : 35] :

$(\overline{A}?)$	\longleftrightarrow	$(A \ \& \ \overline{A})$
$(\overline{A})?$	\longleftrightarrow	$(\overline{A} \ w \ \overline{\overline{A}})$
$(A? \ \& \ B?)$	\longleftrightarrow	$((A \ w \ \overline{A}) \ \& \ (B \ w \ \overline{B}))$
$(A \ \& \ B)?$	\longleftrightarrow	$((A \ \& \ B) \ w \ \overline{(A \ \& \ B)})$
$(A \ \& \ B?)$	\longleftrightarrow	$(A \ \& \ (B \ w \ \overline{B}))$
$(A \ \rightarrow \ B)?$	\longleftrightarrow	$((A \ \rightarrow \ B) \ w \ \overline{(A \ \rightarrow \ B)})$
$(A \ \rightarrow \ B?)$	\longleftrightarrow	$(A \ \rightarrow \ (B \ w \ \overline{B}))$
$(A? \ \rightarrow \ B)$	\longleftrightarrow	$((A \ w \ \overline{A}) \ \rightarrow \ B)$
$(A? \ \rightarrow \ B?)$	\longleftrightarrow	$((A \ w \ \overline{A}) \ \rightarrow \ (B \ w \ \overline{B}))$

Enfin, parmi les autres théorèmes intéressants non mentionnés par BELNAP (ou non rapportés par nous dans le sous-chapitre 2.2 , citons [Harrah 1963 : 36, 40] :

- il peut arriver qu'une question soit vraie, mais n'ait pas de réponse directe consistante, ou n'ait pas de réponse directe vraie. Par exemple, si $Ax_1 \dots x_n$ peut être satisfait dans un domaine infini, mais non dans un domaine fini, $I = ? Ax_1 \dots x_n$ n'a pas de réponse directe consistante. Si le domaine des individus contient une infinité d'individus qui satisfont Ax , alors $I = ? Ax$ a de nombreuses réponses fausses (chacune constituée d'une liste finie), mais aucune réponse directe vraie.

- une condition suffisante pour qu'une question *which* ait au moins une réponse directe fausse est qu'il existe un couple de constantes d'individus (a,b) , tel que l'on ait : $a \neq b$. Soit en effet : $I = ? Ax$, on a : $d(I)_1 = \forall x (Ax \leftrightarrow (x=a))$ et : $d(I)_2 = \forall x (Ax \leftrightarrow (x=b))$; si l'hypothèse : $a \neq b$ est vraie, les deux réponses directes ne peuvent être vraies simultanément.

- si une question a au moins une réponse partielle, alors elle en a une infinité.

- une proposition fausse peut être une réponse directe à une question, mais elle ne peut pas être une réponse partielle.

2.4. SYNTAXE ET SÉMANTIQUE FORMELLES DES QUESTIONS ET DES RÉPONSES

A l'issue de ce trop bref survol des développements récents de la logique érotétique, il convient de dresser un bilan provisoire des apports actuels de cette discipline à une analyse des questions et des réponses. L'extrême diversité des points de départ, et le caractère encore incomplet des théories que nous avons présentées, rendent cette tâche malaisée. D'autant qu'à la difficulté de comparer les théories entre elles, même sur un point particulier bien circonscrit, s'ajoute l'hétérogénéité du vocabulaire et des formalisations. C'est pourquoi nous avons limité notre tentative de synthèse aux classifications et aux propriétés qui nous ont paru les plus assurées, et les plus facilement utilisables pour notre recherche.

La logique érotétique dans son ensemble apparaît comme une articulation entre la syntaxe des questions et des réponses, et leur sémantique. Dans les principaux travaux que nous avons analysés, les deux aspects syntaxique et sémantique étaient toujours clairement distingués, au moins en ce qui concerne les questions. Nous avons conservé cette distinction dans notre bilan, en exposant successivement la syntaxe formelle des questions et des réponses, leur sémantique, et enfin les perspectives d'application que ces travaux nous ont semblé ouvrir.

2.4.1. Syntaxe

Toutes les théories que nous avons passées en revue affirment l'interdépendance des questions et des réponses. Il y a donc sur ce problème un parfait accord entre les théories linguistiques de TESNIÈRE, CHOMSKY, et HARRIS, et la logique érotétique. Ce point peut par conséquent être considéré comme acquis. Par contre, nous avons pu constater qu'il existe de nombreuses divergences, entre les diverses théories logiques, sur la définition des principaux types de questions ou de réponses, ou même sur la structure de base des questions ou des réponses.

Nous commencerons par poser que toute question se compose de deux parties distinctes et complémentaires : le *sujet*, et la *requête*. Si cette terminologie est empruntée à BELNAP, la conception qu'elle recouvre est partagée par ÅQVIST et (au moins implicitement) par KUBIŃSKI. Ainsi que nous l'avons déjà noté, cette conception nous paraît particulièrement bien adaptée à l'objet de notre recherche, puisqu'elle attribue un statut particulier à la *consigne* de la question. D'autre part, en ce qui concerne les principales formes de questions, les divergences entre les différents auteurs que nous avons cités sont considérables. Toutefois, en nous limitant aux points d'accord entre les théories, nous pouvons distinguer trois formes de questions élémentaires d'un usage courant : alternative portant sur une contradiction, question extensive, et question compréhensive (cette classification coïncide semble-t-il avec celle proposée par [Leonard 1967 : 33-34] ; cf. § 2.1.3.b). Ces formes simples peuvent évidemment être combinées entre elles pour donner naissance à des formes composées ; nous renvoyons sur ce point aux auteurs que nous avons cités, principalement à BELNAP et ÅQVIST. Mais à côté de ces formes élémentaires, il existe d'autres formes qui s'en distinguent par le contenu de leur sujet ; nous avons tenté de les classer en moins de dix grandes catégories. D'autre part, il était nécessaire de donner une vue d'ensemble des formes de requêtes distinguées par les logiciens. Enfin, nous rappelons la définition de la principale forme syntaxique de réponse, la réponse directe, et la nature des liens qui l'unissent à la question.

a) Les trois formes de questions élémentaires

La première des formes que nous avons distinguées, l'*alternative portant sur une contradiction*, n'est pas classée à part par tous les logiciens. Ainsi, pour BELNAP, la "question à réponse par *oui* ou *non* régulière" n'est qu'une forme particulière de question *whether*, qui présente toutefois la particularité d'être formalisable de plusieurs manières à peu près équivalentes (cf. [Belnap 1976 : 59], et § 2.2.2.b). De même, pour HARRAH, qui est sur certains points assez proche de BELNAP, cette forme de question n'est qu'une variété de *question disjunctive*, pour laquelle il est cependant possible

d'adopter une écriture condensée particulière (cf. [Harrah 1963 : 34, définition 7.9, § 2]), mais qui ne porte pas de nom spécifique. ÅQVIST, qui considère lui aussi la question à réponse *oui/non* régulière comme un cas particulier de question *whether*, lui donne également les noms de *question "est-ce le cas"* ou de *question whether-or-not* [Åqvist 1965 : 51] ; il introduit en outre un opérateur unaire spécifique de cette forme de question [Åqvist 1965 : 57]. Dans le système de logique érotétique de STAHL, la question à réponse *oui/non* est une *question de vérité unaire* [Stahl 1960 : 297]. Dans celui de KUBIŃSKI enfin, cette forme est traitée à part sous le nom de *question de décision* (cf. [Kubiński 1970 : 27], et § 2.3.2.b).

L'intérêt qu'il y a à considérer l'alternative portant sur une contradiction comme une forme essentielle de question élémentaire, n'est pas seulement de retrouver une catégorie distinguée par la plupart des grammairiens (cf. § 1.1.2.b, 1.2.4.c, 2.1.2.b, etc.), et par la quasi-totalité des précurseurs de la logique érotétique ("question dialectique" d'ARISTOTE, "ἐρώτημα" des Stoïciens, "interrogatio simplex" de JUNGE, etc.). En effet, cette forme de question possède des propriétés spécifiques que ne possèdent pas les autres formes de questions *whether*. En particulier, elle a toujours une réponse vraie, quel que soit l'état de l'univers, puisque sa présupposition est une tautologie ; elle est donc *absolument sûre* (cf. § 2.2.5.c et d). Cette forme de question ne doit pas être confondue avec l'alternative portant sur deux propositions non contradictoires, à laquelle on peut dans certains cas répondre par *oui* ou *non*, mais dont la présupposition n'est pas toujours vraie.

La *question extensive* présente une énumération de propositions au sein de laquelle le répondant choisira la ou les réponses vraies. Cette catégorie correspond par conséquent, pour BELNAP et ÅQVIST, à l'ensemble des *questions whether* autres que l'alternative portant sur des propositions contradictoires. Elle contient, sans toutefois s'y limiter, l'ensemble des *questions conjonctives* et des *questions disjonctives* selon KUBIŃSKI ; l'ensemble des *questions de sélection* d'après STAHL ; l'ensemble des *questions disjonctives*, autres que l'alternative portant sur une contradiction, dans

le système de HARRAH. Dans cette forme de question par conséquent, l'ensemble des réponses directes possibles est présenté *en extension* au répondant (hormis la réponse correctrice évidemment). On peut classer les questions extensives à partir de leur requête, selon le nombre minimum et le nombre maximum de réponses qu'elles demandent, selon qu'elles exigent ou non l'exhaustivité (c'est BELNAP qui est le plus complet sur ces deux points), et enfin selon le nombre de propositions qu'elles présentent comme réponses possibles (ce nombre figure dans l'opérateur interrogatif chez ÅQVIST et KUBIŃSKI, mais non chez BELNAP).

La *question compréhensive*, sous sa forme la plus générale, présente au répondant un sujet constitué de deux éléments distincts : une forme propositionnelle n -aire qui exprime la ou les propriétés que la réponse devra attribuer aux n variables ; et, s'il y a lieu, une série d'expressions (conditions catégorielles) définissant les sous-ensembles du domaine des individus dans lesquels les variables peuvent prendre leurs valeurs (catégories réelles). Les variables non liées à une condition catégorielle peuvent prendre leurs valeurs dans la totalité de l'ensemble des individus. Par conséquent, étant donné la forme propositionnelle exprimant une question compréhensive, la liste des réponses réelles possibles à cette question est l'énumération de toutes les réalisations possibles de la forme propositionnelle, compte-tenu éventuellement des restrictions sur les valeurs que peuvent prendre les variables, restrictions exprimées par les conditions catégorielles. La liste des réponses possibles énoncées sous une forme abrégée se ramène donc à la liste des combinaisons des valeurs possibles de chacune des n variables, définie *en compréhension* par la conjonction des conditions catégorielles (cf. § 2.2.2.d).

Il est clair que les questions compréhensives ainsi définies correspondent assez fidèlement aux *questions which* de BELNAP. Par contre, les définitions les plus proches que proposent les autres auteurs sont beaucoup plus restrictives que celle de BELNAP. Ainsi, la forme générale des *questions what* d'ÅQVIST correspond à celle des *questions which* avec exigence de distinction ; la forme générale des *questions numériques* de KUBIŃSKI, comme d'ailleurs celle des *questions individuelles* de STHAL,

Terminologie proposée :	Alternative portant sur une contradiction	Question extensive	Question compréhensive
BELNAP :	q. à réponse <i>oui/non</i> régulière : $I = ? (\bigvee_1) (A, \bar{A})$	q. <i>whether</i> : $I = ? \rho (A_1, \dots, A_n)$	q. <i>which</i> : $I = ? \rho (C_1 x_1, \dots, C_m x_m // P x_1 \dots x_m)$
ÅQVIST	q. "est-ce le cas", q. <i>whether-or-not</i> : $I = ?_1 (A)$	q. <i>whether</i> : $I = ?_n^\omega (A_1, \dots, A_n)$ avec : $\omega = \neg, u, c$.	q. <i>what</i> : $I = (? \Omega^m x_1, \dots, x_m) P(x_1 \dots x_m)$ avec : $\Omega = B, KB, EB, EKB, C, D, \dots$
KUBIŃSKI :	q. de décision : $[\alpha] \alpha A$	q. conjonctive $I = [\delta^n] \delta^n A_1, \dots, A_n$ q. disjonctive : $I = [\beta^n] \beta^n A_1, \dots, A_n$	q. numérique : $I = \omega_1 / \dots / \omega_{m-1} \Omega_1 / \dots / \Omega_m x_1 \dots x_m P x_1 \dots x_m$ avec : $\Omega_i = k, k \leq, k <, (k), (k \leq), (k <), C$ et : $\omega_i = 1, 2, 3$.
STAHL :	q. de vérité unaire : $I = [f?A]$	q. de sélection : $I = [f_1?A_1] \cap \dots \cap [f_n?A_n]$	q. individuelle : $I = [P x_1 ? \dots x_m ?]$
HARRAH :	q. disjonctive unaire = $(A \bar{A} ?)$: $I = A?$	q. disjonctive : $I = (A_1 \dots A_n ?)$	q. <i>which</i> : $I = ? P x_1 \dots x_m$

Tableau n° 2.12 : Correspondance entre les classes de questions élémentaires selon les auteurs.

correspond à celle des questions *which* sans condition catégorielle ; enfin, la forme générale des *questions which* de HARRAH, à celle des questions *which* de BELNAP avec exigence d'exhaustivité et de distinction. Les questions compréhensives peuvent être classées selon le nombre minimum et le nombre maximum de réponses qu'elles demandent, selon qu'elles exigent ou non l'exhaustivité, et selon qu'elles exigent ou non la distinction ; dans le cas général, on ne peut pas les classer simplement selon le cardinal de l'ensemble de leurs réponses réelles possibles.

Le tableau n° 2.12 rappelle les principales définitions de ces formes élémentaires de questions selon BELNAP, ÅQVIST, KUBIŃSKI, STAHL, et HARRAH.

1) Autres formes de questions

Cette classification des questions élémentaires en trois grandes catégories n'épuise évidemment pas l'ensemble des formes élémentaires distinguées par les logiciens. Parmi celles-ci, il faut se garder de mettre sur le même plan les formes pour lesquelles le calcul est possible, et celles dont la formalisation, en l'état actuel du développement de la logique érotétique, ne débouche sur aucun traitement satisfaisant. Les premières peuvent d'ailleurs toujours être ré-écrites sous la forme d'une combinaison de questions élémentaires, alors que les secondes se situent en marge des grandes théories.

Parmi les formes de questions qu'il est facile de rattacher aux principales théories, les trois plus importantes nous paraissent être :

- la *question conditionnelle*, déjà décrite par WHATELY et M. et A. PRIOR, dans laquelle l'interrogation est subordonnée à une condition

qui lui est extérieure (ex. : "Si vous sortez ce soir, prendrez-vous votre parapluie ?"). Il y a désaccord entre les logiciens sur l'existence d'une réponse directe possible correspondant au cas où la condition n'est pas remplie. Pour BELNAP, si la condition n'est pas réalisée, la question ne demande pas de réponse ; par conséquent, l'ensemble des réponses possibles à une question conditionnelle est le même que pour la question absolue correspondante. De son côté, ÅQVIST accepte la négation de la condition comme étant une pseudo-réponse directe à la question. STAHL défend une thèse tout à fait opposée à celle de BELNAP, et inclut la négation de la condition dans l'ensemble des réponses directes possibles (cette position est également exposée dans [Steinman 1959 : 19-20]). KUBIŃSKI enfin admet également la négation de la condition comme réponse directe possible à la question conditionnelle à antécédent dirimant, mais il analyse cette forme comme la subordination d'une question disjonctive à une question de décision (en particulier, il ne semble pas avoir considéré la possibilité d'une question numérique conditionnelle ; cf. [Kubiński 1970 : 30]).

- la *question hypothétique*, soigneusement distinguée par M. et A. PRIOR de la question conditionnelle, en ce que la condition hypothétique exprimée par l'interrogation fait partie de la question (ex. : "Si vous sortez ce soir, prendriez-vous votre parapluie ?"). BELNAP et KUBIŃSKI s'accordent pour représenter par la relation d'implication (cf. [BELNAP 1976 : 95-97], [Kubiński 1970 : 29]), le lien unissant la condition hypothétique à la proposition sur laquelle porte l'interrogation, et pour considérer, avec M. et A. PRIOR, que refuser l'hypothèse ne dispense pas de répondre à la question.

- la *question "puisque"* (ou : "*étant donné que*"), qui dans son énoncé affirme une condition dont le locuteur et l'interlocuteur admettent tous deux la vérité (ex. : "Puisque vous sortez, prenez-vous votre parapluie ?"). BELNAP et KUBIŃSKI (ce dernier sous le nom de : "question conditionnelle à antécédent non dirimant") donnent de cette forme une analyse identique dans son principe (bien que limitée par KUBIŃSKI à un choix entre deux modalités de réponse possibles ; cf. [Belnap 1976 : 98], [Kubiński 1970 : 28]).

Il convient en outre de citer pour mémoire les très nombreuses formes de *questions composées*, dont ÅQVIST et surtout BELNAP donnent des exemples variés, et qui est le résultat d'une composition (conjonction, union) de questions élémentaires, voire de sujets ayant une requête commune ou de requêtes ayant un sujet commun.

Les formes de questions qui correspondent à des énoncés interrogatifs fréquemment observables, mais pour lesquelles la logique érotétique n'a pas encore, nous semble-t-il, apporté de solution réellement opératoire, peuvent être classées en six catégories :

- *question de causalité*, demandant une explication pouvant être exprimée par un ou plusieurs mots (constantes d'individu), ou bien une ou plusieurs expressions. KUBIŃSKI distingue ainsi trois formes fondamentales de questions de causalité, susceptibles chacune d'avoir pour consigne l'une des sept formes de requêtes applicables aux questions numériques [Kubiński 1970 : 37-38] ; mais l'apport de KUBIŃSKI se limite à une description syntaxique des formes de questions de causalité et de leurs réponses directes, sans essai d'analyse sémantique. Par contre, BROMBERGER s'est efforcé de mener à bien l'analyse sémantique de cette forme de question, sans toutefois y réussir complètement (cf. § 2.3.3.) ; toutefois, la contribution de BROMBERGER a été reprise et simplifiée par BELNAP, qui propose une forme-type de réponse explicative utilisable dans certains cas particuliers (cf. 2.2.3.e).

- *question de description*, qui demande une ou plusieurs formes propositionnelles pouvant s'appliquer à un individu qui est soit désigné par une constante d'individu (ex. : "De quelle couleur est Dick ?" = *question de description* de BELNAP, cf. § 2.2.3.a ; = *question fonctionnelle* de STAHL, cf. § 2.3.4.e), soit décrit par une proposition (ex. : "Quel genre d'homme était le chef militaire français qui a gagné la bataille d'Austerlitz ?" = *question sur des attributs d'individu* de KUBIŃSKI, cf. § 2.3.2.a).

- *question de définition*, qui demande une ou plusieurs formes propositionnelles pouvant s'appliquer à l'ensemble des individus d'une classe définie par une forme propositionnelle (c'est-à-dire aux individus du domaine d'un prédicat unaire). Au sens strict, cette forme correspond à la *question d'équivalence* de BELNAP (ex. : "Qu'est-ce qu'un nombre premier ?") et à la *question de définition portant sur des attributs d'individu* de KUBIŃSKI (ex. : "Qui étaient les huguenots ?", cf. § 2.3.2.a). Au sens large, elle peut englober l'ensemble des *questions what* de BELNAP (cf. § 2.2.3.d), compte-tenu de la difficulté qu'il y a de trouver un équivalent logique aux énoncés de forme : "Qu'est-ce que ... ?" (cf. p.ex., l'analyse de la question : "Qu'est-ce que l'or ?" dans [Löw 1928 : 397]).

- *question d'identité*, qui demande une constante d'individu (ex. : "Qui était l'auteur de *Waverly* ?" [Belnap 1976 : 82]). La *question de définition portant sur des constantes d'individu* de KUBIŃSKI (ex. : "Qu'est-ce que Buenos Aires ?" cf. § 2.3.2.a) est également à ranger dans cette catégorie.

- *question de quantité*, qui demande un quantificateur. Seul ÅQVIST a tenté de fournir une formalisation opératoire de ce type de question (cf. § 2.3.1.c).

- *question sur la valeur de vérité*, qui présente une forme propositionnelle dont les variables liées à l'opérateur d'interrogation peuvent prendre l'une des valeurs *vrai/faux*. KUBIŃSKI en a décrit plusieurs variétés sous le nom de *questions sur la valeur logique* (cf. [Kubiński 1970 : 32-34], et § 2.3.2.b).

Au terme de cette énumération, il subsiste un certain nombre de variétés de questions, analysées par les spécialistes de la logique érotétique, que nous n'avons pas mentionnées dans notre essai de classification. Il s'agit toujours :

- ou bien de légères variantes des formes que nous avons décrites, dues en particulier à des différences dans la seule requête (ex. : deuxième et troisième variétés de "questions numériques" de KUBIŃSKI) ;

- ou bien de formes complexes obtenues, à partir de formes élémentaires, par composition ou quantification ;

- ou bien de formes, au demeurant assez rares, qui ne correspondent pas à des énoncés interrogatifs usités, et que nous n'avons pas réussi à intégrer dans une présentation synthétique (ex. : "questions de vérité n -aires" de STAHL, avec $n > 1$).

c) Analyse du contenu de la requête

Jusqu'ici, les formes de questions que nous avons décrites ne différaient que par la nature de leur sujet, ou par les relations qu'elles entretenaient soit entre elles (questions composées), soit avec des conditions externes (questions relativisées). C'est pourquoi il est nécessaire d'aborder maintenant le problème de la *requête*, d'autant que cette notion présente, pour l'analyse des questions de tests ou de questionnaires, un grand intérêt pratique.

Parmi les logiciens dont nous avons examiné les théories, seuls les trois principaux ont procédé à une analyse détaillée de la requête. C'est probablement BELNAP qui, le premier (dans [Belnap 1963]), a reconnu l'importance de cette notion ; c'est certainement lui qui en a donné la forme la plus achevée, applicable en principe à toutes les variétés de questions. Pour BELNAP, la requête a pour fonction essentielle de déterminer le contenu de la réponse. Elle comporte quatre parties : nombre minimum d'éléments sélectionnés (propositions pour les questions extensives, n -uples d'individus pour les questions compréhensives), nombre maximum d'éléments sélectionnés, exigence d'exhaustivité (affirmée automatiquement dans la réponse), exigence de distinction (également affirmée automatiquement dans la réponse, valable

seulement pour les questions compréhensives). ÅQVIST considère plutôt la requête à la fois sous l'angle de l'information requise, et sous l'angle de l'information apportée par la réponse à celui qui pose la question. D'où une distinction importante entre la satisfaction effective, par le contenu de la réponse, des exigences de la requête, et l'affirmation par la réponse que ces exigences ont été satisfaites. Par ailleurs, si ÅQVIST ne juge pas nécessaire de conserver la possibilité de fixer conjointement un effectif minimum et un effectif maximum distinct pour la taille de la sélection, en revanche il tient compte du cas où la sélection pourrait être vide, indépendamment de toute erreur catégorielle. En outre, ÅQVIST fait figurer, dans la requête des questions compréhensives, le nombre de variables sur lesquelles porte l'interrogation ; et de même, ÅQVIST et KUBIŃSKI incluent dans la requête des questions extensives le nombre de modalités présentées par le sujet de celles-ci. Enfin, dans sa formalisation des "questions numériques complexes", KUBIŃSKI introduit un élément de requête original, relatif aux formes d'appariements possibles entre les arguments de la question ; cet élément ne joue que pour la variété particulière de question compréhensive à n -arguments demandant exactement k n -uples d'individus (le calcul de k étant d'ailleurs lié à la valeur prise par cet élément). L'intérêt de cet élément pour la formalisation des questions d'enquête ou des items d'épreuves psychologiques apparaîtra au chapitre suivant (§ 3.1.2.d).

A partir de l'ensemble de ces analyses, il est facile de dresser la liste des éléments pouvant entrer dans la constitution de la requête d'une question :

(a) nombre de modalités présentées par le sujet (pour les questions extensives) ou nombre de variables liées à l'opérateur interrogatif (pour les questions compréhensives) ;

(b) présomption d'existence d'au moins une modalité de réponse (question extensive) ou un n -uple d'individus (question compréhensive n -aire) satisfaisant à la question ;

- Ⓒ demande d'affirmation, s'il y a lieu, dans la réponse, que la présomption d'existence (de "non vacuité") est satisfaite ;
- Ⓓ nombre minimum d'éléments (modalité de réponse ou n -uple d'individus) requis par la sélection ;
- Ⓔ nombre maximum d'éléments requis pour la sélection ;
- Ⓕ exigence d'exhaustivité de la réponse, exprimée en proportion du nombre de réponses vraies possibles ;
- Ⓖ demande d'affirmation, s'il y a lieu, dans la réponse, que l'exigence d'exhaustivité est satisfaite ;
- Ⓗ exigence de distinction ou de non-redondance (pour les questions compréhensives seulement) ;
- Ⓙ demande d'affirmation, s'il y a lieu, dans la réponse, que l'exigence de distinction est satisfaite (pour les questions compréhensives seulement) ;
- ⓰ règles présidant à l'appariement des individus dans les n -uples (pour les questions compréhensives n -aires, avec $n > 1$), lorsque ces règles ne sont pas entièrement définies par le prédicat de la forme propositionnelle sur laquelle porte l'interrogation.

Ainsi, la requête d'une question extensive peut comporter sept arguments ou paramètres ; celle d'une question compréhensive unaire, neuf ; celle d'une question compréhensive n -aire (avec $n > 1$), dix. Par contre, la requête d'une alternative portant sur une contradiction ne nécessite aucun des dix éléments énumérés ci-dessus, puisque la valeur des arguments Ⓐ , Ⓓ , Ⓔ est fixée à l'avance, la présomption Ⓑ est une propriété inhérente à cette forme de question, les exigences Ⓕ et Ⓗ n'ont pas de sens dans ce cas particulier, et les affirmations Ⓒ , Ⓖ , et Ⓙ correspondant à ces exigences ou à la présomption de non vacuité n'ont pas à être demandées.

Il est aisé, en se fondant sur cette description de la requête, de construire une typologie formelle applicable, à quelques restrictions près,

aux questions extensives et aux questions compréhensives. Pour apprécier la puissance classificatoire de cette typologie, examinons combien de catégories elle permet de définir. Tout d'abord, pour l'exigence d'exhaustivité (f), nous ne considérerons que les deux cas extrêmes : aucune exigence, et exigence d'exhaustivité maximum (cf. [Belnap 1976 : 52]). Nous admettrons ensuite que les couples d'éléments de la requête : (b, c), (f, g), et (h, i), définissent chacun trois cas de figure ; en effet, dans notre perspective, il ne paraît pas utile d'envisager le cas où l'on demande au répondant d'affirmer dans la réponse qu'une exigence non exprimée dans la requête est cependant satisfaite (Notons cependant que cette possibilité est suggérée par [Kubiński 1970 : 64-65], puisque les opérateurs numériques n'expriment pas d'exigence d'exhaustivité ni de distinction, tandis que les réponses directes peuvent exprimer -facultativement- que ces exigences sont satisfaites). D'autre part, les éléments (a), (d) et (e) peuvent prendre, en théorie du moins, une infinité de valeurs. Pour le couple : (d, e), nous nous limiterons à trois cas : le nombre maximum u d'éléments de la sélection n'est pas fixé/il est fixé et est supérieur au nombre minimum ($u > v$)/il est fixé et égal au nombre minimum ($u = v$). Pour l'élément (a), nous considérerons que le nombre n de variables liées dans les questions compréhensives est égal à un , ce qui rend sans objet l'élément de requête (j) ; et nous ne tiendrons pas compte du nombre n de modalités de réponses dans les questions extensives. Malgré toutes ces restrictions, qui limitent évidemment le nombre de types envisageables, on aboutit :

- pour les questions compréhensives unaires, à : $3^4 = 81$ types de base ;
- pour les questions extensives présentant n modalités de réponses, à : $3^3 n = 27n$ types de base (soit : 243 formes possibles, si l'on considère qu'en pratique, on a : $2 < n < 10$) ;
- pour les alternatives portant sur une contradiction, à un type unique (ainsi que nous l'avons montré plus haut).

Soit un total de 109 types de base des requêtes élémentaires, rien que pour les cas usuels.

Éléments de la requête								Formes correspondantes dans les systèmes de logique érotétique étudiés
a	b	c	d	e	f	g		
-	∃1	-	1	1	-	-		BELNAP : q. <i>whether</i> /exemple unique = $? \left(\bigvee_1 - - \right) (A_1, \dots, A_n)$
n	∃1	-	1	1	-	-		ÅQVIST : q. <i>whether</i> /exemple unique = $?_n (A_1, \dots, A_n)$
-	∃1	-	1	n	-	-		BELNAP : q. <i>whether</i> /exemples multiples = $? \left(\bigvee_1 - - \right) (A_1, \dots, A_n)$
n	∃k	-	k	k	-	-		KUBIŃSKI : q. sur la valeur logique/ k exemples = $[1_k^n] \bigvee_k^n A_1, \dots, A_n$
-	∃1	-	1	1	V	-		STAHL : question de sélection = $[f_1 ? A_1] \cap \dots \cap [f_n ? A_n]$
-	∃1	-	1	1	V	+		BELNAP : q. <i>whether</i> /réponse vraie unique = $? \left(\bigvee_1 V - \right) (A_1, \dots, A_n)$
n	-	-	0	n	V	+		KUBIŃSKI : question conjonctive = $[\delta^n] \delta^n A_1, \dots, A_n$
-	∃1	-	1	1	V	+		HARRAH : q. disjonctive à n arguments = $A_1 \dots A_n ?$
n	∃1	-	1	1	V	+		ÅQVIST : q. <i>whether</i> /réponse vraie unique = $?_n^u (A_1, \dots, A_n)$
-	∃1	-	1	n	V	+		KUBIŃSKI : question disjonctive = $[\beta^n] \beta^n A_1, \dots, A_n$
n	∃1	-	1	n	V	+		BELNAP : q. <i>whether</i> /liste exhaustive = $? \left(\bigvee_1 V - \right) (A_1, \dots, A_n)$
n	∃1	-	1	n	V	+		ÅQVIST : q. <i>whether</i> /énumération exhaustive = $?_n^c (A_1, \dots, A_n)$
n	∃k	-	k	k	V	+		KUBIŃSKI : q. sur la valeur logique/liste exhaustive = $[r_k^n] r_k^n A_1, \dots, A_n$

Tableau n° 2.13 : Exemples de requêtes de questions extensives

Afin d'illustrer cette analyse du contenu des formes de requêtes, nous avons procédé à un relevé des exemples les plus significatifs parmi les formes de questions décrites par les auteurs que nous avons cités. Le tableau n° 2.13 présente treize formes de questions extensives, illustrant douze formes de requêtes différentes. Pour l'analyse des éléments de la requête, nous avons dû nous référer non seulement à l'analyse sémantique des opérateurs interrogatifs proposés par les auteurs, mais aussi à la forme des réponses directes à la question. Afin d'alléger la présentation du tableau, nous n'avons conservé que les éléments de requête pouvant figurer dans la requête d'une question extensive (désignés par la lettre qui figure dans l'énumération ci-dessus), et nous avons adopté les symboles suivants :

- $\exists 1$ = présomption d'existence d'au moins une modalité vraie ;
- $\exists k$ = présomption d'existence d'au moins k modalités vraies ;
- \forall = exigence d'exhaustivité maximum ;
- $+$ = demande exprimée ;
- $-$ = élément absent ou non défini dans la requête.

Le tableau n° 2.14 a été conçu selon le même principe pour les questions compréhensives unaires, à ceci près que l'élément \textcircled{a} (nombre de variables liées à l'opérateur interrogatif), toujours égal à 1 par définition, a été omis, et que le symbole \neq a été introduit pour l'exigence de distinction. Les dix-huit exemples sélectionnés dans ce tableau illustrent dix-sept formes de requêtes. En ce qui concerne les requêtes des questions compréhensives à plus d'une variable, les apports de BELNAP et même de KUBIŃSKI sont trop embryonnaires pour justifier un rappel dans ce bref bilan ; ces apports seront toutefois décrits et analysés dans une autre perspective au chapitre suivant (§ 3.1.2.d).

Éléments de la requête										Formes correspondantes dans les systèmes de logique érotétique étudiés
b	c	d	e	f	g	h	i			
31	-	1	-	-	-	-	-	-	-	BELNAP : q. <i>which</i> /exemples multiples = ? $(\bar{1} \quad - \quad -)$ (x//Px)
3k	-	k	-	-	-	-	-	-	-	KUBIŃSKI : q. numérique/lère catégorie/2e forme = k ≤ x Px
3k	-	k	k	-	-	-	-	-	-	KUBIŃSKI : q. numérique/lère catégorie/lère forme = kx Px
31	-	1	-	-	-	-	≠	+	+	BELNAP : q. <i>which</i> /exemples multiples et distincts = ? $(\bar{1} \quad - \quad \neq)$ (x//Px)
3k	+	k	-	-	-	-	≠	+	+	ÅQVIST : q. <i>what</i> /au moins k réponses vraies = $(? \quad \bar{1} \quad x)$ Px
31	-	1	1	1	-	-	-	-	-	BELNAP : q. <i>which</i> /exemple unique = ? $(\bar{1} \quad - \quad -)$ (x//Px)
3k	+	k	k	k	-	-	≠	+	+	ÅQVIST : q. <i>what</i> /énumération exhaustive/E = $(? \quad \bar{1} \quad x)$ Px
-	-	0	-	V	-	-	-	-	-	ÅQVIST : q. <i>what</i> /énumération exhaustive/B = $(? \quad \bar{1} \quad x)$ Px
-	-	0	-	V	+	-	-	-	-	ÅQVIST : q. <i>what</i> /énumération exhaustive/KB = $(? \quad \bar{1} \quad x)$ Px
31	-	1	-	V	+	-	-	-	-	KUBIŃSKI : q. numérique/type C = Cx Px
31	+	1	-	V	-	-	-	-	-	BELNAP : q. <i>which</i> /énumération exhaustive = ? $(\bar{1} \quad V \quad -)$ (x//Px)
31	+	1	-	V	-	-	-	-	-	ÅQVIST : q. <i>what</i> /énumération exhaustive/EB = $(? \quad \bar{1} \quad x)$ Px
31	+	1	-	V	+	-	-	-	-	ÅQVIST : q. <i>what</i> /énumération exhaustive/EKB = $(? \quad \bar{1} \quad x)$ Px
3k	-	k	-	V	+	-	-	-	-	KUBIŃSKI : q. numérique/2e catégorie/2e forme = (k ≤ x) Px
3k	-	k	k	V	+	-	-	-	-	KUBIŃSKI : q. numérique/2e catégorie/lère forme = (k)x Px
31	-	1	-	V	+	≠	+	+	+	BELNAP : q. <i>which</i> /énumération exhaustive et distincte = ? $(\bar{1} \quad V \quad \neq)$ (x//Px)
31	-	1	1	V	+	-	-	-	-	BELNAP : q. <i>which</i> /réponse vraie unique = ? $(\bar{1} \quad V \quad -)$ (x//Px)
3k	+	k	k	V	+	≠	+	+	+	ÅQVIST : q. <i>what</i> /exactement k réponses vraies = $(? \quad \bar{1} \quad x)$ Px

Tableau n° 2.14 : Exemples de requêtes de questions compréhensives unaires.

d) La notion de réponse directe à une question

Dans la quasi-totalité des textes que nous avons analysés, la distinction entre syntaxe et sémantique paraît moins aisée à établir pour les réponses que pour les questions. Cependant, les réponses directes à une question peuvent être dérivées de la question elle-même par des procédures purement syntaxiques. Or, nous l'avons vu, le concept de *réponse directe* à une question est une notion centrale : à partir de ce concept et de la relation d'implication, il est possible de définir un grand nombre de catégories (sémantiques) de réponses. Sur le plan syntaxique, la réponse directe s'oppose aux autres formes de réponses entières (dont nous verrons les variétés au § 2.4.2.), ainsi qu'aux *réponses codées* et aux *répliques*.

Nous avons déjà mentionné (au § 2.1.3.b) le premier postulat de HAMBLIN : "Toute réponse à une question est une proposition. Il n'est nécessaire de postuler ceci que parce que l'on pense parfois que l'on peut répondre à une question avec quelque chose qui est *moins* qu'une proposition ; que, par exemple, si l'on me demande mon nom, je peux répondre en donnant simplement mon nom, ce qui n'est pas une proposition mais un simple nom propre. Il suffit de répondre à cela que, dans le contexte de la question, donner son nom équivaut à énoncer la proposition que *c'est* mon nom" [Hamblin 1958 : 162]. D'où la notion de réponse *abrégée*, ou réponse *codée* ; d'où également les catégories distinguées par DAMOURETTE et PICHON (réponse *anaphorisée* et réponse *zeugmatique* ; cf. § 1.2.4.c), opposées par eux à la réponse *entière*, qui est une phrase énonciative complète. S'il est vrai que toute réponse est une proposition, il n'en est pas moins vrai que toute proposition ne constitue pas une réponse à une question donnée (cf. le deuxième postulat de HAMBLIN). On donnera, avec BELNAP, le nom de *réplique* à tout élément du discours faisant suite à une question, et ne constituant pas une réponse (entière ou codée) à celle-ci.

Parmi les réponses entières possibles à une question, les *réponses directes* (appelées aussi *réponses propres* par AJDUKIEWICZ, *réponses attendues* par STAHL) constituent un sous-ensemble particulièrement intéressant.

En effet, "on peut toujours aboutir à une réponse directe à une question à travers une transformation finie et entièrement déterminée de cette question" [Kubinski 1968 : 185]. Par exemple, nous avons vu (au § 2.3.2.c) qu'il suffisait de deux transformations élémentaires pour passer d'une question numérique formalisée à l'une de ses réponses directes : "Par le moyen de transformations analogues, toute question formalisée peut être transformée en une réponse directe (vraie ou fausse)" [Kubiński 1968 : 186]. Nous en avons décrit la procédure dans les systèmes de KATZ (§ 2.1.2.b) et d'ÅQVIST (§ 2.3.1.d). La démonstration en serait facile à faire dans le système de BELNAP ; elle est triviale dans ceux de STAHL et de HARRAH, puisque ces auteurs identifient les questions à la classe de leurs réponses possibles. Réciproquement, en corrolaire aux deux premiers postulats de HAMBLIN, on peut reconstituer une question à partir de l'ensemble de ses réponses directes possibles.

La meilleure description morphologique de la réponse directe se trouve, à notre avis, chez BELNAP, qui distingue trois éléments constitutifs : la sélection (seul constituant obligatoire), l'affirmation d'exhaustivité, et l'affirmation de distinction ; les éléments constitutifs présents dans la réponse sont liés par la relation de conjonction (cf. [Belnap 1976 : 42]). Cette description est adoptée également par KUBIŃSKI (cf. [Kubinski 1970] : 64-65 , et est compatible avec les conceptions d'ÅQVIST, de HARRAH, et de STAHL. Elle est assez différente par contre des descriptions des linguistes, qui pour la plupart ne prennent pas en considération les affirmations d'exhaustivité et de distinction. La présentation morpho-syntaxique de BELNAP met en évidence la différence qu'il y a entre la *sélection* de la réponse, et les *modalités* de réponse présentées (par exemple) par une question extensive ; la sélection étant, dans le cas général, une conjonction de modalités. Ainsi, le nombre de *réponses directes possibles* non seulement ne se confond pas avec le nombre n de modalités présentées dans le sujet de la question extensive, mais il dépend en outre des valeurs minimum (v) et maximum (u) du nombre de modalités devant être sélectionnées (nous examinerons cette relation de manière détaillée en analysant les formes de consignes observées dans notre corpus ; cf. § 3.1.2.).

Le passage de la question à ses réponses directes est présenté différemment selon les auteurs. Toutefois, les transformations nécessaires dans la plupart des systèmes de logique érotétique sont les suivantes (en utilisant pour les décrire une terminologie proche de celle de BELNAP). Pour les *questions extensives* :

- Ⓐ Sélection, dans l'ensemble des modalités de réponse (propositions) présentées par le sujet de la question, de k modalités (avec : $v \leq k \leq u$ si u est défini, et : $v \leq k$ sinon).
- Ⓑ Constitution du contenu de la sélection S de la réponse, en formant une proposition complexe par conjonction des modalités (propositions) sélectionnées.
- Ⓒ S'il y a lieu (exigence d'exhaustivité maximum), constitution du contenu de l'affirmation d'exhaustivité C , en formant une proposition complexe par conjonction des négations des k modalités (propositions) non sélectionnées.
- Ⓓ Affirmation de S ou, s'il y a lieu, de la conjonction $S \& C$.

Pour les *questions compréhensives* :

- Ⓐ Constitution d'un ensemble de k propositions (avec : $v \leq k \leq u$ si u est défini, et : $v \leq k$ sinon) à partir de la forme propositionnelle $Px_1 \dots x_n$ contenue dans le sujet de la question, par substitution aux variables x_i de constantes $a_{i,j}$ sélectionnées dans le domaine des individus D , ou s'il y a lieu (existence de conditions de catégorie sur les variables) dans le sous-ensemble de D associé au prédicat catégoriel C_i pour chaque x_i .
- Ⓑ Constitution du contenu de la sélection S par conjonction des k propositions formées à partir de la forme propositionnelle du sujet.

ⓐ S'il y a lieu (exigence d'exhaustivité maximum), constitution du contenu de l'affirmation d'exhaustivité S sous la forme d'une proposition complexe exprimant que, pour toute combinaison de valeurs des x_i autre que celles figurant dans la sélection S , la propriété $Px_1 \dots x_n$ n'est pas vérifiée.

ⓓ S'il y a lieu (exigence de distinction), constitution de l'affirmation de distinction D sous la forme d'une proposition complexe exprimant qu'il n'existe pas dans S deux combinaisons de valeurs des x_i se référant au même sous-ensemble d'individus.

ⓔ Affirmation de S ou, s'il y a lieu et selon les cas, des conjonctions $S \& C$, $S \& D$, ou $S \& C \& D$.

Pour les *alternatives portant sur une contradiction*, la réponse directe est obtenue simplement par sélection et affirmation de l'une des deux modalités de réponse possibles ; l'affirmation d'exhaustivité maximum est impliquée par le caractère contradictoire des deux réponses possibles (qui ici se confondent avec les modalités).

2.4.2. Sémantique

Qu'est-ce qu'une question valide ? Quelles sont les conditions de validité ou d'invalidité d'une question ? Une question valide est-elle toujours soluble ? A quoi reconnaît-on qu'une réponse autre qu'une réponse directe peut toutefois répondre effectivement à une question donnée ? Nous l'avons vu, tous ces problèmes sont au coeur de l'analyse sémantique des questions et des réponses. Nous rappelons brièvement ci-après les principales propriétés sémantiques des questions ; nous présentons ensuite une synthèse de la plupart des classifications sémantiques des questions et des réponses que nous avons passées en revue.

a) Propriétés sémantiques des questions

L'analyse sémantique des questions s'appuie sur deux constatations (ou : postulats, si l'on préfère) : l'existence, dans toute interrogation, d'un contenu propositionnel, auquel on peut donner la forme d'une expression logique ; et la possibilité d'exprimer les relations entre une question et diverses expressions sous la forme de relations d'implication entre ces expressions et le contenu propositionnel de la question. Ce deuxième point conduit évidemment à poser le problème de la valeur de vérité de la question elle-même. Il permet en tous cas de définir clairement les notions de validité et de solubilité des questions.

L'existence d'un *contenu propositionnel* de toute question est maintenant communément admise par les logiciens. Nous avons rappelé (au § 2.1.2.a) le passage de la *Logique de Port Royal* dans lequel ARNAULT et NICOLE énoncent l'affirmation contenue dans une question particulière (\approx "Le juge vend-il toujours la justice ?"). Cette remarque a été généralisée depuis ; par exemple : "il est clair que lorsque nous posons une question nous faisons également une affirmation (par exemple, que Marie a porté un et un seul des deux chapeaux rouge et vert), et que les questions, comme les propositions, ont un contenu informationnel" [Belnap 1963 : 129-130]. Ce contenu propositionnel peut être assimilé à ce qu'AJDUKIEWICZ, dans ses premiers travaux sur la logique érotétique, appelle le *datum questionis* (cf. [Kubiński 1958 : 320 n 1], et [Ajdukiewicz 1965 : 86]), ou encore à l'*objet* de la question selon LEONARD (*the topic of concern* = "ce dont il s'agit" dans la question, qui peut être exprimé par une proposition [Leonard 1959 : 178-179]). Le rôle joué, dans le dialogue, par le contenu propositionnel des questions est souligné par HARRAH : "Celui qui interroge exprime ce qu'il sait relativement à un sujet donné en disant : « C'est ceci ou cela » , ou « Certaines choses sont telles et telles » ; l'interlocuteur peut alors décider de répondre à cette proposition par une contre-proposition qui apporte un surcroît d'information, diminuant ainsi l'incertitude du locuteur. Il peut affirmer : « C'est ceci ou non cela » , ou : « Les choses qui sont telles et telles sont A et B » " [Harrah 1961 : 41].

Le contenu propositionnel peut par conséquent être exprimé soit par une proposition, soit par une forme propositionnelle avec quantificateur. La théorie d'ÅQVIST permet à la fois de rendre compte du passage de la question à son contenu propositionnel, et de distinguer le contenu propositionnel de la présupposition de la question. En effet, le noyau de la question selon ÅQVIST est la contrepartie formelle du contenu propositionnel tel que le décrit HARRAH. Par exemple, à la question : "Est-ce Brutus ou Cassius qui est l'(unique) assassin de Jules César ?" correspond le noyau : "Je sais que l'assassin de Jules César est soit Brutus, soit Cassius" [Åqvist 1965 : 61] ; de même, pour la question : "Quels sont les trois nombres premiers compris entre 10 et 20 ?", la formule symbolique du noyau (cf. [Åqvist 1965 : 96 ex. 3b]) peut être paraphrasée en : "Il existe trois nombres dont je sais qu'ils sont distincts, et qu'ils constituent les trois seuls nombres premiers compris entre 10 et 20". Le passage de l'énoncé interrogatif selon ÅQVIST au noyau de la question s'effectuant, nous l'avons vu, par effacement de l'opérateur déontique. Le passage du contenu propositionnel ainsi défini à la présupposition de la question correspond à l'effacement de tous les opérateurs épistémiques. On obtient par conséquent, pour les deux énoncés ci-dessus, les présuppositions : "L'assassin de Jules César est soit Brutus, soit Cassius", et : "Il existe exactement trois nombres premiers distincts compris entre 10 et 20".

La définition des relations formelles entre une question, son contenu propositionnel, sa présupposition, et l'ensemble de ses réponses possibles, fait appel à l'implication logique. Son usage soulève par conséquent le problème de la valeur de vérité d'une question. BELNAP a réussi à résoudre ce problème d'une manière à la fois élégante (sans trop heurter les adversaires d'une assimilation des questions aux expressions) et pratique (aboutissant à une classification sémantique des questions couvrant les cas le plus usuels), en recourant au concept de "quasi-expression". Un examen plus poussé des thèses de BELNAP conduirait peut-être à la conclusion qu'il est possible de se passer de ce concept, et de rester dans les limites du calcul appliqué des prédicats du premier ordre avec identité, en remplaçant la notion de question

par celle de contenu propositionnel de la question (d'ailleurs BELNAP, contrairement à ÅQVIST, ne distingue pas clairement ces deux notions, cf. [Belnap 1976 : 121] ; toutefois, il n'affirme nulle part qu'elles sont identiques, et il précise même que, pour l'interprétation des relations entre quasi-expressions, c'est le contenu propositionnel qui doit être pris en compte plutôt que la question elle-même [Belnap 1976 : 122]). Tel qu'il se présente cependant, le système proposé par BELNAP nous a paru constituer un cadre satisfaisant pour une analyse sémantique des questions.

Comme nous l'avons vu (au § 2.2.5.a), la conception de BELNAP se fonde sur une double relativisation de la valeur de vérité d'une expression (ou d'une quasi-expression). Comme dans toute forme du calcul des prédicats, une expression (ou quasi-expression) donnée A est dite vraie (ou fausse) dans une interprétation M donnée pour une assignation de valeurs α donnée ; c'est-à-dire pour une application donnée α de l'ensemble des variables d'individu de l'expression dans l'ensemble des individus, selon la fonction d'interprétation définie par M (cf. § 2.2.1.a). Mais en outre, la valeur de vérité de l'expression (ou de la quasi-expression) A peut être liée aux valeurs de vérité des expressions (ou quasi-expressions) constituant un ensemble H , dont A peut être ou n'être pas un élément. En conséquence, et à titre d'illustration, si toutes les expressions (ou quasi-expressions) éléments de H sont vraies dans l'interprétation M pour une assignation α , et si A est impliquée par certaines expressions (ou quasi-expressions) éléments de H , alors A est vraie dans l'interprétation M pour l'assignation α *relativement à H* . Tel serait pas exemple le cas de la proposition $A =$ "Ou bien Jean a cessé de fumer, ou bien il fume toujours", si l'ensemble H a pour élément la proposition $B =$ "Jean avait l'habitude de fumer", et si tous les éléments de H sont vrais dans l'interprétation M pour l'assignation α .

b) Classification sémantique des questions

La *validité* d'une question (ou sa *vérité*, selon BELNAP et KUBIŃSKI) dépend de l'existence d'au moins une réponse directe vraie à la question. Aussi cette notion, et celles qui en sont dérivées, sont-elles susceptibles d'être relativisées. Mais cette relativisation peut être effectuée de trois manières :

- soit, selon les règles usuelles du calcul des prédicats, en ne tenant compte que de l'interprétation M ;

- soit en mettant entre parenthèses le problème de la vérité de l'ensemble H , et en ne tenant compte que des relations d'implication concernant (au moins) une expression (ou quasi-expression) de cet ensemble ;

- soit en tenant compte à la fois de l'interprétation M et des relations d'implication de H .

On aboutit de la sorte à quatre systèmes de classification sémantique des questions distincts, selon que le problème de la validité de la question est considéré absolument, relativement à une interprétation M , relativement à un ensemble H d'expressions, ou relativement à la fois à M et à H . Chacun de ces systèmes de classification présente un intérêt pour l'analyse des situations concrètes de questionnement.

Le premier système de classification suppose l'existence de réponses (directes) vraies absolument ou fausses absolument. Ainsi que nous l'avons vu, certaines formes particulières de questions se caractérisent par des propriétés logiques remarquables, telles que le fait de n'avoir que des réponses vraies (sur le modèle des expressions analytiques), ou de n'avoir que des réponses fausses (sur le modèle des expressions contradictoires). Ce sont les formes de questions les moins intéressantes pour le praticien. Nous avons vu d'autre part que Friedrich LÖW détermine l'existence d'une réponse vraie à une question donnée par l'existence d'un état de fait objet de la recherche de la

question (§ 2.3.4.b) ; sur le plan pratique toutefois, cette règle ne permet pas d'affirmer *a priori* qu'une question est ou n'est pas valide absolument, puisqu'il n'est pas toujours possible de s'assurer de l'existence d'un état de fait. La solution proposée par LÖW ne fait donc que reculer la difficulté, et l'on peut craindre d'ailleurs que, posé de cette manière, le problème ne soit pas soluble dans le cas général. Par contre, la réponse apportée par Moritz SCHLICK présente l'avantage de fournir un critère d'évaluation relativement simple de l'existence possible d'une réponse vraie : "Un examen attentif révèle que toutes les différentes manières d'expliquer ce que signifie effectivement une question ne sont en dernière analyse rien d'autre que diverses descriptions de la manière dont la réponse à la question doit être trouvée. Toute explication ou description de la signification d'une réponse consiste, d'une manière ou d'une autre, en règles à appliquer pour la découverte de sa réponse" [Schlick 1935 : 102]. Le fait qu'il ne soit pas possible (pour des raisons contingentes) de recourir à la procédure décrite peut rendre la question insoluble empiriquement (dans un contexte donné), sans pour autant mettre en cause la validité de la question.

Pour être complète, une classification des questions selon leur validité absolue doit tenir compte à la fois de l'existence d'au moins une réponse vraie, et de l'existence d'au moins une réponse fausse. On aboutit ainsi aux quatre types suivants :

- question *valide absolument*, pour laquelle il existe au moins une réponse directe vraie absolument et au moins une réponse directe fausse absolument. C'est la définition même de la question *posée proprement* d'AJDUKIEWICZ (les trois autres types ci-dessous constituant la classe des questions *posées improprement* ; cf. [Ajdukiewicz 1965 : 88]). L'alternative portant sur une contradiction est, par construction, toujours valide absolument.

- question *insoluble absolument*, pour laquelle il n'existe ni réponse directe vraie absolument, ni réponse directe fausse absolument. KATZ a décrit (sous le nom de question *insoluble linguistiquement*) une variété de questions invalides absolument dont la présupposition n'a pas de sens, parce qu'elle contient des termes incompatibles entre eux (contradictaires, contraires, inverses;

cf. [Katz 1972 : 157-171]) : la présupposition est alors ou bien une contradiction, ou bien une proposition qui n'a pas de référent possible. Une forme intéressante de question dont la présupposition est une contradiction est signalée par BELNAP : elle se caractérise par l'existence d'une contradiction entre le sujet de la question et sa requête (exemple : "Parmi les deux phrases suivantes, citez-en au moins trois qui soient vraies" [Belnap 1976 : 133]). Nicholas RESCHER appelle question *illégitime* une question qui n'a aucune réponse possible [Rescher 1974 : 32] ; une telle question est évidemment invalide absolument.

- question *évidente*, dont toutes les réponses possibles sont vraies absolument ; par exemple, parce que toutes ses réponses directes sont des tautologies (question *analytique* selon KATZ).

- question *absurde*, dont toutes les réponses possibles sont fausses absolument. A ce type appartiennent les questions *contradictoire*s de KATZ, dont toutes les réponses directes sont des contradictions.

Cette présentation condensée de la première classification ne rend pas compte de toutes les variétés décrites par les logiciens des questions. En premier lieu, nous avons vu que certains auteurs donnent de la question valide absolument une définition plus restrictive ; c'est en particulier le cas de Friedrich LÖW, pour qui seule est valide absolument une question ayant une et une seule réponse vraie ; c'est également le cas de David HARRAH, qui adopte, pour les questions extensives, un point de vue analogue à celui de LÖW. Ces conceptions supposent toujours, d'une certaine manière, l'exigence d'exhaustivité, ce qui est loin d'être le cas général. Dans le même ordre d'idées, pour Felix COHEN, une question ayant plus d'une réponse directe vraie est considérée comme valide certes, mais *indéterminée* (par analogie avec une équation algébrique). En second lieu, un auteur comme Moritz SCHLICK, qui s'est intéressé au problème de la solubilité des questions, a bien mis en évidence le fait que l'existence d'au moins une réponse directe vraie absolument ne garantit pas la possibilité de découvrir cette réponse : une question valide absolument peut être également insoluble empiriquement au moment où elle est posée, qu'il s'agisse du poids d'HOMÈRE lorsqu'il a commencé à écrire

l'*Iliade*, ou de la nature d'un objet se trouvant sur la face cachée de la lune (en 1935 ; cf. [Schlick 1935 : 99-101]). Nuel D. BELNAP a poussé plus loin l'analyse en citant des exemples de questions valides absolument et insolubles logiquement, comme : "Quels sont tous les nombres premiers ?" (cf. [Belnap 1972 : 339]) ; ou en mentionnant l'impossibilité effective de répondre à une question pourtant valide et soluble, telle que : "Citez un exemple de nombre dont le nom n'a jamais été, n'est pas, et ne sera jamais cité par qui que ce soit" (exemple dû à Nicholas RESCHER ; cf. [Belnap 1976 : 133]). Enfin, le fait qu'une question ait une réponse directe à la fois vraie absolument et effectivement énonçable, ne garantit pas la valeur informative de la réponse : cette valeur est nulle lorsque la réponse vraie se trouve logiquement impliquée par l'énoncé de la question. Cette dernière variété de question est appelée *résolue linguistiquement* (KATZ), *rhétorique* (BELNAP), ou encore *contenant sa propre réponse* ("self-answering").

Les trois autres systèmes de classifications peuvent être calqués sur le premier. En nous limitant à l'essentiel pour le second système, nous ne rappellerons que la dichotomie décrite par BELNAP et KUBIŃSKI :

- question *vraie dans une interprétation M* , pour laquelle il existe au moins une réponse directe vraie dans M . Pour les questions compréhensives, on peut, comme le fait BELNAP, distinguer entre vérité *nominale* et vérité *réelle* [Belnap 1976 : 119].

- question *fausse dans une interprétation M* , pour laquelle il n'existe aucune réponse directe vraie dans M (c'est-à-dire : dont toutes les réponses directes sont fausses dans M).

Le troisième système de classification est fondé sur la *consistance* du contenu propositionnel de la question avec un ensemble H d'expressions. Il est intéressant, d'un point de vue pratique, de faire abstraction de la valeur logique des expressions éléments de H , pour n'étudier que les relations entre une question donnée et cet ensemble d'expressions. En effet, dans une

situation concrète donnée, il n'est pas toujours possible de déterminer si une expression est vraie ou fausse. Dans le contexte par exemple d'un entretien d'enquête, la présupposition d'une question d'opinion, exprimant une croyance de l'auteur du questionnaire, peut être en contradiction avec une croyance de la personne interrogée, voire de l'enquêteur lui-même ; en principe, seul le recul historique permettra de décider du bien fondé de la question. Par contre, il est clair que, dans la situation d'entretien, l'enquêteur doit *assumer* la présupposition de la question, c'est-à-dire faire comme si elle était vraie, ou pouvait être vraie dans certaines circonstances (cf. [Collingwood 1940 : 27-29]). D'ailleurs, nous avons vu que certains auteurs (comme CARNAP et WHEATLEY) considèrent que seules ont une valeur de vérité les questions *théoriques* (ou *d'investigation*) ; les questions *pratiques* (ou *de délibération*) suscitant non une réponse mais une décision, qui, en elle-même, ne peut être jugée vraie ou fausse. Et WHEATLEY range parmi les questions de délibération certains items de questionnaires d'attitude ou d'opinion (ex. : "A votre avis, y aura-t-il dans les vingt prochaines années une troisième guerre mondiale ?" ; cf. [Wheatley 1955 : 51-52]), pour lesquels la personne interrogée doit décider en l'absence d'informations suffisantes et assurées (cf. également sur ce point [Leonard 1967 : 31-32]).

C'est pourquoi l'apport de HARRAH et BELNAP dans ce domaine nous paraît particulièrement important pour l'analyse des questionnaires. BELNAP, rappelons-le, distingue dans cette optique les types de questions suivants :

- question *soluble par H* , dont au moins une réponse directe est impliquée par H .
- question *insoluble par H* , dont aucune réponse directe n'est impliquée par H .
- question *rhétorique relativement à H* , dont au moins une réponse directe est impliquée par l'union de H et du contenu propositionnel de la question , $H \cup \{I\}$.
- question *discutable relativement à H* (ou : *non-rhétorique relativement à H*), dont aucune réponse directe n'est impliquée par $H \cup \{I\}$.

- question *fortement discutable relativement à H*, telle que $H \cup \{I\}$ n'implique ni au moins une réponse directe, ni la négation d'au moins une réponse directe.

BELNAP illustre certaines de ces formes de questions en considérant l'ensemble H_q des connaissances et des croyances de la personne qui pose la question, et l'ensemble H_r des connaissances et des croyances de la personne interrogée. Dans une situation de demande d'information, la question est discutable relativement à H_q ; dans une situation d'examen, elle est rhétorique relativement à H_q . D'autre part, selon que la personne interrogée est capable ou non de fournir la réponse attendue, la question est soluble ou insoluble relativement à H_r .

Le quatrième système de classification sémantique des questions tient compte à la fois d'un ensemble d'expressions H , et de l'ensemble des interprétations M_i dans lesquelles les expressions de H sont toutes vraies. BELNAP construit ce système à partir de deux dichotomies. La première oppose les types suivants :

- question *sûre relativement à H*, pour laquelle, dans toutes les interprétations M_i dans lesquelles H est vrai, il existe au moins une réponse directe vraie impliquée par H .

- question *risquée relativement à H*, pour laquelle il n'existe pas au moins une réponse directe vraie impliquée par H dans toutes les interprétations M_i dans lesquelles H est vrai.

La seconde dichotomie oppose :

- question *stupide relativement à H*, pour laquelle, dans toutes les interprétations M_i dans lesquelles H est vrai, toutes les réponses sont fausses.

- question *possible relativement à H*, pour laquelle il existe au moins une des interprétations M_i dans lesquelles H est vrai, dans laquelle il existe au moins une réponse vraie.

Une même question peut par conséquent être à la fois risquée et possible relativement à H .

c) Classification sémantique des réponses

Si nous nous limitons aux grandes catégories syntaxiques de questions, pour lesquelles il n'y a pas de divergences importantes entre les logiciens (c'est-à-dire les questions extensives et les questions compréhensives), l'analyse sémantique des réponses peut être conduite par référence aux *réponses directes*, sur la base des relations d'implication pouvant exister entre une réponse quelconque à une question et une ou plusieurs réponses directes à la même question, ou leur négation. En nous inspirant pour l'essentiel de BELNAP, nous considérerons huit types de relations :

- $\alpha = \exists d(I), \quad r(I) \rightarrow d(I)$
= la réponse implique au moins une réponse directe.
- $\beta = \forall d(I), \quad r(I) \rightarrow d(I)$
= la réponse implique chacune des réponses directes.
- $\gamma = \exists d(I), \quad r(I) \leftarrow d(I)$
= la réponse est impliquée par au moins une réponse directe.
- $\delta = \forall d(I), \quad r(I) \leftarrow d(I)$
= la réponse est impliquée par chacune des réponses directes.
- $\varepsilon = \exists d(I), \quad r(I) \rightarrow \overline{d(I)}$
= la réponse implique la négation d'au moins une réponse directe.
- $\zeta = \forall d(I), \quad r(I) \rightarrow \overline{d(I)}$
= la réponse implique la négation de chacune des réponses directes.
- $\eta = \exists d(I), \quad r(I) \leftarrow \overline{d(I)}$
= la réponse est impliquée par la négation d'au moins une réponse directe.
- $\theta = \forall d(I), \quad r(I) \leftarrow \overline{d(I)}$
= la réponse est impliquée par la négation de chacune des réponses directes.

Pour une réponse quelconque $r(I)$ donnée, chacune des huit expressions ci-dessus peut être vraie ou fausse (soit en principe 2^8 combinaisons). En réalité, certaines combinaisons (conjonctions d'expressions) sont impossibles parce que contradictoires ; on ne peut avoir par conséquent ni : $\bar{\alpha} \& \beta$, ni $\beta \& \varepsilon$. D'autre part, parmi l'ensemble des combinaisons possibles, toutes n'ont pas retenu l'attention des logiciens. Enfin, bien que cet ensemble de propriétés possibles des réponses à une question ne soit pas suffisant pour une analyse sémantique, il ne semble pas avoir été exploité systématiquement ; en particulier, certaines relations entre ces propriétés devraient déboucher sur la formalisation de théorèmes (comme nous le verrons à propos des réponses partielle et éliminatoire).

C'est pourquoi, en attendant que soit menée à bien une analyse sémantique des réponses comparable, en rigueur et en étendue, à l'analyse syntaxique des questions par exemple, nous ne pouvons que rappeler, dans un cadre formel qui en rende possible la comparaison, les principales catégories sémantiques de réponse mentionnées dans les travaux que nous avons analysés. Pour simplifier l'exposé, nous admettrons :

- que toutes les réponses directes à une question donnée sont réellement distinctes (il n'y a pas deux réponses directes distinctes nominale-ment qui expriment cependant la même chose, en désignant par exemple les mêmes référents) ;

- qu'aucune réponse directe n'est soit une tautologie, soit une contradiction ;

- que toutes les réponses directes sont mutuellement exclusives (troisième postulat de HAMBLIN) ;

- que la réponse quelconque considérée $r(I)$ n'est jamais elle-même une réponse directe $d(I)$;

- que, dans la conjonction de deux expressions existentialisées, il s'agit de la même réponse directe $d(I)$; ou, ce qui revient au même, que le quantificateur existentiel s'applique à la conjonction des deux expressions. Par exemple, la conjonction : $\alpha \& \gamma$ est mise pour : $\exists d(I), r(I) \leftrightarrow d(I)$.

Si nous considérons d'abord les quatre premières relations formelles possibles entre une réponse quelconque et les réponses directes à une question donnée, nous obtenons en premier lieu trois types de réponses de base :

- α & $\bar{\beta}$ & $\bar{\gamma}$ & $\bar{\delta}$ = réponse *indirecte propre* (KUBIŃSKI), caractérisée par l'apport d'une information surabondante par rapport à celle que fournit une réponse directe.

- α & $\bar{\beta}$ & γ & $\bar{\delta}$ = réponse *exactement complète* (BELNAP) ou : réponse *presque directe* (KUBIŃSKI), équivalant strictement sur le plan sémantique à une réponse directe, sans toutefois en avoir les propriétés syntaxiques.

- $\bar{\alpha}$ & $\bar{\beta}$ & γ & $\bar{\delta}$ = réponse *partielle fortement régulière* (BELNAP), qui apporte au locuteur une certaine information, mais moindre toutefois que celle que lui apporterait une réponse directe à la question.

Si l'on accepte d'autre part de ne considérer simultanément que certaines des quatre propriétés possibles, on obtient des catégories de réponse plus larges, pouvant contenir l'une des trois formes de base définies ci-dessus :

- α & $\bar{\beta}$ & $\bar{\delta}$ = réponse *complète* (BELNAP, ÅQVIST), *suffisante* (ÅQVIST), *indirecte* (KUBIŃSKI, KATZ), *indirecte complète* (AJDUKIEWICZ).

- $\bar{\beta}$ & γ & $\bar{\delta}$ = réponse *partielle* (HARRAH) ou : *partielle régulière* (BELNAP).

- $\bar{\delta}$ = réponse *non informative* (BELNAP) ou : *élusion de la question* (KATZ).

Si nous considérons de même les quatre autres propriétés, relatives à la négation des réponses directes, nous obtenons les trois types de base suivants :

- ε & $\bar{\zeta}$ & $\bar{\eta}$ & $\bar{\theta}$ = réponse *éliminatoire fortement régulière* (BELNAP), qui entraîne la négation d'une partie (et d'une partie seulement) de l'ensemble des réponses directes possibles, en fournissant une information surabondante par rapport à celle qu'apporterait la négation d'une réponse directe.

- ε & $\bar{\zeta}$ & η & $\bar{\theta}$ = réponse *exactement éliminatoire* (BELNAP), qui équivaut exactement à la négation d'une (et d'une seule) réponse directe.

- $\bar{\varepsilon}$ & $\bar{\zeta}$ & η & $\bar{\theta}$ = réponse *quasi-éliminatoire régulière* (ainsi dénommée en s'inspirant de BELNAP), qui apporte une certaine information, mais toutefois moindre que celle apportée par la négation d'une réponse directe.

On obtient en outre, à partir de certaines des quatre propriétés relatives à $\overline{d(I)}$, les catégories de réponse plus étendues suivantes :

- ε & $\bar{\zeta}$ & $\bar{\theta}$ = réponse *éliminatoire* (BELNAP).

- $\bar{\zeta}$ & η & $\bar{\theta}$ = réponse *quasi-éliminatoire* (BELNAP).

- ζ = réponse *corrective* (BELNAP, ÅQVIST), *rejet de la question* (KATZ).

Nous ne sommes pas toujours parvenu à insérer dans ce schéma général les variétés de réponse distinguées par les logiciens. Dans certains cas, cela est dû à la pauvreté relative de la description fournie. Par exemple, ce que STAHL appelle une réponse *inattendue* nous a semblé correspondre à la réponse *corrective* de BELNAP et ÅQVIST, mais ceci à partir de l'exemple donné par STAHL, et non selon les propriétés de cette forme de réponse, qui ne sont pas explicitées par l'auteur.

Dans d'autres cas, nous n'avons pas voulu forcer l'interprétation en tirant, des propriétés attribuées aux formes décrites, des conclusions qui auraient permis l'insertion de ces formes dans notre schéma. Tel est le

cas en particulier de la réponse *partielle* telle qu'elle est définie par AJDUKIEWICZ et par KUBIŃSKI. Pour le premier, "une *réponse partielle* est une proposition qui n'implique pas une réponse propre, mais qui élimine certaines réponses propres de l'ensemble des réponses propres possibles. C'est donc une proposition qui [...] implique une disjonction de certaines réponses propres, mais non de toutes" [Ajdukiewicz 1965 : 90]. Comme, pour AJDUKIEWICZ, une réponse *propre* est ce que la plupart des auteurs appelle une réponse *directe*, il semble bien que la réponse *partielle* est ici en réalité une réponse *éliminatoire fortement régulière* ; impression confirmée par l'un des exemples illustrant cette définition : "Qui a fait cela ? - Ce n'est pas moi". Cependant, l'autre exemple (que nous avons cité au § 2.3.4.a : "Qui a découvert l'Amérique ? - Un Italien") suggère plutôt l'assimilation de cette forme à une réponse *partielle fortement régulière*. Cette difficulté de classement est due à une propriété des questions demandant une réponse vraie unique (cas particulier dans lequel AJDUKIEWICZ se place implicitement), que l'on pourrait formuler ainsi : "Toute réponse *partielle fortement régulière* à une question demandant une réponse unique est également une réponse *éliminatoire fortement régulière* à cette question". D'autre part, la réponse *partielle* selon KUBIŃSKI semble bien, en première analyse, correspondre à la réponse *partielle fortement régulière* de BELNAP. En effet, KUBIŃSKI la définit comme une expression qui n'est pas et n'a pas pour conséquence une réponse directe, et telle qu'il existe une autre expression n'étant pas et n'ayant pas pour conséquence une réponse directe, dont la conjonction avec la réponse partielle a pour conséquence (ou équivaut à) une réponse directe [Kubiński 1970 : 66]. Exemple de réponse partielle : "Quelles sont toutes les capitales de district sur l'Oder ? - Wroclaw est une capitale de district sur l'Oder" ; autre exemple de réponse partielle à la même question : "Opole est une capitale de district sur l'Oder et Stettin est une capitale de district sur l'Oder". La conjonction de ces deux réponses partielles équivaut à une réponse directe à la question posée. On peut en tirer la propriété suivante : "Pour toute réponse *partielle fortement régulière*, il existe une réponse *partielle fortement régulière* distincte, telle que leur conjonction soit une réponse exactement complète". En pratique cependant, il n'est pas toujours aisé de formuler en langage naturel cette réponse partielle complémentaire.

Enfin, certaines formes de réponse révèlent l'insuffisance de notre schéma de classification. En effet, les définitions proposées font intervenir les notions :

- d'affirmation et de négation des modalités de réponse (KUBIŃSKI : réponses *positive* et *négative*) ;

- de vérité et de fausseté de la réponse (AJDUKIEWICZ : réponse *exhaustive* = conjonction de toutes les réponses directes vraies).

D'autre part, il existe des formes de réponses se situant à un niveau *métalinguistique* (selon l'expression de HARRAH), qui constituent en quelque sorte des commentaires sur la signification ou la valeur de vérité de l'ensemble des réponses directes possibles à la question posée [Harrah 1961 : 41 ; 1963 : 50-51]. Entrent dans cette catégorie :

- la *pseudo-réponse directe* (ÅQVIST), qui nie une condition ou une présomption de la question posée ;

- la *pseudo-réponse suffisante* (ÅQVIST), qui implique une pseudo-réponse directe ;

- l'*éclaircissement* (HARRAH), qui nie l'une des composantes de la présupposition de la question, sans toutefois nier la présupposition elle-même ;

- l'*aveu d'ignorance* (KATZ), signifiant que la question n'est pas soluble pour la personne interrogée (cf. [Harrah 1961 : 41], [Belnap 1976 : 135]).

En outre, il est possible que ce qu'AJDUKIEWICZ appelle réponses *annulant la présupposition positive* ou la *présupposition négative* de la question se rattache aux réponses métalinguistiques. Mais en raison de la brièveté du texte auquel nous nous réfèrons ([Ajdukiewicz 1965]), il n'est guère aisé de déterminer (par exemple) si la première forme est une réponse métalinguistique, ou simplement une réponse correctrice.

2.4.3. Problèmes en suspens

Pour clore ce bref bilan des apports de la logique érotétique, il nous reste à analyser les lacunes et les insuffisances des théories que nous avons passées en revue, et à déterminer quelles peuvent être, en l'état actuel du développement de ces théories, les applications aux sciences sociales de la logique des questions. Les insuffisances de ces théories sont de deux sortes : insuffisances formelles, dues essentiellement à la relative jeunesse de la logique érotétique, et au fait que l'intérêt qu'elle suscite chez les logiciens reste pour l'essentiel limité à un petit nombre de spécialistes, dispersés dans plusieurs pays (Etats-Unis, Pologne, Suède) ; insuffisances pratiques, liées à la complexité des rapports qu'entretiennent logique et langage naturel. Les possibilités d'application aux sciences sociales découlent du fait que la logique érotétique constitue le cadre formel par excellence des recherches sur les modalités et le contenu de la communication ; dans la perspective qui est la nôtre, cette branche de la logique appliquée convient particulièrement aux situations socialement bien définies que sont les situations d'enquête ou de test psychologique.

a) Insuffisances formelles de la logique érotétique

Le panorama des principaux travaux publiés sur la logique des questions, que nous avons esquissé dans ce chapitre, donne une impression de très forte hétérogénéité. Hétérogénéité du vocabulaire tout d'abord, puisque par exemple une même forme syntaxique s'appelle selon les auteurs question à compléter (LÖW, MORITZ, AJDUKIEWICZ, etc.), question *what* (WHATELY, PRIOR, ÅQVIST), question *which* (BELNAP, HARRAH), question *numérique* (KUBIŃSKI), question *individuelle* (STAHL), etc. Nous avons vu toutefois que, dans la majeure partie des cas, les définitions fournies par les auteurs permettaient de déterminer avec assez de précision la forme dont il s'agissait. Hétérogénéité des systèmes de classification ensuite, entraînant dans certains cas une quasi-impossibilité de recoupement entre les catégories distinguées par deux auteurs. Hétérogénéité également des champs couverts par les théories : si la plupart des logiciens (sauf AJDUKIEWICZ) traite des alternatives à réponse *oui/non* et

des questions compréhensives, la plupart des autres formes de questions ne sont guère abordées que par quelques auteurs, voire souvent par un seul. Hétérogénéité enfin des fondements de la théorie elle-même : calcul des classes pour STAHL, logique modale pour CRESSWELL, calcul des prédicats pour HARRAH, extension aux "quasi-expressions" du calcul des prédicats pour BELNAP, calcul des prédicats muni d'opérateurs épistémiques et déontiques pour ÅQVIST, etc. Ces divergences fondamentales rendent malaisées les comparaisons entre les théories. Elles traduisent en outre l'immaturité de la logique érotétique, qui constitue un domaine de recherches encore relativement peu exploré, et dont l'unification semble par conséquent très problématique pour le moment.

L'existence de vastes zones inexplorées dans le champ de la logique des questions est illustrée par certains passages de deux des principaux ouvrages de logique érotétique. BELNAP, dans les § 2.1. et 2.2. de *The Logic of Questions and Answers* [Belnap 1976 : 78-87], et KUBIŃSKI, dans les § 8 à 14 de la première partie de *Wstęp do logicznej teorii pytań* [Kubiński 1970 : 25-38], décrivent une grande variété de formes de questions, sans pour autant fournir autre chose qu'une formalisation propre à chacune des formes distinguées. En l'absence d'une présentation formelle commune à l'ensemble, ces descriptions constituent presque autant de théories locales sans généralisation possible qu'il y a de variétés de questions dans les langues naturelles ; leur portée théorique est donc pratiquement négligeable en l'état actuel de la logique érotétique. En conséquence, on peut considérer que les seules formes syntaxiques élémentaires qui font l'objet d'une formalisation opératoire, et pour lesquelles il existe des théorèmes et des procédures de calcul, sont l'alternative portant sur une contradiction, la question extensive, et la question compréhensive. Il est donc exclu, pour le moment, de parvenir à traiter efficacement les questions de description, de définition, d'identité, de quantité, sur la valeur de vérité, voire, malgré les efforts de BROMBERGER et de BELNAP, les questions de causalité.

L'éventuelle unification des théories dépend de la possibilité de réduire les nombreuses variétés de questions décrites par les auteurs sinon à une forme unique, du moins à un petit nombre de types de base. Nous avons

vu que, dans sa révision de 1971, ÅQVIST avait considéré que la forme de base des types de questions qu'il avait distinguées était la forme conditionnelle, dont la forme inconditionnelle pouvait être dérivée [Åqvist 1971 : 53-54] ; mais cette révision affecte peu la classification elle-même, dont le nombre de types de base reste constant. Par contre, BELNAP suggère assez clairement une voie à suivre pour aboutir à une forme unique, du moins en ce qui concerne les questions *whether* et les questions *which*. D'une part, il montre que toutes les formes de questions *whether* sont réductibles à la question *whether* demandant un exemple unique [Belnap 1963 : 104]. D'autre part, il laisse entendre qu'une question *which* demandant une énumération exhaustive, dont les conditions catégorielles définissent des ensembles finis d'individus, est équivalente à une question *whether* à réponse *oui/non* universalisée [Belnap 1972 : 338-339] ; il met toutefois en garde contra la tentation de considérer les questions à réponse *oui/non* comme la forme fondamentale de question [Belnap 1976 : 90]. Et enfin, il suggère la possibilité de réduire certaines formes de questions *which* à des combinaisons de questions de HOBSON conditionnelles quantifiées [Belnap 1972 : 339-341], sans cependant parvenir tout à fait à la même possibilité de réduction pour les questions *whether* [Belnap 1976 : 106]. En première analyse donc, les tentatives de BELNAP et d'ÅQVIST conduisent à rechercher le type de base, dont tous les autres types de questions sont dérivés, dans une forme conditionnelle ; en outre, l'utilisation, pour définir ce type, de la question de HOBSON, semble une voie prometteuse en raison de la simplicité (et donc de l'universalité potentielle) de cette dernière forme. Toutefois, même si ces tentatives aboutissent, elles ne pourront dans l'immédiat conduire à l'unification que les parties suffisamment formalisées des théories existantes, c'est-à-dire correspondant aux trois formes fondamentales que nous avons définies au § 2.4.1.a, et aux formes complexes obtenues par combinaison, relativisation, et quantification de ces formes élémentaires.

b) Insuffisances pratiques de la logique érotétique

Lorsque l'on désire utiliser la logique érotétique pour analyser des énoncés interrogatifs, ou des réponses apportées à ces énoncés, le problème central est celui des rapports entre l'expression d'une interrogation en langage naturel, et la formalisation de la question qu'elle pose. Ce problème n'est pas spécifique de la logique érotétique. Nous l'avons rencontré par exemple pour la formalisation des relations d'implication et de présupposition dans les énoncés (§ 2.1.1.c et d). D'autre part, certains auteurs, tels que Willard Van Orman QUINE ou Oswald DUCROT, ont bien étudié les diverses traductions possibles, dans un système formel, des conjonctions *ou* ([Ducrot 1973 : 85-102] ; [Quine 1965 : 40-43]), *parce que* ([Ducrot 1973 : 103-116] ; [Quine 1965 : 53-56]), *mais*, *si*, etc., et des adjectifs indéfinis *aucun*, *tout*, *quelque*, etc. (cf. [Quine 1965 : 43-53, 126-140]). Il ressort de ces exemples que les difficultés de passage d'une langue naturelle à la logique formelle, déjà minutieusement analysées par ARISTOTE dans l'*Organon*, restent d'actualité ; et que ce qu'on appelle les ambiguïtés et les imprécisions du langage constitue un obstacle important à toute tentative de formalisation. Mais peut-être le problème est-il mal posé sous cette forme ; et peut-être pourrait-on également dire que la logique formelle classique est trop pauvre pour s'adapter aux particularités d'une langue naturelle, et permettre de prendre en compte toutes les caractéristiques de celle-ci, y compris la polysémie de la plupart des termes et l'amphibologie de certaines propositions. Quelle que soit la thèse adoptée sur les positions respectives du langage et de la logique, il est clair que leurs relations sont dans presque tous les cas extrêmement complexes, et qu'il est souvent nécessaire d'opter, de façon quasi-arbitraire, entre plusieurs interprétations formelles également possibles d'un énoncé en langage naturelle.

Il n'est guère d'auteurs, parmi ceux que nous avons présentés dans ce chapitre 2, qui n'aient souligné cette difficulté. Ainsi, Friedrich LÖW a montré que l'énoncé : "Qu'est-ce que l'or ?" était susceptible de deux interprétations distinctes [Löw 1928 : 397] ; Tadeusz KUBIŃSKI, que la consigne d'une question numérique à deux arguments : "Citez deux écrivains français et deux de leurs romans" pouvait prendre trois formes différentes, déterminant

trois formes différentes de réponse directe [Kubiński 1970 : 18-19] ; Nuel D. BELNAP, qu'une question à réponse *oui/non* régulière (alternative portant sur une contradiction) pouvait être formalisée de quatre manières, conduisant à deux définitions distinctes de l'ensemble des réponses possibles [Belnap 1963 : 100-101 ; 1976 : 59] ; etc. Mais c'est certainement Lennart ÅQVIST qui a accordé le plus d'attention à ce problème, en proposant par exemple trois formalisations pour la question d'examen : "Qui a composé la Symphonie Héroïque ?" [Åqvist 1969 : 119-120] ; ou encore six formalisations possibles pour la question : "Qui sont les membres de l'Orchestre Philharmonique de Vienne ?" [Åqvist 1965 : 85-94]. Les analyses d'ÅQVIST sont une excellente illustration, dans le domaine des questions, du hiatus existant entre une langue naturelle et un système formel pourtant remarquablement souple (si l'on considère la forme développée des questions chez ÅQVIST).

D'autre part, la logique érotétique ne prend généralement pas en compte (ou ne prend pas en compte d'une manière suffisamment précise ou suffisamment complète) certains traits sémantiques des énoncés, traits qui présentent cependant une grande importance pour le praticien. Par exemple, s'il est relativement facile d'extraire de la représentation formelle d'une question sa présupposition et son contenu propositionnel, il n'est pas possible dans tous les cas de dégager ce que l'énoncé interrogatif non seulement pose et présuppose, mais aussi suppose, suggère, et implique. Par exemple encore, la logique érotétique ne fournit pas d'algorithme permettant de décider si une proposition constitue ou non une réponse à une question donnée : selon quelle procédure formelle considère-t-on que l'affirmation : "Il y a eu une sécheresse exceptionnelle cet été au Soudan" constitue une réponse indirecte à la question : "Y-a-t-il cette année une réserve de denrées alimentaires suffisante pour nourrir la population soudanaise ?", alors que l'affirmation : "Cicéron était orateur" est absolument hors du sujet de la question ? [Hiz 1962 : 254]. Ces problèmes, il est vrai, relèvent peut-être autant de l'analyse linguistique que de la représentation logique des énoncés. Par contre, il est dommage que la logique érotétique ne soit pas outillée pour prendre en compte certaines relations sémantiques telles que les relations temporelles entre propositions, ou encore les relations d'antonymie, d'hyponymie, etc.

(cf.p.ex. [Lyons 1978 : 218-244] ; nous reviendrons sur certaines de ces relations au § 3.1.3.). Cette lacune explique la difficulté qu'éprouvent les logiciens pour traiter par exemple des questions contenant leur propre réponse, ou de certaines causes d'invalidité des questions telles que les erreurs catégorielles.

Enfin, pour conclure sur les rapports qu'entretiennent logique érotétique et langues naturelles, signalons une dernière difficulté, d'ailleurs à notre avis mineure : à partir d'une (quasi-) expression logique, exprimant par exemple une interrogation, une réponse, ou une présupposition, il n'est pas toujours facile d'engendrer l'énoncé en langue naturelle qui lui correspond. Il est vrai que ce problème n'est pas un obstacle important aux applications de la logique érotétique aux tests et aux questionnaires.

c) Perspectives d'application.

Ainsi que nous l'avons signalé au § 2.1.3.b, les domaines d'application de la logique érotétique sont extrêmement variés. En nous limitant au problème qui est le nôtre, c'est-à-dire l'application de la logique érotétique à l'analyse des tests et des questionnaires, nous considérons que l'apport actuel des logiciens concerne plus particulièrement deux aspects de la technique du questionnaire : un aspect théorique, avec la définition précise de notions qui n'avaient pas jusqu'ici fait l'objet d'une étude poussée de la part des psychologues ou des sociologues ; et un aspect pratique, débouchant sur l'analyse de certaines formes d'erreurs dans la rédaction d'un questionnaire, et la formulation de règles permettant de les éviter.

Sur le plan théorique, l'apport de la logique érotétique est avant tout classificatoire. Nous ne rappellerons pas les principales typologies des questions et des réponses décrites dans ce deuxième chapitre ; d'ailleurs, leur

intérêt pour l'analyse des questionnaires est évident, et leur application immédiate. Par contre, il nous paraît important de signaler quelques points qui, s'ils sont généralement jugés secondaires par les logiciens, peuvent se révéler primordiaux pour le spécialiste du questionnaire. En premier lieu, la logique des questions donne un statut précis à la *présupposition* d'une question, et définit ses relations avec l'ensemble des réponses directes possibles et le contenu propositionnel de la question ; ceci conduit à la définition de la procédure d'*assurance* d'une question, et de la notion de *réponse corrective*. Dans la même perspective, des énoncés tels que : "Je ne sais pas", "Je n'ai pas d'opinion sur la question", "Cela dépend", etc. reçoivent également un statut logique bien défini (énoncé métalinguistique remettant en cause la présupposition de la question).

Ensuite, l'analyse sémantique des conditions de validité des questions débouche sur une classification commode des formes de questions utilisées dans les tests et les questionnaires :

- questions *de fait*, pour lesquelles il existe au moins une réponse directe vraie absolument (correspondant à un état de fait), mais ignorée de l'auteur du questionnaire (la question est discutable relativement à H_q , et en principe soluble relativement à H_r).

- questions *de connaissance*, pour lesquelles il existe au moins une réponse directe vraie absolument, connue de l'auteur du questionnaire (la question est rhétorique relativement à H_q), l'objet de l'interrogation étant de déterminer si la personne interrogée connaît elle aussi la réponse vraie (si la question est soluble ou insoluble relativement à H_r). Les questions de niveau scolaire (tests) et les questions de notoriété (enquêtes) entrent dans cette catégorie.

- questions *d'opinion*, pour lesquelles il n'existe ni réponse vraie absolument, ni réponse fausse absolument, mais seulement au moins une réponse vraie relativement aux croyances et aux connaissances de la personne interrogée (la question est en principe soluble relativement à H_r).

Enfin, la logique érotétique fournit une description opératoire des relations possibles entre questions, entre réponses à une question, et entre modalités de réponses présentées dans le sujet d'une question extensive. Les principales relations possibles entre questions présentant un intérêt pour notre recherche sont en premier lieu les relations d'*inclusion*, d'*équivalence*, de *pertinence*, et d'*annulation* (cf. § 2.2.5.e), qui portent sur les rapports sémantiques entre réponses directes à des questions différentes ; nous retrouverons certaines de ces relations à propos de la description de la structure formelle d'un questionnaire, au § 3.2.3. En second lieu, David HARRAH a indiqué, avec les notions d'*ensemble de questions* et de *conjonction des questions* d'un ensemble [Harrah 1961 : 46 D21], une voie à suivre pour constituer une théorie logique du questionnaire ; à notre connaissance toutefois, cette direction de recherches n'a guère été explorée depuis. Les relations les plus intéressantes entre réponses possibles à une question, et également entre modalités de réponse d'une question extensive, nous paraissent être les relations d'*équivalence*, d'*exclusion réciproque*, et d'*implication*, telles que les définit le calcul des propositions. Signalons, pour terminer sur ce point, que ce que HARRAH appelle *combinaison de réponses directes* à un ensemble de questions [Harrah 1961 : 46 D22] est l'équivalent formel d'un *protocole* d'entretien.

Sur le plan pratique, les lacunes et les insuffisances (probablement provisoires) de la logique érotétique risquent de restreindre fortement son champ d'application. Par exemple, il est peu réaliste d'espérer qu'en l'état actuel de son développement, la logique des questions puisse permettre une analyse formelle d'un entretien d'enquête moyennement structuré. En revanche, l'analyse des tests et des questionnaires présente ceci de particulier qu'elle porte en général soit (dans le cas de questionnaires d'enquête ou de tests papier-crayon) sur des questions fermées ou précodées, dont l'ensemble des réponses possibles est fini, bien déterminé, et arrêté *ne varietur* ; soit (dans le cas de certaines épreuves psychologiques de niveau intellectuel

ou de personnalité) sur des questions ouvertes, pour lesquelles les règles de codification des réponses, et la formation technique des codeurs, atteignent un niveau de fiabilité très élevé. En outre, la perspective de l'analyse secondaire des informations recueillies, dans laquelle nous nous sommes placé, ne nous permet aucune action corrective sur les procédures de recueil de l'information. Tout au plus pouvons-nous, à partir de l'examen des questions posées, porter un diagnostic sur la valeur des informations recueillies dans les réponses. Dans cette perspective, et quelles que soient nos réticences théoriques à l'adoption d'un tel point de vue (risque de verbalisme), rien ne nous interdit de considérer les questions de test ou de questionnaire comme appartenant à l'une des trois catégories décrites au § 2.4.1.a. D'ailleurs, bien souvent, les auteurs de test ou de questionnaire tournent les difficultés propres à certaines formes de questions telles que, par exemple, les questions de causalité, en donnant à celles-ci l'apparence de questions extensives (ex. : "Parmi les raisons suivantes, quelles sont celles pour lesquelles vous vous êtes arrêté à cet endroit plutôt qu'à un autre ? [...]") *SOFRES 340 : 9* ; "Le fait que vous ayez poursuivi vos études secondaires jusqu'au bac (ou son équivalent) correspond-il (cochez les trois principales raisons en les numérotant par ordre d'importance) : [...] ?" *OCDE - IUT : 23*). Un tel procédé est certes artificiel, mais le langage utilisé dans la rédaction des tests et des questionnaires n'est pas, malgré l'apparence, le langage de la vie quotidienne, même si l'on s'efforce de le rendre familier.

Tadeusz PAWŁOWSKI est, à notre connaissance, le premier chercheur en sciences sociales à avoir tenté d'appliquer la logique des questions (essentiellement [Belnap 1963], et accessoirement [Kubiński 1966]) au diagnostic des erreurs commises dans la rédaction d'un questionnaire. Cet auteur analyse principalement quatre types d'erreurs relevées dans divers questionnaires d'enquête utilisés en Pologne [Pawłowski 1969 : 104-108], selon que l'erreur porte sur :

- le *sujet* de la question, c'est-à-dire la présentation des modalités de réponse possibles. Par exemple, l'énoncé : "Etes-vous favorable ou défavorable à l'octroi d'allocations et de réductions aux personnes sans emploi ?"

peut être compris (en posant : A = "favorable aux allocations", et R = favorable aux réductions") soit comme présentant l'alternative : $\sigma = (A \& R, \overline{A \& R})$, soit comme proposant : $\sigma' = (A \& R, \overline{A \& R})$, soit encore comme la conjonction de deux interrogatives proposant respectivement : $\sigma_1 = (A, \overline{A})$, et $\sigma_2 = (R, \overline{R})$.

- la *dimension de sélection*, c'est-à-dire le nombre de modalités possibles que la réponse doit présenter. Ainsi, la formulation de la seconde des deux questions suivantes ne permet pas de déterminer s'il s'agit d'une question demandant une réponse vraie ou un exemple unique, des exemples multiples, ou une énumération exhaustive : "Avez-vous eu l'occasion de discuter de l'éducation des enfants ? *Si oui* : en avez-vous discuté avec des amis, au cours d'une réunion publique, à l'école, avec un prêtre, avec d'autres personnes ?"

- l'*exigence d'exhaustivité*. Par exemple, l'énoncé : "A votre avis, l'Union de la Jeunesse Polonaise devrait-elle faire de ses membres : (a) de bons citoyens (b) des défenseurs du socialisme (c) des communistes (d) je ne sais pas", a été compris par les personnes interrogées comme demandant une énumération exhaustive, alors que, semble-t-il, l'auteur du questionnaire n'en attendait qu'une seule réponse.

- la *présupposition* qu'il existe au moins une réponse directe vraie à la question. L'énoncé : "A votre avis, les membres de l'Union de la Jeunesse Polonaise sont-ils supérieurs aux non membres : (a) dans leur formation politique (b) dans leur participation aux activités sociales [...] (i) dans leur souci de la propriété collective ?" n'autorise pas de réponse du type : "Ils ne sont supérieurs aux non membres d'aucun point de vue", qui nie la présupposition de la question.

Les analyses de PAWŁOWSKI montrent l'utilité de la logique érotétique pour l'élaboration de tests ou de questionnaires. Il serait possible en effet d'édicter quelques règles permettant d'éviter les erreurs de logique liées aux relations entre questions, ou entre une question et ses réponses

possibles. Bien que cette préoccupation soit extérieure au champ de recherches que nous nous sommes fixé, nous illustrerons cette possibilité en énonçant les règles à appliquer afin qu'une question soit opératoire. Nous dirons qu'une question est *opératoire* si et seulement si, chaque fois que cette question est effectivement posée, il existe, dans la précodification prévue pour cette question, au moins un code correspondant à l'énoncé fourni par le répondant en réponse à cette question, quel que soit cet énoncé. Le problème n'a évidemment d'intérêt que si la question est fermée ou précodée, puisqu'en cas de question ouverte, la tâche de l'enquêteur se borne à écrire sur le questionnaire l'énoncé fourni par le répondant. Nous supposerons pour simplifier que la question est aisément compréhensible par la personne interrogée, et ne présente ni ambiguïté ni contradiction ; et que l'enquêteur ou le psychologue est entraîné à poser correctement les questions, à interpréter et coder les réponses, et à relancer ou demander des précisions lorsque la réponse fournie lui paraît ambiguë, contradictoire, incomplète, ou hors du sujet (nous écartons ainsi les questionnaires auto-administrés, dans lesquels les questions ne peuvent être qu'ouvertes ou fermées ; cf. § 3.1.1.a). Tout d'abord, deux cas sont à considérer, selon que la disjonction des réponses directes prévues dans la codification est ou non une tautologie : dans le second cas, la précodification devra comprendre, en plus des réponses prévues, la rubrique "*autres*", afin que toutes les réponses directes effectivement fournies puissent être enregistrées. Ensuite, la précodification devra prévoir (au moins) une rubrique pour les réponses métalinguistiques du type : "Je ne sais pas", "Je n'ai pas d'opinion", "Cela dépend", voire : "Je refuse de répondre". Enfin, le problème le plus important, et souvent le plus mal résolu, concerne les réponses correctives, qui nient la présupposition de la question. Bien que les logiciens considèrent qu'une réponse corrective n'est pas une réponse à la question (cf. p.ex. [Maas 1972]), tout énoncé qui ou bien nie la présupposition, ou bien implique la négation de la présupposition, doit être enregistré, et distingué des autres rubriques de la précodification. Deux solutions sont envisageables. La première consiste à *assurer* la question en posant une question-filtre préalable portant sur l'acceptation ou le rejet de la présupposition. La seconde, à inclure dans la précodification une rubrique : "*négation de la présupposition*".

Toutefois, bien que ces deux solutions soient formellement équivalentes, nous verrons au prochain chapitre (§ 3.1.3.c) que leurs effets sur les personnes interrogées ne sont pas identiques, et qu'elles conduiraient très probablement à des taux de réponses différents auprès d'échantillons comparables.

3. QUESTIONNAIRE ET ANALYSE DES DONNÉES

Ce chapitre peut être considéré comme une introduction aux problèmes de l'analyse des informations recueillies à l'aide du questionnaire (épreuve psychologique ou questionnaire d'enquête). Dans notre introduction, nous avons indiqué que ces informations étaient en réalité destinées à l'élaboration des réponses à un autre type de questionnaire, qu'en l'absence de termes consacrés par l'usage, nous avons appelé *questionnaire théorique* (opposé au *questionnaire empirique* administré aux personnes interrogées). L'élaboration des réponses à ce second questionnaire passe par une phase d'analyse (au sens étymologique : ἀνάλυσις = dissolution, décomposition en éléments simples) des informations recueillies, préalablement à leur restructuration (συνθεσις) conformément aux objectifs du questionnaire théorique.

Dans le contexte d'une analyse secondaire de données recueillies par un autre chercheur, il y a évidemment multiplicité potentielle des questionnaires théoriques. C'est pourquoi l'analyse des informations obtenues par l'intermédiaire du questionnaire empirique doit être poussée aussi loin qu'il est possible de le faire, afin d'obtenir des informations élémentaires susceptibles d'être recombinaées entre elles de manières variées. Idéalement, nous pourrions reprendre à notre compte les aspirations de DÉMOCRITTE et de LUCRÈCE, en les transposant du domaine de la matière à celui de l'information, et poser qu'il existe des éléments d'information indivisibles (ἄτροπα), immuables, et qualitativement identiques, dont toute information quelle qu'elle soit est composée. Nous verrons d'ailleurs que Donald Mac KAY défend une thèse analogue (quoique modernisée), en faisant l'hypothèse qu'il existe des *quanta*

d'information ; c'est à l'identification de ces *quanta* dans le contenu des réponses aux questionnaires empiriques que nous nous attacherons dans ce chapitre.

Si ces atomes ou ces quanta d'information existent, c'est bien entendu après l'administration du questionnaire empirique aux personnes interrogées qu'il sera possible de les dénombrer. Nous constaterons d'ailleurs que les fondements de la plupart des techniques élémentaires de statistique descriptive résident dans le dénombrement de ces quanta au sein des informations rassemblées après questionnement. Toutefois, certaines caractéristiques formelles des informations recueillies à l'aide d'un questionnaire sont déterminées a priori par la manière dont les questions sont posées ; en outre, la codification prévue pour les réponses obtenues augmente encore cette détermination. C'est pourquoi l'identification des "particules élémentaires" d'information nous impose l'analyse de la préformation des informations par le questionnaire, à travers chacune des questions qui le composent.

Dans la première partie de ce chapitre, nous examinons les relations que le sociologue ou le psychologue établit entre les *questions* et les *variables*, c'est-à-dire entre les observations ou les faits traduits par les énoncés recueillis, et les concepts servant à relier, structurer, expliquer ces observations. Ces relations dépendent à la fois des intentions du chercheur, de sa problématique, de son système de pensée, et de la nature des observations. Dans le contexte d'une analyse secondaire de données d'enquête, par exemple, on ne peut évidemment pas postuler l'unicité de la problématique, puisqu'il n'y a pas unicité de l'utilisateur des données (encore que les intentions de l'auteur du questionnaire enferment les utilisateurs ultérieurs des informations recueillies dans un certain champ conceptuel dont il n'est pas facile de s'extraire). Quoi qu'il en soit, il y a pluralité potentielle des questionnaires théoriques pour un même ensemble de données. En revanche, les traitements que le chercheur fera subir aux informations recueillies sont fortement déterminés par la nature des faits auxquels renvoient les réponses, et par la culture (au sens ethnologique) à laquelle correspondent les structures du langage dans lequel les questions sont exprimées. Il est donc possible, en partant de l'analyse des caractéristiques techniques des questions, de définir les caractéristiques techniques des réponses obtenues. Ajoutons

toutefois que le traitement des réponses doit être effectué avec précautions, et que leur analyse sémantique est loin d'être simple : nous verrons en effet, à la fin de cette première partie, que, selon l'adage bien connu, les faits traduits par les énoncés sont également trahis par ceux-ci, sans que les biais ainsi introduits soient toujours maîtrisables.

La seconde partie du chapitre 3 vise, au niveau du questionnaire dans son ensemble, à un objectif analogue à celui de la première partie au niveau de la question considérée isolément. En effet, nous y décrivons le passage du questionnaire empirique à la structure de ce que l'on appelle (par abus de langage) la *matrice des données*, c'est-à-dire la présentation normalisée de l'ensemble des réponses obtenues auprès de l'ensemble des personnes interrogées. La structure de la matrice des données dépend étroitement de la structure du questionnaire qui a permis de les recueillir. D'autre part, le développement de l'archivage des données et de l'analyse secondaire a pour effet une forte tendance à la normalisation de la présentation des fichiers. Dans cette perspective, nous serons conduit à définir une procédure de réduction d'un questionnaire quelconque à une forme canonique, aboutissant à une structure de fichier qui soit à la fois la plus simple à traiter, et la plus fidèle possible aux données de base.

Enfin, la troisième partie de ce chapitre aborde le problème général de l'*analyse des données*, c'est-à-dire de l'élaboration des réponses à un questionnaire théorique, à partir du dénombrement des réponses recueillies à l'aide du questionnaire empirique. Une telle analyse se fonde sur le traitement numérique de l'ensemble des réponses empiriques, et a, par conséquent, pour point de départ la matrice des données. Mais elle fait également intervenir un certain nombre d'hypothèses théoriques relatives aux propriétés formelles des variables construites à partir des réponses aux questions empiriques. Ces hypothèses influent sur le mode de traitement numérique choisi, et sur la nature des résultats obtenus. Il est donc important de bien comprendre l'articulation de ces hypothèses avec la définition des unités d'information élémentaires, des *quanta* d'information. L'objet de cette troisième partie sera de montrer d'une part l'usage que la statistique descriptive fait de ces quanta d'information, et d'autre part l'incidence sur les

résultats des traitements statistiques, de variations dans les hypothèses sur les variables. Notre ambition est de contribuer ainsi à clarifier les rapports entre les conceptions *a priori* du chercheur et les résultats que les techniques statistiques peuvent lui apporter, et d'esquisser un cadre théorique permettant de situer les diverses techniques de statistique descriptive les unes par rapport aux autres.

3.1. DES QUESTIONS AUX VARIABLES

La fonction des questions est de permettre au chercheur de recueillir sur le répondant, ou à travers celui-ci, certaines informations qui lui sont nécessaires. L'auteur d'un questionnaire d'enquête cherche en général à décrire une population d'individus à partir des réponses fournies par un échantillon de ceux-ci ; l'utilisateur d'une épreuve psychologique standardisée (test, inventaire de personnalité) vise à comparer un individu à l'ensemble de la population dont il est issu. Les questions posées, et les réponses qu'elles suscitent, sont exprimées dans le langage de la personne interrogée. La nature des informations recueillies dépend par conséquent assez étroitement de la formulation de la question et de sa consigne (nous verrons en outre qu'elle dépend également de certains paramètres de la situation d'entretien). Par contre, l'analyse et l'interprétation de ces informations se font nécessairement dans le langage du chercheur (psychologue ou sociologue), qui les effectue en fonction de son savoir et de ses hypothèses. C'est cette traduction d'un langage (et d'un système de pensée) dans un autre que nous appelons le passage des questions aux variables.

Afin d'analyser ce processus, nous présentons dans ce sous-chapitre la description formelle, du point de vue du psychologue ou du sociologue (par opposition au point de vue du répondant), des principales variétés de questions et de consignes effectivement utilisés dans les tests et les questionnaires. Nous proposons ensuite un cadre sémantique formel dans lequel peut se faire le passage de la question à la variable. Nous rappelons enfin les principales difficultés d'interprétation des réponses aux questions, soulignant ainsi les limites du cadre proposé.

3.1.1. Classification empirique des formes de questions

Il est évidemment possible de classer les questions selon leur *contenu*, c'est-à-dire, en dernière analyse, selon la nature de l'information apportée par les réponses à ces questions. Par exemple, dans les questionnaires d'enquête, on distinguera les questions factuelles des questions d'opinion ou d'attitude ; parmi les questions factuelles, celles qui concernent les renseignements signalétiques, les biens d'équipement possédés, les conditions de travail, la consommation courante, etc. ; parmi les questions non factuelles, celles qui portent sur les croyances religieuses, les opinions politiques, les sentiments à l'égard de telle communauté ou telle personnalité, les attitudes envers un produit donné, etc. Une classification thématique conçue sur ce modèle est en particulier très utile pour les inventaires de libellés de questions préalables à leur standardisation (cf. p.ex. [Patchen 1965], [Debordeaux 1968], [Robinson 1973]). De même, pour les questions utilisées dans les tests psychologiques, on peut adopter par exemple la dichotomie : efficience/personnalité, en rangeant dans la première catégorie les questions d'aptitudes (verbales, numériques, de raisonnement, etc.) et de connaissances (mathématiques, grammaticales, etc.) et dans la seconde les questions d'introversi-on-extraversi-on, d'anxiété, d'autoritarisme, de suggestibilité, etc.

Pour notre propos, qui est lié à l'analyse secondaire de données préalablement recueillies (et, le plus souvent, codifiées), cette forme de classification présente deux inconvénients. En premier lieu, il est clair que le contenu d'une question ne peut pas, en règle générale, être considéré comme univoque. Dans un questionnaire de personnalité, une même question peut servir d'indicateur à des dimensions (ou sous-dimensions) différentes ; c'est évidemment le cas dans les questionnaires de type factoriel, comme le *CAT 16 PF*. On peut également en trouver un exemple dans l' "échelle de fascisme" d'ADORNO, où certains items figurent dans plusieurs "clusters" ; ainsi, l'item 37 ("Si les gens parlaient moins et travaillaient plus, les choses iraient mieux pour tout le monde") participe aux trois dimensions : "conventionalisme", "agression", et "antisubjectivisme" (cf. [Adorno 1950 : 255-256]). Dans un questionnaire d'enquête, classer les questions selon leur contenu

"manifeste" revient à préjuger l'usage que le chercheur fera de l'information ainsi recueillie ; or, rien n'interdit par exemple de considérer les réponses à des questions relatives aux derniers achats effectués ou aux activités culturelles, comme des indicateurs de snobisme ou de conformisme. De même, dans un test de niveau intellectuel, rien n'interdit non plus de tirer des résultats enregistrés, et du comportement global du sujet au cours de l'épreuve, des indications sur la personnalité de celui-ci (cf. p.ex. [Anderson 1951]). On peut, il est vrai, reconnaître une certaine utilité à une forme de classification des questions d'enquête liée à la confiance que l'on peut accorder aux informations qu'elles ont permis de recueillir (à la "validité" de leurs réponses) ; il s'agit de la distinction entre :

① faits actuels, ② faits passés, ③ intentions effectives, ④ intentions hypothétiques ("futur hypothétique"), et ⑤ opinions. En règle générale en effet, on peut considérer qu'en moyenne, les réponses aux questions concernant des faits actuels sont plus "fiabiles" que les réponses aux questions concernant des faits passés, lesquelles sont elles-mêmes plus "fiabiles" que les réponses aux questions portant sur des intentions effectives, etc. Mais cette règle toutefois souffre de nombreuses exceptions, et ne peut guère être considérée que comme un expédient commode, manquant de fondements théoriques. En second lieu, il est loin d'être évident que, si l'on adoptait une classification des questions fondées sur leur contenu, on puisse parvenir à définir des critères d'identification des types de questions, et de classement des énoncés interrogatifs selon ces types, qui soient opératoires.

Nous allons voir par contre qu'il est relativement facile de classer les questions selon leur *forme*. Les auteurs qui traitent de la construction des tests ou de la rédaction des questionnaires proposent ainsi de nombreuses classifications empiriques des diverses formes de questions. Nous commençons tout d'abord par rappeler, dans le premier paragraphe ci-dessous, les deux classifications les plus répandues chez les professionnels de l'enquête psychosociologique ; nous développons ensuite une classification plus détaillée pouvant convenir à la fois aux questions d'enquête et aux items d'épreuves psychologiques.

a) Rappel des classifications usuelles

La plupart des auteurs qui traitent du questionnaire d'enquête distinguent trois catégories de questions, selon la manière dont les réponses sont enregistrées et codifiées :

- questions *ouvertes* (analogues aux questions *which* de BELNAP, aux questions *what* d'ÅQVIST, aux questions *numériques* de KUBIŃSKI, etc.), pour lesquelles l'enquêteur doit en principe noter *in extenso* tout ce que dit la personne interrogée ; la codification de la réponse étant effectuée au moment de la relecture du questionnaire, par une équipe de codeurs spécialisés.

- questions *fermées* (analogues aux questions *whether* de BELNAP et d'ÅQVIST) dans l'énoncé desquelles est présenté un ensemble des réponses possibles au sein duquel la personne interrogée doit choisir sa propre réponse ; en un sens, la codification est effectuée par l'interviewé lui-même.

- questions *précodées*, qui se présentent pour l'interviewé comme des questions ouvertes, mais pour lesquelles l'enquêteur, au lieu de noter textuellement la réponse, doit rechercher parmi une liste de réponses possibles celle qui correspond le mieux à la réponse fournie ; la codification est par conséquent effectuée "à chaud", sur le terrain, par l'enquêteur.

Cette classification doit être conservée à l'esprit lors de l'interprétation des informations recueillies ; elle peut sans inconvénient être appliquée aux questions figurant dans les épreuves psychologiques de notre corpus. Toutefois, s'il s'agit de questionnaires postaux ou administrés collectivement sur table, ou de tests "papier-crayon", la distinction entre question fermée et question précodée n'a évidemment pas de sens, puisque la codification est alors toujours effectuée par la personne interrogée elle-même.

D'autre part, les impératifs du dépouillement automatique des enquêtes ou de certains tests collectifs ont conduit à distinguer, selon les propriétés formelles des réponses après codification, quatre types de questions :

- les questions à *réponses multiples* (ou à codes multiples), telles qu'à une question unique puissent correspondre plusieurs réponses pour un même interviewé (ex. : "Quelles sont les marques françaises auxquelles vous êtes le plus favorable ?" *SOFRES 967 : 16a*).

- les questions de *quantité*, questions à réponse unique pour lesquelles les informations recueillies sont des nombres, sur lesquels il est licite d'effectuer les opérations arithmétiques usuelles (ex. : "A votre avis, combien gagnent, en moyenne par mois, les personnes exerçant les professions ci-dessous ? - Noter en centaines de francs nouveaux" *GEMAS 76 : 36*).

- les questions à *réponses possibles ordonnées*, questions à réponse unique dont l'ensemble des réponses possibles est muni d'une relation d'ordre strict (ex. : "Estimez-vous que cette deuxième consultation aura plus, autant, ou moins d'importance que le référendum ?" *IFOP 58-9 : 23*).

- les questions à *réponses possibles nominales*, questions à réponse unique, pour lesquelles l'ensemble des réponses possibles n'est apparemment pas structuré *a priori* (ex. : "Dans votre travail, êtes-vous considéré comme un coordinateur, un spécialiste, un médiateur ?" *BNIST : 44*).

L'ensemble des réponses possibles aux questions à réponses multiples n'est jamais structuré *a priori* (au sens que nous donnons provisoirement à ce mot).

Cette dernière classification est consacrée par l'usage. Elle rend de grands services pour la saisie des données et leur traitement. Mais il est manifeste qu'elle n'est que le reflet des pratiques les plus usuelles actuellement, et qu'elle ne constitue pas un relevé de tous les cas envisageables (nous reviendrons sur ce point à propos de l'analyse des données). Les divers auteurs qui se sont efforcés de classer les questions selon leur forme ont tenu compte principalement de la manière de présenter l'univers des réponses possibles, ce qui les a conduits à distinguer trois grandes catégories de questions : questions à réponse libre, questions à choix multiples, et alternative. Nous proposons ci-après une synthèse de quelques classifications empiriques selon la présentation des modalités de réponse, relevées chez des auteurs traitant de la rédaction des questions (PAYNE, REMMERS et GAGE,

WESMAN). L'intérêt de ces classifications, conçues dans l'optique du praticien (c'est-à-dire de l'auteur du questionnaire ou du test), est dû à la primauté qu'elles accordent aux considérations formelles sur les problèmes de contenu, et à leur fidélité aux formulations effectivement utilisées. Leurs inconvénients tiennent à leur caractère empirique : désaccords entre classifications, et existence de formes difficilement classables. C'est pourquoi, tout en conservant les grandes lignes des classifications auxquelles nous nous réfèrons, nous avons dû soit distinguer des sous-catégories nouvelles pour augmenter la précision du classement, soit au contraire regrouper des catégories dont la distinction ne paraissait pas essentielle pour notre propos, soit encore ouvrir des rubriques hors classification pour des formes hybrides, ou n'entrant pas, en toute rigueur, dans les catégories définies préalablement. Dans cette optique, les trois formes principales de questions sont les questions à réponse libre, les alternatives, et les questions à choix multiples. Cette trichotomie ne doit pas être confondue avec la classification déjà mentionnée de Henry S. LEONARD (cf. 2.1.3.b), en questions indirectes, directes, et semi-directes [Leonard 1967 : 33-34] (ni par conséquent avec la classification syntaxique inspirée de la logique érotétique que nous avons développée au § 2.4.1.a), en raison de la place accordée par les psychologues et les sociologues à l'alternative portant sur une contradiction (assimilée à l'ensemble des alternatives).

b) Questions à réponse libre

En principe, dans une question à réponse libre, le répondant a toute latitude pour exprimer la réponse qui lui convient sous la forme qui lui convient (les questions ouvertes entrent donc dans cette catégorie). En pratique, l'énoncé de la question exerce sur la personne interrogée une certaine contrainte, qui détermine à la fois la forme de la réponse et sa longueur. A partir des classifications de REMMERS et GAGE, et de A.G. WESMAN (cf. [Remmers 1955 : 79-83] et [Wesman 1971 : 80-91]), nous distinguerons quatre formes principales de questions à réponse libre : l'interrogation, la phrase à compléter, les items d'appariement, et les items à substitution.

L'*interrogation* est exprimée sous la forme d'une question directe, à laquelle la personne interrogée répond, selon le cas, par un nombre, un mot, un groupe de mots (syntagme), une phrase, ou un ensemble de phrases (discours). Dans les formes administrées par un enquêteur ou un psychologue, il peut s'agir de questions précodées.

Les interrogations demandant un *nombre* pour réponse sont naturellement très fréquentes dans les épreuves d'arithmétique (cf. p.ex. : *SUBES*, *CC8*, *NB* : n° 5), et de simple comptage (cf. *BS Z* : 16, 20). Il arrive que l'on en relève dans des tests de connaissance, comme : "Combien faut-il de calories pour transformer en vapeur à 100° huit grammes de glace à 0° ?" [Wesman 1971 : 89], ou comme : "En quelle année l'Amérique a-t-elle été découverte ?" [Remmers 1955 : 81]. On en rencontre également dans les épreuves portant sur des lois de séries (ex. : *NB* : 10, *EAI 5/1* : *SN*). Mais on peut aussi en trouver de nombreux exemples dans les questionnaires d'enquête :

- "Jusqu'à quelle heure dans la soirée [avez-vous suivi le comte-
rendu des résultats électoraux] ?" (IFOP 59-12 : 29b)
- "Quelle est la superficie de son exploitation ?" (IFOP 58-9 : 31d)
- "Nombre de personnes de l'établissement dans lequel vous travaillez"
(CEREQ-SEU : 33)
- "Actuellement, combien de personnes dans votre foyer ont le permis
de conduire ?" (SOFRES 769 : 6)
- "Pendant combien de mois avez-vous gardé cette ... (voiture de
Q. 27) ?" (SOFRES 769 : 30)
- "Le plus souvent, combien de personnes au total transportez-vous à
l'arrière ?" (SOFRES 769 : 59)
- "Quel supplément de prix seriez-vous disposé à payer pour avoir une
boite de vitesses automatique ?" (SOFRES 967 : 34bis)

Les interrogations auxquelles on répond par un simple *mot* sont illustrées, chez les auteurs de classification des questions déjà mentionnés, par des questions de connaissance : "Qui fut le premier Président des Etats-Unis ?" [Remmers 1955 : 79] ; "Qui a inventé la machine à égrener le coton ?" [Wesman 1971 : 89]. Dans les tests et questionnaires de notre corpus, on trouve par exemple :

- "*Comment t'appelles-tu ?*" (BS Z : 7)
- "*Quelle est cette couleur ?*" (BS Z : 17)
- "*De quel instrument jouez-vous ou avez-vous joué ?*" (ORTF 65 : 23)
- "[*Qui est le*] *principal utilisateur de la voiture ?*" (SOFRES 769 : 19)

Mais la frontière entre les questions auxquelles on répond par un mot unique et celles auxquelles on répond par un *syntagme* est floue, d'autant qu'à une question donnée, il est possible de répondre indifféremment par un mot ou un syntagme (voire une courte phrase). Cependant, on rangera plutôt dans la seconde catégorie les énoncés suivants :

- "*Es-tu un petit garçon ou une petite fille ?*" (BS Z : 5)
- "*Qui s'occupera [...] du choix et de l'achat de cette voiture ?*" (SOFRES 769 : 9)
- "*Quel est à votre avis le parti ou la tendance dont les objectifs correspondent le mieux à ceux du général de Gaulle ?*" (IFOP 58-12 : 9)
- "*Quel est le groupe social dont le sort économique vous semble le plus lié au vôtre ?*" (GEMAS 76 : 16)

Remarquons que la frontière entre cette dernière catégorie, et celle des interrogations auxquelles on répond par un nombre, est également floue ;

exemples :

- "*Depuis combien de temps habitez-vous à ... (nom de la localité où se fait l'interview) ?*" (SOFRES 571 : 1)
- "*Combien de temps vous êtes-vous connus et fréquentés avant le mariage ?*" (INED 081 : fiançailles 2)

D'autre part, il arrive que des questions auxquelles on répond par un mot (ou, plus souvent, par un syntagme) soient à réponses multiples :

- "*Tu vas me donner tous les mots que tu connais qui riment avec le mot obéissance*" (BS Z : 45)
- "*Quelles sont toutes les marques de voitures que l'on a déjà eues dans votre foyer ?*" (SOFRES 769 : 36)

REMMERS et GAGE illustrent ce cas particulier par : "Quels sont les deux principaux gaz que l'on trouve dans l'air ?" [Remmers 1955 : 79].

Les interrogations auxquelles on répond par une *phrase complète* sont de la forme : "Expliquez en une courte phrase quel est le lien entre chacun des personnages énumérés ci-dessous, et Abraham LINCOLN [...]" [Remmers 1955 : 79]. Peuvent être rangés dans cette catégorie les énoncés :

- "*Qu'est-ce qu'une chaise ?*" (BS Z : 13, 34)
- "*Quand on a manqué le train, que faut-il faire ?*" (BS Z : 28)
- "*En quoi crayon et plume se ressemblent-ils et en quoi sont-ils différents ?*" (TMI : 1)
- "*[...] Maintenant, dites-moi ce que signifie le proverbe suivant : l'habit ne fait pas le moine*" (TMI : 29)

Ici également, la frontière n'est pas facile à tracer avec les questions

auxquelles on répond par un *discours* de plusieurs phrases. C'est pourquoi les énoncés suivants doivent être rangés dans la catégorie des interrogations auxquelles on répond par *au moins une phrase* :

- "Comment vous êtes-vous connus ? (Donnez quelques détails)"
(INED 081 : fiançailles 1)
- "Pour la France, selon vous, quel est le problème le plus important à l'heure actuelle ?"
(IFOP 58-9 : 1)
- "Parlant du Général de Gaulle, qu'est-ce qui vous plaît ou vous déplaît en lui ?"
(IFOP 58-9 : 16)
- "Par rapport aux autres villes que vous connaissez, quels sont les avantages de X ?"
(SOFRES 571 : 3)
- "Quelles ont été, quelles sont les incidences de votre vie professionnelle sur votre vie familiale (mari et enfants ; pour ceux-ci : santé, développement intellectuel, équilibre affectif, réussite scolaire, etc.) ?"
(F.ING 69 : IX 9)
- "Nommez votre emploi et décrivez-le de la manière la plus détaillée possible"
(CEREQ-SEU : 29a)

Peuvent également être rangées dans cette catégorie les "problèmes d'ingéniosité" du test de TERMAN (*adulte supérieur*, 6°), du TERMAN-MERILL (*forme L* : 14 ans, 4 ; *forme M* : 14 ans, 5) ou du TMI (TMI : 22).

Stanley L. PAYNE propose une intéressante classification empirique des questions de cette catégorie, selon leur fonction dans un questionnaire d'enquête [Payne 1951 : 34-44]. Les principales sous-catégories qu'il cite sont :

- question d'ouverture, qui sert à introduire un sujet (comme les deux premiers énoncés ci-dessus ; ex. : "Il y a des gens qui estiment nécessaire de lire des journaux et qui y consacrent du temps, et il y en a d'autres qui ne lisent pratiquement pas de journaux et qui estiment être suffisamment informés par la radio ; vous personnellement, qu'en pensez-vous ?"
SOFRES 1094 : *ancien lecteur*, 7) ;

- demande de suggestions (ex. : "A votre avis, qu'y aurait-il encore à changer ou à améliorer à X ?" *SOFRES 571 : 32*) ;
- demande de précisions (ex. : "Y a-t-il quelque chose que vous auriez aimé [...] faire, mais que vous n'avez pas pu [...] réaliser dans la vie ?" Si *oui* : "Quoi ?" *GEMAS 76 : 21,22*) ;
- question sur les causes ou les raisons (ex. : "Avez-vous l'intention d'aller voter ce jour-là ?" Si *non* : "Pourquoi ?" *IFOP 58-9 : 12a, 12b* ; "Quelles sont les raisons pour lesquelles vous n'avez pas l'intention d'acheter un modèle à carrosserie sport ?" *SOFRES 967 : 24*) ;
- demande d'arguments (ex. : "Avez-vous constaté qu'un argument soutenant votre position a paru particulièrement convainquant à votre entourage ? Est-ce que je pourrais savoir ce que c'est ?" *IFOP 58-9 : 10*) ;
- question de connaissance ou de mémoire (ex. : "Quelles sont les marques de petits cigares ou de cigarillos que vous connaissez au moins de nom ?" *SOFRES 1061 : 1*) ;
- question sur les sources d'information (ex. : "Comment avez-vous connu Z ? Est-ce que vous en avez entendu parler dans la presse ou par votre entourage, par une proposition d'abonnement ?" *SOFRES 1094 : lecteur, 8*) ;
- question factuelle (ex. : "Quelle est votre profession précise ?" *GEMAS 76 : 97*) ;
- relance (ex. : "Est-ce une des premières choses à améliorer à X, ou bien y a-t-il d'autres améliorations plus urgentes ?" Si *autres* : "Lesquelles ?" *SOFRES 659 : 34, 35*).

Les autres formes de questions à réponse libre ne sont guère utilisées que dans certaines épreuves psychologiques ; dans notre corpus, elles ne s'observent que relativement rarement. C'est pourquoi nous ne leur accordons

pas, dans cette présentation, la même place qu'à l'interrogation. En outre, nous aurons l'occasion de décrire des formes très voisines de celles-ci, à propos des questions à choix multiples ; ces dernières formes étant d'ailleurs, nous semble-t-il, d'un usage sensiblement plus fréquent. Aussi nous contentons-nous de mentionner, pour mémoire en quelque sorte, la phrase à compléter, les items d'appariement, et les items à substitution.

La *phrase à compléter* est utilisée essentiellement dans les épreuves psychologiques. REMMERS et GAGE donnent comme exemples : "Le joule est une unité de _____" ; "La mère a dit à sa fille : va à l'épicerie et _____ moi des pommes" [Remmers 1955 : 82]. WESMAN propose une phrase à compléter par plusieurs mots : "Le corps des insectes est composé de trois parties : _____, _____, _____. Les insectes ont _____ antennes et _____ pattes. Ils respirent à l'aide de _____" [Wesman 1971 : 89]. Les phrases à compléter relevées dans notre corpus ressortissent aux trois variétés suivantes :

- "Deux heures et demie font _____ minutes" (SUBES : page 5, 2)
- "Une personne qui se promenait dans une forêt s'est arrêtée tout à coup très effrayée [...] elle venait de voir à une branche d'arbre un ... Un quoi ?" (BS Z : 48)
- "Je pense que la plupart des filles ..." (STEIN : 25)
- "Ce qu'ils aimaient le plus chez lui c'était ..." (STEIN : 40)

Sur le modèle des deux premiers énoncés ci-dessus, on trouverait d'autres exemples de phrases à compléter dans la brochure d'information sur les tests de l'INOP : tests de lecture silencieuse, de grammaire (TCL), de connaissances littéraires (cf. [Bacher 1975 : 59, 61, 65]). Sur le modèle du troisième, on pourrait évidemment recenser l'ensemble des tests projectifs par complètement de phrases (cf. : [Rotter 1951 : 295-3], [Daston 1968 : 274-288]).

L'item d'*appariement* présente une certaine ressemblance avec la phrase à compléter, en ce sens que la réponse attendue est en général un mot ; la différence tient à ce qu'ici le stimulus n'est plus comme précédemment une phrase, mais également un mot. On peut en distinguer (au moins) deux variétés :

- l'item d'*analogie*, dans lequel on présente au sujet deux termes qui entretiennent entre eux une certaine relation (logico-sémantique) ; on présente ensuite un troisième terme, en demandant au sujet de trouver lui-même un quatrième terme qui soit avec le troisième dans la même relation que les deux premiers entre eux. Par exemple [Remmers 1955 : 102-103] :

- "Animaux : Oxygène. Plantes : _____"
- "Oxygène : O . Chlore : _____"
- "Le verbe est à l'adverbe ce que le nom est à _____".

- l'item d'*association*, dans lequel le sujet doit trouver, pour chaque terme d'une liste qui lui est présentée, un second terme qui entretient avec le premier une certaine relation, définie cette fois-ci explicitement (et non par référence à un exemple). Tel est le cas de l'item [Wesman 1971 : 89] :

- "A côté de chacun des noms de villes ci-après, écrivez le nom de l'Etat dans lequel elles se trouvent.
 Detroit :
 Chicago :
 [.....]"

A cette dernière variété appartiennent les épreuves d'association libre, utilisées pour l'étude des processus mentaux ou comme tests projectifs (cf. p.ex. : [Woodworth 1945 : 464-500], [Rotter 1951 : 279-295], [Anzieu 1960 : 152-156], [Daston 1968 : 264-274]).

L'item de *substitution* est une variante originale de la phrase à compléter, dans laquelle est présenté au sujet un énoncé contenant un élément (souligné) qui en rend la signification fautive ; le sujet est invité à substituer à cet élément fallacieux un autre élément du même type qui rend

la signification de l'énoncé vraie. Exemple [Wesman 1971 : 92] :

- "*L'usage de la vapeur comme force motrice a révolutionné les transports au 17^{ième} siècle*".

L'intérêt pratique de cette manière de poser la question est de soumettre celle-ci à une condition de catégorie qui restreint l'univers de choix de la réponse, sans toutefois définir explicitement cette condition.

c) Alternatives

Une alternative est une "option entre deux choses, entre deux propositions" (Émile LITTRÉ, *Dictionnaire de la langue française*). Par construction, l'alternative est la forme de question qui restreint le plus fortement le choix du sujet, puisque les réponses possibles sont au nombre de deux ; c'est par conséquent aussi la forme de question la moins proche d'un souci d' "objectivité", puisqu'elle n'envisage pas explicitement la possibilité, pour le répondant, soit de nier la présupposition de la question, soit de donner une réponse intermédiaire. C'est pourquoi, dans la pratique, la codification des réponses n'est jamais une dichotomie véritable, mais une liste pouvant atteindre cinq items (comme dans le deuxième exemple ci-dessous), voire plus. Stanley L. PAYNE illustre ce point en proposant, pour des alternatives, les précodifications suivantes [Payne 1951 : 60-61] :

- "*Allez-vous à la chasse, ou non ?*"
 - "*J'y vais*"
 - "*Je n'y vais pas*"
 - *autres réponses*
 - *ne sait pas*
- "*A votre avis, les hommes d'affaires que vous connaissez sont-ils optimistes ou pessimistes au sujet des perspectives commerciales de l'année prochaine ?*"
 - "*optimistes*"
 - "*pessimistes*"
 - "*ni l'un, ni l'autre*"
 - "*certains sont optimistes, d'autres pessimistes*"
 - *sans opinion.*

Aussi, dans la présentation de cette forme très usuelle de questions, n'avons-nous tenu compte que de l'énoncé lui-même, et non de la précodification ; sans cela, nous aurions dû conclure que les véritables alternatives n'existent pas dans la réalité (même pour le sexe, variable dichotomique par excellence, on peut être obligé de prévoir la codification : "sexe non précisé" cf. [Payne 1951 : 59]).

Pour mettre de l'ordre dans notre présentation, il nous a paru commode de distinguer les alternatives auxquelles on répond habituellement par un mot isolé (adverbe, adjectif) ou par une courte locution (ex. : "oui/non", "d'accord/pas d'accord"), et celles auxquelles on répond par une phrase complète, ou un fragment important de phrase. A l'intérieur de chacune de ces deux grandes catégories, nous avons tenu compte, pour décrire les diverses variétés recensées, de la nature des réponses possibles, et aussi de la présentation de séries de questions du même type (c'est-à-dire ayant la même requête, et la même liste de réponses codées possibles).

Les alternatives à *réponse brève* sont classées par WESMAN en questions à réponse par : "vrai/faux", par : "correct/incorrect", et par : "oui/non" [Wesman 1971 : 91] (nous ne tenons pas compte ici de deux autres variétés distinguées par WESMAN, et que nous avons reclassées ailleurs). Cette classification pourrait peut-être suffire pour les épreuves psychologiques ; mais l'analyse des formes de questions utilisées dans les questionnaires d'enquête nous a conduit à l'élargir. D'autre part, il nous a semblé nécessaire de classer à part les questions à réponse par *oui* ou *non*. En effet, cette forme de réponse paraît être de très loin la forme la plus usuelle, et surtout cette forme de question possède en français une propriété linguistique tout à fait remarquable : alors que, pour toutes les autres formes d'alternative (comme nous le verrons), les deux termes de l'alternative doivent être présentés dans l'énoncé interrogatif, pour la question à réponse par *oui* ou *non*, l'un des termes peut être sous-entendu (cf., pour la discussion

de cette propriété, l'analyse de la forme Q 1 au § 1.2.4.d, et les problèmes soulevés par la transformation φ_{z9} au § 1.2.4.e). Sur le plan pratique d'ailleurs, un auteur, réputé expert, comme Stanley L. PAYNE estime qu'il est souvent préférable, dans une question d'enquête, de ne pas expliciter le second terme : "... ou non ?" [Payne 1951 : 58].

Les exemples de question à réponse par *oui* ou *non* dans lesquelles l'alternative est implicite sont certainement, parmi toutes les formes possibles de questions, ceux que l'on trouve en plus grand nombre dans notre corpus. Nous n'en présentons que quelques échantillons :

- "*Recherchez-vous un autre emploi ?*" (CEREQ-SEU : 38)
- "*Ecoutez-vous la radio en voiture ?*" (ORTF 65 : 15)
- "*Trouvez-vous normal de laisser un pourboire au pompiste ?*" (SOFRES 340 : 29)
- "*Estimez-vous que les partis politiques existant avant le 13 mai sont encore capables de faire face à la situation politique actuelle ?*" (IFOP 58-11 : 5b)
- "*L'année où vous avez déposé votre demande d'admission à l'I.U.T., avez-vous fait acte de candidature dans une autre école, institut ou faculté ?*" (OCDE-IUT : 33)
- "*Hors de votre milieu de travail, discutez-vous fréquemment de vos lectures scientifiques et techniques ?*" (BNIST : 9)

Les séries de questions à réponse *oui/non* se rencontrent souvent dans les questionnaires psychologiques. Tel est le cas par exemple du *Questionnaire de personnalité* de BERNREUTER, qui ne contient que des questions de ce type (bien que la précodification admette une troisième modalité de réponse : "Si vous êtes absolument incapable de répondre « oui » ou « non » à la question, alors faites un cercle autour du point d'interrogation" BERN 35 :

instructions). D'autre part, un certain nombre d'échelles d'attitudes ou de questionnaires d'intérêts pour enfants et adolescents sont construits sur le même modèle (cf. p.ex. [Robinson 1973 : 206-217, 508-513]). Dans les questionnaires d'enquête, les exemples en sont moins fréquents, ou plutôt moins systématiques :

- "Si l'on doit pratiquer une politique de participation dans les entreprises, à votre avis, est-ce que cela doit plutôt être (UNE RÉPONSE PAR LIGNE) :

- . une participation des cadres aux décisions importantes concernant l'entreprise
- . une participation des salariés aux décisions importantes concernant l'entreprise
- . une participation des salariés et des cadres aux bénéfices réalisés par l'entreprise"

OUI	NON	?

(IFOP 69-5 : 11)

- "Depuis votre venue ici, vous est-il arrivé d'acheter ...

- . [un] réfrigérateur, une télé, ou une machine à laver
- . un moulin à café électrique ou d'autres petits appareils électro-ménagers
- . des meubles
- . une auto"

OUI	NON

(SOFRES 571 : 23)

- "Dans un ménage, trouvez-vous souhaitable que le mari ...

- . aide à faire la vaisselle
- . donne le biberon au bébé
- . aide parfois aux travaux ménagers
- . fasse de temps en temps le marché"

OUI	NON	mitigé sans avis

(SOFRES 686 : 11)

- "On nous a cité certains inconvénients des autobus de Z : à votre avis, est-ce vrai ou faux ?"

a - ils sont trop lents

b - ils passent rarement

c - ils n'ont pas assez de places assises

[.....]

j - il est difficile d'y voyager avec des enfants

V	F	NSP

(SOFRES 659 : 10)

- "On nous a donné un certain nombre d'opinions sur ... (localité où se fait l'interview). J'aimerais que vous me disiez si, à votre avis, c'est vrai ou faux

- à ... (localité) il n'y a pas grand chose comme distractions

[.....]

- quand on veut un bon médecin, on préfère souvent aller ailleurs qu'à ... (localité)"

vrai	faux	NSP oui et non

(SOFRES 571 : 5)

- "Dans chacune des phrases ci-dessous les mots sont mélangés. Si ce que dit la phrase est vrai, soulignez vrai. Si ce qu'elle dit est faux, soulignez faux.

[.....]

2°) gagne prudent on être à souvent

vrai faux

3°) du nord toutes les routes au sud vont

vrai faux

.....

18°) connaît que l'élève son maître toujours plus de choses"

vrai faux

(NB : 8)

En dehors de notre corpus, il est possible d'en trouver d'autres exemples dans les inventaires de personnalité et les échelles d'attitudes. Ainsi, le *Minnesota Multiphasic Personality Inventory* (MMPI) n'est-il composé que de questions à réponse vrai/faux ; il en est de même pour un certain nombre d'échelles moins connues, comme celle mesurant "la force du moi", de THOMAS

et ZANDER, la "nouvelle échelle de F" de WEBSTER, SANFORD et FREEMAN, l' "échelle de rigidité" de REHFISCH, ou l' "échelle de désirabilité sociale" de CROWNE et MARLOWE (cf. [Robinson 1973 : 132-135, 321-328, 388-392, 727-732]).

Les alternatives de la forme *d'accord/pas d'accord* peuvent être observées de façon isolée ; par exemple :

- *"Certaines personnes disent que les députés sont les principaux responsables du mauvais fonctionnement du régime de la IV^e République. Etes-vous d'accord ou non sur cette façon de voir les choses ?"* (IFOP 58-9 : 26)

Mais elles se rencontrent surtout en séries :

- *"Voici maintenant quelques opinions d'automobilistes concernant les stations-service que l'on rencontre sur la route. Pouvez-vous me dire, pour chacune d'elles, si vous êtes d'accord ou pas d'accord ?*
 - . *ça fait toujours plaisir de s'arrêter dans une belle station*
 - . *l'aspect de la station m'est totalement indifférent*
 - [.....]
 - . *on construit trop de stations en France, c'est une dépense inutile"* (SOFRES 340 : 21)
- *"Pour terminer, voici un certain nombre de jugements sur lesquels les avis diffèrent. Voudriez-vous donner votre opinion en disant si vous vous estimez d'accord ou non ?*
 - a) *pour résoudre le problème du logement, il faut faire entièrement confiance à l'initiative privée, au besoin en libérant complètement le prix des loyers*
 - b) *les fonctionnaires ne devraient pas avoir le droit de faire grève*
 - [.....]
 - g) *l'Etat doit subventionner les écoles libres"*

(IFOP 59-12 : 40)

- "Pouvez-vous me dire si vous êtes plutôt d'accord ou plutôt pas d'accord avec chacune des opinions suivantes :

. Si M. Pompidou est élu Président de la République, il pourra s'appuyer sur une majorité cohérente. Si c'est M. Poher on risque de revenir à l'instabilité politique

[.....]

. M. Poher est l'homme de la réconciliation nationale"

(IFOP 69-5 : 16)

En dehors de notre corpus, certaines échelles d'attitudes utilisent exclusivement cette forme d'alternative : échelles d'anomie de McCLOSKEY et SCHAAR, de SROLE, et de HYMAN et coll. ; échelles d'aliénation de OLSEN, de MIDDLETON, de HORTON et THOMSON ; etc. (cf. [Robinson 1973 : 252-259, 265-267, 279-282, 286-287]).

Les autres variétés d'alternatives à réponse brève utilisent des couples de mots contradictoires ou contraires différents des couples déjà mentionnés ("plaire/déplaire", "autrefois/maintenant", "augmenter/diminuer", etc.). Tel est le cas de la question : "Etes-vous pour ou contre le couvre-feu à neuf heures ?" [Payne 1951 : 58] (PAYNE énumère d'ailleurs divers couples de contraires, en les commentant ; cf. [Payne 1951 : 64-65]). Dans notre corpus, nous avons relevé :

- "Pouvez-vous me dire si ce projet vous plaît plutôt ou vous déplaît plutôt ?" (IFOP 58-9 : 6)

- "L'avenir vous paraît-il plus encourageant ou plus sombre qu'avant le référendum ?" (IFOP 58-11 : 4)

- "Souhaitez-vous, en votant, soutenir ou ne pas soutenir la politique du général de Gaulle ?" (IFOP 58-11 : 26)

- "Lisiez-vous davantage X autrefois ou maintenant ?" (SOFRES 686 : 40)

- "Avait-elle été achetée neuve ou d'occasion ?" (SOFRES 769 : 5bis)

- "Par rapport à votre domicile ou votre lieu de travail, l'agent ou le concessionnaire de votre marque est-il plutôt proche, ou plutôt éloigné ?" (SOFRES 967 : 31)
- "Dans le domaine de la religion, êtes-vous pour ou contre les changements, tels que messes en français, prêtres en civil ... ?" (SOFRES 686 : 19)
- "Dans l'ensemble, lorsque vous achetez de l'essence, est-ce que vous préférez que le pompiste soit un homme ou une femme ?" (SOFRES 340 : 11)
- "Pensez-vous que les différences de fortune ont tendance à augmenter ou à diminuer en France ?" (GEMAS 76 : 44)

Exemples de séries d'alternatives de ce type :

- "En ce qui concerne le levier des vitesses, il peut être placé au volant, au tableau de bord ou au plancher. Voulez-vous me dire, pour chacune de ces solutions, si elle vous convient ou pas ?"

	Convient	Ne convient pas	NSP
... Au volant ?			
... Au tableau de bord ?			
... Au plancher ?"			

(SOFRES 769 : 68)
- "S'il y avait des autobus beaucoup plus pratiques, est-ce que vous prendriez votre voiture moins souvent ou aussi souvent pour ..."

	moins souvent	aussi souvent
. aller faire une visite à des parents ou amis à Z		
[.....]		
. aller faire des achats à Z"		

(SOFRES 659 : 18 B)

réponse est enregistrée. Un moyen simple de distinguer les deux formes fondamentales d'alternatives est de rechercher si le choix est seulement entre deux termes contradictoires ou contraires (comme dans la dernière variété d'alternatives à réponse brève ci-dessus), ou s'il s'étend à au moins deux membres de phrase entre lesquels l'opposition peut être moins fortement marquée (comme dans la première variété ci-dessous).

Les alternatives portant sur un *membre de phrase* sont de la forme :

- "*Pensez-vous qu'un jeune homme a beaucoup d'occasions de se marier, ou bien que son choix est limité par beaucoup de circonstances ?*" (INED 081 : opinions 8a)
- "*Selon vous, vaut-il mieux, en vue d'un mariage considérer surtout l'attrance des personnes, ou bien la situation sociale ?*" (INED 081 : opinions 5)
- "*Estimez-vous que la boîte de vitesses automatique représente une technique d'avenir, qui se répandra de plus en plus, ou qu'il s'agit plutôt d'un équipement particulier, qui ne se répandra pas beaucoup ?*" (SOFRES 967 : 33)
- "*A votre avis, un bon député doit-il, à l'Assemblée Nationale, voter comme son parti lui ordonne de le faire, ou bien doit-il prendre ses décisions lui-même sans tenir compte des positions de son parti ?*" (IFOP 58-12 : 32)
- "*Dans l'ensemble, est-ce que vous pourriez dire, lorsque vous participez à une conversation sur le référendum que c'est vous qui écoutez davantage ou que c'est vous qui essayez davantage de convaincre votre entourage ?*" (IFOP 58-9 : 9)
- "*Avez-vous le sentiment que vous n'êtes bon à rien et que vous ne réussirez jamais rien dans votre vie :*
 - *très rarement ?*
 - *assez souvent ?*" (CAT 16 PF : 65)

Exemples de séries de questions appartenant à cette variété :

- "[...] Pour chaque point que je vais vous citer, dites-moi si vous estimez que les difficultés rencontrées étaient principalement la faute du Gouvernement ou si c'était des difficultés normales ?

 - . le poids excessif des impôts
 - . l'accroissement du chômage
 - [.....]
 - . les désordres dans les lycées et les facultés"
 - (IFOP 69-5 : 3)

On en trouve divers exemples dans les échelles d'attitudes recensées par ROBINSON et SHAVER : échelles de "responsabilité scolaire", échelle de motivation religieuse de WILSON (cf. [Robinson 1973 : 193-205, 697-701]).

Les alternatives portant sur un couple de phrases entières peuvent se rencontrer isolément dans un questionnaire, comme dans les exemples suivants :

- "En ce qui concerne le remplacement de la voiture, voici deux opinions contraires que l'on entend fréquemment chez les automobilistes. Laquelle de ces deux opinions se rapproche le plus de votre point de vue personnel ?
 - A. Aussi longtemps qu'une voiture roule, on a intérêt à la garder
 - B. Mieux vaut remplacer la voiture un peu plus tôt, et ne pas risquer de pépins"
 - (SOFRES 967 : 25)
- "[...] Laquelle de ces deux opinions se rapproche le plus de la vôtre ?
 - A. Pour acheter un break, il faut vraiment en avoir l'usage
 - B. Lorsque les deux versions, berline et break, existent pour un même modèle, on a intérêt à choisir le break"
 - (SOFRES 967 : 36)
- "a) Vous rangez soigneusement vos affaires en classe et à la maison
 - b) S'il y a un peu de désordre autour de vous, cela ne vous fait rien"
 - (E-I/6 : 20)

Mais on les trouve aussi fréquemment en séries :

- "Les alternatives suivantes peuvent se poser pour la France.
Voudriez-vous dire, pour chaque cas, la solution qui vous paraît la meilleure ?

a - engager des négociations avec le F.L.N. et le gouvernement algérien de Ferhat Abbas en vue de l'indépendance
- ou refus de toute négociation et politique d'intégration totale de l'Algérie à la France

[.....]

e - le maintien d'un ensemble franco-africain
- ou l'indépendance la plus rapide possible des territoires d'Outre-Mer" (IFOP 58-11 : 17)

De nombreuses échelles d'attitudes sont constituées uniquement d'alternatives portant sur des phrases entières. Parmi celles-ci, citons les échelles d'adaptation personnelle de SHOSTROM, d'intéro-extérodétermination de GURIN et coll. et de ROTTER, d'hétéronomie de JONES, de "confiance en autrui" de ROSENBERG, ainsi que deux adaptations de "l'échelle de F" par BERKOWITZ et WOLKON, et par SCHUMAN et HARDING [Robinson 1973 : 147-149, 218-224, 227-234, 355-362, 612-614, 329-341].

Avant d'aborder les questions à choix multiples, signalons la possibilité d'autres formes d'alternatives que celles que nous avons relevées dans notre corpus : "grappe" d'alternatives ("*cluster*"), formes mixtes incluant une alternative, et enfin "liste à cocher" ("*check-list*").

WESMAN appelle "*grappe*" d'alternatives une proposition incomplète suivie de plusieurs membres de phrase pouvant la compléter, chaque membre de phrase devant être jugé vrai ou faux [Wesman 1971 : 91]. Exemples :

- "Le volume d'une masse de gaz
 - a - tend à augmenter quand la température augmente
 - b - tend à augmenter quand la pression augmente
 - c - peut être maintenu constant en augmentant la pression et en diminuant la température
 - d - peut être réduit à zéro en augmentant la pression et en diminuant la température"

[Wesman 1971 : 92]

- "La constitution des Etats-Unis détermine :

- le nombre exact de magistrats pour la Cour Suprême

[.....]

- l'âge minimum requis pour pouvoir être élu député ou sénateur"

[Remmers 1955 : 84]

Une forme voisine, pouvant être classée dans la même variété, est décrite par REMMERS et GAGE : il s'agit d'une proposition pouvant être vraie ou fausse, suivie de propositions qui en sont déduites ; si la première est jugée vraie par le sujet, celui-ci doit également juger de la vérité des autres propositions. Exemple :

- "Le Président des Etats-Unis est le commandant en chef des armées

- . Le Président des Etats-Unis peut déclarer la guerre lorsqu'il en voit la nécessité

- . Si l'armée de l'air était séparée de l'armée de terre, le Président en serait le commandant en chef"

[Remmers 1955 : 85]

Les formes *mixtes* sont ici des alternatives suivies d'une question à réponse libre ou à choix multiples. Le premier exemple décrit par REMMERS et GAGE est une proposition comportant un mot souligné ; si la proposition est fausse, il faut trouver le mot correct devant être substitué au mot souligné (réponse libre) :

- "Je vais vous énumérer quelques titres de magazines féminins et vous allez me dire pour chacun s'il vous est déjà arrivé d'en lire ou d'en feuilleter un numéro quelconque même par hasard, que ce soit chez vous ou en-dehors de chez vous ?

1. *Elle*
2. *Marie-Claire*
3. *Confidences*
- [.....]
13. *Vie catholique*
14. *Écho des Françaises"* (SOFRES 686 : 1)

Mais de telles séries d'alternatives peuvent également, ainsi que nous allons le voir, être considérées comme des questions à choix multiples et à réponses multiples.

d) Questions à choix multiples

Les questions à choix multiples (parfois appelées : "questions-cafeteria") sont des énoncés interrogatifs dans lesquels on présente à la personne interrogée une liste finie de réponses possibles. Bien que totalement fermées, les questions à choix multiples sont parfois considérées comme une forme intermédiaire entre les questions à réponse libre et les alternatives. En effet, si le nombre des choix possibles se trouve réduit à deux, nous avons une alternative ; s'il tend vers l'infini, nous avons une question à réponse libre [Payne 1951 : 75]. Mais il s'agit d'autre part d'une forme de question tout à fait spécifique, ne se rencontrant que dans les tests et les questionnaires, et pratiquement exclue de la conversation courante (contrairement à la question à réponse libre et à l'alternative) [Payne 1951 : 75-76]. Les questions à choix multiples sont susceptibles de la même classification interne que les questions à réponse libre : interrogation, phrase à compléter, appariement, substitution. En outre, comme ces questions peuvent être soit à réponse unique, soit à réponses multiples, nous tiendrons compte, dans notre présentation, de cette subdivision supplémentaire.

Les *interrogations* à choix multiples peuvent être classées selon les relations postulées au sein de l'ensemble des réponses possibles proposées au répondant ; les cas les plus usuels étant : absence apparente de structure *a priori* (nous reviendrons sur ce point au § 3.1.3.a), et relation d'ordre. Dans le premier cas, on peut demander à la personne interrogée d'indiquer la réponse qu'elle juge exacte ou correcte (sélection), ou bien au contraire celle qu'elle juge inexacte ou incorrecte (rejet) :

- "Qui a inventé la machine à coudre ?

a. Singer

b. Howe

[.....]

e. Fulton

[Wesman 1971 : 94]

- "Parmi les propositions suivantes, laquelle n'est-elle pas vraie d'un virus ?

a - il ne peut vivre qu'à l'intérieur d'une cellule animale ou végétale

b - il peut se reproduire

c - il est formé d'un grand nombre de cellules vivantes

d - il peut provoquer des maladies" [Wesman 1971 : 95]

Doivent être également considérées comme appartenant à cette variété les questions dans lesquelles on demande à la personne interrogée de sélectionner, parmi les réponses possibles, celle qu'elle estime être la première d'un certain point de vue (la meilleure, la plus différente des autres, la cause des autres, etc.) ou au contraire la dernière (la plus mauvaise, la plus éloignée de sa propre conviction, la conséquence des autres, etc.). En effet, s'il est exact que le répondant doit d'une certaine manière établir une relation d'ordre entre les modalités de réponse, cette relation n'est pas fixée *a priori*.
Exemples :

- *"Quel était l'objectif principal du plan Marshall ?*
 - a. *défense militaire de l'Europe de l'ouest*
 - b. *relèvement économique et industriel de l'Europe de l'ouest*
 - c. *opposition à l'URSS*
 - d. *aide aux personnes sinistrées en Europe"* [Wesman 1971 : 94]

- *"Parmi les quatre événements ou circonstances historiques ci-après, lequel peut être considéré le plus comme étant la cause des trois autres ?*
 - 1. *intolérance des Anglais en matière de religion*
 - 2. *fondation de Jamestown*
 - 3. *liberté religieuse en Amérique*
 - 4. *voyage du Mayflower"* [Remmers 1955 : 94]

- *"Parmi les quatre événements ci-après, l'un est une conséquence des trois autres, qui ont contribué à le susciter. Quel est-il ?*
 - 1. *nauffrage du Lusitania*
 - 2. *guerre sous-marine totale*
 - 3. *entrée en guerre des Etats-Unis en 1917*
 - 4. *invasion de la Belgique par les Allemands"* (d'après [Remmers 1955 : 94])

Dans le second cas, c'est-à-dire lorsque le répondant doit choisir parmi les modalités de réponse ordonnées *a priori*, celles-ci peuvent avoir deux types de structure [Payne 1951 : 94-99] : ordre croissant selon une seule direction ("*questions of degree*"), ou selon deux directions opposées, à partir d'un point neutre origine ("*fold-over technique*"). Nous donnons ci-après un exemple de chacun de ces deux types :

- *"A votre avis, laquelle de ces politiques pensez-vous que les Etats-Unis devraient adopter à l'heure actuelle ?*
 - *entrer immédiatement en guerre contre l'Allemagne et l'Italie*
 - *fournir à la Grande-Bretagne tout le matériel de guerre que nous pouvons, et protéger les convois avec notre marine de guerre*

- *fournir à la Grande-Bretagne tout le matériel de guerre que nous pouvons, mais ne pas utiliser notre marine pour protéger les convois*

- *cesser toute aide à la Grande-Bretagne".*

[Payne 1951 : 96]

- *"Si une autre Société vous proposait un emploi avec la même rémunération, accepteriez-vous ...*

- *sûrement*

- *probablement*

- *probablement pas*

- *sûrement pas ?*

- *(ne sait pas ce qu'il ferait)"*

(adapté de [Payne 1951 : 97])

Les interrogations à choix multiples sans structure d'ordre sur les réponses peuvent être à réponse unique, comme dans les exemples ci-dessous (sélections ou rejets, isolés ou en série) :

- *"Parmi les hommes politiques connus, quel doit être à votre avis celui qui est le plus satisfait des résultats d'ensemble des élections ?"*
(IFOP 58-12 : 7)

- *"Où l'avez-vous vendue ? (ENQUÊTEUR : tendre liste-réponse)*

- *là où j'ai acheté ma nouvelle voiture*

- *à un autre garagiste*

- *à un particulier, un parent, un ami*

- *à la casse*

- *je ne l'ai pas (encore) vendue"*

(SOFRES 967 : 13)

- *"Pierre est à gauche de Jean et à droite de Luc. Qui est au milieu ?*

- *Pierre*

- *Jean*

- *Luc"*

(ECO 5 : RV 1)

- "Parmi les 4 chiffres que l'on vous propose, choisissez le plus vite possible celui qui est ...

[.....]

E. situé entre le double de 1 et la moitié de 10 :

1 2 3 7

[.....]

L. plus petit que le double de 3, et plus grand que la moitié de 6 :

0 2 3 5

(ECO 5 : LN)

- "Parmi les 4 dessins indiqués par des lettres, 3 se ressemblent par quelque chose. Mais un seul dessin est différent des 3 autres. Lequel ?" (ECO 5 : DG)
- "On vous donne un proverbe [...], et on vous demande de trouver parmi les cinq phrases écrites à la suite celle qui ressemble le plus au proverbe" (EAI 5/1 : Pr.)

Exemples d'interrogations à choix multiples, sans structure d'ordre *a priori* sur les réponses, admettant des réponses multiples :

- "Parmi les critiques suivantes qu'on entendait souvent adresser au fonctionnement de la IV^e République, quelles sont celles qui risquent, à votre avis, de s'adresser à la Ve République :

. Trop de partis sont représentés au Parlement

[.....]

. Le Président de la République n'a pas assez de pouvoirs"

(IFOP 58-12 : 34)

- "Quelles sont les deux qualités que vous appréciez le plus chez une femme, parmi les suivantes (Montrer liste)

- économe

- instruite, cultivée

[.....]

- capable de gagner de l'argent"

(SOFRES 686 : 13)

- "Pour quels usages utilise-t-on cette voiture assez régulièrement ?
(tendre liste)
- . Pour aller au travail
 - . Pour les déplacements dans le travail
 - . Pour faire les courses
 - . Pour les loisirs
 - . Pour les week-end
 - . Pour les vacances "
- (SOFRES 769 : 16)
- "Vous trouverez ci-dessous des pensées accompagnées de quatre ou six phrases. Vous ferez une croix dans les cercles correspondant aux deux phrases dont le sens vous paraît se rapprocher le plus (ou s'éloigner le moins) de la pensée"
- (BV 50)

Les interrogations à choix multiples à modalités de réponse ordonnées sont toujours à réponse unique. Exemples de questions dont les modalités sont ordonnées selon une seule dimension :

- "Est-ce que vous vous intéressez beaucoup, moyennement, peu ou pas du tout aux prochaines élections ?" (IFOP 58-11 : 18c)
- "A votre avis, est-il très important, important, ou sans importance, qu'une jeune fille se garde jusqu'au mariage ?" (INED 081 : opinions 7b)
- "Écoutez-vous les émissions régionales à la radio [...] ?
- souvent
 - quelquefois
 - rarement
 - jamais "
- (ORTF 65 : 6)
- "Dans les domaines de la mode, conseils ménagers, éducation des enfants, faites-vous habituellement confiance à ce magazine ?
(montrer liste pour guider les réponses de l'interviewé)
- | | | | |
|---|------|---------------------|--------------------------|
| | mode | conseils
ménager | éducation
des enfants |
| - je fais entièrement confiance | | | |
| - je fais le plus souvent confiance | | | |
| - j'en prends et j'en laisse | | | |
| - il est exceptionnel que je fasse confiance" | | | |
- (SOFRES 686 : 28)

Exemples de questions dont les modalités de réponse sont ordonnées selon deux dimensions à partir d'une origine (pouvant être implicite) :

- "Selon vous, le nombre des partis existant actuellement en France est-il trop grand, convenable, ou trop faible ?"
(IFOP 58-9 : 24)
- "En apprenant les résultats du référendum, quelle a été votre première réaction ? Etiez-vous très satisfait, plutôt satisfait, ou mécontent ?"
(IFOP 58-11 : 2)
- "Votre famille était-elle très favorable ou favorable, ou plutôt défavorable ou très défavorable à votre mariage ?"
(INED 081 : fiançailles 5)
- "Quand vous achetez de l'essence, prenez-vous ...
 - toujours du super
 - principalement du super
 - principalement de l'ordinaire
 - toujours de l'ordinaire ?"
(SOFRES 340 : 2bis)
- "Par rapport aux autres grandes villes comme Z, pensez-vous que Z est dans l'ensemble ...
 - vraiment très agréable à habiter ?
 - assez agréable à habiter, mais sans plus ?
 - pas très agréable ?
 - ou vraiment pas agréable ?"
(SOFRES 659 : 5)
- "Aimeriez-vous vous livrer aux occupations suivantes ? Pour chacune d'elles, faites une croix dans la colonne correspondante, en réservant la colonne « plaît beaucoup » aux occupations qui vous passionneraient, et la colonne « déplaît » à celles qui vous seraient très désagréables.

[.....]

	plaît beaucoup	plaît moyennement	indifférent	déplaît
1. Appartenir à un groupe sportif				
[.....]				
75. Vous entraîner en vue d'un match de basket-ball				

(QI Filles)

Cette forme d'interrogation est très fréquemment utilisée dans la construction des échelles d'attitudes, sous le nom de "technique de LIKERT" (cf. [Cronbach 1949 : 371-373], [Guilford 1954 : 459-460], [Torgerson 1958 : 47], [Coombs 1964 : 314-315], [Oppenheim 1966 : 133-142], etc.) ; on en trouvera de très nombreux exemples dans [Robinson 1973].

Les questions à choix multiples prenant une autre forme que celle de l'interrogation ne se rencontrent que très rarement dans les questionnaires d'enquête (cf. le premier exemple ci-dessous, tiré d'un questionnaire postal). Dans les épreuves psychologiques de notre corpus, nous avons relevé des phrases à compléter, et des items d'appariement, toujours présentés en séries. Exemples à *réponse unique* (sélections et rejets) :

- "Le type de travail que vous effectuez le plus fréquemment est :

- études technico-économiques
- études scientifiques
- études techniques
- essais - expérimentation
- réalisation - pré-série
- fabrication et industrialisation
- commercialisation
- autres"

(BNIST : 37)

- "Soulignez dans les phrases qui suivent le seul mot qui convient pour les compléter :

[.....]

20°) Socrate fut un : politicien - philosophe - savant -
général"

(NB : 1)

- "Dans les phrases suivantes, il manque un mot ou un groupe de mots. Trouvez parmi les 5 mots ou groupes de mots écrits en-dessous de la phrase celui qui convient le mieux [...]"

1. La fraîcheur du vent nous fait ...
frissonner frémir trembler palpiter tressaillir

.....

15. On ne peut être à la fois grand et petit, c'est ...
inconvenant incohérent inadmissible incompatible
inconséquent" (EAI 5/1 : CP)

- "Dans chaque ligne, soulignez le mot qui ne convient pas [...]"

1°) orme - broussailles - érable - chêne - peuplier

[.....]

18°) demander - emprunter - gagner - hériter - prêter"
(NB : 9)

- "On vous donne 3 mots et on vous demande d'en trouver un 4e qui soit au 3e comme le 2e est au 1er [...]"

1. Brûlant est à chaud comme glacé est à ...
gelé chambré congelé engourdi froid

[.....]

15. Film est à disque comme images est à ...
musique enregistrement projection sons écran"
(EAI 5/1 : An.)

Exemple à réponses multiples (en nombre fixé à l'avance) :

- "Dans chaque groupe de mots, soulignez les deux d'entre eux qui indiquent ce que les objets possèdent toujours [...]"

1. Un serpent a toujours :
du poison - un sifflement - des raies - une queue - une
langue

[.....]

20. L'antiphatie suppose toujours :

l'opposition - le dégoût - l'aversion - la peur - la
jalousie - (NB : 4)

Les variétés de questions à choix multiples observées dans notre corpus n'épuisent pas l'ensemble des variétés possibles décrites dans certaines classifications. C'est pourquoi nous signalons brièvement l'existence de quatre variétés supplémentaires : item de substitution à choix multiples, question à possibilité de réponses multiples restreinte, appariement multiple, et question à réponse incomplète.

Un item de *substitution* à choix multiple est évidemment très proche de la forme à réponse libre décrite précédemment. L'unique différence tient à ce que les éléments substituables doivent être choisis dans une liste finie. WESMAN illustre cette variété par une épreuve de correction d'un texte, dans laquelle le sujet doit substituer aux passages soulignés des mots ou groupes de mots tels que la correction, la précision, ou l'élégance du style s'en trouvent accrues [Wesman 1971 : 96].

Une question à *possibilité de réponses multiples restreinte* est une question à choix multiples et à réponses multiples, pour laquelle une précodification autorise seulement certaines des combinaisons possibles de réponses. Exemple :

- "La constitution actuelle [des Etats-Unis]

- a. est le fruit d'un échec antérieur
- b. fut rédigée à Philadelphie au cours de l'été 1787
- c. fut soumise pour adoption à chaque Etat par le Congrès
- d. fut adoptée par le nombre requis d'Etats et entra en vigueur en 1789.

Réponses possibles : 1 - a
 2 - a, b
 3 - a, b, c
 4 - b, c, d
 5 - a, b, c, d

[Wesman 1971 : 97]

Sur cet exemple, un tiers seulement des combinaisons de réponses possibles à été conservé dans la précodification.

L'appariement multiple ("compound matching") est une variante de l'appariement simple, dans laquelle à l'élément stimulus doivent être associés plusieurs éléments réponses, dont chacun provient d'une liste différente.

Exemples :

- " _____ est à eau ce que manger est à _____ "

- | | |
|--------------|---------------|
| 1. continuer | A. conduire |
| 2. boire | B. ennemi |
| 3. pied | C. nourriture |
| 4. fille | D. industrie |

(tiré de Differential Aptitude Tests par [Wesman 1971 : 100]).

- "A la suite de chacune des villes ci-dessous, indiquer le numéro de l'État dans lequel elle se trouve, puis la lettre correspondante à la principale activité industrielle de cette ville".

<u>États</u>	<u>Industries</u>	<u>Villes</u>
1. Illinois	a. abattoirs	
2. Indiana	b. aciéries	Detroit
3. Michigan	c. appareillage électrique	Akron
4. Minnesota	d. aéronautique	St Paul
5. New Jersey	e. automobile	Chicago
6. New York	f. caoutchouc	Schenectady
7. Ohio	g. édition	Pittsburg
8. Pennsylvanie	h. meunerie	
9. Texas	i. téléphone	

[Remmers 1955 : 99]

La question à réponse incomplète est une question à choix multiples et à réponse unique dans laquelle les réponses proposées sont soit des réponses partielles, définies toutefois de telle manière qu'une seule d'entre elles puisse être jugée vraie (premier exemple ci-dessous), soit des réponses portant sur un aspect tout-à-fait secondaire de l'objet de la question (deuxième exemple) [Wesman 1971 : 96-97] :

- "Le nom du disciple de Socrate le plus célèbre commence par une lettre comprise entre :

1. A et E
2. F et J
3. K et O
4. P et T
5. U et Z "

- "Si on calcule la racine carrée de 18, quelle sera la deuxième décimale du résultat ?

3 4 5 6 7 "

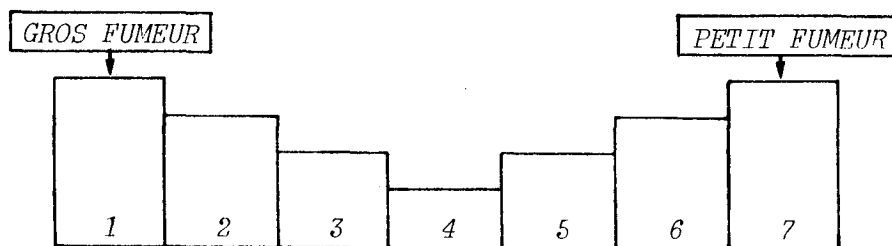
L'intérêt pratique de cette forme de question, dans les tests de connaissance ou d'aptitudes, est d'empêcher le sujet de déterminer la réponse correcte par élimination, dans l'ensemble des réponses proposées, de celles qui sont manifestement fausses.

e) Autres formes de questions

A côté des trois grands types de questions que nous avons distingués, il existe des formes moins courantes que certains auteurs rattachent aux questions à choix multiples. Nous n'examinerons ici que trois formes particulières observées dans notre corpus : les questions de cotation par échelle numérique, les questions d'ordination, et les questions de classification.

Les questions de *cotation* par échelle numérique, quoique pouvant prêter à confusion avec les questions à réponse libre numérique, appartiennent plutôt à la même catégorie que les interrogations dont les modalités de réponse sont ordonnées *a priori* (cf. [Guilford 1954 : 263-269]). La différence avec ces dernières tient à ce que seules sont explicitées les réponses extrêmes, les réponses intermédiaires étant désignées par une valeur numérique comprise entre celles affectées aux réponses extrêmes. Exemples :

- " ... Voici une sorte d'échelle (montrer barème) qui va de 0 à 10 : 0 correspond à une femme pas du tout moderne, 10 à une femme tout à fait moderne, et une note intermédiaire à une position intermédiaire. Où vous placeriez-vous, personnellement, sur cette échelle ?" (SOFRES 686 : 20bis)
- "Je vais vous demander maintenant quelle idée vous vous faites du fumeur de W. Voici une échelle (montrer) sur laquelle on a situé un gros fumeur (montrer), et un petit fumeur. Où situez-vous le fumeur de W. sur cette échelle entre un gros fumeur et un petit fumeur ?"



(SOFRES 1061 : 17a)

Cette technique est évidemment très proche de celle mise au point par OSGOOD sous le nom de "*Semantic Differential*" (cf. [Osgood 1957 : 80-85]). On relèverait divers exemples d'application de cette dernière technique dans les échelles d'auto-évaluation (cf. p.ex. [Robinson 1973 : 104-106, 123-125, 271-274]). Une technique voisine, mais relativement rare, impose au sujet l'estimation d'une *probabilité* (cf. test de "désir de certitude", de BRIM [Robinson 1973 : 409-412], dans lequel le sujet doit non seulement évaluer la probabilité de réalisation, en termes de nombre de chances sur 100, d'une situation ou d'un événement donné, mais encore indiquer, sur une échelle en cinq points, sa confiance dans l'estimation qu'il propose).

Les questions d'*ordination des réponses* sont de la forme :

- "Quel genre de lecture préférez-vous (classer de 1 à 4) :
- littérature (romans)
 - aventure - policiers - reportage - vie pratique
 - sciences et techniques
 - histoire - politique - sociologie - psychologie - philosophie"
- (BNIST : 34)

- "En vue d'un mariage, on peut considérer diverses qualités chez l'autre : la santé, l'aspect physique, les qualités morales, le milieu social, les goûts communs. A votre avis, qu'est-ce qui compte le plus ? Et ensuite ? Classez les autres qualités par ordre d'importance décroissante".

(INED 081 : opinions 3)

- "Parmi les sept éléments suivants, quels sont dans l'ordre les trois dont vous avez tenu compte pour prendre votre décision au moment de voter ? (montrer la carte)

- . liberté politique
- . efficacité politique
- . questions économiques et sociales
- . présence du Général de Gaulle au pouvoir
- . question de politique étrangère
- . problème algérien
- . problèmes des territoires d'Outre-Mer"

(IFOP 58-12 : 30)

Comme on le voit sur le dernier exemple, l'ordination peut être totale ou partielle. Les cas d'ordination totale (reclassement) sont décrits par REMMERS et GAGE (qui distinguent rangement chronologique et rangement logique, tout en signalant d'autres possibilités [Remmers 1955 : 103-106]) et par WESMAN (qui les range parmi les formes les plus complexes de questions à choix multiples [Wesman 1971 : 97]). Nous reviendrons plus en détail sur cette variété au § 3.1.2.b.

Les questions de *classification* enfin peuvent être illustrées par l'énoncé suivant :

- "Voici plusieurs points importants pour le confort d'une voiture. Voulez-vous les classer en trois groupes : les points très importants d'un côté, les points moins importants de l'autre, et au milieu, ceux qui sont moyennement importants.

- a) *moteur silencieux*
- b) *pas de bruit de roulement sur la chaussée*
- c) *position de conduite et disposition des commandes*
- d) *sièges confortables*
- e) *suspension*
- f) *bonne visibilité*
- g) *aisance des places pour les passagers"* (SOFRES 769 : 52)

Ce type de tâche, que l'on rencontre également dans diverses épreuves psychologiques, sera analysé ultérieurement (§ 3.1.2.c).

3.1.2. Classification systématique des formes de consigne

Nous avons déjà vu, en particulier à propos de la logique des questions et des réponses de Nuel D. BELNAP Jr., que l'on peut distinguer dans une question deux parties : le *sujet*, et la *requête* (§ 2.2.1.b). D'une certaine manière, la requête est une *consigne* donnée à la personne interrogée, afin que celle-ci parvienne à une forme de réponse satisfaisante pour celui qui interroge. Nous avons vu également que les formes de requêtes distinguées par les logiciens des questions (cf. § 2.4.1.c) ne permettent pas de couvrir l'ensemble des formes effectivement observées dans notre corpus.

Ce problème de la description formelle des requêtes a attiré l'attention de Clyde H. COOMBS, qui propose une classification simple des techniques de recueil de données [Coombs 1964 : 32-41], dans une perspective qui est plutôt celle de la psychologie expérimentale. En effet, COOMBS ne considère que les situations dans lesquelles l'individu est confronté à un ensemble (généralement fini) de stimuli, parmi lesquels il doit sélectionner celui ou ceux qui lui permettront d'élaborer sa réponse (cf. la description formelle

d'une telle situation expérimentale dans : [Bush 1963]). Par conséquent, la classification des formes de requêtes proposée par COOMBS ne s'applique qu'aux questions fermées (questions extensives, questions *whether*). COOMBS distingue deux familles de techniques de recueil des réponses :

- "*choisir* k/n ", dans laquelle on présente au sujet n stimuli, en lui demandant d'en sélectionner k ;

- "*ordonner* k/n ", dans laquelle on présente au sujet n stimuli, en lui demandant de classer les k premiers selon un ordre décroissant de préférence, de proximité par rapport à un étalon, etc.

Cette classification rend compte de la consigne d'un certain nombre de formes de questions que nous avons décrites. Par exemple :

- alternatives = "*choisir* $1/2$ " = "*ordonner* $1/2$ " ;
- questions à choix multiples à réponse unique = "*choisir* $1/n$ " ;
- questions à choix multiples à réponses multiples en nombre fixé à l'avance = (selon les cas) "*choisir* $2/n$ ", "*choisir* $3/n$ ", etc. ;
- questions de substitution = "*choisir* $1/n$ " ;
- questions d'appariement simple = "*choisir* $1/n$ " ;
- questions de cotation numérique = "*choisir* $1/n$ ", avec n = nombre de valeurs numériques possibles ;
- questions d'ordination = "*ordonner* k/n ", avec : $1 < k < n$ (puisque : "*ordonner* $(n-1)/n$ " et : "*ordonner* n/n " ont la même signification).

L'intérêt de la classification de COOMBS est qu'elle développe un point de vue non pris en compte par la logique érotétique, puisque celle-ci ne traite pas explicitement des questions d'ordination, alors que ces dernières sont d'un usage assez fréquent dans les tests et les questionnaires. Toutefois, cette classification présente à nos yeux deux inconvénients : elle néglige quelques aspects importants de la requête (comme l'exigence

d'exhaustivité) ; et elle laisse de côté certaines des formes de questions que nous avons observées dans notre corpus, ou qui sont décrites dans les textes que nous avons cités. Aussi proposons-nous ci-après une classification systématique des formes de consigne s'inspirant à la fois de la classification de COOMBS (dont elle est une extension), et de la logique érotétique. Nous distinguons quatre formes de consigne : sélection, ordination, classification, et appariement. Les deux premières formes s'appliquent à ce que Claude FLAMENT appelle un "questionnaire simple" ; les deux dernières correspondent à un "questionnaire composé" [Flament 1976 : 62]. Nous verrons toutefois qu'il est possible, au prix d'une légère complexification de l'écriture, de réduire ces quatre formes à la plus simple d'entre-elles : la sélection.

a) Sélection

Les consignes de la forme : "*choisir k/n*", avec $k < n$, s'appliquent à plus de la moitié des questions d'un questionnaire, probablement en raison de la simplicité de la tâche à laquelle elles correspondent. Trois éventualités sont à considérer : k est fixé et est égal à 1 ; k est fixé et est > 1 ; et k n'est pas précisé dans la consigne, et est laissé à la discrétion de la personne interrogée. En outre, la consigne peut exprimer ou non l'exigence d'exhaustivité. On obtient ainsi six variétés de consignes de sélection, assez proches des formes de requêtes décrites par BELNAP. Une bonne illustration de consigne avec exigence d'exhaustivité est donnée par l'énoncé suivant :

- "*Soulignez dans les phrases qui suivent le SEUL mot qui convient pour les compléter :*

EXEMPLE : *La capitale de la France est :*
Lyon Le Havre Paris Bordeaux [...]" (NB : 1)

Cette consigne, qui correspond à la requête d'une question *whether* à réponse vraie unique, pourrait s'écrire : "*choisir \forall 1/4*". En pratique toutefois, il est rare que l'exigence d'exhaustivité soit clairement exprimée ("*choisir tous les ...*") ou clairement omise ("*choisir k exemples de ...*") ; c'est pourquoi nous n'avons pu en tenir compte pour classer les quelques exemples présentés ci-après.

Les consignes de la forme "choisir 1/n" ont pour particularité d'être formellement identiques aux consignes "ordonner 1/n". En outre, elles s'appliquent évidemment aux alternatives, qui ont en principe pour consigne : "choisir 1/2". Nous avons donné suffisamment d'exemples d'alternatives au § 3.1.1.c pour qu'il soit inutile d'en présenter de nouveaux. Par contre, la forme : "choisir 1/n" peut être illustrée par :

- "Trouvez-vous que dans un magazine féminin, la présentation, c'est-à-dire la beauté des photos, des dessins, des couleurs, la clarté de la mise en page ait ...
 - . une importance primordiale
 - . une assez grande importance
 - . peu d'importance
 - . pas d'importance du tout ?" (SOFRES 686 : 24bis)
- "A quel titre poursuivez-vous ces études ? (une seule réponse)
 - . à titre personnel
 - . comme fonctionnaire stagiaire ou élève fonctionnaire
 - . comme salarié envoyé par votre employeur, au moins en partie sur votre temps de travail
 - . comme demandeur d'emploi envoyé par l'ANPE ou l'APEC
 - . formation donnée par l'armée dans le cadre du service militaire" (CEREQ-SEU : 20)
- "Avez-vous voté pour ce candidat surtout ... (MONTRER LA CARTE)
 - . parce qu'il était le moins éloigné de vos opinions
 - . parce qu'il pouvait battre un autre candidat dont vous redoutiez l'élection
 - . pour ces deux raisons à la fois
 - . parce qu'il était sûr d'être élu
 - . parce que vous vouliez manifester votre opposition au candidat que vous n'aimiez pas et qui était sûr d'être élu ?" (IFOP 58-12 : 22)

La forme de consigne de sélection : "choisir k/n" est celle des énoncés suivants :

- "Vous trouverez ci-dessous des pensées accompagnées de quatre ou six phrases. Vous ferez une croix dans les cercles correspondant aux deux phrases dont le sens vous paraît se rapprocher le plus (ou s'éloigner le moins) de la pensée" (BV 50)
- "En matière économique et sociale aucun gouvernement ne peut tout faire en même temps. Dans les prochaines années, quelles sont parmi les suivantes (MONTRER LA CARTE) les trois mesures auxquelles le Gouvernement doit donner la priorité ? [...]" (IFOP 69-5 : 4)

Enfin, la forme : "choisir x/n " (non prise en compte par COOMBS) correspond aux cas où le nombre d'items à sélectionner n'est pas déterminé à l'avance. Exemples :

- "Regardez la première image, la chaise, le balai, le banc et la table. Faites un trait sur tous les objets qui servent pour s'asseoir" (GILLE 44 : 5)
- "Quels reproches parmi les suivants (MONTRER LISTE) faites-vous à X ? [...]" (SOFRES 686 : 42)
- "Qu'est-ce qui vous gêne particulièrement ? (LISTE)
 - . les embouteillages dans N
 - . les embouteillages aux portes de N
 - . les difficultés pour stationner
 - . les dangers pour les piétons
 - . le bruit fatiguant et énervant" (SOFRES 571 : 29bis)
- "Citez les revues que vous lisez au moins une fois par semaine [...]" (BNIST : 16)
- "Une diminution des impôts entraîne une diminution des dépenses de l'Etat. Dans quels domaines parmi les suivants souhaitez-vous que l'Etat dépense moins (MONTRER LA CARTE) [...] ?" (IFOP 69-5 : 5b)
- "Si vous avez hésité [entre formation universitaire et formation en école], qu'est-ce qui a finalement déterminé votre choix ? (ne cochez qu'une seule case)
 - . préférence personnelle pour ce type d'enseignement
 - . meilleurs débouchés
 - . désirs des parents
 - . moindre coût des études" (F.ING 69 : VI 9)

Il est intéressant, lors de la construction d'un test ou d'un questionnaire d'enquête, de calculer le nombre de patrons de réponses distinctes possibles à une question donnée. Cela permet en effet d'évaluer ce que HAMBLIN appelle le *contenu informationnel* de la question [Hamblin 1958 : 166-168], et COOMBS sa *capacité de canal* [Coombs 1964 : 34-36], c'est-à-dire la quantité d'information que peut *a priori* apporter toute réponse à la question (en considérant que toutes les réponses possibles sont équiprobables ; nous reviendrons sur ce point au § 3.1.2.f). En outre, dans les tests de performances ou de connaissances (dans lesquels il y a des réponses vraies et des réponses fausses), cela sert de base au calcul de la probabilité qu'a le

sujet de fournir une réponse exacte en répondant au hasard. Ce nombre n'est pas lié à la condition d'exhaustivité ; il dépend évidemment du nombre d'éventualités présentées n , et de l'effectif k de la sélection. Lorsque k est fixé, le nombre N de réponses directes possibles est égal au nombre de combinaisons sans répétition de n objets pris k à k :

$$N_{\text{choisir } k/n} = \frac{n!}{k! (n-k)!}$$

Lorsque k n'est pas fixé mais ne peut être nul ("*choisir x/n* "), le nombre de réponses distinctes possibles est égal à la somme des nombres de réponses correspondant aux consignes "*choisir k/n* " avec k variant de 1 à n :

$$N_{\text{choisir } x/n} = \sum_{k=1}^{k=n} N_{\text{choisir } k/n} = 2^n - 1$$

(Si l'on pouvait avoir $x = 0$, le nombre de réponses possibles serait alors de : $N = 2^n$). Le tableau n° 3.1. indique le nombre de réponses distinctes possibles à la consigne : "*choisir k/n* ", pour : $1 \leq k \leq n \leq 12$.

Nous n'avons pas trouvé, dans notre corpus, de consigne de sélection analogue, dans son principe, aux formes de requêtes décrites par BELNAP (§ 2.2.1.b), c'est-à-dire avec un nombre minimum v et un nombre maximum distinct u d'items à sélectionner. Il est vrai qu'en pratique les interviewés interprètent souvent la consigne : "Indiquez les k items qui ..." comme : "Indiquez au moins un et au plus k items qui ..." ; en effet, le dénombrement des réponses aux questions à choix multiples dans les questionnaires d'enquête révèle toujours l'existence de répondants qui n'ont donné qu'une seule réponse au lieu (par exemple) des trois qui leur étaient demandées. C'est pourquoi il nous paraît utile de prévoir une forme de consigne plus générale, pouvant s'écrire : "*choisir $\frac{u}{v}/n$* ", et signifiant : "Parmi les n items proposés, sélectionnez-en au moins v et au plus u tels que ...".

$k \backslash n$	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2		1										
3			1									
4				1								
5					1							
6						1						
7							1					
8								1				
9									1			
10										1		
11											1	
12												1

Tableau n° 3.1. Nombre de patrons de réponse distincts possibles à la consigne : "choisir k/n ".

Dans le cas général, le dénombrement des patrons de réponses distinctes possibles obéit à la formule :

$$N_{\text{choisir}} \frac{u}{v/n} = \sum_{k=v}^{k=u} N_{\text{choisir}} \frac{k}{n} = \sum_{k=v}^{k=u} \frac{n!}{k! (n-k)!}$$

Le tableau n° 3.2 indique (pour : $1 < u \leq n \leq 12$) le nombre de réponses distinctes possibles à la consigne "*choisir* $\frac{u}{1/n}$ ", qui correspond à l'interprétation usuelle des consignes de choix multiples dans les enquêtes. Dans la diagonale de ce tableau, on lira le nombre de réponses distinctes possibles à la consigne : "*choisir* x/n " = "*choisir* $\frac{n}{1/n}$ ".

b) Ordination

Avec les consignes d'ordination, nous sortons apparemment du champ exploré par la logique érotétique, puisque tous les travaux que nous avons cités au chapitre 2 se limitent aux consignes de sélection. Bien que les six variétés que nous avons distinguées pour les consignes de sélection soient théoriquement concevables également pour les consignes d'ordination, nous n'avons observé que deux de ces variétés possibles qui soient identifiables sans ambiguïté : "*ordonner* $1/n$ ", et : "*ordonner* k/n ". En effet, dans tous les exemples relevés, l'exigence d'exhaustivité n'est pas clairement exprimée, et le nombre d'items à ordonner est toujours précisé. A ces deux variétés s'ajoute une troisième qui n'a pas d'équivalent observable dans les consignes de sélection : "*ordonner* n/n " (ou, ce qui revient au même : "*ordonner* $n-1/n$ ").

La forme : "*ordonner* $1/n$ " peut être illustrée par les énoncés de questions suivants :

- "*Lisez chaque question et faites une croix devant la MEILLEURE réponse comme le montre l'exemple.*
EXEMPLE : Pourquoi achetons-nous des horloges ?
 1°) *Parce que nous aimons les entendre sonner*
 2°) *Parce qu'elles ont des aiguilles*
 X 3°) *Parce qu'elles nous indiquent l'heure [...]*" (NB : 2)

$\frac{u}{n}$	2	3	4	5	6	7	8	9	10	11	12
2	3										
3	6	7									
4	10	14	15								
5	15	25	30	31							
6	21	41	56	62	63						
7	28	63	98	119	126	127					
8	36	92	162	213	246	254	255				
9	45	129	255	381	465	501	510	511			
10	55	175	395	607	847	967	1012	1022	1023		
11	66	231	561	1023	1485	1815	1980	2035	2046	2047	
12	78	298	793	1585	2509	3301	3796	4016	4082	4094	4095

Tableau n° 3.2. Nombre de patrons de réponse distincts possibles à la consigne : choisir $\frac{u}{n}$.

- "Voici un dessin représentant trois formes possibles pour l'arrière d'une voiture (TENDRE DESSIN). Voulez-vous me dire laquelle des formes S, H ou F vous conviendrait le mieux pour le genre de voiture que vous comptez acheter ?" (UNE SEULE REPONSE) (SOFRES 769 : 62)

La consigne : "ordonner k/n ", avec : $k < n-1$, plutôt rare dans les épreuves psychologiques, s'observe assez couramment dans les questionnaires d'enquête :

- "Parmi les sept éléments suivants (L'ENQUETEUR MONTRE UNE CARTE), quels sont dans l'ordre les trois dont vous tiendrez le plus compte pour prendre votre décision au moment du référendum ? [...]" (IFOP 58-9 : 11)
- "Pourquoi avez-vous entrepris des études supérieures ? (parmi les raisons qui, à cette époque, vous ont incitée à le faire, donnez les deux principales, dans l'ordre d'importance, en portant les chiffres 1 et 2 dans les cases correspondantes) : [...]" (F.ING 69 : VI 5)
- "Quel mot s'associe le mieux pour vous à l'information Scientifique et Technique (en choisir 6 et les classer de 1 à 6) :
théorie réalisation revues [...] création" (BNIST : 29)

Enfin, la consigne : "ordonner n/n " (ou : "ordonner $n-1/n$ ") se rencontre plus fréquemment dans les tests que dans les questionnaires d'enquête :

- "C'est l'histoire d'une fosse que l'on creuse. Classez les images" (GILLE 44 : 51)
- "Chaque ligne comprend 5 mots placés dans un ordre quelconque, et vous devez les ordonner logiquement en commençant par celui qui porte le numéro 1.
EXEMPLE : (A) médiocre (B) bon (C) mauvais (D) moyen
(E) excellent" 1 (EAI 5 : GV)
- "Voici une phrase dont les mots sont tous mélangés, et qui n'a pas de sens. Si les mots étaient rangés en ordre, ça ferait une bonne phrase. Regarde bien, et tâche de me dire comment la phrase doit être [...]" (TERMAN : 12 ans 4°)

- "Votre réaction face à une situation nouvelle est-elle de (classer de 1 à 4) :
 - . la connaître, la comprendre
 - . trouver ce qu'il y a de beau, d'agréable ou de désagréable
 - . de la comparer par similitude avec des situations analogues
 - . de remarquer ce qu'il y a d'insolite et de neuf ?"

(BNIST : 27)

- "Pour quelle(s) raison(s) avez-vous opté pour le Département d'I.U.T. où vous êtes élève (numérotez par ordre d'importance) ? [...] "

(OCDE - IUT : 35)

Le nombre de réponses distinctes possibles pour une consigne d'ordination : "ordonner k/n " est égal au nombre d'arrangements sans répétition de n objets pris k à k :

$$N_{\text{ordonner } k/n} = \frac{n!}{(n-k)!}$$

Le tableau n° 3.3 indique le nombre de patrons de réponses distinctes possibles pour une consigne d'ordination avec : $1 \leq k \leq n \leq 12$. On y vérifie que le nombre de réponses possibles à la consigne : "ordonner n/n " correspond au nombre de permutations sans répétition de n objets, et que c'est également le cas pour : "ordonner $n-1/n$ " :

$$N_{\text{ordonner } n/n} = N_{\text{ordonner } n-1/n} = n!$$

En pratique, dans une enquête par sondage, il arrive assez fréquemment que la personne interrogée n'applique pas rigoureusement la consigne qui lui est imposée, et ordonne un nombre d'items plus grand ou plus petit que le nombre demandé. Dans le premier cas, si l'enquêteur n'est pas intervenu et a enregistré la totalité de la réponse, il suffit évidemment de ne retenir que les k premiers items cités ; dans le second cas, la réponse est incomplète, et obéit en fait à la consigne : "ordonner $\frac{k}{n}$ ". Le nombre de réponses possibles est alors égal à :

n	k											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	2	2										
3	3	6	6									
4	4	12	24	24								
5	5	20	60	120	120							
6	6	30	120	360	720	720						
7	7	42	210	840	2520	5040	5040					
8	8	56	336	1680	6720	20160	40320	40320				
9	9	72	504	3024	15120	60480	181440	362880	362880			
10	10	90	720	5040	30240	151200	604800	1814400	3628800	3628800		
11	11	110	990	7920	55440	332640	1663200	6652800	19953400	39916800	39916800	
12	12	132	1320	11880	95040	665280	3991680	19958400	79833500	239500800	479001600	479001600

Tableau n° 3.3. Nombre de patrons de réponse distincts possibles à la consigne : "Ordonner k/m".

$$N_{\text{ordonner}} \frac{k}{1/n} = \sum_{i=1}^{i=k} N_{\text{ordonner}} \frac{i}{n}$$

à condition que : $k < n$. En effet, étant donné que :

$$N_{\text{ordonner}} \frac{n-1}{n} = N_{\text{ordonner}} \frac{n}{n}$$

on a l'égalité :

$$N_{\text{ordonner}} \frac{n-1}{1/n} = N_{\text{ordonner}} \frac{n}{1/n}$$

Le tableau n° 3.4 donne le nombre de réponses distinctes possibles à cette consigne pour : $1 \leq k < n \leq 12$. Nous n'avons pas trouvé, dans l'ensemble des tests et des questionnaires analysés, d'exemple de consigne de la forme : "*ordonner* $\frac{u}{v/n}$ " , avec : $v > 1$.

c) Classification

Les consignes de classification demandent au répondant de répartir les items en plusieurs classes selon leurs ressemblances et leurs dissemblances, c'est-à-dire de définir une partition sur l'ensemble des items proposés. Dans les exemples que nous avons relevés, le nombre p de classes désirées est toujours précisé ; par contre, selon les cas, les effectifs de toutes les classes, ou de certaines d'entre elles, peuvent être imposés ou laissés libres. Nous symboliserons la consigne détaillée : "*classer* n items en p classes dont les effectifs sont respectivement = e_1, e_2, \dots, e_p " .

(avec : $p < n$, et : $e_1 + e_2 + \dots + e_p = n$) par : "*classer* n/p :

e_1, e_2, \dots, e_p " . Lorsque l'effectif e_i d'une classe i n'est pas précisé dans la consigne, nous le remplacerons par l'inconnue x_i . Si aucun effectif n'est imposé, la consigne s'écrira en abrégé : "*classer* n/p " .

Dans les épreuves psychologiques, les effectifs des classes sont pratiquement toujours fixés, et égaux ; on a par conséquent la relation :

n \ k	k											
	2	3	4	5	6	7	8	9	10	11	12	
2	2											
3	9	9										
4	16	40	40									
5	25	65	205	205								
6	36	156	516	1236	1236							
7	49	259	1099	3619	6659	8659						
8	64	400	2080	8800	28960	69280	69280					
9	81	585	3609	18729	79209	260649	623529	623529				
10	100	820	5860	36100	187300	792100	2606500	6235300	6235300			
11	121	1111	9031	64471	397111	2060311	8713111	28671511	68588311	68588311		
12	144	1464	13344	108384	773664	4765344	24723744	104557344	344058144	823059744	823059744	

Tableau n° 3.4. Nombre de patrons de réponse distincts possibles à la consigne : "ordonner $1^k/n$ ".

$n = p \cdot e_i$, quelle que soit la classe i . Exemples :

- "Les activités que nous proposons à votre choix sont groupées par séries de neuf. Dans chaque série, vous devez indiquer les trois activités que vous préférez et les trois que vous aimez le moins [...]" (QIA/m)
- "[...] Choisissez d'abord les six activités qui vous plairaient le plus et mettez une croix dans la colonne de gauche : « plaît beaucoup » en face de chacune d'elles. Cherchez ensuite les six occupations qui vous plairaient le moins, et mettez une croix dans la colonne de droite « déplaît ». Il reste six activités pour lesquelles vous marquez une croix dans la colonne du milieu : « plaît moyennement »." (QI Garçons)

Le nombre de patrons de réponses distinctes possibles à une consigne de ce type est naturellement égal au nombre de p -partitions de l'ensemble en classes d'effectif $k = \frac{n}{p}$:

$$N_{\text{classer } n/p} : e_i = \frac{n}{p} = \frac{n!}{\left(\frac{n}{p}\right)!^p}$$

Par exemple, dans le premier exemple, où $n = 9$ et $p = 3$, on a $N = 1.680$ réponses distinctes possibles. Dans le second, avec un simple doublement de l'effectif des classes, on aboutit à un nombre de réponses possibles de $N = 17.153.136$. En effet, en passant de $n = 9$ à $n' = 18$, on a multiplié

$$N \text{ par : } \frac{18! - 9!}{(6! - 3!)^3} = 10.210,2 .$$

Il arrive, dans les questionnaires d'enquête, que seuls soient fixés les effectifs des deux classes extrêmes, comme dans l'exemple ci-dessous :

- "Plusieurs explications ont été avancées pour exprimer les raisons qui ont amené bon nombre d'électeurs à voter pour l'U.N.R. Parmi les explications suivantes, pouvez-vous citer les deux qui vous paraissent les plus justes ; les deux qui vous paraissent les plus fausses ?
 - . Soutenir l'action du général de Gaulle
 - . Désir d'union nationale
 - . Lutter contre le communisme
 - . Voter pour des hommes nouveaux
 - . Equivoque politique

- . Favoriser l'intégration de l'Algérie à la France
- . Ignorance politique
- . Battre les candidats des anciens partis
- . Inconscience des dangers du parti unique" (IFOP 58-12 : 10)

Dans ce cas particulier, où les deux classes extrêmes ont le même effectif $e_1 = e_3 = k$, le nombre de réponses distinctes possibles est égal à :

$$N_{\text{classer } n/3 : k, n-2k, k} = \frac{n!}{(n-2k)! (k!)^2}$$

Soit dans cet exemple, avec $n = 9$ et $k = 2$: $N = 756$ réponses possibles. Le nombre de réponses distinctes possibles à une consigne de classification de n éléments, dont les effectifs des p classes sont fixées à l'avance, et valent respectivement e_1, e_2, \dots, e_p , est égal au nombre de permutations avec répétition de n objets dont e_1, e_2, \dots, e_p sont identiques, ou, si l'on préfère, au nombre de (e_1, e_2, \dots, e_p) -partages d'un ensemble de cardinal n :

$$N_{\text{classer } n/p : e_1, e_2, \dots, e_p} = \frac{n!}{e_1! e_2! \dots e_p!}$$

Le cas le plus général, à nombre de classes fixé, s'observe lorsque les effectifs des classes sont laissés entièrement libres :

- "Nous allons résumer les qualités que vous recherchez pour le genre de voiture que vous voulez acheter. Pour cela, vous pouvez classer ces cartons en 3 groupes : d'un côté les points très importants, de l'autre les points peu importants, et au milieu, ceux qui sont entre les deux pour vous
 - a) Economie d'utilisation (essence, réparations, prix des pièces, assurance, vignette)
 - b) Prix d'achat
 - c) Sécurité (tenue de route, freinage)
 - d) Habitabilité (espace intérieur, nombre de places, volume du coffre)
 - e) Nervosité (reprises en général, rapidité au démarrage, facilité pour monter les côtes)
 - f) Robustesse du moteur et des parties mécaniques
 - g) Robustesse de la carrosserie, résistance à l'usure"
- (SOFRES 769 : 75)

- "Y-a-t-il parmi les suivantes (MONTRER LISTE) des rubriques que vous trouvez prendre une place trop importante dans ... (citez le magazine rappelé ci-dessus). Lesquelles ? Y-a-t-il, au contraire, des rubriques qui n'ont pas une place assez importante ? A votre avis, lesquelles ?
- . Mode
 - . Tricot/ouvrage
 - . Maison/décoration
 - [.....]
 - . Beauté
 - . Hygiène, santé" (SOFRES 686 : 25, 25bis)
- "Je vais maintenant vous présenter une série de cartes sur lesquelles sont inscrites toutes les différences pouvant exister entre les individus, et je voudrais que vous les classiez en faisant 3 paquets :
- le premier paquet où vous mettez celles qui vous paraissent jouer un rôle important dans la vie ;
 - le second regroupera celles qui vous semblent jouer un rôle assez peu important ;
 - le troisième enfin comprendra celles dont vous pensez qu'elles ne jouent aucun rôle.
- . différence de fortune
 - . différence de salaire
 - [.....]
 - . différence d'opinion politique
 - . différence de santé" (GEMAS 76 : 71)

Ce cas ne se rencontre guère que dans les questionnaires d'enquête, peut-être en raison du nombre particulièrement élevé des patrons de réponses possibles. En effet, ce nombre est égal au nombre des applications possibles de l'ensemble des items dans l'ensemble des classes, soit :

$$N_{\text{classer } n/p} = p^n$$

Dans le premier exemple cité ci-dessus, on a par conséquent : $N = 3^7 = 2.187$ réponses possibles. Pour le second, dans lequel nous n'avons pas fait figurer l'ensemble des 16 rubriques proposées, on aurait : $N = 3^{16} = 43.046.721$ réponses distinctes possibles. Pour le troisième enfin, dans lequel il y avait 28 rubriques, on aurait : $N = 3^{28} > 22.876$ milliards de réponses distinctes possibles.

Pour terminer ce panorama des consignes de classification utilisées dans les tests et les questionnaires, signalons qu'il existe des épreuves psychologiques dans lesquelles on demande au sujet de fournir, s'il le peut, plusieurs classifications d'un même ensemble d'items. Tel est le cas du test d' "abstraction" pour adultes (*ABS*) mis au point par le Service de Recherches de l'INOP, dans lequel on présente au sujet des séries de 12 ou de 8 éléments (graphiques ou verbaux), à classer en deux groupes égaux. Il y a donc *a priori* :

$$N = \frac{12!}{6!6!} = 924 \text{ bipartitions possibles pour } n = 12, \text{ et } N = \frac{8!}{4!4!} =$$

70 bipartitions possibles pour $n = 8$. Si le nombre des bipartitions qu'un sujet peut citer n'était pas limité, on aurait donc en principe :

$$N_{12} = 2^{294} - 1 \simeq 10^{278} \text{ réponses distinctes possibles pour } n = 12, \text{ et :}$$

$N_8 = 2^{70} - 1 \simeq 10^{21}$ réponses distinctes possibles pour $n = 8$. Comme les consignes de correction de l'épreuve nous apprennent qu'il n'y a pas plus de 10 réponses correctes par série de 12 items (*ABS : S5*), et de 15 réponses correctes par série de 8 (*ABS : V9*), on peut calculer un nombre de réponses distinctes possibles plus réaliste en assimilant l'épreuve (respectivement) à la sélection d'un sous-ensemble de 1 à 10 classifications sur 924 pour $n = 12$, et d'un sous-ensemble de 1 à 15 classifications sur 70 pour $n = 8$. Soit, en appliquant la formule proposée au § 3.1.2.a, un nombre de réponses distinctes possibles égal à :

$$N_{12} = N_{\text{choisir}} \binom{10}{1} / 924 = \frac{925!}{11! \cdot 914!} \simeq 10^{25}$$

pour les séries de 12 items, et à :

$$N_8 = N_{\text{choisir}} \binom{15}{1} / 70 = \frac{71!}{16! \cdot 55!} \simeq 3 \times 10^{15}$$

pour les séries de 8 items. Pour faciliter la passation et la correction du test, une liste de 28 réponses possibles pour $n = 12$, et de 30 réponses possibles pour $n = 8$, est présentée au sujet, qui doit choisir la ou les bonnes réponses parmi celles-ci. Nous reviendrons au § 3.1.2.f sur les contraintes introduites par cette présentation.

d) Appariement

Nous avons précédemment cité quelques exemples de questions d'appariement simple (§ 3.1.1.b) et multiple (§ 3.1.1.d). Les consignes d'appariement simple s'observent fréquemment :

- "Soulignez le mot qui convient.
EXEMPLE : Oreille sert à entendre comme oeil sert à table - voir - main - jouer [...]" (NB : 7)
- "Regardez la première image, la pomme, l'auto, la cerise, et la balance. Deux choses peuvent aller ensemble. Quelles sont-elles ? [...]" (GILLE 44 : 8)

Malgré leur formulation, qui fait appel à un mécanisme d'association ou d'analogie, elles ne se distinguent guère des consignes de sélection. C'est pourquoi nous ne traiterons, dans ce paragraphe, que des formes d'appariement multiple.

En principe, une consigne d'appariement multiple porte sur deux ensembles disjoints d'items à apparier. Comme nous le verrons sur les exemples que nous avons recensés, le rôle joué par chacun de ces ensembles n'est pas symétrique : l'un d'eux est en quelque sorte imposé à la personne interrogée (il contient des items *stimuli*), tandis que l'autre est simplement mis à sa disposition pour qu'elle en sélectionne certains éléments (il contient des items *réponses*). Nous conviendrons donc de considérer le premier comme l'ensemble de départ d'une correspondance dont le second est l'ensemble d'arrivée ; l'objet de la consigne étant de demander à la personne interrogée d'établir entre les deux ensembles une correspondance d'un certain type (c'est cette correspondance qui constitue la réponse du sujet). Cette convention nous fournit un schéma de classification commode pour les consignes d'appariement : nous analyserons en premier lieu les consignes demandant une application, en commençant par les applications bijectives ; après avoir passé en revue les autres formes d'applications envisageables, nous terminerons par les correspondances qui ne sont pas des applications. D'autre part, si nous appelons m le cardinal de l'ensemble de départ D , n le cardinal de l'ensemble d'arrivée A , et k le cardinal du sous-ensemble C de $D \times A$ définissant la correspondance demandée, la consigne : "apparier $k/m,n$ " signifiera : "définir k relations d'association binaires, dont chacune a pour premier élément l'un des m items

de l'ensemble de départ, et pour second élément l'un des n items de l'ensemble d'arrivée". Dans le cas où la relation demandée est une application, nous précisons de quel type d'application il s'agit, en utilisant l'initiale de chacun des types spécifiés, et la lettre A pour une application qui n'est ni injective, ni surjective. Comme, par ailleurs, dans toute application, nous avons par définition : $k = m$, la forme générale d'une consigne demandant une application quelconque sera : "*appariar (A) m/m,n*".

Une consigne demandant un appariement bijectif s'écrit par conséquent : "*appariar (B) n/n,n*". Bien que cette forme de consigne soit assez courante dans certaines formes de jeux pour enfants (p.ex. deviner à quel personnage appartient tel objet), nous n'en avons pas trouvé d'exemple dans les tests papier-crayon que nous avons examinés. Dans les épreuves individuelles, les planches à encastrement ("form boards") dans lesquelles à chaque découpe de la planche correspond un objet et un seul (ce que P.R. BIZE appelle "planches à encastrement simple" [Bize 1958 : I, VI 77]), illustrent bien la consigne bijective, ainsi d'ailleurs que les items d'identification et de discrimination du TERMAN-MERRILL :

- "[...] *Trois petits morceaux de bois, de mêmes formes et de mêmes grandeurs peuvent entrer dans ces encoches. On présente le tout en place à l'enfant [...]. On dit alors à l'enfant : « Maintenant remets les morceaux dans les trous ».*"
(TM 37 : L 2ans 1°)
- "*Montrez la carte et dites : « Montre-moi ...*
 - a) ... *avec quoi nous faisons la cuisine*
 - b) ... *ce que nous prenons quand il pleut*
 - [...]
 - f) ... *qui attrape les souris »*" (TM 37 : L 4ans 4°)
- "*Le matériel comprend deux cartes avec des dessins d'animaux identiques, mais diversement ordonnés [...]. « Tu vois tous ces animaux ? Trouve m'en un autre exactement pareil à celui-ci » (Montrer)."*
(TM 37 : M 3,5ans 3°)

Dans le premier exemple, $n = 3$; dans le second, $n = 6$; dans le troisième : $n = 12$. En outre, dans ce dernier, l'ensemble de départ est également ensemble d'arrivée, et la correspondance demandée est l'application identique.

Le nombre de patrons de réponses possibles distincts à une consigne bijective est évidemment égal au nombre de bijections possibles :

$$N_{\text{appariier (B) } n/n,n} = n!$$

Les consignes d'appariement injectif sont de la forme : "appariier (I) $m/m,n$ ", avec : $m < n$. Pratiquement absentes de notre corpus, elles semblent assez fréquentes dans les tests de connaissance conçus par les pédagogues américains :

- "Quel est :

- | | |
|---|-----------------------|
| 1. le premier président des U.S. | A. Alexander Hamilton |
| 2. le chef de la justice de la cour suprême | B. John Marshall |
| 3. le secrétaire du Trésor | C. George Washington |
| | D. Andrew Jackson" |
- (cité par [Bize 1958 : I, VI 35]).

- "En face de chaque phrase décrivant une découverte scientifique, écrivez la lettre précédant le nom de l'auteur de cette découverte.

- | | |
|--|--------------------|
| 1. A découvert la circulation du sang | a. Louis Pasteur |
| 2. A réalisé l'étude statistique de l'hérédité humaine | b. Gregor Mendel |
| 3. A effectué des expériences décisives sur les mécanismes de l'hérédité | c. Francis Galton |
| | d. Robert Koch |
| | e. William Harvey" |

[Wesman 1971 : 98]

On trouverait d'autres exemples de consignes injectives dans [Remmers 1955 : 100-102]. Le nombre de réponses distinctes possibles est ici égal au nombre d'injections possibles (c'est-à-dire au nombre d'arrangements de n objets pris m à m

$$N_{\text{appariier (I) } m/m,n} = \frac{n!}{(n-m)!}$$

Les consignes d'appariement qui demandent une application surjective sont de la forme : "appariier (S) $m/m,n$ ", avec : $m > n$. Elles peuvent être illustrées par les exemples suivants, tirés de questionnaires d'enquête exclusivement :

- "Voici des dessins de voitures (tendre dessins R et S)
Regardez bien ces deux voitures R et S et dites-moi, à votre avis ...
- dans laquelle des deux il y a plus de place pour les passagers ?
- et pour les bagages ?
- laquelle est la plus courte ?
- vous préférez laquelle des deux ?" (SOFRES 769 : 77)
- "Si la décision finale était d'abattre certaines rues ou certains quartiers, voulez-vous me dire lesquels, parmi ceux que je vais vous citer, seraient à abattre, lesquels seraient à conserver ?
a) Rue des Clercs
[.....]
g) Quartier du Pontiffroy et de Chambière" (SOFRES 659 : 38)

Nous n'avons pas trouvé d'exemple indiscutable de cette forme de consigne dans les tests et les questionnaires psychologiques que nous avons recensés. Un exemple fictif (obtenu par modification d'une question de connaissance citée par [Wesman 1971 : 98-99]) pourrait être :

- "En face de chacune des oeuvres musicales ci-dessous, écrivez la lettre qui correspond au nom de son auteur.
·Concertos Brandebourgeois a. J.S. Bach
·Capriccio Italien b. Beethoven
·Symphonie Héroïque c. Moussovgsky
·Une nuit sur le Mont Chauve d. Tchaïkovsky"
·Casse-noisette
·Toccata et fugue en ré mineur

L'application demandée est bien une surjection dans la mesure où chaque compositeur cité est l'auteur d'au moins une des oeuvres proposées. Mais ceci n'est pas exprimé dans la formulation de la question, et, pour le sujet, la correspondance demandée apparaît comme une application sans propriété particulière. C'est peut-être la raison pour laquelle l'appariement par surjection ne se rencontre pratiquement pas dans les épreuves psychologiques, dans lesquelles en principe l'uniformisation des conditions de passation est plus poussée que dans les questionnaires d'enquêtes : la légère diminution du nombre des réponses possibles serait acquise au prix d'un accroissement de la difficulté de compréhension de la consigne. Le nombre de réponses distinctes

possibles à une consigne de ce type est égal au nombre de surjections possibles d'un ensemble de cardinal m dans un ensemble de cardinal $n < m$:

$$N_{\text{appariier (S) } m/m,n} = n! S(m,n)$$

où : $S(m,n)$ est un nombre de STIRLING de seconde espèce (cf. [Comtet 1970 : II 38-39, 58 § 8(1°)]). A titre d'illustration, les valeurs de N pour les trois exemples cités seraient :

$$N_{\text{appariier (S) } 4/4,2} = 2! \times 7 = 14$$

$$N_{\text{appariier (S) } 7/7,2} = 2! \times 63 = 126$$

$$N_{\text{appariier (S) } 6/6,4} = 4! \times 65 = 1560 \text{ réponses distinctes possibles.}$$

Les consignes d'appariement demandant une application qui ne soit ni surjective, ni injective, s'écrivent : "appariier (A) $m/m,n$ ". Les cubes de KOHS en sont une illustration très connue : si l'on considère l'ensemble de départ de l'application comme étant constitué par les m cases du modèle proposé (avec, selon les figures, $m = 4$, $m = 9$, ou $m = 16$), et l'ensemble d'arrivée par les $n = 6$ faces distinctes des cubes servant à élaborer la réponse, il apparaît clairement que, pour un modèle donné, certaines faces peuvent n'être pas utilisées, alors que d'autres le sont pour plusieurs cases. Autre exemple tiré d'un test de connaissances :

- "Chaque série de questions ci-dessous est composée de cinq réponses possibles, désignées chacune par une lettre, et de trois phrases numérotées. Pour chaque phrase numérotée, indiquez la réponse qui vous paraît convenir le mieux. Dans chaque série, toute réponse peut être utilisée une ou plusieurs fois, ou ne pas être utilisée du tout.
- A. période cambrienne
 - B. période dévonienne
 - C. période jurassique
 - D. période précambrienne
 - E. ère tertiaire

1. *Les êtres vivants les plus nombreux sont les trilobites*
 2. *Les premières algues apparaissent*
 3. *Les pantotheria y sont très nombreux"*
- (cité par [Thorndike 1971 : 180])

Le nombre d'appariements possibles de cette forme est évidemment égal au nombre total d'applications d'un ensemble de cardinal m sur un ensemble de cardinal n . C'est-à-dire que l'on a l'égalité :

$$N_{\text{appariier}}(A) \ m/m,n = N_{\text{classer}} \ m/n = n^m$$

(soit, dans l'exemple ci-dessus : $N_{\text{appariier}}(A) \ 3/3,5 = 5^3 = 125$ réponses distinctes possibles). En pratique, on n'utilise évidemment pas une formulation du type : "*classer m/n*" lorsque $m < n$ (dans l'exemple ci-dessus, on aurait : "ranger chacune des trois questions dans l'une des cinq catégories A,B,C,D,E"). A cette réserve linguistique près, les deux formes de consigne sont équivalentes, en ceci qu'elles conduisent en principe à des résultats identiques.

Enfin, il arrive qu'une consigne d'appariement ne demande pas une application, mais une correspondance sans propriété particulière. Nous en avons vu un exemple, de la forme : "*appariier 1/4,4*", au § 3.1.1.d ("_____ est à eau ce que manger est à _____", tiré de [Wesman 1971 : 100]). Dans une telle consigne, de forme : "*appariier k/m,n*", le nombre de relations binaires possibles entre l'ensemble de départ et l'ensemble d'arrivée est égal au produit $m.n$; par conséquent, le nombre de réponses distinctes possibles est ici égal au nombre de combinaisons sans répétition possibles de $m.n$ objets pris k à k :

$$N_{\text{appariier}} \ k/m,n = \frac{(m.n)!}{k! \ ((m.n)-k)!}$$

(soit, dans notre exemple, $N = 16$ réponses possibles). Il peut même arriver que le nombre de relations à établir soit laissé à la discrétion du répondant, comme dans l'exemple d'interrogation bi-nucléaire proposé par BELNAP (cf. § 2.2.2.d, et [Belnap 1976 : 76-77]), que l'on pourrait reformuler ainsi :

- "Quels garçons sont les frères de quelles filles ?

<u>Garçons</u>	<u>Filles</u>
Emile	Annie
François	Béatrice
Georges	Colette
Henri	Denise "

Selon les règles d'écriture adoptées pour les consignes de sélection, la consigne devient ici : "appariéer $x/m.n$ ", puisque le nombre de couples liés par la relation citée peut varier de 0 à $m.n$. Le nombre de réponses distinctes possibles est alors égal à :

$$N_{\text{appariéer } x/m,n} = 2^{m.n}$$

soit dans notre exemple $N = 2^{16} = 65.536$ réponses distinctes possibles. On pourrait de même, sur le modèle des consignes de sélection, étudier sans difficulté la forme générale : "appariéer $u/m,n$; mais cette forme ne nous a pas parue usitée dans les tests et les questionnaires.

Il reste à examiner les consignes d'appariement portant sur plus de deux ensembles distincts. Nous en avons présenté au § 3.1.1.d., un exemple tiré de [Remmers 1955 : 99], dans lequel on demandait au sujet d'associer, à chacune des six villes des États-Unis qui lui étaient proposées, le nom de l'État dans lequel celle-ci est située, ainsi que le nom de la principale activité industrielle qui y est installée ; le nombre des noms d'États était de 9, celui des types d'industrie de 9 également. Il est clair qu'une telle consigne peut en réalité être considérée comme la combinaison, en une seule question, de deux consignes d'appariement, ayant toutes deux pour ensemble de départ la liste des six agglomérations. La consigne d'appariement villes-États est évidemment de type injectif (chaque ville ne peut être que dans un seul État). Si l'on fait l'hypothèse qu'il en est de même pour l'appariement villes-industries, le nombre de réponses distinctes possibles est égal à :

$$N = N_{\text{appariéer } (I) 6/6,9} \times N_{\text{appariéer } (I) 6/6,9} = 3.657.830.400 .$$

Si par contre nous supposons qu'une industrie peut être associée à plus d'une agglomération, nous avons :

$$N = N_{\text{apparié (I) 6/6,9}} \times N_{\text{apparié (A) 6/6,9}} = 32.141.551.550$$

patrons de réponses distincts possibles (soit près de neuf fois plus).

e) Réduction à la sélection

Nous avons déjà noté que la consigne : "ordonner 1/n" était formellement équivalente à : "choisir 1/n". En réalité, il est possible de trouver d'autres équivalences du même type (certaines sont d'ailleurs suggérées par la formulation même des questions). Aussi allons-nous montrer que toute consigne peut être ré-écrite sous la forme d'une séquence de consignes de sélection. Cette possibilité de réduire toute forme de consigne à la sélection facilitera la comparaison de la puissance (ou de la capacité érotétique) de chacune de ces formes.

Toute consigne d'ordination de k éléments peut être considérée comme une série de k sélections successives d'un élément unique. C'est d'ailleurs ainsi que sont présentées les consignes d'ordination suivantes :

- "[...] Vous devez d'abord considérer les trois activités du groupe et décider laquelle vous plaît le plus. [...]. Vous devez ensuite décider quelle est, dans le même groupe, celle des trois activités qui vous plaît le moins.
[...]" (CHAUDAGNE : consigne)
- "Dans les domaines suivants, quel est celui que vous aimeriez voir bénéficier en premier lieu de vos sacrifices ; et ensuite (en classer 3 par ordre d'importance) ?
 - . recherche scientifique et progrès technique
 - . logement
 - . Algérie
 - . Afrique Noire
 - . enseignement" (IFOP 58-9 : 33c)

- "Voici maintenant un certain nombre de qualités que l'on peut désirer trouver chez un pompiste (TENDRE LISTE). Pouvez-vous me dire quelle est la qualité que vous considérez comme la plus importante ; et celle que vous mettez en second, et celle que vous mettez en troisième, enfin celle que vous mettez en quatrième ?"
- . bien stylé
 - . s'y connaît bien en voiture
 - . souriant, bonne tête
 - . a toujours une tenue impeccable
 - . travaille vite
 - . sympathique
 - . respectueux
 - . donne des conseils techniques
 - . vous accueille avec un mot aimable"

(SOFRES 340 : 27)

Dans le premier exemple ci-dessus, l'ordination de trois activités, en raison de l'intérêt que celles-ci présentent pour le sujet, est obtenue à l'aide de deux consignes de sélection successives. Si nous convenons d'utiliser le point et virgule (;) comme opérateur de relation indiquant la succession, la consigne du premier exemple s'écrit : "ordonner 2/3" = "choisir 1/3 ; choisir 1/2". Dans les exemples suivants, l'ordination de k éléments sur n (avec : $k < n$) est obtenue par k consignes de sélection d'un élément. D'où l'équivalence formelle :

$$\text{ordonner } k/n = \text{choisir } 1/n ; \text{ choisir } 1/n-1 ; \dots ; \text{ choisir } 1/n-k+1.$$

A titre de contrôle, on peut vérifier que le nombre de patrons de réponses distincts possibles à la consigne : "ordonner k/n " est bien égal au produit des nombres de réponses possibles à k consignes de sélection d'un élément sur un ensemble de respectivement n , $(n-1)$, ... , $(n-k+1)$:

$$N_{\text{ordonner } k/n} = \frac{n!}{(n-k)!} = n \times (n-1) \times \dots \times (n-k+1).$$

De même, toute consigne de classification de n éléments en p classes distinctes peut être considérée comme équivalente à une séquence de $p-1$ consignes de sélection. Certains des exemples cités au § 3.1.2.c

utilisent d'ailleurs cette équivalence formelle dans la présentation d'une consigne de classification à effectifs imposés (IFOP 58-12 : 10) ou libres (SOFRES 686 : 25, 25bis). A la limite, il est même possible de tenir pour équivalentes une consigne de la forme : "classer $n/2$ " et une série de n alternatives (de forme : "choisir $1/2$ ") ; exemples :

- "Voici différents avantages qu'une station peut présenter ; voulez-vous me dire, pour chacun d'eux, si votre endroit habituel le présente ou non ? [...]" (SOFRES 340 : 6c)
- "Avez-vous été satisfait ou mécontent d'apprendre la réélection de Messieurs :
 - . Guy Mollet
 - . Georges Bidault
 - . Jacques Soustelle
 - . Paul Reynaud
 - . Félix Gaillard ?" (IFOP 58-12 : 5)

En utilisant les mêmes conventions d'écriture que précédemment, nous avons les équivalences formelles :

$$\begin{aligned} \text{Classer } n/p : e_1, e_2, \dots, e_p &= \\ \text{choisir } e_1/n ; \text{ choisir } e_2/n-e_1 ; \dots ; \text{ choisir } e_{p-1}/n-(e_1+\dots+e_{p-2}) \end{aligned}$$

lorsque les effectifs des classes sont fixés, et :

$$\begin{aligned} \text{classer } n/p &= \\ \text{choisir } x_1/n ; \text{ choisir } x_2/n-x_1 ; \dots ; \text{ choisir } x_{p-1}/n-(x_1+\dots+x_{p-2}) \end{aligned}$$

(avec certains x_i pouvant être nuls) lorsque les effectifs des classes sont laissés entièrement libres. On peut vérifier aisément que le nombre de patrons de réponses possibles à : "classer $n/p : e_1, e_2, \dots, e_p$ " est bien égal au produit des nombres de réponses possibles aux $p-1$ consignes de sélection successives :

$$N_{\text{classer } n/p : e_1, e_2, \dots, e_p} = \frac{n!}{e_1! e_2! \dots e_p!} =$$

$$\frac{n!}{e_1! (n-e_1)!} \times \frac{(n-e_1)!}{e_2! (n-(e_1+e_2))!} \times \dots \times \frac{(n-(e_1+\dots+e_{p-2}))!}{e_{p-1}! (n-(e_1+\dots+e_{p-1}))!}$$

puisque : $n-(e_1+\dots+e_{p-1}) = e_p$.

Les consignes d'appariement peuvent, pour la plupart d'entre-elles, être réduites à une consigne d'ordination ou de classification ; par transitivité, elles sont donc également réductibles à la sélection. L'appariement bijectif de deux ensembles de cardinal n équivaut à l'ordination totale de l'un des ensembles ; en effet, si l'on définit sur chaque ensemble une relation d'ordre (arbitraire), toute bijection correspond à un reclassement de l'ensemble d'arrivée (permutation) :

$$\text{appariement (B) } n/n, n = \text{ordonner } n/n$$

En suivant le même raisonnement, toute injection d'un ensemble de cardinal m dans un ensemble de cardinal n équivaut à l'ordination de m éléments d'un ensemble de n éléments :

$$\text{appariement (I) } m/m, n = \text{ordonner } m/n$$

De même, une application quelconque équivaut, comme nous l'avons déjà noté, à une classification :

$$\text{appariement (A) } m/m, n = \text{classer } m/n$$

Cette équivalence est bien illustrée par l'énoncé suivant (qui illustre également l'équivalence entre classement et sélection) :

- Voici une liste de quelques journaux hebdomadaires. Dans cette liste, y en a-t-il que vous lisez régulièrement, quelquefois, jamais ?

- . Paris-Match
- . Jours de France
- . Canard Enchaîné
- . Express
- . La France Catholique
- . La Terre
- . France-Observateur
- . Carrefour
- . Rivarol
- . Elle
- . Témoignage Chrétien"

(IFOP 58-9 : 38c)

Dans cet exemple en effet, on peut voir soit une consigne de la forme : "appariier (A) 11/11,3" ou : "classer 11/3", soit une série de 11 consignes : "choisir 1/3". Enfin, au paragraphe précédent, nous avons considéré qu'une consigne d'appariement demandant une correspondance quelconque, entre deux ensembles de cardinal m et n , pouvait être réduite à une consigne de sélection dans l'ensemble produit de cardinal $m.n$:

$$\text{appariier } k/m,n = \text{choisir } k/(m.n)$$

Seule, la consigne de surjection pose quelques problèmes, en imposant que toute réponse apparaisse au moins une fois (c'est-à-dire que tout élément de l'ensemble D soit image d'au moins un élément de l'ensemble A). On peut toutefois considérer cette forme de consigne comme la succession d'une sélection de n éléments de D ("choisir n/m "), d'une bijection de ces éléments sur A ("appariier (B) $n/n,n$ ", équivalant formellement à une ordination totale), et d'une application quelconque des $m-n$ éléments restant de D sur A ("appariier (A) $(m-n)/(m-n),n$ "). Cette représentation simpliste ne se prête guère, cependant, au calcul du nombre de réponses *distinctes* possibles, ce qui marque les limites du système d'écriture que nous avons adopté. En effet, ce système ne permet pas de rendre compte des contraintes exercées par la bijection sur les réponses possibles à la dernière application, contraintes qui se traduiraient ici par l'existence de patrons de réponses possibles non distincts. Nous reviendrons au paragraphe suivant sur ces problèmes

de contraintes sur les réponses.

Auparavant, notons que la réduction à la sélection permet de formaliser des consignes n'entrant pas dans notre système de classification. Par exemple :

- *"Voici une liste de 4 voitures. Parmi ces 4 voitures, quelles sont à votre avis ...*
 - . *les 2 plus spacieuses*
 - . *les 2 plus robustes*
 - [.....]
 - . *les 2 qui offrent la meilleure sécurité"*
- (SOFRES 967 : 19)

Dans cette consigne d'appariement, l'ensemble de départ comporte 9 items (qualités), et l'ensemble d'arrivée 4 (voitures). Il ne peut évidemment s'agir d'une application, puisque chaque élément de l'ensemble de départ a en principe deux images. Si l'on écrit cette consigne sous la forme d'une demande de correspondance ("*apparier 18/9,4*"), on néglige le fait que chaque élément de D doit avoir exactement deux images ; si l'on essaye de l'écrire sous la forme de deux applications successives ("*apparier (A) 9/9,4*"), on ne peut exprimer l'exigence de distinction des deux images d'un même élément (Ces difficultés soulignent à nouveau les insuffisances de notre système). Par contre, on peut traduire cette consigne par une série de 9 consignes de sélection, de la forme : "*choisir 2/4*" (ou éventuellement "*choisir $\frac{2}{0/4}$* ").

f) Puissance d'une question

Toute question a pour finalité d'apporter de l'information au locuteur. Or, "tout acte de communication n'apporte de l'information que dans la mesure où il réduit une situation d'ignorance ou d'incertitude relativement à l'état de choses considéré" ; "information et incertitude sont deux concepts étroitement liés, et si nous pouvons mesurer l'incertitude, nous pouvons mesurer l'information en tant que diminution de l'incertitude" [Garner 1962 : 3]. L'information apportée par une assertion dépend de la sélection effectuée par

l'énonciation, parmi l'ensemble des assertions possibles dans un contexte donné (cf. [Mac Kay 1961 : 205]), c'est-à-dire en dernière analyse de la question à laquelle elle répond. Pour évaluer l'incertitude du locuteur, il faudrait pouvoir affecter une probabilité d'apparition à chacune des réponses possibles à la question posée. En l'absence d'information sur ce à quoi s'attend le locuteur, ou encore dans le cas général (quel que soit le locuteur), on fera l'hypothèse de l'incertitude maximum ; c'est-à-dire que l'on supposera :

- que toutes les réponses possibles sont indépendantes les unes des autres (il n'existe pas de réponse qui soit impliquée par une autre réponse ; toutes les réponses s'excluent mutuellement) ;

- que chaque réponse a *a priori* autant de chances que les autres d'être fournie (s'il y a n réponses possibles, chacune a une probabilité d'apparition égale à : $1/n$).

La quantité d'information apportée par la réalisation d'un événement est égale à l'incertitude que l'on avait *a priori* sur sa réalisation ; elle est d'autant plus grande que cet événement était considéré comme moins probable (cf. [Bar-Hillel 1953 : 149]). On peut donc adopter comme mesure de l'information apportée par un événement une fonction monotone croissante de l'inverse de sa probabilité d'apparition, en posant en outre que la réalisation d'un événement certain n'apporte aucune information. Ceci nous conduit à deux propositions intéressant la théorie des questions :

- une réponse à une *question de HOBSON* n'apporte pas d'information (l'incertitude initiale étant nulle, l'information véhiculée par la réponse est également nulle) ;

- dans l'hypothèse de réponses équiprobables, l'incertitude *a priori* est d'autant plus grande que le nombre des réponses possibles est plus élevé ; l'information apportée par une réponse dépend ainsi du nombre des réponses distinctes possibles.

Si l'on exige en outre que les quantités d'information apportées par des événements indépendants soient additives (cf. [Bar-Hillel 1953 : 150], [Yaglom 1957 : 34]), on aboutit à la définition de la quantité d'information $H(x_i)$ apportée par la réalisation d'un événement x_i :

$$H(x_i) = \log_a \frac{1}{\text{prob}(x_i)}$$

Dans le cas d'une question présentant n réponses possibles équiprobables, on a par conséquent, pour chaque réponse r_i , la quantité d'information :

$$H(r_i) = \log_a n$$

Il est usuel de prendre comme étalon la quantité d'information apportée par la réponse à une alternative ("*choisir 1/2*") dont les deux membres avaient la même probabilité d'apparition. On choisit alors : $a = 2$, et l'unité de mesure de la quantité d'information est : $\log_2 2 = 1$ bit. HAMBLIN avait déjà, en 1958, suggéré cette démarche pour mesurer le "contenu logique" des questions [Hamblin 1958 : 167-168], c'est-à-dire pour évaluer leur *puissance* en termes de capacité de discrimination entre des éventualités dont on ignore les probabilités (et que l'on suppose équiprobables). Nous aboutissons ainsi à une mesure de la puissance H d'une consigne quelconque, de la forme :

$$H_{\text{consigne}} = \log_2 N_{\text{consigne}}$$

A partir de cette consigne, il est facile d'énoncer un certain nombre de règles, du type :

- le remplacement de : "*choisir k/n*" par : "*ordonner k/n*" augmente la puissance de la question de : $\log_2 k!$;

- le remplacement de : "*ordonner k/n*" par : "*ordonner n/n*" augmente la puissance de la question de : $\log_2 (n-k)!$;

- le remplacement de : "ordonner k/n " par : "ordonner $k+1/n$ ", si :
 $k < n-1$, augmente la puissance de la question de : $\log_2(n-k)$.

A titre d'illustration, nous indiquons ci-dessous (en ordre croissant) la puissance de diverses formes de questions, pour $n = 7$ et $k = 3$:

$$H_{\text{choisir } 3/7} = \log_2 35 = 5,129$$

$$H_{\text{choisir } \frac{3}{1}/7} = \log_2 63 = 5,977$$

$$H_{\text{choisir } x/7} = \log_2 127 = 6,989$$

$$H_{\text{ordonner } 3/7} = H_{\text{apparier}(I) 3/3,7} = \log_2 210 = 7,714$$

$$H_{\text{ordonner } \frac{3}{1}/7} = \log_2 259 = 8,017$$

$$H_{\text{apparier}(A) 3/3,7} = \log_2 343 = 8,422$$

$$H_{\text{apparier } 3/3,7} = \log_2 1330 = 10,377$$

$$H_{\text{apparier}(S) 7/7,3} = \log_2 1806 = 10,819$$

$$H_{\text{classer } 7/3} = H_{\text{apparier}(A) 7/7,3} = \log_2 2187 = 11,095$$

$$H_{\text{ordonner } 7/7} = H_{\text{apparier}(B) 7/7,7} = \log_2 5040 = 12,299$$

$$H_{\text{ordonner } \frac{7}{1}/7} = \log_2 8659 = 13,080$$

On voit, sur cet exemple numérique, que la consigne d'ordination totale de l'ensemble des items est la consigne la plus puissante, en particulier lorsqu'elle est interprétée par le répondant comme admettant des réponses incomplètes (c'est-à-dire comme demandant au moins une ordination

partielle). On remarquera que la notion de puissance d'une question est identique à celle de capacité de canal utilisée par COOMBS ; mais cet auteur n'utilise en fait que la capacité de canal maximum, qui est la somme des capacités de canal de la consigne considérée, pour toutes les combinaisons possibles des éléments d'un ensemble donné de stimuli [Coombs 1964 : 34-36]. Ce point de vue s'explique dans un contexte expérimental, où il est usuel de présenter plusieurs fois un même stimulus, associé à d'autres stimuli variables selon les présentations (méthodes des comparaisons par paires, des triades, etc.), voire de présenter plusieurs fois le même stimulus isolé, ou le même ensemble de stimuli (mesures de seuils par exemple). A notre connaissance, dans la situation d'enquête ou de test, cette situation ne se présente pratiquement jamais.

La puissance d'une question peut être diminuée par l'introduction de contraintes au niveau de l'enregistrement des réponses du sujet (cette pratique ne se rencontre guère que dans certains tests de niveau scolaire ou intellectuel). Nous en avons vu un exemple au § 3.1.1.d (question sur la Constitution des États-Unis, tirée de [Wesman 1971 : 97]), dans lequel, à la suite d'une consigne de forme : "*choisir x/4*", on présente au répondant 4 patrons de réponse seulement, sur les 15 possibles. Nous avons de même indiqué, au § 3.1.2.c, que tel était également le cas du test d'abstraction ABS. Un autre exemple intéressant est le suivant :

- "*Dans cette page on a mélangé les phrases d'un texte. Lisez d'abord toutes les phrases pour connaître l'histoire. Puis, vous allez classer les 4 phrases du début et ensuite les 6 phrases de la fin [...]. Voici la fin : [...].*

. La 1ère est :	A	B	C	D
. La 2ème est :	A	B	E	F
. La 3ème est :	B	C	E	F
. La 4ème est :	C	D	E	F
. La 5ème est :	B	C	D	E
. La 6ème est :	C	D	E	F "

(ECO 5 : SV)

L'objectif d'une telle contrainte sur les réponses possibles est évidemment de faciliter l'épreuve en réduisant l'incertitude du sujet. Il n'est cependant pas certain que cette présentation rende l'épreuve plus aisée, dans la mesure où le temps alloué au sujet (enfant en classe de cinquième) est limité à 8 minutes pour les deux classements successifs. Si par exemple l'enfant a choisi comme premières phrases de la fin du récit DBF, il ne peut poursuivre le classement (la lettre A n'apparaît plus) et doit tout recommencer ; de plus, cette réponse sera comptée comme fautive, alors que ce classement partiel est correct. Sans pousser plus avant la critique de cette présentation (et celle des procédures de correction de l'épreuve), examinons dans quelle mesure ces contraintes réduisent la puissance de la consigne initiale. Celle-ci est égale à :

$$H_{\text{ordonner 6/6}} = \log_2 720 = 9,492$$

Le calcul du nombre de réponses distinctes possibles à une épreuve dans laquelle s'exerce une contrainte sur les réponses autorisées est plus complexe qu'en l'absence de contrainte ; d'autant que, sur cet exemple, la répartition des lettres figurant les réponses possibles pour chaque rang n'est pas équilibrée. En particulier, on ne peut certainement pas considérer simplement cette question comme une série de six consignes de forme : "choisir 1/4". La solution est celle du calcul du nombre de permutations à positions interdites (cf. [Comtet 1970 : II 34-35]), fondée sur le calcul du permanent (cf. [Ryser 1963 : 27-32], [Comtet 1970 : II 29-31]) d'une matrice carrée dont les lignes désignent les rangs et les colonnes les items, et dans laquelle les 0 indiquent les positions interdites, et les 1 les positions permises (calcul du nombre de dérangements). Soit, pour notre exemple :

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

On aboutit ainsi à 64 réponses distinctes possibles, soit une puissance de : $H = \log_2 64 = 6$ bits. La puissance de la question a donc été diminuée de 3,492 bits par les contraintes imposées sur les réponses.

3.1.3. La construction des variables

Ainsi que nous l'avons dit dans notre introduction, l'analyse secondaire de données d'enquête impose au chercheur l'utilisation d'indicateurs qu'il n'a pas lui-même construits. Aussi, avant d'entreprendre la construction des variables nécessaires à sa recherche, doit-il se livrer à une analyse sémantique approfondie des indicateurs qu'il a retenus. Formellement, nous pouvons considérer qu'un indicateur (ici : une réponse) est le résultat de l'application d'une consigne sur un ensemble de modalités (de réponse) ; c'est le schéma développé par BELNAP pour les questions extensives (questions *whether*). Lorsqu'il s'agit de réponses à une question ouverte, il est certain que ce schéma est psychologiquement inadéquat (du moins en ce qui concerne le comportement de la personne interrogée ; il est adapté par contre au comportement des codeurs). Mais, formellement, les indicateurs correspondants se présentent comme le résultat de l'application d'une consigne de sélection (selon les cas : "*choisir 1/n*" ou : "*choisir k/n*") à l'ensemble des catégories d'analyse de contenu utilisées pour coder les réponses à la question. C'est pourquoi, pour définir les relations sémantiques postulées entre indicateurs, il nous suffira de définir les relations existant entre les modalités de réponses, et les relations introduites par l'application de la consigne.

La difficulté d'une telle analyse réside dans les ambiguïtés du langage, et dans l'absence d'une formalisation satisfaisante des relations sémantiques répertoriées par les linguistes. En outre, il est bien connu que la correspondance entre les déclarations des personnes interrogées et la réalité qu'elles expriment varie fortement selon le contexte dans lequel

sont faites les déclarations, et selon la personnalité du répondant. C'est sans doute pourquoi seules sont prises en compte, dans le traitement statistique des données, quelques formes de relations sémantiques aisément formalisables. L'ensemble des autres formes de relation est en revanche utilisé par le chercheur pour interpréter les résultats de l'analyse statistique ; mais la procédure d'utilisation de ces relations dans l'interprétation des réponses au questionnaire n'a pas, à notre connaissance, fait l'objet d'une formalisation opératoire.

Nous commençons donc par décrire les relations sémantiques entre modalités de réponse, qui peuvent en principe être effectivement prises en compte dans un dépouillement d'enquête. Nous rappelons ensuite comment ces relations, et celles qui sont introduites par l'application de la consigne, servent de matériaux pour la construction des variables. Enfin, nous illustrons les limites de cette procédure en évoquant quelques-uns des problèmes d'interprétation des réponses bien connus des spécialistes de l'enquête.

a) Les relations sémantiques entre les modalités de réponse

Rappelons tout d'abord que, dans le cas général, l'ensemble des modalités de réponse ne se confond pas avec l'ensemble des réponses possibles : l'ensemble des modalités de réponses est l'ensemble des matériaux à partir desquels, en appliquant la procédure exprimée par la consigne, l'interlocuteur doit élaborer sa réponse. Ce n'est par conséquent que dans le cas d'une consigne de sélection unique ("*choisir 1/n*") que ces deux ensembles sont identiques. Encore faut-il considérer que l'ensemble des réponses effectivement utilisé dans l'analyse des informations recueillies englobe non seulement l'ensemble des réponses possibles *stricto sensu* (tel qu'il peut être défini à partir des modalités de réponses proposées et de la consigne), mais également d'autres formes d'énoncés correspondant à l'aveu d'ignorance, au refus de répondre, à la négation de tout ou partie de la présupposition, etc. (ainsi que nous l'avons indiqué à la fin du chapitre 2). D'autre part, l'ensemble des réponses possibles au sens large (tel que nous venons de le

décrite) ne se confond pas non plus, dans le cas général, avec l'ensemble des précodifications. En effet, ce n'est guère que dans les questions à consigne de sélection unique qu'il y a identité entre ces deux ensembles. Pour les autres formes de consigne, la précodification comporte l'énumération des modalités proposées, plus la liste des réponses métalinguistiques possibles ; l'enregistrement de la réponse se faisant par application de l'ensemble des modalités sur l'ensemble {choisi, non choisi} pour les consignes de sélection, l'ensemble {1,2,..., k} pour les consignes d'ordination, l'ensemble des catégories proposées pour les consignes de classification, etc.

Les relations sémantiques entre les modalités de réponse proposées à la personne interrogée (dans le cas de questions fermées) ou entre les catégories d'analyse de contenu de la réponse (dans le cas de questions ouvertes ou précodées) sont en général celles qui sont admises par la majorité des locuteurs sains d'esprit et connaissant bien la langue. Dans la mesure où toute question impose une sélection (cf. p.ex. sur ce point l'analyse de la requête chez BELNAP, § 2.2.1.b), elle présuppose l'existence de différences entre certaines au moins des modalités de réponse présentées à l'interlocuteur. Ces différences entrent dans le cadre des relations sémantiques décrites par les linguistes entre lexèmes ou expressions ; en effet, dans les cas assez rares où les modalités de réponse sont des phrases complètes distinctes, la relation de sens perçue entre ces phrases peut presque toujours être réduite à une relation entre deux mots ou deux expressions.

En nous inspirant pour l'essentiel de John LYONS, nous distinguerons quatre formes principales de relations entre lexèmes (ou expressions) [Lyons 1978 : 218-257] :

- *synonymie*, caractérisée par l'interchangeabilité des lexèmes sans modification du sens de l'énoncé. On peut distinguer des degrés de synonymie, selon que le remplacement d'un lexème par un synonyme ne conserve que le sens descriptif de l'énoncé, ou qu'il préserve également son sens expressif, ses connotations. Formellement, la relation de synonymie

est une relation d'équivalence.

- *hyponymie*, correspondant à une organisation hiérarchique du vocabulaire, traduisible en termes d'inclusion de classes (ex. : "tulipe" est hyponyme de "fleur" = la classe des tulipes est incluse dans la classe des fleurs). Formellement, la relation d'hyponymie peut être assimilée à l'implication (la synonymie devient alors une hyponymie réciproque).

- *relation partie-tout*, formellement assez proche de la relation d'hyponymie, dans la mesure où elle reflète également une organisation hiérarchique du vocabulaire, mais sémantiquement différente (ex. : "bras/corps").

- *contraste*, caractérisé par une relation d'incompatibilité relativement à un contexte donné. Formellement, la relation de contraste est assez difficile à définir : en premier lieu, le contraste n'existe qu'à l'intérieur de ressemblances ("tulipe" et "cheval" sont (linguistiquement) incompatibles, mais le contraste n'a pas de sens puisque le premier mot dénote une fleur, et le second un animal) ; en second lieu, le contraste ne se réduit pas à l'incompatibilité au sens logique (alors que "lundi" et "mardi" entretiennent entre eux une relation de contraste, "Il est venu lundi" n'est pas logiquement incompatible avec : "Il est venu mardi"). Cette relation étant fondamentale pour l'analyse statistique des réponses aux questions, nous allons examiner plus en détail les principales variétés de contrastes observables en français.

LYONS distingue les contrastes binaires, qu'il nomme *oppositions*, et les contrastes non binaires. Les oppositions englobent quatre variétés de contraste binaire, correspondant à des couples de lexèmes antonymes, complémentaires, converses, contraires. Ces qualificatifs étant, dans l'usage courant, assez imprécis (ils sont même parfois, dans certains contextes, traités comme des synonymes), LYONS leur donne une définition plus spécifique :

- les *antonymes* ont entre eux une relation d'opposition privative : l'un des lexèmes dénote une propriété positive, et l'autre l'absence de cette propriété (exemples : "animé/inanimé" ; "moral/amoral" ; "bon/mauvais").

- les *complémentaires* sont liés par une relation d'opposition équilibrée : chaque lexème dénote une propriété positive, mais les deux propriétés sont logiquement contradictoires (exemple : "mâle/femelle").

- les *converses* sont des couples de noms relationnels à deux arguments, exprimant chacun une relation (ex. : "mari/femme" ; "père/fils") telle que l'affirmation de l'une pour un couple d'arguments $\langle X, Y \rangle$ équivaut à l'affirmation de l'autre pour le couple $\langle Y, X \rangle$ (ex. : "X est le mari de Y" \equiv "Y est la femme de X").

- les *contraires* entretiennent entre eux une opposition directionnelle, exprimant l'idée de mouvement dans des directions opposées (contraires *antipodaux* ; ex. : "monter/descendre" ; "devant/derrière" ; "dessus/dessous" ; "Nord/Sud") ou simplement différentes (contraires *orthogonaux* ; ex. : "Nord/Ouest").

Ces oppositions ne sont d'ailleurs pas exclusives ; et, par exemple, la relation d'opposition directionnelle antipodale peut s'appliquer, dans l'espace des concepts, à des converses (termes de parenté vus sous la représentation spatiale de l'arbre généalogique) ou à des complémentaires (ex. : "homme/femme", "homme/garçon", selon la dimension considérée).

Les *contrastes non binaires* sont des relations unissant les lexèmes dans des ensembles à plus de deux éléments. Parmi ceux-ci, nous citerons les contrastes fondés sur :

- la relation d'*incompatibilité* au sens strict (exemple : l'ensemble des noms de fleurs : si X est une pivoine, alors X n'est ni une rose, ni une tulipe, ni ..., ni une violette).

- la relation d'*ordre cyclique*, comme pour les ensembles de mots qui dénotent des périodes temporelles (jours de la semaine, mois de l'année, saisons).

- la relation d'*ordre sériel*, correspondant à un ordre strict total sur l'ensemble des lexèmes (exemples : les niveaux d'excellence ou de chaleur ; les grades ; etc.).

Ces notions recourent partiellement les distinctions que nous avons faites à propos des interrogations à choix multiples (au § 3.1.1.d) entre ensembles de modalités de réponse sans structure *a priori*, et ensembles ordonnés, ainsi que la distinction usuelle entre échelles nominales et échelles ordinales. Toutefois, avant de nous prononcer définitivement sur ces diverses relations sémantiques, il nous faut examiner plus complètement la nature des relations d'ordre entre lexèmes.

Dans ce domaine, les travaux d'Edward SAPIR nous paraissent revêtir une grande importance pour la perspective qui est la nôtre. Cet auteur aboutit en effet, sur le plan sémantique, à des propositions pratiquement équivalentes à certaines de celles auxquelles nous a conduit la recherche des informations de base utilisées dans l'analyse des données (cf. § 3.3.1.a). En particulier, SAPIR affirme que la gradation (*grading*), en tant que processus psychologique, précède la mesure et le dénombrement ; que le sens absolu de certains antonymes (comme : "chaud/froid", "grand/petit", "beaucoup/peu") est une illusion, due au fait linguistique que la gradation implicite contenue dans ces termes n'est pas marquée de façon formelle [Sapir 1944 : 93] ; et enfin, que la relation d'égalité peut être considérée comme un point de transition ou d'équilibre entre deux relations d'ordre opposées [Sapir 1944 : 105-106]. Aussi pourrions-nous constater, dans la troisième partie de ce chapitre, que le texte de SAPIR sur la gradation est un texte charnière entre l'analyse sémantique et l'analyse statistique des réponses aux questions, tout comme les conceptions de Zellig S. HARRIS et

celles de C.L. HAMBLIN, sur les relations entre questions et réponses, sont des théories charnières entre la syntaxe et la sémantique des questions. C'est pourquoi nous présentons l'essentiel (dans notre optique) des conclusions de SAPIR d'une manière relativement détaillée.

Edward SAPIR commence par faire remarquer que les jugements apparemment absolus sont en réalité soit des jugements relatifs, soit des jugements dont la signification repose d'abord sur une comparaison :

- "Les jugements de quantité en termes d'unité de mesure ou de dénombrements présupposent toujours, explicitement ou implicitement, des jugements de gradation préalables" [Sapir 1944 : 93]. Le sens du mot "quatre" se définit par référence à certains mots ("cinq", "six"), qui sont *plus* que "quatre", et à d'autres ("deux", "trois") qui sont *moins* que lui.

- "Pour ne prendre qu'un exemple, *beaucoup* n'englobe aucune classe de jugements centrés sur une quantité ou une norme données qui seraient applicables à toute forme d'expérience, dans le sens où *rouge* ou *vert* sont applicables à toute expérience pouvant impliquer la couleur ; c'est au contraire un terme purement relatif, qui perd toute signification lorsqu'il est privé de sa connotation de «plus que» et «moins que». *Beaucoup* signifie un nombre quelconque, défini ou non, qui est *plus* qu'un autre nombre pris comme point de départ" [Sapir 1944 : 93-94].

La gradation peut être plus ou moins explicite ; l'auteur distingue quatre degrés d'explicitation de la gradation [Sapir 1944 : 94-95] :

- *gradation implicite et virtuelle*. Exemples : "maison", "rouge" (il n'y a pas deux maisons exactement semblables à tous points de vue, et on peut opposer deux maisons quelconques en termes de "plus" et de "moins" sur de nombreux traits inhérents au concept de "maison" ; de même, deux "rouges" peuvent être comparés selon leur intensité, leur étendue, leur distance par rapport à une couleur étalon, etc.).

- *gradation implicite par quantification*. Exemples : "la moitié de la maison" ; "la maison toute entière" ; "trois maisons".

- *quantification par gradation implicite*. Exemples : "beaucoup de gens" ; "peu de lait" ; "une grande maison".

- *gradation explicite et quantification implicite*. Exemples : "plus de gens (que ...)" ; "une plus grande maison".

Les deux premiers degrés dans l'explicitation de la gradation présentent peu d'intérêt pratique, l'un parce que la gradation reste potentielle et n'est pas utilisée, l'autre parce que la quantification explicite est un processus déjà bien connu et bien analysé. Les deux derniers par contre méritent une analyse plus poussée.

L'étude de la gradation peut être entreprise dans une perspective logique, psychologique, ou linguistique. Chacun de ces points de vue débouche sur une classification distincte des variétés de gradation [Sapir 1944 : 95-103]. La perspective *logique* permet de distinguer deux types principaux, selon que la gradation se réfère :

- à une *norme* (exemples : "moyen", "à une distance normale" ; "pire", "moins loin" ; "meilleur", "supérieur") ;

- à des *termes de comparaison* ("relativement meilleur" ; "à une plus grande distance").

Dans chacun de ces types, la gradation peut être explicite ou implicite.

L'orientation *psychologique* conduit à une classification plus détaillée, comprenant cinq types principaux de gradation (dont chacun peut prendre la forme explicite ou implicite) :

- *série ouverte*, traduisant une dimension psychologique unique. Dans le cas de gradation explicite, les termes expriment un ordre strict total

("a meilleur que b" = "b moins bon que a") ; lorsque la gradation est implicite, ils expriment simplement une opposition sur une seule dimension ("a est près" = "b est loin").

- *séries fermées conjointes*, correspondant à deux échelles opposées avec une zone de transition ou de recouvrement. Exemple de gradation de ce type : nuances de couleurs allant du vert au jaune. Lorsque l'on parcourt l'ensemble du vert au jaune, on a, dans le cas de gradation explicite, des relations d'ordre sur le vert ("a est plus vert que b" = "b est moins vert que a"), puis des relations mixtes ("c est plus vert que d" = "d est plus jaune que c"), et enfin des relations d'ordre sur le jaune ; dans le cas de gradation implicite, on passe de différents degrés de vert à des "vert tirant sur le jaune" ou "vert jaunâtre", puis à des "jaune tirant sur le vert" ou "jaune verdâtre", et enfin à des degrés de jaune.

- *série ouverte interprétée en termes de séries fermées conjointes*, avec deux zones de transition. Lorsque la gradation est explicite, on a successivement la première zone de sériation ("a est moins loin que b" = "b est plus loin que a"), la première zone intermédiaire ("c est moins loin que d" = "d est plus loin que c" = "c est plus près que d" = "d est moins près que c"), la seconde zone de sériation ("e est plus près que f" = "f est moins près que e"), et la seconde zone de transition ("g est plus près que h" = "h est moins près que g" = "g est moins loin que h" = "h est plus loin que g"), créant ainsi un cycle. Lorsque la gradation est implicite, on a par exemple le cycle : "bon", "pas mauvais", "mauvais", "pas bon". Selon l'auteur, ce type de gradation est probablement le plus naturel pour les jugements subjectifs [Sapir 1944 : 100].

- *séries fermées disjointes*, pour lesquelles la zone intermédiaire est une zone d'indifférence, où aucun qualificatif ne s'applique. Exemple : nuances de bleu et nuances de jaune, sans recouvrement (qui correspondrait ici à des nuances de vert).

- *série ouverte interprétée en termes de séries fermées disjointes*. Exemples : gradation en termes de : "froid/chaud", ou de : "jeune/vieux", sans qualificatif pour les cas intermédiaires. L'auteur illustre ce type de

jugement par l'énoncé : "*a* est meilleur que *b*, mais tous deux sont bons ; *c* appartient à un tout autre genre, et est en fait mauvais ; quant à *d*, il n'est ni l'un ni l'autre, et ne présente par conséquent aucun intérêt".

La perspective *linguistique* enfin débouche sur une classification en deux types :

- gradation *explicite*, qui peut utiliser des séries de termes *abstraits* ("plus/moins") ou *spécialisés* ("meilleur/pire" ; "supérieur/inférieur").

- gradation *implicite*, utilisant des termes *abstraits* ("beaucoup/peu") ou *spécialisés* ("brûlant/chaud/frais/froid/glacial").

Oswald DUCROT a analysé, sous le nom d'*échelles argumentatives*, certaines formes linguistiques de gradation en français, et en particulier les relations de gradation entre des locutions ou des adverbes tels que : "même", "presque", "à peine", "aussi", "très", "peu", "pas", "pas du tout", "plus que", "beaucoup plus que", "aussi", etc. [Ducrot 1973 : 225-260]. Il a en outre montré la possibilité d'établir des relations entre la gradation et l'implication [Ducrot 1973 : 260-283], relations dont nous parlerons au paragraphe suivant.

La notion de "terme gradable" est liée à la notion de mouvement ou de changement. Par exemple, DUCROT fait remarquer que l'énoncé : "Je suis encore fatigué de mon voyage" pose : "Je suis fatigué" et présuppose : "Cette fatigue va ultérieurement disparaître". De même : "Le tonneau est encore à moitié vide" introduit l'idée que l'on est en train de remplir le tonneau, alors que : "Le tonneau est encore à moitié plein" introduit l'idée que l'on est en train de le vider [Ducrot 1973 : 252-253]. C'est pourquoi, à partir de la constatation que la gradation s'accompagne souvent d'impressions kinesthésiques, Edward SAPIR propose de "définir «égal à» [...] comme un point de passage ou d'équilibre plus ou moins temporaire entre «plus que» et «moins que» , ou comme un point d'arrivée dans une échelle sur laquelle le terme

objet de la gradation est en augmentation ou en diminution constante" [Sapir 1944 : 105]. On aboutit ainsi à cinq types d'égalité, selon qu'il y a une impression kinesthésique explicite ou implicite, correspondant à un mouvement croissant ou décroissant, ou qu'il n'y a pas d'impression kinesthésique. En raisonnant de la même manière pour les jugements de supériorité et d'infériorité, on obtient les quinze types fondamentaux de jugements de gradation figurant dans le tableau n° 3.5.

L'ensemble de ces analyses nous permet de décrire assez complètement les relations existant entre les modalités de réponses correspondant à une question. Si nous passons en revue les exemples d'énoncés de questions cités au début de ce chapitre, nous constatons que :

- la relation de *synonymie* n'apparaît jamais entre les modalités présentées : lorsqu'une certaine synonymie est perceptible (et ce peut être l'objet de la question que de proposer de l'identifier), elle n'est jamais totale, et les modalités apparaissent d'abord comme toutes distinctes les unes des autres.

- il peut arriver qu'une relation d'*implication* soit présentée explicitement entre certaines modalités (exemples : questions sur la cause ou la conséquence, tirée de [Remmers 1955 : 94], au § 3.1.1.d ; modalité de réponse : "pour ces deux raisons à la fois", *IFOP 58-12* : 22, au § 3.1.2.a) ; mais cette forme de relation n'apparaît que très rarement.

- dans les alternatives, on rencontre surtout la relation logique de *contradiction* ("oui/non" ; "d'accord/pas d'accord"), et la relation linguistique d'*antonymie* ("plaire/déplaire" ; "neuve/d'occasion" ; "proche/éloigné" ; "pour/contre" ; "bien/mal"). Mais il arrive aussi que les termes en jeu soient des complémentaires ("homme/femme") ou des contraires antipodaux ("augmenter/diminuer"). On observe également parfois des couples de termes extraits d'ensembles plus grands, munis de la relation linguistique d'incompatibilité

Types de jugement	Moins	Égal	Plus
<u>Dynamique explicite</u> <i>Croissant</i> : "Quelle distance a-t-il déjà parcourue ?" <i>Décroissant</i> : "De combien de temps dispose-t-il encore pour terminer ?"	$\rightarrow q$: "Il n'a pas encore parcouru 5 kms" $\leftarrow q$: "Il lui reste moins de 5 heures"	$\rightarrow q$: "Il a parcouru juste 5 kms" $\leftarrow q$: "Il lui reste exactement 5 heures"	$q \rightarrow$: "Il a parcouru plus de 5 kms" $q \leftarrow$: "Il lui reste encore plus de 5 heures"
<u>Dynamique implicite</u> <i>Croissant</i> : "Combien avait-il parcouru lorsqu'il s'est arrêté ?" <i>Décroissant</i> : "Combien pouvait-il encore soulever lorsqu'il a dû abandonner ?"	$\uparrow q$: "Il s'est arrêté à moins de 5 kms de son point de départ" $\uparrow q$: "Il s'était affaibli au point de ne même plus pouvoir soulever 5 kgs"	$\uparrow q$: "Il s'est arrêté exactement à 5 kms de son point de départ" $\uparrow q$: "Il s'était affaibli au point de ne plus plus pouvoir soulever que 5 kgs"	$q \uparrow$: "Il s'est arrêté à plus de 5 kms de son point de départ" $q \uparrow$: "Il s'était affaibli au point d'avoir du mal à soulever plus de 5 kgs"
<u>Non dynamique</u> "A quelle distance de la mer se trouve la ville de N ?"	$a < q$: "N est à moins de 5 kms de la mer"	$a = q$: "N est juste à 5 kms de la mer"	$a > q$: "N est à plus de 5 kms de la mer"

Tableau n° 3.5 : Classification des formes de base de jugements de gradation, selon [Sapir 1944 : 106-108].

("faute du gouvernement/difficultés normales") ou de la relation d'ordre ("moins souvent/aussi souvent" ; "autrefois/maintenant" ; "beaucoup d'occasions/choix limité").

- dans les questions à choix multiples, les trois types de relation les plus fréquents sont l'*incompatibilité* (énumération d'items variés, sans autre relation sémantique entre eux ; cf. IFOP 58-12 : 7 , SOFRES 967 : 13 , SOFRES 686 : 13 , etc. au § 3.1.1.d), et la gradation en *série ouverte implicite* ("souvent/quelquefois/rarement/jamais"), ou en *séries fermées disjointes implicites* ("toujours du super/principalement du super/principalement de l'ordinaire/toujours de l'ordinaire"). Il arrive toutefois que l'on rencontre d'autres formes de gradation, comme la série ouverte implicite interprétée en termes de séries fermées disjointes ("plaît beaucoup/plaît moyennement/indifférent/déplaît") ou les séries fermées conjointes implicites ("échelle en 7 degrés "gros fumeur/petit fumeur", SOFRES 1061 : 17a).

- en règle générale, toutes les relations de *gradation* entre modalités de réponses sont *implicites* (nous allons voir que les relations de gradation explicites observables dans les réponses résultent de l'application des consignes d'ordination). En outre, les sociologues qui ont une certaine expérience de l'enquête emploient assez rarement les séries fermées conjointes, ou les séries ouvertes interprétées en termes de séries fermées conjointes. Ils leur préfèrent soit les séries ouvertes, soit les séries fermées disjointes ; ils évitent ainsi l'accumulation des réponses dans la zone de transition ("oui et non", "les deux à la fois"), considérée comme apportant peu d'informations utilisables.

b) Des modalités de réponse aux variables

La construction des variables, à partir des modalités de réponse correspondant à une ou plusieurs questions, passe par plusieurs étapes. Au départ, on dispose pour chaque question de l'ensemble des modalités, muni d'une structure sémantique ; d'une consigne dont l'application

modifiera les relations sémantiques entre certaines des modalités ; et d'un ensemble de réponses métalinguistiques correspondant au refus de l'interviewé d'appliquer la consigne. L'application systématique de la consigne sur les modalités engendre l'ensemble des réponses linguistiques possibles à la question ; les propriétés formelles de la consigne et la nature des relations sémantiques entre les modalités déterminent des relations sémantiques entre les réponses possibles. Si l'on ajoute, à l'ensemble de ces réponses possibles au sens strict, l'ensemble des réponses métalinguistiques, on obtient l'ensemble des réponses possibles au sens large, au sein duquel sont définies un certain nombre de relations sémantiques, dont celles qui existent dans chacun des deux ensembles initiaux. En principe, chacune de ces réponses possibles au sens large est un indicateur ; et les relations sémantiques entre indicateurs correspondant à une même question doivent être considérées comme données lorsque l'on entreprend une analyse secondaire. Ce sont donc ces relations qu'il nous importe surtout d'analyser. A ces relations, qui sont en quelque sorte imposées au chercheur, il est toujours possible d'en ajouter d'autres, soit *a priori* à partir de considérations théoriques, soit *a posteriori*, en fonction de certains résultats numériques. Ces opérations de modification des relations entre réponses sont celles de la construction des variables. Comme l'analyse des données porte, en fait, sur les variables, et non directement sur les réponses aux questions, la construction des indicateurs et des variables est un chaînon important entre le questionnaire et l'analyse statistique des informations qu'il a permis de recueillir.

Dans l'esprit de l'auteur du questionnaire, l'ensemble des modalités doit apparaître à la personne interrogée comme muni d'une relation binaire totale, qui peut être :

- soit une relation de *non-équivalence* (relation antiréflexive, symétrique, intransitive), signifiant que toutes les modalités sont considérées *a priori* comme distinctes les unes des autres (c'est ce que nous avons appelé, au § 3.1.1.d, absence apparente de structure *a priori*) ;

- soit une relation d'*ordre strict* (relation antiréflexive, antisymétrique, transitive), correspondant à la gradation de SAPIR ou à l'échelle argumentative de DUCROT.

Dans le second cas, s'il ne s'agit pas d'une série ouverte, mais de séries fermées (disjointes ou conjointes), correspondant à ce que Stanley L. PAYNE appelle *fold-over technique* [Payne 1951 : 97], il y a apparemment deux sous-ensembles disjoints de modalités de réponse munis chacun d'une relation d'ordre strict totale (par exemple : "très favorable/.../plutôt favorable" et : "plutôt défavorable/.../très défavorable"). On fait généralement l'hypothèse que l'ensemble des modalités est perçu par le répondant comme totalement ordonné selon une dimension unique (ici : de "favorabilité"), l'un des sous-ensembles de termes ou d'expressions ayant une coloration négative.

L'application de la consigne sur l'ensemble des modalités, pour engendrer l'une des réponses possibles (au sens strict), a pour effet de modifier la structure postulée sur l'ensemble des modalités. Lorsque celles-ci sont simplement considérées *a priori* comme non équivalentes, l'application de la consigne se traduit par le remplacement de la totalité ou de la plus grande partie des relations de non-équivalence par des relations d'équivalence, ou des relations d'ordre strict. Plus spécifiquement, si l'on considère les $\frac{n(n-1)}{2}$ paires de modalités distinctes :

- une consigne de sélection quelconque, de forme : "*choisir k/n*", engendre $\frac{k(k-1)}{2}$ relations d'équivalence entre les k modalités sélectionnées, $\frac{(n-k)(n-k-1)}{2}$ relations d'équivalence entre les $(n-k)$ modalités non sélectionnées, et $k(n-k)$ relations d'ordre strict traduisant la préférence du répondant pour les modalités sélectionnées.

- une consigne d'ordination totale, de forme : "*ordonner n/n*" ou : "*ordonner n-1/n*" détermine $\frac{n(n-1)}{2}$ relations d'ordre strict se substituant aux relations de non-équivalence.

- une consigne d'ordination partielle, de forme : "*ordonner k/n*" avec $k < n-1$, détermine $\frac{k(k-1)}{2}$ relations d'ordre strict entre les k modalités sélectionnées, $\frac{(n-k)(n-k-1)}{2}$ relations d'équivalence entre les $(n-k)$ modalités non sélectionnées, et $k(n-k)$ relations d'ordre strict entre les k modalités sélectionnées et les $(n-k)$ modalités non sélectionnées.

- une consigne de classification quelconque, de forme : "*classer n/p : e₁, ..., e_p*" engendre $\frac{e_i(e_i-1)}{2}$ relations d'équivalence entre les e_i modalités rangées dans chacune des p classes i . Pour chaque paire de modalités n'appartenant pas à la même classe, si les critères de classification sont simplement non équivalents, les relations de non-équivalence entre modalités sont conservées ; si les critères de classification sont totalement ordonnés (ce qui correspond au cas le plus fréquent), elles sont remplacées par des relations d'ordre strict. En règle générale, si deux critères de classification X et Y sont dans la relation : $X R Y$, et si une modalité a est associée à X et une modalité b à Y , alors a et b se trouvent placés par l'opération de classification dans la relation $a R b$.

Lorsque les modalités de réponse sont ordonnées *a priori*, la consigne est pratiquement toujours de la forme : "*choisir 1/n*" (comme nous l'avons noté au § 3.1.1.d). Elle a donc pour effet de sélectionner une modalité unique, considérée d'un certain point de vue, par le répondant, comme "meilleure" que toutes les autres (correspondant le mieux à sa situation, son opinion, son goût, etc.). La relation d'ordre partiel introduite par la sélection unique (équivalente à : "*ordonner 1/n*") se substitue donc, d'une certaine manière, à la relation d'ordre total initiale. Il importe toutefois de ne pas confondre ces deux relations, dont les valeurs sémantiques sont évidemment très différentes. La particularité de l'effet d'une consigne de sélection unique sur un ensemble de modalités totalement ordonnées est de sélectionner, en fait, plus d'une modalité de réponse. Pour comprendre ce phénomène, il faut revenir

à l'analyse que les linguistes font de la gradation. "On peut dire que les jugements de «plus que» et «moins que» sont fondés sur des perceptions d'«enveloppement». Si A peut être «enveloppé par» B, soit réellement soit en imagination, de sorte qu'il paraisse enfermé dans les limites de celui-ci au lieu de les déborder, il est estimé être «moins que» B, tandis que B est estimé être «plus que» A" [Sapir 1944 : 93]. Cette représentation en termes d'inclusion de classes est parallèle à la définition de l'ordre en théorie des ensembles (cf. p.ex. [Halmos 1960 : 31-35]), et justifie les remarques de DUCROT sur les relations d'implication dans les échelles argumentatives (cf. [Ducrot 1973 : 260-283]). En d'autres termes, dans un ensemble de modalités totalement ordonné, choisir l'une de ces modalités c'est sélectionner également toutes celles qui lui sont "inférieures" (toutes celles qu'elle contient, selon la perspective de SAPIR, ou toutes celles qui sont également vraies *ipso facto*, dans l'optique de DUCROT). C'est pourquoi nous considérerons que sélectionner, dans un ensemble ordonné de n modalités, la modalité de rang r , c'est en fait sélectionner les $n - r + 1$ modalités correspondant à ce rang (la modalité explicitement sélectionnée, et toutes celles qui lui sont "inférieures").

Pour construire un indicateur à partir des réponses à une question, il faut définir les relations sémantiques sur l'ensemble des réponses possibles. En ne considérant tout d'abord que l'ensemble des réponses possibles au sens strict, nous examinerons trois cas particuliers qui recouvrent la quasi-totalité des cas observables couramment :

① les modalités sont seulement non équivalentes, et la consigne n'a pas de paramètre indéterminé (le nombre k de modalités à sélectionner ou à ordonner, ou bien le nombre p de classes et leurs effectifs e_i sont fixés). Dans ces conditions, sans apport d'informations supplémentaires, on doit considérer toutes les réponses comme ayant le même statut ; il est impossible de les ordonner d'une manière significative (l'ordre lexicographique n'a pas de signification psychologique ou sociologique), et la seule affirmation que l'on puisse faire à leur sujet est qu'elles sont distinctes. La seule relation que l'on puisse définir sur l'ensemble des réponses

possibles est donc la relation de non équivalence.

② les modalités sont seulement non équivalentes, et la consigne a au moins un paramètre indéterminé (autre que n : le nombre k de modalités à sélectionner ou à ordonner varie de v à u ; les effectifs des classes de la partition demandée ne sont pas fixés). Il est possible alors de définir sur certains couples de réponses une relation d'ordre strict, lorsqu'une réponse en "contient" ou en implique une autre. Pour reprendre un exemple classique, si une armoire mesure au moins 1,60 m, il est également vrai qu'elle mesure au moins 1,55 m, 1,50 m, etc. La réponse : "L'armoire mesure (au moins) un mètre soixante" contient ou implique d'autres réponses, telles que : "L'armoire mesure au moins un mètre cinquante". Les relations binaires entre réponses distinctes qui ne sont pas des relations d'ordre sont des relations de non-équivalence.

③ les modalités sont totalement ordonnées, et la consigne est de la forme : "*choisir* 1/ n ". L'ensemble des n réponses possibles au sens strict est totalement ordonné.

D'autre part, pour obtenir l'ensemble des réponses possibles au sens large, il faut ajouter, aux réponses engendrées par l'application de la consigne, l'ensemble des réponses méta-linguistiques (qui, pour la plupart des logiciens, ne sont pas des réponses au sens strict). On peut en principe considérer que ces réponses sont simplement non équivalentes entre elles, ainsi qu'avec chacune des réponses possibles au sens strict. L'union de ces deux ensembles engendre l'ensemble des réponses possibles au sens large, entre lesquelles sont définies des relations sémantiques appartenant à l'un des trois types que nous avons mentionnés : équivalence, non-équivalence, et ordre strict. Pour tout couple de réponses possibles au sens large, une et une seule de ces relations est toujours définie. Les principales propriétés de ces relations sémantiques sont résumées dans le tableau n° 3.6. Nous verrons toutefois, dans la suite de ce paragraphe, que des relations plus spécifiques pourront, dans certains cas, être postulées entre les réponses ; mais ces relations n'apparaîtront que comme des variétés de ces trois types de base.

$$R_i = \{a, b, c, \dots, r, \dots, t\}$$

= ensemble des réponses possibles au sens large à une question i .

$\forall a \in R_i, \forall b \in R_i, \forall c \in R_i$, on a :

- | | | |
|-----|---|--|
| [1] | $a \equiv a$ | |
| [2] | $a \equiv b \Leftrightarrow b \equiv a$ | |
| [3] | $a \neq b \Leftrightarrow b \neq a$ | |
| [4] | $a > b \Leftrightarrow b < a$ | |
| [5] | $a \equiv b \ \& \ b \equiv c \Rightarrow a \equiv c$ | |
| [6] | $a > b \ \& \ b > c \Rightarrow a > c$ | |
| [7] | $a > b \ \& \ b \equiv c \Rightarrow a > c$ | |
| [8] | $a \equiv b \ \& \ b > c \Rightarrow a > c$ | |
| [9] | $a \neq b \ \& \ b \equiv c \Rightarrow a \neq c$ | |

Tableau n° 3.6. Propriétés des relations sémantiques définies dans l'ensemble des réponses possibles (au sens large) à une question.

Deux exemples simples illustreront la procédure de passage des modalités de réponse aux indicateurs. Le premier est une consigne de sélection de forme : "choisir 2/4", appliquée à un ensemble de modalités non ordonnées :

- "[...] Parmi les secteurs suivants, quels sont les deux qui doivent être modernisés en premier ?"
- . Industrie
 - . Commerce de détail
 - . L'Agriculture
 - . Le travail administratif" (IFOP 69-5 : 6)

Les quatre modalités présentées : $M = \{I, C, A, T\}$ sont non équivalentes les unes par rapport aux autres. Si la consigne est appliquée strictement, il y a 6 réponses distinctes possibles : $\{I, C\}, \{I, A\}, \{I, T\}, \{C, A\}, \{C, T\}, \{A, T\}$,

toutes également non équivalentes les unes aux autres. La codification prévoyant la réponse métalinguistique : "Je ne sais pas", l'ensemble des réponses possibles au sens large comporte 7 éléments, tous non équivalents les uns avec les autres (cf. Figure n° 3.7). Si la consigne est interprétée comme équivalant à : "*choisir* $\frac{2}{1/4}$ ", il y a 10 réponses distinctes possibles, ordonnables partiellement selon la relation d'inclusion ou d'implication (cf. Figure n° 3.8).

Le second exemple est une consigne de sélection unique appliquée à un ensemble de modalités ordonnées :

- "*Les conditions de reprise qui vous ont été faites étaient-elles, selon vous ...*"
 - . *les mêmes qu'ailleurs*
 - . *moins bonnes*
 - . *ou meilleures qu'ailleurs ?* (SOFRES 967 : 14)

Les trois réponses possibles sont ordonnées strictement, comme les modalités auxquelles elles correspondent (puisque, quand $k=1$, l'ensemble des réponses possibles et l'ensemble des modalités sont identiques) : "meilleures" > "les mêmes" > "moins bonnes". La codification prévoit la réponse métalinguistique : "Ne sait pas, ça dépend", non équivalente à chacune des réponses possibles au sens strict. On obtient donc un ensemble de 4 réponses possibles au sens large, dont 3 sont strictement ordonnées (cf. Figure n° 3.9).

Les relations sémantiques que nous venons de décrire nous paraissent être (en l'état actuel du développement de la linguistique) les seules que l'on puisse postuler entre les réponses possibles (au sens large) à une question. Les informations qu'elles apportent sont donc en principe les seules informations qui soient en quelque sorte imposées au chercheur qui entreprend une analyse secondaire. La construction de variables à partir de

	{I,C}	{I,A}	{I,T}	{C,A}	{C,T}	{A,T}	<i>NSP</i>
{I,C}	≡	≠	≠	≠	≠	≠	≠
{I,A}	≠	≡	≠	≠	≠	≠	≠
{I,T}	≠	≠	≡	≠	≠	≠	≠
{C,A}	≠	≠	≠	≡	≠	≠	≠
{C,T}	≠	≠	≠	≠	≡	≠	≠
{A,T}	≠	≠	≠	≠	≠	≡	≠
<i>NSP</i>	≠	≠	≠	≠	≠	≠	≡

Figure n° 3.7. Exemple de relations binaires dans l'ensemble des réponses possibles au sens large à une question de sélection de forme : "choisir 2/4"

	{I}	{C}	{A}	{T}	{I,C}	{I,A}	{I,T}	{C,A}	{C,T}	{A,T}	<i>NSP</i>
{I}	≡	≠	≠	≠	<	<	<	≠	≠	≠	≠
{C}	≠	≡	≠	≠	<	≠	≠	<	<	≠	≠
{A}	≠	≠	≡	≠	≠	<	≠	<	≠	<	≠
{T}	≠	≠	≠	≡	≠	≠	<	≠	<	<	≠
{I,C}	>	>	≠	≠	≡	≠	≠	≠	≠	≠	≠
{I,A}	>	≠	>	≠	≠	≡	≠	≠	≠	≠	≠
{I,T}	>	≠	≠	>	≠	≠	≡	≠	≠	≠	≠
{C,A}	≠	>	>	≠	≠	≠	≠	≡	≠	≠	≠
{C,T}	≠	>	≠	>	≠	≠	≠	≠	≡	≠	≠
{A,T}	≠	≠	>	>	≠	≠	≠	≠	≠	≡	≠
<i>NSP</i>	≠	≠	≠	≠	≠	≠	≠	≠	≠	≠	≡

Figure n° 3.8. Exemple de relations binaires dans l'ensemble des réponses possibles au sens large à une question de sélection de forme : "choisir $\frac{2}{1} / 4$ "

	"mêmes"	"moins bonnes"	"meilleures"	"NSP, ça dépend"
"mêmes"	≡	>	<	≠
"moins bonnes"	<	≡	<	≠
"meilleures"	>	>	≡	≠
"NSP, ça dépend"	≠	≠	≠	≠

Figure n° 3.9 Exemple de relations binaires dans l'ensemble des réponses possibles au sens large à une question de sélection unique appliquée à des modalités totalement ordonnées.

ces indicateurs fait intervenir le "savoir implicite" du chercheur, c'est-à-dire l'ensemble de ses connaissances, de ses habitudes de pensée, de ses savoir-faire [Grémy 1976 : 7] ; ce "savoir implicite", qui est par définition extérieur aux données, détermine les informations supplémentaires que le chercheur va introduire dans les données qu'il aura recueillies, pour les interpréter et les synthétiser. La construction des variables implique par conséquent une interprétation des indicateurs par apport d'informations étrangères, c'est-à-dire une modification des relations sémantiques entre indicateurs qui étaient contenues en germe dans l'énoncé de la question.

Le "savoir implicite" du sociologue ou du psychologue interprétant des données recueillies à l'aide d'un questionnaire peut être considéré comme formé de deux sortes d'informations : des informations d'ordre technique concernant la situation d'enquête, le mécanisme de la réponse aux questions, l'incidence de la formulation de la question sur le comportement du répondant, etc. ; et des informations d'ordre théorique, relatives à l'objet de

la recherche lui-même. Comme exemples d'informations d'ordre technique, citons la tendance à répondre affirmativement aux alternatives, l'incidence sur les réponses de la "désirabilité sociale" (*social desirability* ; *soziale Wünschbarkeit*), et les résultats des recherches expérimentales sur la formulation des questions ; nous reparlerons de cet aspect du "savoir implicite" du spécialiste du questionnaire au paragraphe suivant. Les informations d'ordre théorique comprennent aussi bien l'ensemble des résultats de recherches antérieures que l'on peut considérer comme assurés, que des hypothèses ou des conjectures sur la "nature des choses", ou sur certains mécanismes psychologiques ou sociaux non encore élucidés. Ces informations supplémentaires sont étroitement liées à la problématique propre à chaque chercheur.

Les opérations de construction de variables constituent l'unique procédure de passage du questionnaire théorique, exprimant les interrogations du chercheur, au questionnaire empirique qui a servi à recueillir les données. Ainsi que nous l'avons signalé dans notre introduction, le questionnaire théorique peut prendre de nombreuses formes. Il se concrétise notamment dans les plans de dépouillement des données successifs, chaque plan de dépouillement étant élaboré à partir des informations apportées par le dépouillement précédent. Cela nous conduit à distinguer la construction de variables *a priori*, antérieurement à tout dépouillement, et la construction *a posteriori*, qui tient compte d'informations apportées par le dénombrement des réponses aux questionnaires. A ce second type appartiennent en particulier toutes les techniques ayant pour but la recherche de *dimensions* ou celle de *facteurs*. En effet, le critère de dimensionnalité est l'absence ou la rareté constatée, dans les données, de certaines configurations de réponses ; de même, la condition d'existence de facteurs (au sens de l'analyse factorielle) est l'existence de corrélations élevées entre certains indicateurs ou ensembles d'indicateurs, c'est-à-dire, en dernière analyse, l'absence ou la rareté de certaines configurations de réponses. Nous renvoyons, pour l'analyse de ces problèmes, à [Coombs 1964] et à [Flament 1976]. La construction de variables *a priori* par contre est en général peut traitée, ou présentée comme "naturelle". C'est pourquoi il nous paraît nécessaire d'en rappeler brièvement les particularités.

La procédure de construction de variables *a priori* peut être décomposée en deux familles de techniques distinctes : la première a pour but de condenser l'ensemble des réponses possibles à une question afin de réduire le nombre d'indicateurs ; la seconde vise à combiner en une seule variable plusieurs indicateurs correspondant à des questions différentes. En pratique, les techniques appartenant à la première famille ont pour effet de réduire le nombre de relations de non équivalence. Il est usuel par exemple (sans être "naturel" pour autant), dans une question de sélection multiple ($k > 1$), de considérer comme appartenant à une même classe toutes les réponses ayant sélectionné une même modalité ; dans une question d'ordination multiple ($k > 1$), de ranger dans une même classe les réponses citant une même modalité au premier ou au second rang ; etc.. C'est en effet très souvent sous cette forme que sont dénombrées les réponses aux questions à réponses multiples (nombre total de citations pour chaque modalité considérée isolément), et aux questions d'ordination (nombre de citations en premier ou en second pour chaque modalité, à l'exclusion de toute autre information). Devant de telles pratiques, il y aurait d'ailleurs lieu de s'interroger sur l'utilité d'introduire dans les questionnaires des questions très puissantes (disons, pour fixer les idées, d'une puissance supérieure à $= \log_2 20 = 4,322$ bits), pour n'extraire ensuite qu'une faible part de l'information qu'elles ont permis de recueillir. Dans les questions de connaissance (pour lesquelles il existe une ou plusieurs "bonnes réponses"), il est par contre légitime de définir, à ce stade de l'analyse, des relations d'équivalence entre des réponses également "bonnes", également "mauvaises", ou également "passables", ainsi que des relations d'ordre entre les classes d'équivalence mutuellement exclusives ainsi définies ; mais il importe d'être conscient qu'il s'agit là de la conséquence de l'introduction d'informations supplémentaires, ignorées en principe des répondants, et non contenues (même implicitement) dans les modalités présentées. Parmi les techniques appartenant à la seconde famille, nous ne citerons pour mémoire que les plus connues : construction d'indices de types variés, et élaboration de typologie par "réduction de l'espace d'attribut". Ces techniques sont également utilisées *a posteriori*.

A notre connaissance, l'analyse de la construction de variables *a priori* n'a jamais été très clairement distinguée de la construction de variables à partir du dénombrement des réponses aux questions (indicateurs). Nous verrons cependant, dans la troisième partie de ce chapitre, que ces démarches sont de natures radicalement différentes, et qu'il convient de traiter d'une manière distincte les informations que chacune d'elles apporte au chercheur.

c) Problèmes d'interprétation des réponses.

Les relations sémantiques prises en compte dans la formalisation des indicateurs et la construction des variables sont évidemment loin d'épuiser l'ensemble des informations véhiculées par les réponses ; ce sont toutefois les seules relations sémantiques relativement stables. Les professionnels de l'enquête sociologique et les spécialistes des épreuves psychologiques savent bien que la signification d'une réponse peut varier selon le contexte socio-culturel du répondant, la nature des interactions en jeu dans la situation d'entretien, et la forme linguistique de l'énoncé à valeur de question. L'objet de ce paragraphe est de rappeler brièvement les termes dans lesquels se posent les principaux problèmes d'interprétation des réponses.

Les relations entre le langage et la représentation de l'univers ont été souvent décrites par les anthropologues (cf. p.ex. [Whorf 1956], [Lévi-Strauss 1962]). Ces différences interculturelles entraînent évidemment des difficultés de traduction d'une langue dans une autre (cf. [Mounin 1963]), et la nécessité pour l'enquêteur à l'étranger d'adapter non seulement son questionnaire, mais également sa stratégie (cf. p.ex. : "*Interviewing Frenchmen*" [Lerner 1956]). Un exemple particulièrement frappant de l'incidence sur le questionnement des différences de langage (et donc de forme de pensée) est celui des questions de causalité. Le mot français : "pourquoi ?"

(anglais : "why ?") n'a pas d'équivalent strict dans certaines langues ; par exemple, en russe, on distingue la recherche de la cause ("почему?") de celle du but poursuivi ("зачем?") ; en grec ancien, il était semble-t-il usuel de distinguer l'intention ("εἰς τί ;"), l'objectif visé ("ἵνα τί ;"), la cause émotionnelle ("τί παθών ;") ou rationnelle ("τίνι λόγῳ ;"), etc. Il est clair que, selon la langue, le degré de précision dans la question ne pourra être identique (Remarquons incidemment que, pour un sociologue américain, l'exigence de précision et de réalisme dans la question paraît être un trait typiquement français [Lerner 1956 : 191]). Ces différences inter-culturelles se retrouvent entre les sous-cultures d'une même civilisation, tant sur le plan proprement linguistique que sur celui de l'acceptation de l'entretien d'enquête. Ainsi, les Américains de bas statut socio-économique et culturel semblent accepter plus fréquemment que les Américains de statut élevé d'être interrogés sur des problèmes politiques [Sjoberg 1954 : 425] ; le contenu des réponses est d'autant plus précis, riche, et structuré que le niveau socio-économique et culturel est plus élevé : "Vraisemblablement, [l'Américain de bas statut social] n'a l'habitude de parler de tels sujets, et d'une manière aussi détaillée, qu'à des interlocuteurs avec lesquels il partage une quantité importante d'expériences et de symboles, de sorte qu'il n'a pas besoin de faire très attention aux techniques de communication" [Schatzman 1955 : 337].

L'incidence des paramètres de la situation d'entretien sur les réponses a été souvent relevée par les psychologues. Alfred BINET avait déjà fait remarquer que les réponses données aux tests par des conscrits étaient moins bonnes lorsque l'examen avait lieu en présence d'un officier. Depuis, de nombreuses études expérimentales ont été entreprises afin d'évaluer les biais introduits par la situation d'enquête sur les réponses aux tests de niveau, aux épreuves projectives, aux questionnaires d'enquête. Nous avons déjà, dans notre introduction, mentionné certaines de ces recherches appliquées au test de RORSCHACH. A titre d'autres exemples, FRANZEN et WILLIAMS

ont évalué la part de la variance dans les réponses imputable aux différences entre enquêteurs ; comme cette part est plus élevée pour les questions de formulation imprécise, ou portant sur un futur hypothétique, il semblerait que la principale source de variations soit l'énoncé de la question lui-même [Franzen 1956]. D'autres recherches ont décelé l'existence d'influence des différences de statut social entre l'enquêteur et la personne interrogée, se traduisant par la tendance à l'acquiescement, et la tendance à donner des réponses "socialement désirables" (cf. p.ex. [Lenski 1960], [Phillips 1970], [Carr 1971]). Cependant, après avoir passé en revue les résultats d'un grand nombre de recherches de ce type, SUDMAN et BRADBURN concluent : "Bien que les travaux exposés dans ce chapitre aient présenté de nombreux exemples d'effets sur les réponses dus aux caractéristiques respectives de l'enquêteur et de l'enquêté, on est en fin de compte frappé par la minceur des résultats", l'essentiel de ceux-ci montrant que l'influence ne s'exerce que sur les questions dont le contenu est lié aux caractéristiques (sexe, ethnie, etc.) selon lesquelles l'enquêteur et l'enquêté se différencient (ou même s'opposent) [Sudman 1974 : 137].

Toutefois, les mécanismes de la situation d'entretien, et leurs effets sur le contenu des réponses, ne sont guère connus que d'une manière empirique. Diverses recherches théoriques ont pourtant été entreprises dans ce domaine. A titre d'exemples, mentionnons brièvement trois directions en cours d'exploration actuellement : la théorisation du processus de questionnement et de ses conséquences sur les réponses, la formalisation du dialogue dans l'optique de l'analyse des systèmes, et l'analyse linguistique de la conversation. Dans la première direction, J.W. GETZELS a proposé un modèle du processus question-réponse en trois étapes : ① réaction interne à la question (identique quel que soit le contexte) ; ② évaluation des exigences liées au contexte de l'entretien (s'affirmer, se présenter sous un jour favorable, etc.) ; ③ expression d'une réponse adaptée au contexte [Getzels 1954]. Le modèle de GETZELS permet de prévoir la direction du biais dans la réponse introduit par l'interaction enquêteur-enquêté. La "théorie de la question", de Kurt HOLM, est une tentative plus ambitieuse (inspirée par l'analyse factorielle) pour décomposer le contenu d'une question en une somme pondérée de

trois contenus distincts : le premier correspondant à la dimension visée, le second aux dimensions étrangères aux objectifs de la recherche, et le troisième résultant de la "désirabilité sociale" (se traduisant en particulier par la tendance à répondre affirmativement) [Holm 1974a]. L'intérêt du modèle de HOLM est qu'il débouche sur un ensemble de règles pratiques concernant la formulation des questions et l'analyse de leurs réponses, et est prolongé par une "théorie de la batterie de questions" [Holm 1974b] ; mais il présente l'inconvénient pratique de ne s'appliquer effectivement qu'aux échelles numériques.

La formalisation du dialogue dans une perspective systémique commence, à notre connaissance, avec l'élaboration de programmes interactifs de réponses aux questions. Par exemple, le programme *BASEBALL* a été conçu pour analyser diverses formes d'interrogation concernant des rencontres de baseball, et pour fournir les informations requises par la question [Green 1963]. Cette forme de recherche peut faire intervenir des opérations d'analyse linguistique relativement complexe (cf. p.ex. [Lindsay 1963]). Bien que leur champ d'application soit apparemment la documentation automatique, les recherches de ce type sont essentiellement des travaux sur l'intelligence artificielle. Par contre, la théorie des communications de David HARRAH est explicitement la modélisation du dialogue, dont l'interview est un cas particulier. Dans ce modèle, un émetteur S (*sender*) et un récepteur R (*receiver*) disposent d'un langage formalisé commun L , dont la logique des questions est un sous-ensemble (cf. § 2.3.4.d). Le récepteur R est caractérisé par quatre paramètres : le langage L , l'ensemble k de ses connaissances (exprimées par des propositions de L), l'ensemble Q des questions qu'il se pose (propositions de L exprimant ses centres d'intérêt, à rapprocher de ce que nous avons appelé *questionnaire théorique*), et une fonction I d'évaluation de la valeur des messages reçus [Harrah 1963 : 6-10, 45, 54-60]. Un événement E est décrit par un message m (exprimé dans le langage L) destiné à R au moment t , et un indice i correspondant à la succession temporelle des événements [Harrah 1963 : 61-62]. Lorsque R reçoit un message, celui-ci a en général pour effet de modifier k , Q , et I , c'est-à-dire de modifier

l' "état d'esprit" de *R*. Le modèle proposé par HARRAH permet de rendre compte de deux processus de communication distincts, l'*entretien* et la *controverse* ; il s'applique en outre aux situations à plus de deux interlocuteurs [Harrah 1963 : 62-70]. L'intérêt et les limites de la théorie des communications de HARRAH sont que celle-ci propose un modèle de l'homme rationnel (qui rappelle par certains côtés l'*homo œconomicus*), susceptible de clarifier des problèmes tels que : "Comment un homme rationnel pose-t-il des questions ?", ou : "Quel est le comportement qui garantit à un récepteur rationnel le minimum de risques dans telle situation ?" (cf. [Harrah 1963 : 9-10]) ; mais ce modèle ne peut déboucher sur des propositions utilisables par les praticiens du questionnaire qu'après que ses implications concrètes aient fait l'objet d'expérimentation dans des situations de communication réelles. Tel est également le cas de la famille de modèles mathématiques du dialogue proposée par C.L. HAMBLIN, qui a pour point de départ l'ensemble des événements de communication verbale (*locution-events*) définis par un message, son émetteur, et son rang d'apparition dans la séquence temporelle du dialogue [Hamblin 1971 : 130-131]. L'auteur introduit dans ses modèles diverses notions intéressantes, comme celles de *soutien* d'une assertion (*commitment*), et de *retrait* du soutien, cette dernière ne devant pas être confondue avec le fait de nier l'assertion en question [Hamblin 1971 : 136-137]. Les différents modèles proposés conduisent notamment à la définition de plusieurs types de relations sémantiques entre les messages. Signalons à ce propos l'essor récent des recherches logiques et linguistiques sur le dialogue, dont on trouvera une bibliographie dans [Ficht 1978 : 112-114].

L'analyse linguistique de la conversation s'inscrit dans la lignée des travaux inspirés par J.L. AUSTIN, et en particulier par sa distinction, parmi les "actes de discours", des actes locutoires, des actes illocutoires, et des actes perlocutoires (cf. [Austin 1962 : 98-103], et notre § 1.1.1.a). Poser une question est un acte illocutoire ; un énoncé à valeur de question est caractérisé (comme nous l'avons vu au premier chapitre) soit par un verbe performatif de la forme : "je vous demande si ...", soit plus fréquemment par une modification de l'ordre des mots, une intonation (signalée graphiquement par le point d'interrogation), des morphèmes particuliers ("est-ce que", "-ne", "ли"), voire l'inachèvement d'une phrase en suspens (cette

description s'applique pour l'essentiel également à l'anglais ; cf. [Searle 1965 : 44], [Sadock 1974 : 9-10]). John SEARLE décrit cinq règles sémantiques régissant l'usage du questionnement [Searle 1969 : 62-66] :

- Règle du *contenu propositionnel*, qui est soit une proposition quelconque (questions extensives), soit une forme propositionnelle quelconque (questions compréhensives).

- Règles *préparatoires*, définissant la situation préalable au questionnement : le locuteur ignore la réponse, et ni le locuteur ni l'interlocuteur ne savent à l'avance si ce dernier aurait fourni ultérieurement l'information demandée sans qu'on l'en prie.

- Règle de *sincérité*, exprimant l'intention du locuteur : celui-ci désire effectivement obtenir l'information demandée.

- Règle *essentielle* : l'énoncé de la question a effectivement valeur de tentative du locuteur pour obtenir de son interlocuteur l'information recherchée.

La règle essentielle ne s'applique que si la règle du contenu propositionnel et les règles préparatoires sont satisfaites [Searle 1969 : 63] ; ce qui implique que la présupposition de la question soit vraie [Katz 1977 : 97]. Le problème central pour nous, dans cette perspective, est celui de la *force illocutoire* de l'énoncé à valeur de question, c'est-à-dire de la pression exercée par le locuteur sur l'interlocuteur pour en obtenir une réponse. Sur ce point, l'apport des linguistes reste purement descriptif : "Exécuter des actes illocutoires est s'engager dans une forme de comportement régie par des règles" [Searle 1965 : 40] ; "Un acte illocutoire est la force conventionnelle d'un énoncé" [Sadock 1974 : 9]. Il débouche sur la formulation, par David GORDON et George LAKOFF, des "postulats de conversation" [Gordon 1971], qui décrivent un certain nombre de contextes pouvant avoir une influence sur le sens perçu de l'énoncé. Sur le plan sémantique, cette théorie vise à expliquer comment l'énonciation d'une chose peut servir à communiquer autre chose, c'est-à-dire pourquoi le sens contextuel de l'énoncé diffère de son sens

littéral ; d'où des relations sémantiques intéressantes, comme celles d'équivalence conversationnelle, d'implication conversationnelle, etc. (cf. [Gordon 1971 : 53-54]). Sur le plan psychosociologique, elle esquisse un cadre formel qui pourrait susciter un renouvellement de la problématique des recherches sur la situation d'entretien (à rapprocher peut être de la notion de "rites d'interaction" chez GOFFMAN).

Le problème des biais dans les réponses introduits par la forme de l'énoncé interrogatif est, à nos yeux, le problème actuellement le plus important pour l'interprétation sémantique des indicateurs. Ce point de vue est d'ailleurs partagé par certains spécialistes de l'enquête : "Pour beaucoup de ceux qui travaillaient dans la branche recherches, il devint bientôt évident que les erreurs ou les biais attribuables à l'échantillonnage et aux méthodes d'administration du questionnaire étaient relativement faibles en comparaison des autres formes de variations - et plus particulièrement des variations attribuables aux différentes manières de formuler les questions" [Stouffer 1950 : 709]. C'est pourquoi de nombreuses recherches expérimentales ont été entreprises sur ce sujet aux États-Unis depuis 1940 environ ; à notre connaissance, et en nous limitant volontairement aux travaux publiés, elles n'ont pas d'équivalent en langue française. Il faut toutefois souligner que les premiers travaux expérimentaux sur ce problème nous paraissent remonter (au moins) à Alfred BINET et ses recherches sur la suggestibilité. Etudiant les mécanismes de la suggestion par questionnaire standardisé, BINET déclare : "J'ai fait cette étude surtout pour me rendre compte de l'influence des mots et des phrases employés pour suggestionner. Il y a là une question de grammaire et de syntaxe qu'il m'a paru utile d'élucider. J'ai indiqué plus haut que les moindres nuances d'inflexion de la voix ont une grande influence sur l'effet d'une suggestion ; [...] mais il est beaucoup plus facile de peser la valeur de chaque mot, en remplaçant un mot par un autre, en employant divers tours de phrase, et en ayant soin de toujours faire lire au sujet des questions

écrites afin d'éviter les variations dans l'accentuation de la voix. Ce procédé des questions écrites permet d'éliminer une bonne partie de ce qui est indéfini et immesurable dans l'action personnelle" [Binet 1900 : 296]. Les travaux de BINET et de ses continuateurs immédiats se situent dans l'optique de la critique du témoignage en psychologie judiciaire. La procédure expérimentale consiste à présenter aux sujets un stimulus (carton sur lequel sont fixés divers objets [Binet 1900 : 250-253] ; série de courtes scènes filmées [Muscio 1916 : 356-357]), puis à le leur faire décrire en répondant à un questionnaire écrit, dont les questions comportent diverses variétés de biais dont on désire mesurer et comparer les effets : présupposition fautive ; forme alternative/forme compréhensive ; article défini/article indéfini ; tournure objective ("Y avait-il un ... ?")/tournure subjective ("Avez-vous vu un ... ?") ; formulation positive ("Y avait-il ... ?")/formulation négative ("N'y avait-il pas... ?"). Les résultats de ces expériences sont extrêmement riches. Retenons seulement que, selon BINET, la suggestion introduite par la formulation n'est en général pas perçue par les sujets, et que son influence est aussi grande sur des jeunes gens relativement instruits (élèves d'école normale) que sur des enfants (élèves de cours moyen et de cours supérieur), la suggestion la plus forte étant obtenue avec des alternatives dont la présupposition est fautive [Binet 1900 : 322-329]. Ce dernier point est corroboré pour les tournures objectives par Bernard MUSCIO (cf. [Muscio 1916 : 376]), qui note en outre que l'effet de suggestibilité est accru par le passage de la tournure objective à la tournure subjective, par la substitution de l'article indéfini à l'article défini, et par l'introduction de la négation [Muscio 1916 : 376-385].

Le texte d'Alfred BINET que nous avons cité au paragraphe précédent résume la problématique des expériences qui seront entreprises quarante ans plus tard aux États-Unis, dans la perspective de l'étude systématique du questionnaire d'enquête. Toutefois, le détail de la procédure a évolué, en ce sens qu'il ne s'agit plus d'évaluer les écarts des réponses à une norme objective (les caractéristiques du stimulus), mais seulement de décrire les variations dans les réponses induites par les variations dans les questions.

En outre, on tend à ne plus se contenter d'observer ces effets dans des conditions de laboratoire, sur des populations relativement homogènes (comme c'est encore le cas pour l'expérience de MENEFEE sur l'influence des stéréotypes [Menefee 1936]), mais à expérimenter en vraie grandeur, en quelque sorte *in vivo*. La procédure utilisée consiste à administrer, à deux échantillons comparables de la même population, deux formes parallèles (ou plus ; cf. p.ex. [Blankenship 1940 : 357-361]) d'un même questionnaire ; la seule différence entre ces formes parallèles concerne évidemment l'objet de l'expérience (place de la question dans le questionnaire, variations dans l'énoncé interrogatif, type de consigne, ordre de présentation des modalités de réponses). Cette méthode, connue sous le nom de *split-ballot technique* [Cantril 1944 : 23], a permis de très nombreuses observations, conduisant à l'énonciation de lois empiriques sur les effets de contexte, et l'incidence de la présentation des modalités de réponse et de la formulation de l'énoncé interrogatif. Nous rappelons brièvement les principales conclusions auxquelles ces expériences ont abouti.

Les recherches empiriques sur la formulation des questions dans les questionnaires d'enquêtes ont permis d'identifier trois groupes d'effets sur les réponses de la structure du questionnaire, ou de la forme de la question :

- *effet de contexte*. La distribution des réponses à une question peut varier selon la position de celle-ci par rapport à d'autres questions sur le même sujet (cf. divers exemples dans [Cantril 1944 : 28-29], [Noelle-Neumann 1970 : 198-199]). Elle varie à plus forte raison lorsqu'une information sur le sujet est apportée explicitement au répondant avant que la question lui soit posée [Hyman 1972 : 206-207]. Chose plus étonnante, elle peut également varier de manière significative selon les informations contenues dans les questions posées après la question étudiée [Noelle-Neumann 1970 : 193-194].

- *effets dus à la structure de la question* (type de consigne ; modalités présentées). En dehors des effets qui résultent des mérites respectifs de l'alternative et de la question à choix multiples (cf.

[Blankenship 1940 : 398-399], [Gallup 1941], [Cantril 1944 : 30-34]), l'effet le plus intéressant sur la distribution des réponses est dû à l'explicitation de certaines modalités implicites, telles que le second membre d'une alternative, l'absence d'opinion, ou les réponses intermédiaires (cf. les exemples cités par [Cantril 1944 : 35-37], [Payne 1951 : 7-8, 56, 62-63], [Presser 1975 : 19-23]).

- *effets dus au choix des mots de la question.* La distribution des réponses à une question est affectée par l'emploi de mots à forte charge affective (cf. p.ex. [Menefee 1936], [Roper 1940], [Rugg 1941], [Cantril 1944 : 42-43]), ou la référence explicite à une personnalité (cf. [Roper 1940], [Cantril 1940], [Gallup 1941], [Cantril 1944 : 38-42]). Mais on peut également ranger dans cette catégorie les effets imprévus dus à une formulation défectueuse de la question, le plus souvent par méconnaissance de la manière dont le problème se pose pour les personnes interrogées (cf. p.ex. [Payne 1951 : 10-13]).

D'autres effets sur les réponses recueillies ont également été décrits. Par exemple, celui de la précodification des réponses, celui de la variété des tâches demandées au répondant, ou celui de la question supplémentaire (relance) attachée à une modalité de réponse [Noelle-Neumann 1970 : *passim*] ; ou encore les effets dus à une formulation de question délibérément orientée (cf. [Cantril 1944 : 44-46], [Payne 1951 : 180-181], [Adams 1956]). D'autre part, dans un texte souvent cité, Lee J. CRONBACH a dressé un bilan, pour les épreuves psychologiques, de l'ensemble des tendances qui poussent une personne "à faire aux items de tests des réponses différentes de celles qu'ils auraient faites si le même contenu leur avait été présenté sous une forme différente" [Cronbach 1946 : 491]. Ce bilan est en parfait accord avec celui que présentent les sociologues de l'enquête, et en particulier CANTRIL. Enfin, les auteurs qui traitent des biais dans les réponses concluent en général leur exposé par des suggestions pratiques relatives à la construction des tests et des questionnaires (suggestions qui seraient ici hors de notre propos).

L'intérêt de l'analyse sémantique des énoncés à valeur de question pour l'interprétation de résultats d'enquête est manifeste. Avant même que se développe l'analyse secondaire, Stanley L. PAYNE en avait présenté une illustration dans la comparaison de résultats d'enquêtes apparemment discordants [Payne 1946]. Malheureusement, les effets décrits dans la littérature consacrée à ce sujet sont loin d'être systématiques (cf. divers contre-exemples dans [Gallup 1941 : 265-266], et [Cantril 1944 : 45]). En outre, l'explication qui en est donnée est relativement pauvre, et peut se résumer en une phrase : l'incidence sur les réponses de la formulation de la question est pratiquement nulle si l'opinion est déjà solidement cristallisée sur le thème qui est l'objet de la question ; cette influence croît en raison de l'indétermination du public sur ce thème [Cantril 1944 : 45]. Il faut par conséquent souhaiter que soient entreprises d'autres recherches systématiques sur les facteurs pouvant influencer sur la distribution des réponses aux questions (cf. les suggestions de recherches futures, dans [Sudman 1974 : 145-147]). En ce qui concerne les sources de variations liées à la formulation des questions, le cadre théorique est bien tracé (cf. p.ex. [Litwak 1956]), et la méthodologie parfaitement au point. Trois approches ont en effet été explorées : l'expérimentation en laboratoire, bien illustrée par les travaux d'Alfred BINET, de Bernard MUSCIO, et plus récemment de Patricia WRIGHT (cf. [Wright 1972]) ; l'analyse secondaire d'enquêtes différentes réalisées à la même époque sur des échantillons comparables (cf. les travaux de PAYNE, du *Projet Simulmatics*, etc., déjà cités) ; et enfin la technique du *split-ballot*, largement utilisée depuis 1940. Si des recherches systématiques du même type pouvaient être entreprises sur la formulation des questions d'enquête en langue française, on trouverait un modèle très détaillé de la conception des questionnaires expérimentaux dans [Blankenship 1940 : 357-361], et une procédure intéressante d'analyse des résultats (par croisement des réponses aux questions testées avec diverses variables) dans [Schuman 1974] et [Presser 1975]. Notons à ce propos que les réticences des instituts de sondages d'opinion ou d'études de marché à l'égard de ce type de recherches en France pourraient peut-être être levées en considérant l'intérêt pratique de l'information apportée par la technique du *split-ballot* au sujet du degré de cristallisation de l'opinion sur un thème donné.



3.2. DU QUESTIONNAIRE A LA "MATRICE DES DONNÉES"

Les questionnaires utilisés en psychologie, en ethnologie, en démographie, ou en sociologie, se présentent sous des formes extrêmement variées. Cette variété de forme est due évidemment à la diversité des théories qui ont présidé à la conception des questionnaires, et aux différences entre les objectifs visés par leurs auteurs. Mais elle est due également à d'autres facteurs plus aisés à classer et à décrire, qui dépendent des conditions d'administration du questionnaire. Ainsi, les questionnaires de personnalité ou d'intérêts, les tests d'aptitudes ou de niveau, sont des instruments d'investigation destinés à une utilisation répétée sur un grand nombre de cas individuels échelonnés dans le temps ; alors que les questionnaires d'opinion ou d'étude de marché sont en général construits pour les besoins d'une étude ponctuelle, qui ne sera en principe pas réitérée sous la même forme (sauf dans le cas de recherches visant à décrire l'évolution d'un phénomène dans le temps). De même, les questionnaires postaux et la plupart des formulaires administratifs sont destinés à être remplis par le répondant seul ; les questionnaires de personnalité ou d'intérêts, et les tests dits "papier-crayon", à être administrés collectivement sur table, et remplis par le répondant sous le contrôle du psychologue ; les questionnaires d'opinion ou d'étude de marché et certains tests de niveau, à être remplis par un spécialiste (enquêteur, psychologue) à partir des réponses fournies (en général verbalement) par le répondant.

Les facteurs liés aux conditions d'utilisation présentent, par rapport à ceux qui dépendent du contenu et des objectifs du questionnaire, l'avantage d'être immédiatement perceptibles d'après le libellé du questionnaire et, s'il y a lieu, des consignes d'utilisation qui l'accompagnent. Ces facteurs constituent le *mode d'emploi* du questionnaire : ils déterminent le "vécu" du questionnement, c'est-à-dire l'adaptation, aux caractéristiques individuelles du répondant, d'un outil cependant standardisé. Ce mode d'emploi précise les règles à appliquer pour :

- *définir à qui s'adresse la question* (exemples : "si oui à la question 5, demander [...]"; "à poser seulement à ceux qui ont des enfants de moins de 6 ans"; "à tout le monde"; "seulement aux possesseurs de télévision couleur"; "si le chef de famille est cultivateur").

- *amener la personne interrogée à une bonne compréhension de la question posée* (exemples : "avant de répondre, lisez attentivement la liste des opinions qui vous sont proposées"; "enquêteur : avant de poser la Q 12, reportez-vous à la Q 1 : si l'interviewé a choisi un modèle sport, dire [...]"; "montrer à l'interviewé les dessins F, G, et H").

- *formuler correctement la réponse* (exemples : "une seule réponse par ligne"; "plusieurs réponses possibles"; "parmi ces trois professions, quelle est celle qui vous paraît [...]"; "classez ces cartons en trois paquets : d'un côté, ceux qui correspondent à des points que vous jugez très importants [...]"; "si le sujet répond [...], lui demander : *précisez* ou *expliquez*").

- *enregistrer la réponse sur le formulaire* (exemples : "cochez le carré correspondant à votre réponse"; "inscrivez le chiffre choisi dans la case prévue à cet effet"; "soulignez l'adjectif qui vous paraît le mieux s'appliquer à [...]"; "si l'interviewé déclare avoir écouté pendant 10 minutes au moins, cocher 1/4 d'heure; 20 minutes au moins, cocher 1/2 heure"; "noter intégralement tout ce que dit l'interviewé").

- *décider si la réponse obtenue est satisfaisante* (exemples : "insister pour obtenir *trois* titres de films"; "temps maximum : une minute pour chaque item").

- *déterminer la question à poser ensuite* (exemples : "passer à la question 9"; "si une seule marque citée, poser Q 15, et passer ensuite à Q 19; si plusieurs marques, poser Q 16"; "si vous ne possédez pas de jardin, passez aux renseignements signalétiques en dernière page"; "proposer les phrases de 3 en 3, quitte à revenir en arrière en cas d'échec").

- *arrêter l'entretien* ("si non fumeur, fin d'interview"; "on arrêtera l'épreuve dès que 2 ou 3 phrases consécutives auront été incorrectement répétées").

Nous appellerons *structure* d'un questionnaire l'ensemble des règles qui définissent à qui ou dans quelles conditions les questions doivent être posées, dans quel ordre elles doivent l'être, sous quelle forme, et quand ou dans quelles conditions il convient d'arrêter l'interrogation. Nous allons voir que, si l'on excepte les cas les plus simples, la structure d'un questionnaire d'enquête est, du moins superficiellement, assez différente de celle d'un test psychologique, en particulier d'un test de niveau. Nous verrons ultérieurement qu'il est toutefois possible de réduire la structure de tout questionnaire (au sens large) à une forme canonique, qui servira de base pour la définition de la "matrice des données".

3.2.1. Structure d'un questionnaire d'enquête

Bien que les questionnaires d'enquête puissent être administrés dans des conditions très variées, nous allons voir que les différences de structure entre un questionnaire postal, un formulaire administratif, un questionnaire administré collectivement sur table, et un questionnaire administré par un enquêteur, sont en réalité assez minces. Pour décrire les principales caractéristiques des questionnaires d'enquête que nous avons analysés, nous commençons par la présentation matérielle du questionnaire ; nous abordons ensuite les problèmes des aiguillages, et des répétitions de questions ; nous terminons par la distinction entre le rôle structurel et le rôle informatif des questions.

a) Présentation matérielle du questionnaire

En règle générale, le questionnaire est constitué d'une sorte de cahier, dont les pages doivent en principe être consultées dans l'ordre de leur apparition, comme dans un livre. Ceci distingue le questionnaire du guide d'entretien par exemple, où l'ordre des points abordés peut ne pas dépendre du support matériel, mais de l'ordre d'énonciation par la personne interrogée ; le guide d'entretien est en fait plus proche matériellement du "livre

brouillé" de l'enseignement programmé, que du livre traditionnel. Il existe toutefois de nombreuses exceptions à cette règle de lecture séquentielle. L'une d'elles a pour unique objectif la facilitation du travail de l'enquêteur. Elle consiste à lui faire consulter d'abord les rectos à partir du début, puis les versos en commençant par la fin (cf. p.ex. : *SOFRES 686*, *SOFRES 1061*). Cette lecture, qui s'apparente à celle de l'écriture en boustrophédon, rend moins malaisée la manipulation du cahier au cours de l'interview.

Les autres exceptions sont liées à la structure logique du questionnaire ; elles reposent toutes sur l'existence d'*aiguillages* (cf. § 3.2.1.b), qui ont pour effet de sélectionner à la fois un sous-ensemble de questions, et un sous-ensemble d'interviewés à qui celles-ci doivent être posées. Matériellement, les aiguillages peuvent se traduire par l'existence d'un premier élément de questionnaire valable pour tous les interviewés, suivi de sous-questionnaires différents, dont un seulement doit être posé à une personne donnée. Tel est le cas par exemple d'une enquête réalisée pour le compte d'un hebdomadaire ("Enquête image presse", *SOFRES 1094*), enquête pour laquelle l'enquêteur disposait d'un court questionnaire visant à déterminer quels quotidiens et quels hebdomadaires la personne interrogée lisait ou avait lus régulièrement ; à partir des réponses à ce premier questionnaire (de couleur blanche), il déterminait lui-même le questionnaire à poser ensuite (bleu : lecteur de l'hebdomadaire en question ; rose : ancien lecteur ; vert : non lecteur). L'utilisation d'une *feuille de contact*, préalablement à l'administration du questionnaire proprement dit, constitue un cas particulier de rupture de la lecture séquentielle, puisque les réponses aux questions de la feuille de contact conditionnent l'administration ultérieure du questionnaire lui-même. Par exemple, dans une enquête auprès des fumeurs de cigares ou de cigarillos (*SOFRES 1061*), la feuille de contact (rose) sert à sélectionner les fumeurs de cigares ou de cigarillos, auxquels seront posées les questions du questionnaire (vert), et accessoirement à décrire sommairement les non fumeurs. Nous verrons au paragraphe suivant les autres formes d'aiguillages qui ne se traduisent pas par une modification matérielle du questionnaire.

Indépendamment de l'existence d'aiguillages, un questionnaire peut être divisé en parties distinctes, dont chacune correspond à un thème ou un domaine donné. Le questionnaire administré par enquêteur, utilisé par Alain GIRARD dans sa recherche sur *Le choix du conjoint* (INED 081) se subdivise en quatre sous-questionnaires, portant respectivement sur les renseignements signalétiques des conjoints, les fiançailles, le mariage, et les opinions générales sur le mariage. De même, pour son étude *Qui sont les femmes ingénieurs en France ?*, Geneviève de PESLOUAN a construit un questionnaire postal (F.ING 69) qui comporte neuf parties distinctes : état-civil, études, activité professionnelle, activités diverses, caractéristiques du mari, opinions sur les études suivies, sur la carrière, sur l'entreprise, sur l'activité professionnelle. Plus généralement, la quasi-totalité des questionnaires d'enquête présente une partie bien distincte du reste du questionnaire, constituée par l'ensemble des questions concernant les renseignements signalétiques de la personne interrogée (âge, sexe, état-civil, profession, etc.). Cette partie, relativement normalisée, se trouve soit (le plus souvent) en fin de questionnaire, soit tout à fait au début de celui-ci (cf. p. ex. : INED 081, F.ING 69, CEREQ - SEU, OCDE - IUT, RNUR 70). D'autre part, la subdivision du questionnaire en parties distinctes peut servir de repères aisément identifiables en cas d'aiguillages importants : dans le questionnaire "10.000 chefs de ménage" (SOFRES 769), la première partie est posée aux non possesseurs d'automobile, la seconde aux automobilistes, la troisième (opinions) à tous ceux qui ont l'intention d'acheter une voiture neuve dans les trois années à venir, et les renseignements signalétiques à tous.

La distinction de parties dans un questionnaire interfère souvent avec la numérotation des questions. Chaque question en effet est située, dans la suite des questions qui constituent le questionnaire, par référence à un ensemble totalement ordonné de "mots", qui sont les "numéros" des questions. D'un auteur à l'autre, la plus grande fantaisie semble présider à la définition de cet ensemble. Par exemple, dans le questionnaire INED 081, la numérotation recommence à 1 dans chacune des parties, et les sous-questions sont désignées par un nombre entier suivi d'une lettre (ex. : *opinions 7b*) ; dans F.ING 69, le système utilisé est identique, à ceci près que les parties

sont numérotées en chiffres romains (ex. : VII 8 c) ; dans *CEREQ - SEU*, la numérotation (nombre éventuellement suivi d'une lettre) est continue, sans tenir compte des subdivisions (ex. : 29 d) ; dans *SOFRES 769*, la numérotation ne tient pas compte des subdivisions, mais les questions sont repérées soit par des nombres éventuellement suivis d'une lettre minuscule (ex. : 50 h) ou d'un adverbe de répétition (ex. : 42 *ter*), soit par des lettres majuscules s'il s'agit des renseignements signalétiques (ex. : G). En l'absence de conventions admises par tous, les systèmes de numérotation des questions dépendent beaucoup des habitudes propres à un institut (voire à un chercheur) donné, et aussi des remaniements successifs que le questionnaire a subis (ainsi s'expliquent la plupart des *bis* et *ter* relevés). Il faut toutefois noter que nous n'avons jamais constaté d'ambiguïtés dans les systèmes de numérotation des questions à l'intérieur de notre corpus : les "mots" ou "numéros" forment bien un ensemble totalement ordonné (ordre strict), et la disposition matérielle des questions est toujours conforme à l'ordre ainsi défini (malgré la diversité des mises en page).

Le questionnaire peut être complété par trois types de documents, qui sont souvent indispensables à sa bonne utilisation :

- une *consigne générale* destinée à l'enquêteur (ou au répondant s'il s'agit d'un questionnaire postal ou d'un formulaire administratif), exposant les buts de l'enquête et son intérêt, la manière de répondre au questionnaire, et les consignes particulières à certaines questions jugées délicates.

- le *matériel* devant être présenté à la personne interrogée à l'occasion de questions données : feuilles comportant une énumération des réponses possibles, ou le libellé de questions trop longues ou trop complexes pour être facilement comprises ; cartons à classer ; dessins, schémas, ou photographies illustrant une question ; parfois même, objets à présenter à la personne interrogée.

- une *grille de codification* pour les réponses aux questions ouvertes, destinée soit à la relecture (cf.p.ex. : *INED 081*), soit à l'enquêteur (ex. : *SOFRES 1061*), soit au répondant (ex. : *CEREQ - SEU*).

Il peut arriver cependant que les consignes d'administration, les textes à présenter à l'interviewé, et la codification des réponses, figurent sur le questionnaire lui-même. Sur ce point également, il n'existe pas de conventions ou de normes admises par tous.

b) Les aiguillages

Comme leur nom l'indique, les aiguillages servent à orienter le répondant au questionnaire vers certaines questions plutôt que d'autres. L'ordre normal de présentation des questions étant celui de leur numérotation, un aiguillage est par conséquent une rupture de séquence. En outre, comme la rupture dépend généralement d'une condition préalable (caractéristique particulière du répondant, ou réponse à une question antérieure), on peut considérer que tout aiguillage est une *rupture de séquence conditionnelle*. Les aiguillages sont en effet le moyen de transformer la relation d'ordre strict totale sur les questions, due à la numérotation, en une relation d'ordre strict partielle (ainsi que nous le verrons au § 3.2.3.a).

Nous avons signalé, au paragraphe précédent, l'existence d'aiguillages servant à déterminer quel questionnaire devait être soumis à la personne interrogée ; en voici un exemple (SOFRES 1094) :

Q 3 - *Et maintenant, voici une liste de journaux ou de revues paraissant toutes les semaines. Pour chaque titre, pouvez-vous m'indiquer si, personnellement, vous avez l'habitude de le lire ou de le feuilleter, chez vous ou au-dehors*

- *régulièrement*
- *assez souvent*
- *rarement*
- *jamais*

(Présenter liste)

[.....]

Si l'interviewé n'est pas un lecteur de X

Q 6c - Vous est-il arrivé de lire X de façon assez régulière, c'est-à-dire au moins une fois par mois, depuis 3 ans ; depuis le changement de format ?

- oui
- non

ENQUÊTEUR : Si l'interviewé est un lecteur actuel de X (voir Q 3 lit régulièrement ou assez souvent) passer au questionnaire lecteur de couleur bleue.

Si l'interviewé est un ancien lecteur de X depuis 3 ans (oui à Q 6c) passer au questionnaire ancien lecteur de couleur rose.

Si l'interviewé est un non lecteur de X passer au questionnaire non lecteur de couleur verte.

De même, un aiguillage peut avoir pour fonction d'orienter l'interviewé vers telle ou telle partie du questionnaire, voire exceptionnellement de définir qui doit être interrogé ensuite, comme dans l'exemple ci-dessous (SOFRES 769) :

Q 1 - Possède-t-on actuellement dans votre foyer une ou plusieurs voitures ?

[.....]

- oui, 1 voiture
 - oui, 2 voitures
 - oui, 3 voitures ou plus
 - non, aucune voiture
- } → Passer à 2e partie Q. 11
- Poser 1^{re} partie Q. 2 à Q. 10

[.....]

Q 9 - Qui s'occupera (dire s'est occupé si la voiture est déjà en commande) du choix et de l'achat de cette voiture ?

- l'interviewé
 - une autre personne du foyer (préciser)
- } → Passer aux instructions en tête de la 3e partie Q. 40 ou Q. 41

[.....]

TROISIÈME PARTIE : OPINIONS

ENQUÊTEUR : Cette partie doit être posée à la personne qui va s'occuper du choix et de l'achat de la voiture. Selon la réponse à Q 9 ou à Q 39, vous devez donc poser cette partie ...

... ou au chef de ménage

... ou à « une autre personne du foyer »

Q 40 - [.....]

Dans ce questionnaire, la question 9 (de la première partie, posée aux non automobilistes) et la question 39 (de la seconde partie, posée aux automobilistes) sont rigoureusement identiques, et ont pour objet de définir le responsable de l'achat projeté d'une voiture neuve ; c'est à cette dernière personne que sera posée la troisième partie du questionnaire. Enfin, nous avons également mentionné les aiguillages susceptibles d'annuler l'interview dès la première question, comme dans les deux cas suivants :

A - Vous arrive-t-il d'acheter, pour les fumer vous-même, des cigares, des petits cigares ou des cigarillos ?

- oui → Poser le questionnaire

- non → Fin d'interview

(SOFRES 1061)

Depuis combien de temps habitez-vous à Z ou dans les communes voisines ?

- moins d'un mois → Arrêter l'interview

- 1 mois à moins d'un an

- 1 an à moins de 3 ans

- 3 ans à moins de 6 ans

- 6 ans et plus

} → Continuer l'interview

(SOFRES 659 : début)

Les aiguillages sont toutefois d'un usage extrêmement courant dans le corps même d'un questionnaire ou d'une partie de questionnaire. Le cas le plus simple se présente lorsqu'une réponse donnée suscite une question supplémentaire et une seule, visant à compléter ou préciser la réponse ; exemples :

22 a - *Croyez-vous qu'il y ait en France un danger de dictature ?*

22 b - *(Si « oui » ou « peut-être ») quand et sous quelle forme pourrait se manifester ce danger ?* (IFOP 58 - 9)

4 - *Y a-t-il eu un mariage religieux ?*

- *oui*
- *non*

(Si oui) Dans quelle religion ? (INED 081 : mariage)

8 a - *Estimez-vous que la nouvelle Assemblée Nationale reflète-t-elle l'opinion du pays ?*

8 b - *(Si non) En quoi la nouvelle Assemblée Nationale ne reflète-t-elle pas l'opinion du pays ?* (IFOP 58 - 12)

(On notera que le dernier exemple ci-dessus est une réalisation simple de ce que ÅQVIST appelle une chaîne de questions avec accumulation de corrections ; cf. § 2.3.1.d). Dans certains cas, plusieurs réponses distinctes conduisent à une même question supplémentaire :

Q 41 - [.....] *Vous-même, quelle solution pensez-vous que vous adopterez le plus souvent (tendre la liste) :*

- *je me garerai en plein centre [...] en payant 500 F.* → *Passer à Q 43*

- *je me garerai à la sortie de l'autoroute en payant 200 F.*

- *je me garerai dans un parking gratuit à M*

- *je ne prendrai plus ma voiture pour aller à Z*

Q 42 - *Quel prix accepteriez-vous de payer pour vous garer en plein centre ?* (SOFRES 571)

Il peut arriver que plusieurs questions successives soient nécessaires après l'aiguillage pour obtenir les informations désirées ; par exemple :

- Q 11 - Est-ce que, depuis 15 jours, vous êtes sorti de N ? Je ne parle pas des sorties pour un travail régulier, mais de toutes les autres ?
- non —————> Passer à Q 15
- oui
- Q 12 - Depuis 15 jours, dans quelle(s) localité(s) êtes-vous allé ? [...]
- Q 13 - Combien de fois y êtes-vous allé depuis 15 jours ? [...]
- Q 14 - Y alliez-vous avec un but précis, ou simplement pour vous y promener ? [...] (SOFRES 571)

Il peut également arriver qu'à chaque réponse (ou chaque groupe de réponses) à la question correspondant à l'aiguillage (*question - filtre*) soit affectée une question ultérieure distincte (ou des séries distinctes de questions) :

- 6 a - Un certain nombre d'électeurs n'ont pas voté le 28 septembre parce qu'ils étaient absents, malades, ou pour tout autre raison. Vous-même, avez-vous voté ?
- 6 b - (Si non) Pourquoi ?
- 7 - (Si oui) Est-ce que je pourrais vous demander comment vous avez voté ? (IFOP 58 - 11)
- 49 - Qu'envisagez-vous de faire immédiatement après avoir terminé vos études à l'I.U.T. ?
- entreprendre un travail professionnel
 - faire votre service militaire
 - poursuivre d'autres études
 - je ne sais pas
 - autre chose (Quoi ? _____).

- 50 - Si vous envisagez de travailler professionnellement dès la fin de vos études à l'I.U.T., savez-vous déjà :
- a) dans quelle branche d'activité vous exerceriez ? [...]
 - b) quel serait votre type d'emploi ? [...]
 - c) combien vous espérez gagner pour commencer ? [...]
 - d) combien vous espérez gagner au bout de 5 ans ? [...]
 - e) dans quelle région vous comptez travailler ? [...]
- 51 - Si vous envisagez de poursuivre d'autres études, dès la fin de vos études à l'I.U.T., précisez :
- a) le type d'études [...]
 - b) leur durée [...]
 - c) quelles sont les raisons de votre projet de poursuivre d'autres études ? [...]
- 52 - Si vous n'envisagez pas de travailler professionnellement dès la fin de vos études à l'I.U.T., savez-vous ce que vous ferez lorsque vous exercerez une activité professionnelle :
- a) dans quelle branche d'activité vous exerceriez [...]
 - b) quel serait votre type d'emploi [...]
 - c) combien vous espérez gagner pour commencer [...]
 - d) combien vous espérez gagner au bout de 5 ans [...]
 - e) dans quelle région vous comptez travailler [...]
- (OCDE - IUT)

Dans ce dernier exemple, la réponse : "travail professionnel" à la Q 49 aiguille vers la Q 50 ; la réponse : "autres études", vers la Q 51 ; et l'ensemble des réponses : "service militaire", "autres études", "ne sait pas", et "autre chose", vers la Q 52. Par conséquent, si l'on répond : "autres études" à la Q 49, on se voit poser les deux questions 51 et 52 ; dans tous les autres cas, on se voit poser soit la Q 50, soit la Q 52.

On appelle *question - filtre* une question telle que la réponse donnée à cette question par la personne interrogée a pour effet de déterminer quelles questions lui seront posées ensuite. Nous venons de voir quelques exemples simples de questions - filtres dans les extraits de questionnaire ci-dessus

(SOFRES 769 : 1 ; SOFRES 1061 : A ; IFOF 58 - 9 : 22a ; etc.) ; nous allons examiner maintenant diverses situations qui sont soit plus complexes quoique d'un usage fréquent (questions filtres en série ; combinaison de plusieurs questions filtres ; questions filtres à effet différé), soit d'un usage relativement rare mais intéressant pour la technique du questionnaire (question - filtre ayant pour unique fonction la modification de l'énoncé d'une question ; aiguillage apparemment sans question - filtre).

Les questions - filtres en série ont pour effet de multiplier le nombre de voies possibles, c'est-à-dire les ensembles distincts de questions devant être posées à des personnes de caractéristiques différentes. Les aiguillages successifs forment donc une arborescence ou un treillis. Exemple :

- 17 - Y a-t-il eu un second tour de scrutin dans votre circonscription ?
- 18 a - (Si « oui ») Avez-vous voté au second tour ?
- 18 b - (Si « non ») Pourquoi ?
(Si l'enquêté n'a pas voté au 2e tour, passer à la question 23)
- A ceux qui ont répondu « oui » à la question 18 a
- 19 - Le candidat pour qui vous avez voté au premier tour s'est-il représenté au second tour ?
(Si « non », passer à la question 21)
- 20 a - (Si « oui ») Avait-il des chances d'être élu ?
- 20 b - Avez-vous voté pour lui ?
- 20 c - (Si « non ») Pourquoi n'avez-vous pas voté pour lui ?
- 21 - Pourrais-je vous demander pour quel candidat vous avez voté au second tour ?
(Si l'enquêté n'a pas voulu répondre à la question 21, posez la question 22 en la faisant précéder de la formule : « En tout cas, voudriez-vous me dire ... »)

22 - Avez-vous voté pour ce candidat surtout ...

- parce qu'il était le moins éloigné de vos opinions
- parce qu'il pouvait battre un autre candidat dont vous redoutiez l'élection
- pour ces deux raisons à la fois
- parce qu'il était sûr d'être élu
- parce que vous vouliez manifester votre opposition au candidat que vous n'aimiez pas et qui était sûr d'être élu ?

23 - [.....]

(IFOP 58 - 12)

Il peut arriver que des questions - filtres en série constituent une chaîne de questions avec accumulation de corrections (§ 2.3.1.d). Dans l'exemple ci-dessous, c'est évidemment le cas pour les questions 4, 6 et 7 ; si la question 3 avait été remplacée par sa négation ("Etes-vous dépourvu d'un poste ..."), on aurait même une chaîne de quatre questions avec accumulation de corrections (cas assez rare dans le questionnaire d'enquête) :

Q 3 - Y a-t-il dans votre foyer un poste de radio en état de fonctionner ?

- oui (Passer à Q 4)
- non

[.....]

(Si la personne n'a pas la radio, passer à Q 22)

[.....]

Q 4 - Parmi ces stations de radio, lesquelles écoutez-vous ?

- France-Inter
- France-Culture
- France-Musique
- Monte-Carlo
- Radio Andorre
- Luxembourg
- Europe n° 1
- autres
- ne sait pas

[.....]

(Si n'écoute pas France-Inter, passer à Q 8).

[.....]

Q 6 - *Ecoutez-vous les émissions régionales à la radio [...] ?*

- souvent
- quelquefois
- rarement
- jamais
- ne sait pas

Passer à Q 8

Q 7 - *Dans l'ensemble, trouvez-vous ces émissions régionales ... ?*

- excellentes
- très bonnes
- bonnes
- moyennes
- mauvaises
- ne sait pas

(ORTF 65)

Une combinaison de questions - filtres se distingue du cas précédent en ce que les questions - filtres successives n'augmentent pas le nombre de voies distinctes, mais effectuent une sélection en quelque sorte par tris successifs suivis de fusions. Dans l'exemple ci-après, deux questions filtres successives (Q 3 et Q 3a) sont utilisées pour déterminer les automobilistes ayant un "point de vente habituel", automobilistes auxquels sont posées les questions 4 et suivantes (décrivant ce point de vente et les avantages qu'il présente) :

Q 3 - *Avez-vous un endroit où vous achetez régulièrement votre essence, c'est-à-dire un endroit où vous faites au moins la moitié de vos achats d'essence ?*

- oui (passer à Q 4)
- non

Q 3a - *(Si non à Q 3) Parmi les différents endroits où vous achetez votre essence, y en a-t-il un où vous allez un peu plus souvent qu'aux autres ?*

- oui
- non (passer à Q 7)

Q 4 - *(Si oui à Q 3 ou à Q 3a) Cet endroit est-il situé ...*

- près de chez vous ?
- près de votre lieu de travail ?
- sur le trajet maison-travail ?
- autre (préciser).

(SOFRES 340)

Une question - filtre à effet différé peut se présenter comme une question ordinaire, non suivie immédiatement d'un aiguillage. On s'aperçoit toutefois qu'il s'agit d'une question - filtre au fait que, dans le cours du questionnaire, on découvre un aiguillage utilisant les réponses à cette question.

Exemple :

Q 2 - Lisez-vous ou feuillotez-vous X :

- (1) régulièrement
- (2) assez souvent
- (3) de temps en temps
- (4) rarement

ou (5) pratiquement jamais ?

[.....]

Aux lectrices de X seulement (c'est-à-dire 1 ou 2 ou 3 à Q 2)

(Pour les non-lectrices, passer à Q 42)

Parlons maintenant plus précisément de X.

Q 33 - Avez-vous constaté un changement depuis quelque temps dans X
ou au contraire estimez-vous qu'il n'a pas changé ?

(SOFRES 686)

Evidemment, il arrive fréquemment qu'une question - filtre soit à la fois à effet immédiat et à effet différé. Tel est le cas de la combinaison de questions - filtres présentée à l'alinéa précédent ("point de vente habituel"), à laquelle on se réfère plus loin dans le questionnaire pour déterminer la question la mieux appropriée parmi les deux suivantes (entre lesquelles existe d'ailleurs une relation d'inclusion : Q 7a contient Q 7b ; cf. § 2.2.5.e) :

Si OUI à Q 3 ou Q 3a

Q 7a - La deuxième fois que vous avez acheté de l'essence, était-ce ...

- à votre endroit habituel (Passer à Q 11)

- dans un garage

- sur la piste d'une station service

- au bord du trottoir, à une pompe

- autre (préciser)

- NSP

} → (Passer à Q 7c)

Si NON à Q 3 et Q 3a

Q 7b - *La dernière fois que vous avez acheté de l'essence, était-ce ...*

- *dans un garage*
- *sur la piste d'une station service*
- *au bord du trottoir, à une pompe*
- *autre (préciser)*
- *NSP*

(SOFRES 340)

Enfin, nous avons relevé un exemple de question à réponses multiples jouant le rôle d'une question - filtre à effet différé :

Q 4 - *Parmi ces stations de radio, lesquelles écoutez-vous ?*

[.....]

Si la personne écoute Radio Luxembourg, demander :

Q 18 - *Écoutez-vous Radio Luxembourg régulièrement ou quelquefois, aux heures suivantes ? [...]*

Si la personne écoute Inter-Grandes Ondes, demander :

Q 19 - *Écoutez-vous Inter-G.O. régulièrement [...]*

Si la personne écoute Europe N° 1, demander :

Q 20 - *Écoutez-vous Europe N° 1 régulièrement [...]*

Si la personne écoute Inter-Petites Ondes, demander :

Q 21 - *Écoutez-vous Inter-P.O. régulièrement [...]*

(ORTF 65)

Les formes de questions-filtres que nous venons de présenter sont d'un usage relativement courant dans les questionnaires d'enquête. Les deux formes que nous analysons maintenant sont plus rares. La première, la *question filtre modifiant un énoncé* de question, peut soit se présenter explicitement comme un aiguillage entre deux questions dont le texte diffère en un de ses éléments (cf. l'exemple *SOFRES 340 : 7a, 7b* ci-dessus), soit n'apparaître que dans le corps de la question, comme une simple variante (cf. ci-dessus l'exemple *SOFRES 769 : 9* au début de ce paragraphe ; ou encore l'exemple *IFOP 58-12 : 21-22*, dans l'alinéa sur les questions-filtres en série). Exemples de simples variantes :

40 a - Quel est le parti politique dont vous vous sentez le plus proche ?

[.....]

40 c - Etes-vous membre de ce parti (ou : d'un parti, si l'enquêté n'a pas voulu répondre à la question 40 a) ?

(IFOP 58 - 9)

Q 7 - Y a-t-il actuellement une voiture en commande dans votre foyer ?

- oui → Passer à Q 8

- non → Poser Q 7bis

Q 7bis - Pensez-vous que vous-même, ou un autre membre de votre foyer, achèterez une voiture au cours des 3 prochaines années ?

- oui, sûrement

- c'est possible

- non, je ne pense pas

- non, sûrement pas

} → Passer à Q 10

Q 8 - Cette voiture sera-t-elle (ou : est-elle) neuve ou d'occasion ?

(SOFRES 769)

Q 1 - Voudriez-vous tout d'abord me préciser le modèle de la voiture que vous avez achetée [...]

Q 2 - *En plus de cette ... (voiture de Q 1), de quelle autre
voiture dispose-t-on dans votre foyer ? [...]*
(SOFRES 967)

Q 47 - *Voici encore une liste d'endroits que l'on peut trouver plus
ou moins agréables à habiter, à conditions de logement
égales bien entendu. Vous-même, lequel de ces cinq endroits
trouveriez-vous le plus agréable à habiter ?*

[.....]

Q 51 - *Vous m'avez cité, comme lieu préféré. Voyez-vous toutefois
des inconvénients à habiter (lieu préféré) ?*
(SOFRES 659)

On voit sur ces exemples que, dans les phrases ainsi modifiables, tout se passe comme si l'un des éléments de l'énoncé (en principe autre que celui sur lequel porte l'interrogation) était une variable, dépendant d'autres paramètres du questionnaire ; le rôle de la question - filtre étant alors d'affecter une valeur (constante) à cette variable.

Nous avons appelé *aiguillages sans question - filtre apparente* les aiguillages qui ne se réfèrent pas à une question posée explicitement dans le questionnaire, c'est-à-dire une question pour laquelle les réponses ont été ensuite enregistrées par l'enquêteur, ou le répondant lui-même. De tels aiguillages sont assez rares dans les questionnaires d'enquête, car l'information apportée par la réponse à une question - filtre est en général conservée. Dans le premier exemple ci-après, il s'agit d'une enquête politique en trois vagues successives, sur le modèle de l'enquête par panel. Cependant, aucune des personnes interrogées ne l'a été pour l'ensemble des trois vagues : certaines ont répondu aux vagues 1 et 2, d'autres aux vagues 1 et 3, d'autres aux vagues 2 et 3 ; d'autres encore n'ont répondu qu'à la vague 1 ou à la vague 2. La question - filtre implicite du questionnaire de la vague 2 (IFOP 58 - 11) élimine les personnes ayant déjà répondu à la première vague, et auxquelles par conséquent il n'est pas utile de poser à nouveau les questions concernant leur affiliation ou certaines attitudes politiques, et leurs caractéristiques socio-économiques (renseignements signalétiques). Les autres exemples sont suffisamment clairs pour se passer de commentaires.

Les questions qui suivent ne sont pas à poser dans la forme A
(personnes déjà interrogées avant le référendum)

44 - Certaines personnes disent qu'il y a actuellement en France une menace de crise économique. Cette menace vous paraît-elle très grave, peu grave, inexistante ? (IFOP 58 - 11)

(si la personne interrogée est autre que le chef de famille)

31 a - Quelle est la profession du chef de famille ? (IFOP 58 - 9)

(Demander à la personne si elle a des enfants de 6 à 15 ans. Si non, passer à la Q 14)

Q 13 - Dans l'ensemble, comment trouvez-vous les émissions de radio destinées aux enfants de 6 à 15 ans ? (ORTF 65)

c) Les répétitions de questions

Pour pouvoir affirmer qu'une question donnée se trouve répétée dans un questionnaire, il faut non seulement que l'énoncé interrogatif qui l'exprime soit rigoureusement identique dans toutes les occurrences présumées de ladite question, mais encore (s'il s'agit d'une question fermée ou précodée) que la liste des réponses possibles à la question soit également rigoureusement identique dans tous les cas. Cette dernière condition interdit de considérer comme une répétition (au sens strict) les questions fermées 7 a et 7 b du questionnaire *SOFRES 340* citées au paragraphe précédent, puisque la réponse : "à votre endroit habituel", qui figure parmi les réponses possibles à Q 7a, n'est pas une réponse possible à Q 7b.

En principe, l'auteur d'un questionnaire d'enquête n'éprouve pas le besoin de poser deux fois la même question à une même personne au cours d'une interview unique (encore que cela puisse se concevoir, par exemple, à des fins de contrôle). Aussi est-il naturel de ne rencontrer les répétitions de questions que dans deux situations :

- interviews réitérées d'une même personne, en vue d'apprécier l'évolution de ses connaissances, de ses croyances, de ses attitudes, de ses

comportements, etc. (recherche sur panel ; évaluation de l'impact d'une publicité par la méthode test - retest).

- dans un questionnaire donné, au cours d'une enquête donnée, distinction de sous-parties dont chacune s'adresse à des personnes de caractéristiques différentes (sélectionnées par un aiguillage), sans qu'il soit possible de poser toutes les questions destinées à l'ensemble des interviewés dans la partie commune du questionnaire.

Nous excluons des répétitions de questions la ré-utilisation d'une question, sous une forme rigoureusement identique, dans des enquêtes différentes auprès de personnes différentes, comme dans les études longitudinales sur échantillons distincts ("baromètres" de notoriété ou de popularité, p.ex.).

La première situation décrite ci-dessus sort évidemment du cadre de l'analyse d'un questionnaire. La seconde apparaît comme une conséquence de la manière dont l'auteur du questionnaire se représente le déroulement de l'interview. En effet, considérons par exemple le questionnaire "10.000 chefs de ménage" (SOFRES 769), déjà mentionné aux § 3.2.1.a et b. Les première et seconde parties de ce questionnaire s'adressent l'une aux non possesseurs de voiture, et l'autre aux automobilistes, tandis que la quatrième partie (renseignements signalétiques) s'adresse à tous les interviewés. On pourrait penser que la première partie ne comporte que les questions à poser aux seuls non automobilistes, et la seconde, que les questions à poser aux seuls automobilistes ; les questions communes aux deux sous-échantillons se trouvant regroupées dans une partie commune située soit avant, soit après ces deux parties. C'est effectivement le cas pour les renseignements signalétiques. Par contre, pour les questions relatives au nombre de personnes vivant au foyer qui ont leur permis de conduire, et à l'éventualité d'acheter une voiture au cours des trois prochaines années, l'auteur a jugé préférables de les faire figurer dans chacune des deux parties, plutôt que de risquer de rompre le fil de l'entretien en les insérant au début du questionnaire ou parmi les renseignements signalétiques, ou bien de compliquer la tâche de l'enquêteur en augmentant le nombre (déjà élevé) des aiguillages. On a en conséquence les identités suivantes : Q 6 = Q 36bis, Q 7 = Q 37, Q 7bis = Q 37bis, Q 8 = Q 38, Q 9 = Q 39.

L' "enquête image presse" (SOFRES 1094), déjà citée, constitue un exemple particulièrement intéressant de répétitions de questions. Nous avons vu qu'après une série de questions sur ses habitudes de lecture de quotidiens ou d'hebdomadaires (Q 1 à Q 6c, questionnaire blanc), la personne interrogée se voyait proposer une forme différente de questionnaire selon qu'elle était lecteur, ancien lecteur, ou non lecteur d'un hebdomadaire X, objet de l'enquête. La comparaison des trois formes du questionnaire révèle l'existence de répétitions de questions soit dans les trois versions, soit dans deux seulement d'entre elles (à l'exclusion toutefois du couple : lecteur/non lecteur). Le tableau n° 3.10 présente simultanément une table de concordance des questions répétées (horizontalement), et la description de l'enchaînement des thèmes abordés dans chaque forme du questionnaire. Pour simplifier le tableau, nous avons considéré comme équivalentes la question 12a du questionnaire "lecteur" ("J'aimerais maintenant que vous me disiez ce que vous aimez dans X [...]"), et la question 20a du questionnaire "ancien lecteur" ("J'aimerais maintenant que vous me disiez ce que vous aimiez dans X [...]"), ainsi que les questions 12b "lecteur" et 20b "ancien lecteur", qui présentent la même légère différence de formulation ; mais il est bien clair qu'il s'agit là d'une infraction à la règle énoncée au début de ce paragraphe.

Il n'est pas sans intérêt d'élargir le concept de répétition de questions en créant la notion de *répétition partielle*. Le cas des questions 12a "lecteur" et 20a "ancien lecteur" de l'exemple précédent en est une illustration, puisque ces énoncés sont *identiques à un élément près*, qui est le temps du verbe de l'interrogative indirecte. Il en est de même pour les questions SOFRES 340 : 7a, 7b, qui ne diffèrent que par l'une des réponses possibles proposées à l'interviewé. Ces deux exemples ressortissent d'ailleurs à la même situation, visant à l'adaptation de l'énoncé interrogatif à l'interlocuteur, sans que la signification attachée par l'auteur du questionnaire aux réponses possibles soit modifiée. Remarquons d'ailleurs que, dans les deux cas mentionnés, les questions répétées partiellement dans le questionnaire ne

Thèmes abordés	Forme du questionnaire		
	<i>Bleu</i> Lecteur	<i>Rose</i> Ancien lecteur	<i>Vert</i> Non lecteur
Historique du lecteur de X	Q 7 - Q 8b		
Opinions sur les hebdomadaires et attentes		Q 7 - Q 11b	Q 7 - Q 11b
Historique de l'abandon de lecture de X		Q 12 - Q 16	
Connaissance de X			Q 12 - Q 16
Description de X à un ami	Q 9	Q 17	
Image des lecteurs de X	Q 10 - Q 11d	Q 18 - Q 19d	Q 17 - Q 18d
Qualités et défauts de X	Q 12a, Q 12b	Q 20a, Q 20b	
Image de X	Q 13a - Q 15	Q 21a - Q 23	Q 19a - Q 21
"X est un journal bourgeois ..."	Q 16	Q 24	
Attentes du lecteur de X	Q 17		
Opinions sur un journal politique	Q 18, Q 19	Q 25, Q 26	Q 22, Q 23
Habitudes de lecture et centres d'intérêt	Q 20 - Q 48		
Jugements sur X et son contenu		Q 27 - Q 29	
Réactions à la présentation d'un numéro de X			Q 24 - Q 30
Attitudes vis-à-vis de l'abonnement	Q 49a - Q 49d	Q 30a - Q 30d	Q 31a - Q 31d
Renseignements signalétiques	Q A - Q Fb	Q A - Q Fb	Q A - Q Fb

Tableau n° 3.10 - Exemple de répétition de questions : "Enquête image presse" (SOFRES 1094)

sont jamais posées à un même interviewé. C'est pourquoi il serait peut être légitime d'assimiler ces exemples aux répétitions de questions au sens strict, à condition toutefois de définir avec précision les règles relatives à cette assimilation.

Il existe par contre de nombreux cas de répétitions partielles de questions auprès d'une même personne au cours d'un même interview, indépendamment de tout problème d'adaptation de la manière d'interroger aux caractéristiques spécifiques de l'interlocuteur. On peut en distinguer deux variétés, selon que le nombre de répétitions est fixe, ou variable. Exemples de série de répétitions partielles en nombre fixe :

166 - Combien de fois depuis un mois avez-vous pris un repas :
 a - chez des amis
 b - dans la famille (RNUR 70)

Q 23 - Voici une liste de choses que l'on peut trouver dans une station service. Pour chacune d'elles, je voudrais que vous me disiez si vous la trouvez très importante, assez importante, peu importante, ou pas importante du tout ? [...]
 (SOFRES 340)

8 a - Dans votre famille, est-ce qu'on a plutôt la même position ou pas la même position que vous sur le projet de constitution ?

8 b - Et vos collègues, ont-ils plutôt la même position ou pas la même position que vous ?

8 c - Et vos amis, ont-ils plutôt la même position ou pas la même position que vous ? (IFOP 58 - 9)

6 - Comment étiez-vous classé au point de vue travail scolaire en classe de :

- 6e
- 3e
- terminale

(OCDE - IUT)

Q 5 - Lorsque vous écoutez France-Inter, est-ce plutôt sur les G.O. ?
les P.O. ? la M.F. ?

- le matin
- entre 12 h et 14 h
- dans l'après-midi (jusqu'à 19 h)
- entre 19 h et 20 h 30
- après 20 h 30

(ORTF 65)

Lorsque la réponse à chaque question de la série est du type *oui/non*, il s'agit le plus souvent d'une question dite "à réponses multiples" :

16 - Pouvez-vous me dire si vous êtes plutôt d'accord ou plutôt pas d'accord avec chacune des opinions suivantes :
[.....] (IFOP 69-5)

IV 7 - Lisez-vous régulièrement :

- un quotidien ?
- un hebdomadaire ?
- des revues techniques ou professionnelles ?
- des ouvrages techniques ou professionnels ?
- plus de 5 livres non techniques par an ?

(F.ING 69)

Enfin, on rencontre parfois des répétitions partielles en nombre fixe concernant un ensemble de questions, comme dans les deux exemples ci-dessous. On remarquera, dans le second exemple, la présence d'un aiguillage à l'intérieur de l'ensemble de questions soumis à répétition :

10 - Indiquez avec le plus de précision possible :

a - la profession de votre père, de votre mère, de vos
grands-parents paternels : grand-père, grand-mère,
maternels : grand-père, grand-mère

b - les diplômes qu'ils ont obtenus (OCDE - IUT)

Q 22 - a - quand êtes-vous allé au cinéma pour la dernière fois ?

b - où était-ce ?

- à N (passer à Q 22d)

- autre ville (préciser)

c - vous est-il arrivé d'aller au cinéma à N ?

si oui : quand pour la dernière fois ?

d - Avec qui étiez-vous (à N) ?

e - Etait-ce l'après-midi ou le soir (à N) ?

ENQUÊTEUR : mêmes questions à propos des points suivants :

- allé au théâtre

- allé au concert

- allé au café

[.....]

(SOFRES 571)

Les répétitions partielles en nombre variable dépendent toujours d'un aiguillage, qui détermine le nombre et la nature des répétitions. Dans le premier exemple ci-dessous, les réponses à la première partie de la question à réponses multiples Q 4 déterminent pour quelles station de radio (de la liste présentée à l'interviewé) sera posée la question sur la fréquence de l'écoute ; dans le second exemple, le nombre de répétitions de la Q 39 dépend évidemment du nombre de modèles de voitures cités en notoriété spontanée (Q 38) ou suggérée (Q 38bis) :

Q 4 - Parmi ces stations de radio, lesquelles écoutez-vous ?

Pour chaque station écoutée, demander :

l'écoutez-vous :

- tous les jours ou presque ?

- souvent (2, 3 fois par semaine) ?

- quelquefois (moins de 2 fois par semaine)

(ORTF 65)

Q 38 - En ce qui concerne, par exemple, la marque Z, quels sont tous les modèles de voitures Z que vous connaissez ? (Enquêteur : ne rien suggérer).

Q 38bis - Toujours parmi les modèles Z, avez-vous entendu parler de la ... (énoncer successivement chacun des modèles de la liste qui n'ont pas été cités à Q 38) ?

Q 39 - (A poser pour chacun des modèles cités à Q 38 et Q 38bis [...]).
Combien coûte actuellement, selon vous, la ... ?

(SOFRES 967)

Dans les deux exemples ci-dessus, le nombre maximum possible des répétitions peut être déterminé à partir du texte du questionnaire (nombre de stations de radio figurant sur la liste présentée ; nombre total de modèles fabriqués par la marque Z). Il arrive toutefois que ce nombre maximum ne soit pas explicitement fixé par l'auteur du questionnaire ; c'est généralement le cas par exemple pour les renseignements signalétiques concernant l'ensemble des frères et soeurs de l'interviewé (*OCDE - IUT : 9*), ou l'ensemble des personnes vivant au foyer (*RNUR 70 : 1*). Autres exemples :

Q 1 - Possède-t-on actuellement dans votre foyer une ou plusieurs voitures ?

- oui, 1 voiture

- oui, 2 voitures

- oui, 3 voitures ou plus

- non, aucune voiture

Passer à Q 11

[.....]

Q 11 - Marque et modèle des voitures actuellement possédées par le foyer ? (SOFRES 769)

III 13 - Entreprises dans lesquelles vous avez travaillé [...]

III 14 - Postes successifs que vous avez occupés [...]

(F. ING 69)

Toutefois, s'il est exact que l'énoncé des questions et les consignes données à l'enquêteur dans les exemples ci-dessus ne permettent pas de définir le nombre maximum possible des répétitions, en pratique, les exigences de la mise en page du questionnaire ou du plan de carte limitent ce nombre. Ainsi, le questionnaire *OCDE - IUT* prévoit seulement cinq lignes pour l'énumération des frères et soeurs ; le questionnaire *RNUR 70*, dix lignes pour les personnes vivant au foyer ; le questionnaire *SOFRES 769*, trois colonnes pour l'ensemble des voitures possédées par le foyer ; le questionnaire *F. ING 69* enfin, cinq lignes pour les entreprises, et sept pour les postes successifs. Cependant, cette manière de procéder laisse subsister, dans les cas extrêmes (familles nombreuses, multi-possesseurs de voitures, changements d'emploi fréquents), une certaine incertitude dans les informations recueillies (nombre exact de frères et soeurs ; critères de sélection des personnes citées ; etc.).

d) Rôle informatif et rôle structurel des questions

La nécessité de distinguer entre le rôle informatif et le rôle structurel d'une question est bien illustrée par l'existence de questions - filtres à effet différé. De telles questions assument, dans le questionnaire, deux fonctions nettement séparées : d'une part recueillir une information utile pour l'auteur du questionnaire, et d'autre part permettre à l'enquêteur (ou au répondant lui-même, si le questionnaire est auto-administré) de sélectionner les questions à poser à un moment donné, distinct du moment du recueil de l'information. La possibilité de distinguer entre ces deux rôles est manifeste : il existe des questions qui n'ont qu'une fonction informative, et elles constituent d'ailleurs la majorité des questions d'un questionnaire ; il existe également, quoique beaucoup plus rarement il est vrai, des aiguillages sans question - filtre apparente, dans lesquels la sélection des questions à poser est déterminée par une information recueillie en quelque sorte hors questionnaire. Si en France la coutume n'est pas de matérialiser ces deux fonctions séparément dans le cas de questions - filtres à effet immédiat, aux Etats-Unis par contre il arrive que l'on demande à l'enquêteur à la fois d'enregistrer la réponse (informative) à la question - filtre et d'indiquer sur le dictionnaire, à un "point de contrôle" (*checkpoint*), la branche de l'aiguillage choisie à partir de cette réponse :

A 12 - *Besides voting, people sometimes take part in several other types of political activities that are related to national politics and the government in Washington. Please look at this list of activities and tell me if you have done any of them in the last two or three years.*

[.....]

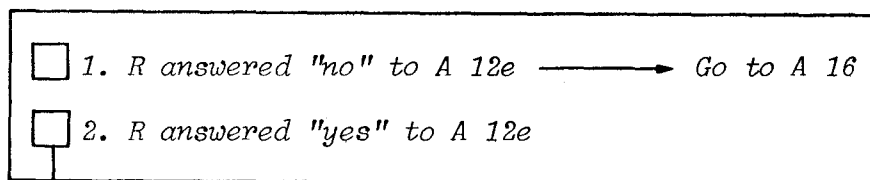
A 13 - *INTERVIEWER CHECKPOINT*

- | |
|--|
| <input type="checkbox"/> 1. <i>R answered "no" to all activities → Turn to p. 6, A 16</i>
<input type="checkbox"/> 2. <i>R answered "yes" to one or more activities</i> |
|--|

↓
For each item answered "yes" in A 12 ask appropriate item below

[.....]

A 15 - INTERVIEWER CHECKPOINT (see activity A 12e)



A 15a - [...]

(SRC 76 Pre)

Le rôle *structurel* des questions, ou plus précisément des informations recueillies dans le questionnaire, est analogue au rôle des instructions de rupture de séquence conditionnelles dans les langages de programmation. Selon les cas, la fonction de l'aiguillage peut être :

- d'adapter la formulation des questions à certaines des caractéristiques de l'interviewé (exemple : mise au féminin ou au pluriel de mots ou d'expressions déterminés) ;

- de compléter certains énoncés interrogatifs à partir du contenu des réponses antérieures de la personne interrogée (ex. : *SOFRES 967 : Q 2 ; SOFRES 659 : Q 51*) ;

- d'éviter la redondance entre questions (ex. : *SOFRES 571 : Q 42*, évitant la réponse : "Mais je viens de vous le dire !" à la Q 43 devenue sans objet pour ceux qui ont déclaré à la Q 42 : "Je me garerai en plein centre [...] en payant 500 F.")

- de sélectionner une sous-population de répondants au sujet desquels les informations demandées présentent un intérêt plus grand pour le chercheur qu'au sujet d'autres catégories de répondants (exemples : dans l' "enquête image presse" *SOFRES 1094*, questions sur les habitudes de lecture de l'hebdomadaire habituel, posées aux seuls lecteurs de X ; dans le "sondage radio-télévision" *ORTF 65*, question 31 : "Dans l'ensemble, comment trouvez-vous les émissions de TV destinées aux enfants de 6 à 15 ans ?", posée aux seuls téléspectateurs ayant des enfants) ;

- de prévenir une réponse correctrice, niant l'un des présupposés de la question.

Cette dernière fonction des aiguillages est certainement la fonction la plus importante pour l'analyse ultérieure des informations recueillies. C'est en effet une fonction d'assurance des questions qui les suivent, selon la méthode appelée par ÅQVIST : *méthode de WHATELY - PRIOR* (cf. § 2.3.1.a). Par exemple :

- 5 - Y a-t-il eu contrat de mariage ?
 (Si oui)
 Sous quel régime les deux conjoints sont-ils mariés ?
 (INED 081 : mariage)
- 12 a - Le référendum doit avoir lieu le 28 septembre. Avez-vous l'intention d'aller voter ce jour là ?
 [.....]
- 12 c - (Si "oui") Savez-vous déjà comment vous voterez au référendum ?
- 13 a - (Si "oui" ou "?" à la question 12 c) Est-ce que je pourrais vous demander s'il y a des chances pour que vous votiez plutôt oui ou plutôt non au référendum ? (IFOP 58 - 9)

On peut d'ailleurs postuler que, dans tout questionnaire d'enquête bien construit et non conçu à des fins manipulatoires, les questions non factuelles (attitudes, opinions, comportements projetés ou hypothétiques) sont toujours assurées. D'où l'importance, pour les questions dont la précodification ne prévoit pas de réponse corrective, du contenu des questions-filtres dont elles dépendent.

Si l'on poursuit l'analogie entre le rôle structurel des informations recueillies au cours de l'enquête (dans le questionnaire ou en-dehors) et les instructions de rupture de séquence dans un programme informatique, il est tentant d'établir un parallèle entre les répétitions partielles de questions et les boucles de programmes. On peut en effet considérer la partie commune aux énoncés répétés comme une sorte de forme propositionnelle, dont les variables liées à l'opérateur de répétition prennent leurs valeurs dans autant d'ensembles qu'il y a de variables liées. Chaque ensemble est totalement ordonné par

une relation d'ordre strict. Pour reprendre un exemple simple (à une seule variable), la forme propositionnelle correspondant à la question *SOFRÉS 571 : Q 22a* déjà citée serait : "Quand êtes-vous allé au x pour la dernière fois ?", x prenant ses valeurs dans l'ensemble totalement ordonné : {cinéma, théâtre, concert, café}. Dans de tels cas, la poursuite ou l'arrêt de la répétition partielle de la question est déterminé par un aiguillage dépendant de la réponse donnée par l'enquêteur (ou le répondant dans un questionnaire auto-administré) à une question de la forme : "Reste-t-il des valeurs que x doit prendre et qu'il n'a pas encore prises ?", avec comme consignes : "si oui, réitérer la série en donnant à x la valeur suivante ; si non, continuer en séquence". Il est clair que, dans cette perspective, une question implicite de ce type n'a aucune valeur informative.

En conclusion des analyses qui précèdent, nous dirons que le rôle structurel d'une information, relativement à un questionnaire d'enquête, est de permettre de déterminer quelle question (ou série de questions) il y a lieu de poser à un moment donné de l'entretien, et éventuellement sous quelle forme il y a lieu de la poser ; la règle étant qu'en l'absence d'information à fonction structurelle, les questions sont posées dans l'ordre dans lequel elles se présentent (ordre des "numéros" des questions). La distinction entre les rôles informatif et structurel conduit à définir trois catégories de questions :

- les questions *purement informatives*, pour lesquelles les réponses sont enregistrées dans le questionnaire, sans toutefois que l'information apportée par celles-ci puisse modifier l'ordre et le mode de présentation des questions ;

- les questions *purement structurelles*, qui sont posées exclusivement hors questionnaire, et pour lesquelles les réponses ont pour unique fonction de modifier l'ordre ou le mode de présentation des questions du questionnaire ;

- les *questions-filtres*, qui ont à la fois un rôle informatif et un rôle structurel.

Les questions purement structurelles peuvent ne pas s'adresser à la personne interrogée, mais à l'enquêteur. Elles n'appartiennent pas au questionnaire proprement dit, mais sont toutefois présentes (au moins implicitement) dans la situation d'entretien.

D'autre part, dans la mesure où un aiguillage détermine quelles questions doivent être posées à un interviewé donné, il détermine également *ipso facto* à qui doit être posée une question donnée. Nous appellerons *questions omnibus* les questions destinées à être posées à l'ensemble des personnes interrogées, et questions *non-omnibus* celles qui, à la suite d'un aiguillage, ne sont posées qu'à une partie seulement des personnes interrogées. Par commodité de langage, nous dirons également qu'une question non-omnibus est *dominée* par l'aiguillage dont elle dépend, ou (ce qui revient au même) que celui-ci la *domine*. Remarquons à ce propos que, si l'on ne considère que les questions figurant effectivement dans le questionnaire (ce qui écarte les questions purement structurelles), on peut recenser des questions purement informatives omnibus, des questions purement informatives non-omnibus, des questions-filtres omnibus, et des questions-filtres non omnibus.

3.2.2. Structure d'un test de niveau

Alors qu'il est extrêmement rare de trouver un questionnaire d'enquête sans rupture de séquence (toutefois, dans notre corpus, les questionnaires *BNIST* et *REL 55* n'ont chacun qu'un seul aiguillage), la grande majorité des épreuves psychologiques sont purement séquentielles. C'est en particulier le cas des inventaires de personnalité (cf.p.ex. : *BERGER*, *BERN 35*, *CAT 16 PF*, *E-I/6*, etc.), des questionnaires d'intérêt (cf.p.ex. : *QIA/m*, *QI filles* et *garçons*; etc.), des échelles d'attitudes en général (cf. : [Debaty 1967], [Robinson 1973]), et même de certains tests de niveau administrés collectivement (cf.p.ex. : *ABS*, *BV 8*, *BV 50*, *EAI 5/1* et *5/2*, *ECO 5*, *SUBES*, etc.). La structure de ces épreuves est par conséquent très simple, et ne justifie pas un effort d'analyse particulier. Les tests dits "de niveau intellectuel" (ou :

"d'intelligence") administrés individuellement, par contre, présentent souvent, malgré leur apparente simplicité, une structure particulièrement complexe. Un bon exemple en est l'"échelle métrique de mesure de l'intelligence" (cf. [Binet 1909 : 113]) mise au point par Alfred BINET et Th. SIMON (cf. [Binet 1911]). Ses deux révisions américaines successives dites "de Stanford" [Pichot 1954 : 26-27], le test de TERMAN, et le TERMAN - MERRILL (cf. [Terman 1937] et [Cesselin 1952]), ses révisions françaises par René ZAZZO en 1949 (*BS Z*) et en 1966 [Zazzo 1966], ont approximativement la même structure, que nous décrivons ci-après. D'autres tests de niveau individuels, comme le WECHSLER-BELLEVUE (cf. [Wechsler 1944]) ou la révision du TERMAN-MERRILL par le Centre d'Application de l'I.N.O.P. (*TMI*), constituent des "échelles par points" [Pichot 1954 : 28-29], dans lesquelles toutes les questions sont posées séquentiellement ; l'appréciation de la réussite globale au test se fait par addition des points obtenus à chaque épreuve particulière (On notera que cette solution avait été déjà utilisée par BINET et SIMON avant la mise au point définitive du test qui porte leur nom : cf. [Binet 1907 : 97-99], et qu'elle a été reprise par ZAZZO pour sa "nouvelle échelle métrique d'intelligence" : cf. [Zazzo 1966 : II 11]). Ces tests conçus pour être d'un maniement plus facile que ceux du type BINET-SIMON, ont en conséquence une structure plus pauvre.

Pour une recension de l'ensemble des caractéristiques servant à décrire la structure d'une épreuve psychologique (et par conséquent le déroulement de sa passation), les tests du type BINET-SIMON apparaissent donc comme les plus intéressants. Nous décrivons ci-après la révision ZAZZO 1949 de ce test de niveau intellectuel (*BS Z*) ; sa présentation matérielle d'abord, puis l'ensemble des règles qui définissent sa structure : consignes d'appréciation de la réponse, et règles d'enchaînement des sous-questions et des questions jusqu'à l'arrêt de l'épreuve.

a) Présentation matérielle du BINET-SIMON

Cette épreuve étant vraisemblablement le test de niveau le plus connu, il n'est pas indispensable de fournir une description détaillée de sa présentation matérielle. D'autre part, la révision faite par René ZAZZO en 1949 concerne

essentiellement la correspondance entre une question, et l'âge auquel il est "normal" (au sens fort pour SIMON, au sens d'une norme statistique pour ZAZZO ; cf. [Zazzo 1946 : 36-39]) de répondre correctement à la question. C'est pourquoi elle n'affecte pas la présentation matérielle de l'épreuve, si ce n'est en proposant une numérotation des réponses correctes aux questions, inconnue des auteurs, et pourtant commode pour une description de la structure du test.

Les consignes d'application de l'épreuve figurent dans la brochure [Binet 1911]. Celle-ci expose les règles générales d'administration du test (qui définissent sa *structure* au sens où nous l'entendons), et, pour chaque question, les règles particulières qu'il convient d'appliquer pour la poser au sujet et pour évaluer la réponse de celui-ci. Le matériel nécessaire comporte en outre :

- divers objets usuels (à faire nommer par le sujet) ;
- un assortiment de pièces de monnaie ;
- cinq boîtes d'aspect identique mais de poids différents ;
- deux cartes blanches, dont l'une est coupée selon la diagonale ;
- un livret contenant les dessins à montrer au sujet ;
- une feuille de notation des réponses ;
- une feuille de niveau.

(Le psychologue a en outre besoin d'un crayon, d'une paire de ciseaux, de feuilles de papier blanc, et d'un chronomètre).

La feuille de notation des réponses (ou "feuille d'interrogatoire") est, ainsi que le rappelle la page de consignes générales qui lui est annexée, un "*questionnaire* pour la prise de niveau intellectuel d'un enfant". Ce questionnaire se divise en deux parties distinctes (cette division étant la principale modification apportée par SIMON à la version de 1911 du test ; cf. [Binet 1911 : 36 n 1]) : la première comporte neuf questions (ou séries de sous-questions) ayant pour fonction de permettre une première évaluation approximative du niveau du sujet (avec ZAZZO, nous les appellerons "questions

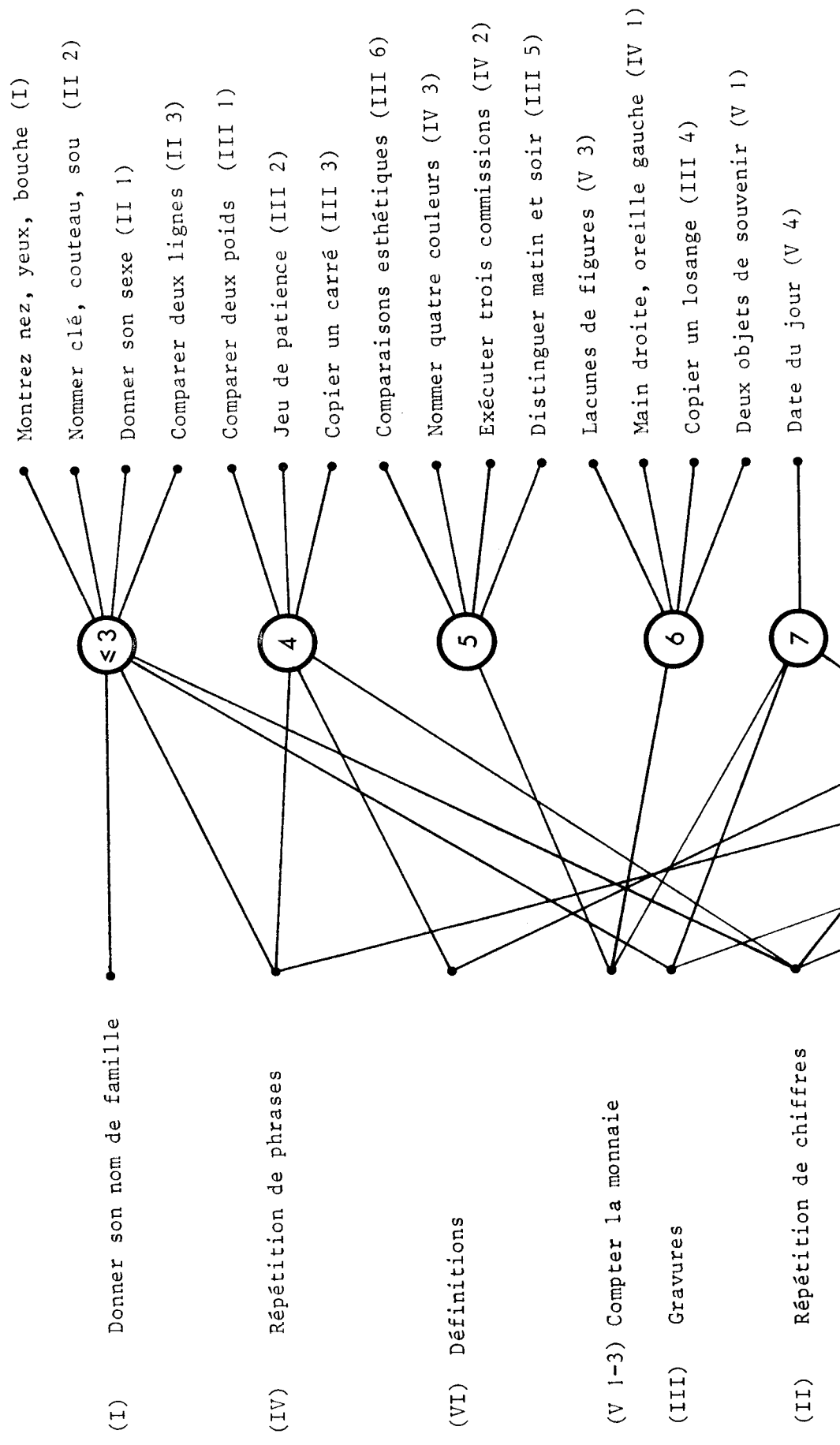
de sondage") ; la seconde comprend trente-trois questions, classées par catégories dont chacune correspond à un "âge mental". Dans la version d'origine, ces questions ne sont pas numérotées sur la feuille (elles le sont dans le livret ; cf. tableau n° 3.11 ci-après). La "feuille d'interrogatoire" sert à noter intégralement les réponses du sujet aux questions posées (il n'y a pas de précodification).

La feuille de niveau compte cinquante-quatre items, regroupés en catégories correspondant à un "âge mental". La version originale de cette feuille présente onze âges mentaux (de "3 ans" à "adultes"), sans numérotation des items ; la révision ZAZZO définit également onze "âges mentaux" (de "3 ans et moins" à "supérieur à 12 ans"), mais classe les items, numérotés cette fois, de manière très sensiblement différente (déclassement de certains des items inférieurs à 12 ans, surclassement de certains des items classés "12 ans", regroupement de treize items dans une seule catégorie appelée : "supérieur à 12 ans"). Point important pour l'analyse de la structure du test, la numérotation des items correspond à un classement, du plus facile au plus difficile [Zazzo 1966 : I 55]. Dans les deux versions du test utilisées actuellement, les vingt "items de sondage" (réponses aux neuf "questions de sondage") sont signalés sur la feuille de niveau par un astérisque.

La feuille de niveau sert à évaluer le niveau intellectuel du sujet à partir de ses réponses enregistrées sur la feuille d'interrogatoire. Mais sa fonction la plus importante (pour l'analyse de la structure du test) est de présenter, à chaque étape de la passation, un bilan des épreuves réussies jusqu'ici par le sujet, bilan qui servira à déterminer quelle question doit être posée ensuite, et à quel moment il convient d'arrêter la passation. Les items de la feuille de niveau peuvent être considérés comme une codification des réponses jugées acceptables aux questions de la feuille d'interrogatoire. A chaque question correspond donc au moins un item de la feuille de niveau (à certaines questions toutefois peuvent correspondre jusqu'à quatre items) ; d'autre part, à tout item correspond exactement une question. L'application de l'ensemble des questions sur l'ensemble des items est par conséquent une injection. Comme par ailleurs la classification des items en "âges mentaux" constitue une partition de l'ensemble des items, l'application de l'ensemble

AUTRES QUESTIONS

QUESTIONS DE SONDAGE



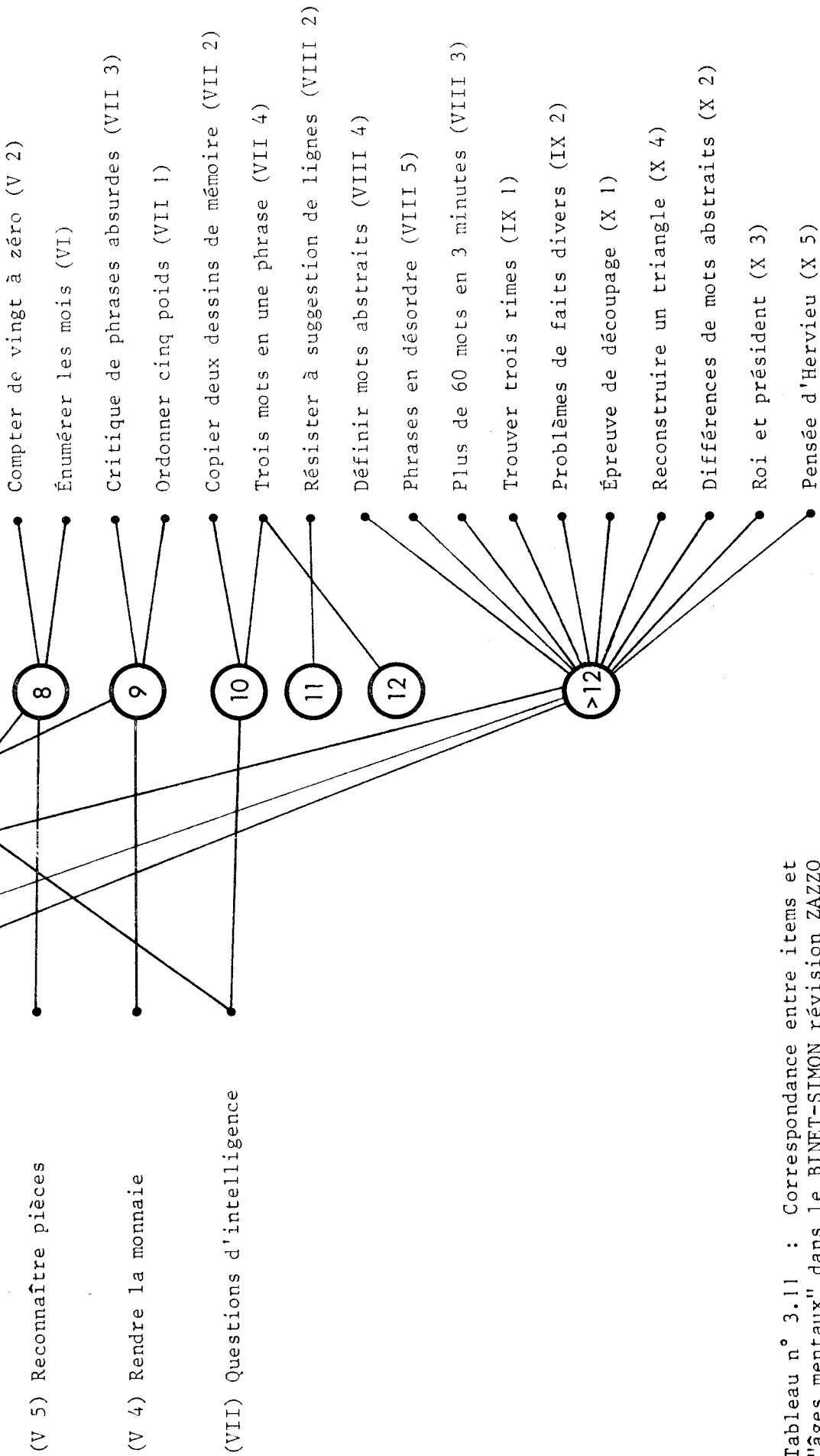


Tableau n° 3.11 : Correspondance entre items et "âges mentaux" dans le BINET-SIMON révision ZAZZO

des items sur l'ensemble des "âges mentaux" est une surjection. Enfin, la composition de ces deux applications, l'application de l'ensemble des questions sur l'ensemble des âges mentaux, n'a pas de propriétés particulières ; elle est représentée par le tableau n° 3.11, dans lequel les "questions de sondage" ont été figurées à part des autres questions.

b) Règles d'appréciation des réponses

Les relations établies par le psychologue entre les questions de la feuille d'interrogatoire et les items de la feuille de niveau sont le résultat d'une analyse du contenu des réponses fournies par le sujet. Cette procédure de codification, qui s'effectue pendant le déroulement même de l'épreuve, constitue une différence fondamentale entre les questionnaires d'enquête et les tests de niveau du type BINET-SIMON. Nous avons vu que, dans les questionnaires d'enquête, la plupart des questions sont soit fermées, soit précodées. Lorsque les questions sont précodées, la codification de la réponse ne présente généralement pas de difficulté pour l'enquêteur. Enfin, s'il y a dans le questionnaire quelques questions ouvertes (non précodées), il est tout à fait exceptionnel qu'elles jouent le rôle d'une question-filtre (leur fonction est en général purement informative) ; l'enquêteur n'a donc pas à procéder lui-même et sur le champ à l'analyse de la réponse qui lui est fournie. Dans le BINET-SIMON, toutes les questions sont ouvertes, et toutes ont un rôle structurel. Il est donc nécessaire que le psychologue effectue l'*analyse* de la réponse dès que celle-ci lui est donnée. En outre, il doit également *évaluer* cette réponse par rapport à certaines normes de réussite qui lui sont imposées pour des raisons de comparabilité des résultats individuels (même si, historiquement, ces normes ont été dégagées empiriquement).

Formellement, il est possible de distinguer quatre procédures de codification des réponses, selon d'une part que la question est unique ou au contraire se subdivise en sous-questions, et d'autre part selon que la codification compte une ou plusieurs variétés de réponses acceptables. Le cas le plus simple est celui de la question unique à réponse acceptable unique. Il

est illustré par la question : "Comment t'appelles-tu ?", à laquelle (éventuellement après relance) le sujet doit donner son nom de famille (sondage I, *BS Z* : 7) ; par les épreuves de comparaison de deux lignes (II 3, *BS Z* : 6) ou de deux poids (III 1, *BS Z* : 9), pour lesquelles il n'y a qu'une seule réponse correcte ; ou encore par l'énumération des mois de l'année, pour laquelle on tolère un oubli ou une inversion (VI, *BS Z* : 32). Dans tous les exemples cités, le résultat de l'évaluation est une alternative : réussite/échec, se traduisant sur la feuille de niveau par un + ou un - placé devant l'item correspondant (encore qu'il soit possible de recourir à une appréciation en cinq points, cf. [Binet 1911 : 33], mais sans que ce gain de précision dans la notation ait un effet sur le déroulement du test ou sur l'évaluation finale de l' "âge mental").

Un ensemble de sous-questions peut donner lieu à une seule évaluation globale de type dichotomique. Tel est le cas par exemple de l'épreuve de nomination de trois objets usuels, dans laquelle on montre successivement à l'enfant une clé, un canif fermé, et un sou, en lui demandant : "Qu'est-ce que c'est que ça ? Comment cela s'appelle-t-il ?" (II 2, *BS Z* : 2) ; la réussite à l'épreuve exige que les noms des trois objets soient cités par l'enfant. Il en est de même pour les trois comparaisons esthétiques (III 6, *BS Z* : 15), la désignation de quatre couleurs (IV 3, *BS Z* : 17), et la résolution de deux problèmes de faits divers (IX 2, *BS Z* : 48). Dans tous ces exemples, il était nécessaire, pour que l'épreuve soit réussie, que les réponses à toutes les sous-questions aient été correctes. Il arrive cependant que le système d'évaluation soit plus tolérant, n'exigeant par exemple que deux réponses exactes sur trois pour les comparaisons de deux objets en souvenir (V 1, *BS Z* : 24), les trois essais d'ordination de cinq poids (VII 1, *BS Z* : 36), la résistance à une suggestion de lignes (VIII 2, *BS Z* : 40), les définitions de mots abstraits (VIII 4, *BS Z* : 42), et les phrases en désordre (VIII 5, *BS Z* : 43) ; trois réponses exactes sur quatre pour les lacunes de figures (V 3, *BS Z* : 21) ; ou trois bonnes réponses sur cinq pour la critique de phrases absurdes (VII 3, *BS Z* : 35).

La réponse à une question unique peut faire l'objet d'une évaluation en plusieurs points. Ainsi, la question suivante :

"*Bien, voici trois mots (les écrire) ; PARIS, FORTUNE, RUISSEAU. Eh bien, fais une phrase dans laquelle tu mettras ces trois mots-là. Tu peux ajouter d'autres mots, mais il faut que ces trois-là soient dans la même phrase.*" (VII 4)

donne lieu à un classement des réponses en trois groupes principaux (cf. [Binet 1911 : 91-92]) :

- il existe trois idées distinctes (ex. : "Paris est une ville ; une personne a une fortune ; le ruisseau coule") → *échec* ;

- il existe deux idées (ex. : "A Paris, il y a des ruisseaux et des hommes qui ont une grande fortune") → *semi-réussite* de niveau 10 ans (BS Z : 37) ;

- il existe une idée unique (ex. : "A Paris, dans un ruisseau, j'ai trouvé une fortune") → *réussite* de niveau 12 ans (BS Z : 41).

Enfin, pour certaines questions, qui sont en fait des *ensembles de sous-questions*, les réponses sont évaluées sur des échelles de cotation en trois, quatre, voire cinq niveaux de réussite différents (et strictement hiérarchisés). Ces questions sont toujours des "questions de sondage". Nous avons déjà vu (§ 2.2.3.d) que les réponses aux questions de définition de mots abstraits (sondage VI) sont classées en : "définitions par l'usage" (BS Z : 13) et : "définitions supérieures à l'usage" (BS Z : 34). Pour chacun des cinq mots proposés ("chaise", "cheval", "fourchette", "table", "maman"), la réponse est donc évaluée en trois points (échec / usage / > usage) ; pour l'ensemble de l'épreuve, le niveau retenu est celui qui est majoritaire (c'est-à-dire : "qui se rencontre dans trois d'entre elles" [Binet 1911 : 58], ce qui laisse subsister une ambiguïté pour certaines des configurations de réponses possibles). De même, pour les trois gravures à décrire (sondage III), on distingue les réponses par énumération (BS Z : 3), par description (BS Z : 25) , et par interprétation (BS Z : 46) ; le niveau retenu étant "le type de réponse qui a eu la majorité" [Binet 1911 : 47], on obtient une cotation de l'épreuve en quatre niveaux (avec les mêmes ambiguïtés que pour les définitions de mots abstraits). Pour la série de huit sous-

questions appelée : "questions d'intelligence" (sondage VII), la cotation de l'ensemble de l'épreuve obéit aux règles suivantes [Binet 1911 : 63] :

- réussite à au moins trois des cinq dernières sous-questions
 → réussite aux *questions difficiles*, de niveau 10 ans (BS Z : 39) ;

- sinon : réussite à deux au moins des trois premières sous-questions → réussite aux *questions faciles*, de niveau 9 ans pour la version de BINET et SIMON, de niveau 7 ans pour la révision ZAZZO (BS Z : 28) ;

- sinon → *échec*.

Enfin, les questions de répétition donnent lieu à une cotation en quatre points pour les phrases (selon le nombre de syllabes répétées correctement : *échec/BS Z : 8/11/49*), et en cinq points pour les chiffres (selon la longueur de la série répétée correctement, avec deux échecs possibles par série : *échec/BS Z : 4/10/30/47*).

c) Aiguillages internes à une sous-question

Pour analyser les consignes qui président au déroulement de la passation du test, il est commode de diviser la difficulté et de procéder en trois étapes : étudier d'abord les aiguillages à l'intérieur d'une sous-question (ou d'une question simple) ; examiner ensuite les règles internes à une série de sous-questions constituant une question unique, c'est-à-dire donnant lieu à une évaluation globale de l'ensemble des réponses recueillies ; analyser enfin les règles d'enchaînement des questions entre elles pour l'ensemble du test. C'est pourquoi, dans ce paragraphe, nous nous limitons à la description de la structure d'une sous-question (ou d'une question simple). Nous traitons les autres points dans les deux paragraphes qui suivent.

Il arrive souvent qu'une sous-question (ou une question simple) se subdivise elle-même en plusieurs énoncés, dont toutefois la plupart sont optionnels ; soit que le psychologue ait le choix entre deux énoncés différents pour le sujet, mais équivalents pour lui-même, soit qu'il dispose d'énoncés facultatifs de relance, de correction, ou de demande d'éclaircissement.

Dans tous les cas, l'usage de ces énoncés dépend d'un aiguillage préalable. Pour plus de clarté, nous avons classé les aiguillages internes à une sous-question (ou à une question simple) en sept catégories : ceux qui modifient l'énoncé de la question ; ceux qui suscitent une relance, pour incompréhension de la question elle-même, pour incompréhension de la procédure à suivre, pour réponse ambiguë ou obscure, pour réponse insuffisante, et pour réponse fautive ; ceux enfin qui déterminent l'arrêt de l'épreuve imposée par la sous-question.

Les *aiguillages modifiant l'énoncé* de la question elle-même ne se trouvent, dans le BINET-SIMON, qu'en deux exemplaires :

- l'épreuve "donner son sexe" (II 1, *BS Z* : 5) a pour consigne : Si le sujet est un petit garçon, dire : "Es-tu un petit garçon ou une petite fille ?" ; si c'est une fille, dire : "Es-tu une petite fille ou un petit garçon ?".

- l'épreuve "distinguer matin et soir" (III 5, *BS Z* : 19) : "A présent, sommes-nous le matin ou le soir ?", est suivie de la consigne : "Placer à la fin de la question le moment de la journée où l'on n'est pas".

Le but de la modification des énoncés est, dans les deux cas, "d'éviter le danger d'une bonne réponse par écholalie" [Binet 1911 : 65 ; 71 n 1].

Les *relances pour incompréhension de la question* ont pour but de préciser l'objectif de l'épreuve, en décrivant, de manière plus détaillée que dans l'énoncé primitif, le résultat escompté de la tâche que le sujet doit accomplir. La relance est déclenchée soit par l'absence de réponse, comme dans les deux premiers exemples ci-après, soit par une mauvaise réponse, comme dans les deux derniers :

- "Loger trois mots en une phrase" (VII 4, *BS Z* : 37 et 41) a pour consigne de relance : "Quelques enfants déclarent qu'ils ne comprennent pas [...]. Nous complétons souvent : Tu peux ajouter d'autres mots, mais il faut que les trois mots se trouvent dans la même phrase. D'autres enfants comprennent, mais ils n'arrivent pas à faire une phrase quelconque ou bien une phrase

capable de les satisfaire. Comme ces derniers pourraient être des puristes, il faut insister" [Binet 1911 : 90-91].

- "Comparer deux objets en souvenir" (V 1, BS Z : 24) : "Quelle différence y a-t-il entre une mouche et un papillon ?", est complété par : "Si l'enfant ne répond pas à la question : quelle différence ? essayer - après avoir indiqué son premier échec - la formule : En quoi n'est-ce pas pareil ?" (cf. [Binet 1911 : 79 n 1]).

- pour la répétition de chiffres (sondage II, BS Z : 4, 10, 30, 47), on peut être amené à préciser au sujet qu'il ne s'agit pas de répéter les chiffres au fur et à mesure que le psychologue les prononce, mais à la fin de la série ; ou bien qu'il faut les répéter dans l'ordre dans lequel ils ont été prononcés [Binet 1911 : 40-41].

- pour les phrases à répéter (sondage IV, BS Z : 8, 11, 49), "s'il y a lieu on rappellera la consigne, à savoir qu'on exige le mot à mot, et le même ordre dans les mots" [Binet 1911 : 49].

Signalons encore un exemple de *fausse relance* de ce type : dans l'épreuve "Trouver trois rimes" (IX, 1, BS Z : 45), après avoir demandé : "Tu sais ce que c'est qu'une rime ?" (version ZAZZO) ou : "Tu sais ce qu'on appelle les rimes ?" (version originale), on fournit l'explication au sujet quelle que soit sa réponse [Binet 1911 : 104-105].

Les *relances pour incompréhension de la procédure* à mettre en oeuvre concernent ici les épreuves avec manipulation d'objets :

- pour "compter quatre objets" (sondage V 1, BS Z : 16), dans la version ZAZZO, si l'enfant répond sans manier les objets, il faut insister pour qu'il compte effectivement avec ses doigts ; dans la version originale, cette règle ne s'appliquait que si en outre la réponse donnée était fausse [Binet 1911 : 50].

- pour "comparer deux poids" (III 1, *BS Z* : 9), l'enfant doit être incité à prendre en main les deux poids s'il ne le fait pas spontanément [Binet 1911 : 67].

- pour "reconnaître toutes les pièces de notre monnaie" (sondage V5, *BS Z* : 31), si le sujet retourne la pièce pour en lire la valeur (ce qui rend la réussite trop facile), on doit l'en empêcher.

- pour le "jeu de patience" (III 2, *BS Z* : 12), si le sujet retourne l'une des deux pièces du puzzle (ce qui rend la réussite impossible), il faut la remettre à l'endroit et lui dire qu'il suffit de rapprocher les morceaux l'un de l'autre.

Les *relances pour réponse ambiguë* sont nécessaires lorsque la formulation par le sujet de sa réponse est telle qu'elle ne permet pas au psychologue de déterminer si l'épreuve est réussie ou non. Le besoin d'éclaircissements se fait surtout sentir pour les "questions d'intelligence" (sondage VII, *BS Z* : 28, 39), la "définition de mots abstraits" (VIII 4, *BS Z* : 42), et la "critique de phrases absurdes" (VII 3, *BS Z* : 35) [Binet 1911 : 60 n2, 103]. La formulation de la relance est laissée à l'initiative du psychologue dans la majeure partie des cas.

Les *relances pour réponse insuffisante* (mais non fausse) sont en général normalisées, en ce sens qu'elles sont déclenchées par des types de réponses bien déterminés (dont la liste est fournie), et que leur formulation est précisée à l'avance. Exemples :

- "donner son nom de famille" (sondage I, *BS Z* : 7) : "si l'enfant donne son prénom seulement, insister pour avoir son nom de famille. Ex. : Roger, mais Roger comment ?"

- "gravures" (sondage III, *BS Z* : 3, 25, 46) : si l'enfant ne répond pas (version ZAZZO), ou s'il dit : "c'est une image", "c'est une carte postale", on reprend la question sous la forme : "Qu'est-ce que tu vois là ?".

- "définitions" (sondage VI, *BS Z* : 13, 34) : si, à la question : "Qu'est-ce que c'est qu'une chaise ?", l'enfant répond : "C'est une chaise", on insiste en disant : "C'est une chaise, c'est entendu ; mais qu'est-ce que c'est qu'une chaise ?".

- "date du jour" (V4, *BS Z* : 27) : "Noter d'abord la première réponse de l'enfant. Puis, s'il ne donne pas d'emblée la date entière, faire compléter ce qui manque : quel jour de la semaine ? quel mois ? le combien ? de quelle année ?".

- "lacunes de figures" (V3, *BS Z* : 21) : si pour l'un des trois premiers dessins (visages), à la question : "Qu'est-ce qui manque à cette figure-là ?", l'enfant répond : "Le corps", "Le ventre", "Les pieds", etc., on ajoute : "Oui, mais à la figure ?".

Les *relances pour réponses fausses* sont plutôt rares. En effet, les auteurs du test insistent à de nombreuses reprises sur l'obligation de ne pas trop aider l'enfant (afin de conserver aux épreuves leur caractère normalisé), et de lui donner toujours l'impression que l'on est satisfait de sa réponse. Toutefois, comme nous l'avons signalé plus haut, dans la version originale du test, en cas de réponse fausse aux épreuves de comptage d'objets ou de pièces de monnaie (sondage VI, 2, 3 ; *BS Z* : 16, 20, 26), il fallait insister pour que l'enfant compte avec ses doigts et à haute voix [Binet 1911 : 50-51].

Les *aiguillages déterminant l'arrêt de l'épreuve* n'ont ici de sens que dans les cas d'épreuve en temps limité ; dans les autres cas, il n'y a pas d'aiguillage, et l'épreuve est terminée après l'obtention de la réponse attendue (correcte ou incorrecte). Les épreuves du BINET-SIMON ne sont en général pas explicitement limitées par le temps, sauf évidemment celle qui consiste à citer plus de soixante mots en trois minutes (VIII 3, *BS Z* : 42). Toutefois, il est conseillé de limiter le temps imparti à 10 secondes pour "compter neuf francs" (sondage V3, *BS Z* : 26) ; à 20 secondes pour le comptage à rebours de vingt à zéro (V2, *BS Z* : 29) ; à une minute pour "trouver trois rimes" (IX 1, *BS Z* : 45). Notons, à propos des problèmes de durée de l'épreuve, qu'il peut y

avoir fixation d'un délai *minimum*, comme pour les "questions d'intelligence", où les auteurs recommandent de laisser au sujet au moins vingt secondes de réflexion [Binet 1911 : 59].

d) Aiguillages internes à une question

La règle d'enchaînement des sous-questions dans une même série est, en principe, la continuation en séquence à partir de la première sous-question de la série. Les exceptions à cette règle dans le BINET-SIMON sont peu nombreuses. Cependant, l'analyse de leur structure conduit à dégager cinq types d'aiguillages différents, selon que ceux-ci déterminent la sous-question par laquelle commencer, l'arrêt de la série, ou l'enchaînement des sous-questions ; et, dans ce dernier cas, selon qu'ils orientent vers une sous-question située en amont, une sous-question située en aval, ou une réitération de la même sous-question.

La *détermination de la première sous-question* à poser pour l'épreuve de répétition de phrases (sondage IV, BS Z : 8, 11, 49) est régie par la consigne suivante : "Si l'enfant reste muet, intimidé, on l'amorce en lui faisant répéter les phrases les plus courtes. Sans quoi on commencera plutôt par une phrase telle que la mémoire des chiffres en rend la répétition probable" [Binet 1911 : 49]. En clair, cela implique une certaine relation d'équivalence entre les résultats à l'épreuve de répétition de chiffres (sondage II, BS Z : 4, 10, 30, 47) et ceux à l'épreuve de répétition de phrases, de telle sorte que le psychologue puisse faire commencer cette dernière épreuve à un niveau un peu inférieur à celui qu'il estime devoir être atteint par le sujet. La connaissance de cette relation est supposée acquise par le psychologue au cours de son expérience professionnelle antérieure, puisqu'elle n'est pas explicitée dans le manuel. Il est possible toutefois de lui donner une forme plus objective en comparant les niveaux de réussite aux deux épreuves, jugés "normaux" pour un âge donné. Dans la version ZAZZO, on a les équivalences : 2 chiffres = 6 syllabes (≤ 3 ans), 3 chiffres = 10 syllabes (4 ans), et 7 chiffres = 26 syllabes (> 12 ans) ; la version originale du test divergeant de sa révision pour la seconde équivalence. Si l'on appelle S le nombre de syllabes répétées, et C le nombre de chiffres répétés, on a par conséquent la

relation linéaire : $S = 4C - 2$. On peut donc considérer que la longueur de la première phrase à répéter, soumise à un sujet non intimidé ayant préalablement réussi à répéter C chiffres, sera fixée en pratique à : $S = 4C - 6$ ou : $S = 4C - 8$.

L'*aiguillage vers une sous-question située en aval*, différente de la sous-question suivant immédiatement, a pour effet d'accélérer le déroulement de l'épreuve en fonction des réussites précédentes. L'épreuve de répétition de chiffres est composée, dans l'ordre, de trois séries de 2 chiffres, trois séries de 3 chiffres, etc., jusqu'à 3 séries de 7 chiffres, dont l'ensemble constitue l'épreuve proprement dite (il n'est pas interdit, bien au contraire, d'essayer des séries plus longues : cf. [Binet 1911 : 40 n 1] ; mais le résultat n'en est pas pris en compte). La consigne est que dès qu'un nombre déterminé est répété, on passe à une série comportant un chiffre de plus : "ce n'est qu'en cas d'échec qu'on essaie de nouveau avec un même nombre de chiffres" [Binet 1911 : 40]. Au cours de l'épreuve de répétition de phrases, pour laquelle on dispose de quinze "phrases" comportant respectivement 2, 4, 6, etc. jusqu'à 30 syllabes, "on n'essaiera pas toutes les phrases, mais on ira de deux en deux ou de trois en trois" [Binet 1911 : 49].

L'*aiguillage vers une sous-question située en amont*, différente de celles qui viennent d'être posées, est la conséquence, dans la répétition de phrases, de la consigne ci-dessus : en cas d'échec, il est nécessaire en effet de revenir en arrière, afin de "fixer la limite de répétition correcte d'une phrase après une seule audition" [Binet 1911 : 49].

La *réitération d'une sous-question* ne se rencontre, dans le BINET-SIMON, que pour les épreuves de pesée. Pour la comparaison de deux poids (III 1, BS Z : 9), il est possible de faire procéder à nouveau à la pesée des deux boîtes : "pour s'assurer que cette désignation n'est pas l'effet du hasard, on peut d'ailleurs avoir besoin de recommencer" [Binet 1911 : 67] (Dans la révision ZAZZO, la réitération est systématique). Pour l'ordination de cinq poids (VII 1, BS Z : 36), le sujet fait obligatoirement trois essais.

On risquerait d'en déduire que la réitération de l'épreuve est déterminée par une sorte de compteur, imposant mécaniquement la répétition de la question, indépendamment du résultat des essais. Si l'on sait que l'épreuve est jugée réussie lorsque deux essais ont été réussis, on comprend qu'en pratique le processus de réitération est un peu plus subtil ; et que, selon le pronostic qu'il fait sur le "niveau mental" du sujet, l'examineur pourra arrêter l'épreuve soit après deux échecs consécutifs, soit après deux réussites consécutives, afin de passer plus rapidement aux autres questions.

La *détermination de l'arrêt de la série*, sans que toutes les sous-questions aient été posées, dépend en général du nombre des échecs enregistrés précédemment :

- pour les répétitions de chiffres, on arrête l'épreuve lorsqu'il y a eu trois échecs successifs pour un même nombre de chiffres.

- pour les répétitions de phrases, on arrête également après trois échecs consécutifs (mais ici les phrases à répéter ont toutes une longueur différente, croissant de 2 en 2 syllabes).

- pour les "questions d'intelligence", on arrête l'épreuve en cas d'échec aux trois premières sous-questions, les cinq dernières étant d'une difficulté beaucoup plus grande.

Par contre, il peut arriver que le facteur temps s'ajoute au facteur échec pour déterminer l'arrêt de l'épreuve : pour l'ensemble des trois essais successifs d'ordination de cinq poids, la version originale du test préconise d'arrêter l'épreuve au bout de trois minutes [Binet 1911 : 85].

e) Règles d'enchaînement des questions

En principe, "il n'est nullement nécessaire d'observer l'ordre hiérarchique, suivant lequel les questions sont disposées dans le tableau qui sert d'échelle" (c'est-à-dire : la "feuille de niveau") [Binet 1911 : 35]. En pratique toutefois, l'usage s'est établi de procéder à l'administration du test en deux étapes :

① "Comme certaines épreuves servent, par des complications progressives, à des âges différents (ainsi les questions sur les images ou la répétition des chiffres ...), ces épreuves sont particulièrement commodes pour déterminer rapidement la zone d'opération. Il y a donc avantage à y procéder d'abord et à les parcourir entièrement dès qu'on les aborde [...] Reportant au fur et à mesure sur la feuille de niveau de l'enfant examiné les résultats obtenus, on voit d'un coup d'oeil où se fait le passage des signes + aux signes -. Or c'est évidemment dans cette zone de transition que se trouvera le niveau mental du sujet. C'est donc là qu'il le faut chercher.

② Sur une deuxième feuille d'interrogatoire, les questions sont classées par âge. Guidé par les résultats de la première feuille on choisira les épreuves destinées à compléter ce que nous appellerons volontiers « l'âge de base », où l'on a chance de voir le sujet réussir toutes les épreuves de cet âge. Ensuite, par tâtonnements (et quelquefois en revenant en arrière), on fixera la limite où le sujet s'arrête définitivement" [Binet 1911 : 36-37].

Ainsi, comme certains des questionnaires d'enquête que nous avons (partiellement) décrits, le BINET-SIMON se subdivise en deux sous-questionnaires : le premier, qui est en principe administré intégralement à tous les sujets, oriente ceux-ci, d'après leurs réponses, vers telle ou telle partie du second ; ce dernier, qui au contraire du précédent n'est jamais administré en entier, peut être considéré à la rigueur comme une collection de neuf sous-questionnaires (un par "âge mental" estimé à partir des questions de sondage). Mais cette manière de voir, quoique tout à fait fondée, risque de conduire à une confusion. En effet, chacun de ces neuf sous-questionnaires ne se confond pas avec la liste des questions qui correspondent à l' "âge mental" considéré ; il comprend non seulement l'ensemble de ces questions, mais également (au minimum) celles des "âges mentaux" qui précèdent ou suivent immédiatement l' "âge mental" considéré. En raison de la grande souplesse des règles qui président à l'ordre d'administration des questions de la seconde partie du questionnaire, il serait peu éclairant de comparer la structure générale du BINET-SIMON à celle d'un questionnaire d'enquête même de structure relativement compliquée, comme celui de l' "Enquête image presse" (*SOPRES 1094* ; cf. § 4.1.1.c, et tableau n° 4.1).

Il est plus simple de considérer la seconde partie du test comme un sous-questionnaire unique, dont l'administration est régie par un ensemble de consignes assez complexes.

L'administration de la feuille de sondage est séquentielle (chaque question peut selon SIMON, doit selon ZAZZO, être posée dans l'ordre de son apparition sur la feuille), sauf en un point : l'épreuve "rendre la monnaie" (sondage V 4, *BS Z* : 33) n'est pas proposée au sujet s'il a échoué aux épreuves antérieures sur la monnaie (On notera l'imprécision de la consigne originale : s'agit-il à la fois des épreuves de comptage d'objets ou de pièces, et de reconnaissance des pièces de monnaie, ou seulement de celles qui utilisent nécessairement des pièces, ou seulement de celle de reconnaissance ? Ce qui complique encore le problème, c'est que les enfants de 4 ans d'âge réel réussissent à "rendre la monnaie" légèrement mieux qu'à "reconnaître toutes les pièces de notre monnaie" ; cf. [Zazzo 1966 : I 59]). Le rôle structural de la feuille de sondage est de déterminer la question de la seconde partie du test à poser en premier. La consigne d'origine est relativement floue, et autorise de nombreuses variantes. Dans la révision ZAZZO de 1949, la procédure à appliquer est entièrement déterminée : après avoir reporté sur la feuille de niveau les résultats de la passation des questions de sondage, il suffit de repérer l' "âge mental" le plus élevé tel que tous les items de sondage de cet âge et des âges inférieurs ont été réussis par le sujet ; la première question à poser est celle qui commence la série correspondant à cet "âge mental de base".

L'avantage du classement des items du BINET-SIMON par ordre croissant de difficulté simplifie beaucoup l'administration de la seconde partie du questionnaire. En principe, il devrait même suffire de partir de l'item de sondage le plus difficile que le sujet ait réussi, et d'arrêter la passation au premier échec rencontré, cet échec déterminant le niveau atteint par le sujet. Une telle procédure se révélerait cependant tout à fait irréaliste pour au moins trois groupes de raisons :

- il peut arriver qu'un enfant échoue à une épreuve, qu'il serait en d'autres circonstances capable de réussir, par cause de distraction, timidité, perturbation de l'examen, etc. ; réciproquement, un enfant peut réussir une

épreuve un peu supérieure à son niveau s'il y a été préalablement entraîné (cf. les manuels d'entraînement aux tests d'aptitude ou de niveau, comme p.ex. [A. 1965]).

- l'ordre de difficulté "moyen" (pour l'ensemble de l'échantillon sur lequel la révision ZAZZO a été étalonnée) ne coïncide pas tout à fait avec l'ordre de difficulté pour une classe d'âge donnée ; par exemple, l'item qui occupe le 19e rang dans l'échelle globale ("distinguer matin et soir", III 5) se classe au 12e rang seulement de difficulté pour les enfants de 3 ans (cf. [Zazzo 1966 : I 61-62]).

- dans toute échelle composite comme le BINET-SIMON, les aptitudes mises en jeu peuvent différer d'une épreuve à l'autre. En conséquence, l'ordre de difficulté des épreuves peut n'être pas identique pour deux sujets, même si ceux-ci ont en définitive le même "âge mental" ; par exemple, l'épreuve "énumérer les mois" (IV, BS Z : 32) est plus facile qu'en moyenne pour les débiles, et plus difficile qu'en moyenne pour les dyslexiques (cf. [Zazzo 1966 : I 56, 62-63]).

C'est pourquoi la procédure d'application du test consiste à :

① définir l' "âge de base", c'est-à-dire la catégorie (de la feuille de niveau) la plus élevée pour laquelle on n'a relevé aucun échec aux items de sondage ; c'est par définition la catégorie d'âge immédiatement inférieure à la catégorie du premier échec enregistré.

② poser au sujet toutes les questions correspondant à cet âge de base.

③ en cas d'échec à au moins une question de cette catégorie, poser les questions de la catégorie immédiatement inférieure, jusqu'à ce que tous les items d'une catégorie d'âge aient été réussis.

④ dans tous les cas, poser toutes les questions supérieures à l'âge de base dans l'ordre croissant de difficulté (on arrête l'épreuve selon les spécifications ci-après).

Les règles pour arrêter l'épreuve sont assez différentes, dans la révision ZAZZO de 1949, des règles de la version originale de BINET et SIMON .

Elles considèrent trois cas :

- si le sujet a échoué à au moins un des items inférieur à 10 ans, arrêter l'épreuve après échec à tous les items de deux groupes d'âge successifs ;

- sinon (si tous les items jusqu'à 9 ans inclus ont été réussis), et si aucun item de sondage supérieur à 12 ans n'a été réussi, arrêter l'épreuve après échec à cinq items successifs ;

- sinon (si au moins un item de sondage supérieur à 12 ans a été réussi), poser toutes les questions inférieures à l'item réussi le plus élevé, et continuer l'application du test jusqu'à l'enregistrement de cinq échecs successifs (ou jusqu'à épuisement des questions).

3.2.3. Description formelle des questionnaires

Nous considérerons qu'un questionnaire au sens large est constitué par (cf. [Flament 1976 : 60-61]) :

- un ensemble fini de t questions $Q = \{q_i\}$, avec $i \in K$ et $K = \{1, 2, \dots, t\}$;

- une famille $\{R_i\}$, $i \in K$, de t ensembles finis R_i , dont chacun est l'ensemble des réponses possibles au sens large à la question q_i ;

- des règles de questionnement (sur lesquelles nous reviendrons au § 3.3.1.).

Ce cadre formel nous permet de reprendre de façon plus systématique et plus rigoureuse la description de la structure des tests et des questionnaires d'enquête, en distinguant trois types de relations binaires : les relations entre questions dépendant des règles d'administration du questionnaire, les relations entre les questions et les ensembles de réponses possibles, et enfin les relations entre questions liées à des considérations sémantiques. Il nous a paru intéressant de traduire ces relations en termes de *graphes*. Par exemple, le graphe d'une relation binaire R entre questions est le couple $G_R = (Q, \Gamma_R)$, dont le premier élément est l'ensemble Q des questions (appelé l'ensemble des *sommets* du graphe), et le second l'ensemble des couples de Q^2 vérifiant la relation R établie dans Q (les éléments de Γ_R sont appelés *arcs*). Le *graphe topologique* est la représentation d'un graphe dans un plan, en figurant les sommets par des points, et les arcs par des flèches reliant les sommets pour lesquels la relation est vérifiée ; tous les graphes correspondant aux relations que nous avons ainsi définies sont des 1-graphes d'ordre $|Q|$ (nous adoptons pour l'essentiel la terminologie de [Berge 1970] et de [Picard 1972]). Il aurait été agréable d'illustrer chacune des relations que nous avons définies en présentant le graphe topologique de cette relation pour un questionnaire réel. Malheureusement, les dimensions d'un tel graphe interdisent sa présentation en une double page au maximum, même pour un questionnaire très court. C'est pourquoi nous avons construit un questionnaire fictif d'une dizaine de questions seulement, dont la structure est assez complexe pour nous fournir au moins un exemple de chacune de ces relations ; ce questionnaire, tout à fait artificiel, et sans intérêt pratique autre que celui d'être l'occasion d'illustrer les plus importantes des relations décrites, est présenté dans le tableau n° 3.12.

a) Relations entre questions dépendant de la consigne

Il est possible de définir dans Q^2 cinq types de relations simples distincts : relations de position, de précession, de dominance, d'implication, et d'incompatibilité. Nous en donnons ci-après brièvement la définition, ainsi que les principales propriétés formelles.

- Q 1. Croyez-vous en l'existence de Dieu ?
- oui
 - non
 - ne sait pas, sans opinion] → *Passer à Q 7*
- Q 2. Assistez-vous régulièrement aux offices religieux ?
- oui → *Passer à Q 4*
 - non
 - ne peut pas dire
- Q 3. Dans quelles circonstances vous arrive-t-il d'assister à un office religieux ? (*Présenter liste. Plusieurs réponses possibles*)
- à l'occasion de fêtes religieuses (Pâques, Noël, ...)
 - à l'occasion de cérémonies privées (mariages, baptêmes, enterrements, ...)
 - quand j'en ressens le besoin
 - cela ne m'arrive jamais → *Passer à Q 5*
- Q 4. Parmi les raisons figurant sur cette liste (*Montrer*), laquelle s'applique le mieux à votre cas personnel ?
(*Une seule réponse possible*)
- "Ce que m'apporte la pratique religieuse, c'est avant tout :
- un retour sur moi-même
 - un sentiment d'exaltation et d'oubli de soi
 - une impression de communion avec la Divinité
 - la conscience d'appartenir à la communauté des croyants"
 - aucune de ces raisons
- *Passer à Q 10*
- Q 5. Avez-vous reçu une éducation religieuse ?
- oui
 - non
 - ne peut pas dire
- Q 6. Avez-vous déjà assisté à des offices religieux ?
- oui, plusieurs fois
 - oui, une fois ou deux
 - non, absolument jamais] → *Passer à Q 10*

Q 7. Assistez-vous régulièrement aux offices religieux ?

- oui
- non
- ne peut pas dire] → Passer à Q 9

Q 8. Parmi les raisons figurant sur cette liste (*Montrer*), laquelle s'applique le mieux à votre cas personnel ?
(*Une seule réponse possible*)

"Ce que m'apporte la pratique religieuse, c'est avant tout :

- | | | | | |
|-------|---|---|---|---------------|
| Liste | } | - un retour sur moi-même | → | Passer à Q 10 |
| | | - la conscience d'appartenir à une communauté | | |
| | | - la satisfaction d'accomplir un devoir | | |
| | | - le sentiment de donner l'exemple à mon entourage" | | |
| | | - aucune de ces raisons | | |

Q 9. Avez-vous reçu une éducation religieuse ?

- oui
- non
- ne peut pas dire

Q 10. Nous allons maintenant passer en revue tous les membres de votre famille qui vivent actuellement avec vous.
Pour chacun d'eux, je voudrais que vous me disiez s'il a été baptisé ou non.

Degré de parenté	Baptême
- vous-même	oui non NSP
-	oui non NSP
-	oui non NSP
-	oui non NSP

Q 11. Accepteriez-vous d'être interrogé plus longuement sur diverses questions touchant aux problèmes religieux ?

- oui
- non
- NSP

Fin d'interview

Tableau n° 3.12 : Questionnaire d'enquête fictif destiné à illustrer les relations constituant la structure formelle d'un questionnaire au sens large.

La relation de *position* correspond à l'indication des questions (et en général à leur situation sur le support matériel du questionnaire). C'est une relation d'ordre strict total (antiréflexive, antisymétrique, transitive), qui a pour fonction de permettre l'identification des questions par référence à un ensemble d'étiquettes totalement ordonné. Si, dans le questionnaire, une question q_i est placée avant une question q_j , nous avons nécessairement : $i < j$, et réciproquement. Ceci est un corollaire de la règle, implicite dans les consignes d'administration de tout questionnaire, qui veut qu'en l'absence d'indication contraire (aiguillage ou saut inconditionnel), les questions soient posées dans l'ordre de leur numérotation. Le graphe de la relation de position : $G_{<} = (Q, \Gamma_{<})$ est un graphe complet. Il peut être intéressant sur le plan pratique de définir également une relation de *position immédiate* entre deux questions q_i et q_j : on dira que q_i est placée immédiatement avant q_j si et seulement si q_i est placée avant q_j , et s'il n'existe pas de question q_k telle que q_i soit placée avant q_k et q_k avant q_j . Cette relation est antiréflexive, antisymétrique, et antitransitive. Son graphe se réduit au chemin hamiltonien $\mu = [q_1, q_2, \dots, q_t]$, dont la fermeture transitive est le graphe de la relation de position au sens large. Cette relation n'a pas une grande importance théorique. Par contre, son intérêt pratique est certain, puisqu'elle sert à nommer les questions, et à les ordonner totalement (comme l'ordre lexicographique ordonne totalement les mots d'un dictionnaire).

La relation de *précession* exprime l'ordre dans lequel les questions sont effectivement posées. On dira que q_i précède q_j , et l'on écrira : $q_i P q_j$, si la question q_i est effectivement posée avant la question q_j . Cette relation est irréflexive, asymétrique, et transitive.

En effet, lorsqu'une question est réitérée au sein d'un même questionnaire, la relation peut être réflexive (dans notre exemple, on a : $q_{10} P q_{10}$) ; dans le cas général toutefois, la relation est antiréflexive. De même, lorsqu'une série de questions $[q_i, \dots, q_k]$ est réitérée, on peut avoir simultanément : $q_i P q_k$ et : $q_k P q_i$; mais dans le cas général, la relation est antisymétrique. Enfin, on a toujours, quels que soient i, j , et k :

$$q_i P q_j \quad \& \quad q_j P q_k \quad \Rightarrow \quad q_i P q_k$$

Dans un questionnaire purement séquentiel, la relation de précession coïncide avec la relation de position. Dans un questionnaire sans retour en arrière (mais avec éventuellement des aiguillages vers l'aval), la relation de précession est antiréflexive et antisymétrique ; si une question q_i est posée avant une question q_j , sa position dans le questionnaire est avant q_j . Plus brièvement, puisque l'antisymétrie entraîne l'antiréflexivité, on peut écrire (en posant : $q_j \bar{P} q_i = q_j$ ne précède pas q_i) :

$$\forall q_i \in Q, \forall q_j \in Q (q_i P q_j \quad \& \quad q_j \bar{P} q_i \Rightarrow i < j)$$

Dans le cas général (questionnaire avec sauts vers l'aval et retours en arrière), il n'y a pas de relation systématique entre position et précession. Le graphe de la relation de précession est un graphe transitif, quasi-fortement connexe inférieurement ; dans le cas général, il n'est pas planaire.

Il nous sera très commode de disposer d'une relation de *précession immédiate*, notée P_i , telle que : $q_i P_i q_j$ signifie : lorsque q_i et q_j sont posées à une même personne, q_i peut être posée immédiatement avant q_j . Cette relation est irréflexive et asymétrique comme la relation de précession au sens large, et pour les mêmes raisons. Elle est de plus intransitive. En effet, dans la plupart des cas, on constate que la précession immédiate est antitransitive. Il faut se garder toutefois d'en conclure qu'il s'agit d'une propriété de P_i . En effet, dans les aiguillages obtenus par combinaison de questions-filtres, l'antitransitivité n'est pas vérifiée : dans l'exemple

utilisé au § 3.2.1.b (SOFRES 340 : 3-4), on a simultanément les relations :

$$\left\{ \begin{array}{l} q_3 \text{ Pi } q_{3a} \\ q_{3a} \text{ Pi } q_4 \\ q_3 \text{ Pi } q_4 \end{array} \right.$$

Ce cas particulier justifie la précaution prise dans la définition de la relation de précession immédiate ($q_i \text{ Pi } q_j \Leftrightarrow q_j$ peut être posée immédiatement avant q_i). Dans un questionnaire purement séquentiel, la relation de précession immédiate coïncide avec la relation de position immédiate. Dans un questionnaire sans retour en arrière, si q_i précède immédiatement q_j , i est avant j :

$$q_i \text{ Pi } q_j \quad \& \quad q_j \overline{\text{Pi}} q_i \Rightarrow i < j$$

Dans le cas général, la précession immédiate implique la précession au sens large ; quels que soient i et j , on a :

$$q_i \text{ Pi } q_j \Rightarrow q_i \text{ P } q_j$$

Le graphe G_{Pi} de la relation de précession immédiate a une racine et une seule (la première question posée) ; il peut par contre avoir plusieurs sommets terminaux : il est donc en général seulement quasi-fortement connexe inférieurement. C'est un graphe partiel de G_{P} , puisque, de la relation d'implication ci-dessus, il s'ensuit que : $\Gamma_{\text{Pi}} \subset \Gamma_{\text{P}}$; en outre, G_{P} est la fermeture transitive stricte de G_{Pi} .

La relation de précession immédiate a une grande importance pratique. En effet, le graphe topologique correspondant à G_{Pi} est l'*organigramme* du questionnaire (cf. figure n° 3.13). Lorsque G_{Pi} a plusieurs sommets terminaux, on peut lui préférer le graphe $G'_{\text{Pi}} = (Q \cup \text{fin}, \Gamma_{\text{Pi}})$, obtenu en ajoutant à Q un sommet supplémentaire, appelé *fin*, et qui est tel que,

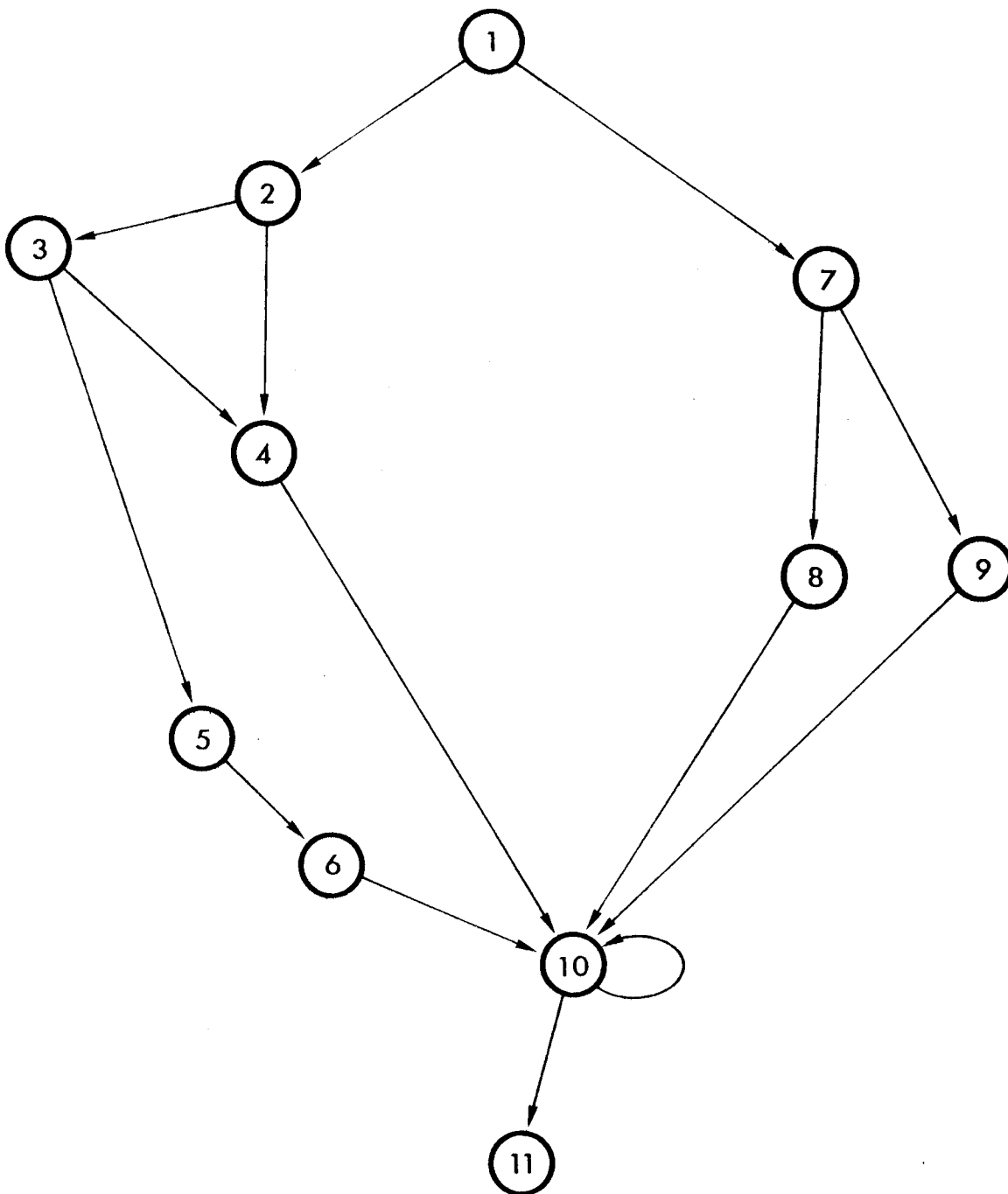


Figure n° 3.13 : Graphe topologique de la relation de précession immédiate dans le questionnaire fictif du tableau n° 3.12.

pour tout sommet terminal q_t de G_{P_i} , on a dans G'_{P_i} la relation : $q_t \text{ Pi fin}$. Ce graphe a donc un seul sommet terminal ; G'_{P_i} est donc quasi-fortement connexe. L'organigramme ainsi défini met en évidence certaines propriétés formelles d'un questionnaire ; en particulier, lorsque G'_{P_i} :

- se compose d'un *seul chemin* de longueur $|Q|$ unissant la racine au sommet terminal, le questionnaire est purement séquentiel ;
- comporte des *circuits*, le questionnaire présente des retours en arrière (réitérations de questions) ;
- est un *treillis*, le questionnaire comporte des questions-filtres.

En outre, chaque chemin dans G_{P_i} ou dans G'_{P_i} , qui a pour origine la première question q_1 et pour extrémité un sommet terminal, est un *scénario* de l'administration du questionnaire : il correspond à une variante dans l'administration des questions, selon les réponses aux questions-filtres de la personne interrogée. L'ensemble des scénarii distincts représente par conséquent l'ensemble des déroulements possibles de l'entretien d'enquête ; il peut être figuré par un arbre de racine q_1 (cf. figure n° 3.14). La notion de scénarii est nécessaire pour définir le contexte d'une question pour un répondant donné.

La relation de *dominance* (déjà évoquée au § 3.2.1.d) régit les rapports entre une question-filtre et les divers sous-ensembles de questions vers lesquelles elle oriente les personnes interrogées. Nous dirons qu'une question-filtre q_i *domine* une question quelconque q_j , et nous écrirons : $q_i \text{ D } q_j$, si et seulement si le fait que la question q_j soit posée à l'interviewé dépend de la réponse de celui-ci à q_i . La relation de dominance est antiréflexive, antisymétrique, et transitive. En effet, une question ne peut se dominer elle-même :

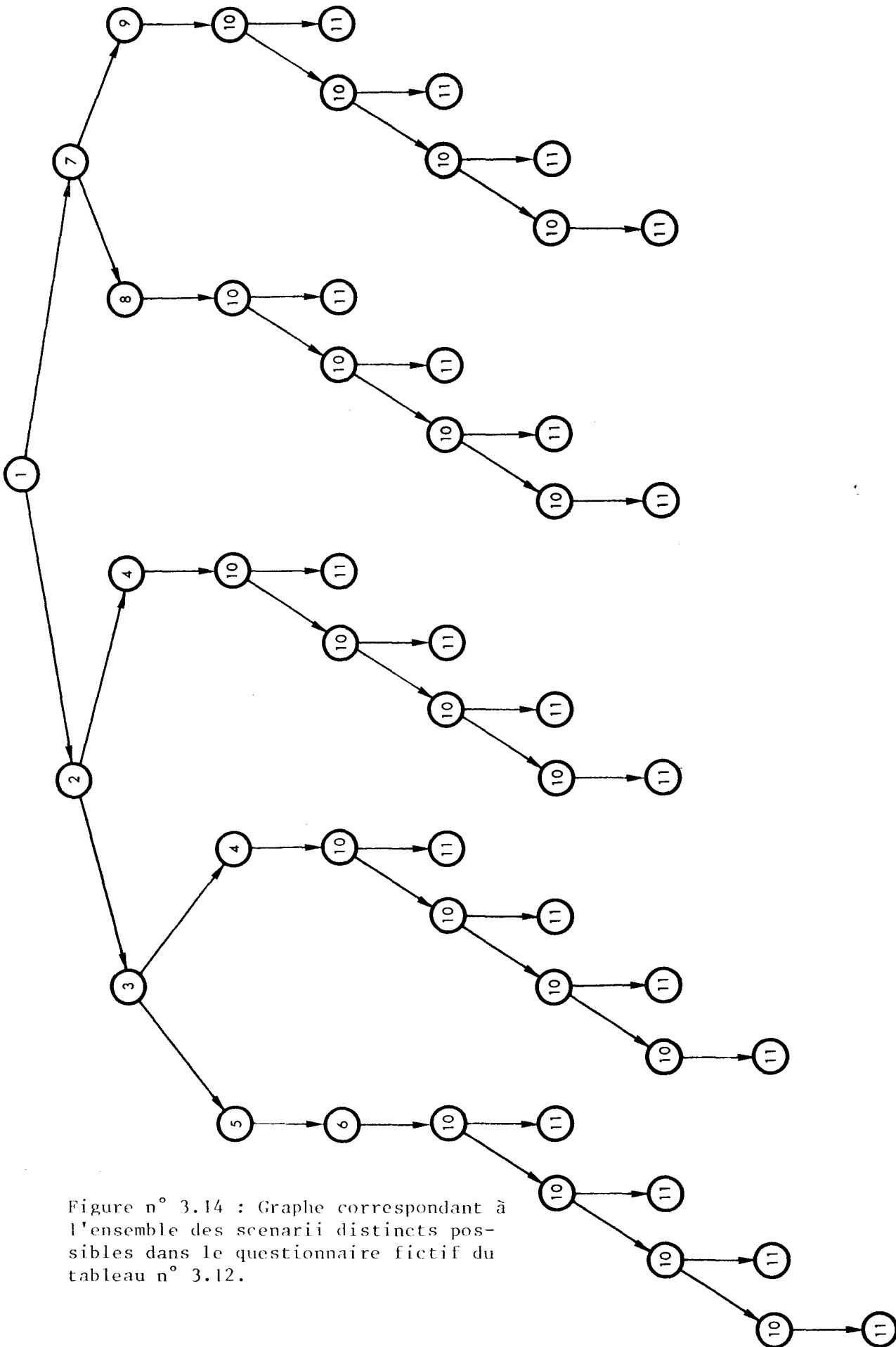


Figure n° 3.14 : Graphe correspondant à l'ensemble des scénarii distincts possibles dans le questionnaire fictif du tableau n° 3.12.

$$\forall q_i \in Q \quad (q_i \bar{D} q_i)$$

Si une question en domine une autre, elle ne peut être dominée par celle-ci :

$$\forall q_i \in Q, \forall q_j \in Q \quad (q_i D q_j \Rightarrow q_j \bar{D} q_i)$$

Si q_i domine q_j qui domine q_k , alors q_i domine q_k :

$$\forall q_i \in Q, \forall q_j \in Q, \forall q_k \in Q \quad (q_i D q_j \ \& \ q_j D q_k \Rightarrow q_i D q_k)$$

Cette relation entretient avec les relations de position et de précession les rapports suivants :

- si q_i domine q_j , q_i est placée avant q_j :

$$\forall q_i \in Q, \forall q_j \in Q \quad (q_i D q_j \Rightarrow i < j)$$

- si q_i domine q_j , alors q_i est posée avant q_j :

$$\forall q_i \in Q, \forall q_j \in Q \quad (q_i D q_j \Rightarrow q_i P q_j)$$

Le graphe $G_D = (Q, \Gamma_D)$ de la relation de dominance (cf. figure n° 3.15) est par conséquent un graphe partiel du graphe de position $G_<$ (on a : $\Gamma_D \subset \Gamma_<$) et du graphe de précession G_P (on a : $\Gamma_D \subset \Gamma_P$). Il s'ensuit que G_D est un graphe partiel de leur réunion : $G_D \subset G_< \cup G_P$. En outre, puisque l'on a, pour tout couple de sommets, la propriété :

$$q_i D q_j \Rightarrow i < j \ \& \ q_i P q_j$$

on en déduit que :

$$\Gamma_D \subset \Gamma_< \cap \Gamma_P$$

Le graphe de dominance est donc un graphe partiel de l'intersection des graphes de position et de précession :

$$G_D \subset G_< \cap G_P$$

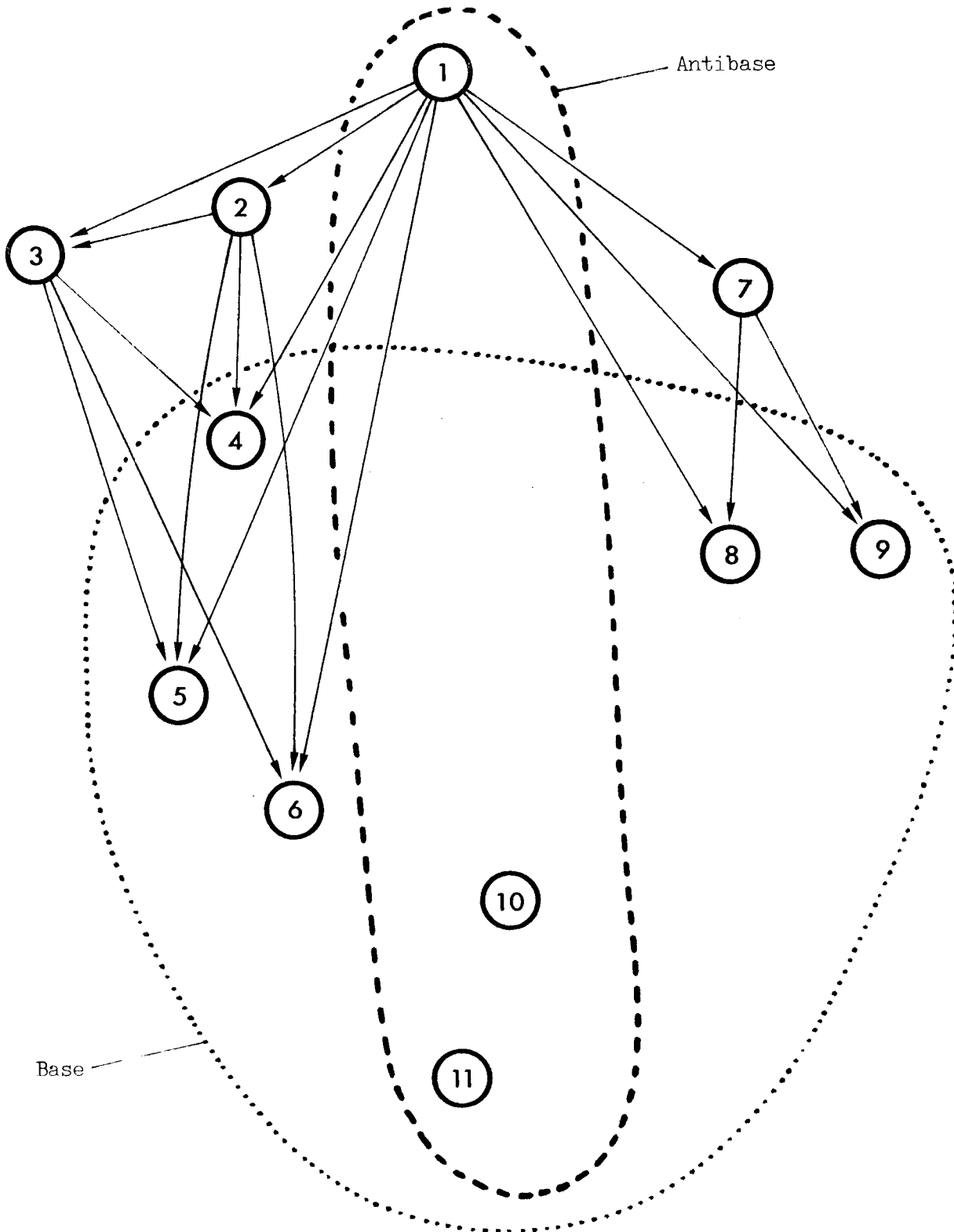


Figure n° 3.15 : Graphe topologique de la relation de dominance dans le questionnaire fictif du tableau n° 3.12.

Les propriétés du graphe de dominance nous ramènent à la classification des questions selon leur rôle dans le questionnaire (cf. § 3.2.1.d). En effet, G_D étant transitif et antisymétrique, il est sans circuit. Or, tout graphe fini sans circuit admet une *base* unique ; tout graphe strictement transitif admet un *noyau* ; G_D a donc une base unique et un noyau. En outre, G_D étant à la fois sans circuit et fermé transitivement, il y a identité de sa base et de son noyau. Enfin, G_D étant un graphe fini sans circuit, il admet une *antibase* unique. On montre facilement que :

- le noyau (et la base) de G_D est l'ensemble $S \subset Q$ des questions purement informatives du questionnaire. En effet, l'ensemble S des questions purement informatives est stable (intérieurement) puisque, par définition, une question purement informative ne peut en dominer une autre ; l'ensemble S est absorbant (stable extérieurement), puisque toute question-filtre (n'appartenant pas à S par définition) domine au moins une question purement informative (élément de S). Donc S est bien la base (et le noyau) de G_D . Dans notre exemple, $S = \{ q_4, q_5, q_6, q_8, q_9, q_{10}, q_{11} \}$.

- l'antibase de G_D est l'ensemble $T \subset Q$ des questions omnibus. En effet, par définition, il n'existe pas de relation de dominance entre les éléments de T ; et toute question élément de Q n'appartenant pas à T est dominée par au moins une question-filtre indépendante (non dominée elle-même). Dans notre exemple, $T = \{ q_1, q_{10}, q_{11} \}$.

A partir de la base S et de l'antibase T de G_D , et de leurs complémentaires dans Q (respectivement \bar{S} et \bar{T}), nous pouvons définir les huit sous-ensembles de Q suivants :

- S = ensemble des questions purement informatives ;
- T = ensemble des questions omnibus (indépendantes) ;
- \bar{S} = ensemble des questions-filtres ;
- \bar{T} = ensemble des questions dominées ;
- $\bar{S} \cap \bar{T}$ = ensemble des questions-filtres indépendantes ;
- $\bar{S} \cap T$ = ensemble des questions-filtres dominées ;
- $S \cap T$ = ensemble des questions purement informatives indépendantes ;
- $S \cap \bar{T}$ = ensemble des questions purement informatives dominées.

Dans notre exemple, il y a une question-filtre indépendante : q_1 ; trois questions-filtres dominées : q_2 , q_3 , q_7 ; deux questions purement informatives indépendantes (omnibus) : q_{10} , q_{11} ; et cinq questions purement informatives dominées : q_4 , q_5 , q_6 , q_8 , q_9 .

Certaines propriétés de G_D nous renseignent sur les particularités du questionnaire correspondant. Par exemple :

- si G_D est connexe, toutes les questions indépendantes (omnibus) sont des questions-filtres. En effet, s'il existait une question indépendante q_i qui ne soit pas question-filtre, il existerait au moins un couple de questions $(q_i, q_j) \in Q^2$ avec $i > j$, tel que : $q_i \bar{D} q_j$. Or, nous avons par définition : $\forall j > i, q_j \bar{D} q_i$; comme G_D est fermé transitivement, cela signifie qu'il existerait au moins un couple de questions qui ne sont liées par aucune chaîne, c'est-à-dire que le graphe ne serait pas connexe (la réciproque de ce théorème n'est pas vérifiée).

- si G_D est quasi-fortement connexe inférieurement, la première question posée domine toutes les autres, et réciproquement. En effet, la propriété de G_D signifie qu'il existe une question qui domine toutes les autres (racine du graphe) ; puisque : $\forall (q_i, q_j) \in Q^2$ ($q_i D q_j \Rightarrow q_i P q_j$), c'est nécessairement la première question posée qui est la racine de G_D .

On pourrait évidemment définir une relation de *dominance immédiate* D_i . Cette relation est antiréflexive, antisymétrique, intransitive. Son graphe G_{D_i} est un graphe partiel de G_{P_i} ; mais, dans le cas général, il n'a pas pour fermeture transitive G_D . Cette relation ne semble pas présenter d'intérêt pratique.

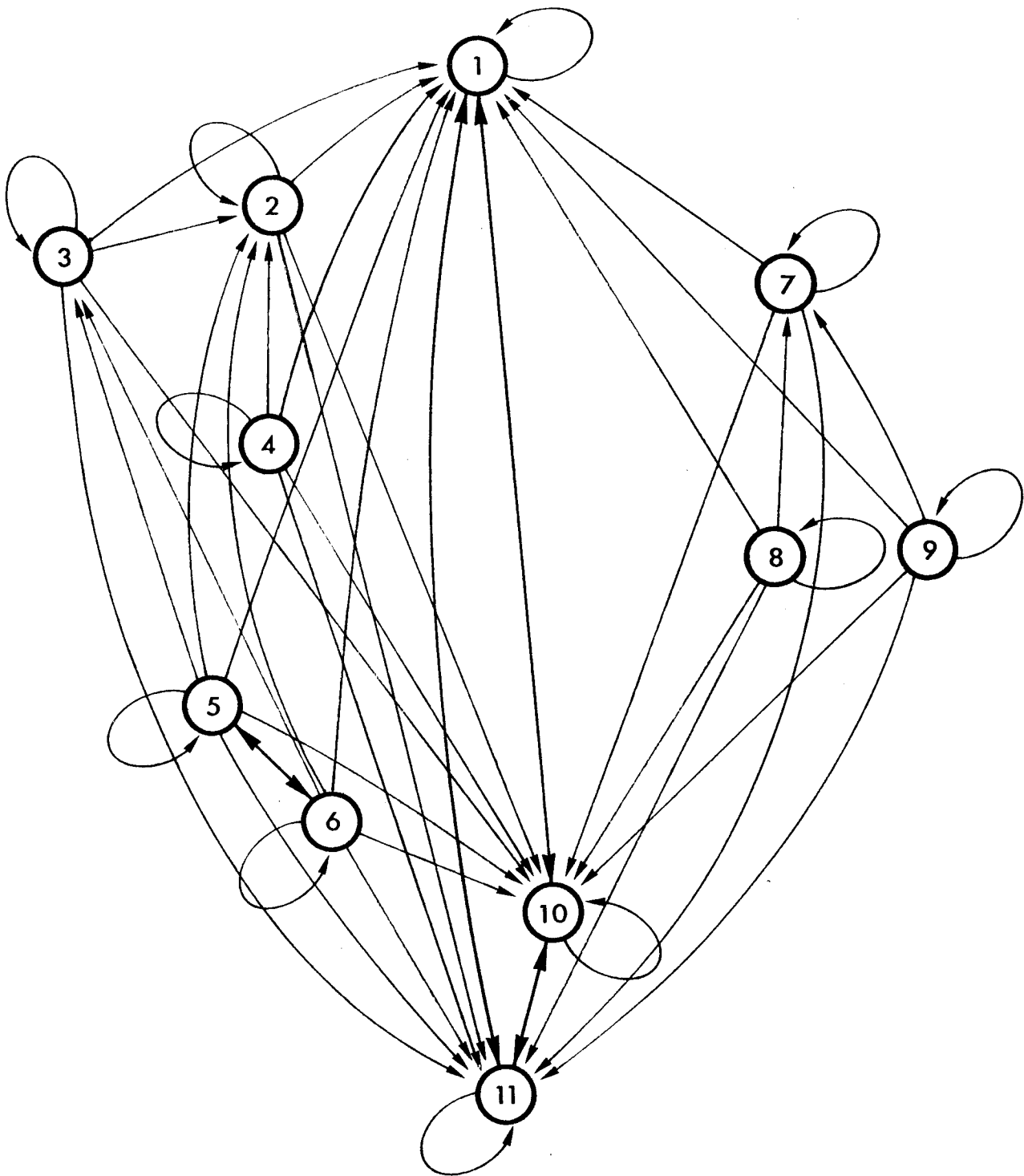


Figure n° 3.16 : Graphe topologique de la relation d'implication dans le questionnaire fictif du tableau n° 3.12.

La relation d'*implication* : $q_i \rightarrow q_j$ signifie : si l'on pose q_i à un répondant, alors on doit poser aussi q_j . C'est une relation réflexive, asymétrique, transitive. Par définition, une question omnibus est impliquée par toutes les questions. L'ensemble des questions q_k telles que $q_i \rightarrow q_k$ constitue le contexte global (amont et aval) de la question q_i . Le graphe topologique de la relation d'implication correspondant à la structure de notre questionnaire fictif est représenté par la figure n° 3.16. On y remarque quatre couples de questions pour lesquelles l'implication est réciproque : (q_1, q_{10}) , (q_1, q_{11}) , (q_5, q_6) , et (q_{10}, q_{11}) .

La relation d'*incompatibilité* : $q_i \mid q_j$, signifie : q_i et q_j ne peuvent jamais être posées à une même personne au cours d'un entretien. Cette relation est antiréflexive, symétrique, intransitive (cf. le graphe topologique de la figure n° 3.17).

b) Relations entre questions et réponses

Soit une question quelconque : $q_i \in Q$, et l'ensemble : $R_i = \{ r_{i,l} \}$, $l \in L = \{ 1, \dots, l, \dots, m \}$ des m réponses $r_{i,l}$ à cette question. Dans le contexte de l'analyse secondaire des réponses au questionnaire, R_i est nécessairement fini (puisque les réponses aux questions ouvertes ont été codées selon un ensemble fini de catégories). Dans le produit cartésien $R_i \times Q$, nous définissons trois types de relations : les relations d'appartenance, de pointage, et de sélection. Les graphes correspondant à ces relations sont des 1-graphes bipartis, d'ordre $|R_i| + |Q|$. Si nous considérons l'ensemble des réponses à l'ensemble des questions, les relations sont définies dans le produit cartésien : $(\bigcup_{i \in K} R_i) \times Q$. Les graphes correspondant sont des 1-graphes bipartis d'ordre : $\sum_{i \in K} |R_i| + |Q|$. D'autres représentations sont toutefois possibles. En particulier, comme nous le verrons, au prix d'une identification de l'ensemble des réponses possibles R_i à la question q_i elle-même, les relations de pointage et de sélection peuvent être définies dans Q^2 , le graphe correspondant à ces relations étant alors un p -graphe d'ordre $|Q|$, avec :

$$p = \max_{i \in K} |R_i|.$$

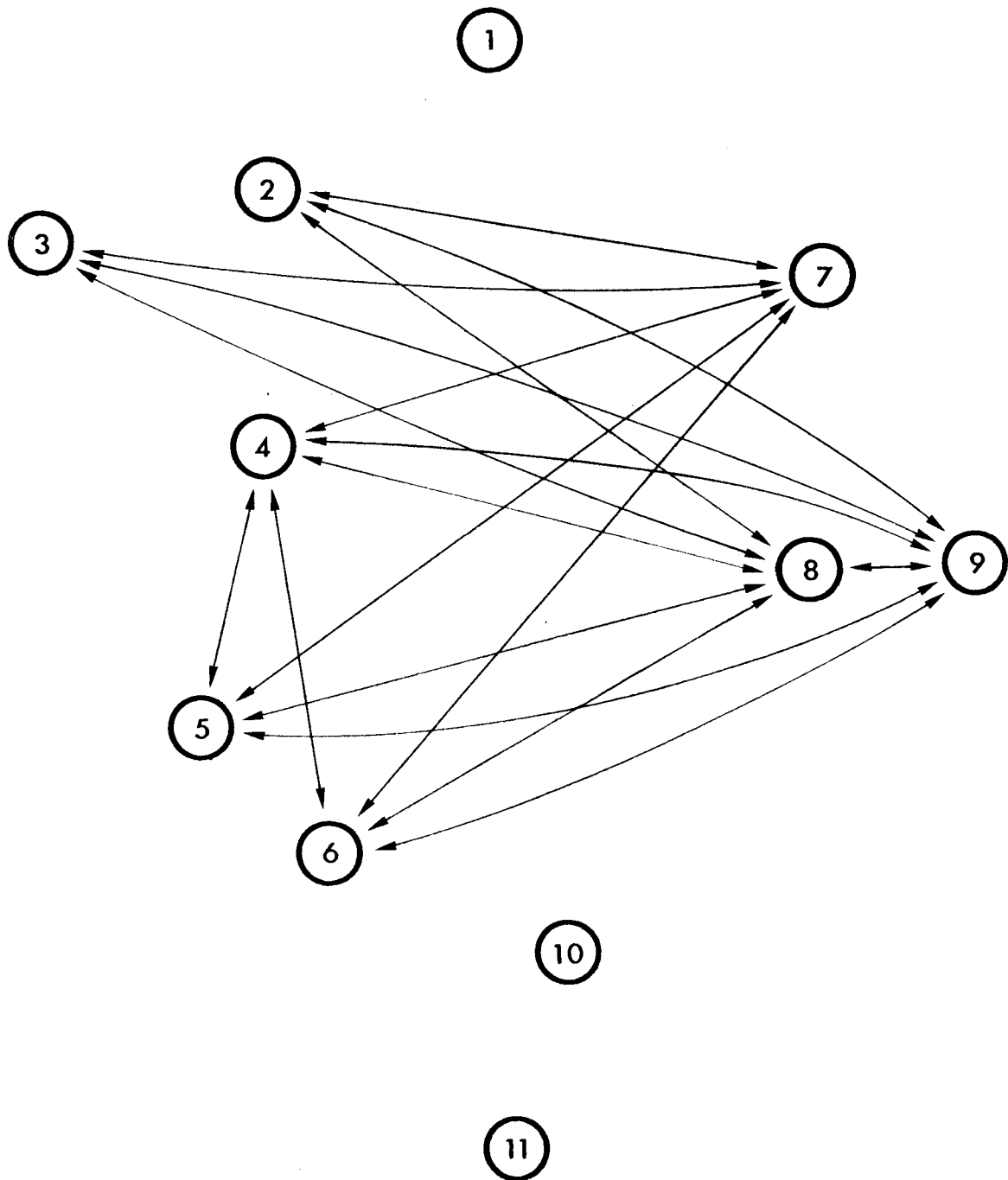


Figure n° 3.17 : Graphe topologique de la relation d'incompatibilité dans le questionnaire fictif du tableau n° 3.12.

La relation d'*appartenance* définit les liens qui unissent une question et les éléments de l'ensemble de ses réponses possibles. La correspondance entre la question q_i et les réponses $r_{i,l} \in R_i$ est telle que :

- pour toute question q_i il existe un ensemble de réponses possibles R_i et un seul ;
- à deux questions q_i et q_j distinctes correspondent deux ensembles R_i et R_j distincts ;
- à deux questions q_i et q_j identiques correspondent deux ensembles R_i et R_j identiques (nous reviendrons sur cette notion d'identité au § 3.2.3.c).

Il y a donc correspondance bi-univoque entre une question et l'ensemble de ses réponses possibles. Nous appelons *appartenance* d'une réponse $r_{i,l} \in R_i$ à une question q_i la composée de la relation injective $r_{i,l} \rightarrow R_i$ et de la bijection $R_i \leftrightarrow q_i$. Nous dirons donc qu'une réponse $r_{i,l}$ appartient à une question q_i , si et seulement si $r_{i,l} \in R_i$. Une telle définition est évidemment un abus de langage, à moins que l'on admette que $R_i = q_i$. Cette thèse a été (nous l'avons vu aux § 2.3.4.c et d) soutenue par Gerold STAHL et David HARRAH ; elle a été vigoureusement combattue par C.L. HAMBLIN (cf. [Hamblin 1963]). Sans prétendre trancher le débat, remarquons avec André WARUSFELD que la relation d'égalité est souvent utilisée abusivement par les mathématiciens, à la place de la relation de simple équivalence, sans que toutes les propriétés d'un élément soient vérifiées pour l'élément qui lui est prétendument égal [Warusfeld 1966 : 357] ; si l'on adopte le point de vue de HAMBLIN, il reste toutefois possible, par conséquent, de considérer R_i et q_i comme identiques d'un certain point de vue. Si l'on pose, pour alléger l'écriture : $\mathcal{R} = \bigcup_{i \in K} R_i$, le graphe $G_{\mathcal{E}^*} = (\mathcal{R}, Q, \Gamma_{\mathcal{E}^*})$ est une forêt. En effet, chacune des composantes connexes de $G_{\mathcal{E}^*}$ est un arbre : elle est quasi-fortement connexe supérieurement (puisque par définition : $\forall r_{i,l}, r_{i,l} \in^* q_i$), et sans cycle

(puisque E^* est antiréflexive et antitransitive). L'inverse du graphe d'appartenance, $G_{E^*}^{-1}$, est une forêt dont les composantes connexes sont des arborescences de racine q_i .

La relation de *pointage* est définie entre l'ensemble R_i des réponses possibles à une question q_i , et l'ensemble des questions qui sont posées immédiatement après q_i . Nous dirons qu'une réponse $r_{i,l} \in R_i$ pointe vers q_j , et nous écrirons : $r_{i,l} \mapsto q_j$, si et seulement si $q_i \text{ Pi } q_j$ et q_j est posée obligatoirement à toute personne ayant répondu $r_{i,l}$ à q_i . Réciproquement, si une réponse au moins de R_i pointe vers q_j , q_i est un des prédecesseurs immédiats de q_j .

Il peut arriver que toutes les réponses à une question q_i pointent vers une même question q_j ; dans ce cas, q_i est purement informative, et elle est l'un des prédecesseurs immédiats de q_j :

$$\begin{aligned} & (q_i, q_j) \in Q^2 \quad \& \quad \forall r_{i,l} \in R_i \quad (r_{i,l} \mapsto q_j) \\ \Rightarrow & q_i \text{ Pi } q_j \quad \& \quad \forall q_k \in Q \quad (q_i \bar{D} q_k) \end{aligned}$$

Il arrive également que certaines des réponses possibles à une question ne pointent vers aucune autre question. Sauf omission ou erreur dans la construction du questionnaire, ce cas correspond à une consigne du type : "si réponse x , fin d'interview". Par commodité d'écriture, nous introduirons une question fictive (déjà utilisée pour la relation de précession immédiate) notée *fin* $\notin Q$, et nous écrirons :

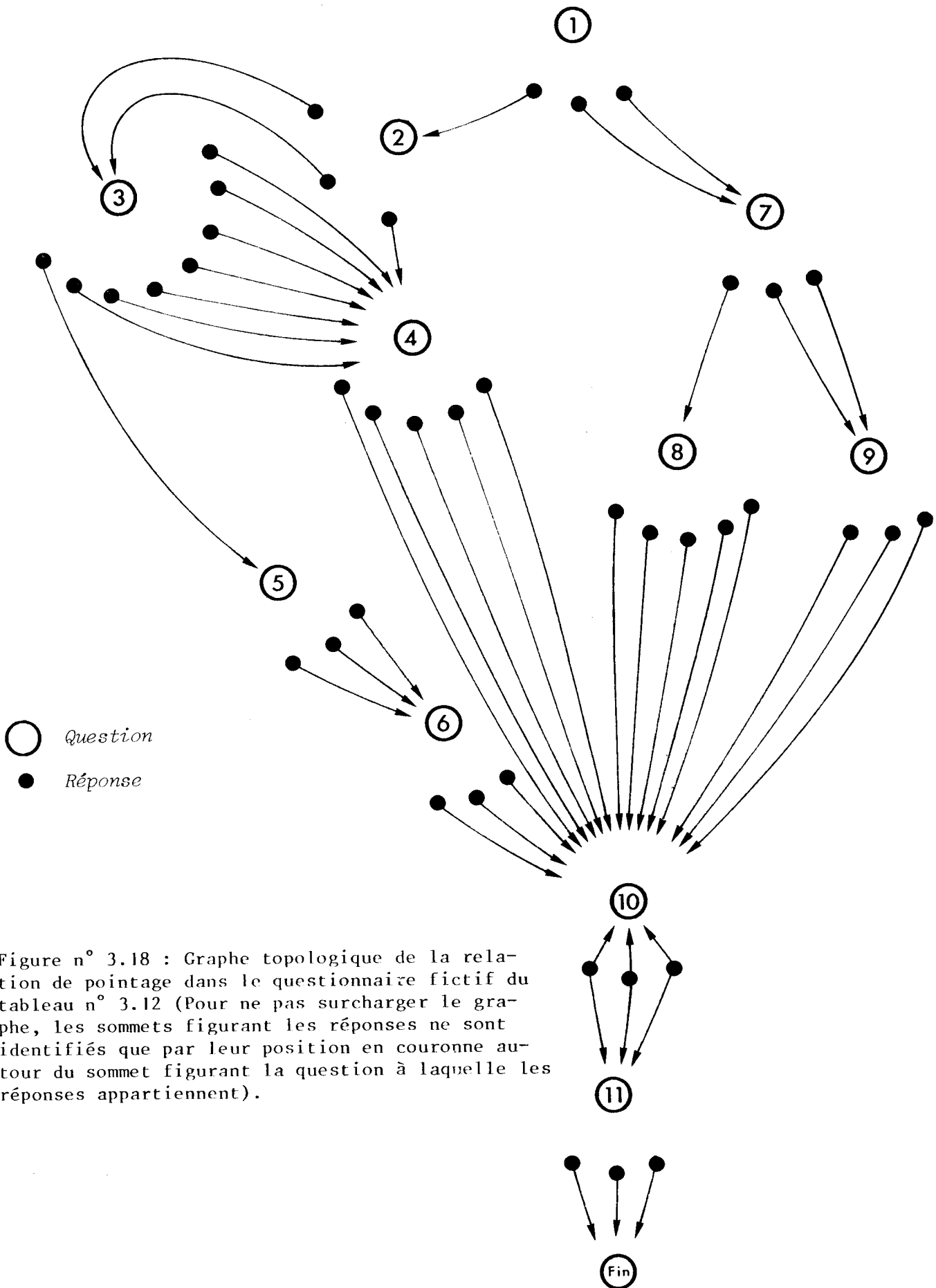
$$\forall r_{i,l} \in R_i \quad (\forall q_j \in Q \quad (r_{i,l} \not\mapsto q_j)) \Rightarrow r_{i,l} \mapsto \text{fin})$$

La réciproque n'est vraie que dans les questionnaires sans retour en

arrière. En effet, dans un questionnaire comportant la réitération de certaines questions, une même réponse peut pointer vers deux questions différentes, l'une en amont, et l'autre en aval. Enfin, dans tout questionnaire, il existe au moins une question telle que toutes les réponses possibles à cette question pointent vers *fin* ; cette question est un sommet terminal dans G_p (et par conséquent aussi dans G_{p_i}). Dans un questionnaire sans retour en arrière, une telle question n'a pas de successeur (puisque $fin \notin Q$).

Le graphe correspondant à la relation de pointage est un graphe biparti $G_{\mapsto} = (\mathcal{R}, Q \cup fin, \Gamma_{\mapsto})$. Dans un questionnaire sans retour en arrière, G_{\mapsto} est sans cycle : G_{\mapsto} est alors une forêt dont chaque arbre, sauf l'arbre dégénéré correspondant à q_1 , est quasi-fortement connexe supérieurement. Dans ces conditions en effet, la relation \mapsto définit une application de \mathcal{R} dans $(Q - q_1) \cup fin$ qui est surjective, puisque tout élément de l'ensemble d'arrivée est l'image d'au moins une réponse élément de l'ensemble de départ \mathcal{R} . Notre exemple illustre le cas général d'un questionnaire avec réitération de question, dans lequel il peut arriver qu'une réponse pointe vers deux questions distinctes (ici : q_{10} et q_{11}). Avec $|Q| = 11$ et $|\mathcal{R}| = 42$, le graphe G_{\mapsto} est d'ordre 54 (cf. figure n° 3.18).

D'autres définitions du graphe associé à la relation de pointage sont envisageables. Par exemple, il est commode de définir cette relation sur \mathcal{R}^2 , en posant $R_i = q_i$; ainsi, lorsqu'une réponse $r_{i,l}$ pointe vers q_j , on considère qu'elle pointe vers toutes les $r_{j,l} \in R_j$. Le graphe correspondant $G'_{\mapsto} = (\mathcal{R}, \Gamma'_{\mapsto})$ est un 1-graphe connexe ; dans le cas général, il n'est pas planaire. Sa base $S \subset \mathcal{R}$ est l'ensemble des dernières réponses que peuvent donner les personnes interrogées ; son antibase $T \subset \mathcal{R}$ est l'ensemble des premières réponses que peuvent donner les personnes interrogées. L'ensemble des chemins $\mu[x,y]$ tels que $x \in T$ et $y \in S$ correspond à l'ensemble des configurations de réponses possibles au



questionnaire (patrons de réponses). Dans notre exemple, S est constituée des réponses à q_{11} , et T , des réponses à q_1 ; bien qu'il n'y ait que 20 scenarii distincts (cf. figure n° 3.14), le nombre de chemins distincts de T à S est égal à 41.400 (en raison de la réitération éventuelle de q_{10} ; si q_{10} n'était pas réitérée, le nombre de patrons de réponses ne serait plus que de 1035; si en outre q_3 était à réponse unique, ce nombre se trouverait réduit à 675). Remarquons que si l'on définit une relation d'équivalence entre les réponses à une même question, la réduction de G'_{\rightarrow} qui en résulte engendre un p -graphe d'ordre $|Q|$ (avec $p \leq \max_{i \in K} |R_i|$), dont le squelette est le graphe de précession immédiate G'_{Pi} . On pourrait d'ailleurs représenter économiquement ce graphe réduit sans perte d'information, sous la forme d'un l -graphe valué $G''_{\rightarrow} = (Q \cup \text{fin}, \Gamma_{\rightarrow}, V_{\rightarrow})$ dans lequel la valuation de chaque arc (q_i, q_j) ou (q_i, fin) est le nombre de réponses $r_{i,l} \in R_i$ pointant vers q_j ou vers fin . Il s'agit en fait ici d'une valuation de G'_{Pi} .

La relation de *sélection* exprime le rapport existant entre une réponse $r_{i,l} \in R_i$ à une question-filtre q_i , et une question q_j dominée par q_i , telle que la réponse $r_{i,l}$ a sélectionné la population à qui q_j devait être posée; nous dirons alors que $r_{i,l}$ *sélectionne* q_j , et nous écrirons : $r_{i,l} S q_j$. Le cas le plus simple se rencontre lorsque la question dominée q_j est le successeur immédiat de la question-filtre qui la domine q_i . Dans une telle situation, la relation de précession immédiate implique qu'il existe au moins une réponse à q_i qui pointe vers q_j (et réciproquement); la relation de dominance implique que cette réponse sélectionne q_j :

$$\forall (q_i, q_j) \in Q^2 \quad (q_i D q_j \ \& \ q_i Pi q_j \Rightarrow$$

$$\exists r_{i,l} \in R_i \quad (r_{i,l} \mapsto q_j \Rightarrow r_{i,l} S q_j))$$

Lorsque la question dominée q_j n'est pas un successeur immédiat de la question-filtre q_i , il existe une question q_k distincte de q_j , telle que q_k est à la fois dominée et précédée immédiatement par q_i , et que q_k précède q_j . Dans ces conditions, toute réponse $r_{i,l}$ qui pointe vers q_k sélectionne q_k et sélectionne également q_j :

$$\begin{aligned} \forall (q_i, q_j) \in Q^2 (q_i D q_j \ \& \ q_i \overline{P} q_j \Rightarrow \\ \exists q_k \in Q (q_i D q_k \ \& \ \exists r_{i,l} \in R_i (r_{i,l} \mapsto q_k)) \Rightarrow \\ r_{i,l} S q_j) \end{aligned}$$

Dans le cas général, on dira que, pour qu'une réponse $r_{i,l}$ à une question q_i sélectionne une question q_j , il faut et il suffit que q_j soit dominée par q_i , et que $r_{i,l}$ pointe soit vers q_j , soit vers une question q_k dominée par q_i et prédécesseur de q_j :

$$\begin{aligned} \forall (q_i, q_j) \in Q^2 (r_{i,l} \in R_i \ \& \ r_{i,l} S q_j \Leftrightarrow \\ q_i D q_j \ \& \ (r_{i,l} \mapsto q_j \vee \exists q_k \in Q (q_i D q_k \ \& \\ q_k P q_j \ \& \ r_{i,l} \mapsto q_k))) \end{aligned}$$

Le graphe de la relation de sélection est un l-graphe biparti $G_S = (\mathcal{R}, Q, \Gamma_S)$; il n'est en général pas planaire (cf. figure n° 3.19). C'est un graphe partiel de la fermeture transitive du graphe de la relation de pointage. On peut également définir la relation de sélection sur \mathcal{R}^2 , en admettant que lorsqu'une réponse $r_{i,l}$ sélectionne q_j , elle sélectionne de ce fait toutes les réponses $r_{j,k} \in R_j$. Le graphe $G'_S(\mathcal{R}, \Gamma_S)$ ainsi obtenu est un l-graphe qui, dans le cas général, est non planaire. Sa base contient toutes les réponses qui ne sélectionnent aucune question, et

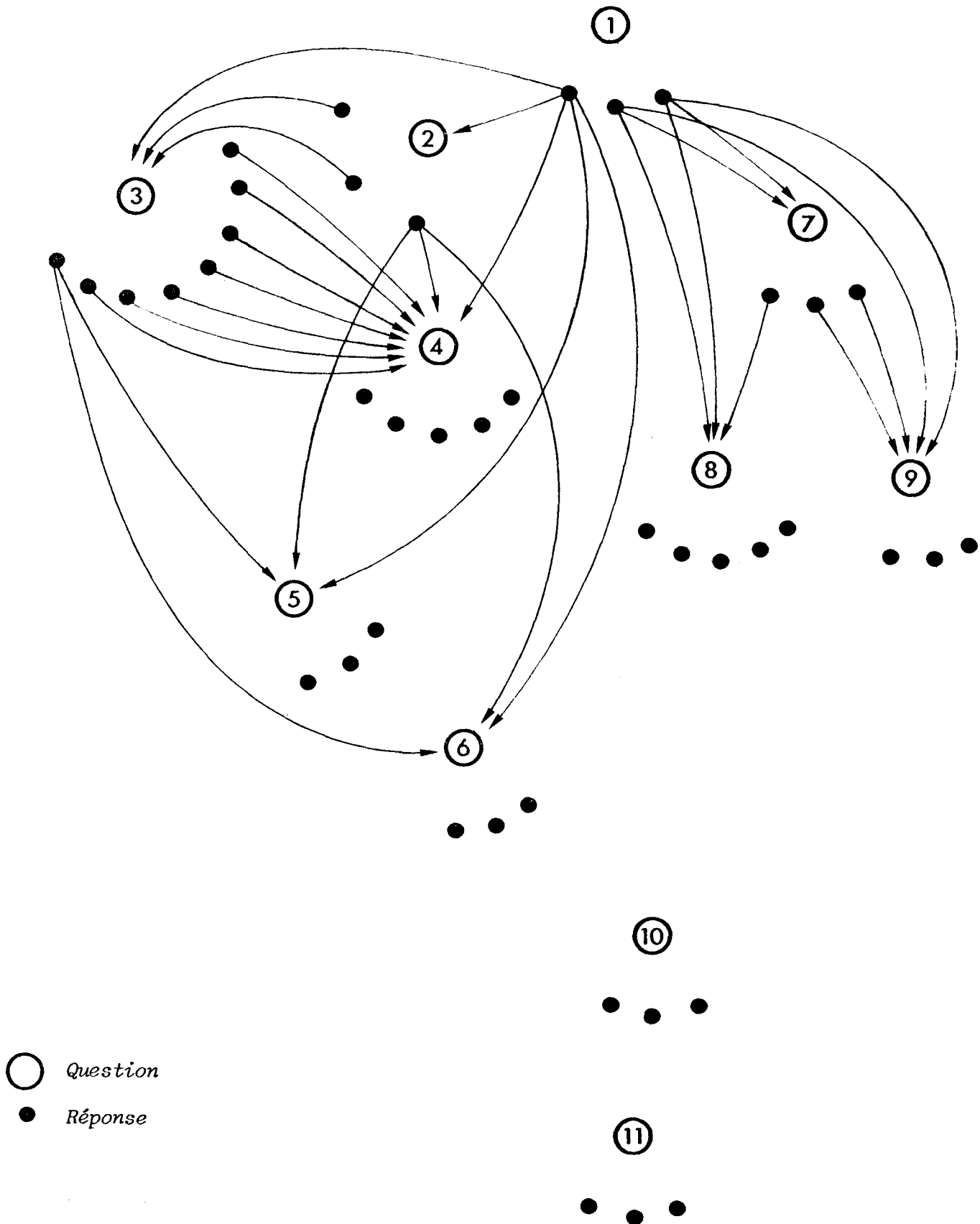


Figure n° 3.19 : Graphe topologique de la relation de sélection dans le questionnaire fictif du tableau n° 3.12.

son antibase toutes les réponses qui ne sont pas sélectionnées (c'est-à-dire l'ensemble des réponses à l'ensemble des questions omnibus). Si l'on définit en outre une relation d'équivalence entre les réponses à une même question, la réduction de G'_S engendre un p -graphe d'ordre $|Q|$ (avec : $p \leq \max_{i \in K} |R_i| - 1$), dont le squelette est le graphe de dominance G_D . On peut engendrer le graphe $G''_S = (Q, \Gamma_S, V_S)$ en valuant G_D selon le nombre d'arcs dans G'_S .

c) Relations sémantiques entre questions

La possibilité de définir des relations sémantiques entre questions dépendra des progrès futurs de l'analyse sémantique des questionnaires. A l'heure actuelle, il serait certainement prématuré d'étudier l'incidence structurelle de relations qui pourtant présentent un grand intérêt théorique et pratique, comme les relations de présupposition ou d'implication logique. Il est facile par contre de définir une relation d'identité sémantique entre questions, relation qui nous sera utile dans les opérations de réduction d'un questionnaire à une forme canonique.

La relation d'*identité sémantique* entre deux questions correspond à la notion de répétition de questions (cf. § 3.2.1.c). Nous dirons que deux questions q_i et q_j sont sémantiquement identiques, et nous écrirons $q_i \equiv q_j$, si et seulement si elles ont toutes deux le même énoncé et le même ensemble de réponses possibles. C'est une relation réflexive, symétrique, et transitive définie dans Q^2 . Dans notre exemple, nous avons : $q_2 \equiv q_7$, et $q_5 \equiv q_9$ (cf. le graphe topologique de cette relation, figure n° 3.20) ; mais q_4 et q_8 ne sont pas identiques, car si leur énoncé est le même, l'ensemble de leurs réponses possibles varie selon la question. L'intérêt pratique de cette notion tient à ce que si deux questions sont à la fois équivalentes et incompatibles (si l'on a : $q_i \mid q_j$ & $q_i \equiv q_j$), alors il s'agit en réalité d'une unique question posée à deux sous-populations distinctes (pour des raisons liées à la facilité de l'interview).

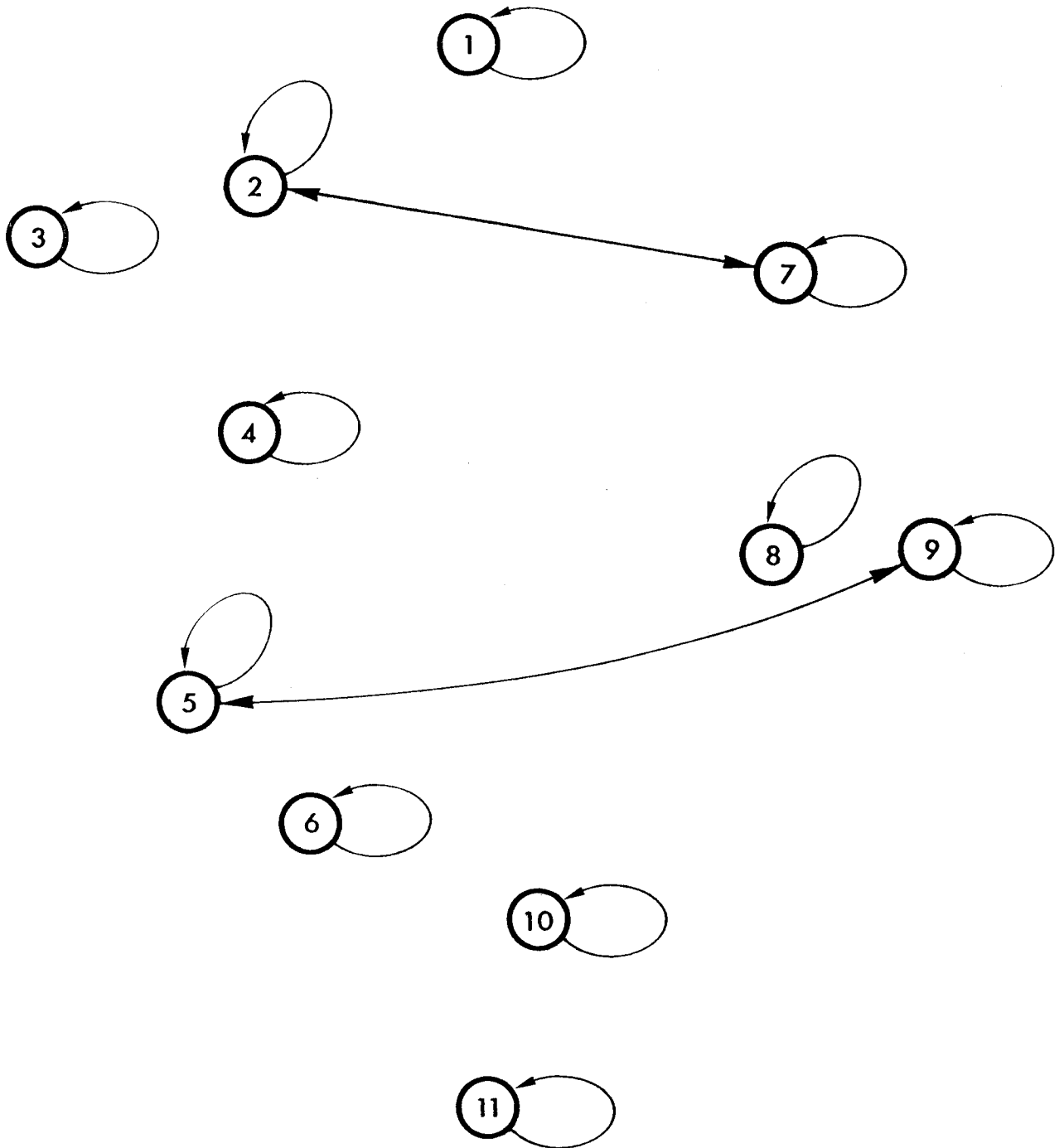


Figure n° 3.20 : Graphe topologique de la relation d'identité sémantique entre les questions du questionnaire fictif du tableau n° 3.12.

D'autres relations d'identité sémantique pourraient être définies. Par exemple, une relation d'identité partielle entre questions ayant le même ensemble de réponses possibles, mais dont les énoncés interrogatifs diffèrent légèrement (cf. la notion de répétition partielle, au § 3.2.1.c). Ou encore une relation d'identité sémantique entre réponses, définie sur \mathcal{R}^2 . L'intérêt théorique ou pratique de ces définitions ne nous a pas paru évident.

3.2.4. Réduction d'un questionnaire à une forme canonique.

Nous disposons maintenant du cadre formel nécessaire à la réduction de tout questionnaire à une forme canonique. L'utilité d'une telle réduction est d'aboutir à une présentation uniforme des informations recueillies, quelle que soit la structure du questionnaire, et quelles que soient les particularités du répondant. Une présentation uniforme facilite d'ailleurs l'archivage des données selon un format fixe identique pour tous les enregistrements d'un fichier, et fournit une base commode pour l'analyse statistique ultérieure de ces informations. C'est sur cette base que nous définirons ce qu'on appelle la "matrice des données".

La forme canonique d'un questionnaire est ce que Claude FLAMENT appelle un *questionnaire formel*, c'est-à-dire un système classificatoire constitué d'un ensemble de questions dont chacune est un facteur de classification des personnes interrogées (cf. [Flament 1976 : 52-63]). Par conséquent, tout élément q_i^* de ce questionnaire formel Q^* doit permettre de définir une partition de l'ensemble U des personnes interrogées. Il faut donc que toute question formelle $q_i^* \in Q^*$ soit telle qu'à toute personne interrogée corresponde toujours une réponse formelle $r_{i,l}^* \in R_i^*$ et une seule. Cette règle est impérative. On peut, par souci de commodité, et selon les besoins, lui adjoindre d'autres règles : règle d'économie, afin d'éviter une certaine redondance dans les données ; ou règle d'optimalité du codage, selon les caractéristiques du support physique des informations et la nature du traitement statistique ultérieur. Il y a donc plusieurs formes canoniques possibles pour un même questionnaire concret.

Nous présentons ci-après les trois grandes étapes de la réduction d'un questionnaire quelconque à une forme canonique déterminée : confusion des questions identiques, réduction des questions à une forme canonique, réduction de la structure du questionnaire à une structure purement séquentielle. Evidemment, pour un questionnaire donné, certaines de ces étapes sont inutiles, ou peuvent être écourtées. La procédure décrite vise en effet à la plus grande généralité pour le traitement d'un questionnaire considéré isolément.

a) Confusion des questions identiques incompatibles

Cette première opération, d'ailleurs facultative, ne s'applique qu'aux questionnaires dans lesquels deux questions ou plus ont exactement le même énoncé, et le même ensemble de réponses possibles au sens large ; nous en avons vu plusieurs exemples au § 3.2.1.c (SOFRES 769 ; SOFRES 1094), ainsi que dans notre questionnaire fictif. La raison de cette opération est avant tout une raison économique ; mais elle contribue de plus à la clarification du système classificatoire recherché.

Lorsque deux questions q_i et q_j sont à la fois identiques et incompatibles (c'est-à-dire posées à des sous-populations disjointes), la première opération consiste à les confondre en une question unique q_{ij} . Cette opération, répétée autant de fois qu'il est nécessaire, engendre un premier questionnaire formel, Q_1^* , dont le nombre de questions est en principe inférieur à celui de Q .

b) Réduction des questions à une forme canonique

Cette étape comporte deux opérations distinctes. La première opération est la constitution de l'ensemble R_i des réponses possibles au sens large à la question q_i , à partir de l'ensemble M_i des modalités qu'elle présente, de sa requête ρ_i , et de l'ensemble Φ_i des réponses métalinguistiques prévues par la codification. La procédure suivie varie évidemment selon la nature de la requête. Dans le cas le plus simple : $\rho_i = \text{"choisir } 1/n\text{"}$,

on a : $R_i = M_i \cup \Phi_i$. Dans les autres cas, la procédure est plus complexe.

Afin d'alléger l'écriture, posons au préalable :

- $\mathcal{P}_k (M_i)$ = ensemble des k -parties de M_i (c'est-à-dire des parties de cardinal k) ;
- $\mathcal{E}_k (M_i)$ = union des ensembles des permutations correspondant à chacune des k -parties de M_i ;
- $\mathcal{A}(M_i , K_i)$ = ensemble des applications de M_i dans K_i .

A titre d'exemple, en posant : $|M_i| = n$, si :

- $\rho_i = \text{"choisir } k/n\text{"}$, alors : $R_i = \mathcal{P}_k (M_i) \cup \Phi_i$;
- $\rho_i = \text{"choisir } \frac{u}{v}/n\text{"}$, alors : $R_i = \left(\bigcup_{v < k < u} \mathcal{P}_k (M_i) \right) \cup \Phi_i$;
- $\rho_i = \text{"ordonner } k/n\text{"}$, alors : $R_i = \mathcal{E}_k (M_i) \cup \Phi_i$;
- $\rho_i = \text{"ordonner } \frac{u}{v}/n\text{"}$, alors : $R_i = \left(\bigcup_{v < k < u} \mathcal{E}_k (M_i) \right) \cup \Phi_i$;
- $\rho_i = \text{"classer } n/p\text{"}$, (avec : $K_i =$ ensemble des classes dans lesquelles les éléments de M_i doivent être rangés, et : $|K_i| = p$), alors :
 $R_i = \mathcal{A}(M_i , K_i)$.

Il est possible, au prix de quelques complications dans l'écriture des formules, de passer ainsi en revue l'ensemble des types de consignes que nous avons décrits. Or, nous avons vu que, pour la plupart d'entre eux, le nombre des réponses possibles est en général trop élevé pour que R_i puisse être traité directement par le chercheur (d'autant que parfois il excèdera même le nombre des personnes interrogées). De plus, une codification qui aurait pour éléments les réponses possibles à toute question, quel que soit le type de sa consigne, risquerait de se révéler extrêmement coûteuse. C'est pourquoi il nous paraît utile de prévoir une seconde opération, ayant pour but de coder plus économiquement l'ensemble des réponses possibles à une

question. En pratique, on peut le plus souvent passer directement à cette seconde phase, sans avoir auparavant procédé à l'énumération des éléments de R_i (c'est d'ailleurs là l'une des raisons pour lesquelles nous n'avons pas donné un plus grand développement à la description de la première phase). Toutefois, sur le plan théorique, chacune des formes canoniques possibles a le même statut, et présente le même intérêt. La première phase a pour aboutissement une forme canonique (que nous appellerons la forme *alpha* : q_i^α), qui constitue en quelque sorte un type-idéal (au sens de WEBER) par rapport auquel pourront être définies d'autres formes, d'un usage plus commode. Cette forme est constituée d'une consigne de sélection unique appliquée à l'ensemble des réponses possibles à q_i . Soit, en posant $|R_i| = N$, et en utilisant la symbolique de BELNAP (cf. § 2.2.1.b) :

$$q_i^\alpha = ? \rho_i^\alpha \sigma_i^\alpha$$

$$\rho_i^\alpha = \text{"choisir } 1/N\text{"}$$

$$\sigma_i^\alpha = R_i$$

La seconde opération de réduction de q_i , à une forme canonique qui soit d'une utilisation commode, est une phase de division de q_i en une combinaison de questions élémentaires. Nous avons vu (au § 3.1.2.e) que toute consigne pouvait être réduite à une séquence de consignes de sélection ; en outre, toute consigne de sélection multiple peut être réduite à une séquence de consignes de sélection unique. La forme canonique q_i^* que nous nous proposons de définir pour une question quelconque q_i est constituée d'une séquence de questions de sélection unique. L'avantage pratique de cette forme est sa grande simplicité. Pour chaque question élémentaire $q_{i,l}$ figurant dans cette séquence, il nous faut définir l'ensemble $\sigma_{i,l}$ des modalités présentées. Plusieurs solutions sont possibles. L'une d'entre elles (que nous appellerons la forme *oméga* : q_i^ω) nous paraît constituer un autre type-idéal de forme canonique ; d'une certaine manière, cette forme est à l'opposé de la forme *alpha*. Compte-tenu des inconvénients pratiques de

ces deux formes, la forme canonique que nous retiendrons, q_i^* , sera une forme intermédiaire entre q_i^α et q_i^ω .

La forme canonique *oméga* consiste en une séquence de questions élémentaires de consigne $\rho_{i,l}^\omega = \text{"choisir } 1/2\text{"}$, et de sujet $\sigma_{i,l}^\omega = \{0, 1\}$.

L'ensemble des modalités de réponses au sens large de q_i est l'union : $M_i \cup \Phi_i$; posons : $|M_i \cup \Phi_i| = m$ (En conservant la notation adoptée pour la description des formes de consignes : $|M_i| = n$; nous avons par conséquent ici : $m \geq n$). Dans le cas le plus simple, où q_i est une question de sélection (quelconque), il y a m questions élémentaires (une par modalité) ; la ou les modalités choisies prennent la valeur 1, les autres modalités prenant la valeur 0. Dans le cas général, le nombre de questions élémentaires est un multiple de m . Ainsi :

- si $\rho_i = \text{"choisir } k/n\text{"}$, ou : $\text{"choisir } \frac{u}{v}/n\text{"}$, le nombre de questions élémentaires est égal à m . En outre, le nombre total de 1 est égal à k pour la première consigne, et compris dans l'intervalle $[v, u]$ pour la seconde.

- si $\rho_i = \text{"ordonner } k/n\text{"}$ (avec : $k < n$), le nombre de questions élémentaires est égal à $k.m$. Une première série de m questions identifie la modalité classée en premier ; une deuxième série, la modalité classée en second ; et ainsi de suite, jusqu'à la k -ième série. En outre, une modalité sélectionnée dans une série ne peut plus être sélectionnée dans une série ultérieure.

- si $\rho_i = \text{"ordonner } \frac{u}{v}/n\text{"}$, on a de même $u.m$ questions élémentaires, en u séries de m .

- si $\rho_i = \text{"classer } n/p : e_1, e_2, \dots, e_p\text{"}$, ou : $\text{"classer } n/p\text{"}$, il y a $p.m$ questions élémentaires. La première série de m questions identifie les modalités classées dans la première catégorie, et la p -ième série, les modalités classées dans la p -ième catégorie (Par souci d'économie, on pourrait se limiter à $p-1$ séries de questions élémentaires ; mais la

redondance permet un contrôle des données). Comme pour les consignes d'ordination, toute modalité ne peut être citée qu'une fois.

- si $\rho_i = \text{"apparié } k/m,n\text{"}$ (quel que soit le type de correspondance escompté), il y a $m.n$ questions élémentaires. Le nombre total de 1 pour l'ensemble des questions élémentaires est égal à k . En outre, s'il s'agit d'une application, tout élément de l'ensemble de départ doit être cité une fois et une seule (les autres contraintes possibles, plus spécifiques, se déduisent des propriétés des diverses formes d'application).

Les deux formes canoniques extrêmes que nous avons décrites ne correspondent pas aux pratiques habituelles en matière de codification et d'enregistrement des réponses à un questionnaire. A l'heure actuelle, ces pratiques restent encore très influencées par l'ère de la carte perforée (de 80 colonnes décimales ou duodécimales, telle qu'elle a été définitivement fixée en 1935). C'est pourquoi la forme *alpha* présente en principe un nombre de modalités beaucoup trop élevé, tandis que la forme *oméga* présente en moyenne un nombre de modalités trop faible. De plus, aucune de ces formes ne permet un codage de l'information réellement économique : q_i^{ω} est redondant par rapport à q_i^{α} , comme en témoignent les contraintes sur les réponses que nous avons dû introduire dans nos descriptions ; et q_i^{α} utilise $|R_i|$ positions binaires, alors qu'un codage optimal n'en utiliserait que $\log_2 |R_i|$ (arrondi à l'entier immédiatement supérieur). Cependant, compte-tenu des progrès prévisibles des techniques informatiques, il n'est pas impossible que l'on adopte un jour l'une ou l'autre de ces deux formes canoniques ; voire, si le rapport : coût du traitement/coût de la lecture des données continue à décroître, que l'on adopte soit le codage binaire le plus économique correspondant à la forme *alpha*, soit même une codification binaire correspondant à l'ensemble des configurations de réponses possibles à l'ensemble des q_i^{α} (c'est-à-dire à : $\prod_i \sigma_i^{\alpha}$). Ce dernier type de codage est celui qui réduirait le plus le volume nécessaire à l'archivage d'une enquête (dans notre exemple fictif, on passerait de 50 bits par répondant à 24 bits) ; il exige par contre un logiciel de codification-décodification spécifique (et non réalisé actuellement).

Si l'on s'en tient aux habitudes de pensée actuelles, on peut définir une forme canonique q_i^* qui soit compatible avec les formats de données usuels sur cartes perforées. En prenant en compte l'ensemble des contraintes imposées par les différents logiciels utilisés en sciences humaines, on est conduit à se limiter aux codes numériques sur une ou deux colonnes à perforation unique, et aux quantités (variables numériques enregistrées telles quelles). De plus, il paraît souhaitable que le sujet $\sigma_{i,l}^*$ de chaque question élémentaire de la forme canonique q_i^* ne présente pas plus d'une dizaine de modalités (voire, dans certains cas, une vingtaine) afin de pouvoir être facilement appréhendé dans sa totalité par l'esprit humain (exception faite évidemment pour les variables numériques, et éventuellement pour certaines variables d'un usage courant, comme le code à deux chiffres de la catégorie socio-professionnelle selon l'INSEE). Ces considérations nous ont amené à préconiser une forme canonique relativement bâtarde, qui présente les avantages d'être d'une utilisation assez facile (compte-tenu des pratiques courantes actuellement), et de bien se prêter à l'enregistrement des données en image-carte. Toute consigne ρ_i peut être considérée comme une procédure qui vise à établir une correspondance entre deux ensembles : l'ensemble M_i des modalités présentées, et un autre ensemble E_i , qui peut être l'ensemble binaire : {sélectionné, non sélectionné}, l'ensemble des rangs de 1 à k , ou de 1 à u , l'ensemble des p catégories de classement des modalités, ou encore un deuxième ensemble de modalités pour les consignes d'appariement. Cette correspondance est donc définie dans le produit cartésien $M_i \times E_i$. En dehors de quelques cas particuliers (très rares en pratique) de consignes d'appariement, il est toujours possible de choisir l'ensemble de départ de la correspondance demandée de telle sorte que celle-ci soit une application : pour une consigne d'ordination, ce sera l'ensemble des rangs ; pour les autres formes (exception faite des appariements qui ne sont pas des applications), ce sera l'ensemble M_i des modalités.

La règle de base de la réduction d'une question q_i à la forme canonique q_i^* est que l'ensemble de départ de l'application définit les questions élémentaires $q_{i,l}^*$ dont la séquence constitue q_i^* , et que l'ensemble d'arrivée définit le sujet $\sigma_{i,l}^*$ de ces questions élémentaires. Par

exemple, pour $\rho_i = \text{"ordonner } k/n\text{"}$, il y aura k questions élémentaires ayant pour sujet $\sigma_i = M_i$, plus une question "métalinguistique" ayant en principe pour sujet Φ_i . En toute rigueur, la question métalinguistique $q_{i,k+1}$ n'a d'utilité que si la personne interrogée n'a fourni aucune réponse linguistique (correspondant aux k premières questions élémentaires). Cette question est donc une question conditionnelle. Plutôt que d'introduire dans la série de questions élémentaires $q_{i,1}$ une question-filtre dominant $q_{i,k+1}$, ce qui aurait pour conséquence de complexifier la structure du questionnaire, nous préférons anticiper sur l'étape suivante (§ 3.2.4.c, transformation δ), et ajouter au sujet de la question métalinguistique la réponse : "question non posée" (abrégée en : "qnp" ; nous avons donc : $\sigma_{i,k+1} = \Phi_i \cup \{qnp\}$). Pour $\rho_i = \text{"classer } n/p\text{"}$, il y aura : $|\sigma_i| = m$ questions élémentaires dont le sujet sera l'ensemble des p catégories de classification, plus une question "métalinguistique", de sujet $\Phi_i \cup \{qnp\}$. Les seules exceptions à cette règle de base sont les questions de sélection (ou d'ordination) unique, dont la forme originale n'est pas modifiée, et les questions d'appariement non réductibles à une application entre les deux ensembles de modalités, pour lesquelles il y aura autant de questions élémentaires que de correspondances possibles, plus une question "métalinguistique" (à la question métalinguistique près, qui n'est pas décomposée en questions binaires, nous retrouvons ici la forme *oméga*). Le tableau n° 3.21 présente la forme q_i^* correspondant à q_i pour chacune des formes de consigne ρ_i décrites au § 3.1.2.

Nous appellerons Q_2^* le résultat de cette seconde étape dans la réduction d'un questionnaire à une forme canonique. Bien que cette étape soit obligatoire dans la procédure de réduction, son résultat reste très lié aux contraintes des logiciels utilisés dans les sciences sociales. Sous la forme que nous proposons, il est compatible avec les programmes de dépouillement d'enquête ou d'analyse des données n'admettant que des format d'entrée de type *FORTTRAN* (ce qui par exemple est le cas des programmes *OSIRIS*). Pour des logiciels moins rigides en entrée (tels que *DAPHNE* ou *NLT*, qui admettent une codification binaire, à douze positions par colonne), il conviendrait sans

Consigne originale ρ_i	Questions élémentaires $q_{i,l}^*$ constituant q_i^*		
	nombre total	sujet $\sigma_{i,l}^*$ des questions élémentaires de base	sujet de la question "métalinguistique"
<i>choisir</i> 1/n <i>choisir</i> k/n <i>choisir</i> $\frac{u}{v}$ /n <i>choisir</i> x/n	1 } n + 1	$R_i = M_i \cup \Phi_i$ } { cité , non cité }	pas de question métalinguistique } $\Phi_i \cup \{qnp\}$
<i>ordonner</i> 1/n <i>ordonner</i> k/n <i>ordonner</i> $\frac{u}{v}$ /n <i>ordonner</i> n-1/n	1 k + 1 u + 1 n + 1	$R_i = M_i \cup \Phi_i$ } $\sigma_i = M_i$	pas de question métalinguistique } $\Phi_i \cup \{qnp\}$
<i>classer</i> n/p : e_1, e_2, \dots, e_p <i>classer</i> n/p	} n + 1	} { classe 1, ..., classe p }	} $\Phi_i \cup \{qnp\}$
<i>appariier</i> (B)n/n,n <i>appariier</i> (I)m/m,n <i>appariier</i> (S)m/m,n <i>appariier</i> (A)m/m,n <i>appariier</i> k/m,n <i>appariier</i> x/m,n	n + 1 } m + 1 } (m.n)+1	} ensemble d'arrivée de l'application (de cardinal n) } { cité , non cité }	} $\Phi_i \cup \{qnp\}$ } $\Phi_i \cup \{qnp\}$

Tableau n° 3.21 : Description des questions élémentaires $q_{i,l}^*$, de consigne :

$\rho_{i,l}^* = \text{"choisir } 1/|\sigma_{i,l}^*|"$, constituant la forme canonique q_i^* d'une question q_i .

doute de définir une procédure légèrement différente, dont le résultat serait une forme canonique plus condensée.

c) Mise du questionnaire sous une forme séquentielle

A ce stade du processus de réduction, le questionnaire Q_2^* a approximativement encore la même structure que le questionnaire original Q : en particulier, il présente les mêmes questions-filtres, et les mêmes circuits (boucles de réitération des questions). En conséquence, certaines questions peuvent d'être pas posées à toutes les personnes interrogées, tandis que d'autres questions sont posées plusieurs fois à un même répondant. Or, pour que la forme canonique d'un questionnaire quelconque puisse constituer un système classificatoire, il faut qu'à tout couple : *question-interviewé* corresponde une réponse et une seule. Nous devons donc transformer Q_2^* en une forme Q_3^* qui présente cette propriété ; or, la seule forme qui satisfasse à cette condition est la forme purement séquentielle.

En vue de réduire un questionnaire concret quelconque à une forme séquentielle, il est commode de décrire la structure de ce questionnaire selon deux dimensions. La première concerne les aiguillages ; elle comporte quatre niveaux partiellement hiérarchisés :

- questionnaire sans aiguillage (la relation de dominance n'est définie pour aucun couple de questions) ;

- questionnaire avec questions-filtres organisées en arborescence seulement (si l'on a : $q_i \text{ D } q_j$ & $q_i \text{ D } q_k$, on n'a jamais : $q_j \text{ D } q_k$, ni : $q_k \text{ D } q_j$) ;

- questionnaire avec questions-filtres en treillis seulement (si : $q_i \text{ D } q_j$ & $q_i \text{ D } q_k$, on a toujours : $q_j \text{ D } q_k \vee q_k \text{ D } q_j$) ;

- questionnaire avec questions-filtres en arborescence et questions-filtres en treillis.

La seconde dimension porte sur la réitération de questions ou de séries de questions ; elle comporte également quatre niveaux partiellement hiérarchisés :

- questionnaire sans circuit (on n'a jamais : $q_i \text{ P } q_j$ & $q_j \text{ P } q_i$) ;
- questionnaire dont tous les circuits sont parcourus un nombre de fois déterminé à l'avance (réitérations des questions en nombre fixe) ;
- questionnaire dont tous les circuits sont parcourus un nombre de fois qui dépend des caractéristiques ou des réponses de la personne interrogée (réitérations des questions en nombre variable) ;
- questionnaire comportant à la fois des réitérations en nombre fixe et des réitérations en nombre variable.

	Pas d'aiguillage	Questions-filtres en arborescence seulement	Questions-filtres en treillis seulement	Questions-filtres en arborescence et en treillis
Pas de circuit	Forme <i>O</i>	Forme <i>A</i>	Forme <i>B</i>	Forme <i>AB</i>
Itérations en nombre fixe seulement	Forme <i>C</i>	Forme <i>AC</i>	Forme <i>BC</i>	Forme <i>ABC</i>
Itérations en nombre variable seulement	Forme <i>D</i>	Forme <i>AD</i>	Forme <i>BD</i>	Forme <i>ABD</i>
Itérations en nombre fixe et en nombre variable	Forme <i>CD</i>	Forme <i>ACD</i>	Forme <i>BCD</i>	Forme <i>ABCD</i>

Tableau n° 3.22 : Typologie des questionnaires en vue de leur réduction à la forme *O* (structure purement séquentielle).

En combinant ces deux dimensions, on obtient seize types de questionnaires, que nous avons arbitrairement désignés par des symboles alphabétiques (Tableau n° 3.22). La forme purement séquentielle est la forme O . Notre questionnaire fictif est de la forme ABD : on y trouve à la fois des questions-filtres en arborescence (q_1 suivie de q_7), des questions-filtres en treillis (q_2 et q_3 dominant q_4), et un circuit à nombre de réitérations variable (q_{10} est répétée un nombre de fois qui varie selon l'effectif des membres de la famille de l'interviewé qui vivent avec lui).

Dans le cas général, pour réduire à la forme O un questionnaire quelconque, on utilise quatre espèces distinctes de transformations :

⊙ *Suppression des itérations en nombre variable.* Pour cela, on commence par déterminer le nombre maximum possible d'itérations. S'il s'agissait de préparer l'analyse primaire d'un questionnaire non préalablement codifié, il serait nécessaire de rechercher à quelle situation concrète se réfère ce maximum (dans notre questionnaire fictif, quel est le nombre maximum possible de membres d'une même famille vivant dans un même foyer, en France, au moment de l'enquête). Dans la perspective d'une analyse secondaire, il suffit de relever le nombre maximum de fois où la réponse à la question réitérée peut matériellement être enregistrée (4 fois dans notre exemple fictif), que ce soit à cause de la précodification des réponses, ou à la suite d'une analyse de contenu ; ce nombre est nécessairement fixé préalablement au report des réponses sur un support lisible par ordinateur. Ce nombre i ayant été déterminé, on introduit $i-1$ questions-filtres purement structurales pour commander l'éventuelle réitération (dans notre exemple fictif : "Reste-t-il des personnes de la famille de l'interviewé et vivant avec lui au sujet desquelles la q_{10} n'a pas encore été posée ? Si *non*, passer à q_{11} "). On reproduit enfin $i-1$ fois la séquence à réitérer, après chaque apparition de la question-filtre structurale. S'il s'agit en réalité d'une répétition partielle (ce qui correspond au cas le plus fréquent), on remplace évidemment, dans chaque répétition de la question, l'élément variable par sa valeur constante correspondant à cette répétition (cf. les exemples cités au § 3.2.1.c).

On a donc substitué aux réitérations une séquence de questions comportant des questions-filtres en arborescence.

⊙ *Suppression des itérations en nombre fixe.* Il s'agit en règle générale de répétition partielle d'une question, ou d'une série de questions (ce dernier cas correspondant à ce que FLAMENT appelle un "questionnaire [composé] en produit cartésien" [Flament 1976 : 61-62]). Comme pour la transformation précédente, il suffit de répéter en séquence la ou les questions réitérées, en remplaçant les éléments variables de l'énoncé par l'élément constant qui convient (cf. p.ex. : "Quand êtes-vous allé *au cinéma* pour la dernière fois ? [...] Quand êtes-vous allé *au théâtre* pour la dernière fois ? [...] " *SOFRES 571 : 22a*, cité au § 3.2.1.c). Au circuit est ainsi substituée une série de questions purement séquentielle.

⊙ *Suppression des questions-filtres formant treillis.* On peut considérer les ensembles de questions-filtres formant treillis comme une manière commode de poser une question unique. Dans l'exemple cité au § 3.2.1.b, il s'agissait de définir le point d'achat d'essence habituel en deux questions successives (*SOFRES 340 : 3-3a*). Dans notre exemple fictif (en supposant pour simplifier que l'opération facultative de confusion des questions identiques n'a pas été effectuée), les questions 2 et 3 ont pour fonction structurelle de distinguer entre pratiquants réguliers, pratiquants occasionnels, et non pratiquants (aux deux premières catégories sera ensuite posée la q_4 ; à la dernière, seront posées les q_5 et q_6). Pour substituer au treillis ainsi formé une simple arborescence, il suffit de fusionner en une question-filtre unique les questions-filtres formant combinaison. S'il y a plus de deux questions à fusionner, cette transformation peut être effectuée en plusieurs étapes successives portant chacune sur un couple de questions voisines : (q_i, q_j) tel que : $q_i \text{ Pi } q_j \vee q_j \text{ Pi } q_i$ (de préférence en commençant par la question qui domine toutes les autres). Chaque étape comprend deux opérations simultanées :

(a) Constitution de l'ensemble $R_{i,j}$ des réponses possibles à la question $q_{i,j}$ résultant de la fusion de q_i et de q_j . Appelons E l'ensemble des réponses à q_i qui pointent vers q_j : $E = \{r_{i,l} \in R_i : r_{i,l} \mapsto q_j\}$. L'ensemble des réponses possibles à $q_{i,j}$ est l'union des réponses à q_i qui ne pointent pas vers q_j , et du produit cartésien de l'ensemble des réponses à q_i qui pointent vers q_j et de l'ensemble des réponses à q_j :

$$R_{i,j} = (R_i - E) \cup (E \times R_j)$$

(b) Reformulation des réponses ainsi formées. Pour chaque réponse $r_{i,j,m} \in E \times R_j$ qui est constituée d'un couple de réponses : $(r_{i,k}, r_{j,l})$, on procède à la conjonction de l'énoncé de $r_{i,k}$ (qui est un élément de la présupposition de q_j) et de l'énoncé de $r_{j,l}$; on obtient ainsi l'énoncé de $r_{i,j,m}$.

Pour illustrer ceci, reprenons l'exemple simple du point de vente habituel (*SOFRES 340 : 3-3a*), cité au § 3.2.1.b. L'ensemble des réponses $R_{3,3a}$ à la question résultant de la fusion sera constitué de :

- la réponse à la question 3 qui ne pointe pas vers 3a : "a un endroit où il fait au moins la moitié de ses achats d'essence" ;

- la conjonction de la réponse à la question 3 qui pointe vers 3a, et de la première réponse à 3a : "n'a pas un endroit où il fait au moins la moitié de ses achats d'essence, mais en a un où il va un peu plus souvent qu'aux autres" ;

- la conjonction de la réponse à la question 3 qui pointe vers 3a, et de la seconde réponse à 3a : "n'a pas d'endroit où il fasse au moins la moitié de ses achats d'essence, ou bien où il aille un peu plus souvent qu'aux autres".

Notre exemple fictif est un peu plus complexe, dans la mesure où deux réponses à q_2 pointent vers q_3 , et où q_3 est à réponses multiples. q_3 a dû être réduite à une forme canonique q_3^* , constituée d'un ensemble de quatre questions élémentaires $q_{3,1}^*$, pour lesquelles l'opération de fusion avec q_2 devra être exécutée. A titre d'exemple, le résultat de cette opération pour q_2^* et $q_{3,1}^*$ donnera l'ensemble des cinq réponses possibles :

- "assiste régulièrement aux offices religieux" ;
- "n'assiste pas régulièrement, mais assiste à l'occasion de fêtes religieuses" ;
- "n'assiste ni régulièrement, ni à l'occasion de fêtes religieuses" ;
- "ne peut pas dire qu'il assiste régulièrement, mais assiste à l'occasion de fêtes religieuses" ;
- "ne peut pas dire qu'il assiste régulièrement, et n'assiste pas à l'occasion de fêtes religieuses".

On constate sur cet exemple que $q_{2.3,1}^*$ n'est plus une question-filtre, puisque toutes ses réponses pointent vers q_4 . Le rôle structurel de q_3 est rempli par $q_{2.3,4}^*$, qui a pour réponses possibles (abrégées) :

- "assiste régulièrement" (passer à q_4) ;
- "n'assiste pas régulièrement, ni exceptionnellement" (passer à q_5) ;
- "ne peut pas dire qu'il assiste régulièrement et n'assiste pas exceptionnellement" (passer à q_5).

En règle générale, le résultat de cette transformation est donc la substitution, à un aiguillage en treillis, d'un aiguillage en arborescence.

(δ) *Suppression des questions-filtres en arborescence.* Cette transformation a pour effet d'effacer en quelque sorte le contenu structurel des questions, et, par cette opération, de rendre effectivement classificatoire toute question informative. Avant cette transformation, aucune question de Q_2^* dominée n'est classificatoire, puisqu'il existe au moins une catégorie de personnes interrogées à qui la question n'a pas été posée. Pour rendre la question classificatoire, il suffit d'étendre l'ensemble Φ_i de ses réponses métalinguistiques en ajoutant l'indication : "question non posée". On peut en outre préciser la raison de cette omission. Par exemple, dans notre questionnaire fictif, l'ensemble des réponses possibles à $q_{2.3,4}^*$, qui est dominée par q_1 , sera augmenté de la réponse métalinguistique : "question non posée : déclare ne pas croire en Dieu, ou ne pas avoir d'opinion sur ce point". La raison de l'omission de $q_{2.3,4}^*$ (dans le questionnaire original, la raison de l'omission de q_2 et de q_3) est la disjonction des contenus sémantiques des réponses à q_1 qui ne pointent pas vers q_2 . Lorsqu'une question q_i est dominée par plusieurs questions-filtres ($q_j, q_k, \text{etc.}$), la raison de l'omission est la conjonction des disjonctions des réponses qui ne pointent ni vers q_i , ni vers q_j tel que : $q_k D q_i$ & $q_k D q_j$ & $q_j D q_i$. Par exemple, dans notre questionnaire fictif où q_9 est dominée par q_1 et q_7 , R_9 s'accroît de la réponse métalinguistique : "question non posée : déclare ne pas croire en Dieu ou n'avoir pas d'opinion sur ce point, et déclare ne pas assister régulièrement aux offices religieux ou ne pas pouvoir répondre sur ce point". Remarquons que cette transformation a pour effet secondaire d'intégrer dans les réponses aux questions les informations nécessaires à l'accomplissement de la fonction des questions-filtres purement structurelles (ce que nous avons appelé, au § 3.2.1.b *in fine*, les "aiguillages sans question-filtre apparente"). Dans la réduction d'un test de niveau à une forme canonique, on serait ainsi conduit à ajouter, après l'indication : "question non posée", des informations du type : "réussite antérieure à k questions de même difficulté" ; "échec antérieur à k

questions plus faciles" ; "dépassement du temps imparti à la série de questions" ; etc.

Les quatre transformations que nous venons de décrire ne peuvent pas être appliquées dans un ordre quelconque. De plus, parmi tous les enchaînements possibles de transformations qui aboutissent à la forme O , certains sont plus courts que d'autres. La figure n° 3.23 est le graphe des procédures les plus rapides pour réduire à une structure purement séquentielle chacune des formes de questionnaire décrites dans le tableau n° 3.22 (exception faite évidemment de la forme O). Pour certaines formes, plusieurs enchaînements de même longueur sont possibles (il y en a trois pour la forme $ABCD$, et deux pour les formes ACD , ABD , et CD) ; des recherches ultérieures seraient nécessaires pour comparer l'efficacité de chacune de ces procédures concurrentes.

A l'issue de l'ensemble de ces transformations, on a engendré, à partir d'un questionnaire concret quelconque Q , un questionnaire formel simple Q_3^* , que nous appellerons *forme canonique de Q* , et que nous abrègerons en : Q^* . Ainsi que nous l'avons signalé, cette forme canonique n'est pas la seule possible ; selon les conditions d'archivage des données, et aussi selon les techniques d'analyse statistique envisagées, on peut lui préférer d'autres formes canoniques, telles que Q^α ou Q^ω par exemple.

d) Questionnement et "matrice des données"

Nous disposons maintenant d'un système classificatoire $Q^* = \{ q_i^* \}$, avec $i \in \Lambda = \{ 1, \dots, \lambda \}$. A chaque question canonique $q_i^* \in Q^*$ correspond l'ensemble des réponses possibles R_i^* . Comme nous l'avons fait pour Q au § 3.2.3.b, nous pouvons définir pour Q^* l'ensemble des réponses possibles à chacune des questions : $\mathcal{R}^* = \bigcup_{i \in \Lambda} R_i^*$. Une

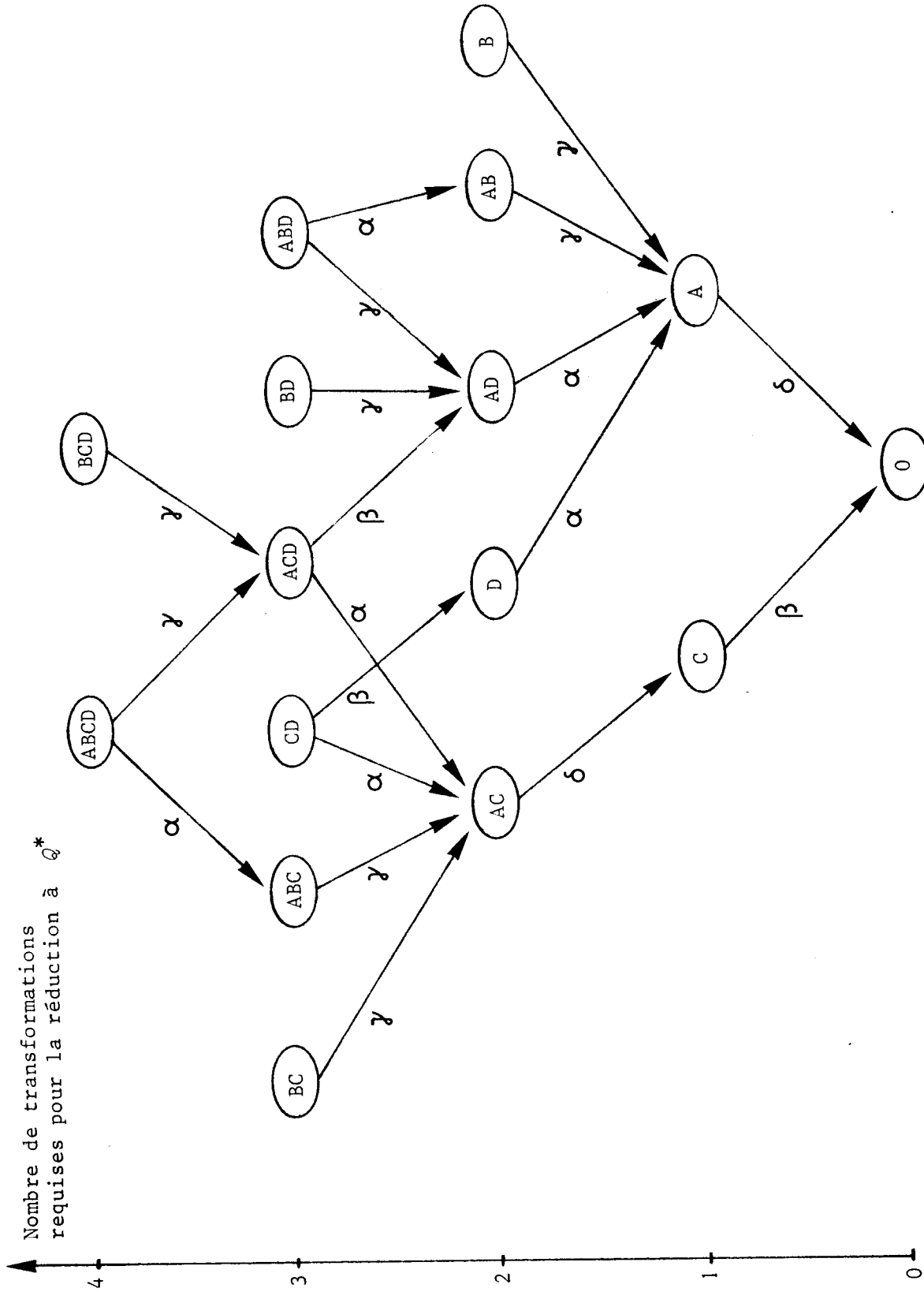


Figure n° 3.23 : Graphe de la procédure de réduction d'un questionnaire à une forme canonique.

notion plus utile est l'ensemble des configurations de réponses possibles à Q^* (ensemble des patrons de réponse) : $R^* = \prod_{i \in \Lambda} R_i^*$.

Administrer le questionnaire concret Q à une population d'individus U (en général, un échantillon d'une population plus importante), c'est en dernière analyse appliquer U sur R^* . Il faut en effet considérer le questionnaire concret Q comme dérivé du questionnaire formel Q^* (et non l'inverse, comme l'analyse qui précède pourrait le laisser croire) : pour l'essentiel, la réduction à une forme canonique est une démarche inductive visant à reconstituer, à partir de Q , la forme Q^* qui l'a engendré. Cela se conçoit mieux si l'on replace le questionnaire concret dans le processus global de la recherche. Nous avons, dans notre introduction (§ 0.1.d), décrit schématiquement ce que nous avons appelé le *questionnaire théorique* du chercheur, et les rapports de celui-ci avec l'observation ou l'enquête sur le terrain. L'objectif de l'enquête étant de détecter dans la population étudiée des régularités ou des concomitances, le questionnaire empirique doit nécessairement permettre de caractériser chaque individu de la population pour toute dimension incluse dans le questionnaire : c'est là précisément une propriété du questionnaire formel Q^* , que ne possède plus (en règle générale) le questionnaire concret Q . La dérivation de Q à partir de Q^* s'explique évidemment par des considérations liées à la stratégie de l'entretien, puisque l'omission (grâce aux aiguillages) de certaines questions, inutiles dans un contexte donné, réduit la durée moyenne de l'entretien et permet de ne pas trop indisposer le répondant. Nous appellerons donc *questionnement* l'application $\pi : U \rightarrow R^*$, qui à tout individu $u \in U$ fait correspondre un patron de réponse $r^* \in R^*$ et un seul (cf. [Flament 1976 : 64]).

Le résultat d'un questionnement donné (pour une population U et un questionnaire Q^* déterminés) peut être présenté de plusieurs manières. En nous inspirant pour l'essentiel de [Flament 1976 : 64-66] et de [Rouanet 1977 : 29-30], nous distinguerons préalablement :

- l'ensemble $R^* = \prod_i R_i^*$ des configurations de réponses possibles au questionnaire Q^* , appelé également *espace d'observation* ;

- l'ensemble $\Pi^* = (R^*)^U$ de toutes les applications possibles π de U sur R^* .

Ces ensembles ne dépendent évidemment que des caractéristiques de Q^* . Si l'on considère les effets du questionnement π , on peut définir un sous-ensemble $R_{i,U}^*$ de l'ensemble R_i^* des réponses possibles à une question q_i^* , tel que toute réponse de $R_{i,U}^*$ ait été donnée une fois par au moins un individu $u \in U$. Nous distinguerons ensuite :

- l'ensemble $R_U = \prod_i R_{i,U}^*$ des configurations de réponses possibles pour la population U , compte-tenu du fait que certaines réponses prévues dans Q^* peuvent n'avoir jamais été données (cet ensemble est appelé *espace de description*) ;

- l'ensemble $\Pi_U = (R_U)^U$ de toutes les applications possibles π de U sur R_U ;

- l'ensemble P_U de toutes les configurations de réponses effectivement observées sur la population U , c'est-à-dire tel que :
 $\forall p \in P_U (\exists u \in U (\pi : u \rightarrow p))$.

Certains de ces ensembles ont entre eux des relations d'inclusion :
 $\Pi \subset \Pi^*$, et : $P_U \subset R_U \subset R^*$.

Ce que les statisticiens des sciences sociales appellent "*matrice des données*" est l'ensemble $\mathcal{M} = U \times Q^*$, dans lequel l'élément $m_{i,j}$, correspondant à l'individu i et à la question (formelle) j , prend ses valeurs dans R_j^* . Autrement dit, l'élément $m_{i,j} \in \mathcal{M}$ est la réponse donnée par l'individu i à la question j (cf. p.ex. [Galtung 1967 : 9-11]). Le choix de l'expression "*matrice des données*" (*data matrix*) est doublement malheureux : ainsi que nous l'avons rappelé dans l'introduction, les

informations recueillies grâce au questionnaire ne sont pas "données", mais construites ; et le tableau rectangulaire par lequel on représente généralement le produit cartésien $U \times Q^*$ n'est ordinairement pas une matrice, car les nombres qui y figurent ne sont le plus souvent que des symboles commodes, sans aucune propriété numérique. C'est pourquoi, dans les sciences humaines expérimentales, on appelle souvent \mathcal{M} un *protocole*, ou une *famille d'observations* (cf. [Barbut 1973 : I 26], [Flament 1976 : 64], [Rouanet 1977 : 29]) ; le mot "protocole" évoquant (en français comme en anglais) l'idée d'un formulaire administratif uniforme servant à enregistrer des informations. Nous avons toutefois cru bon de conserver l'expression "matrice des données", malgré ses défauts évidents, parce que l'usage des sociologues l'a consacrée, et parce qu'elle est en pratique dépourvue d'ambiguïté (ce qui n'est pas le cas pour le mot protocole, qui peut, selon les contextes, désigner également l'ensemble des consignes d'application d'un test, l'ensemble des réponses données par un sujet unique, etc.).

La "matrice des données" constitue la représentation la plus courante des résultats du questionnement, et un bon point de départ pour aborder le problème de l'analyse des données (au sens large). Il s'agit toutefois d'une représentation abstraite, puisqu'une enquête par sondage relativement modeste porte au minimum sur un échantillon d'un millier de personnes, et que la forme canonique du questionnaire utilisé comporte en général beaucoup plus de deux cents questions ; la "matrice des données" correspondante aurait donc plus de deux cent mille cases élémentaires, ce qui suffit à lui ôter tout intérêt pratique. D'autres représentations abstraites sont parfois utilisées, telles que l'énumération ou le dénombrement des patrons de réponse effectivement observés ; elles se déduisent facilement des différents ensembles que nous avons définis (et en particulier de R^* , R_{ij} , P_{ij}). Les ensembles "théoriques" Π^* et Π_{ij} servent d'univers de référence pour l'interprétation des données observées.

3.3. PRÉALABLES A L'ANALYSE DES DONNÉES

"L'objectif de toute procédure statistique appliquée à des données est de condenser ces données en une représentation simplifiée acceptable, susceptible de conduire à des conclusions «signifiantes», interprétables" [Rouanet 1977 : 20]. Nous pouvons dans cette optique, considérer l'analyse des données au sens large comme un ensemble de méthodes visant à présenter la "matrice des données" \mathcal{M} sous une forme réduite, plus facile ainsi à appréhender et à manipuler. En effet, si l'on classe le plus souvent les techniques d'analyse des données soit selon la théorie mathématique sous-jacente (cf. p.ex. [Bertier 1975]), soit selon le niveau de mesure exigé des variables à traiter, il est également possible de les classer selon le type de condensation qu'elles opèrent sur la matrice des données. On obtient ainsi trois grandes familles de méthodes, selon que leur objectif principal est la réduction du nombre de lignes de \mathcal{M} (c'est-à-dire du nombre d'individus ; ex : taxinomie numérique), du nombre de colonnes (c'est-à-dire du nombre de variables ; ex : analyse factorielle en composantes principales), ou simultanément du nombre de lignes et du nombre de colonnes (ex : analyse hiérarchique).

A la base de toutes ces méthodes se trouve l'évaluation d'une ressemblance ou d'une proximité entre lignes (indice de similarité ou de dissimilarité entre individus, p.ex.), ou entre colonnes (indices de liaison, de corrélation, de contingence, etc.). Dans la très grande majorité des cas, ces divers indices sont définis entre deux ensembles (d'individus ou de variables) ; ils se réduisent, en dernière analyse, à une mesure de la variabilité intergroupe, comme on le voit bien par exemple pour la mesure de l'information transmise (cf. [Attneave 1959 : 43-51], ou pour le rapport de corrélation η (cf. [Morice 1954 : 344-348]). On peut donc penser qu'à partir du moment où l'on est parvenu à jeter un pont théorique entre les considérations logico-sémantiques que nous avons développées sur le questionnaire, et les mesures de dispersion et de liaison entre variables, on est parvenu par cela même à définir un lien unissant théorie du questionnaire et analyse des données.

Notre objectif, dans les pages qui suivent, est de contribuer à l'édification de ce pont théorique. Nous montrons tout d'abord comment il est possible de traduire les propriétés formelles que nous avons postulées sur les variables (au § 3.1.3.b) en termes de "quanta d'information", au prix il est vrai d'une formalisation mathématique relativement complexe (du moins pour un non mathématicien). Nous verrons ensuite qu'à partir de cette formalisation, il est assez facile de reconstituer les principaux coefficients usuels de dispersion et de liaison.

3.3.1. Information structurelle et niveaux de mesure

L'hypothèse de l'existence de quanta d'information n'est pas récente ; selon S. S. STEVENS, dès 1926, E. G. BORING avait suggéré la possibilité de découvrir, à partir des mesures de seuils différentiels, des "points critiques" dans un continuum sensoriel, ce qui pourrait conduire à l'adoption d'une théorie quantique de la discrimination sensorielle [Stevens 1940 : 583 n 1]. Mais c'est à notre connaissance Donald M. Mac KAY qui (en 1950) a donné à cette hypothèse sa forme la plus achevée, et la plus directement utilisable. Nous commençons donc par présenter les grandes lignes de la conception de Mac KAY, et en particulier sa distinction fondamentale entre les deux types de contenu de l'information. Nous définissons ensuite les opérateurs \mathcal{S} et \mathcal{D} , qui sont à la base de la formalisation que nous proposons. Nous décrivons enfin les relations existant entre ces opérateurs et les niveaux de mesure distingués par STEVENS et par TORGERSON.

a) Questionnaire formel et quanta d'information

Par définition, le questionnaire formel Q^* , qui correspond à un questionnaire concret Q , est un système de classification des individus que l'on se propose d'interroger. Le parallélisme entre cette conception du questionnaire (que l'on trouve par exemple chez [Barbut 1973 : I 27] et

[Flament 1976 : 61]), et celle de l'appareillage expérimental du physicien selon Mac KAY (cf. § 0.1.d), va nous permettre d'appliquer, aux informations apportées par le questionnement, les analyses de Mac KAY sur les résultats d'expérimentations physiques. Ces analyses nous conduiront à distinguer, dans ces informations, un contenu structurel, ou *a priori*, et un contenu métrique, ou *a posteriori*, chacun de ces contenus étant composés d'informations élémentaires indifférenciées.

Nous l'avons vu dans notre introduction, selon Mac KAY, la première phase d'une expérience est la conception d'un système de classification des événements que l'on se propose d'observer. Les résultats de l'expérience sont exprimés par une proposition scientifique, c'est à dire par une expression logique fondée sur des données finies. Cette expression est susceptible d'être traduite sous la forme d'une configuration de propositions élémentaires (ou "atomiques") sans perte d'information (Mac KAY se réfère ici explicitement à WITTGENSTEIN ; cf. [Wittgenstein 1921 : § 2.0201]). "Toute proposition atomique rend compte d'un fait si élémentaire qu'il ne peut plus être décomposé, et qu'il n'a par conséquent que la caractéristique « vrai » ou « faux » ; son existence est son seul attribut". [Mac Kay 1950 : 292]. "La proposition la plus élémentaire qui rende compte d'une observation établit l'existence d'une *relation de coïncidence* entre deux entités. D'un autre côté nous définissons une *grandeur* en disant qu'elle occupe un certain intervalle sur une échelle" [Mac Kay 1950 : 293]. Cette échelle est composée d'intervalles significatifs minimum (cf. les quanta de discrimination sensorielle, in [Stevens 1940]), correspondant à l'unité d'échelle appropriée à ce type d'observation ; la grandeur peut être alors caractérisée par le nombre d'intervalles qu'elle recouvre sur l'échelle. Ce nombre constitue une information *métrique*, liée à l'opération de mesure. L'unité d'information métrique, le *metron*, exprime une relation de présence élémentaire (recouvrement d'un intervalle de l'échelle). Le contenu métrique de l'information apportée par une observation est en principe "la reproductibilité de la configuration d'événements" décrite par la proposition [Mac Kay 1950 : 295] ; en d'autres termes, il résulte de l'association d'un nombre avec chacune des catégories possibles définies par la structure de l'expérimentation [Mac Kay 1950 : 291].

L'information structurelle apportée par une expérience est liée au nombre de dimensions indépendantes (de "degrés de liberté") permises par la structure de l'expérimentation. "Le plan d'une expérience est essentiellement la définition *a priori* d'une structure, d'un ensemble de catégories dans les termes seuls desquelles le résultat peut être décrit. Tous les événements de l'expérience doivent trouver place dans l'une ou l'autre de ces catégories, quoiqu'évidemment chaque catégorie ne se trouve pas nécessairement réalisée dans une expérience donnée. Comme chaque catégorie indépendante nous permet d'introduire une mesure de différenciation - c'est à dire de forme ou de structure - dans la description d'un résultat, nous pouvons considérer que sa connaissance nous apporte de l'*information a priori* ou *structurelle*" [Mac Kay 1950 : 296]. L'unité d'information structurelle, le *logon*, est ce qui nous permet d'énoncer une proposition décrivant une caractéristique du résultat indépendante *a priori* des autres caractéristiques. A titre d'exemple, la capacité structurelle d'un microscope, pour une région donnée du plan focal, est son pouvoir de résolution (exprimé en *logons* par cm^2). Remarquons que Warren S. TORGENSON a ultérieurement, mais de manière indépendante semble-t-il (Mac KAY n'est jamais cité, et ne figure pas dans la bibliographie), développé une distinction semblable à celle des composantes structurelle et métrique de l'information scientifique : "La classification des échelles de mesure a été, en un certain sens, fondée sur la *quantité* d'information, au sujet de la propriété mesurée, que représentaient les nombres [...]. Il est également nécessaire de distinguer entre ce que l'on pourrait appeler les *espèces* d'information représentées par les nombres" [Torgerson 1958 : 21]. Un axiome important concernant l'information structurelle est énoncé comme suit : "*Les bornes de l'observation scientifique sont celles de notre vocabulaire logique. Si un phénomène peut être défini (en termes de propositions atomiques de la méthode scientifique), alors il peut en principe être observé*" [Mac Kay 1950 : 305].

La théorie quantique de l'observation scientifique due à Mac Kay renoue explicitement avec la conception kantienne de la connaissance, et avec la description du monde selon WITTGENSTEIN (cf. p.ex. : "Chaque chose se trouve pour ainsi dire dans un espace d'états de choses possibles" ;

"La possibilité de son occurrence dans un état de choses constitue la forme de l'objet" ; "La configuration des objets forme l'état de choses" [Wittgenstein 1921 : § 2.013, 2.0141, 2.0272]). Elle débouche d'autre part sur la théorie statistique de l'information, en assimilant l'information métrique à l'inverse de l'incertitude [Mac Kay 1950 : 294], et en montrant que l'entropie et l'information métrique sont des quantités équivalentes [Mac Kay 1950 : 301]. En ce sens, elle constitue un lien entre le contenu logico-sémantique d'un système classificatoire, et l'analyse statistique que l'on peut réaliser sur les classements d'objets ainsi obtenus.

En appliquant presque terme à terme les propositions de Mac KAY aux résultats du questionnement $\pi : U \rightarrow R^*$, il convient donc de distinguer l'information structurelle contenue dans le questionnaire avant même son administration, et l'information métrique qui résulte du questionnement. Cette dernière est liée au dénombrement des réponses, c'est à dire à une application $d : R^* \rightarrow \mathbb{N}_0$ telle qu'à chaque élément de R^* soit associé le nombre de fois où cette configuration de réponses a été donnée par un répondant. A partir de l'exemple donné par Mac KAY, il est clair que l'information métrique, sous sa forme canonique, résulte de l'application $\pi' : U \rightarrow S^*$, de la population interrogée U sur la structure "atomique" (en composantes élémentaires, ou logons) S^* de R^* ; sa mesure est, comme précédemment, une application $d' : S^* \rightarrow \mathbb{N}_0$ (Cette mesure étant, en pratique, dérivable du dénombrement des réponses observées d , sans toutefois se confondre avec lui). Notre première tâche doit donc être de définir la structure "atomique" de R^* , et pour cela d'identifier l'information correspondant aux faits (possibles) tellement élémentaires qu'ils ne peuvent plus être décomposés en éléments plus petits. Il ne semble pas qu'en l'état actuel de son développement, la linguistique soit parvenue à identifier des unités sémantiques réellement indécomposables : les exemples présentés pour illustrer la notion de *sème* ou de *trait sémantique*, ainsi que les définitions qui en sont données (cf. p.ex. : [Prieto 1966 : 43-45], [Lyons 1978 : 257-271]), ne nous paraissent pas correspondre au niveau d'information le plus primitif (tel que nous le décrivons ci-après).

C'est pourquoi nous avons cru utile de développer une formalisation correspondant à ce qui nous a paru constituer les "atomes" de l'information structurelle contenue dans un questionnaire.

Selon Mac KAY, un logon exprime l'existence d'une possibilité de différenciation la plus élémentaire possible. En adaptant cette notion au problème du questionnement, on peut définir le logon comme le quantum d'information qui suffit à différencier deux configurations de réponses possibles dans R^* ; c'est par conséquent la plus petite distance possible dans l'espace d'observation, l'unité de mesure du questionnaire. Corrélativement, le metron est le quantum d'information qui suffit à différencier deux individus de notre population U . Cette manière de présenter l'information qu'apporte l'administration d'un questionnaire s'accorde avec la définition que donne Lee J. CRONBACH des tests mentaux ("Un test peut être défini comme une procédure systématique pour comparer le comportement de deux personnes ou plus" [Cronbach 1949 : 11]). Dans les paragraphes qui suivent, nous tentons de donner un contenu opératoire à cette notion de quantum d'information structurelle ou métrique.

b) Les opérateurs de relation \mathcal{S} et \mathcal{D} .

L'information structurelle doit, par définition, être recherchée dans les relations sémantiques qu'entretiennent entre elles les configurations de réponses possibles $r^* \in R^*$; nous définirons donc le *logon* comme la relation la plus élémentaire possible unissant deux patrons de réponses quelconques (r_k^*, r_l^*) dans $(R^*)^2$. Par construction, chaque ensemble de réponses R_i^* à une question formelle q_i^* définit sur R^* une partition. En reprenant les considérations développées au § 3.1.3.b sur la construction des variables, nous dirons que, si r_k^* et r_l^* appartiennent à une même partie A_i définie par q_i^* , elles sont *équivalentes selon q_i^** , et nous écrirons : $r_k^* \equiv_i r_l^*$; si chacune d'elles appartient à une partie distincte (par exemple : $r_k^* \in A_i$ & $r_l^* \in B_i$), nous dirons qu'elles sont *non-*

équivalentes selon q_i^* , et nous écrirons : $r_k^* \neq_i r_l^*$. Nous voyons par conséquent que, pour définir l'information structurelle élémentaire (*logon*) contenue dans une question formelle q_i^* , nous devons considérer non seulement un couple de patrons de réponses quelconques $(r_k^*, r_l^*) \in (R^*)^2$, mais également un couple de parties $(A_i, B_i) \in (\mathcal{P}(R_i^*))^2$. Le quantum d'information structurelle de q_i^* doit donc être défini sur le produit cartésien : $(R^*)^2 \times (\mathcal{P}(R_i^*))^2$.

Avant de généraliser ceci à l'ensemble des partitions sur R^* induites par l'ensemble des questions formelles q_i^* , il faut nous assurer que cette formalisation tient compte de toutes les informations sémantiques que nous avons retenues pour la construction des indicateurs et des variables. En effet, nous l'avons rappelé à plusieurs reprises, ce qui intéresse le chercheur est moins l'analyse des réponses aux questions (telles qu'elles sont condensées dans \mathcal{M}) que l'analyse des indicateurs et des variables auxquels celles-ci correspondent. Si nous confrontons nos développements sur la construction des variables (§ 3.1.3.b) à ceux qui concernent la réduction à une forme canonique d'une question concrète (§ 3.2.4.b), force nous est de constater que certains éléments faisant partie intégrante du contenu sémantique de la question originale n'ont pas été retenus dans les questions formelles qui en constituent la forme canonique : ces éléments sont les relations d'ordre strict entre les modalités de réponses présentées, soit que ces relations aient été contenues dans la structure même de l'ensemble des modalités, soit qu'elles aient été imposées au répondant par une consigne d'ordination (au moins partielle) des modalités. Quoi qu'il en soit, ces relations d'ordre n'ont pas été conservées dans la réduction à la forme canonique Q^* ayant servi à construire la "matrice des données". Peut-être eût-il convenu de modifier la procédure de réduction à une forme canonique d'un questionnaire, afin d'intégrer ce type d'information. Ainsi que nous l'avons souligné, nous avons préféré opter pour une forme bâtarde mieux adaptée aux pratiques actuelles des sociologues, tout en conservant la possibilité de concevoir ultérieurement, dans le même cadre théorique

général, d'autres formes canoniques répondant à d'autres besoins des utilisateurs. A l'heure actuelle, il semble que les spécialistes des sciences sociales se contentent de réintroduire ces informations ordinales, qui ne figurent pas explicitement dans \mathcal{M} , au moment du choix des modes de traitement statistique des données (problème des "niveaux de mesure").

Il est clair cependant que, si nous voulons décrire la totalité de l'information structurelle contenue dans les variables qui peut être prise en compte dans l'analyse des données, il nous faut élaborer un cadre formel dans lequel les relations d'ordre strict sur les modalités sont intégrées (Sur l'importance des relations d'équivalence et d'ordre strict pour les fondements de la "mesure" dans les sciences de l'homme, cf. p.ex. les opérateurs binaires P et C dans [Torgerson 1958 : 27-30], ou les relations de dominance et de coïncidence dans [Coombs 1964 : 16]). Cela nous conduit à un approfondissement de notre analyse, jusqu'à l'explicitation des relations sémantiques internes à chacune des réponses possibles. En nous limitant aux propriétés structurelles du questionnaire, et sans tenir compte de la possibilité de construction ultérieure de variables à partir des réponses, deux cas sont à considérer :

- la réponse exprime une simple partition des modalités présentées (exemples : consignes de classification ; consignes de sélection de modalités non ordonnées), et son contenu structurel se réduit à un ensemble de relations d'équivalence et de non équivalence ;

- la réponse exprime une ordination au moins partielle des modalités présentées (exemples : consignes d'ordination ; consignes de sélection de modalités totalement ordonnées), et son contenu structurel comporte aussi des relations d'ordre strict entre les modalités.

Dans le premier cas, l'unité d'information élémentaire entre deux modalités peut être l'une des relations d'équivalence ou de non équivalence déjà évoquées ; ce cas ne pose donc pas de problème particulier. Dans le second cas, le problème est de tenir compte de chacune des relations d'ordre strict entre

les modalités figurant dans la sélection contenue dans la réponse. S'il s'agit d'une consigne d'ordination, la taille de la sélection est fixée par la consigne ; s'il s'agit d'une consigne de sélection unique dans un ensemble ordonné, la taille de la sélection est définie par le rang de la modalité sélectionnée explicitement (cf. § 3.1.3.b).

La formalisation que nous présentons ci-après permet de définir une unité d'information plus élémentaire que les relations d'équivalence, de non équivalence, et d'ordre strict, que nous avons utilisées jusqu'ici ; cette unité d'information permet par conséquent de rendre compte de ces relations moins primitives. Définie sur le produit cartésien : $(R^*)^2 \times (\mathcal{P}(R^*))^2$, cette unité correspondra au quantum d'information structurelle, ou *logon*. Définie sur le produit cartésien : $U^2 \times (\mathcal{P}(U))^2$, elle correspondra à l'unité d'information métrique, ou *metron*. C'est pourquoi, afin de conserver à notre formalisation son caractère de généralité, nous la présentons à partir d'un ensemble quelconque E .

Soient un ensemble E quelconque, deux parties de E quelconques : $A \subset E$ et : $B \subset E$, et deux éléments de E quelconques : $x \in E$ et : $y \in E$. Nous définirons sur $E^2 \times (\mathcal{P}(E))^2$ les relations :

$$\mathcal{S}(A, B, x, y) \stackrel{\text{df}}{=} x \in A \quad \& \quad y \in B \quad \& \quad x \in B$$

$$\mathcal{D}(A, B, x, y) \stackrel{\text{df}}{=} x \in A \quad \& \quad y \in B \quad \& \quad x \notin B$$

La relation \mathcal{S} présente les propriétés suivantes :

① Pour A et B fixés, \mathcal{S} est un préordre dans E^2 . En effet :

- \mathcal{S} est réflexif dans E^2 :

$$\forall x \in E, \forall (A, B) \in (\mathcal{P}(E))^2 : \mathcal{I}(A, B, x, x)$$

(on a : $x \in A$ & $x \in B$).

- \mathcal{I} est transitif dans E^2 :

$$\forall x \in E, \forall y \in E, \forall z \in E, \forall (A, B) \in (\mathcal{P}(E))^2 :$$

$$\mathcal{I}(A, B, x, y) \quad \& \quad \mathcal{I}(A, B, y, z) \Rightarrow \mathcal{I}(A, B, x, z)$$

puisque, par définition :

$$\mathcal{I}(A, B, x, y) \quad \& \quad \mathcal{I}(A, B, y, z) =$$

$$x \in A \quad \& \quad x \in B \quad \& \quad y \in A \quad \& \quad y \in B \quad \& \quad z \in B$$

$$\mathcal{I}(A, B, x, z) = x \in A \quad \& \quad x \in B \quad \& \quad z \in B$$

② Pour x et y fixés, \mathcal{I} est un préordre dans $(\mathcal{P}(E))^2$. En effet :

- \mathcal{I} est réflexif dans $(\mathcal{P}(E))^2$:

$$\forall A \in \mathcal{P}(E), \forall (x, y) \in E^2 : \mathcal{I}(A, A, x, y)$$

(on a : $x \in A$ & $y \in A$)

- \mathcal{I} est transitif dans $(\mathcal{P}(E))^2$:

$$\forall A \in \mathcal{P}(E), \forall B \in \mathcal{P}(E), \forall C \in \mathcal{P}(E), \forall (x, y) \in E^2 :$$

$$\mathcal{I}(A, B, x, y) \quad \& \quad \mathcal{I}(B, C, x, y) \Rightarrow \mathcal{I}(A, C, x, y)$$

puisque, par définition :

$$\mathcal{I}(A, B, x, y) \quad \& \quad \mathcal{I}(B, C, x, y) =$$

$$x \in A \quad \& \quad x \in B \quad \& \quad y \in B \quad \& \quad x \in C \quad \& \quad y \in C$$

$$\mathcal{I}(A, C, x, y) = x \in A \quad \& \quad x \in C \quad \& \quad y \in C$$

③ \mathcal{I} est réflexif dans $E^2 \times (\mathcal{P}(E))^2$:

$$\forall A \in \mathcal{P}(E), \forall x \in E : \mathcal{I}(A, A, x, x)$$

(on a : $x \in A$)

La relation \mathcal{D} présente les propriétés suivantes :

① Pour A et B fixés, \mathcal{D} est antiréflexif et antisymétrique dans E^2 :

- \mathcal{D} est antiréflexif :

$$\forall x \in E, \forall (A, B) \in (\mathcal{P}(E))^2 : \overline{\mathcal{D}(A, B, x, x)}$$

puisque : $\mathcal{D}(A, B, x, x) \Rightarrow x \in B \text{ \& } x \notin B$

- \mathcal{D} est antisymétrique :

$$\forall (x, y) \in E^2, \forall A \in \mathcal{P}(E), \forall B \in \mathcal{P}(E) :$$

$$\overline{\mathcal{D}(A, B, x, y) \text{ \& } \mathcal{D}(A, B, y, x)}$$

puisque : $\mathcal{D}(A, B, x, y) \text{ \& } \mathcal{D}(A, B, y, x) \Rightarrow$

$$x \in B \text{ \& } x \notin B \text{ \& } y \in B \text{ \& } y \notin B$$

② Pour x et y fixés, \mathcal{D} est antiréflexif et antisymétrique dans $(\mathcal{P}(E))^2$:

- \mathcal{D} est antiréflexif :

$$\forall A \in \mathcal{P}(E), \forall (x, y) \in E^2 : \overline{\mathcal{D}(A, A, x, y)}$$

puisque : $\mathcal{D}(A, A, x, y) \Rightarrow x \in A \text{ \& } x \notin A$

- \mathcal{D} est antisymétrique :

$$\forall (x, y) \in E^2, \forall A \in \mathcal{P}(E), \forall B \in \mathcal{P}(E) :$$

$$\overline{\mathcal{D}(A, B, x, y) \text{ \& } \mathcal{D}(B, A, x, y)}$$

puisque : $\mathcal{D}(A, B, x, y) \ \& \ \mathcal{D}(B, A, x, y) \Rightarrow$

$$x \in A \ \& \ x \notin A \ \& \ x \in B \ \& \ x \notin B$$

③ Pour \mathcal{D} , la transitivité n'a de sens ni dans E^2 , ni dans $(\mathcal{P}(E))^2$.

En effet :

- pour A et B fixés, on a, quels que soient x, y , et z :

$$\mathcal{D}(A, B, x, y) \ \& \ \mathcal{D}(A, B, y, z) \Rightarrow y \in B \ \& \ y \notin B$$

- pour x et y fixés, on a, quels que soient A, B , et C :

$$\mathcal{D}(A, B, x, y) \ \& \ \mathcal{D}(B, C, x, y) \Rightarrow x \in B \ \& \ x \notin B$$

Enfin, \mathcal{P} et \mathcal{D} sont mutuellement incompatibles :

$$\forall (A, B) \in (\mathcal{P}(E))^2, \forall (x, y) \in E^2 : \overline{\mathcal{P}(A, B, x, y) \ \& \ \mathcal{D}(A, B, x, y)}$$

puisque, par définition :

$$\mathcal{P}(A, B, x, y) \ \& \ \mathcal{D}(A, B, x, y) \Rightarrow x \in B \ \& \ x \notin B$$

Il nous reste à montrer que ces deux opérateurs de relation sont bien susceptibles de rendre compte des relations d'équivalence, de non équivalence, et d'ordre strict. La relation d'équivalence est définie sur un couple d'éléments de E appartenant à une même classe d'équivalence. Si l'on a x est équivalent à y relativement à A : $x \equiv_A y$, on a nécessairement :

$$x \in A \ \& \ y \in A = \mathcal{P}(A, A, x, y) = \mathcal{P}(A, A, y, x)$$

La relation de non équivalence, par contre, est définie sur un couple d'éléments de E appartenant à deux classes disjointes. Si x est non-équivalent

à y , relativement au couple de parties A et B , c'est que : $A \cap B = \emptyset$ (en raison de la transitivité de la relation d'équivalence), et que soit : $x \in A$ & $y \in B$, soit : $x \in B$ & $y \in A$. En prenant comme exemple la première éventualité, on a par conséquent :

$$x \in A \quad \& \quad x \notin B \quad \& \quad y \notin A \quad \& \quad y \in B = \\ \mathcal{D}(A, B, x, y) \quad \& \quad \mathcal{D}(B, A, y, x)$$

Dans la seconde éventualité, on aurait eu : $\mathcal{D}(A, B, y, x)$ & $\mathcal{D}(B, A, x, y)$. Enfin, la relation d'ordre strict est définie sur un couple d'éléments de E et un couple de parties dont l'une est incluse strictement dans l'autre : si $A \subset B$, et si $x \in B$ & $y \in \bar{A} \cap B$, alors $y > x$. On a par conséquent :

$$x \in A \quad \& \quad x \in B \quad \& \quad y \notin A \quad \& \quad y \in B = \\ \mathcal{P}(A, B, x, y) \quad \& \quad \mathcal{D}(B, A, y, x)$$

Ces opérateurs de relations présentent d'autres propriétés intéressantes, et permettent de rendre compte d'autres formes de relations binaires que les trois que nous venons d'examiner. Leur étude plus approfondie sortirait du cadre d'une recherche sur le questionnaire. Aussi allons-nous maintenant examiner plutôt comment les relations \mathcal{P} et \mathcal{D} s'intègrent dans les théories de la mesure propres aux sciences de l'homme.

c) Les "niveaux de mesure" dans les sciences de l'homme.

De nombreux auteurs traitant des problèmes de la mesure dans les sciences de l'homme se réfèrent en premier lieu à la théorie de N.R. CAMPBELL sur la mesure dans les sciences physiques (cf. p.ex. : [Torgerson 1958 : 13-14, 21-22], [Reuchlin 1962 : 15-21], [Suppes 1963 : 15-17]). Selon ces auteurs, CAMPBELL définit la mesure comme l'assignation de numéros ("*numerals*") pour présenter les propriétés des systèmes matériels, en vertu

des lois qui gouvernent ces propriétés. Si ces systèmes peuvent être ordonnés selon la propriété à mesurer, sans toutefois que les numéros correspondant à ces systèmes soient susceptibles d'être additionnés, on parlera de grandeurs *intensives*, ou *qualitatives* ; si l'addition est possible, il s'agit de nombres au sens strict, et l'on parlera alors de grandeurs *extensives*, ou *quantitatives*. D'autre part, si les grandeurs peuvent être mesurées directement, à partir des lois qui relient entre eux plusieurs exemplaires des systèmes matériels examinés (grandeurs *de type A*), on parlera de *mesure fondamentale* (exemples : mesures de longueur ou de volume). Si par contre les grandeurs ne peuvent qu'être déduites de lois reliant la propriété mesurée à d'autres propriétés (grandeurs *de type B*), on parlera de *mesure dérivée* (exemples : mesures de densité, obtenues par calcul du rapport : masse/volume ; et toutes les mesures de grandeurs très importantes ou très petites à l'échelle humaine).

Les conceptions de CAMPBELL ne sont évidemment pas applicables telles quelles aux sciences de l'homme ; mais elles constituent une base de départ qui a facilité les recherches ultérieures de STEVENS, COOMBS, SUPPES, etc. En ce qui concerne les formes de mesure (ou les types de grandeurs), TORGERSON a défini une troisième forme qu'il appelle la *mesure par décision arbitraire* ("*measurement by fiat*" [Torgerson 1958 : 21-22]), fondée sur des relations présumées entre des observations et un concept donné. Cette définition s'applique en particulier à de nombreuses mesures en sociologie (indice de statut socio-économique, p.ex.) et en psychologie (cf. : "l'intelligence, c'est ce que mesure mon échelle"). Comme le souligne STEVENS, les psychologues s'efforcent souvent "d'évaluer les divers aspects du comportement par le moyen de ce que l'on peut appeler des *indicateurs*. Ceux-ci sont des *effets* ou des *covariations* qui sont liés aux dimensions psychologiques par des lois *inconnues*. Cette procédure est inévitable au stade actuel de notre développement" [Stevens 1951 : 47]. Dans cette forme de mesure, les caractéristiques (liées au concept) ont une signification intrinsèque directe, et une signification opératoire indirecte seulement, par le biais des indicateurs (qui n'ont qu'une signification opératoire). Une autre forme de mesure est également proposée par SUPPES et ZINNES, sous le nom de *mesure*

par index ("pointer measurement" [Suppes 1963 : 20-22]) : c'est la mesure obtenue par lecture directe sur un instrument validé (Cette forme, qui pourrait peut-être s'appliquer par extension aux enquêtes répétitives, ou à l'administration des épreuves psychologiques standardisées, nous a cependant paru moins intéressante que les trois formes précédentes).

En ce qui concerne la définition de la mesure, la conception de CAMPBELL présente l'avantage d'englober des opérations qu'avant lui on ne considérait généralement pas comme des mesures proprement dites : entre dans cette catégorie par exemple l'évaluation de la dureté des minéraux (échelle de dureté de Friedrich MOHS). En définissant la mesure comme une application d'un ensemble d'observations dans un ensemble de numéros, c'est-à-dire d'étiquettes numériques, CAMPBELL a posé les bases d'une théorie applicable à la plupart des sciences de l'homme (comme l'a bien vu STEVENS), même si lui-même n'a considéré que les propriétés ordinales et les propriétés cardinales des nombres désignés par ces étiquettes. En outre, cette théorie établit une distinction importante entre d'une part le caractère direct ou indirect de l'observation (mesure fondamentale / mesure dérivée), et d'autre part la forme du résultat de la mesure (ordination / mesure au sens strict), dépendant de la nature des grandeurs considérées (qualitatives / quantitatives). La relation existant entre ces deux dimensions est due à ce que seules les grandeurs quantitatives peuvent être mesurées directement. Enfin, la conception de CAMPBELL pose que la notion de grandeur mesurable est inséparable de la notion d'ordre sur les caractéristiques de cette grandeur. D'ailleurs, l'égalité y est définie à partir de la relation d'ordre strict :

$$(a \succ b \quad \& \quad b \succ a \quad \& \quad \forall c ((a > c \Rightarrow b > c) \quad \& \quad (c > a \Rightarrow c > b))) \\ \Rightarrow a = b$$

La relation d'ordre est donc une relation primitive dans cette conception de la mesure.

A la suite de CAMPBELL, les spécialistes des sciences de l'homme ont développé diverses théories de la mesure adaptées à leurs besoins. Ces théories sont trop connues pour qu'il soit nécessaire de les exposer en détail (on en trouvera d'ailleurs un résumé assez complet et très clair dans [Reuchlin 1962 : 51-65]) ; aussi nous contenterons-nous d'un bref rappel historique. La première en date de ces théories paraît être la "théorie des échelles de mesure" de S.S. STEVENS, qui distingue quatre niveaux de mesure hiérarchisés : échelle nominale, échelle ordinale, échelle d'intervalles, échelle de rapports (cf. [Stevens 1946]). La valeur pratique de cette théorie se mesure à sa diffusion : les trois premiers niveaux sont très largement utilisés pour classer les multiples techniques statistiques applicables aux sciences de l'homme, et pour définir les conditions sur les variables que leurs fondements mathématiques imposent (cf. p.ex. : [Siegel 1956 : 21-30], [Reuchlin 1976]). Cette trichotomie correspond en effet à trois grandes familles de techniques statistiques, applicables à des données sur lesquelles on postule des propriétés mathématiques simples. On peut d'ailleurs retrouver cette classification trichotomique des données dans des domaines scientifiques assez éloignés des sciences de l'homme, comme la biologie (cf. [Lison 1958 : 1-2]) : données d'énumération (classes d'équivalence), de sériation (équivalence et ordre strict), de mesure proprement dite (équivalence, ordre strict, et distance). STEVENS ayant défini la mesure comme "l'attribution de numéros aux choses en vue de rendre compte des faits et des conventions à propos des faits" [Stevens 1946 : 680], cette trichotomie des niveaux de mesure correspond en fait à la trichotomie : numéro / nombre ordinal / nombre cardinal. L'apport de STEVENS par rapport à CAMPBELL est d'avoir considéré déjà comme une mesure la simple attribution de noms aux choses, et d'avoir introduit entre les grandeurs intensives (échelle ordinale) et les grandeurs extensives (échelle de rapports) un niveau intermédiaire, l'échelle d'intervalles (sans origine naturelle).

Warren S. TORGERSON a restructuré la classification de STEVENS d'une part en écartant du champ de la mesure les échelles nominales, d'autre part en systématisant la description des trois niveaux restant. Pour cela, TORGERSON considère deux dimensions indépendantes applicables à des

grandeurs au sens de CAMPBELL : définition ou non d'une distance entre les grandeurs, et définition ou non d'une origine naturelle constituant le point zéro de la mesure [Torgerson 1958 : 15-16]. Il introduit ainsi un niveau de mesure supplémentaire : celui des échelles ordinales avec une origine naturelle. Clyde H. COOMBS a enrichi les classifications de STEVENS et de TORGERSON, en introduisant en particulier deux niveaux intermédiaires : échelle partiellement ordonnée (entre échelle nominale et échelle ordinale), et échelle métrique ordonnée (entre échelle ordinale et échelle d'intervalles) [Coombs 1953 : 540-552]. D'autres systèmes de classification des niveaux de mesure dans les sciences de l'homme ont été proposés, en particulier par COOMBS et par SUPPES, et leurs collaborateurs ; ces systèmes visent à une meilleure adaptation, aux propriétés observables des données, de la structure formelle destinée à en rendre compte. En outre, on peut trouver, dans la formalisation très générale des structures numériques développée par Stig KANGER, un cadre théorique permettant de distinguer, si besoin est, de nouveaux types d'échelle de mesure [Kanger 1972].

La présentation de ces théories sortirait du cadre de notre propos, qui est d'établir un lien entre notre formalisation du contenu sémantique des réponses possibles aux questions d'un questionnaire (information structurelle), et les techniques usuelles d'analyse des données. Or, pour l'essentiel, ces techniques trouvent leur place dans les trois premiers niveaux de la classification de STEVENS. Toutefois, par souci de systématisation, nous avons adopté un compromis entre cette classification simple mais commode, et la classification de TORGERSON. Nous distinguerons donc six "niveaux de mesure", dont cinq correspondent à l'union de ces deux classifications, le sixième (N_0) étant ajouté par symétrie :

- N = échelle nominale, pour laquelle sont définies, sur l'ensemble des couples d'éléments, les seules relations d'équivalence. Exemple : ensemble des réponses possibles au sens strict à une question sur la région habitée.

- N_o = échelle nominale avec origine, analogue formellement à N , avec en plus la définition d'un élément considéré comme point de référence. Exemple : ensemble des réponses possibles au sens strict à un item d'un test de connaissance ou de niveau, dans lequel il existe une réponse vraie et une seule.

- O = échelle ordinale, pour laquelle sont définies, sur l'ensemble des couples d'éléments, les relations d'équivalence et d'ordre strict. Exemple : ensemble des réponses possibles au sens strict à un item d'attitude de la forme : "peu favorable, ..., très favorable".

- O_o = échelle ordinale avec origine fixe, analogue formellement à O , avec en plus un point zéro. Exemple : ensemble des réponses possibles au sens strict à un item de la forme : "ne plait pas du tout, plait un peu, ..., plait énormément".

- M = échelle d'intervalles (ou : échelle métrique), pour laquelle sont définies des distances entre chaque couple d'éléments. Exemple : repères temporels dans un questionnaire autobiographique.

- M_o = échelle métrique avec origine fixe (ou : échelle de rapports), analogue formellement à M avec en plus un point zéro. Exemple : ensemble des réponses aux questions portant sur des quantités dénombrables telles que le revenu, le nombre d'enfants, etc.

Nous verrons cependant, dans les pages qui suivent (§ 3.3.2a et b), que les échelles avec origine fixe ont peu retenu l'attention des statisticiens des sciences de l'homme. Elles présentent pourtant une grande importance pratique, en raison par exemple de l'intérêt qu'il y aurait à mesurer la dispersion des réponses à un item de niveau par rapport à la "bonne" réponse (c'est-à-dire dans N_o plutôt que dans N), ou à mesurer la liaison entre une prévision électorale et le vote lui-même dans $M_o \times M_o$, plutôt que dans $M \times M$.

d) Représentation des niveaux de mesure par les opérateurs \mathcal{Y} et \mathcal{D}

Les opérateurs de relation définis au § 3.3.1.b permettent de rendre compte, en termes de *logons*, de l'information structurelle correspondant à chacun des niveaux de mesure que nous avons retenus. En principe, il devrait suffire d'utiliser l'opérateur \mathcal{D} , qui exprime une relation élémentaire de divergence ou de dissemblance entre deux individus. En effet, si l'on admet que l'objectif de l'administration d'un test ou d'un questionnaire est d'établir des différences entre les individus, c'est bien la relation exprimée par l'opérateur \mathcal{D} qui remplit cette fonction ; c'est donc elle qui constitue l'unité d'information structurelle appelée *logon*. Nous avons cependant cru bon de conserver l'opérateur \mathcal{Y} , qui exprime une relation élémentaire d'analogie ou de similitude entre deux individus, pour une raison essentiellement pratique : dans certains cas, il est plus facile d'écrire la formule d'un indice de dispersion ou de liaison en utilisant l'opérateur \mathcal{Y} , qu'en utilisant l'opérateur \mathcal{D} . Cette commodité d'écriture ne doit cependant pas masquer le fait qu'il n'existe qu'un type d'information structurelle élémentaire. Ceci est mis en évidence par la complémentarité des deux opérateurs dans leur domaine de définition : pour tout quadruplet $(A, B, x, y) \in E^2 \times (\mathcal{P}(E))^2$ tel que $x \in A$ & $y \in B$, on a nécessairement soit $\mathcal{Y}(A, B, x, y)$, soit $\mathcal{D}(A, B, x, y)$.

Dans ce paragraphe, nous présentons donc la traduction de chacun des six niveaux de mesure décrits au paragraphe précédent, en termes de composantes informationnelles de type \mathcal{Y} ou \mathcal{D} . Ainsi que nous le verrons, chaque niveau de mesure, défini sur un ensemble E d'indicateurs, sera caractérisé par un sous-ensemble Y de $\mathcal{P}(E)$. Pour des raisons qui apparaîtront dans le cours de notre développement, nous aurons recours à deux formes de traduction distinctes, selon que l'échelle a une origine fixe ou non. Dans le premier cas (échelle avec origine fixe), le principe de la traduction est d'indiquer, pour chaque quadruplet élément de $E^2 \times Y^2$, s'il appartient au domaine de définition de ces relations, et, si oui, quelle est la nature de l'information qui lui correspond. Dans le second cas (échelle sans origine fixe), il suffirait de considérer les triades de

forme (A, x, y) , avec $A \in Y$, et d'admettre que la relation correspondante est :

- une relation \mathcal{S} si $y \in A$ & $x \in A$
- une relation \mathcal{D} si $y \in A$ & $x \notin A$
- non définie si $y \notin A$.

Plutôt que de créer de nouveaux opérateurs à trois arguments afin de rendre compte des échelles sans origine fixe, nous avons préféré appliquer la procédure décrite pour les échelles avec origine, mais en identifiant les relations \mathcal{S} et \mathcal{D} dans le produit cartésien $E^2 \times (\{E\} \times Y)$. Cette dernière procédure est bien équivalente à la précédente, puisque, pour tout quadruplet $(E, A, x, y) \in E^2 \times (\mathcal{P}(E))^2$, on a toujours $x \in E$, quel que soit x . D'autre part, nous présenterons également les niveaux de mesure de façon condensée, en dénombrant les relations \mathcal{S} et les relations \mathcal{D} dans E^2 .

Nous définirons une *échelle nominale sans origine fixe* par l'ensemble des parties de E de cardinal 1 :

$$N = \mathcal{P}_1(E)$$

Chaque partie élément de N contient donc un indicateur et un seul ; l'union des parties de N est égale à E , et l'intersection de deux parties distinctes quelconques de N est toujours vide. A titre d'exemple, soit :

$$E = \{w, x, y, z\}$$

$$N = \{A, B, C, D\} = \{\{w\}, \{x\}, \{y\}, \{z\}\}$$

L'énumération des relations \mathcal{S} et \mathcal{D} dans $E^2 \times (\{E\} \times N)$, et leurs dénombrements dans E^2 , sont représentés par la figure n° 3.24. En généralisant ces résultats, on peut dire qu'une échelle nominale sans origine fixe est caractérisée par les propriétés suivantes :

ABCD

WW	S
WX	D
WY	D
WZ	D
XW	D
XX	S
XY	D
XZ	D
YW	D
YX	D
YY	S
YZ	D
ZW	D
ZX	D
ZY	D
ZZ	S

\mathcal{D}					\mathcal{S}				
	W	X	Y	Z		W	X	Y	Z
W	0	1	1	1	W	1	0	0	0
X	1	0	1	1	X	0	1	0	0
Y	1	1	0	1	Y	0	0	1	0
Z	1	1	1	0	Z	0	0	0	1

Figure n° 3.24 : Enumération dans $E^2 \times (\{E\} \times N)$ et dénombrement dans E^2 des relations \mathcal{S} et \mathcal{D} correspondant à l'échelle nominale sans origine fixe $N = \{ \{w\}, \{x\}, \{y\}, \{z\} \}$.

- pour tout couple d'indicateurs $(x, y) \in E^2$, il y a toujours une relation \mathcal{S} ou \mathcal{D} et une seule ;
- pour tout couple composé d'éléments identiques, il s'agit d'une relation \mathcal{S} ;
- pour tout couple composé d'éléments différents, il s'agit d'une relation \mathcal{D} .

En conséquence, si l'on pose : $|E| = K$, dans une échelle nominale à K indicateurs et K classes, il y a K relations \mathcal{S} et $K(K-1)$ relations \mathcal{D} . Une autre propriété de N , qui n'apparaît pas ici, est que l'énumération des relations \mathcal{S} et \mathcal{D} dans $E^2 \times (\{E\} \times N)$, et cette même énumération dans $E^2 \times N^2$, se traduisent par des dénombrements identiques dans E^2 . Toutes ces propriétés sont spécifiques de l'échelle nominale sans origine fixe.

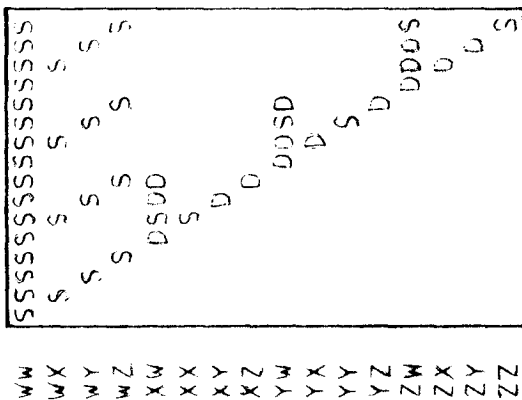
Nous définirons une *échelle nominale avec origine fixe* par l'ensemble N_o constitué de la partie de E contenant l'élément origine o et lui seul, et de toutes les parties de E de cardinal 2 dont l'un des éléments est l'origine o :

$$N_o = \{o\} \cup \{X : X \in \mathcal{P}_2(E) \ \& \ \forall y \in E, X = \{o, y\}\}$$

Cette définition se justifie par l'analyse de l'exemple donné au paragraphe précédent : dans une question de test de connaissances ou de niveau à consigne de sélection unique, et pour laquelle il n'y a qu'une réponse vraie, le sujet qui sélectionne une réponse fausse exprime implicitement le jugement que cette dernière est "supérieure" à la réponse vraie. Dans la terminologie de COOMBS, il s'agit en fait d'un cas particulier d'échelle partiellement ordonnée. Si, à titre d'exemple, nous reprenons le même ensemble d'indicateurs E que précédemment, nous avons maintenant :

$$N_o = \{A, B, C, D\} = \{\{w\}, \{w, x\}, \{w, y\}, \{w, z\}\}$$

AAAABBBBCCCCDDDD
 ABCDABCCDABCCDABCD



WW
 WX
 WY
 WZ
 XW
 XX
 XY
 XZ
 YW
 YX
 YY
 YZ
 ZW
 ZX
 ZY
 ZZ

\mathcal{Y}						\mathcal{D}			
w	x	y	z	w	x	y	z		
16	4	4	4	0	0	0	0		
1	1	0	0	3	0	1	1		
1	0	1	0	3	1	0	1		
1	0	0	1	3	1	1	0		

Figure n° 3.25 : Énumération dans $E^2 \times N^2$ et dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} correspondant à l'échelle nominale avec origine fixe $N_0 = \{ \{w\}, \{w,x\}, \{w,y\}, \{w,z\} \}$.

avec $A = \{w\}$ = origine de l'échelle. La figure n° 3.25 présente l'énumération des relations élémentaires dans $E^2 \times N_o^2$, et leur dénombrement dans E^2 . On voit sur cet exemple qu'une échelle nominale avec origine fixe ainsi formalisée se décompose en une échelle nominale simple :

$$N = \{B, C, D\} = \{ \{x\}, \{y\}, \{z\} \}$$

et une origine $A = \{w\}$ telle que tout autre indicateur lui soit supérieur ; en effet, pour tout couple $(y, B) \in E \times N_o$, tel que $y \in B$, avec $y \neq w$ et $A \neq B$, on a :

$$\mathcal{S}(A, B, w, y) \quad \& \quad \mathcal{D}(B, A, y, w)$$

En conséquence, le dénombrement dans E^2 fait apparaître $(K-1)^2 + (K-1)(K-2) = (K-1)(2K-3)$ relations \mathcal{D} , et $K^2 + 2(K-1) + K(K-1) = K^2 + (K-1)(K+2)$ relations \mathcal{S} . L'énumération des relations \mathcal{D} dans $E^2 \times N_o^2$ et dans $E^2 \times (\{E\} \times N_o)$ donnent lieu à des dénombrements identiques dans E^2 ; cette identité ne se vérifie pas pour les relations \mathcal{S} .

Nous poserons qu'une *échelle ordinale sans origine fixe* est un ensemble O de K parties distinctes non vides de E (avec $|E| = K$, comme précédemment), totalement ordonnée par une relation d'inclusion stricte. Par conséquent, pour toute partie de cardinal c tel que $1 < c < K$, il existe toujours une partie qui ne diffère de celle-ci que par un élément en plus, et une autre partie qui n'en diffère que par un élément en moins. De plus, la partie de cardinal K est identique à l'ensemble E . La partie de cardinal 1 sera appelée l'*origine arbitraire* de l'échelle, nécessaire à sa description, mais sans valeur sémantique particulière. A titre d'exemple, soit :

$$O = \{A, B, C, D\} = \{ \{w\}, \{w, x\}, \{w, x, y\}, \{w, x, y, z\} \}$$

L'énumération des relations correspondantes dans $E^2 \times (\{E\} \times O)$ et leur dénombrement dans E^2 est représenté par la figure n° 3.26. On y constate

ABCD

WW	SSSS
WX	SSS
WY	SS
WZ	S
XW	DSSS
XX	SSS
XY	SS
XZ	S
YW	DDSS
YX	DSS
YY	SS
YZ	S
ZW	DDDS
ZX	DDS
ZY	DS
ZZ	S

\mathcal{Y}	w	x	y	z	\mathcal{D}	w	x	y	z
w	4	3	2	1	w	0	0	0	0
x	3	3	2	1	x	1	0	0	0
y	2	2	2	1	y	2	1	0	0
z	1	1	1	1	z	3	2	1	0

Figure n° 3.26 : Enumération dans $E^2 \times (\{E\} \times O)$ et dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} correspondant à l'échelle ordinale sans origine fixe $O = \{ \{w\}, \{w,x\}, \{w,x,y\}, \{w,x,y,z\} \}$.

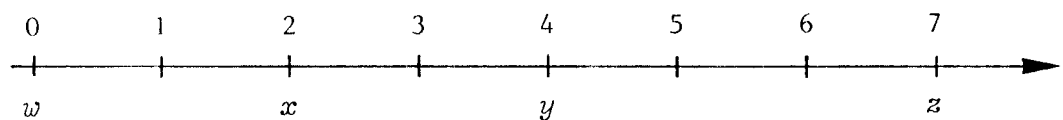
qu'il y a $\sum_{i=1}^{i=n} (K-i) i$ relations \mathcal{D} , qui expriment l'ensemble des relations d'ordre qui caractérisent l'échelle ordinaire sans origine fixe.

Nous dirons qu'une *échelle ordinaire avec origine fixe* est un ensemble O_o analogue à l'ensemble O défini à l'alinéa précédent ; la différence entre O et O_o réside dans la procédure d'énumération et de dénombrement des relations \mathcal{S} et \mathcal{D} , qui se déroule dans $E^2 \times O_o^2$. Pour illustrer cette définition, reprenons notre exemple précédent :

$$O_o = \{ \{w\}, \{w,x\}, \{w,x,y\}, \{w,x,y,z\} \}$$

en posant : $A = \{w\}$ = origine de l'échelle. L'énumération et le dénombrement des relations \mathcal{S} et \mathcal{D} , représentés par la figure n° 3.27, peuvent être comparés à ceux de l'échelle O correspondante (figure n° 3.26). On y constate que l'effet de l'origine fixe se traduit, dans le produit cartésien $E^2 \times O_o$, par la multiplication par r du nombre de relations \mathcal{S} et \mathcal{D} correspondant à l'indicateur de rang r dans l'échelle. De plus, comme on a ici : $D = E$, il y a identité des énumérations dans $E^2 \times (\{E\} \times O)$ et dans $E^2 \times (\{D\} \times O)$.

Nous définirons une *échelle métrique sans origine fixe* ("échelle d'intervalles") comme un ensemble M de parties non vides de E non nécessairement distinctes, tel qu'il existe un sous-ensemble $O \subset M$, contenant un exemplaire et un seul de toutes les parties éléments de M , tel que O soit totalement ordonné strictement. Pour comprendre cette définition, prenons l'exemple de l'échelle d'intervalles ci-dessous (graduée à partir d'une origine arbitraire $A = \{w\}$) à laquelle correspondent quatre indicateurs (réponses possibles), figurés par des lettres :



AAAABBBBCCCCDDDD
 ABCDABCDABCDABCD

SSSSSSSSSSSSSSSS
SSS SSS SSS SSS
SS SS SS SS
S S S S
DSSSDSSSDSSSS
SSS SSS SSS
SS SS SS
S S S S
DDSSDDSS
DSS DSS
SS SS
S S
DDDS
DDS
DS
S

WX
 WY
 WZ
 XW
 XX
 XY
 XZ
 YW
 YX
 YY
 YZ
 ZW
 ZX
 ZY
 ZZ

\mathcal{Y}				\mathcal{D}			
W	X	Y	Z	W	X	Y	Z
16	12	8	4	0	0	0	0
9	9	6	3	3	0	0	0
4	4	4	2	4	2	0	0
1	1	1	1	3	2	1	0

Figure n° 3.27 : Enumération dans $E^2 \times O_0^2$ et dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} correspondant à l'échelle ordinale avec origine fixe $O_0 = \{ \{w\}, \{w,x\}, \{w,x,y\}, \{w,x,y,z\} \}$.

Exprimé en unités de l'échelle, l'intervalle qui sépare deux indicateurs est selon les cas égal à 2 ou à 3 . Si l'on voulait représenter cette échelle en termes d'inclusion d'ensembles, on pourrait écrire, en remplaçant le nom des indicateurs manquants par un astérisque :

$$M = \{ \{w\} , \{w,*\} , \{w,*,x\} , \{w,*,x,*\} , \{w,*,x,*,y\} , \\ \{w,*,x,*,y,*\} , \{w,*,x,*,y,*,*\} , \{w,*,x,*,y,*,*,z\} \}$$

Si l'on réécrit cette formule en supprimant les astérisques, on obtient :

$$M = \{ \{w\} , \{w\} , \{w,x\} , \{w,x\} , \{w,x,y\} , \{w,x,y\} , \\ \{w,x,y\} , \{w,x,y,z\} \}$$

Dans cette formule, c'est la répétition d'un même élément de $\mathcal{P}(E)$ qui rend compte de la longueur des intervalles (en termes d'unités de l'échelle). La figure n° 3.28 présente l'énumération et le dénombrement des relations \mathcal{P} et \mathcal{D} correspondant à cette formalisation. En comparant cette figure à la figure n° 3.26, on constate que, selon cette formalisation, l'échelle ordinale est analogue à une échelle métrique à intervalles égaux à l'unité. L'introduction entre les indicateurs d'intervalles, de longueur égale à l unités, se traduit par la multiplication par l du nombre de relations \mathcal{D} correspondant à ces indicateurs.

Nous dirons enfin qu'une *échelle métrique avec origine fixe* ("échelle de rapports") est un ensemble M_0 analogue à l'ensemble M défini précédemment, mais pour lequel l'énumération et le dénombrement des relations \mathcal{P} et \mathcal{D} est entrepris dans $E^2 \times M_0^2$. La figure n° 3.29 représente le résultat de ces énumérations et dénombrements pour l'ensemble $M_0 = M$.

ABCDEFGH

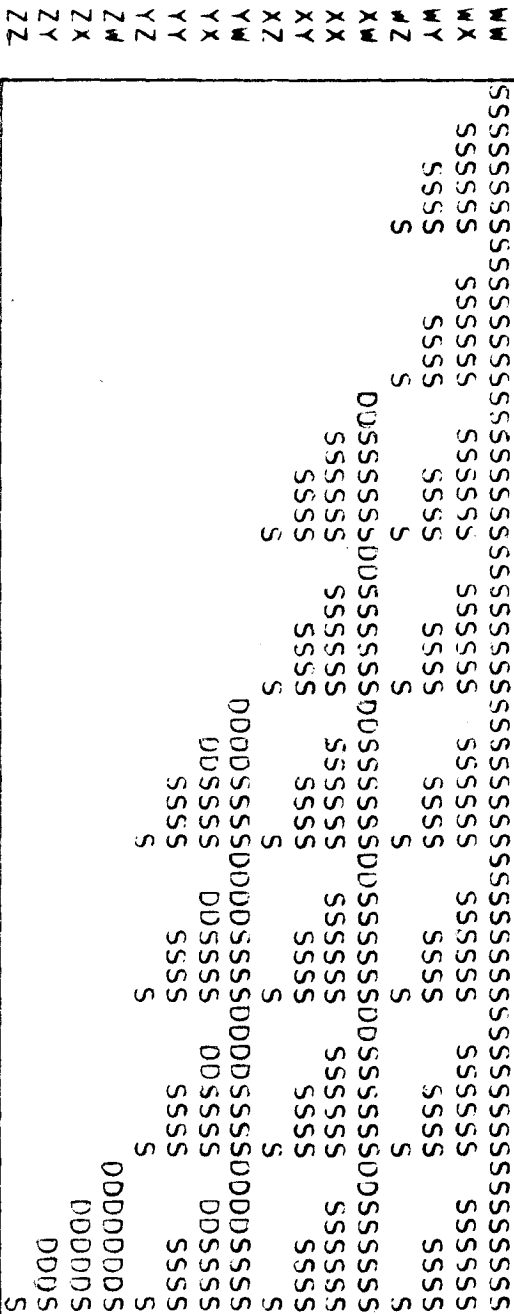
SSSSSSSS
SSSSSS
SSSS
S
DDSSSSSS
SSSSSS
SSSS
S
DDDDSSSS
DDSSSS
SSSS
S
DDDDDDDS
DDDDDS
DDDS
S

WM
WX
WY
WZ
XW
XX
XY
XZ
YW
YX
YY
YZ
ZW
ZX
ZY
ZZ

\mathcal{J}				\mathcal{D}			
w	x	y	z	w	x	y	z
8	6	4	1	0	0	0	0
6	6	4	1	2	0	0	0
4	4	4	1	4	2	0	0
1	1	1	1	7	5	3	0

Figure n° 3.28 : Enumération dans $E^2 \times (\{E\} \times M)$ et dénombrement dans E^2 des relations \mathcal{J} et \mathcal{D} correspondant à l'échelle d'intervalles $M = \{ \{w\}, \{w\}, \{w,x\}, \{w,x\}, \{w,x,y\}, \{w,x,y\}, \{w,x,y,z\}, \{w,x,y,z\} \}$.

AAAAAABRRRBRBRBCCCCCCCCDDDDDDDEEEEEEEFFFFFFGGGGGGHHHHHH
 ABCDEFGHABRCDEFGHABRCDEFGHABRCDEFGHABRCDEFGHABRCDEFGH



\mathcal{Y}	w	x	y	z
w	64	48	32	8

	x	y	z
x	36	36	24
y	16	16	16
z	1	1	1

\mathcal{D}	w	x	y	z
w	0	0	0	0
x	12	0	0	0
y	16	8	0	0
z	7	5	3	0

Figure n° 3.29 : Enumération dans $E^2 \times M_0^2$ et dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} correspondant à l'échelle de rapports $M_0 = \{ \{w\}, \{w\}, \{w,x\}, \{w,x\}, \{w,x,y\}, \{w,x,y\}, \{w,x,y,z\}, \{w,x,y,z\} \}$.

La formalisation que nous avons développée permet donc de rendre compte des six niveaux de mesure retenus au § 3.3.1.c ; cette formalisation sera en quelque sorte validée par l'application que nous en ferons au calcul des mesures de dispersion et de liaison. Auparavant, il faut noter que ce cadre formel s'adapte mal à la classification que nous avons utilisée ; il y a donc là un domaine de recherches pouvant déboucher soit sur un système de formalisation des quanta d'information structurelle qui permette de mieux traiter du problème des échelles de mesure, soit sur une remise en question de la définition de ces échelles et de leur classification. Dans cette seconde perspective, remarquons qu'il est possible de traduire en termes de logons n'importe quelle forme de question tirée d'un questionnaire. Soit par exemple une question de sélection unique dans un ensemble de cinq modalités constituant (selon la terminologie de SAPIR, cf. § 3.1.3.a) deux séries fermées conjointes :

v = "très favorable à V"
 w = "plutôt favorable à V"
 x = "hésite entre les deux"
 y = "plutôt favorable à Z"
 z = "très favorable à Z"

La précodification prévoit en outre deux réponses métalinguistiques :

t = "sans opinion"
 u = "refuse de se prononcer"

Dans notre système, cette question peut s'écrire :

$$P = \{ \{v,w,x\} , \{w,x\} , \{x\} , \{x,y\} , \{x,y,z\} , \{t\} , \{u\} \}$$

Les figures n° 3.30 et 3.31 présentent l'énumération dans $E^2 \times P^2$ et le dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} correspondant à cette question, en faisant l'hypothèse que le chercheur considère la zone de recouvrement des deux séries conjointes comme un point de référence fixe

\mathcal{P}	T	U	V	W	X	Y	Z
T	1	0	0	0	0	0	0
U	0	1	0	0	0	0	0
V	0	0	1	1	1	0	0
W	0	0	2	4	4	0	0
X	0	0	5	10	25	10	5
Y	0	0	0	0	4	4	2
Z	0	0	0	0	1	1	1

\mathcal{D}	T	U	V	W	X	Y	Z
T	0	1	1	2	5	2	1
U	1	0	1	2	5	2	1
V	1	1	0	1	4	2	1
W	2	2	0	0	6	4	2
X	5	5	0	0	0	0	0
Y	2	2	2	4	6	0	0
Z	1	1	1	2	4	1	0

Figure n° 3.31 : Dénombrement dans E^2 des relations \mathcal{P} et \mathcal{D} correspondant à l'échelle partiellement ordonnée avec origine fixe $C = \{w\}$: $P = \{ \{v,w,x\}, \{w,x\}, \{x\}, \{x,y\}, \{x,y,z\}, \{t\}, \{u\} \}$.

$C = \{x\}$. Cette représentation est différente de celle que l'on adopterait en admettant l'hypothèse (plus forte) que ces deux séries conjointes de gradation sont l'expression linguistique d'une seule dimension sous-jacente : les cinq modalités de réponse linguistiques constitueraient alors une échelle totalement ordonnée.

Nous voyons sur ce dernier exemple que la formalisation proposée permet au chercheur non seulement de rendre compte de questions dont l'ensemble des réponses possibles ne s'intègre pas dans les "échelles de mesure" usuelles, mais également de modifier *ad libitum* la structure de la variable construite à partir de l'ensemble des réponses à une question. Pour prendre un exemple simple, à partir de la liste des catégories socio-professionnelles proposées comme réponses possibles à une question de sélection unique du type : "Quelle était la profession de votre père quand vous avez cessé de fréquenter régulièrement l'école ?" il est aisé de construire diverses "échelles" intermédiaires entre l'échelle nominale et l'échelle ordinale. On peut ainsi définir soit un ordre partiel sur les indicateurs, soit des distances entre certains indicateurs seulement, même sans relation d'ordre ; de cette manière, il est possible de tenir compte de théories complexes concernant la hiérarchie sociale, ou les distances sociales entre CSP.

Cette formalisation pourrait donc ouvrir un domaine de recherches encore peu exploré. Comme ces préoccupations, malgré l'intérêt qu'elles présentent, sortent du cadre de cet exposé, nous n'avons pas cru nécessaire de les développer plus complètement. Par contre, afin de faciliter le contrôle de nos assertions à ce sujet, nous présentons en annexe II un programme *FORTTRAN* effectuant automatiquement le dénombrement des relations \mathcal{P} et \mathcal{D} pour toute forme de variable, quel que soit son niveau de complexité. Ce programme permet ainsi de mesurer l'information structurelle contenue dans une variable quelconque.

3.3.2. Information métrique et indices statistiques

L'information métrique est apportée par le questionnement $\pi : U \rightarrow R^*$, et résumée habituellement par le résultat de l'application $d : R \rightarrow \mathbb{N}_0$, telle qu'à chaque réponse canonique $r^* \in R^*$ corresponde le nombre d'individus élément de U qui ont donné cette réponse. L'information métrique élémentaire, exprimée en termes de *metrons*, est apportée par l'application $\pi' : U \rightarrow S^*$, en appelant comme précédemment S^* la structure élémentaire de R^* en termes de *logons*. On peut la résumer par le résultat de l'application $d' : S^* \rightarrow \mathbb{N}_0$, telle qu'à chaque *logon* élément de S^* corresponde le nombre de couples d'individus ayant, par leur réponse, affirmé ce logon. L'information métrique est donc une *quantité*, qui, pour chaque logon, est comprise entre zéro et $|U|^2$.

Cette quantité est, en soi, un résumé, puisque le dénombrement des individus ayant affirmé un logon donné ne tient pas compte de l'ensemble des autres traits qui caractérisent chacun de ces individus. L'analyse de l'information métrique apportée par le questionnement consiste à étudier les distributions de metrons correspondant à plusieurs logons considérés simultanément. Il faut pour cela tenir compte des relations définies par le chercheur entre les indicateurs que constituent les réponses élémentaires aux questions, et les variables qu'il a construites. Ces relations, pour une variable donnée, se déduisent de la structure ("échelle de mesure") définie sur chaque variable ; on peut par conséquent identifier facilement les logons correspondant à cette structure, pour une variable donnée. Après questionnement et dénombrement des metrons obtenus par le questionnement, on dispose, pour chaque variable, d'une distribution des quantités d'information métrique correspondant à chaque logon (distribution dite "univariée"). Il est commode de résumer l'information apportée par cette distribution à l'aide d'une quantité exprimant la *variété* de cette distribution (quantité de *dispersion*). Si l'on s'intéresse aux distributions métriques correspondant à deux variables considérées simultanément, on souhaitera distinguer les quantités exprimant la variété de chacune des variables, et la quantité exprimant la variété liée, ou

conjointe, de la distribution "bivariée" (quantité de *liaison* entre variables). Enfin, si l'on s'intéresse à plus de deux variables considérées simultanément (distribution dite "multivariée"), il est semble-t-il toujours possible de décomposer l'information métrique en ses composantes univariées et bivariées (voire en composantes plus complexes ; cf. p.ex. l'analyse multidimensionnelle de l'incertitude, in [Garner 1962 : 98-112], et [Grémy 1966 : 461-465]).

En vue de faciliter les comparaisons ultérieures, il est souvent commode de présenter les quantités ainsi définies sous la forme d'indices numériques, dont les bornes de variation sont fixées de façon uniforme. En règle générale, le calcul d'un indice a pour résultat la transformation d'une quantité en un nombre réel compris, selon les cas, dans l'intervalle $[0, 1]$, ou dans l'intervalle $[-1, +1]$. Parmi les conditions auxquelles doit répondre un bon indice, telles qu'elles ont été définies par G.U. YULE et M.G. KENDALL (cf. [Morice 1954 : 9]), deux nous paraissent particulièrement importantes pour notre propos ; elles posent qu'un bon indice doit :

- dépendre de toutes les observations de la distribution considérée ;

- se prêter aux calculs algébriques ultérieurs.

Dans les paragraphes suivants, nous ne considérons que des indices répondant à ces deux exigences : cette règle nous a conduit, en particulier, à négliger certains indices de dispersion utilisables au niveau ordinal sans origine fixe (p.ex. : écart interquartile), malgré l'usage fréquent qui est fait de ces indices.

a) Mesures de dispersion

Appelons $s_{x y}$ le dénombrement dans E^2 des metrons de la forme $\mathcal{Y}(A, B, x, y)$ pour tous les couples (A, B) de parties de E définissant l'échelle de mesure utilisée, et $d_{x y}$ le dénombrement dans E^2 des metrons de la forme $\mathcal{D}(A, B, x, y)$ pour tous les couples (A, B) de l'échelle. Ce dénombrement se fonde évidemment sur la comparaison des réponses de deux individus. Par exemple, si nous considérons la question prise comme illustration d'une variable complexe à la fin du § 3.3.1.d, supposons qu'une des personnes interrogées soit très favorable à V (réponse v), alors qu'une autre hésite entre V et Z (réponse x). En nous reportant à la figure n° 3.3.1, nous constatons que, dans l'hypothèse où cette question a x pour origine, il y a cinq relations \mathcal{Y} de x à v et une de v à x , et quatre relations \mathcal{D} de v à x . L'information métrique apportée par ces deux réponses est donc égale à :

$$\left\{ \begin{array}{l} d_{x v} = 0 \\ d_{v x} = 4 \\ s_{x v} = 5 \\ s_{v x} = 1 \end{array} \right.$$

Ce raisonnement doit être répété pour tout couple de répondants. Par conséquent, en pratique, le nombre de metrons apportés par l'ensemble des réponses v et x à cette question est égal au nombre d'unités élémentaires (logons) recensées entre ces réponses dans le tableau n° 3.3.1, multiplié par le produit des effectifs des individus ayant répondu v et de ceux ayant répondu x . Si par exemple 141 personnes ont répondu v , et 92 x , l'information métrique apportée par leurs réponses est égale à :

$$\left\{ \begin{array}{l} d_{x v} = 0 \times 141 \times 92 = 0 \\ d_{v x} = 4 \times 141 \times 92 = 51\ 888 \\ s_{x v} = 5 \times 141 \times 92 = 64\ 860 \\ s_{v x} = 1 \times 141 \times 92 = 12\ 972 \end{array} \right.$$

Pour calculer l'information métrique correspondant par exemple aux relations \mathcal{D} , apportée par une distribution de réponses $F = \{n_v, n_w, \dots, n_t, n_u\}$, il suffit de multiplier case par case le dénombrement de ces relations \mathcal{D} structurelles dans E^2 par $F^2 = \{n_v^2, n_v n_w, n_v n_x, \dots, n_u n_t, n_u^2\}$.

Nous dirons que toute mesure de la *quantité de dispersion* de la distribution a nécessairement l'une des quatre formes suivantes :

$$\left\{ \begin{array}{l} \textcircled{a} = \varphi(f(\pi_{x y} d_{x y})) \\ \textcircled{b} = \psi(f(\pi_{x y} s_{x y})) \\ \textcircled{c} = \psi(g(\pi_{x y} d_{x y})) \\ \textcircled{d} = \varphi(g(\pi_{x y} s_{x y})) \end{array} \right.$$

où $\pi_{x y}$ est un coefficient de pondération attaché au couple d'indicateurs (x, y) , φ et f des fonctions monotones croissantes, ψ et g des fonctions monotones décroissantes. En effet, les statistiques de dispersion mesurent l'hétérogénéité de la distribution ; or, plus une distribution est hétérogène, plus le nombre de relations \mathcal{D} est élevé, et, corrélativement, plus le nombre de relations \mathcal{Y} est faible. Il est donc assez naturel de choisir comme mesure de la dispersion une statistique qui soit une fonction croissante du nombre de relations \mathcal{D} , ou une fonction décroissante du nombre de relations \mathcal{Y} , qui constituent l'information métrique induite par le questionnement.

Au niveau nominal sans origine, la mesure de dispersion la plus usitée paraît être l'entropie :

$$H = \sum_i p_i \log_2 \frac{1}{p_i} = \log_2 N - \frac{1}{N} \sum_i n_i \log_2 n_i$$

avec, dans le contexte de l'administration d'un questionnaire :

$$N = |U| = \text{nombre total d'individus ayant répondu ;}$$

$$n_i = \text{nombre d'individus éléments de } U \text{ ayant donné la réponse } i ;$$

$$p_i = \frac{n_i}{N} = \text{fréquence relative de la réponse } i .$$

C'est une quantité variant de zéro (dispersion nulle) au logarithme du nombre K d'éventualités de la variable (nombre de réponses possibles au sens large) : $\log_2 K$ (dispersion maximum correspondant à l'équirépartition). Dans notre système de formalisation, c'est une mesure de dispersion de forme $\textcircled{b} = \psi(f(\pi_{x x} s_{x x}))$, avec :

$$\pi_{x x} = \frac{\log_2 s_{x x}}{2 N \sqrt{s_{x x}}} = \frac{\log_2 s_{x x}}{2 N \cdot n_x}$$

$$f(\xi) = \sum_x \xi_x$$

$$\psi(\eta) = \log_2 N - \eta$$

Au niveau nominal sans origine, la relation \mathcal{Y} n'est définie que sur les couples identiques (x, x) , comme nous l'avons vu au § 3.3.1.d. Il est possible de tirer de la quantité de dispersion ainsi mesurée un indice variant de zéro à un, en posant :

$$h = \frac{H}{\log_2 K} \quad (\text{cf. [Attneave 1959 : 9]}).$$

En permutant les fonctions φ et ψ dans les formules des mesures de dispersion définies au début de ce paragraphe, on obtient des mesures de l'homogénéité de la distribution. Par exemple, en remplaçant, dans la formule de l'entropie ci-dessus, la fonction ψ par une fonction $\varphi(\xi) = \xi$, on aboutit à une mesure d'homogénéité variant de $\log_2 N - \log_2 K$ à $\log_2 N$. On constate ainsi que : $H =$ homogénéité maximum possible moins homogénéité observée.

Le khi deux comparant une distribution observée à l'équirépartition est une mesure de l'homogénéité de la distribution, de forme

$$\textcircled{b^{-1}} = \varphi(f(\pi_{x x} \quad s_{x x})), \text{ avec :}$$

$$\pi_{x x} = 1$$

$$f(\xi) = \sum_i \xi_i$$

$$\varphi(\eta) = \frac{K}{N} \eta - N$$

Soit la formule :

$$\chi^2 = \sum_i \frac{(n_i - \frac{N}{K})^2}{\frac{N}{K}} = \frac{K}{N} (\sum_i s_i i) - N$$

Cette mesure d'homogénéité varie de zéro (équirépartition) à $N(K - 1)$. On pourrait en tirer une mesure de la quantité de dispersion (variété) :

$$V = K(N - \frac{1}{N} \sum_i s_i i)$$

ou encore un indice de dispersion variant dans l'intervalle $[0, 1]$:

$$v = \frac{K(N - \frac{1}{N} \sum_i s_i i)}{N(K-1)}$$

Au niveau ordinal, nous n'avons pas trouvé, dans les ouvrages courants de statistique appliquée aux sciences de l'homme, de mesure de la dispersion qui satisfasse aux deux conditions de YULE que nous avons rappelées. Il serait facile d'appliquer à ce niveau de mesure les formules générales que nous avons proposées au début de ce paragraphe. Il conviendrait toutefois, au préalable, de trancher le problème du "*rang vide*". En effet, historiquement, il semble que les premiers travaux sur les échelles ordinales se soient inspirés de procédures d'ordination stricte totale de tous les individus ou objets ; dans le contexte des tests et des questionnaires, cela reviendrait à ordonner strictement l'ensemble des répondants. Il y aurait donc bijection entre les répondants et les modalités présentées (rangs), ces dernières devant être en nombre égal à celui des individus interrogés. Cette manière de voir inspire d'ailleurs largement la présentation des coefficients de corrélation par rangs qui est faite dans les manuels : la présence d'ex aequos y est en effet considérée comme une anomalie, qu'il est nécessaire de corriger, et non comme un résultat naturel. Remarquons d'ailleurs que, dans le cas d'une application bijective des répondants sur des modalités ordonnées strictement, il n'y a pas lieu d'envisager une mesure de la dispersion, puisque celle-ci est toujours maximum par définition.

Dans un questionnaire, si l'on demande par exemple aux répondants de choisir l'une des dix modalités ordonnées strictement, on s'expose évidemment à trouver non seulement des réponses ex aequos, mais également à l'éventualité que l'une des modalités présentées ne soit sélectionnée par aucun répondant. Si l'on considère la question comme le moyen d'ordonner (au sens large) l'ensemble des répondants, le problème se pose évidemment de savoir si l'on doit conserver l'ensemble des modalités (en acceptant un "*rang vide*"), ou si l'on doit ne considérer que les modalités sélectionnées au moins une fois (ou au moins k fois, selon un seuil arbitraire fixé par le chercheur). L'information métrique obtenue, et servant de base à la mesure de la dispersion, n'est pas la même dans les deux cas (pour une même distribution des réponses). L'information structurelle, qui sert de base au dénombrement des metrons, compte 165 relations \mathcal{D} pour dix modalités

ordonnées, contre 120 pour la même échelle réduite à neuf modalités. Sur les 45 logons supprimés par la réduction de l'ensemble des modalités, certains correspondent à des couples de modalités conservées après réduction ; par conséquent, la conservation ou l'élimination du "rang vide" a une incidence sur l'information métrique, et donc sur la mesure de la dispersion.

Au niveau métrique sans origine fixe, la mesure de dispersion usuelle est la variance (et sa racine carrée, l'écart-type de la distribution). Cette statistique est de la forme $\textcircled{a} = \varphi(f(\pi_{x y} d_{x y}))$, avec :

$$\pi_{x y} = \lambda_{x y} = \text{nombre de logons dans } E^2 \text{ pour le couple } (x, y)$$

$$f(\xi) = \sum_i \xi_i$$

$$\varphi(\eta) = \frac{1}{N^2} \eta$$

soit la formule de la variance :

$$\sigma^2 = \frac{1}{N^2} \sum_x \sum_y \lambda_{x y} d_{x y}$$

Signalons toutefois que cette formule simple ne vaut que pour l'échelle d'intervalles M telle que nous l'avons définie au § 3.3.1.d. En particulier, un déplacement de l'origine arbitraire de l'échelle conduirait à des résultats différents, et imposerait une complexification de la fonction φ .

Ce bref survol des principales statistiques de dispersion suffit à montrer que l'unité élémentaire d'information apportée par le questionnement permet bien de rendre compte des mesures statistiques usuelles qui

satisfont aux conditions de YULE. En outre, il souligne à la fois la pauvreté de l'hétérogénéité de l'arsenal statistique courant dans ce domaine. Il serait intéressant, en considérant que l'unité de mesure structurelle (logon) est la relation élémentaire de type \mathcal{D} , d'étudier systématiquement les propriétés des mesures de dispersion obtenues en remplaçant, dans les formules (a) et (c), les fonctions φ , ψ , f , et g , par des fonctions simples usuelles, et en adoptant par exemple pour $\pi_{x y}$ soit la valeur $\pi_{x y} = 1$, soit $\pi_{x y} = \lambda_{x y}$, où $\lambda_{x y}$ est le nombre de logons correspondant, dans l'échelle considérée, à chaque couple d'indicateurs (x, y) . En s'inspirant de la formule de la variance, on obtient ainsi par exemple au niveau nominal une mesure de dispersion variant de zéro à $\frac{K-1}{K}$ (pour : $\pi_{x y} = \lambda_{x y} = 1$).

b) Mesures de liaison entre variables

Ainsi que nous l'avons noté, les mesures de liaison entre deux variables ont pour fondement une décomposition de la dispersion analogue à l'analyse de la variance ou de l'incertitude. En principe donc, nous pourrions considérer que notre objectif, qui était d'établir un lien entre les quanta d'information et les techniques statistiques usuelles, a été atteint au paragraphe précédent. Nous voudrions toutefois présenter quelques résultats supplémentaires, concernant les couples de variables. Malheureusement, cette recherche, qui constitue un prolongement naturel de notre recherche sur le questionnaire, n'en est encore qu'à ses débuts. C'est pourquoi les résultats que nous présentons, et en particulier ceux qui concernent le χ^2 et le r de BRAVAIS-PEARSON, ne sont que des résultats parcellaires.

Parmi les mesures de liaison entre deux variables, l'information transmise T , tirée de la théorie de l'information, ne pose aucun problème nouveau : il suffit de calculer l'entropie H_1 et H_2 pour chacune des distributions marginales, l'entropie $H_{1,2}$ pour la distribution correspondant au croisement des deux variables, et à appliquer la formule [Attneave 1959 : 48] :

$$T_{1,2} = H_1 + H_2 - H_{1,2}$$

Pour l'ensemble des autres mesures, il convient de définir préalablement les incidences structurelles du croisement de deux variables quelconques. Ainsi que nous le verrons, dans tous les cas que nous avons étudiés, seules les relations \mathcal{D} sont utilisées pour le calcul de la quantité de liaison. C'est pourquoi nous ne considérerons que les dénombrements de ces relations dans les ensembles d'indicateurs.

Soient deux variables quelconques, Y_i et Y_j , et soient E_i et E_j les ensembles d'indicateurs sur lesquels sont définies ces variables. Posons : $(a,b) \in E_i^2$ et $(x,y) \in E_j^2$, et appelons comme précédemment $\lambda_{a b}$ (respectivement : $\lambda_{x y}$) le dénombrement dans E_i^2 (respectivement : E_j^2) des relations \mathcal{D} structurelles (*logons*) de a à b (respectivement : de x à y). Nous définirons le résultat du croisement de Y_i avec Y_j comme une variable bidimensionnelle $Y_{i,j}$; nous appellerons *ax* l'élément (a,x) de $E_{i,j} = E_i \times E_j$, et $\lambda_{ax by}$ le dénombrement dans $E_{i,j}^2$ des relations \mathcal{D} structurelles de *ax* à *by*. Nous dirons que la structure produite par le croisement de Y_i et de Y_j se traduit dans $E_{i,j}^2$ par le résultat de l'application de la règle suivante :

$$\lambda_{ax by} = \lambda_{a b} + \lambda_{x y}$$

Pour illustrer cette règle sur un cas particulièrement simple, considérons les deux variables nominales dichotomiques :

$$N_1 = \{ \{a\}, \{b\} \}$$

$$N_2 = \{ \{x\}, \{y\} \}$$

Le résultat de leur croisement dans le produit cartésien des ensembles d'indicateurs correspondants est représenté par la figure n° 3.32. On

	<i>a</i>	<i>b</i>	
<i>a</i>	0	1	
<i>b</i>	1	0	

$(E_1)^2$

	<i>x</i>	<i>y</i>
<i>x</i>	0	1
<i>y</i>	1	0

$(E_2)^2$

	<i>ax</i>	<i>ay</i>	<i>bx</i>	<i>by</i>
<i>ax</i>	0	1	1	2
<i>ay</i>	1	0	2	1
<i>bx</i>	1	2	0	1
<i>by</i>	2	1	1	0

$(E_1 \times E_2)^2$

Figure n° 3.32 : Exemple du dénombrement des relations \mathcal{D} structurelles résultant du croisement de deux variables nominales dichotomiques.

voit sur cet exemple que certains couples éléments de $(E_1 \times E_2)^2$ ont entre eux deux relations structurelles. Ces couples sont de la forme (ax, by) ou (ay, bx) . Nous verrons qu'ils jouent un rôle important dans toutes les mesures de liaison entre deux variables. Leur caractéristique est d'être composés de deux couples élémentaires (a, x) et (b, y) , tous deux éléments de $E_1 \times E_2$, et tels que $a \neq b$ & $x \neq y$. En outre, si l'on considère les quatre couples possibles dans $(E_1 \times E_2)^2$ à partir de deux couples d'indicateurs élémentaires, on relève entre eux deux types d'opposition (figure n° 3.33) :

Ⓐ opposition par permutation des éléments de $(E_{1,2})^2 = (E_1 \times E_2)^2$, correspondant à des cases symétriques par rapport à la diagonale principale ($\lambda_i \quad i = 0$).

Ⓑ opposition par permutation des éléments appartenant à l'un des couples élémentaires initiaux soit dans E_1 , soit dans E_2 .

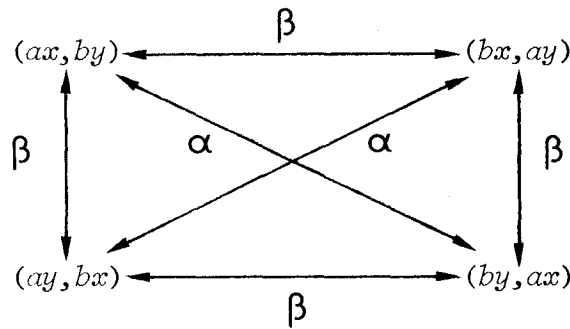


Figure n° 3.33 : Relations d'opposition entre éléments de $(E_i \times E_j)^2$ obtenus à partir d'un couple d'éléments de E_i distincts et d'un couple d'éléments de E_j distincts.

Une mesure de quantité d'association entre deux variables dichotomiques se fonde sur l'opposition de type Ⓑ. Appelons comme précédemment $d_{ax by}$ l'information métrique (apportée par le questionnement) relative au couple (ax, by) : $d_{ax by}$ est donc égal au produit des $\lambda_{ax by}$ par

le produit des effectifs n_{ax} et n_{by} des individus ayant répondu respectivement a & x , et b & y :

$$d_{ax by} = \lambda_{ax by} \times n_{ax} \times n_{by}$$

La quantité d'association entre les variables est ici égale à :

$$L_N = d_{ax by} - d_{ay bx}$$

On peut en tirer par exemple le coefficient d'association de YULE :

$$Q = \frac{d_{ax by} - d_{ay bx}}{d_{ax by} + d_{ay bx}}$$

ou encore le coefficient de contingence de PEARSON :

$$\Phi^2 = \frac{N (d_{ax by} - d_{ay bx})^2}{d_a b \times d_x y}$$

d'où, dans le cas particulier de variables dichotomiques, la valeur du χ^2 mesurant l'écart à l'hypothèse d'indépendance :

$$\chi^2 = N \Phi^2$$

La formule générale du coefficient de PEARSON (et donc du χ^2 correspondant) est plus complexe, et reste à déterminer. Outre ce problème important, les recherches ultérieures pourraient également prendre pour objet l'examen, dans le cadre de cette formalisation, de l'indice d'association proposé par LEIK et GOVE (cf. [Leik 1971 : 284-290]). La démarche de ces auteurs présente en effet certaines analogies avec celle que nous avons esquissée ici .

Au niveau ordinal, la règle de dénombrement des relations \mathcal{D} dans $(E_i \times E_j)^2$ s'applique de la même manière. La figure n° 3.34 illustre sa mise en oeuvre pour deux variables ordinales sans origine fixe :

$$O_1 = \{ \{a\} , \{a,b\} , \{a,b,c\} \}$$

$$O_2 = \{ \{x\} , \{x,y\} , \{x,y,z\} \}$$

Si, comme précédemment, nous ne considérons dans E_{ij}^2 que les couples issus d'un quadruplet de $(E_i \times E_j)^2$ tel que tous ses éléments soient distincts, nous pouvons classer l'ensemble de ces couples en deux sous-ensembles disjoints :

- le sous-ensemble $C \subset E_{ij}^2$, composés des couples pour lequel il y a convergence des relations \mathcal{D} initiales dans E_i^2 et dans E_j^2 .

Comme exemple d'élément de C , on peut citer le couple (by, ax) , auquel correspondent dans E_i^2 , le couple (b, a) , et dans E_j^2 , le couple (y, x) .

On voit que pour chacun des deux couples initiaux ("univariés"), il existe au moins une relation \mathcal{D} .

- le sous-ensemble $D \subset E_{ij}^2$, composé des couples pour lequel il y a divergence des relations \mathcal{D} initiales. Comme exemple d'élément de D , citons le couple (bx, ay) , pour lequel il existe une relation $\mathcal{D}_{b \ a}$ dans E_i^2 , mais pas de relation \mathcal{D} entre x et y dans E_j^2 .

Lorsqu'elles sont définies, les relations d'oppositions entre couples appartenant à un même sous-ensemble (C ou D) sont du type $\textcircled{\alpha}$; les relations d'opposition entre un élément de C et un élément de D sont du type $\textcircled{\beta}$.

Cette description se rapproche de l'analyse de l'association entre variables ordinales présentée par Theodore R. ANDERSON et Morris ZELDITCH Jr. [Anderson 1958 : 142-155]. Ces auteurs en effet dénombrent

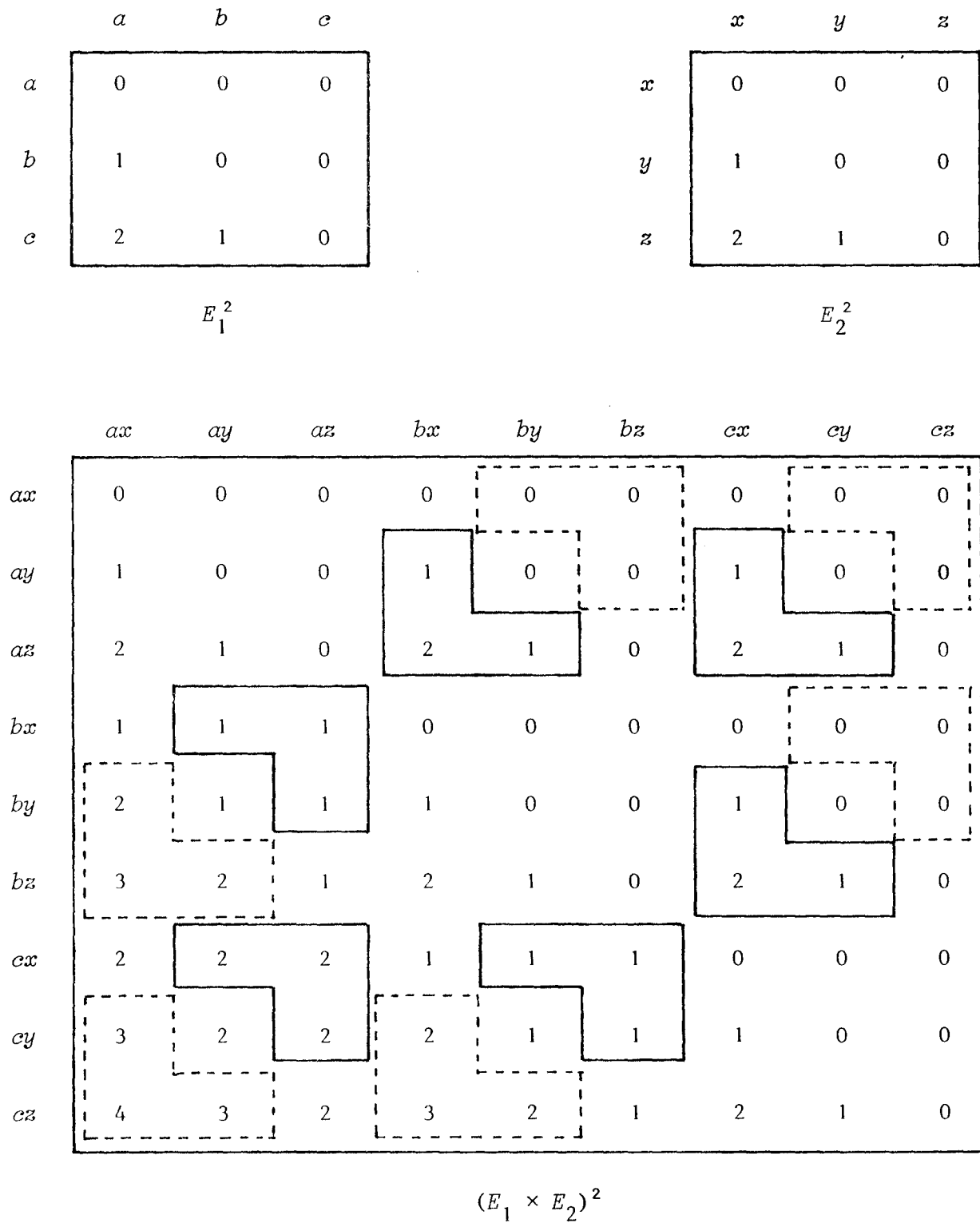


Figure n° 3.34 : Exemple de dénombrement des relations \mathcal{D} structurales résultant du croisement de deux variables ordinales.

les couples (d'indicateurs ou d'individus) pour lesquels les deux variables donnent le même ordre, et ceux pour lesquels les deux variables donnent un ordre différent. Or, la procédure de dénombrement des relations \mathcal{D} dans E_{ij}^2 , appliquée aux échelles ordinales, aboutit à classer dans C (même ordre sur les deux variables) tous les couples (ax, by) tels que $a > b$ & $x > y$ ou $a < b$ & $x < y$, et dans D (ordre différent sur les deux variables) tous les couples (cw, dz) tels que $c < d$ & $w > z$ ou $c > d$ & $w < z$. ANDERSON et ZELDITCH montrent ensuite que de nombreux coefficients d'association entre variables ordinales se fondent sur la différence entre ces deux dénombrements. En s'inspirant de leur démonstration, nous dirons qu'une mesure possible de la quantité de liaison entre deux variables ordinales est de la forme :

$$L_0 = A - B$$

$$\text{avec : } A = \sum_{(ax, by) \in C} \frac{d_{ax by}}{\lambda_{ax by}}$$

$$B = \frac{1}{2} \sum_{(cw, dz) \in D} \frac{d_{cw dz}}{\lambda_{cw dz}}$$

Cette quantité de liaison sert de base pour le calcul de divers indices d'association entre variables ordinales, tels que les coefficients d de SOMER, γ de GOODMAN et KRUSKAL, et τ de KENDALL, la différence entre ces coefficients réside dans la quantité par laquelle est divisée L_0 (cf. p.ex. : [Anderson 1958 : 152], [Kohout 1974 : 224-232]). A titre d'exemples, le γ de GOODMAN et KRUSKAL peut s'écrire :

$$\gamma = \frac{A - B}{A + B}$$

et la formule générale du τ de KENDALL (avec possibilité d'ex-aequos) :

$$\tau = \frac{L_0}{\frac{N(N-1)}{2}}$$

(avec N = nombre d'individus ayant répondu aux deux questions).

Le ρ de SPEARMAN sans ex-aequos obéit à une logique voisine de celle du coefficient γ (quoique rarement mise en parallèle avec celle-ci). Il se fonde sur la quantité de liaison :

$$L'_O = A' - B'$$

$$\text{avec : } A' = \sum_{(ax, by) \in C} d_{ax by}$$

$$B' = \sum_{(cw, dz) \in D} d_{cw dz}$$

Nous avons ainsi :

$$\rho = \frac{A' - B'}{A' + B'}$$

Le ρ avec ex-aequos peut, d'autre part, être considéré comme un coefficient de liaison entre variables métriques (échelles d'intervalles), puisque chaque rang calculé pour les ex-aequos est un rang moyen, pouvant n'être pas un nombre entier ; il n'y a plus alors l'équidistance entre indicateurs qui caractérise l'échelle ordinale.

Au niveau métrique, le coefficient de liaison le plus utilisé reste le r de BRAVAIS-PEARSON, en dépit du rôle symétrique qu'il impose aux deux variables, et de l'hypothèse de linéarité de la liaison. Ainsi que nous l'avons dit au début de ce paragraphe, nous ne sommes pas parvenu, pour l'instant, à définir la formule générale de ce coefficient, malgré le lien évident qui existe entre notre formalisation et la méthode des moindres carrés. Signalons simplement que la formule proposée ci-dessus pour le calcul de la quantité de liaison L'_O peut, dans le cas d'un tableau 2×2 , s'appliquer aussi bien au calcul du χ^2 qu'à celui

du coefficient de corrélation r (les dénombrements des métrons étant évidemment différents, à distribution égale, selon la structure des variables). Ce résultat partiel suggère d'ailleurs une parenté possible entre les deux indices.

c) Perspectives de recherches.

Pour conclure cette présentation des premiers résultats d'une recherche qui, quoi qu'extérieure à notre propos, en constitue la suite obligée, signalons cinq directions de recherches possibles :

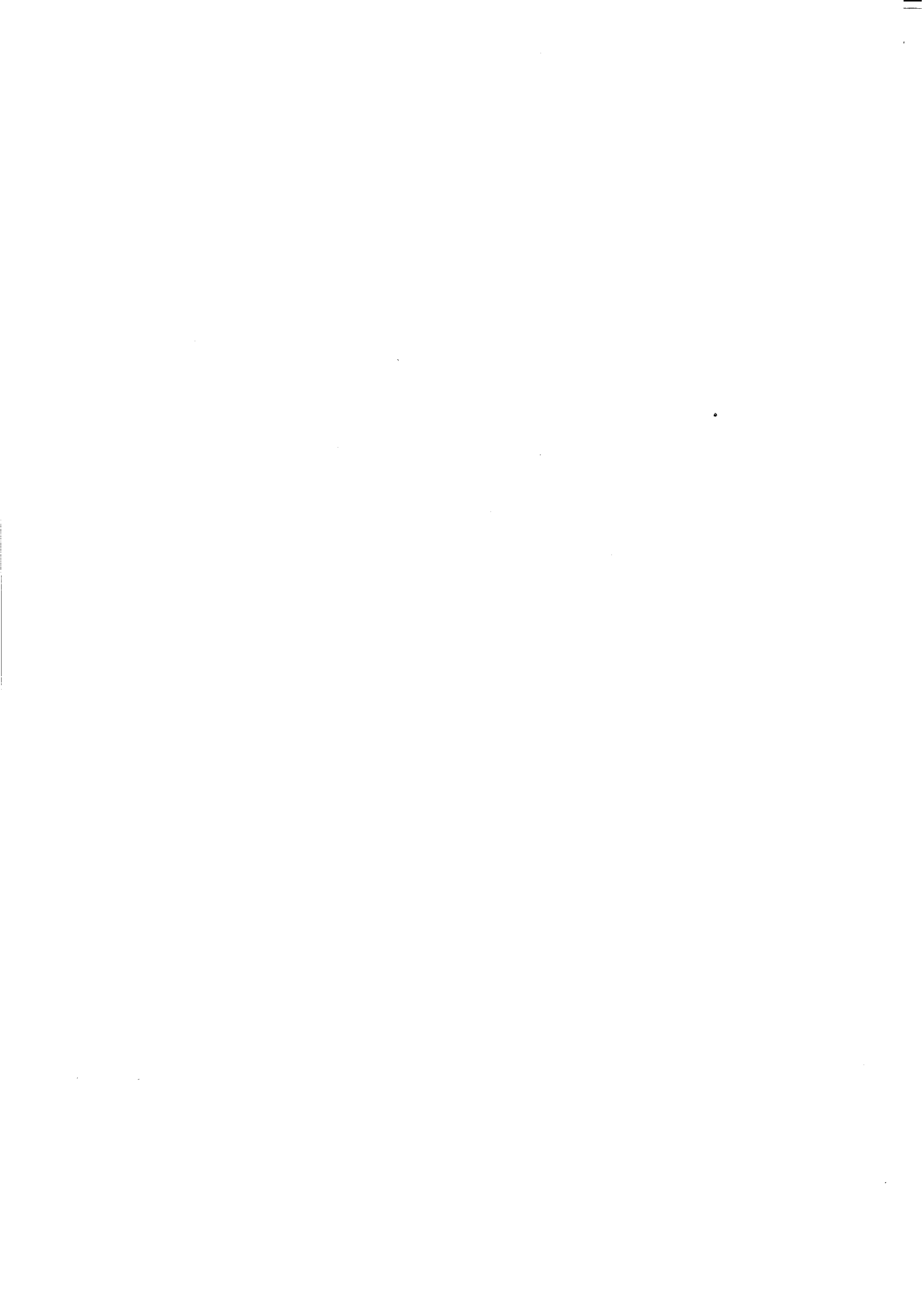
① Poursuite des travaux actuels sur les coefficients statistiques usuels de dispersion et de liaison, en vue d'élaborer une théorie susceptible de rendre compte de la plupart d'entre eux, sans restriction à des cas particuliers. Peut-être ces travaux déboucheraient-ils sur une simplification de l'appareil théorique (logons définis dans $E^2 \times \mathcal{P}(E)$ seulement), et une complexification du dénombrement des métrons pour les distributions à deux variables et plus.

② Sur le modèle des coefficients dont la théorie rend compte actuellement, élaboration de nouveaux indices de liaison et de dispersion, adaptés à divers niveaux de mesure, et intégrés dans une formalisation commune (cf. la tentative faite en ce sens dans [Leik 1971]). A titre d'exemples, citons l'absence de mesures de dispersion satisfaisant aux conditions de YULF au niveau ordinal, ou de mesures de liaison pour certains couples de variables de niveaux différents, comme le couple ordinal-nominal.

③ Extension de la démarche suivie aux mesures de distance ou de ressemblance entre individus. Il serait intéressant, par exemple, de traduire en termes de quanta d'information les divers indices de similarité utilisés en taxinomie numérique (cf. [Lerman 1970 : 18]), et de procéder à une étude systématique de la formalisation ainsi obtenue.

④ Redéfinition du système des "échelles de mesure" à partir de la formalisation proposée pour l'analyse de l'information structurelle.

⑤ Mise au point de mesures de dispersion et de liaison pouvant s'appliquer à n'importe quelle structure de variables, et permettant de comparer entre elles des distributions correspondant à des variables de structures différentes.



C O N C L U S I O N S

A l'issue de cet exposé, il est clair que les précautions que nous avons prises lorsque nous en avons arrêté le titre étaient justifiées : il s'agit bien d'une recherche théorique (et non d'un travail orienté vers la construction des questionnaires), et les résultats présentés ne sont guère que des jalons en vue de l'élaboration d'une théorie du questionnaire. C'est pourquoi l'essentiel de ces conclusions consistera moins en un bilan de la recherche entreprise qu'en une description des suites possibles à lui donner. Nous voudrions toutefois auparavant expliquer certaines des insuffisances ou des lacunes de cet exposé.

Au cours de notre rédaction, nous avons souvent été tenté d'abandonner le rôle du chercheur pour celui de l'enseignant, et d'adopter l'optique de l'auteur de manuel plutôt que celle du rédacteur d'un compte-rendu de recherches en cours. Notre texte y aurait vraisemblablement gagné en clarté, grâce à la description d'exemples plus nombreux et plus complets, à la multiplication des tableaux synoptiques ou de synthèse, à l'abandon de la terminologie propre à chaque auteur au profit d'un vocabulaire uniformisé, etc. Il y aurait certainement perdu en concision, car de tels développements auraient probablement nécessité un volume double ou triple du volume actuel. Le souci d'être bref nous a également amené à ne mentionner que sommairement certains résultats pouvant être entièrement déduits des thèmes présentés, et à préférer une formule ou un tableau à un long développement en langage naturel. En procédant ainsi, nous nous sommes exposé aux reproches que nous avons nous même adressés à KUBIŃSKI à propos de sa trop grande concision. Nous dirons à notre décharge que l'ouvrage de KUBIŃSKI se présente comme une introduction à une théorie déjà très élaborée, alors que notre texte n'est que le compte-rendu de l'état

actuel d'une recherche qui devrait en principe se poursuivre. Il ne nous a pas paru souhaitable de donner à des résultats provisoires, sujets par conséquent à révision dans la suite de notre recherche, la forme qui aurait convenu à des acquis plus assurés. Enfin, nous aurions sans doute pu faire l'économie de certains développements, en passant sous silence des travaux théoriques ou des résultats expérimentaux pouvant apparaître comme moins féconds ou de moindre portée. Par exemple, dans la perspective d'un enseignement sur la théorie du questionnaire, on pourrait limiter l'approche linguistique aux seules conceptions de HARRIS, qui nous ont paru être à la fois les plus rigoureuses et les plus compatibles avec les recherches des logiciens sur les questions. C'eût été renoncer à parler d'une oeuvre profondément originale comme celle de TESNIÈRE, ou d'une orientation de recherches extrêmement répandue, comme celle issue des travaux de CHOMSKY. Un tel choix, en linguistique comme d'ailleurs en logique épistémologique, nous a paru prématuré. Nous n'avons enfreint systématiquement notre règle de recherche de la concision que pour la présentation, à l'appui de nos développements, d'extraits de questionnaires d'enquête ou d'épreuves psychologiques. Il nous a souvent semblé indispensable en effet, soit de justifier la nécessité de considérations théoriques relativement complexes et abstraites en montrant qu'elles correspondent à un besoin pratique, soit de vérifier le bien-fondé de certaines conclusions en les confrontant à des exemples concrets. En outre, le peu d'intérêt que portent certains chercheurs universitaires en sciences sociales à la technologie de l'enquête par questionnaire, et les idées simplistes qu'ils contribuent à diffuser au sujet de cet outil de recueil d'informations sur le terrain dont ils n'ont pas la pratique, nous ont d'autant plus incité à tenter d'en montrer la complexité.

Il serait d'ailleurs certainement très utile et très intéressant de consacrer un manuel entier au seul questionnaire d'enquête, présenté selon l'optique du praticien. A notre connaissance, un tel manuel n'existe, en langue française, qu'à l'état embryonnaire ; il serait pourtant susceptible de rendre de grands services aux chercheurs en sciences sociales qui

n'ont pas la possibilité d'apprendre les règles empiriques de l'élaboration d'un questionnaire par la pratique dans un organisme de sondage. Un tel manuel apparaîtrait comme une image en creux de notre propre recherche, en donnant une place prépondérante aux problèmes que nous avons volontairement négligés. Il devrait en particulier traiter d'une manière détaillée, illustrée de nombreux exemples, des principes élémentaires de formulation d'une question (forme syntaxique, connotations des mots utilisés, etc.) ; de la décomposition d'un problème en une série de questions destinées à en aborder les divers aspects (p.ex. : questions d'opinions suivies de questions portant sur un comportement hypothétique de la personne interrogée) ; des questions standardisées sur un thème donné (p.ex. : séries de questions permettant de cerner la classe sociale de référence, la pratique religieuse, le comportement de lecture d'un périodique) ; de la progression dans les thèmes abordés successivement dans le questionnaire ; de la conception générale du questionnaire selon le mode d'administration adopté ; de l'organisation logique du questionnaire (bon usage des questions-filtres, décomposition en sous-questionnaires, organigramme) ; de la présentation matérielle du questionnaire, et de son effet sur les réponses recueillies. En outre, un tel manuel pourrait signaler l'existence d'effets parasites souvent mal contrôlés (incidence sur les réponses de la formulation de la question, de son contexte, et de la situation d'entretien), et présenter les moyens de contrôler ou d'atténuer ces effets. Enfin, il pourrait mentionner plus brièvement certaines techniques de recueil d'information par questionnaire plus élaborée, les unes classiques (échelles d'attitudes, items de type projectif), les autres encore au stade expérimental (étude des processus de décision par la méthode des "tableaux d'information" ; exploration simultanée de plusieurs dimensions par la méthode des scénarios ; recensement de comportements jugés dévalorisants ou culpabilisants par la technique dite "des réponses randomisées"; etc.). Toutefois, si l'on veut qu'un tel manuel soit autre chose qu'une simple énumération de recettes empiriques, il semble indispensable de disposer, préalablement à sa conception, d'une part d'un cadre théorique cohérent, d'autre part d'un ensemble de résultats expérimentaux portant sur la

formulation des questions en français. D'où l'importance pratique de ces deux directions de recherches.

Comme nous l'avons signalé dans le cours de notre exposé, l'étude des questionnaires utilisés dans les sciences de l'homme constitue un champ de recherches étendu, dans lequel subsistent de vastes domaines encore peu explorés. Il nous a paru intéressant de recenser les principaux de ces domaines, ainsi que certains domaines adjacents dont l'exploration peut avoir une incidence sur la problématique du questionnement. Compte-tenu de l'optique adoptée dans cette recherche, qui est celle de l'analyse secondaire de résultats d'enquêtes, nous mentionnerons brièvement tout d'abord trois domaines relativement extérieurs à notre propos : celui des liens entre le questionnaire et la problématique de son auteur, celui de l'incidence de la situation d'enquête sur les réponses, et enfin celui des techniques statistiques d'analyse des données. Nous aborderons ensuite plus complètement les problèmes logico-linguistiques posés par les questions, ceux liés aux formes de consignes, et enfin ceux qui concernent la structure des questionnaires.

Dans le domaine de la *problématique de l'auteur d'un questionnaire* concret, il y aurait nous semble-t-il, à entreprendre des recherches d'ordre épistémologique, concernant en particulier ce que nous avons appelé le "questionnaire théorique" du chercheur : comment s'élaborent et se succèdent, au cours du processus de recherche, les différents "questionnaires théoriques" ; quelles transformations ont été nécessaires pour passer d'un questionnaire théorique au questionnaire concret qu'il a engendré ; quelles sont les règles qui président à la construction de variables *a priori*, c'est-à-dire à partir de la seule conjonction du questionnaire théorique et du questionnaire concret. Les recherches empiriques sur la *situation d'enquête* ou la situation d'entretien sont très nombreuses, et ont apporté au sociologue de multiples informations utiles pour la construction du questionnaire, la formation des enquêteurs, le contrôle de certains biais, etc. Peut-être serait-il possible d'intégrer ces informations dans un modèle formel du type de celui de HARRAH sur la communication, de celui de

WEJLAND sur l'entretien, ou de ceux de HAMBLIN ou de EGLI sur le dialogue, afin d'élaborer une théorie formelle de la situation d'enquête qui soit plus réaliste. A défaut, il serait intéressant de confronter le comportement rationnel de l' "*homo eroteticus*" défini par ces modèles avec les comportements réels des enquêteurs et des répondants. Enfin, dans le champ de l'*analyse des données*, nous souhaiterions voir se développer les recherches statistiques liées au dénombrement des relations \mathcal{D} dans les réponses, en vue de mettre au point un ensemble d'outils adaptés aux différents types de variables pouvant être construites à partir des réponses à un questionnaire.

Le domaine des *recherches logiques et linguistiques* sur les questions reste très ouvert. Rappelons en premier lieu que les travaux des linguistes sur l'interrogation et ceux des logiciens sur les questions se poursuivent, et que de nouveaux résultats sont publiés chaque année. Ainsi, un ouvrage récemment composé sous la direction de HIŽ rassemble entre autres des contributions de HARRIS, de KARTTUNEN, et de HINTIKKA sur ce sujet. Il y a donc d'abord un travail de recherche bibliographique, et de confrontation des diverses théories, à poursuivre non seulement en logique érotétique, mais également en linguistique. Si peu de logiciens français s'intéressent aux questions, il existe en revanche de nombreux linguistes qui ont traité de l'interrogation en français (cf. p.ex. les travaux récents de J. MILNER, ou A. BORILLO). Peut-être les sociologues pourraient-ils profiter de ce contexte favorable pour susciter des recherches linguistiques sur la syntaxe et le vocabulaire de la langue utilisée dans les questionnaires, ou pour requérir la participation de linguistes à des recherches expérimentales sur l'incidence des variations dans la formulation des questions d'enquête sur la distribution des réponses. D'autre part, une étroite collaboration entre linguistes et spécialistes du questionnaire serait nécessaire pour approfondir deux points qui nous paraissent essentiels pour la constitution future d'une théorie du questionnaire : la définition de transformations qui permettraient de passer d'une question compréhensive à une question extensive,

et l'analyse de la gradation dans la perspective des travaux de SAPIR et de DUCROT. Enfin, il serait utile d'examiner les possibilités d'application, au texte des questionnaires, des diverses procédures existantes de traitement automatique des textes en langues naturelles.

Les recherches relatives aux *formes de consignes* concernent le psychologue au moins autant que le sociologue. Dans la perspective ouverte par COOMBS, il est possible de procéder à une étude systématique des formes de consignes et de leurs propriétés : nombre de réponses possibles pour un nombre de modalités donné, puissance, capacité de canal, effets des contraintes sur les réponses possibles, etc. En outre, il pourrait être intéressant d'examiner les conséquences, sur le résultat de l'application d'une consigne, des hypothèses faites par le chercheur sur l'ensemble des modalités, ou sur l'ensemble des réponses possibles. Ce type de recherches porte d'ailleurs en fait sur la construction de variables *a priori*, mais dépend très étroitement de la forme de la consigne. Parmi les hypothèses faites couramment dans la pratique des tests et des questionnaires, figure celle de l'équivalence théorique de certaines réponses (équipondération des bonnes réponses, ou des réponses exprimant une même attitude latente) ; dans cette optique, il serait utile, pour chaque forme de consigne, de calculer la probabilité de fournir un ensemble donné de "bonnes réponses" en répondant "au hasard", en dénombrant les patrons de réponses correspondant à un score de réussite ou d'aptitude donné. A titre d'exemple différent, mais toujours dans la même optique, il serait possible d'étudier les effets, au niveau de la variable, d'une consigne de sélection multiple portant sur des modalités que le chercheur considère comme ordonnées. D'autre part, une extension envisageable de la formalisation des consignes usuelles consisterait à élaborer une véritable algèbre des consignes, en introduisant, dans notre notation inspirée de celle de COOMBS, le *nom* des ensembles auxquels la consigne s'applique (et non plus seulement leur cardinal). Il y a là une direction de recherches formelles pouvant déboucher sur une théorie des batteries de questions (cf. HOLM), et donc sur la constructions d'échelles (au sens de GUTTMAN).

La *structure des questionnaires* enfin offre de très nombreux thèmes de recherches. Au niveau de la théorie, signalons d'abord qu'il est facile de réduire l'ensemble des relations non sémantiques entre questions, ou entre questions et réponses, à la relation de pointage (\mapsto) ; le fait que nous n'ayons pas perçu l'intérêt pratique de cette réduction ne permet en rien de juger de son intérêt théorique. Remarquons ensuite que nous n'avons pas tiré directement parti, dans notre formalisation, des travaux de C.F. PICARD et surtout de la thèse de B. BOUCHON sur la réalisation de questionnaires, bien que leur problématique nous ait certainement inspiré. En particulier, la tentative faite par B. BOUCHON, d'appliquer la notion de "questionnaire flou" à la formalisation de la procédure de mise au point (par pré-tests successifs) d'un questionnaire d'enquête, mérite d'être suivie par les sociologues ; elle pourrait conduire à l'automatisation partielle de l'élaboration de certains types de questionnaires. Dans l'optique de l'archivage et de l'analyse secondaire de données d'enquêtes, on peut envisager de définir d'autres formes canoniques du questionnaire, mieux adaptées à divers besoins spécifiques. Sur le plan pratique, il serait relativement facile de mettre au point un langage formel de description d'un questionnaire, distinguant entre les rôles informatif et structurel des questions, et précisant la signification exacte donnée par les enquêteurs ou les répondants à des consignes verbales souvent imprécises ; ce langage pourrait, à l'instar du langage *ALGOL* pour les algorithmes numériques, servir à la fois à la communication entre spécialistes, et à la définition de la structure logique d'un fichier archivé. A partir d'une telle description, il deviendrait possible de concevoir un logiciel de codage-décodage des informations archivées, pouvant en particulier réaliser automatiquement la mise sous forme canonique d'un questionnaire quelconque, et le stockage des réponses recueillies sous la forme la plus économique possible (codage binaire des informations, et enregistrements de longueur variable).

Les possibilités de recherches sur le questionnaire que nous venons de décrire ne sont pas les seules concevables ; mais ce sont les directions qui nous paraissent les plus faciles à explorer dans l'immédiat, compte-tenu de l'état d'avancement des travaux antérieurs. Il faut noter que seule l'expérimentation sur la formulation des questions nécessite des moyens importants (à l'échelle d'un budget de recherches en sciences humaines) ; pour toutes les autres formes de recherches que nous avons recensées, le travail à temps partiel d'une petite équipe de chercheurs peut suffire. Par contre, deux principes nous paraissent devoir impérativement présider à toute recherche dans la perspective présentée ici : le respect absolu des données d'observation (c'est-à-dire des éléments d'un corpus de questionnaires concrets suffisamment étendu et varié), et la pluridisciplinarité de l'équipe de recherches.

A N N E X E I
LISTE DES TESTS ET QUESTIONNAIRES CITÉS

(Les références précédées d'un astérisque indiquent les tests ou questionnaires constituant le corpus d'énoncés de questions utilisé pour l'analyse morpho-syntaxique présentée au chapitre 1).

- *ABS "ABS" (Test collectif d'"abstraction" pour adultes) INOP.
- *BERGER *Questionnaire caractérologique*, par Gaston BERGER (Paris, PUF, 1950).
- *BERN 35 "Questionnaire de personnalité de Bernreuter". Adaptation française de "The Personality Inventory", par Robert G. BERNREUTER (Stanford University Press, 1935), par le Centre de Psychologie Appliquée, 1951.
- *BNIST Enquête postale sur l'information scientifique et technique, menée par le Bureau National pour l'Information Scientifique et Technique (Ministère de l'Industrie et de la Recherche) et Bertin et Cie.
- *BS Z Révision, par René ZAZZO, de l'échelle de BINET-SIMON (cf. [Binet 1911]), dit "Binet-Simon 1949" (cf. [Zazzo 1949 ; 1966 : I 55-56]).
- *BV 8 "BV 8. Test du Pr. R. BONNARDEL" (Collectif) Editions Scientifiques et Psychotechniques, 1967.
- *BV 50 "BV 50. Test du Pr. BONNARDEL" (Collectif) Editions Scientifiques et Psychotechniques, 1965.

- *CAT 16 PF Questionnaire de personnalité : "Test 16 PF de R.B. Cattell". Edition française. (Institute for Personality and Ability Testing, 1950). Centre de Psychologie Appliquée, 1951.
- *CC 8 Epreuves de calcul (Epreuve Collective de Connaissances). Service de Recherches de l'INOP.
- *CEREQ-SEU Enquête postale : "Etudiants de la promotion sortie en 1975". CEREQ - Secrétariat d'Etat aux Universités, 1976.
- *CHAUDAGNE "Questionnaire d'intérêts pour garçons de 13 à 16 ans" établi par H. CHAUDAGNE. Service de recherches de l'INOP.
- *D 48 "Test D 48" (Collectif). Centre de Psychologie Appliquée, 1969.
- *EAG 7/6 Epreuve d'anxiété pour garçons (Collective). Service de Recherches de l'INOP.
- *EAI 5/1 "Epreuves Analytiques d'Intelligence (5e)" (Collectives). Cahier 1 (verbal-numérique) Service de Recherches de l'INOP.
- *EAI.5/2 "Epreuves Analytiques d'Intelligence (5e)" (Collectives). Cahier 2 (numérique-spatial) Service de Recherches de l'INOP.
- *ECO 5 "E. VALIN. Epreuve composite d'intelligence pour l'orientation scolaire au niveau de la fin du Cycle d'Observation". Service de Recherches de l'INOP.
- *E-I/6 "Questionnaire d'extraversion-introversion pour élèves de 6e". Service de Recherches de l'INOP, 1974.
- *FABLES Technique projective en psychanalyse infantile, publiée dans [Düss 1950 : 16-18].

- *F.ING 69 Enquête postale sur les femmes ingénieurs en France. Questionnaire publié dans : PESLOUAN, Geneviève de : *Qui sont les femmes-ingénieurs en France ?* (Paris, PUF, 1974), pp. 139-157.
- *GEMAS 76 "Enquête sur la perception des inégalités". CNRS, Groupe d'Etude des Méthodes Appliquées à la Sociologie (GEMAS).
- GILLE 44 "Test mosaïque" de René GILLE. (Epreuve collective utilisée dans l'enquête INED-INOP de 1944) in : [Gille 1950 : 84-94].
- *IFOP 58-9 Enquête d'opinion du 3 septembre 1958. Questionnaire publié dans : *Le référendum de septembre et les élections de novembre 1958* (Paris, A. Colin, 1960), 185-187.
- *IFOP 58-11 Enquête d'opinion du 3 novembre 1958. Questionnaire publié dans : *Le référendum de septembre et les élections de novembre 1958* (Paris, A. Colin, 1960), 188-190.
- *IFOP 58-12 Enquête d'opinion du 12 décembre 1958. Questionnaire publié dans : *Le référendum de septembre et les élections de novembre 1958* (Paris, A. Colin, 1960), 191-193.
- *IFOP 69-5 Enquête d'opinion du 12 mai 1969. (Elections présidentielles). Questionnaire F 991.
- *INED 081 Enquête sur le choix du conjoint. INED, Section de Psychosociologie, 1959. Questionnaire publié dans [Girard 1964 : 43-45].
- *KRX 25 "K/R X 25" (Collectif). Test établi d'après le principe du "K2" par Victor Kouteynikoff. 1955.
- *NB "Test de niveau. Forme B. D'après Lewis M. Terman" (épreuve individuelle) INOP.

- **NIV. INT* "Echelle collective de niveau intellectuel" (INED-INOP. Enquête de 1965). Présentée dans : [Benedetto 1969].
- **OCDE-IUT* "Les étudiants des I.U.T.". Enquête réalisée par le Centre d'Etudes Sociologiques (CNRS) pour l'Organisation de Coopération et de Développement Economiques. (Administration collective dans les I.U.T.).
- **ORTF 65* "Sondage Radio-Télévision". ORTF, mai 1965.
- **QIA/m* "Questionnaire d'intérêts. 2e cycle de l'enseignement secondaire". Garçons. Service de Recherches de l'INOP.
- **QI Filles* "Questionnaire d'intérêts - Filles (version expérimentale en cours d'étude)". Service de Recherches de l'INOP.
- **QI Garçons* "Questionnaire d'intérêts pour garçons de 10 à 12 ans". Service de Recherches de l'INOP.
- **REL 55* "Etude religieuse. Diocèse d'Autun". 1955. Spécimen de bulletin pour recensement de messés, publié dans [Maître 1972 : 62-63].
- RNJR 70* "Enquête Loisirs et Mode de vie". Régie Nationale des Unises Renault. Commissariat Général du Plan - CREDOC-SEDES. (janvier 1970).
- **SOFRES 340* "Etude UGP". SOFRES, contrat n° 340 (1963).
- **SOFRES 571* Etude "Attraction de Metz". SOFRES, contrat n° 571 (novembre 1964).
- **SOFRES 659* "Etude Metz-Urbanisme". SOFRES, contrat n° 659 (avril 1965).
- **SOFRES 686* Enquête "MODEC". SOFRES, contrat n° 686 (1965).

- *SOFRES 769 Enquête "10.000 chefs de Ménage". SOFRES, contrat n° 769 (1965).
- *SOFRES 967 "Etude VPF, février 1967". SOFRES, contrat n° 967.
- SOFRES 1061 Etude "Reinitas". SOFRES, contrat n° 1061 (mars 1967).
- SOFRES 1094 Enquête Image Presse. SOFRES, contrat n° 1094. Code enquête : XK.
- SRC 76 Post Survey Research Center (Univ. of Michigan) "1976 Post-election Study" (P 495456, automne 1976) (questionnaire conçu pour lecture optique des codifications portées par l'enquêteur).
- SRC 76 Pre Survey Research Center (Univ. of Michigan) "1976 Pre-election Study" (P 495453, automne 1976).
- *STEIN Phrases à compléter de STEIN (adaptation française du *Stein Sentence Completion Test*, par D. BONNET, Thèse de Doctorat de Médecine, 1953).
- *SUBES "Tests de niveau scolaire de J. et M. Subes. Orthographe et calcul". Centre de Psychologie Appliquée, 1954.
- *TACHES "Taches d'encre". cf. : BENASSY, M. ; CHAUFFARD, C.: "Une épreuve de tâches d'encre envisagée sous l'aspect quantitatif". *Année Psychologique*, 45-46 (1944-45).
- *TERMAN "Consignes du test de TERMAN". INOP, 1957.
- TM 37 Test de Terman-Merrill [Terman 1937]. Adaptation française : CESSÉLIN, Félix : *Comment évaluer le niveau intellectuel* (Paris, Bourrelier, 1952).
- *TMI "T.M.I. (Terman-Merrill I.N.O.P.). Adaptation et révision de l'échelle de niveau mental de L.M. Terman et M.A. Merrill". Travail du Centre d'Application de l'INOP, 1956.

A N N E X E II
MESURE DE L'INFORMATION STRUCTURELLE

Le dénombrement dans E^2 des relations \mathcal{Y} et \mathcal{D} structurelles est assez fastidieux à effectuer manuellement, surtout si la variable étudiée est de structure particulièrement complexe. Il est d'ailleurs préférable, dans ce cas, de procéder au préalable à l'énumération des relations dans $E^2 \times (\mathcal{P}(E))^2$, ou dans $E^2 \times (\{E\} \times \mathcal{P}(E))$, ce qui constitue une tâche relativement longue, et présentant des risques d'erreurs certains pour peu que le nombre d'indicateurs soit élevé. Par contre, le dénombrement dans $(E_i \times E_j)^2$ des relations structurelles résultant d'un croisement de variables, ou encore le calcul de la quantité d'information métrique engendrée par le questionnement, sont des opérations assez simples à réaliser ou à programmer, à condition de disposer préalablement des résultats du dénombrement des logons correspondant aux variables étudiées.

C'est pourquoi nous avons cru utile de présenter dans cette annexe un court programme *FORTTRAN* qui effectue le dénombrement automatique des relations \mathcal{Y} et \mathcal{D} , dans $E^2 \times (\mathcal{P}(E))^2$ et dans $E^2 \times (\{E\} \times \mathcal{P}(E))$, pour n'importe quelle structure de variable. Ce programme est aussi simple que possible. Nous avons en effet préféré la transparence de l'algorithme à la performance, d'autant que l'exécution du programme pour une variable demande généralement moins de dix centièmes de seconde de temps CPU sur IBM 370/168 (le temps de compilation, par contre, est de l'ordre de deux secondes). La version que nous présentons ne traite que des variables de 10 indicateurs au plus, comportant 25 sous-ensembles au maximum ; elle peut traiter un nombre quelconque de variables. Il est évidemment facile d'en augmenter les capacités en modifiant le dimensionnement et les formats.

Dans ce programme, les données sont constituées par une description des variables à analyser. Cette description obéit aux règles suivantes :

- chaque variable occupe une carte et une seule ;
- la description de la variable commence en colonne 1, se termine par un astérisque, et ne comporte aucun blanc avant l'astérisque ;
- chaque modalité est identifiée par un caractère alphabétique quelconque ;
- chaque sous-ensemble est décrit par l'énumération (sans séparateur) des indicateurs qui en sont les éléments ;
- les sous-ensembles sont séparés par une barre oblique ("slash" : /) ;
- la fin du fichier des données est signalée au programme par une carte portant un astérisque en colonne 1.

Indicateurs et sous-ensembles peuvent être présentés dans un ordre quelconque. En sortie, les indicateurs sont reclassés selon l'ordre alphabétique.

A titre d'exemple de description d'une variable, considérons la question présentée à la fin du § 3.3.1.d, dont la structure peut s'écrire :

$$P = \{ \{v,w,x\}, \{w,x\}, \{x\}, \{x,y\}, \{x,y,z\}, \{t\}, \{u\} \}$$

Cette structure sera codée dans les colonnes 1 à 20 comme suit :

VWX/WX/X/XY/XYZ/T/U*

Les sorties comportent trois pages par variable :

① le rappel de la carte données, et son interprétation sous la forme de relations d'appartenance entre les modalités (reclassées et désignées par leur nom) et les sous-ensembles (numérotés dans l'ordre de leur apparition sur la carte) ;

② la présentation dans E^2 des résultats des dénombrements dans $E^2 \times (\mathcal{P}(E))^2$;

③ la présentation dans E^2 des résultats des dénombrements dans $E^2 \times (\{E\} \times \mathcal{P}(E))$.

C'est cette dernière page qui a été utilisée pour la figure n° 3.31, correspondant à la variable ci-dessus.

```

COMMON NT(25,10),A(10),NK,NL
DATA STR/'*'/
2 CONTINUE
CALL LECT
IF(A(1).EQ.STR)GOTO 1
CALL DENBR
GOTO 2
1 STOP
END
C LECTURE DE LA DESCRIPTION DE LA VARIABLE
SUBROUTINE LECT
COMMON NT(25,10),A(10),NK,NL
DIMENSION E (80)
DATA SLH/'/',STR/'*'/
101 FORMAT(80A1)
102 FORMAT (/5X,A1,2X,25I3)
103 FORMAT (/8X,25I3)
104 FORMAT('1DONNEES : ',80A1/1X,7(1H-)/)
READ (5,101) E
A(1)=E(1)
IF(E(1).EQ.STR)GOTO 9
WRITE(6,104)E
DO 7 I=1,25
DO 7 J=1,10
7 NT(I,J)=0
NT(1,1)=1
NL=1
NK=1
DO 1 I=2,80
B=E(I)
IF(B.EQ.SLH)GOTO 5
IF(B.EQ.STR)GOTO 2
DO 3 J=1,NL
IF(B.EQ.A(J))GOTO 4
3 CONTINUE
NL=NL+1
A(NL)=B
NT(NK,NL)=1
GOTO 1
4 NT(NK,J)=1
GOTO 1
5 NK=NK+1
1 CONTINUE
2 CONTINUE
CALL CLASS
WRITE(6,103)(I,I=1,NK)
DO 6 J=1,NL
6 WRITE(6,102) A(J),(NT(I,J),I=1,NK)
9 RETURN
END

```

C RECLASSEMENT ALPHABETIQUE DES INDICATEURS

```

SUBROUTINE CLASS
COMMON NT(25,10),A(10),NK,NL
DIMENSION D(25)
NB=NL-1
DO 1 K=1,NB
NI=NL-K
N=0
DO 2 I=1,NI
J=I+1
IF(A(I).GT.A(J))GOTO 2
B=A(I)
A(I)=A(J)
A(J)=B
DO 3 L=1,NK
D(L)=NT(L,I)
NT(L,1)=NT(L,J)
NT(L,J)=D(L)
3 CONTINUE
N=N+1
2 CONTINUE
IF (N.EQ.0)GOTO 4
1 CONTINUE
4 RETURN
END

```

C DENOMBREMENT DANS E2 DES RELATIONS S ET D

```

SUBROUTINE DENBR
COMMON NT(25,10),A(10),NK,NL
INTEGER DOR(10,10),DSO(10,10),SOR(10,10),SSO(10,10)
50 FORMAT ('1RELATIONS SANS URIGINE'//1X,22(1H-)//)
51 FORMAT(/3X,A1,15I5)
53 FORMAT(4X,15(4X,A1))
54 FORMAT('+',60X,A1,15I5)
55 FORMAT('+',61X,15(4X,A1))
57 FORMAT('////',3X,'RELATIONS D',47X,'RELATIONS S')
58 FORMAT('1RELATIONS AVEC ORIGINE'//1X,22(1H-)//)
70 FORMAT('////1X,'NOMBRE TOTAL DE RELATIONS : D =',
AI4,' S =',I4,' D+S =',I4)
DO 9 I=1,NL
DO 9 J=1,NL
DOR(I,J)=0
DSO(I,J)=0
SOR(I,J)=0
SSO(I,J)=0
9 CONTINUE
C (NT(M,I).EQ.1) = X ELEMENT DE A
C (NI(M,J).EQ.1) = Y ELEMENT DE A
C (NT(N,I).EQ.1) = X ELEMENT OF B
C (NT(N,J).EQ.1) = Y ELEMENT OF B

```

```

DO 1 J=1,NL
DO 1 N=1,NK
IF (NT(N,J).EQ.0)GOTO 1
DO 10 I=1,NL
IF (NT(N,I).EQ.0)DSO(I,J)=DSO(I,J)+1
IF (NT(N,I).EQ.1)SSO(I,J)=SSO(I,J)+1
DO 10 M=1,NK
IF (NT(M,I).EQ.0)GOTO 10
IF (NT(N,I).EQ.0)DOR(I,J)=DOR(I,J)+1
IF (NT(N,I).EQ.1)SOR(I,J)=SOR(I,J)+1
10 CONTINUE
1 CONTINUE
WRITE (6,50)
WRITE (6,53) (A(I),I=1,NL)
WRITE (6,55) (A(I),I=1,NL)
DO 8 I=1,NL
WRITE (6,51) A(I), (DSO(I,J),J=1,NL)
WRITE (6,54) A(I), (SSO(I,J),J=1,NL)
8 CONTINUE
WRITE (6,57)
CALL NTG(DSO,ND)
CALL NTG(SSO,NS)
NR=ND+NS
WRITE (6,70) ND,NS,NR
WRITE (6,58)
WRITE (6,53) (A(I),I=1,NL)
WRITE (6,55) (A(I),I=1,NL)
DO 7 I=1,NL
WRITE (6,51) A(I), (DOR(I,J),J=1,NL)
WRITE (6,54) A(I), (SOR(I,J),J=1,NL)
7 CONTINUE
WRITE (6,57)
CALL NTG(DOR,ND)
CALL NTG(SOR,NS)
NR=ND+NS
WRITE (6,70) ND,NS,NR
RETURN
END
C TOTALISATION DANS E-
SUBROUTINE NTG (NTAB,NS)
COMMON NT(25,10),A(10),NK,NL
DIMENSION NTAB(10,10)
NS=0
DO 1 K=1,NL
DO 1 L=1,NL
1 NS=NS+NTAB(K,L)
RETURN
END

```



B I B L I O G R A P H I E

Chacune des références citées dans cette bibliographie est précédé de la partie fixe du renvoi à l'oeuvre en question, telle qu'elle apparaît dans notre texte. Cette partie fixe est constituée en principe du nom de l'auteur, suivi de la date de publication de l'oeuvre. En pratique, pour la définition de la partie fixe des renvois, nous avons appliqué les règles suivantes :

- lorsque le nom de l'auteur ne figure pas dans la publication, il est remplacé par la lettre A ("anonyme") suivie d'un point ;
- si l'ouvrage a plusieurs auteurs, le nom mentionné dans le renvoi est celui du premier auteur cité dans la publication ;
- s'il s'agit d'un ouvrage collectif publié sous l'égide d'un organisme, nous avons choisi d'indiquer dans le renvoi le nom ou le sigle de l'organisme (ex. : [Académie], [Ined], [Src]) plutôt que la mention "collectif" ;
- le nom figurant dans le renvoi ne tient pas compte d'éventuelles homonymies : [Katz] est mis pour le linguiste Jerrold J. KATZ et pour le psychologue social Daniel KATZ (cf. également [Garner], [Miller], [Phillips], etc.) ;
- la graphie du nom est, à quelques particularités typographiques près, la graphie française ou francisée consacrée par l'usage ;
- l'ordre de classement des références, et par conséquent des renvois, est conforme à l'ordre lexicographique français (par exemple, ÅQVIST figure à la lettre A , et non en fin d'alphabet comme le voudrait l'usage

suédois ; de même, **ΦΑΤΚΙΗ** est classé à la lettre *F* , conformément à sa transcription en caractères latins : **FATKIN** ; etc.) ;

- la date citée est en règle générale celle de la première édition du texte original de la version que nous avons utilisée (lorsque nous nous référons à une version remaniée, c'est donc la date de l'édition princeps de cette version qui figure dans le renvoi) ;

- en cas de désaccord entre la date d'édition mentionnée dans l'ouvrage et celle du copyright, nous avons fait figurer la plus ancienne de ces deux dates (cf. [Kubiński 1970]) ;

- lorsque ces règles ne suffisent pas à distinguer deux références, en général parce que nous citons deux textes du même auteur publiés la même année, nous avons fait suivre la date d'une lettre minuscule indiquant l'ordre probable de parution de ces textes au cours de l'année ;

- pour les auteurs de l'antiquité, nous n'avons pas mentionné de date, mais seulement, lorsque cela s'est révélé nécessaire pour distinguer deux ouvrages du même auteur, les abréviations usuelles des ouvrages en question.

La référence indiquée à la suite du renvoi est toujours (dans la mesure du possible) celle de la première édition connue du texte original (éventuellement après remaniement de celui-ci). Lorsque le texte que nous avons utilisé n'est pas ce texte original, mais une ré-édition ou une traduction de ce texte, nous avons fait précéder du signe → la référence de cette ré-édition ou traduction.

Dans notre texte, pour tout renvoi qui porte, non sur la totalité de l'ouvrage, mais seulement sur un passage de celui-ci, nous avons ajouté à la partie fixe du renvoi l'indication des pages où ce passage se trouve. En règle générale, la pagination indiquée est celle du texte que nous avons utilisé. Toutefois, pour certaines oeuvres, afin de faciliter la consultation d'autres éditions, nous avons adopté les règles suivantes :

- lorsqu'un ouvrage est divisé en courts paragraphes numérotés, nous avons préféré indiquer le numéro de ce paragraphe, précédé du signe § (ex. : [Bacon 1620], [Wittgenstein 1921]) ;

- lorsque la pagination de l'édition originale (ou d'une édition ultérieure faisant autorité) a été conservée dans la ré-édition ou la traduction que nous avons utilisée, c'est cette pagination qui figure dans le renvoi (ex. : [Austin 1962], [Chomsky 1965]) ; pour les auteurs de l'antiquité, nous nous sommes conformé à l'usage qui est de citer le numéro de page, suivi de l'indication de la colonne (désignée par la lettre a ou b) de la première édition imprimée de l'oeuvre (la seule exception à cette règle concerne DIOGENE LAËRCE, l'édition utilisée par nous ne comportant pas ces informations).

L'indication de la pagination peut être précédée de celle de la tomaison, et suivie de celle de divisions particulières internes au texte cité. Dans un même renvoi entre crochets, lorsque nous citons plusieurs passages d'une même oeuvre, les indications de pagination sont séparées par une virgule ; lorsque nous citons plusieurs oeuvres d'un même auteur, les indications de date et de pagination sont séparées par un point-virgule.

[A. 1965]

(Anonyme) :
Preparation for MILLER Analogies Test
(Chicago, Henry Regnery, 1965)

[Académie 1932]

ACADÉMIE FRANÇAISE :
Grammaire de l'académie française
(Paris, Firmin-Didot, 1932)

[Adams 1956]

ADAMS, J. Stacy :
"An Experiment on Question and Response Bias"
Public Opinion Quaterly, 20, 3 (automne 1956),
593-598

- [Adorno 1950] ADORNO, T.W. ; FRENKEL-BRUNSWIK, Else ;
LEVINSON, Daniel J. ; SANFORD, R. Nevitt :
The Authoritarian Personality
(New-York, Harper & Row, 1950)
- [Ajdukiewicz 1965] AJDUKIEWICZ, Kazimierz :
Logika Pragmatyczna (Varsovie, 1965)
→ Traduit par Olgierd WOJTASIEWICZ :
Pragmatic Logic
(Dordrecht, Reidel, 1974)
- [Albou 1968] ALBOU, Paul :
Les questionnaires psychologiques
(Paris, PUF, 1968)
- [Alleton 1973] ALLETON, Viviane :
Grammaire du chinois
(Paris, PUF, 1973)
- [Anderson 1951] ANDERSON, Gladys L. :
"Qualitative Aspects of the Stanford-Binet"
in : ANDERSON, Harold H. ; ANDERSON, Gladys L. :
*An Introduction to Projective Techniques and
Other Devices for Understanding the Dynamics
of Human Behavior*
(Englewood Cliffs, Prentice-Hall, 1951), 581-603
- [Anderson 1958] ANDERSON, Theodore R. ; ZELDITCH, Morris, Jr. :
*A Basic Course in Statistics with Sociological
Applications*
(New-York, Rinehart & Winston, 1958)
- [Anzieu 1960] ANZIEU, Didier :
Les méthodes projectives
(Paris, PUF, 1960)

- [Åqvist 1965] ÅQVIST, Lennart :
A New Approach to the Logical Theory of Interrogatives. Part 1 : Analysis
 (Uppsala, Almqvist & Wiksells, 1965)
- [Åqvist 1969] ÅQVIST, Lennart :
 "Scattered Topics in Interrogative Logic",
 in DAVIS, J.W. ; HOCKNEY, D.J. ; WILSON, W.K.,
 (Eds) : *Philosophical Logic*.
 (Dordrecht, D. Reidel, 1969), 114-121
- [Åqvist 1971] ÅQVIST, Lennart :
 "Revised Foundations for Imperative - Epistemic
 and Interrogative Logic"
Theoria, 37 (1971), 33-73
- [Aristote] ARISTOTE :
Organon
 Traduit par J. TRICOT - 5 tomes
 (Paris, J. Vrin, 1950 à 1962)
- [Arnault 1683] ARNAULT, Antoine ; NICOLE, Pierre :
La logique ou l'art de penser
 (5e édition, Paris, Guillaume Després, 1683)
 → Réédité avec une introduction de Louis MARIN
 (Paris, Flammarion, 1970)
- [Attneave 1959] ATTNEAVE, Fred :
*Applications of Information Theory to Psychology :
 A Summary of Basic Concepts, Methods, and Results*
 (New York, Henry Holt, 1959)
- [Austin 1962] AUSTIN, John Langshaw :
How to do things with words
 (Oxford University Press, 1962)
 → Traduit par Gilles LANE :
Quand dire, c'est faire
 (Paris, seuil, 1970)
 * (avec indication de la pagination de
 l'original)

- [Bacher 1975] BACHER, F. ; DEMANGEON, M. ; HUTEAU, M. ; LARCEBEAU, S. ; PELNARD-CONSIDÈRE, J. :
Brochure d'information sur les tests de l'I.N.O.P.
(Paris, INOP, 1975)
- [Bacher 1978] BACHER, Françoise :
L'organisation et l'interprétation des enquêtes en psychologie
Thèse d'état pour le Doctorat ès lettres et sciences humaines, Université Paris V, 1978
- [Bacon 1605] BACON, Francis (Lord VERULAM) :
The Dignity and Advancement of Learning (1605)
→ réédité avec introduction et notes par James Edward CREIGHTON
(New-York, Willey, 1900)
- [Bacon 1620] BACON, Francis (Lord VERULAM) :
Novum Organum (1620)
→ réédité avec introduction et notes par James Edward CREIGHTON
(New-York, Willey, 1900)
- [Barbut 1973] BARBUT, Marc ; D'ADHEMAR, Claude ; LECLERC, Bruno ; JULLIEN, Pierre :
Mathématiques élémentaires. Applications à la statistique et aux sciences sociales. 2 volumes
(Paris, PUF, 1973-1974)
- [Bar-Hillel 1953] BAR-HILLEL, Yehoshua ; CARNAP, Rudolf :
"Semantic Information"
British Journal for the Philosophy of Science,
4 (1953-54), 147-157
- [Barker 1974] BARKER, John A. :
"Knowledge, Ignorance and Presupposition"
Analysis, 35, 2 (décembre 1974), 33-45

- [Bell 1970] BELL, Charles G. :
 "The Joys and Sorrows of Secondary Data Use"
 in : BISCO, Ralph L. (Ed.) :
Data Bases, Computers, and the Social Sciences
 (New-York, Wiley, 1970), 52-60
- [Belnap 1963] BELNAP, Nuel D., Jr. :
An Analysis of Questions. Preliminary Report
 (Santa Monica, Cal., System Development Corporation, 1963)
- [Belnap 1966] BELNAP, Nuel D., Jr. :
 "Questions, Answers, and Presuppositions"
The Journal of Philosophy, 63, 20 (octobre
 1966), 609-611.
- [Belnap 1969] BELNAP, Nuel D., Jr. :
 "Åqvist's Corrections-Accumulating Question-
 Sequences"
 in DAVIS, J.W. ; HOCKNEY, D.J. ; WILSON, W.K.,
 (Eds) : *Philosophical Logic*, (Dordrecht,
 D. Reidel, 1969), 122-134
- [Belnap 1972] BELNAP, Nuel D., Jr. :
 "S-P Interrogatives"
Journal of Philosophical Logic, 1 (1972)
 331-346
- [Belnap 1976] BELNAP, Nuel D., Jr. ; STEEL, Thomas B., Jr. :
The Logic of Questions and Answers
 (New Haven, Yale University Press, 1976)
- [Benedetto 1969] BENEDETTO, Pierre :
 "Echelle collective de niveau intellectuel"
 in [Ined 1969 : 33-50]

- [Benveniste 1958] BENVENISTE, Émile :
 "De la subjectivité dans le langage"
Journal de Psychologie, juillet-septembre 1958
 → réédité dans :
Problèmes de linguistique générale
 (Paris, Gallimard, 1966), 258-266
- [Benveniste 1962] BENVENISTE, Émile :
 "Pour l'analyse des fonctions casuelles : le
 génitif latin"
Lingua, 11 (1962)
 → réédité dans :
Problèmes de la linguistique générale
 (Paris, Gallimard, 1966), 140-148
- [Benveniste 1963] BENVENISTE, Émile :
 "La philosophie analytique et le langage"
Les études philosophiques, 1 (janvier-mars 1963)
 → réédité dans :
Problèmes de linguistique générale
 (Paris, Gallimard, 1966), 267-276
- [Berge 1970] BERGE, Claude :
Graphes et hypergraphes
 (Paris, Dunod, 1970)
- [Bernard 1865] BERNARD, Claude :
Introduction à l'étude de la médecine expérimentale
 (Paris, 1865)
 → réédition (Paris, Garnier-Flammarion, 1966)
- [Berthier 1971] BERTHIER, Nicole ; BERTHIER François :
Le sondage d'opinion
 (Paris, Bordas, 1971)
- [Bertier 1975] BERTIER, Patrice ; BOUROCHE, Jean-Marie :
Analyse des données multidimensionnelles
 (Paris, PUF, 1975)

- [Bickerton 1975] BICKERTON, Derek :
 "Two levels of logical presupposition"
*Papers from the eleventh regional meeting of
 the Chicago Linguistic Society*
 (Université de Chicago, 1975), 48-59
- [Binet 1900] BINET, Alfred :
La suggestibilité
 (Paris, Schleicher, 1900)
- [Binet 1907] BINET, Alfred ; SIMON, Th. :
Les enfants anormaux
 (Paris, Colin, 1907)
- [Binet 1909] BINET, Alfred :
Les idées modernes sur les enfants
 (Paris, Flammarion, 1909)
 → réédité par Flammarion en 1973, avec une
 préface de Jean PIAGET
- [Binet 1911] BINET, Alfred ; SIMON, Th. :
 "La mesure du développement de l'intelligence
 chez les jeunes enfants"
Bulletin de la Société Alfred Binet, 70 (1911)
 → réédité en fascicule sous le même titre
 (Paris, Bourrelier, 1954)
- [Bize 1958] BIZE, P.R. :
Sélection et orientation professionnelles
 Cours du Conservatoire National des Arts et
 Métiers - Edité en plusieurs fascicules
 (Paris, CDU-SEDES, 1958-1960)
- [Bizos 1949] BIZOS, M. :
Syntaxe grecque
 (Paris, Vuibert, 1949)

- [Blanché 1957] BLANCHÉ, Robert :
Introduction à la logique contemporaine
(Paris, Armand Colin, 1957)
- [Blankenship 1940] BLANKENSHIP, Albert B. :
"The Influence of the Question form upon the
Response in a Public Opinion Poll"
The Psychological Record, 3, 23 (mars 1940),
347-422
- [Bloch 1945] BLOCH, Oscar ; GEORGIN, René :
*Grammaire française (Classe de 4e et classes
supérieures)*
(Paris, Hachette, 1945)
- [Bloomfield 1961] BLOOMFIELD, Leonard :
Language
(New-York, Holt, Rinehart & Winston, 1961)
→ Traduit par Janick GAZIO :
Le langage
(Paris, Payot, 1970)
- [Blum 1971] BLUM, Y. ; BRISSON, J. :
"Implication et publicité"
Langue française, 12 (décembre 1971), 83-89
- [Boudon 1967] BOUDON, Raymond :
L'analyse mathématique des faits sociaux
(Paris, Plon, 1967)
- [Boudon 1969] BOUDON, Raymond :
Les méthodes en sociologie
(Paris, PUF, 1969)
- [Bourdieu 1968] BOURDIEU, Pierre ; CHAMBOREDON, Jean-Claude ;
PASSERON, Jean-Claude :
Le métier de sociologue
(Paris, Mouton-Bordas, 1968)

- [Brochard 1912] BROCHARD, Victor :
Etudes de philosophie ancienne et de philosophie moderne
 (Paris, Félix Alcan, 1912)
- [Bromberger 1966a] BROMBERGER, Sylvain :
 "Why-Questions"
 COLODNY, Robert G. (Ed) :
Mind and Cosmos, vol. III
 (Pittsburg, University Press, 1966), 68-111
 → réédité dans :
 BRODY, Baruch A. :
Readings in the Philosophy of Science
 (Englewood Cliffs, Prentice Hall, 1970), 66-87
- [Bromberger 1966b] BROMBERGER, Sylvain :
 "Questions"
The Journal of Philosophy, 63, 20 (octobre 1966),
 597-606.
- [Brunot 1926] BRUNOT, Ferdinand :
La pensée et la langue. Méthode, principes et plan d'une théorie nouvelle du langage appliquée au français. Troisième édition revue
 (Paris, Masson, 1926)
- [Brunschvicg 1922] BRUNSCHVICG, Léon :
L'expérience humaine et la causalité physique
 (Paris, Alcan, 1922)
 → réédition (Paris, PUF, 1949)
- [Bush 1963] BUSH, Robert R. ; GALANTER, Eugene ; LUCE, R. Duncan :
 "Characterization and Classification of Choice Experiments"
 in : LUCE, R. Duncan ; BUSH, Robert R. ; GALANTER, Eugene (Ed) :
Mathematical Psychology
 (New-York, Wiley, 1963), Tome 1, 77-102

- [Campbell 1970] CAMPBELL, Angus :
 "Some Questions about the New Jerusalem"
 in : BISCO, Ralph L. (Ed) :
Data Bases, Computers, and the Social Sciences
 (New-York, Wiley, 1970), 42-51
- [Cannell 1968] CANNELL, Charles F. ; KAHN, Robert L. :
 "Interviewing"
 in LINDZEY, Gardner ; ARONSON, Elliot :
The Handbook of Social Psychology - 2e édition
 (Reading, Mass., Addison-Wesley, 1968)
 tome 2, 526-595
- [Cannell 1975] CANNELL, Charles F. ; LAWSON, Sally A. ;
 HAUSSER, Doris L. :
A Technique for Evaluating Interviewer Performance
 (Ann Arbor, Survey Research Center, Institute for
 Social Research, 1975)
- [Cantril 1940] CANTRIL, Hadley :
 "Experiments in the Wording of Questions"
Public Opinion Quarterly, 4, 2 (juin 1940), 330-332
- [Cantril 1944] CANTRIL, Hadley :
Gauging Public Opinion
 (Princeton University Press, 1944)
- [Capt 1956] CAPT, Katherine Gordon :
 "The Questionnaire and Other Reporting Forms
 as Aids in Field Exploration"
 in YOUNG, Pauline V. :
Scientific Social Surveys and Research
 (Englewood Cliffs, Prentice-Hall, 1956), 176-204
- [Carnap 1952] CARNAP, Rudolf :
 "Meaning Postulates"
Philosophical Studies, 3 (1952), 65-73
 → réédité in :
*Meaning and Necessity. A Study in Semantics and
 Modal Logic* (The University of Chicago Press,
 édition augmentée, 1956), 222-229

- [Carnap 1966] CARNAP, Rudolf :
Philosophical Foundations of Physics
 (Basic Books, 1966)
 → Traduit par Jean-Mathieu LUCCIONI et
 Antonia SOULEZ :
Les fondements philosophiques de la physique
 (Paris, A. Colin, 1973)
- [Carr 1971] CARR, Leslie G. :
 "The SROLE Items and Acquiescence"
American Sociological Review, 36, 2 (avril 1971),
 287-293
- [Cartwright 1959] CARTWRIGHT, Ann :
 "The families and individuals who did not cooperate on a sample survey"
Milbank Memorial Fund Quaterly, 37, 4, (octobre 1959), 347-368
- [Cesselin 1952] CESSELIN, Félix :
Comment évaluer le niveau intellectuel.
Adaptation française du test Terman-Merrill
 2e édition remaniée
 (Paris, Bourrelier, 1952)
- [Chaumier 1976] CHAUMIER, Jacques :
Les banques de données
 (Paris, PUF, 1976)
- [Chauvineau 1974] CHAUVINEAU, Jean :
La logique moderne
 6e édition revue et corrigée
 (Paris, PUF, 1974)
- [Chevry 1962] CHEVRY, Gabriel R. :
Pratique des enquêtes statistiques
 (Paris, PUF, 1962)

- [Chomsky 1957] CHOMSKY, Noam :
Syntactic Structures
 (La Haye, Mouton, 1957)
 → Traduit par Michel BRAUDEAU :
Structures syntaxiques
 (Paris, Seuil, 1969)
- [Chomsky 1958] CHOMSKY, Noam :
 "A Transformational Approach to Syntax"
*Proceeding of the Third Texas Conference on
 Problems of Linguistic Analysis in English, 1958*
 (Austin, University of Texas, 1962), 124-158
 → réédité dans :
 FODOR, Jerry A. ; KATZ, Jerrold J. :
*The Structure of Language. Readings in the
 Philosophy of Language*
 (Englewood Cliffs, Prentice-Hall, 1964)
 211-245
- [Chomsky 1962] CHOMSKY, Noam :
 "Current Issues In Linguistic Theory"
 Exposé au neuvième congrès international des
 linguistes
 (Cambridge, Mass., 1962)
 → édité dans :
 FODOR, Jerry A. ; KATZ, Jerrold J. :
*The Structure of Language. Readings in the
 Philosophy of Language*
 (Englewood Cliffs, Prentice-Hall, 1964), 50-118
- [Chomsky 1965] CHOMSKY, Noam :
Aspects of the Theory of Syntax
 (Cambridge, MIT Press, 1965)
 → Traduit par Jean-Claude MILNER :
Aspects de la théorie syntaxique
 (Paris, Seuil, 1971)
 * (avec indication de la pagination de l'original)
- [Chomsky 1972] CHOMSKY, Noam :
Studies on Semantics in Generative Grammar
 (La Haye, Mouton, 1972°)
 → Traduit par Bernard CERQUIGLINI :
Questions de sémantique
 (Paris, Seuil, 1975)

- [Clemans 1971] CLEMANS, William V. :
 "Test Administration"
 in : THORNDIKE, Robert L. (Ed) :
Educational Measurement
 (Washington, American Council on Education,
 1971), 188-201
- [Cohen 1929] COHEN, Felix S. :
 "What is a Question ?"
The Monist, 39 (1929), 3, 350-364
- [Collart 1975] COLLART, Jean :
Grammaire du latin
 (Paris, PUF, 1975)
- [Collingwood 1940] COLLINGWOOD, R.G. :
An Essay on Metaphysics
 (Oxford, Clarendon Press, 1940)
- [Combès 1971] COMBÈS, Michel :
Fondements des mathématiques
 (Paris, PUF, 1971)
- [Comtet 1970] COMTET, Louis :
Analyse combinatoire - 2 tomes
 (Paris, PUF, 1970)
- [Conrad 1948] CONRAD, H.S. :
 "Investigating and Appraising Intelligence and
 Other Aptitudes"
 in : ANDREWS, T.G. :
Methods of Psychology
 (New-York, Wiley, 1948), 498-538
 → Traduit par Pierre OLÉRON :
 "L'examen et l'évaluation de l'intelligence et
 des autres aptitudes"
 in : ANDREWS, T.G. :
Méthodes de la psychologie
 (Paris, PUF, 1952), 612-662

- [Coombs 1953] COOMBS, Clyde H. :
 "Theory and Methods of Social Measurement"
 in : FESTINGER, Leon ; KATZ, Daniel (Ed.) :
Research Methods in the Behavioral Sciences
 (New-York, The Dryden Press, 1953)
 → Traduit par Honoré LESAGE :
 "La mesure dans les sciences sociales :
 théories et méthodes",
*Les méthodes de recherche dans les sciences
 sociales*
 (Paris, PUF, 1963), 538-611
- [Coombs 1964] COOMBS, Clyde H. :
A Theory of Data
 (New-York, Wiley, 1964)
- [Coyaud 1972] COYAUD, Maurice :
Linguistique et documentation
 (Paris, Larousse, 1972)
- [Creighton 1898] CREIGHTON, James Edwin :
An Introductory Logic
 (New-York, Mac Millan, 1898)
- [Cresswell 1963] CRESSWELL, M.J. :
 "The Logic of Interrogatives"
 in CROSSLEY, J.N. ; DUMMETT, M.A.E., (Eds) :
*Formal Systems and Recursive Functions,
 Proceedings of the Eighth Logic Colloquium
 (Oxford, July 1963)* (Amsterdam, North-Holland,
 1965), 8-11
- [Cronbach 1946] CRONBACH, Lee J. :
 "Response Sets and Test Validity"
Educational and Psychological Measurement, 6
 (1946), 475-494
- [Cronbach 1949] CRONBACH, Lee J. :
Essential of Psychological Testing
 (New-York, Harper, 1949)

- [Cronbach 1971] CRONBACH, Lee J. :
"Test Validation"
in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
(Washington, American Council on Education,
1971), 443-507
- [Crystal 1971] CRYSTAL, David :
Linguistics
(Harmondsworth, Penguin Books, 1971)
- [Dainville 1952] DAINVILLE, R.P. François de :
"Un dénombrement inédit au XVIIIe siècle.
L'enquête du Contrôleur Général ORRY. 1745"
Population, 7, 1 (janvier 1952), 49-68
- [Damourette 1931] DAMOURETTE, Jacques ; PICHON, Édouard :
*Des mots à la pensée. Essai de grammaire de la
langue française - 8 volumes*
(Paris, Editions d'Artrey, 1931-1971)
- [Daston 1968] DASTON, Paul G. :
"Word Association and Sentence Completion
Techniques"
in : RABIN, Albert I. :
*Projective Techniques in Personality Assessment.
A Modern introduction.*
(New-York, Springer, 1968), 264-289
- [Daval 1963] DAVAL, Roger ; BOURRICAUD, François ; DELAMOTTE,
Yves ; DORON, Roland :
Traité de psychologie sociale - 2 volumes
(Paris, PUF, 1963)
- [Debaty 1967] DEBATY, Pol :
La mesure des attitudes
(Paris, PUF, 1967)

- [Debordeaux 1968] DEBORDEAUX, Danièle ; SAINT-PAUL, Dominique :
La normalisation des questionnaires
(Paris, Société d'Economie et de Mathématiques Appliquées, décembre 1968)
- [Delaunay 1974] DELAUNAY, Marie-Dominique :
"La métalangue est dans la langue"
Informatique et sciences humaines, 23 (décembre 1974), 57-63
- [Descartes] DESCARTES, René :
Oeuvres
Publiées par Charles ADAM et Paul TANNERY - 11 volumes
(Paris, Léopold Cerf, 1897-1909)
- [Deutsch 1970] DEUTSCH, Karl W. :
"The Impact of Complex Data Bases on the Social Sciences"
in : BISCO, Ralph L. (Ed.) :
Data Bases, Computers, and the Social Sciences
(New-York, Wiley, 1970), 19-41
- [Diogène Laërce] DIOGENE LAËRCE :
Vie, doctrines et sentences des philosophes illustres
Traduction de Robert GENAILLE - 2 tomes
(Paris, Garnier Flammarion, 1965)
- [Dubois 1969] DUBOIS, Jean :
Grammaire structurale du français : la phrase et ses transformations
(Paris, Larousse, 1969)
- [Dubois 1970] DUBOIS, Jean ; DUBOIS-CHARLIER, Françoise :
Éléments de linguistique française : syntaxe
(Paris, Larousse, 1970)

- [Ducrot 1972] DUCROT, Oswald :
Dire et ne pas dire
(Paris, Hermann, 1972)
- [Ducrot 1973] DUCROT, Oswald (avec la collaboration de
M.C. BARBAULT et J. DEPRESLE) :
La preuve et le dire - Langage et logique
(Paris, Mame, 1973)
- [Durieux 1975] DURIEUX, Jean Loup ; PICHEREAU, Alain :
Bases de données et sciences humaines
(Paris, STERIA - Automatismes et Ingénierie,
novembre 1975)
- [Durkheim 1895] DURKHEIM, Émile :
Les règles de la méthode sociologique
(Paris, 1895)
→ réédition (Paris, PUF, 1947)
- [Durkheim 1897] DURKHEIM, Émile :
Le suicide
(Paris, Alcan, 1897)
→ réédition (Paris, PUF, 1967)
- [Düss 1950] DUSS, Louisa :
La méthode des fables en psychanalyse infantile
(Paris, l'Arche, 1950)
- [Egli 1976] EGLI, Urs ; SCHLEICHERT, Hubert :
"Bibliography of the Theory of Questions and
Answers"
in : [Belnap 1976], 155-200
- [Ennis 1964] ENNIS, Philip H. :
"Data Galaxies in the Social Sciences"
American Behavioral Scientist, 7, 10 (juin
1964), 17-20
→ réédité in :
WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
(Englewood, Information Handling Services,
1977), 5-12

- [Esmonin 1954] ESMONIN, Edmond :
 "Quelques données inédites sur VAUBAN et les
 premiers recensements de population"
Population, 9, 3 (juillet 1954), 507-512
- [Eysenck 1947] EYSENCK, H.J. :
Dimensions of Personality
 (Londres, Routledge & Kegan Paul, 1947)
 → Traduit par Madame D. MAZÉ :
Les dimensions de la personnalité
 (Paris, PUF, 1950)
- [Fair 1975] FAIR, Frank :
 "J.J. Katz' Logic of Questions : New Departure
 or Dead End ?"
Philosophical Studies, 27, 4 (avril 1975),
 283-290
- [Fatkin 1970] ФАТКИН, Л. :
 "Теория информации"
Философская Энциклопедия, Том 5
 (Москва, 1970), 210-213
- [Ferrand 1967] FERRAND, Marcel :
Pour comprendre le russe
 (Paris, Doin, 1967)
- [Ficht 1978] FICHT, Heribert :
 "Supplement to a Bibliography on the Theory
 of Questions and Answers"
Linguistische Berichte, 55 (juin 1978), 92-114
- [Fillmore 1971] FILLMORE, Charles J. :
 "Types of lexical information"
 in : STEINBERG, Danny D. ; JAKOBOVITS, Leon
 A. (Eds) :
*Semantics. An Interdisciplinary Reader in
 Philosophy, Linguistics and Psychology*
 (Cambridge University Press, 1971), 370-392

- [Flament 1976] FLAMENT, Claude :
L'analyse booléenne de questionnaire
(Paris, Mouton, 1976)
- [Fodor 1974] FODOR, J.A. ; BEVER, T.G. ; GARRET, M.F. :
The Psychology of Language. An Introduction to Psycholinguistics and Generative Grammar
(New-York, Mc Graw-Hill, 1974)
- [Fontanier 1827] FONTANIER, Pierre :
Traité général des figures du discours autres que les tropes
(Paris, De Maire-Nyon, 1827)
→ réédité dans :
Les figures du discours
(Paris, Flammarion, 1968), 269-493
- [Foulet 1921] FOULET, Lucien :
"Comment ont évolué les formes de l'interrogation"
Romania, 47 (1921), 243-348
- [Franzen 1956] FRANZEN, Raymond ; WILLIAMS, Robert :
"A Method for Measuring Error Due to Variance Among Interviewers"
Public Opinion Quaterly, 20 (1956), 587-592
- [Frege 1892] FREGE, Gottlob :
"Ueber Sinn und Bedeutung"
Zeitschrift für Philosophie und Philosophische Kritik, 100 (1892)
→ Traduit par Claude IMBERT :
"Sens et dénotation"
in : *Écrits logiques et philosophiques*
(Paris, Seuil, 1971), 102-126
- [Freudenthal 1960] FREUDENTHAL, Hans :
LINCOS : Design of a Language for Cosmic Inter-course - Part I
(Amsterdam, North-Holland, 1960)

- [Galichet 1963] GALICHET, Georges :
*Méthodologie grammaticale - Etude psychologique
des structures*
2e édition revue et augmentée
(Paris, PUF, 1963)
- [Gallup 1941] GALLUP, George :
"Question Wording in Public Opinion Polls"
Sociometry, 4, 1 (février 1941), 259-268
- [Galtung 1967] GALTUNG, Johan :
Theory and Methods of Social Research
(London, George Allen & Unwin, 1967)
- [Garner 1962] GARNER, Wendell R.
*Uncertainty and Structure as Psychological
Concepts*
(New-York, Wiley, 1962)
- [Garner 1971] GARNER, Richard :
"Presupposition in Philosophy and Linguistics"
in : FILLMORE, C.J. ; LANGENDOEN, D.T. (Eds) :
Studies in Linguistic Semantics
(New-York, Holt Rinehart & Winston, 1971), 22-42
- [Geach 1950] GEACH, P.T. :
"Russell's Theory of Descriptions"
Analysis, 10 (1950), 84-88
→ réédité in MACDONALD, Margaret (Ed) :
Philosophy and Analysis
(Oxford, Basil Blackwell, 1954), 32-36
- [Genot 1973] GENOT, Gérard :
Grammaire de l'italien
(Paris, PUF, 1973)
- [Georgin 1952] GEORGIN, René :
Guide de langue française
Nouvelle édition revue et corrigée
(Paris, André Bonne, 1952)

- [Gerhan 1975] GERHAN, David ; WALKER, Loretta :
"A Subject Approach to Social Science Data Archives"
R.Q., hiver 1975, 132-149
→ réédité in :
WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
(Englewood, Information Handling Services, 1977), 25-44
- [Getzels 1954] GETZELS, J.W. :
"The Question-Answer Process : A Conceptualisation and Some Derived Hypotheses for Empirical Examination"
Public Opinion Quarterly, 18, 1 (printemps 1954), 80-91
- [Ghiglione 1978] GHIGLIONE, Rodolphe ; MATALON, Benjamin :
Les enquêtes sociologiques. Théories et pratique
(Paris, Armand Colin, 1978)
- [Gille 1950] GILLE, René :
"Les documents de base : le questionnaire et le test mosaïque"
in : [Ined 1950 : 71-96]
- [Girard 1964] GIRARD, Alain :
Le choix du conjoint. Une enquête psychosociologique en France
(Paris, PUF, 1964)
- [Goodenough 1949] GOODENOUGH, Florence L. :
Mental Testing. Its History, Principles, and Applications
(New-York, Rinehart, 1949)

- [Gordon 1971] GORDON, David ; LAKOFF, George :
 "Conversational Postulates"
*Papers from the Seventh Regional Meeting of
 the Chicago Linguistic Society*
 (Univ. of Chicago, 1971), 63-85
 → Traduit par Marie-Lise BEFFA, in :
 "Logique et langage",
Langages, n° 30 (juin 1973), 32-55
- [Grémy 1966] GRÉMY, Jean-Paul :
 "L'analyse multidimensionnelle de l'incertitude :
 une étape vers la simulation des phénomènes so-
 ciaux"
Metra, 5 (1966), 3, 453-470
- [Grémy 1970a] GRÉMY, Jean-Paul :
*Rôle de la phase qualitative précédant
 l'enquête par sondage*
 Communication au groupe de travail sur les
 méthodes qualitatives. Ronéotypé
 (Centre d'Etudes Sociologiques, avril 1970)
- [Grémy 1970b] GRÉMY, Jean-Paul :
*La simulation sur ordinateur des phénomènes
 sociaux*
 (Paris. Institut des Sciences Humaines
 Appliquées, 1970)
- [Grémy 1975] GRÉMY, Jean-Paul :
 "Informatique et développement des sciences
 humaines",
*Actes du colloque international DGRST-CNRS
 "Informatique et Sciences Humaines"*, Marseille,
 11-13 décembre 1975
 in : *Informatique et Sciences Humaines*, 40-41
 (mars-juin 1979), 369-377.
- [Grémy 1976] GRÉMY, Jean-Paul ; LE MOAN, Marie-Joëlle :
*Analyse de la démarche de construction de
 typologies dans les sciences sociales*
 (Paris, ADISH-DGRST, septembre 1976)
 → réédité in :
Informatique et Sciences Humaines
 n° 35 (décembre 1977)

- [Grémy 1978] GRÉMY, Jean-Paul :
Rapport de mission à l'Inter-University Consortium for Political and Social Research
 (Paris, LEMTAS, mars 1978)
- [Green 1963] GREEN, Bert F.; WOLF, Alice K. ; CHOMSKY, Carol ; LAUGHERY, Kenneth :
 "Baseball : An Automatic Question Answerer"
 in : FEIGENBAUM, Edward A. ; FELDMAN, Julian (Ed.) :
Computers and Thought
 (New-York, Mac Graw Hill, 1963), 207-216
- [Gross 1967] GROSS, Maurice ; LENTIN, André :
Notions sur les grammaires formelles
 (Paris, Gauthier-Villars, 1967)
- [Gross 1968] GROSS, Maurice :
Grammaire transformationnelle du français. Syntaxe du verbe
 (Paris, Larousse, 1968)
- [Gross 1976] GROSS, Maurice :
 "Problèmes de représentation du lexique et de la syntaxe d'une langue naturelle"
 Communication au Séminaire de l'Institut de Programmation de Paris, le 10 juin 1976 (non publié)
- [Gross 1977] GROSS, Maurice :
Grammaire transformationnelle du français. Syntaxe du nom
 (Paris, Larousse, 1977)
- [Guilford 1954] GUILFORD, J.P. :
Psychometric Methods - 2e édition refondue
 (New York, Mc Graw-Hill, 1954)

- [Guiraud 1958] GUIRAUD, Pierre :
La grammaire
(Paris, PUF, 1958)
- [Guiraud 1962] GUIRAUD, Pierre :
La syntaxe du français
(Paris, PUF, 1962)
- [Hagood 1941] HAGOOD, Margaret Jarman ;
Statistics for Sociologists
(New-York, Henry Holt, 1941)
- [Halmos 1960] HALMOS, Paul R. :
Naive Set Theory
(New-York, Van Nostrand, 1960)
→ Traduit par GARDELLE, J. :
Introduction à la théorie des ensembles
(Paris, Mouton-Gauthier-Villars, 1967)
- [Hamblin 1958] HAMBLIN, C.L. :
"Questions"
The Australasian Journal of Philosophy, 36,
3 (décembre 1958), 159-168
- [Hamblin 1963] HAMBLIN, C.L. :
"Discussion : Questions aren't Statements"
Philosophy of Science, 30, 1 (janvier 1963),
62-63
- [Hamblin 1971] HAMBLIN, C.L. :
"Mathematical Models of Dialogue"
Theoria, 37 (1971), 130-155
- [Hare 1949] HARE, R.M. :
"Imperative Sentences"
Mind, 58, n° 229 (janvier 1949), 21-39

- [Harrah 1961] HARRAH, David :
 "A Logic of Questions and Answers"
Philosophy of Science, 28, 1 (janvier 1961),
 40-46
- [Harrah 1963] HARRAH, David :
Communication : a Logical Model
 (Cambridge, Mass., M.I.T. Press, 1963)
- [Harrah 1965] HARRAH, David :
 "A Model of Semantic Information and Message
 Evaluation"
 in : MASSARIK, Fred ; RATOOSH, Philburn (Eds) :
Mathematical Explorations in Behavioral Science,
 (Homewood, Ill., R.D. Irwin and The Dorsey Press,
 1965), 56-65
- [Harrah 1966] HARRAH, David :
 "Question Generators"
The Journal of Philosophy, 63, 20 (octobre 1966),
 606-608
- [Harrah 1967] HARRAH, David :
 Compte-rendu de [Åqvist 1965]
 in : *Journal of Symbolic Logic*, 32 (1967),
 403-404
- [Harrah 1972] HARRAH, David :
 Compte-rendu de [Belnap 1963]
 in : *Journal of Symbolic Logic*, 37, 2 (juin
 1972), 420-421
- [Harris 1946] HARRIS, Zellig Sabetai :
 "From Morpheme to Utterance",
Language, 22 (1946), 3, 161-183
 → réédité dans :
 [Harris 1970 : 100-125]

- [Harris 1951] HARRIS, Zellig Sabbetai :
Methods in Structural Linguistics
 (Chicago, The Univ. of Chicago Press, 1951)
- [Harris 1952] HARRIS, Zellig Sabbetai :
 "Discourse Analysis"
Language, 28 (1952), 1-30
 → Traduit par Françoise DUBOIS-CHARLIER :
 "Analyse du discours"
Langages, 13 (mars 1969), 8-45
- [Harris 1954] HARRIS, Zellig Sabbetai :
 "Distributional Structure"
 in : MARTINET, André ; WEINREICH, Uriels (Eds) :
Linguistics Today
 (New-York, Linguistic Circle of New-York, 1954),
 26-42.
- [Harris 1955] HARRIS, Zellig Sabbetai :
 "From Phoneme to Morpheme"
Language, 31 (1955), 2, 190-222
 → réédité dans [Harris 1970 : 32-67]
- [Harris 1957] HARRIS, Zellig Sabbetai :
 "Co-occurrence and Transformation in
 Linguistic Structure"
Language, 33 (1957), 3, 283-340
 → réédité dans [Harris 1970 : 390-457]
- [Harris 1964] HARRIS, Zellig Sabbetai :
 "The Elementary Transformations"
Transformations and Discourse Analysis Papers,
 54 (1964)
 → réédité dans [Harris 1970 : 482-532]
- [Harris 1965] HARRIS, Zellig Sabbetai :
 "Transformational Theory"
Language, 41 (1965), 3, 363-401
 → réédité dans [Harris 1970 : 532-577]

- [Harris 1966] HARRIS, Zellig Sabbetai :
 "Algebraic Operations in Linguistic Structure",
 Communication au Congrès International de
 Mathématiques, Moscou, 1966
 → réédité dans [Harris 1970 : 603-611]
- [Harris 1968] HARRIS, Zellig Sabbetai :
Mathematical Structures of Language
 (New-York, Wiley, 1968)
 → Traduit par Catherine FUCHS :
Structures mathématiques du langage
 (Paris, Dunod, 1971)
- [Harris 1970] HARRIS, Zellig Sebbetai :
*Papers in Structural and Transformational
 Linguistics*
 (Dordrecht, D. Reidel, 1970)
- Harris 1976 HARRIS, Zellig Sabbetai :
Notes du cours de syntaxe
 Traduit par Maurice GROSS
 (Paris, Seuil, 1976)
- [Heim 1977] HEIM, Kathleen M. :
 Compte-rendu de :
 LUCCI, York ; ROKKAN, Stein ; MEYERHOFF, Eric :
A Library Center of Survey Research Data
 (New-York, Columbia University School of
 Library Service, 1957)
 in : *IASSIST Newsletter*, 1, 3 (mai 1977), 25-29
- [Hempel 1966] HEMPEL, Carl G. :
Philosophy of Natural Science
 (Englewood Cliffs, Prentice Hall, 1966)
 → Traduit par Bertrand SAINT-SERNIN :
Éléments d'épistémologie
 (Paris, Colin, 1972)

- [Henrysson 1971] HENRYSSON, Sten :
 "Gathering, Analyzing, and Using Data on
 Test Items"
 in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
 (Washington, American Council on Education,
 1971), 130-159
- [Hiz 1962] HIŻ, Henry :
 "Questions and Answers"
The Journal of Philosophy, 59, 10 (mai 1962),
 253-265
- [Holm 1974a] HOLM, Kurt :
 "Theorie der Frage"
*Kölner Zeitschrift für Soziologie und
 Sozialpsychologie*, 26 (1974), 1, 91-114
- [Holm 1974b] HOLM, Kurt :
 "Theorie der Fragebatterie"
*Kölner Zeitschrift für Soziologie und
 Sozialpsychologie*, 26 (1974), 2, 316-341
- [Hurrell 1964] HURRELL, Paul Manson :
 "Discussion : Interrogatives, Testability, and
 Truth-value"
Philosophy of Science, 2 (1964), 173-182
- [Hyman 1972] HYMAN, Herbert H. :
*Secondary Analysis of Sample Surveys :
 Principles, Procedures, and Potentialities*
 (New-York, Wiley, 1972)
- [Ined 1950] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES :
*Le niveau intellectuel des enfants d'âge sco-
 laire. Une enquête nationale dans l'enseignement
 primaire*, présentée par le Pr. Georges HEUYER,
 le Pr. Henri PIÉRON, Mme Henri PIÉRON et Alfred
 SAUVY
 (Cahier n° 13, Paris, PUF, 1950)

- [Ined 1954] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES :
Le niveau intellectuel des enfants d'âge scolaire. La détermination des aptitudes. L'influence des facteurs constitutionnels, familiaux et sociaux
 (Cahier n° 23, Paris, PUF, 1954)
- [Ined 1969] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES :
*Enquête nationale sur le niveau intellectuel des enfants d'âge scolaire **
 (Cahier n° 54, Paris, PUF, 1969)
- [Ined 1973] INSTITUT NATIONAL D'ÉTUDES DÉMOGRAPHIQUES :
*Enquête nationale sur le niveau intellectuel des enfants d'âge scolaire ***
 (Cahier n° 64, Paris, PUF, 1973)
- [Javeau 1971] JAVEAU, Claude :
L'enquête par questionnaire. Manuel à l'usage du praticien
 (Bruxelles, Editions de l'Institut de Sociologie, 1971)
- [Jespersen 1924] JESPERSEN, Otto :
The Philosophy of Grammar
 (Londres, George Allen & Unwin, 1924)
 → Traduit par Anne-Marie LÉONARD :
La philosophie de la grammaire
 (Paris, Editions de Minuit, 1971)
- [Johnson 1836] JOHNSON, Alexander Bryan :
A Treatise on Language
 (New-York, 1836)
 → réédité par David RYNIN
 (New-York, Dover, 1968)

- [Junge 1638] JUNGE, Joachim :
Joachimi JUNGII Logica Hamburgensis
 (Hambourg, 1638)
 → réédité par Rudolf W. MEYER
 (Hambourg, J.J. Augustin, 1957)
 * avec indication de la pagination originale
- [Kalinowski 1972] KALINOWSKI, Georges :
La logique des normes
 (Paris, PUF, 1972)
- [Kammeyer 1971] KAMMEYER, Kenneth C.W. ; ROTH, Julius A. :
 "Coding Responses to Open-ended Questions"
 in COSTNER, Herbert L. :
Sociological Methodology 1971
 (San Francisco, Jossey-Bass, 1971), 60-78
- [Kanger 1972] KANGER, Stig :
 "Measurement : An Essay in Philosophy of
 Science"
Theoria, 38 (1972), 1-44
- [Kant 1781] KANT, Emmanuel :
Kritik der reinen Vernunft
 → Traduit par Jules BARNI (sur la 2e édition) :
Critique de la raison pure
 (Paris, Joseph Gibert, 1946)
- [Kant 1800] KANT, Emmanuel :
Logik
 (publication posthume par JAESCHE en 1800)
 → Traduit par Louis GUILLERMIT :
Logique
 (2e édition, Paris, Vrin, 1970)
- [Karttunen 1973] KARTTUNEN, Lauri :
 "La logique des constructions anglaises à
 complément prédicatif"
 traduit par Maurice BOREL
Langages, 30 (juin 1973), 56-80

- [Katz 1953] KATZ, Daniel :
 "Field Studies"
 in : FESTINGER, Leon ; KATZ, Daniel (Eds) :
Research Methods in the Behavioral Sciences
 (New-York, The Dryden Press, 1953)
 → Traduit par Honoré LESAGE :
 "Les études sur le terrain"
*Les méthodes de recherche dans les sciences
 sociales*
 (Paris, PUF, 1959), 68-118
- [Katz 1964] KATZ, Jerrold J. ; POSTAL, Paul M. :
An Integrated Theory of Linguistic Descriptions
 (Cambridge, Mass., M.I.T. Press, 1964)
 → Traduit par Jean-Yves POLLOCK :
Théorie globale des descriptions linguistiques
 (Paris, Mame, 1973)
- [Katz 1966] KATZ, Jerrold J. :
The Philosophy of Language
 (New-York, Harper & Row, 1966)
 → Traduit par Janick GAZIO :
La philosophie du langage
 (Paris, Payot, 1971)
- [Katz 1972] KATZ, Jerrold J. :
Semantic Theory
 (New-York, Harper and Row, 1972)
- [Katz 1977] KATZ, Jerrold J. :
Propositional Structure and Illocutionary Force
 (New-York, Crowell, 1977)
- [Keenan 1971] KEENAN, Edward Louis :
 "Two Kinds of Presupposition in Natural
 Language"
 in : FILLMORE, C.J. ; LANGENDOEN, D.T. (Eds) :
Studies in Linguistic Semantics
 (New-York, Holt Rinehart & Winton, 1971), 44-52

- [Kiefer 1973] KIEFER, Ferenc :
 "On Presuppositions"
 in : KIEFER, Ferenc ; RUWET, Nicolas (Eds) :
Generative Grammar in Europe
 (Dordrecht, Reidel, 1973), 218-242
 → Traduit par Laurent DANON-BOILEAU,
 in : *Essais de sémantique générale*,
 (Paris, Mame, 1974), ch. 3 (81-116)
- [Kiparsky 1970] KIPARSKY, Paul ; KIPARSKY, Carol :
 "Fact"
 in : BIERWISH, M. ; HEIDOLPH, K. (Eds) :
Progress in Linguistics
 (La Haye, Mouton, 1970)
 → réédité in : STEINBERG, Danny D. ; JAKOBOVITS,
 Leon A. (Eds) :
*Semantics. An Interdisciplinary Reader in Philo-
 sophy, Linguistics and Psychology*
 (Cambridge University Press, 1971), 345-369
- [Kohout 1974] KOHOUT, Frank J. :
*Statistics for Social Scientists. A coordinated
 Learning System*
 (New-York, Wiley, 1974)
- [Kubiński 1958] KUBIŃSKI, Tadeusz :
 "An Essay in Logic of Questions",
*Atti del XII Congresso Internazionale di
 Filosofia (Venezia, 12-18 Settembre 1958)*
 (Florence, Sansoni, 1960), vol.5, 315-322
- [Kubiński 1966] KUBIŃSKI, Tadeusz :
 "Przegląd niektórych zagadnień logiki pytań"
Studia Logica, 18 (1966), 105-137
- [Kubiński 1967] KUBIŃSKI, Tadeusz :
 "Some Observations about a Notion of Incomplete
 Answer"
Studia Logica, 21 (1967), 39-42

- [Kubiński 1968] KUBIŃSKI, Tadeusz :
 "The Logic of Questions"
 in : KLIBANSKY, Raymond (Ed.) :
La Philosophie Contemporaine
 (Florence, La Nuova Italia Editrice, 1968)
 Tome 1, 185-189
- [Kubiński 1970] KUBIŃSKI, Tadeusz :
Wstęp do logicznej teorii pytań
 (Varsovie, Państwowe Wydawnictwo Naukowe, 1971)
 (© 1970)
- [Lakoff 1970] LAKOFF, George :
 "Linguistics and Natural Logic"
Synthèse, 22 (1970), 1-2
 → Traduit par Judith MILNER et Joëlle SAMPY :
Linguistique et logique naturelle
 (Paris, Klincksieck, 1976)
- [Lakoff 1971 a] LAKOFF, George :
 "On Generative Semantics"
 in : STEINBERG, Danny D. ; JAKOBOVITS, Leon A.
 (Eds) :
*Semantics. An Interdisciplinary Reader in Philo-
 sophy, Linguistics and Psychology*
 (Cambridge University Press, 1971), 232-296
- [Lakoff 1971 b] LAKOFF, George :
 "Presupposition and Relative Well-Formedness"
 in : STEINBERG, Danny D. ; JAKOBOVITS, Leon A.
 (Eds) :
*Semantics. An Interdisciplinary Reader in Philo-
 sophy, Linguistics and Psychology*
 (Cambridge University Press, 1971), 329-340
- [Langacker 1972] LANGACKER, Ronald W. :
Fundamentals of Linguistic Analysis
 (New-York, Harcourt Brace Jovanovich, 1972)

- [Langendoen 1971a] LANGENDOEN, D. Terence :
 "Presupposition and Assertion in the Semantic Analysis of Nouns and Verbs in English"
 in : STEINBERG, Danny D. ; JAKOBOVITS, Leon A. (Eds) :
Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology
 (Cambridge University Press, 1971), 341-344
- [Langendoen 1971b] LANGENDOEN, D. Terence ; SAVIN, Harris B. :
 "The Projection Problem for Presuppositions"
 in : FILLMORE, C.J. ; LANGENDOEN, D.T. (Eds) :
Studies in Linguistic Semantics
 (New-York, Holt Rinehart & Winston, 1971), 54-60
- [Lazarsfeld 1935] LAZARSFELD, Paul F. :
 "The Art of Asking Why. Three Principles Underlying the Formulation of Questionnaires"
National Marketing Review, 1, 1 (été 1935)
- [Lazarsfeld 1958a] LAZARSFELD, Paul F. :
 "Evidence and Inference in Social Research"
Daedalus, 87, 4 (automne 1958), 99-130
- [Lazarsfeld 1958b] LAZARSFELD, Paul F. ; THIELENS, Wagner :
 "Deux mesures de l'éminence" (extrait de
The Academic Mind, Glencoe, Free Press, 1958)
 in : BOUDON, Raymond ; LAZARSFELD, Paul F. (Eds) :
Le vocabulaire des sciences sociales
 (Paris, Mouton, 1967), 69-73
- [Le Bidois 1968] LE BIDOIS, Georges ; LE BIDOIS, Robert :
Syntaxe du français moderne. Ses fondements historiques et psychologiques - 2 tomes
 (Paris, Picard, 1968)
- [Lécuyer 1963] LECUYER, Bernard-Pierre :
La recherche sociale empirique en France sous l'Ancien Régime
 (Séminaire d'histoire de la sociologie empirique en France, Document n° 8)
 (Paris, Ecole Pratique des Hautes Etudes, 1963)

- [Leech 1974] LEECH, Geoffrey :
Semantics
(Harmondsworth, Penguin Books, 1974)
- [Leik 1971] LEIK, Robert K. ; GOVE, Walter R. :
"Integrated Approach to Measuring Association"
in : COSTNER, Herbert L. (Ed.) :
Sociological Methodology 1971
(San Francisco, Jossey-Bass, 1971), 279-301
- [Lejeune 1937] LEJEUNE, Michel :
"Conditions générales des changements linguistiques"
in : *Encyclopédie française. Tome 1*
(Paris, Société de gestion de l'Encyclopédie Française, 1937), IIe partie, chapitre II
(fascicule 1-34)
- [Lenski 1960] LENSKI, Gerhard E. ; LEGETT, John C. :
"Caste, Class, and Deference in the Research Interview"
The American Journal of Sociology, 65, 5
(mars 1960), 463-467
- [Leonard 1959] LEONARD, Henry S. :
"Interrogatives, Imperatives, Truth, Falsity and Lies"
Philosophy of Science, 26, 3 (juillet 1959),
172-186
- [Leonard 1961] LEONARD, Henry S. :
"A Reply to Professor Wheatley"
Philosophy of Science, 28, 1 (janvier 1961),
55-64
- [Leonard 1967] LEONARD, Henry S. :
Principles of Reasoning
(New-York, Dover Publications, 1967)
* (nouvelle édition revue de : *Principles of Right Reason*, New-York, Henry Holt, 1957)

- [Lerman 1970] LERMAN, I.C. :
Les bases de la classification automatique
(Paris, Gauthier-Villars, 1970)
- [Lerner 1956] LERNER, Daniel :
"Interviewing Frenchmen"
The American Journal of Sociology, 62, 2
(septembre 1956), 187-194
- [Lévi-Strauss 1962] LEVI-STRAUSS, Claude :
La pensée sauvage
(Paris, Plon, 1962)
- [Lewis 1932] LEWIS, Clarence Irving ; LANGFORD, Cooper Harold :
Symbolic Logic
(New-York, The Century, 1932)
- [Lindsay 1963] LINDSAY, Robert K. :
"Inferential Memory as the Basis of Machines
Which Understand Natural Language"
in : FEIGENBAUM, Edward A. ; FELDMAN, Julian (Ed.):
Computers and Thought
(New-York, Mac Graw Hill, 1963), 217-233
- [Lison 1958] LISON, L. :
Statistique appliquée à la biologie expérimentale
(Paris, Gauthier-Villars, 1958)
- [Litwak 1956] LITWAK, Eugene :
"A Classification of Biased Questions"
American Journal of Sociology, 62, 2
(septembre 1956), 182-186
- [Lord 1950] LORD, Edith :
"Experimentally Induced Variations in Rorschach
performance"
Psychological Monographs, 64 (1950), 10
→ réédité dans : HIRT, Michael :
Rorschach Science. Readings in Theory and Method
(Glencoe, Free Press, 1962), 101-143

- [Loubet del Bayle 1978] LOUBET DEL BAYLE, Jean-Louis :
Introduction aux méthodes des sciences sociales
 (Toulouse, Privat, 1978)
- [Löw 1928] LÖW, Friedrich :
 "Logik der Frage"
Archiv für die Gesamte Psychologie, 66 (1928),
 3-4, 357-436
- [Lyons 1968] LYONS, John :
An Introduction to Theoretical Linguistics
 (Cambridge University Press, 1968)
 → Traduit par Françoise DUBOIS-CHARLIER et
 David ROBINSON :
*Linguistique générale. Introduction à la lin-
 guistique théorique*
 (Paris, Larousse, 1970)
- [Lyons 1978] LYONS, John :
Eléments de sémantique
 Traduit par Jacques DURAND et Eliane KOSKAS
 (Paris, Larousse, 1978)
- [Maas 1972] MAAS, Utz :
 "Ein Problem der Fragelogik : Sind zurückgewie-
 sene Präsuppositionen Antworten ?"
Linguistische Berichte, 19 (1972), 69-73
- [Mac Kay 1950] Mac KAY, Donald M. :
 "Quantal Aspects of Scientific Information"
Philosophical Magazine, 41, 314 (mars 1950),
 289-311
- [Mac Kay 1961] Mac KAY, Donald M. :
 "The Informational Analysis of Questions and
 Commands"
 in : CHERRY, Colin (Ed.) :
Information Theory : Fourth London Symposium
 (Londres, Butterworth's, 1961)
 → réédité in : BUCKLEY, Walter (Ed.) :
*Modern Systems Research for the Behavioral
 Scientist*
 (Chicago, Aldine, 1968), 204-208

- [Mac Kay 1969] Mac KAY, Donald M. :
Information, Mechanism and Meaning
(Cambridge, Mass., M.I.T. Press, 1969)
- [Maingueneau 1976] MAINGUENEAU, Dominique :
Initiation aux méthodes de l'analyse du discours. Problèmes et perspectives
(Paris, Hachette, 1976)
- [Maitre 1972] MAITRE, Jacques :
Sociologie religieuse et méthodes mathématiques
(Paris, PUF, 1972)
- [Makovelski 1967] МАКОВЕЛЬСКИЙ, Александр Осипович :
История Логики (Москва, Наука, 1967)
→ traduit par Geneviève DUPOND :
MAKOVELSKI, Alexandre :
Histoire de la logique
(Moscou, Editions du Progrès, 1978)
- [Mallet 1845] MALLET, Charles Auguste :
Histoire de l'Ecole de Mégare et des Ecoles d'Elis et d'Erétrie
(Paris, V^e Maire-Nyon, 1845)
- [Martinotti 1978] MARTINOTTI, Guido :
"Data Processing, Government and the Public :
Reflections on the Italian Case"
International Social Sciences Journal, 30
(1978), 1, 146-163
- [Materna 1977] MATERNA, Pavel :
Compte-rendu de [Kubiński 1970]
in : *Journal of Symbolic Logic*, 42 (1977),
426-428
- [Mauger 1968] MAUGER, G. :
Grammaire pratique du français d'aujourd'hui.
Langue parlée. Langue écrite
8e édition revue (Paris, Hachette, 1968)

- [Mayo 1956] MAYO, Bernard :
 "Deliberative Questions : a Criticism"
Analysis, 16, 3 (janvier 1956), 58-63
- [Menefee 1936] MENEFEE, Selden C. :
 "The Effect of Stereotyped Words on Political
 Judgments"
American Sociological Review, 1 (1936), 614-621
- [Miller 1963] MILLER, George A. :
 "What is Information Measurement ?"
American Psychologist, 8 (1963), 3-11
 → réédité in : Buckley, Walter (Ed.) :
*Modern System Research for the Behavioral
 Scientist*
 (Chicago, Adeline, 1968), 123-127
- [Miller 1968] MILLER, Daniel R. :
 "Prediction of Behavior by Means of the
 Rorschach Test"
 in : MILLON, Theodore (Ed.) :
Approaches to Personality
 (New-York, Pitman, 1968), 475-487
- [Miller 1970] MILLER, Robert J. ; ROBERTS, Noreen J. :
 "Working with Complex Data Files : I.
 Development and Analysis of Complex Data Files
 in a Regional Transportation Study"
 in : BISCO, Ralph L. (Ed.) :
Data Bases, Computers, and the Social Sciences
 (New York, Wiley, 1970, 83-111
- [Milner 1973] MILNER, Jean-Claude :
Arguments linguistiques
 (Paris, Mame, 1973)
- [Mirambel 1969] MIRAMBEL, André :
Grammaire du grec moderne
 (Paris, PUF, 1969)

- [Moreau 1975] MOREAU, René :
Introduction à la théorie des langages
(Paris, Hachette, 1975)
- [Morgan 1938] MORGAN, Christiana D. ; MURRAY, Henry A. :
"Thematic Apperception Test"
in : MURRAY, Henri A. (Ed.) :
Explorations in Personality
(New-York, Oxford University Press, 1938),
530-545
- [Morice 1954] MORICE, E. ; CHARTIER, F. :
*Méthode statistique. 2e partie : Analyse
statistique*
(Paris, Imprimerie Nationale, 1954)
- [Moritz 1940] MORITZ, Manfred :
"Zur Logik der Frage"
Theoria, 6 (1940), 123-149
- [Moser 1958] MOSER, C.A. :
Survey Methods in social investigations
(Londres, Heinemann, 1958)
- [Mounin 1963] MOUNIN, Georges :
Les problèmes théoriques de la traduction
(Paris, Gallimard, 1963)
- [Mounin 1972] MOUNIN, Georges :
La linguistique du XXe siècle
(Paris, PUF, 1972)
- [Mucchielli 1967] MUCCHIELLI, Roger (Ed.) :
Le questionnaire dans l'enquête psycho-sociale
(Librairies Techniques/Editions Sociales Fran-
çaises, 1967)

- [Muscio 1916] MUSCIO, Bernard :
"The Influence of the Form of a Question"
British Journal of Psychology, 8, 3
(septembre 1916), 351-389
- [Nahoum] NAHOUM, Charles :
L'entretien psychologique
(Paris, PUF, 1958)
- [Nasatir 1973] NASATIR, David :
*Data Archives for the Social Sciences :
Purposes, Operations and Problems*
(Paris, UNESCO, 1973)
- [Noelle-Neumann 1970] NOELLE-NEUMANN, Elisabeth :
"Wanted : Rules for Wording Structured
Questionnaires"
Public Opinion Quaterly, 34 (1970), 2, 191-201
- [Oppenheim 1966] OPPENHEIM, A.N. :
Questionnaire Design and Attitude Measurement
(Londres, Heinemann, 1966)
- [Osgood 1957] OSGOOD, Charles E. ; SUCI, George J. ;
TANNENBAUM, Percy H. :
The Measurement of Meaning
(Urbana, University of Illinois Press, 1957)
- [Osgood 1959] OSGOOD, Charles E. :
"The Representational Model and Relevant
Research Methods"
in : POOL, Ithiel de Sola :
Trends in Content Analysis
(Urbana, University of Illinois Press, 1959),
33-88

- [Panfilov 1963] ПАНФИЛОВ, В.З.:
*Грамматика и Логика. Грамматическое и
 логико-грамматическое членение простого
 предложения* (Москва, изд. Академии Наук, 1963)
 → translated by H.A. VLADIMIRSKY :
 PANFILOV, V.Z. :
Grammar and Logic
 (La Haye, Mouton, 1968)
- [Patchen 1965] PATCHEN, Martin :
*Some Questionnaire Measures of Employee
 Motivation and Morale*
 (Ann Arbor, Institute for Social Research, 1965)
- [Pawłowski 1969] PAWŁOWSKI, Tadeusz :
 "Theory of Questions and Its Applications in
 the Social Sciences"
The Polish Sociological Bulletin, 2 (1969),
 95-109
- [Payne 1946] PAYNE, Stanley L. :
 "Some Opinion Research Principles Developed
 Through Studies of Social Medicine"
Public Opinion Quarterly, 10, 1 (printemps 1946),
 93-98
- [Payne 1951] PAYNE, Stanley L. :
The Art of Asking Questions
 (Princeton, N.J., Princeton University Press,
 1951)
- [Pérard 1947] PÉRARD, Albert ; TERRIEN, Jean :
Les mesures physiques
 (Paris, PUF, 1947)
- [Philipp 1974] PHILIPP, Marthe :
Grammaire de l'allemand
 (Paris, PUF, 1974)

- [Phillips 1966] PHILLIPS, Bernard S. :
Social Research. Strategy and Tactics
(New-York, Macmillan, 1966)
- [Phillips 1970] PHILLIPS, Derek L. ; CLANCY, Kevin J. :
"Response Biases in Field Studies of Mental
Illness"
American Sociological Review, 35, 3 (juin
1970), 503-515
- [Phillipson 1968] PHILLIPSON, Michael :
"The Classification, Storage and Retrieval
of Survey Data"
Extrait de : *Making Fuller Use of Survey Data*
(Londres, Political and Economic Planning
Report, 1968), 38-52
→ réédité in :
WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
(Englewood, Information Handling Services, 1977),
211-223
- [Piaget 1967] PIAGET, Jean :
"Les deux problèmes principaux de l'épistémologie
des sciences de l'homme"
in : PIAGET, Jean (Ed.) :
Logique et connaissance scientifique
(Encyclopédie de la Pléiade, XXII, Paris-
Gallimard, 1967), 1114-1150
- [Piaget 1970] PIAGET, Jean :
Epistémologie des sciences de l'homme
(Paris, Gallimard, 1970)
- [Picard 1972] PICARD, Claude-François :
Graphes et questionnaires. 2 tomes
(Paris, Gauthier-Villars, 1972)
- [Pichot 1954] PICHOT, Pierre :
Les tests mentaux
(Paris, PUF, 1954)

- [Piéron 1951] PIÉRON, Henri :
 "Méthodes expérimentales. Psychométrie"
 in : PIÉRON, Henri (sous la direction de) :
Traité de psychologie appliquée,
Livre II : Méthodologie psychotechnique
 (Paris, PUF, 1951), 123-241
- [Pottier 1972] POTTIER, Bernard :
Grammaire de l'espagnol
 (Paris, PUF, 1972)
- [Presser 1975] PRESSER, Stanley ; SCHUMAN, Howard :
 "Question Wording as an Independent Variable
 in Survey Analysis : a First Report"
Proceedings of the Social Statistics Section
 (American Statistical Association, 1975), 16-25
- [Prieto 1966] PRIETO, Luis J. :
Messages et signaux
 (Paris, PUF, 1966)
- [Prior 1955] PRIOR, Mary ; PRIOR, Arthur :
 "Erotetic Logic"
The Philosophical Review, 64 (1955), 43-59
- [Quine 1961] QUINE, Willard Van Orman :
From a Logical Point of View
 Second Edition, Revised
 (New-York, Harper & Row, 1961)
- [Quine 1965] QUINE, Willard Van Orman :
Elementary Logic
 (New-York, Harper & Row, 1965)
 → traduit par Jean LARGEAULT et Bertrand
 SAINT-SERNIN :
Logique élémentaire
 (Paris, A. Colin, 1972)

- [Ragon 1889] RAGON, E. :
Grammaire grecque - 30e édition
(Paris, J. de Gigord, 1938)
- [Régnier 1974] RÉGNIER, André :
La crise du langage scientifique
(Paris, Anthropos, 1974)
- [Remmers 1955] REMMERS, H.H. ; GAGE, N.L. :
Educational Measurement and Evaluation
Revised edition
(New-York, Harper, 1955)
- [Rescher 1955] RESCHER, Nicholas :
Compte-rendu de [Prior 1955]
in : *Journal of Symbolic*, 20 (1955), 302
- [Rescher 1967] RESCHER, Nicholas :
"Avicenna on the Logic of Questions"
Archiv für Geschichte Philosophie, 49 (1967),
1, 1-6
- [Resher 1974] RESHER, Nicholas
"Logic, Applied"
in : *Encyclopaedia Britannica*
(Chicago, W. Beaton, 1974),
Macropaedia, vol. 11
- [Reuchlin 1957] REUHLIN, Maurice :
Histoire de la psychologie
(Paris, PUF, 1957)
- [Reuchlin 1962] REUHLIN, Maurice :
Les méthodes quantitatives en psychologie
(Paris, PUF, 1962)

- [Reuchlin 1976] REUCHLIN, Maurice :
Précis de statistique
(Paris, PUF, 1976)
- [Ribot 1904] RIBOT, Théodule :
"Sur la valeur des questionnaires en psychologie"
Journal de Psychologie Normale et Pathologique
1, 1 (janvier-février 1904), 1-10
- [Robinson 1973] ROBINSON, John P. ; SHAVER, Philipp R. :
Measures of Social Psychological Attitudes
(Ann Arbor, Institute for Social Research, 1973)
- [Roper 1940] ROPER, Elmo :
"Wording Questions for the Polls"
Public Opinion Quarterly, 4, 1 (mars 1940),
129-130
- [Rorschach 1921] RORSCHACH, Hermann :
Psychodiagnostik
(Berne, Hans Huber, 1921)
→ traduit d'après la 4e édition (1940)
par OMBREDANE, A. :
Psychodiagnostic, 3^e édition révisée et corrigée
(Paris, PUF, 1962)
- [Rotter 1951] ROTTER, Julian B. :
"Word Association and Sentence Completion
Methods"
in : ANDERSON, Harold H. ; ANDERSON, Gladys L. :
*An Introduction to Projective Techniques and
Other Devices for Understanding the Dynamics
of Human Behavior*
(Englewood Cliffs, Prentice-Hall, 1951), 279-311
- [Rouanet 1977] ROUANET, Henri ; LÉPINE, Dominique :
*Introduction à l'analyse des comparaisons pour
le traitement des données expérimentales.*
Numéro spécial de *Informatique et Sciences
Humaines*, 33-34 (juin 1977)

[Rugg 1941]

RUGG, Donald :
 "Experiments in Wording Questions : II"
Public Opinion Quaterly, 5, 1 (mars 1941), 91-92

[Russel 1905a]

RUSSEL, Bertrand :
 "The Existential Import of Propositions"
Mind, 14 (juillet 1905), 398-401
 → réédité dans :
Essays in Analysis, choix de textes de RUSSEL
 présenté par Douglas LACKEY
 (New-York, George Braziller, 1973), 98-102

[Russel 1905b]

RUSSEL, Bertrand :
 "On denoting"
Mind, 14 (octobre 1905), 479-493
 → réédité dans :
Logic and Knowledge, choix de textes de RUSSEL
 présenté par Robert C. MARSH,
 (New-York, Capricorn Books, 1971), 41-56

[Russel 1940]

RUSSEL, Bertrand :
An Inquiry into Meaning and Truth
 (Londres, Allen & Unwin, 1940)
 → traduit par Philippe DEVAUX :
Signification et vérité
 (Paris, Flammarion, 1969)

[Russel 1957]

RUSSEL, Bertrand :
 "Mr Strawson on Referring"
Mind, 66 (juillet 1957), 385-389
 → réédité dans :
Essays in Analysis, choix de textes de RUSSEL
 présenté par Douglas LACKEY
 (New-York, George Braziller, 1973), 120-126

- [Ruwet 1967] RUWET, Nicolas
Introduction à la grammaire générative
 (Paris, Plon, 1967)
- [Ruwet 1972] RUWET, Nicolas :
Théorie syntaxique et syntaxe du français
 (Paris, Seuil, 1972)
- [Ryser 1963] RYSER, H.J. :
Combinatorial Mathematics
 (New-York, Wiley, 1963)
 → traduit par P. CAMION :
Mathématiques combinatoires
 (Paris, Dunod, 1969)
- [Sadock 1974] SADOCK, Jerrold M. :
Towards a Linguistic Theory of Speech Acts
 (New-York, Academic Press, 1974)
- [Sapir 1944] SAPIR, Edward :
 "Grading, a Study on Semantics"
Philosophy of Science, 11 (1944), 2, 93-116
- [Saussure 1915] SAUSSURE, Ferdinand De :
Cours de linguistique générale
 publié par C. BALLY et A. SECHEHAYE, avec la
 collaboration de A. RIEDLINGER
 (Nouvelle édition (avec la pagination de
 l'édition originale), Paris, Payot, 1972)
- [Schatzmann 1955] SCHATZMANN, Leonard ; STRAUSS, Anselm :
 "Social Class and Modes of Communication"
American Journal of Sociology, 60, 4 (janvier
 1955), 329-338

- [Schlick 1935] SCHLICK, Moritz :
 "Unanswerable Questions ?"
The Philosopher, 13 (1935), 98-104
- [Schuman 1974] SCHUMAN, Howard ; DUNCAN, Otis Dudley :
 "Questions about Attitude Survey Questions"
 in : COSTNER, Herbert L. (Ed.) :
Sociological Methodology 1973-1974
 (San Francisco, Jossey-Bass, 1974), 232-251
- [Scott 1968] SCOTT, William A. :
 "Attitude Measurement"
 in : LINDZEY, Gardner ; ARONSON, Elliot :
The Handbook of Social Psychology - 2e édition
 (Reading, Addison-Wesley, 1968), vol. 2,
 204-273
- [Searle 1965] SEARLE, John R. :
 "What is a Speech Act ?"
 in : BLACK, Max :
Philosophy in America
 (Allen & Unwin, 1965), 221-239
 → réédité in : SEARLE, John R. (Ed.) :
The Philosophy of Language
 (Oxford University Press, 1971), 39-53
- [Searle 1969] SEARLE, John R. :
Speech Acts. An Essay in the Philosophy of Language
 (Cambridge University Press, 1969)
- [Semerari 1968] SEMERARI, Giuseppe :
 "Del Domandare"
Il Problema della Domanda
Archivio di Filosofia
 (Padova, CEDAM, 1968), 27-40
- [Sextus Empiricus AM] SEXTUS EMPIRICUS :
Adversus Mathematicos
 Texte établi par Hermann MUTSCHMANN et J. MAU
Sexti Empirici : Opera : vol. III
 (Leipzig, Teubner, 1954)

- [Sextus Empiricus HP] SEXTUS EMPIRICUS :
Πυρρωνείων Ἐπιτομῶσεων
Texte établi par Hermann MUTSCHMANN et J. MAU
Sexti Empirici Opera : Vol. I
(Leipzig, Teubner, 1958)
- [Siegel 1956] SIEGEL, Sidney :
Nonparametric Statistics for the Behavioral Sciences
(New-York, Mac Graw-Hill, 1956)
- [Sjoberg 1954] SJOBERG, Gideon :
"A Questionnaire on Questionnaires"
Public Opinion Quaterly, 18 (1954), 423-427
- [Sokolowski 1971] SOKOLOWSKI, Robert :
"Scientific and Hermeneutic Questions
in Aristotle"
Philosophy and Rhetoric, 4 (1971), 4, 242-261
- [Sperantia 1935] SPERANTIA, Eugeniu :
"Remarques sur les propositions interrogatives.
Projet d'une «Logique du Problème»"
*Actes du Congrès International de Philosophie
Scientifique (Paris, 1935), Fascicule VII*
(Paris, Herman, 1936), 18-28
- [Sperantia 1937] SPERANTIA, Eugeniu :
"La métaphysique implicite dans les postulats
de toute pensée possible"
*Travaux du IXe Congrès international de Philo-
sophie, VIII : Analyse réflexive et transcendante*
(Paris, Herman, 1937), 80-85
- [Src 1976] SURVEY RESEARCH CENTER :
Interviewer's Manual
Edition révisée
(Ann Arbor, Institute for Social Research, 1976)

- [Stahl 1956] STAHL, Gerold :
 "La logica de las preguntas"
Anales de la Universidad de Chile,
 n° 102 (2e trimestre 1956), 71-75
- [Stahl 1960] STAHL, Gerold :
 "Un développement de la logique des questions"
 (Conférence prononcée le 25 novembre 1960 à
 l'Institut d'Histoire des Sciences et des Tech-
 niques de l'Université de Paris)
*Revue Philosophique de la France et de
 l'étranger*, 153 (1963), 293-301
- [Stahl 1969] STAHL, Gerold :
 "The Effectivity of Questions"
Noûs, 3, 2 (mai 1969), 211-218
- [Stanley 1971] STANLEY, Julian C. :
 "Reliability"
 in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
 (Washington, American Council on Education,
 1971), 356-442
- [Steinman 1959] STEINMAN, Martin, Jr :
 "Questions and Answers"
Graduate Review of Philosophy, 1 (1959), 17-28
- [Stenius 1969] STENIUS, Erik :
 "Mood and Language-Game"
 in DAVIS, J.W. ; HOCKNEY, D.J. ; WILSON, W.K.
 (Eds) ;
Philosophical Logic
 (Dordrecht, D. Reidel, 1969), 251-271
- [Stevens 1940] STEVENS, S.S. ; VOLKMAN, J. :
 "The Quantum of Sensory Discrimination"
Science, 92, 2399 (20 décembre 1940), 583-585

- [Stevens 1946] STEVENS, S.S. :
"On the Theory of Scales of Measurement"
Science, 103, 2684 (7 juin 1946), 677-680
- [Stevens 1951] STEVENS, S.S. :
"Mathematics, Measurement, and Psychophysics"
in : STEVENS, S.S. (Ed.) :
Handbook of Experimental Psychology
(New-York, Wiley, 1971), 1-49
- [Stewart 1967] STEWART, Douglas :
"Social Implications of Social Science Data
Archives"
(System Development Corporation, 1967)
→ réédité in : WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
(Englewood, Information Handling Services, 1977)
13-24
- [Stoetzel 1973] STOETZEL, Jean ; GIRARD, Alain :
Les sondages d'opinion publique
(Paris, PUF, 1973)
- [Stouffer 1950] STOUFFER, Samuel A. ; GUTTMAN, Louis ; SUCHMAN,
Edward A. ; LAZARSFELD, Paul F. ; STAR, Shirley
A. ; CLAUSEN, John A. :
Measurement and Prediction
(Princeton University Press, 1950)
→ réédition (Gloucester, Peter Smith, 1973)
- [Strawson 1950] STRAWSON, P.F. :
"On Referring"
Mind, 59 (1950), 320-344
→ réédité dans :
Logico-Linguistic Papers
(Londres, Methuen, 1971), 1-27

- [Strawson 1964] STRAWSON, P.F. :
 "Identifying Reference and Truth-values"
Theoria, 30 (1964), 96-118
 → réédité dans :
Logico-Linguistic Papers
 (Londres, Methuen, 1971), 75-95
- [Sudman 1974] SUDMAN, Seymour ; BRADBURN, Norman M. :
Response Effects in Surveys. A Review and Synthesis
 (Chicago, Aldine, 1974)
- [Suppes 1963] SUPPES, Patrick ; ZINNES, Joseph L. :
 "Basic Measurement Theory"
 in : LUCE, R. Duncan ; BUSH, Robert R. ;
 GALANTER, Eugene (Ed.) :
Mathematical Psychology
 (New-York, Wiley, 1963), Tome 1, 1-76
- [Tarski 1941] TARSKI, Alfred :
*Introduction to Logic and to the Methodology
 of Deductive Sciences*
 (New-York, Oxford University Press, 1941)
 → traduit par Jacques TREMBLAY :
Introduction à la logique
 (Paris, Gauthier-Villars, 1971)
- [Teller 1974] TELLER, Paul :
 "On Why-Questions"
Noûs, 8, 4 (novembre 1974), 371-380
- [Tellier 1974] TELLIER, André :
Grammaire de l'anglais
 (Paris, PUF, 1974)
- [Terman 1937] TERMAN, Lewis M. ; MERRILL, Maud A. :
Measuring Intelligence
 (Boston, Houghton Mifflin, 1937)

- [Tesnière 1959] TESNIÈRE, Lucien :
Eléments de syntaxe structurale
(Paris, Klincksieck, 1959)
- [Thomson 1960] THOMSON, A.J. ; MARTINET, A.V. :
A Practical English Grammar
(Londres, Oxford University Press, 1960)
- [Thornike 1971] THORNIKE, Robert L. :
"Reproducing the Test"
in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
(Washington, American Council on Education,
1971), 160-187
- [Tinkelman 1971] TINKELMAN, Sherman N. :
"Planning the Objective Test"
in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
(Washington, American Council on Education,
1971), 46-80
- [Torgerson 1958] TORGERSON, Warren S. :
Theory and Methods of Scaling
(New-York, Wiley, 1958)
- [Ullmo 1969] ULLMO, Jean
La pensée scientifique moderne
(Paris, Flammarion, 1969)
- [Van Fraassen 1968] VAN FRASSEN, Bas C. :
"Presupposition, implication and self-reference"
Journal of Philosophy, 65, 5 (1968), 136-152
- [Vax 1970] VAX, Louis
L'empirisme logique
(Paris, PUF, 1970)

- [Vendler 1967] VENDLER, Zeno :
 "Singular Terms"
 extrait de : *Linguistics in Philosophy*
 (Cornell University Press, 1967), 33-69
 → réédité in : STEINBERG, Danny D. ;
 JAKOBOVITS, Leon A. (Eds) :
*Semantics. An Interdisciplinary Reader in
 Philosophy, Linguistics and Psychology*
 (Cambridge University Press, 1971), 115-133
- [Vendryes 1923] VENDRYES, Joseph :
*Le langage. Introduction linguistique à
 l'histoire*
 (Paris, La Renaissance du livre, 1923)
 → réédition (Paris, Albin Michel, 1968)
- [Veyrenc 1968] VEYRENC, Charles-Jacques :
Grammaire du russe
 (Paris, PUF, 1968)
- [Wagner 1962] WAGNER, R.L. ; PINCHON, J. :
Grammaire du français classique et moderne
 2e édition revue et corrigée
 (Paris, Hachette, 1962)
- [Waismann 1965] WAISMANN, F. :
The Principles of Linguistic Philosophy
 (Londres, Macmillan, 1965)
- [Wandruszka 1970] WANDRUSZKA, Mario :
 "Réflexions sur la polymorphie de l'interro-
 gation française"
Revue de Linguistique Romane, 34 (1970),
 n° 133-134, 65-77
- [Warusfeld 1966] WARUSFELD, André :
Dictionnaire raisonné de mathématiques
 (Paris, Seuil, 1966)

- [Wechsler 1944] WECHSLER, David :
The Measurement of Intelligence
 3e édition augmentée
 (Baltimore, Williams & Wilkins, 1944)
- [Weiser 1975] WEISER, Ann :
 "How to Not Answer a Question :
 Purposive Devices in Conversational Strategy"
 in : *Papers from the Eleventh Regional Meeting
 of the Chicago Linguistic Society*
 (University of Chicago, 1975), 649-660
- [Wesman 1971] WESMAN, Alexander G. :
 "Writing the Test Item"
 in : THORNDIKE, Robert L. (Ed.) :
Educational Measurement
 (Washington, American Council on Education,
 1971), 81-129
- [Whately 1848] WHATELY, Richard :
Elements of Logic
 9e édition révisée
 (Londres, John W. Parker, 1848)
- [Whately 1857] WHATELY, Richard :
Elements of Rhetoric
 réédition d'après la 7e édition
 (Londres, John W. Parker, 1857)
- [Wheatley 1955] WHEATLEY, J.M.O. :
 "Deliberative questions"
Analysis, 15, 3 (janvier 1955), 49-60
- [Wheatley 1961] WHEATLEY, J.M.O. :
 "Note on Professor Leonard's Analysis of
 Interrogatives, etc."
Philosophy of Science, 28, 1 (janvier 1961),
 52-54

- [White 1977] WHITE, Howard D. :
 "Libraries and Access to Social Science Data"
 in : WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
 (Englewood, Information Handling Services, 1977),
 175-194
- [Whorf 1956] WHORF, Benjamin Lee :
Language, Thought and Reality
 (Cambridge, MIT Press, 1956)
 → traduit par Claude CARME :
Linguistique et anthropologie
 (Paris, Denoël/Gonthier, 1969)
- [Wilcox 1967] WILCOX, Allen R. ; BOBROW, Davis B. ;
 BWY, Douglas P. :
 "System SESAR : Automating an Intermediate
 Stage of Survey Research"
American Behavioral Scientist, 10, 5 (janvier
 1967), 8-11
 → réédité in : WHITE, Howard D. (Ed.) :
Reader in Machine-Readable Social Data
 (Englewood, Information Handling Services,
 1977), 224-231
- [Wittenborn 1949] WITTENBORN, J.R. :
 "Statistical Tests of Certain Rorschach
 Assumptions. Analyses of Discrete Responses"
Journal of Consulting Psychology, 13 (1949),
 257-267
 → réédité dans : HIRT, Michael (Ed.) :
Rorschach Science. Readings in Theory and Methods
 (Glencoe, Free Press, 1962), 147-162
- [Wittgenstein 1921] WITTGENSTEIN, Ludwig :
Logisch-Philosophische Abhandlung
 Annalen der Naturphilosophie, 1921
 → traduit par Pierre KLOSSOWSKI :
Tractatus Logico-philosophicus
 (Paris, Gallimard, 1961)

- [Wittgenstein 1953] WITTGENSTEIN, Ludwig :
Philosophical Investigations
 Texte allemand, et traduction anglaise de
 G.E.M. ANSCOMBE
 (Oxford, Basil Blackwell, 1953)
- [Woodworth 1945] WOODWORTH, Robert S. :
Experimental Psychology
 4e édition
 (New-York, Holt, 1945)
 → traduit par A. OMBREDANE et I. LÉZINE :
Psychologie expérimentale
 (Paris, PUF, 1949)
- [Wright 1972] WRIGHT, Patricia :
 "Some Observations on how People Answer
 Questions About Sentences"
*Journal of Verbal Learning and Verbal
 Behavior*, 11 (1972), 188-195
- [Yaglom 1957] ЯГЛОМ, А.М. ; ЯГЛОМ, И.М. :
Вероятность и информация
 (Москва, 1957)
 → traduit par Wladimir MERCOUROFF :
Probabilité et Information
 (Paris, Dunod, 1959)
- [Zazzo 1946] ZAZZO, René :
Intelligence et quotient d'âge
 (Paris, PUF, 1946)
- [Zazzo 1949] ZAZZO, René :
 "Rectification expérimentale d'un test
 Binet-Simon"
Enfance, 2, 4 (septembre 1949), 366
- [Zazzo 1951] ZAZZO, René :
 "Introduction à la méthode des tests"
 in : *Psychiatrie sociale de l'enfant*
 (Paris, Centre International de l'Enfance,
 1951), 447-455

- [Zazzo 1966] ZAZZO, René ; GILLY, Michel ; VERBA-RAD, Mina :
Nouvelle échelle métrique de l'intelligence
2 volumes
(Paris, Colin, 1966)
- [Zeisel 1968] ZEISEL, Hans :
Say it with figures
5e édition révisée
(New-York, Harper & Row, 1968)
- [Zwicky 1971] ZWICKY, Arnold M. :
"On reported Speech"
in : FILLMORE, C.J. ; LANGENDOEN, D.T. (Eds) :
Studies in Linguistic Semantics
(New-York, Holt Rinehart & Winston, 1971),
72-77
- [Zuber 1972a] ZUBER, Ryszard :
"A propos de la question dite générale"
Dialectica, 26 (1972), 2, 131-137
- [Zuber 1972b] ZUBER, Ryszard :
Structure prépositionnelle du langage
(Association Jean Favard, Documents de linguistique quantitative n° 17, 1972 ; distribué par Dunod).
- [Zulliger 1953] ZULLIGER, Hans :
Der Diapositiv-Z-Test
(Berne, Hans Huber, 1953)
→ traduit par P. PAILLET et F. SALOMON :
Le test Z collectif, 2e édition revue et augmentée
(Berne, Hans Huber, 1957)



TABLE DES MATIÈRES

INTRODUCTION

	1	
<u>0.1. LE QUESTIONNAIRE DANS LES SCIENCES DE L'HOMME</u>	5	
a) Le questionnaire dans l'enquête sociologique	5	
b) Le questionnaire en psychologie appliquée	13	
c) Questionnaire d'enquête et test psychologique	25	
d) La place du questionnaire dans la recherche	34	
<u>0.2. L'ANALYSE SECONDAIRE ET LES BANQUES DE DONNÉES</u>	39	
a) Les banques de données pour les sciences sociales	40	
b) Analyse secondaire et théorie du questionnaire	47	
<u>0.3. OBJECTIFS DE LA RECHERCHE ENTREPRISE</u>	55	
<table border="1" style="margin-left: 0;"><tr><td style="padding: 2px;">1. QU'EST-CE QU'UNE QUESTION ?</td></tr></table>	1. QU'EST-CE QU'UNE QUESTION ?	61
1. QU'EST-CE QU'UNE QUESTION ?		
<u>1.1. POLYMORPHIE DES ÉNONCÉS EXPRIMANT UNE QUESTION</u>	63	
<u>1.1.1. Fausses interrogatives, et vraies questions sans interrogation</u>	64	
a) Modalités de la phrase et fonctions de l'énoncé	64	
b) Fausses questions	67	
c) Injonctives à valeur de question	69	
d) Énonciatives à valeur de question	72	
e) Pseudo-interrogatives à valeur de question	77	

<u>1.1.2. Interrogatives directes</u>	78
a) Marques morphologiques de l'interrogation directe	79
b) Portée de l'interrogation	83
c) Enoncé simple ou disjonctif	91
<u>1.1.3. Interrogatives indirectes</u>	95
a) Marques morphologiques de l'interrogation indirecte	95
b) Modalité de la phrase exprimant l'interrogation	99
<u>1.1.4. Autres caractères morpho-syntaxiques</u>	101
a) Propositions subordonnées circonstancielles	101
b) Mode et temps du verbe de l'interrogation	103
c) Désignation du locuteur et de l'interlocuteur	105
<u>1.2. RÉDUCTION DE L'INTERROGATION A UNE FORME CANONIQUE</u>	119
<u>1.2.1. La syntaxe structurale de Lucien TESNIÈRE</u>	120
a) La phrase verbale énonciative simple	121
b) L'interrogation directe	127
c) L'interrogation indirecte et la phrase complexe	131
d) Vue d'ensemble de l'interrogation selon TESNIÈRE	135
e) Bilan de l'apport de TESNIÈRE à l'analyse de l'interrogation	143
<u>1.2.2. Les grammaires génératives</u>	145
a) La perspective générative en linguistique	146
b) La syntaxe du français de DUBOIS et DUBOIS-CHARLIER	151
c) L'interrogation selon DUBOIS et DUBOIS-CHARLIER	160
d) Apport des modèles génératifs à l'analyse de l'interrogation	170

<u>1.2.3. Les grammaires transformationnelles</u>	174
a) La démarche de Z.S. HARRIS	174
b) Les éléments de la phrase : opérateurs et arguments	177
c) Les transformations élémentaires	180
d) L'interrogation selon HARRIS	185
e) Apport des grammaires transformationnelles à l'analyse de l'interrogation	190
<u>1.2.4. Le substrat commun aux énoncés à valeur de question</u>	195
a) La notion de transformation et ses implications	196
b) Le problème de la portée de l'interrogation	201
c) Le problème des relations entre questions et réponses	203
d) Description formelle des grandes catégories de questions	207
e) La forme canonique des énoncés à valeur de question	223
2. QUESTIONS ET RÉPONSES	239
<u>2.1. DE LA LINGUISTIQUE A LA LOGIQUE DES QUESTIONS</u>	243
<u>2.1.1. Les relations sémantiques entre énoncés déclaratifs</u>	243
a) Le problème linguistique des présupposés d'un énoncé	243
b) Les différents types de présupposés	249
c) Les essais de formalisation de la relation de présupposition	257
d) Les autres types de relations sémantiques entre énoncés	266
<u>2.1.2. Les relations sémantiques entre questions, réponses, et présupposés</u>	275
a) Les apports de la linguistique	275
b) L'analyse logico-sémantique des questions de Jerrold J.KATZ	284
c) Insuffisance des analyses de type linguistique	293

<u>2.1.3. La logique érotétique : brève présentation historique</u>	295
a) Les précurseurs (de l'antiquité au dix-neuvième siècle)	295
b) Naissance et développement de la logique des questions	304
<u>2.2. LA LOGIQUE DES QUESTIONS ET DES RÉPONSES DE NUEL D. BELNAP</u>	315
<u>2.2.1. Les concepts de base</u>	316
a) Les fondements assertoriques	316
b) La structure des questions élémentaires	320
c) Le concept de réponse directe à une question	323
<u>2.2.2. Types de questions fondamentaux</u>	325
a) Les quatre formes distinguées de questions <i>whether</i>	325
b) Variété et réductibilité des questions <i>whether</i>	330
c) Les six formes distinguées de questions <i>which</i>	335
d) Caractères généraux des questions <i>which</i>	343
<u>2.2.3. Autres formes de questions simples</u>	346
a) Les questions de description	347
b) Les questions d'identité	348
c) Les questions de quantité	348
d) Les questions <i>what</i>	349
e) Les questions de causalité	352
f) Irréductibilité à une forme unique	354
<u>2.2.4. Les questions complexes</u>	355
a) Composition de questions simples	355
b) Relativisation des questions absolues	358
c) Quantification des questions simples	363
d) Autres formes de questions complexes	365

<u>2.2.5. Analyse sémantique des questions et des réponses</u>	368
a) Valeur de vérité d'une question	368
b) Présupposition d'une question	371
c) Classifications sémantiques des réponses	373
d) Classifications sémantiques des questions	379
e) Relations sémantiques entre questions	384
<u>2.3. BRÈVE PRÉSENTATION D'AUTRES SYSTÈMES DE LOGIQUE ÉROTÉTIQUE</u>	389
<u>2.3.1. La théorie logique des interrogatives de Lennart ÅQVIST</u>	389
a) Les questions <i>whether</i>	392
b) Les questions <i>what</i> demandant une énumération exhaustive	397
c) Autres types de questions <i>what</i>	402
d) Les réponses aux questions	408
<u>2.3.2. La théorie logique des questions de Tadeusz KUBIŃSKI</u>	418
a) Les questions numériques	419
b) Les autres types de questions	428
c) Relations entre questions, réponses, et présuppositions	437
<u>2.3.3. L'analyse des questions de causalité de S. BROMBERGER</u>	449
a) Insuffisances de la théorie de HEMPEL	450
b) Règles générales et lois de déviance	452
c) Les réponses aux questions de causalité	455
d) Insuffisances de la théorie de BROMBERGER	457
<u>2.3.4. Survol de quelques autres formes de logique érotétique</u>	461
a) Kazimierz AJDUKIEWICZ	462
b) Friedrich LÖW	465
c) Gerold STAHL	471
d) David HARRAH	475

<u>2.4. SYNTAXE ET SÉMANTIQUE FORMELLES DES QUESTIONS ET DES RÉPONSES</u>	481
<u>2.4.1. Syntaxe</u>	481
a) Les trois formes de questions élémentaires	482
b) Autres formes de questions	486
c) Analyse du contenu de la requête	490
d) La notion de réponse directe à une question	497
<u>2.4.2. Sémantique</u>	500
a) Propriétés sémantiques des questions	501
c) Classification sémantique des questions	504
d) Classification sémantique des réponses	510
<u>2.4.3. Problèmes en suspens</u>	516
a) Insuffisances formelles de la logique érotétique	516
b) Insuffisances pratiques de la logique érotétique	519
c) Perspectives d'application	521
<div style="border: 1px solid black; padding: 5px; display: inline-block;">3. QUESTIONNAIRE ET ANALYSE DES DONNÉES</div>	529
<u>3.1. DES QUESTIONS AUX VARIABLES</u>	533
<u>3.1.1. Classification empirique des formes de questions</u>	534
a) Rappel des classifications usuelles	536
b) Questions à réponse libre	538
c) Alternatives	546
d) Questions à choix multiples	561
e) Autres formes de questions	572

<u>3.1.2. Classification systématique des formes de consignes</u>	575
a) Sélection	577
b) Ordination	582
c) Classification	587
d) Appariement	593
e) Réduction à la sélection	600
f) Puissance d'une question	605
<u>3.1.3. La construction des variables</u>	611
a) Les relations sémantiques entre les modalités de réponse	612
b) Des modalités de réponse aux variables	623
c) Problèmes d'interprétation des réponses	635
<u>3.2. DU QUESTIONNAIRE A LA "MATRICE DES DONNÉES"</u>	647
<u>3.2.1. Structure d'un questionnaire d'enquête</u>	649
a) Présentation matérielle du questionnaire	649
b) Les aiguillages	653
c) Les répétitions de questions	666
d) Rôle informatif et rôle structurel des questions	674
<u>3.2.2. Structure d'un test de niveau</u>	678
a) Présentation matérielle du BINET-SIMON	679
b) Règles d'appréciation des réponses	684
c) Aiguillages internes à une sous-question	687
d) Aiguillages internes à une question	692
e) Règles d'enchaînement des questions	694
<u>3.2.3. Description formelle des questionnaires</u>	698
a) Relations entre questions dépendant de la consigne	699
b) Relations entre questions et réponses	713
c) Relations sémantiques entre questions	722

<u>3.2.4. Réduction d'un questionnaire à une forme canonique</u>	724
a) Confusion des questions identiques incompatibles	725
b) Réduction des questions à une forme canonique	725
c) Mise du questionnaire sous une forme séquentielle	733
d) Questionnement et "matrice des données"	740
<u>3.3. PRÉALABLES À L'ANALYSE DES DONNÉES</u>	745
<u>3.3.1. Information structurelle et niveaux de mesure</u>	746
a) Questionnaire formel et quanta d'information	746
b) Les opérateurs de relation \mathcal{Y} et \mathcal{D}	750
c) Les "niveaux de mesure" dans les sciences de l'homme	757
d) Représentation des niveaux de mesure par les opérateurs \mathcal{Y} et \mathcal{D}	763
<u>3.3.2. Information métrique et indices statistiques</u>	779
a) Mesures de dispersion	781
b) Mesures de liaison entre variables	787
c) Perspectives de recherche	796
<u>CONCLUSIONS</u>	799
<u>Annexe I</u> : Liste des tests et questionnaires cités	807
<u>Annexe II</u> : La mesure de l'information structurelle	813
<u>Bibliographie</u>	819