Étude comparative des pronoms démonstratifs neutres anglais et français à l'oral : référence indexicale, structure du discours et formalisation en grammaire notionnelle dépendancielle

Laurie Buscail

▶ **To cite this version:**

Laurie Buscail. Étude comparative des pronoms démonstratifs neutres anglais et français à l'oral : référence indexicale, structure du discours et formalisation en grammaire notionnelle dépendancielle. Linguistique. Université Toulouse le Mirail - Toulouse II, 2013. Français. NNT : 2013TOU20058 . tel-00965362

HAL Id: tel-00965362
https://theses.hal.science/tel-00965362

Submitted on 25 Mar 2014

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 de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 2 Le Mirail (UT2 Le Mirail)

Présentée et soutenue par :
Laurie Buscail
Le vendredi 4 octobre 2013

Titre :
Etude comparative des pronoms démonstratifs neutres anglais et français à l'oral : référence indexicale, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle

---

École doctorale et discipline ou spécialité :
ED CLESCO : Linguistique anglaise

Unité de recherche :
CLLE-ERSS UMR5263

Directeur(s) de Thèse :
Monsieur Jacques Durand, Professeur Emérite à l'Université Toulouse 2 Le Mirail & Membre Senior de l'IUF
Madame Nathalie Rossi-Gensane, Professeur à l'Université Lumière Lyon 2

Rapporteurs :
Monsieur Nicolas Ballier, Professeur à l'Université Paris Diderot (Paris 7)
Madame Mireille Bilger, Professeur à l'Université de Perpignan Via Domitia

Autres membres du jury :
Madame Anne Condamines, Directrice de Recherche au CNRS & CLLE-ERSS (UMR5263)
Madame Fabienne Toupin, Professeur à l'Université de Tours

Sommaire

Remerciements ***4***

Résumé/abstract ***6***

Chapitre I. Introduction générale et présentation des chapitres ***7***

Chapitre II. Les démonstratifs en anglais : analyses structurale, générative et énonciatives ***10***

Chapitre III. Les démonstratifs en français : analyses structurale, générative et énonciatives ***44***

Bilan des chapitres II et III. Les ressemblances et les divergences entre démonstratifs anglais et français ***79***

Chapitre IV. La référence et l'indexicalité ***82***

Chapitre V. Représentation des expressions référentielles en Grammaire Notionnelle Dépendancielle ***129***

Chapitre VI. Les corpus oraux en linguistique : avantages, limites et choix des données ***149***

Chapitre VII. Retour sur les démonstratifs anglais : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle ***182***

Chapitre VIII. Retour sur les démonstratifs en français : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle ***235***

Chapitre IX. Perspectives de recherches ***278***

Chapitre X. Conclusion générale ***300***

Bibliographie ***303***

Conventions typographiques ***315***

Glossaire ***317***

Index ***340***

Table des matières ***342***

Table des figures ***349***

Table des tableaux ***351***

Annexes : corpus et conventions PAC/PFC ..... ***CD***

Remerciements

Une page (ou deux…) de remerciements n'est certainement pas suffisante à exprimer toute ma gratitude envers les personnes qui sont devenues, au fil de ces quatre années, des membres de ma famille. Je vais, néanmoins, faire de mon mieux.

Tout d'abord, je tiens à remercier, du fond du cœur, mes co-encadrants de thèse, Monsieur Jacques Durand et Madame Nathalie Rossi-Gensane. Merci d'avoir repris ma thèse en main, à neuf mois de l'échéance et au bord de l'abandon. Je ne l'oublierai jamais ! Nathalie : merci pour nos longues conversations (2h30 au téléphone la dernière fois, je n'y crois toujours pas !) qui m'ont, à chaque fois, mis beaucoup de baume au cœur. Nos prochains coups de fil seront au sujet de notre chapitre pour OUP, on n'a pas d'autre choix que de s'y mettre, et vite ! Jacques : c'est à ce moment précis, où je désire exprimer ma gratitude, que je réalise à quel point les mots sont parfois insuffisants. Merci pour nos débats entre linguistique et philosophie, merci pour votre disponibilité sans faille, pour votre bonne humeur, votre enthousiasme communicatif, votre bienveillance, votre aide, vos encouragements. Je vous dois tout et bien plus.

Merci à l'ensemble des équipes PAC et PFC, sans qui la réalisation de cette thèse n'aurait pas été possible. En particulier, merci à Chantal Lyche, qui m'a fait découvrir la méthode de fabrication du poiré domfrontais, merci à Julien Eychenne, qui me fait confiance au point de mettre au monde un bébé SAXO. Un grand merci à Sylvain Detey et à Isabelle Racine : c'est grâce à vous que mon pharmacien s'est enrichi, de par ma consommation phénoménale d'aspirine pendant les périodes de codage des voyelles nasales. C'est grâce à vous également que chaque Journée PFC se transforme en colonie de vacances (en dehors des heures de conférences, bien entendu !) et demeure un souvenir impérissable. Isabelle : merci de me faire prendre l'accent suisse chaque été !

Merci à l'ensemble du laboratoire CLLE-ERSS, pour son accueil chaleureux et pour le soutien dont il fait preuve à l'égard des doctorants. Un grand merci aussi à l'IUT GEA de Perpignan, qui m'a accueillie à bras ouverts en tant qu'ATER pour ma dernière année de thèse. Je serai encore parmi vous à la rentrée 2013, mais il ne faut jamais remettre au lendemain ce que l'on peut faire le jour même : Gilles, Hélène, Laurent, Marie-Françoise,

Michel, Michèle, Rose-Marie, je vous aime ! Par ailleurs, merci aussi à mes étudiants de me donner moins de vingt-cinq ans.

Un très grand merci à ma famille : maman, Leslie, Nicolas (Niko pour les intimes). Merci d'avoir supporté mes sautes d'humeur, mon caractère irascible, mes envolées linguistiques, mon enfermement, mon introspection. Je vous aime du fond du cœur et j'admire votre patience. Merci également à Sylvain Navarro, mon cher jumeau de doctorat, avec qui la collaboration administrative et amicale fut d'une aide précieuse, surtout pendant les dernières semaines.

Enfin, je tiens à remercier très chaleureusement les membres du jury : Monsieur Nicolas Ballier, Madame Mireille Bilger, Madame Anne Condamines et Madame Fabienne Toupin. Merci à Monsieur Nicolas Ballier et à Madame Mireille Bilger d'accepter d'être les rapporteurs de cette thèse. Mireille : merci de votre accueil au sein de l'équipe de Perpignan et de m'avoir intégrée, Henry Tyne et vous-même, au comité d'organisation de l'AFLS2013.

Résumé/abstract

Résumé : Cette thèse explore le fonctionnement indexical des pronoms démonstratifs anglais *this, that* et *it* d'une part, et des pronoms démonstratifs français, en particulier *ça*, d'autre part, en vue d'une comparaison entre ces deux systèmes. L'ensemble des phénomènes référentiels et discursifs observés sont ramenés à certaines caractéristiques syntactico-sémantiques propres à chaque démonstratif, alors formalisées dans le cadre de la Grammaire Notionnelle Dépendancielle. Les occurrences de *this, that, it* et *ça* analysées étant issues de conversations orales spontanées et enregistrées selon le protocole des projets PAC et PFC, notre étude apporte un questionnement sur les avantages et les limites des grands corpus oraux pour les recherches en linguistique théorique.

Abstract: This thesis focuses from a comparative perspective on the indexical functioning of both English and French demonstrative pronouns, namely *this, that* and *it* on the one hand, and *celui-ci/là, ceci, cela* and *ça* on the other – with particular emphasis on the neuter pronoun *ça*. The overall referential and discursive phenomena which are examined are tied to a number of syntactic and semantic features and representations characterizing each demonstrative item. Our treatment is then formalized within the framework of Notional Dependency Grammar. As all our crucial examples concerning *this, that, it* and *ça* are extracted from authentic face to face conversations from the PAC and PFC projects, the present study leads to a discussion of the potential advantages and limits of large spoken corpora for research in theoretical linguistics.

Chapitre I. Introduction générale et présentation des chapitres

Cette thèse de doctorat explore le fonctionnement indexical des unités grammaticales que sont les pronoms démonstratifs anglais *this, that* et *it* d'une part, et les pronoms démonstratifs du français, en particulier *ça*, d'autre part, en vue d'une comparaison entre ces deux systèmes. Notre étude est une double contribution, à la fois empirique et théorique. Au niveau empirique, nos travaux s'inscrivent dans les projets PAC (Phonologie de l'Anglais Contemporain) et PFC (Phonologie du Français Contemporain), puisque toutes les occurrences analysées sont extraites de conversations spontanées enregistrées dans le cadre de ces programmes. L'exploitation de ces bases de données mènera par là même à une réflexion sur les avantages et les limites des grands corpus oraux pour la recherche en syntaxe et en sémantique. Au niveau théorique, notre étude part d'une volonté d'harmonisation des différentes analyses proposées pour les démonstratifs. Considérant que l'anaphore est dérivée de la deixis et que ces procédures indexicales sont des phénomènes *ad hoc* observables en discours, nous proposons de ramener l'ensemble des manifestations référentielles et discursives de *this, that, it* et *ça* à certains traits minimaux propres à chacun de ces pronoms. De là, nous soumettons une représentation possible des sytèmes des pronoms démonstratifs anglais et français en Grammaire Notionnelle Dépendancielle, cadre théorique qui considère et formalise la structure syntactico-sémantique interne aux éléments langagiers.

Les chapitres II et III sont construits selon le même schéma : ils offrent une vue d'ensemble des grands courants linguistiques, du structuralisme à l'énonciation, tout en présentant le traitement des démonstratifs proposé par ces théories. Ces deux chapitres étant introductifs, ce qui pourrait apparaître comme des incohérences ou des inexactitudes terminologiques est volontairement conservé afin d'exposer les points de vue adoptés par ces différents cadres théoriques, de la manière la plus factuelle possible. Qui plus est, l'inventaire des analyses proposé dans ces chapitres permet de dégager certaines propriétés des démonstratifs, des syntagmes nominaux et du langage en général ; en effet, bien que ces analyses puissent s'avérer insuffisantes, notamment en raison de l'exclusion de certains paramètres indispensables à la description fine des éléments linguistiques, elles sont à la base de la compréhension du fonctionnement langagier. Les trois grands courants exposés dans les chapitres II et III sont le structuralisme, la Grammaire Générative et les linguistiques de l'énonciation. La section sur l'analyse structurale dresse un inventaire des formes démonstratives, dégage les propriétés distributionnelles des démonstratifs et apporte un

éclairage terminologique et catégoriel sur ces items. La section concernant les analyses génératives distingue les notions de compétence et de performance essentielles à la compréhension de ces traitements résolument syntaxiques. Néanmoins, nous verrons que l'évolution de ces théories suscite des hypothèses qui sortent du cadre purement syntaxique et dessine une vision cognitive de la référence. La section sur les linguistiques de l'énonciation distingue les notions de 'langue' et de 'parole' ou de 'langue' et de 'discours', le postulat général de ces théories étant que les énoncés constituent la trace visible d'opérations cognitives sous-jacentes. Ainsi, chaque item langagier serait doté d'un pôle d'invariance, de sens ou de forme selon le modèle théorique, capable de rendre compte de tous les emplois observés en discours. Bien que les analyses structurales, génératives et énonciatives aient des prémisses théoriques pour le moins divergentes, il n'en reste pas moins qu'elles s'accordent sur la fonction majeure des démonstratifs, à savoir celle de concourir à un mode de détermination particulier. Les chapitres II et III sont suivis d'un bilan (Bilan des chapitres II et III), en préambule à la délimitation du cadre théorique que nous choisissons d'adopter pour nos analyses.

Le chapitre IV s'attache à présenter les notions essentielles à l'étude des démonstratifs anglais et français proposée dans les chapitres VII et VIII. Nous argumentons en faveur d'une théorie uniforme de la référence qui ne distingue pas la désignation du 'réel' – ou 'extralinguistique' – de celle du 'texte'. Pour cela, il est nécessaire de (re)définir, d'affiner, voire de modifier, les concepts de 'référence', de 'texte', de 'contexte', de 'discours', d''anaphore' et de 'deixis'. A la description référentielle est également indispensable la présentation des différents ordres d'entités proposés par Lyons, qui classe les objets-de-discours en fonction de leur visée sémantique plus ou moins abstraite. Enfin, nous incluons la mémoire, selon trois types concentriques et perméables, à notre modèle référentiel.

Le chapitre V introduit le cadre théorique qu'est la Grammaire Notionnelle Dépendancielle de John Anderson, issue de la tradition de Tesnière et ancrée dans la grammaire des cas. Ce modèle propose une description syntactico-sémantique de la structure interne aux éléments langagiers, ce qui permet de rendre compte du comportement distributionnel, référentiel et discursif de ces derniers. La représentation formelle proposée par la Grammaire Notionnelle Dépendancielle sera appliquée à *this*, *that*, *it* et *ça* dans les chapitres VII et VIII.

Le chapitre VI propose une définition des différents registres langagiers, à l'oral et à l'écrit, et problématise la finalité des corpus oraux et l'utilité de ces derniers pour l'analyse en syntaxe et en sémantique. Nous soulevons les avantages et les insuffisances des corpus disponibles à ce jour, en nous basant, principalement, sur les projets PAC et PFC, dont sont issues les occurrences de *this*, *that*, *it* et *ça* analysées dans les chapitres VII et VIII. Enfin, nous évoquons les difficultés et les solutions possibles à la création d'un analyseur de l'oral.

Les chapitres VII et VIII proposent une analyse contrastive des démonstratifs anglais et français selon les paramètres définis dans le chapitre IV. Nous explorons la référence indexicale de *this, that, it* et *ça* par la description fine du mécanisme opéré par ces indexicaux lors d'un accès à la référence via les différents contextes – situationnel immédiat, discursif, savoirs encyclopédiques et partagés –, mais aussi par la mise en exergue de leurs fonctions en tant que signaux de co-construction du discours. Les différents phénomènes observables en discours produits par *this, that, it* et *ça* sont alors ramenés à deux paramètres minimaux intrinsèques à chacun de ces indexicaux : le fonctionnement syntactico-sémantique propre aux démonstratifs et la sphère cognitive et discursive définie par chacun d'eux. De là, nous proposons une représentation de la structure interne des pronoms démonstratifs anglais *this*, *that* et *it* d'une part, et des pronoms démonstratifs français *celui-ci/là*, *ceci*, *cela* et *ça* d'autre part, en Grammaire Notionnelle Dépendancielle.

Le chapitre IX offre des pistes de recherches possibles afin d'approfondir l'étude synchronique des démonstratifs, notamment par la prise en compte de l'évolution diachronique de ces items, de l'accentuation prosodique, du lien potentiel entre intonation et référence, et de l'observation des indexicaux dans les différentes langues du monde.

Enfin, le chapitre X présente les conclusions de cette thèse.

Chapitre II. Les démonstratifs en anglais : analyses structurale, générative et énonciatives

Ce chapitre est à la fois introductif et présentatif de ce que peut offrir la littérature existante portant sur les démonstratifs anglais. C'est délibérément que ne sont pas exposées ici les théories cognitives issues du courant américain, les théories fonctionnalistes, ainsi que les écoles alliant ces deux versants, à la faveur des cadres épistémologiques 'classiques' – ici compris comme 'fondamentaux' – qui permettent de dégager certaines propriétés des démonstratifs, du syntagme nominal (désormais SN) et du langage en général. Ainsi, ce chapitre présente les démonstratifs en tant qu'objets syntagmatiques, syntaxiques et opératoires par le parcours des courants de pensée structural (section 1), génératif (section 2) et énonciatif (section 3).

Le système des démonstratifs en anglais est composé de *this* et *that* pour les formes au singulier, *these* et *those* pour celles au pluriel. Si cet ensemble restreint comprend à la fois les déterminants et les pronoms démonstratifs, il est à noter qu'un troisième élément issu du vieil-anglais, *yon* (« that thing over there » ; cf. Britain, 2007 ; Kortmann & Upton, 2008), perdure aujourd'hui dans les textes poétiques ainsi que dans certaines variétés de l'anglais. *This*, *that*, *these* et *those* appartiennent à la classe grammaticale des formes en *th-*, dont la consonne d'attaque est invariablement la fricative voisée /ð/. Notons ici que *that* [ðæt] est distingué de *that* [ðət], dont la forme phonologique correspond au complémenteur – *I didn't know* ***that*** *he was coming ; he is the only student* ***that*** *I know*. Nous pourrons constater en section 3.2.3.1 que cette particularité du /ð/ à l'initiale, loin d'être anodine, est un élément central de l'analyse des démonstratifs dans certaines théories énonciatives ; mais, tout d'abord, nous souhaitons dresser un portrait structural des démonstratifs en anglais.

1. L'analyse structurale des démonstratifs en anglais

Les courants structuraliste et distributionnaliste (Bloomfield, 1933) permettent l'émergence du SN, ayant le nom comme noyau, auquel sont adjoints[1] un ou plusieurs constituants, dont les déterminants. Les démonstratifs sont eux communément divisés en deux

[1] 'Adjoint' est ici accepté dans son sens le plus général, sans distinction aucune entre 'adjoint/facultatif' et 'dépendant/obligatoire'.

catégories : les déterminants démonstratifs, qui accompagnent un groupe[2] lexical, et les pronoms démonstratifs, qui, comme ce terme l'indique, remplaceraient un nom ou un SN. En anglais, déterminants et pronoms partagent les mêmes formes, et sont par conséquent non distinguables d'un point de vue morphologique :

(1) *Give me **that** book* (déterminant)
(2) *Give me **that*** (pronom)

Ceci n'est pas le cas dans toutes les langues ; en français, par exemple, les déterminants démonstratifs (*ce*, *cette*) se différencient des pronoms (*celui-ci*, *celui-là ; ceci*, *cela ; ça*). Dans certaines langues comme le turc, les deux catégories partagent la même forme (*bu*), mais se distinguent par leur comportement flexionnel : les pronoms démonstratifs sont marqués suivant le cas et le genre, tandis que les déterminants sont, eux, invariables, et accompagnent un nom fléchi (exemples issus de Kornfilt, 1997 : 312, 315) :

(3) *Ali* ***bu**n-u* *unut-ami-yor*.
Ali *this*-ACC *forget-not*-PROG
('Ali ne peut pas oublier ceci')
(4) ***Bu*** *gazete-yi*
This *newspaper*-ACC
('ce journal')

Pourrions-nous donc supposer que déterminants et pronoms démonstratifs appartiennent à la même catégorie en anglais ? Nous serions encline à répondre que la réponse est non. En effet, tout comme l'anglais, d'autres langues ne distinguent pas leurs déterminants et pronoms démonstratifs, d'un point de vue morphologique. Le tuscarora[3], par exemple, possède deux démonstratifs, *kyè:ni:kə:* (*this*/*these*) et *hè:ni:kə:* (*that*/*those*), utilisés en tant que pronoms et déterminants. Cependant, lors de ce dernier emploi, le démonstratif et le groupe nominal qu'il accompagne sont reliés de manière faible, pourrait-on dire : la position qu'ils occupent l'un par rapport à l'autre est flexible (exemples (5a) et (5b)), et une rupture prosodique peut

[2] 'Groupe lexical' ou 'groupe nominal' désigne le « nom + modifieurs éventuels », tandis que 'syntagme' nominal signale l'ensemble complet « déterminant + nom + modifieurs éventuels ». Il est à noter qu'en anglais, le déterminant peut être l'élément zéro *ø*.

[3] Le tuscarora est une langue iroquoienne du Nord parlée dans l'État de New York et en Ontario sur la réserve des Six-Nations.

survenir lorsque tous deux sont juxtaposés (exemple (5c)) (exemples issus de Mithun, 1987 : 184, 186) :

(5) a. ***hè:ni:kə:*** *àha:Θ*
that horse
('ce cheval')

b. *uʔné:wa:k* ***hè:ni:kə:***
ghost that
('ce fantôme')

c. *waʔtkahà:hi:Θ* ***hè:ni:kə: (…)*** *ruyaʔkwàhehr*
it.met.it that (…) he.body.carries
('il a rencontré ce dinosaure')

En anglais, la situation est tout autre : le déterminant démonstratif occupe une position fixe par rapport au groupe lexical qu'il accompagne, et tous deux partagent le même contour mélodique. De plus et à l'inverse des pronoms, les déterminants requièrent la présence d'un élément lexical, sont défocalisés, voire clitiques dans certaines langues – les articles en français par exemple –, et peuvent présenter un phénomène d'incorporation, comme l'élément *au* qui est composé de la préposition *à* associée à l'article *le*. Par conséquent, déterminants et pronoms démonstratifs en anglais sont en relation paradigmatique avec d'autres éléments qui appartiennent à des catégories différentes : les déterminants démonstratifs occupent la même place syntaxique que les articles et les déterminants possessifs, tandis que les pronoms démonstratifs relèvent du même paradigme que d'autres pronoms. Pour toutes ces raisons, nous pourrions trouver légitime de penser que déterminants et pronoms démonstratifs en anglais appartiennent à deux catégories différentes, bien que non distinguables d'un point de vue morphologique. Nous reviendrons sur la catégorisation des déterminants et pronoms démonstratifs lors de leur analyse syntactico-sémantique dans les chapitres VII et VIII.

D'autre part, les déterminants démonstratifs ne sont plus considérés comme des adjectifs (bien que cette appellation persiste chez certains linguistes, comme dans les écrits de Kleiber par exemple, peut-être par respect de la tradition[4]). En effet, les déterminants

[4] Le terme 'déterminant' en tant que « mot précédant le substantif, de même genre et de même nombre que lui, qui le caractérise, en précise la valeur dans la phrase » fait son entrée dans la neuvième édition de la *Grammaire de l'Académie Française* (1992).

indiquent un mode d'accès à la référence*, mais ne modifient pas le référent* en tant qu'altération de sens*. Ces deux catégories se doivent par conséquent de rester distinctes. En anglais, les déterminants démonstratifs se distinguent des autres déterminants – article défini, déterminants possessifs – par la marque du nombre selon le nom qu'ils accompagnent – *this/that* + nom au singulier, *these/those* + nom au pluriel – et, comme nous l'avons déjà souligné, par leur capacité à agir seuls en tant que SN ; de cette dernière propriété se dégagent les déterminants et pronoms démonstratifs, mentionnés comme 'pronoms' et 'pronoms sans nom'[5] (Jespersen, 1933), ou comme 'adjoints' et 'têtes' (Strang, 1968). Notons ici la présence dialectale de *them* en instance de *those*, attestée dans de nombreuses régions de l'Angleterre (cf. Britain, 2007 ; Kortmann & Upton, 2008).

THIS et *THAT*[6], déterminants et pronoms, appartiennent à la même catégorie indexicale que les adverbes *here* et *there*, *now* et *then*, qui sont des déictiques* (cf. entre autres Jespersen, 1933 ; Strang, 1968). Du grec ancien *deiktikos* ('action de montrer'), les déictiques sont intimement liés à l'acte énonciatif afin de pouvoir être référentiellement interprétés. En d'autres termes, *here* et *there* ne pourront désigner un « ici » et un « là » qu'en fonction d'un centre locatif (*hic*[7]), *now* et *then* seront interprétables grâce à un ancrage temporel (*nunc*) ; les déictiques n'ayant pas de signifié fixe, ces items sont parfois appelés 'embrayeurs' – 'shifters' en anglais. A ce système *hic* et *nunc* s'ajoute un troisième point de référence, le locuteur/scripteur *je*. Ainsi, ce *hic/nunc/je* représente un espace vectoriel dont le locuteur/scripteur *je* est le 'centre déictique', ou *origo* (Bühler, 1934[2009]). Au sein de cet espace déictique, *THIS* et *THAT* auront la fonction prototypique de désigner des « choses », ostension souvent accompagnée d'un geste déictique, généralement un signe du doigt, de la tête, ou du regard. Il est à noter que *THIS* marque la proximité, et *THAT* l'éloignement (cf.

[5] Dans sa théorie des rangs, Jespersen distingue parties du discours et syntaxe, et propose une classification selon trois rangs – rang 1, rang 2, rang 3 : « Dans la combinaison *extremely hot weather* ('temps extrêmement chaud'), le dernier mot, *weather* ('temps'), qui représente évidemment l'idée principale, peut être désigné comme de rang 1 ; *hot* ('chaud'), qui définit *weather*, est alors de rang 2, et *extremely* ('extrêmement'), qui définit *hot*, de rang 3. (...) Il est inutile de distinguer plus de trois rangs, car il n'existe aucun trait, formel ou autre, qui permette de distinguer ces mots d'un rang inférieur des mots de rang 3 » (Jespersen, 1971). Ainsi, le pronom démonstratif serait de rang 1, tandis que le déterminant démonstratif serait de rang 2.

[6] *THIS* et *THAT* écrits en lettres majuscules signalent que nous incluons les formes déclinées sans les énumérer.

[7] Tout au long de cette thèse, *hic/nunc*, *here/there* et *now/then* seront tour à tour employés pour renvoyer aux repères spatio-temporels.

entre autres Jespersen, 1933[8] ; Tellier, 1971 ; Halliday & Hasan, 1976). Cette distance relative est souvent perçue comme relevant du domaine spatial, mais peut également être décrite en termes d'affectivité ou de rejet en faveur du référent évoqué (Zandvoort, 1970). Strang (1968) avance l'idée de *this* comme étant le membre marqué* de la paire *this/that*, notamment pour des raisons de substitution – en théorie, chaque occurrence de *this* peut être remplacée par *that*, mais pas l'inverse. Nous mentionnons ici cette notion de marque plus ou moins distale sans plus l'expliciter, mais nous aurons le loisir d'approfondir cette relation de 'distance' dans les chapitres VII et VIII. Notons cependant que l'association de *THIS*/*THAT* (ou d'autres déictiques tels que les pronoms personnels de troisième personne) et du geste déictique ne désigne pas forcément l'« objet » physiquement montré (le *demonstratum*), mais constitue parfois un point de départ à l'identification du référent. Mentionnons ici les exemples bien connus de Nunberg (1993 : 42) et de Kleiber (1994), reproduits en (6) et en (7) :

(6) ***The ham sandwich*** *is sitting at table 7.*

(7) [A s'adresse à B, désignant du doigt le chapeau d'un homme, ce dernier étant absent]
Il *avait une grosse tête.*

The ham sandwich en (6) réfère au « client du restaurant », *il* en (7) désigne un chapeau (le *demonstratum*) mais réfère à « l'homme qui portait ce chapeau » ; ces cas d'ostension sont appelés 'deferred reference', depuis les travaux de Quine (1971). La notion de deixis* va de pair avec celle d'anaphore* : si la deixis sert à désigner un « objet », l'anaphore est utile pour reprendre un « objet » déjà mentionné. Une conception traditionnelle de l'anaphore la définit comme étant un rapport de dépendance entre deux unités linguistiques, la première – appelée 'antécédent'*, 'interprétant' (Ducrot & Todorov, 1972 : 358), 'référent' (Dubois, 1965), ou encore 'source sémantique' (Tesnière, 1959[1988] : 87) – étant nécessaire à l'interprétation de la seconde – l''anaphorique'*. Selon cette conception, la deixis et l'anaphore seraient en opposition, la deixis – également connue sous le nom d''exophore'* – référant à l''extralinguistique', l'anaphore – parfois nommée 'endophore'* – étant une manifestation textuelle. Il existe cependant une approche différente de la deixis et de l'anaphore, avancée dans les travaux de Lyons en particulier (1977), où l'anaphorique ne reprend pas un

[8] Jespersen remarque que la notion de proximité ou de distance se reflète en phonologie, *this* [ðɪs] comportant une voyelle haute fermée, *that* [ðæt] une voyelle basse ouverte, tout comme *ici* [isi] et *là* [la] en français (Jespersen, 1933 : 157).

antécédent textuel, mais réfère à l'entité discursive générée par cet antécédent textuel. Ainsi, la deixis et l'anaphore réfèrent toutes deux à des objets du discours, mais la distinction entre univers extralinguistique et texte* disparaît : l'anaphore, contrairement à la deixis, présuppose l'existence d'une entité* discursive au sein de l'univers du discours*. L'anaphore serait donc un procédé dérivé de la deixis, celle-ci étant première, tant au niveau du développement historique de quelque langue donnée, qu'à celui de l'acquisition du langage. Nous reviendrons sur les notions de deixis et d'anaphore dans le chapitre IV. *THIS* et *THAT* peuvent donc être employés comme anaphoriques. Néanmoins, Jespersen remarque que l'effet de 'monstration' est annihilé lors d'une reprise effectuée par l'un ou l'autre des démonstratifs, au profit d'une simple fonction de 'représentation', qui peut être celle d'un objet stable ou d'une action, auquel cas le pronom démonstratif « remplace une proposition » (Jespersen, 1933 : 158).

2. Le traitement des démonstratifs anglais dans un cadre génératif moderne

La linguistique générative émerge dans les années 50 avec les travaux de Chomsky (*Syntactic Structures*, 1957) et va par de nombreux côtés à l'encontre du distributionnalisme de Bloomfield. Si ce dernier est fortement influencé par le behaviourisme, Chomsky envisage la capacité de langage comme innée. Ce raisonnement naît de plusieurs observations, parmi lesquelles la faculté qu'ont les enfants de produire des phrases construites selon des règles syntaxiques précises, sans que personne ne les leur enseigne. Ainsi, Chomsky pense qu'en tout être humain est intégré le savoir tacite d'un ensemble fini de règles formelles à partir desquelles est productible une infinité de phrases[9] bien formées. Dès lors, Chomsky distingue la 'compétence' langagière, qui est le savoir-faire linguistique d'un locuteur idéal, de la 'performance', qui est le résultat observable de la production langagière. Ainsi, l'énoncé perçu par un destinataire quelconque est le produit d'une série de transformations appliquées à la structure de base, origine de cet énoncé. Ces structures – 'structure de surface' correspondant à l'output perçu, 'structure profonde' correspondant à l'agencement fondamental – ainsi que les transformations opérées sont représentées de manière formelle, sous forme linéaire ou sous forme d'arbre syntaxique. La conception chomskyenne de la production langagière trouve ses fondements dans la logique mathématique, plus particulièrement dans la théorie des automates et celle des systèmes formels. Ainsi, toute

[9] Nous employons ici le terme de 'phrase' en référence aux analyses génératives. Nous débattrons de la forme canonique de la phrase et des modifications terminologiques et conceptuelles à apporter dès lors que nous passons à l'étude de données authentiques dans le chapitre VI.

phrase jugée syntaxiquement incorrecte en sortie est mal formée et exclue (‘ruled out’). La théorie développée par Chomsky connaît plusieurs étapes successives, mais a toujours conservé la syntaxe au centre de ses préoccupations, dans le but ultime d’extraire des règles communes à toutes les langues du monde – la bien nommée Grammaire Universelle* (GU). Néanmoins, la Théorie X-barre (Jackendoff, 1977), au cœur de la Théorie du Gouvernement et du Liage* (Chomsky, 1981), est une version remaniée et améliorée de l’analyse en constituants immédiats proposée par la Théorie Standard* (Chomsky, 1965), ou TS, et la Théorie Standard Etendue* (Chomsky, 1972), ou TSE, de la Grammaire Générative (désormais GG). En effet, la TS et la TSE stipulent que les unités syntaxiques complexes – syntagmes, propositions, phrases – sont composées d’unités plus simples, le dernier niveau étant constitué des items lexicaux ; très vite, il est constaté que ces règles de réécriture* ne permettaient pas de rendre compte du fait des principes universaux régissant les langues naturelles. La Théorie du Gouvernement et du Liage est une théorie des principes universaux qui alimentent la GU et des paramètres qui satisfont ces principes. Ainsi, le principe de projection permet de rendre compte du fait que les informations lexicales – sous-catégorisation, rôle sémantique – sont conservées dans les dérivations syntaxiques. De plus, la GU est composée d’un jeu de six modules (cf. Fodor, 1983) – dont le module X-barre, constitué des règles de réécriture telles que présentes dans la TS et la TSE – connectés par un ensemble de paramètres. Quatre niveaux de représentation linguistique sont pris en compte dans ce modèle : la ‘structure-D’ ou ‘structure profonde’ qui schématise la structure fondamentale de la phrase, la ‘structure-S’ ou ‘structure de surface’ qui figure la structure superficielle de la phrase, la ‘forme phonétique’ et la ‘forme logique’, sorties de la structure-S. Le module X-barre génère la structure profonde de la phrase.

2.1. L’analyse des syntagmes dans la Théorie X-barre ‘classique’

Dans un premier temps, et ce afin de bien cerner ce qu’est la Théorie X-barre dans son ensemble, regardons les exemples suivants, qui illustrent quatre types de syntagmes :

(8) *The picture of John* (syntagme nominal)
(9) *Hardly know this girl* (syntagme verbal)
(10) *Very important to me* (syntagme adjectival)
(11) *Just after the break* (syntagme prépositionnel)

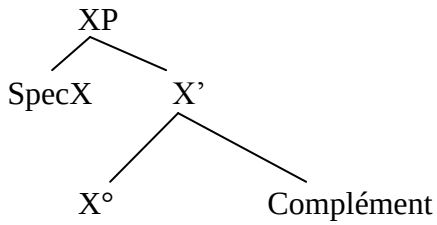
A partir de ces exemples, la Théorie X-barre révèle trois propriétés :

- Un syntagme est organisé autour d'une tête lexicale, élément recteur du syntagme : le nom *picture* pour le syntagme nominal, le verbe *know* pour le syntagme verbal, l'adjectif *important* pour le syntagme adjectival, la préposition *after* pour le syntagme prépositionnel.
- A la nature de la tête lexicale correspond le nom du syntagme : syntagme nominal (NP), syntagme verbal (VP), syntagme adjectival (AP), syntagme prépositionnel (PP).
- Les syntagmes ont tous la même structure endocentrique : une tête lexicale précédée d'un spécifieur – *the, hardly, very, just* – et suivie d'un complément – *of John, this girl, to me, the break.* Si l'on généralise la nature de la tête – N, V, A, ou P – à une variable X, nous obtenons la structure suivante : « XP = SpecX + X + Complément ».

Cette première représentation formelle des syntagmes ne satisfaisant pas à la structure à niveaux requise par la Théorie X-barre, « XP = SpecX + X + Complément » est ainsi formulée :

(12) a. XP → SpecX + X'
b. X' → X° + Complément

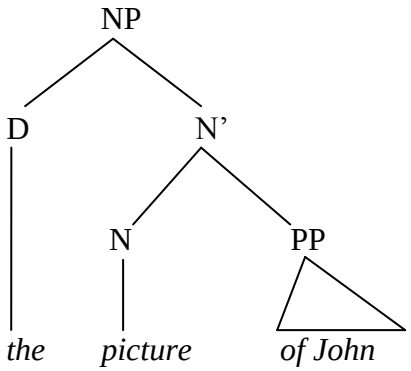
où X° représente la tête du syntagme, XP la projection maximale de la tête, et X' la projection intermédiaire de la tête. Le spécifieur et le complément sont des projections maximales, ces valeurs pouvant être nulles. La structure en (12) peut être représentée sous forme d'arborescence :



Ainsi, le SN en (8) aura la structure suivante :

(13) a. NP → D + N'
b. N' → N (PP)

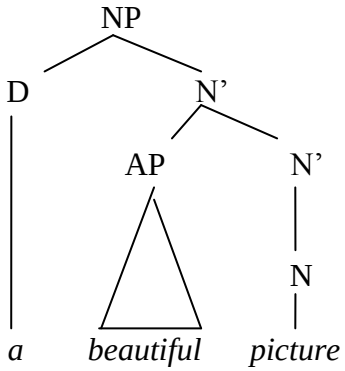
ou, sous forme d'arborescence :



Il est à noter que des constituants non obligatoires – des modifieurs – peuvent s'ajouter à deux niveaux différents :

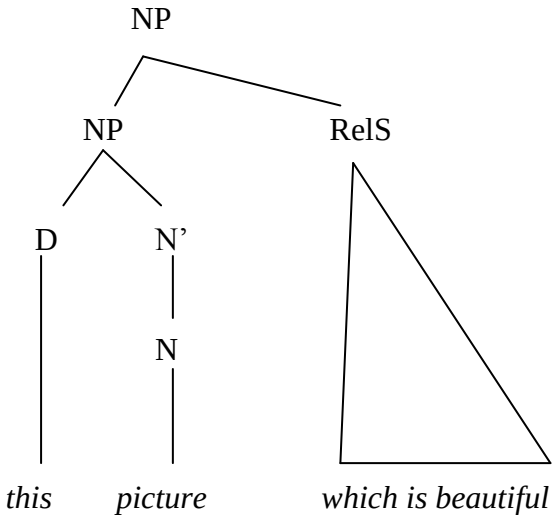
- ajout d'un modifieur au niveau X' :

(14) *A **beautiful** picture* (ici, un syntagme adjectival)



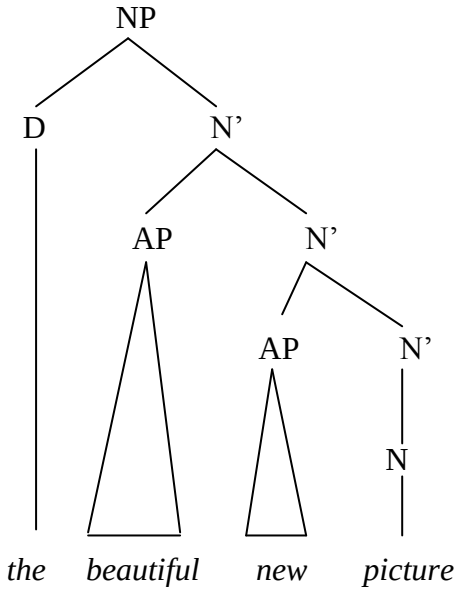
- Ajout d'un modifieur au niveau XP :

(15) *This picture, **which is beautiful*** (ici, une proposition relative)



Les modifieurs, contrairement à certains spécifieurs et aux compléments, sont récursifs et optionnels : leur présence, unique ou multiple, n'est pas requise par la tête – ils ne constituent pas un 'argument'. C'est pour cette raison qu'un syntagme adjectival AP, par exemple, est à la fois fille de N' et sœur de N' :

(16) *The **beautiful new** picture*



Cette analyse des syntagmes soulève plusieurs problèmes, notamment la différence entre spécifieurs et modifieurs, qui ne semble pas nette ; en effet, *very* dans *very important to me* est considéré comme spécifieur, mais nous ne voyons pas ce qui le démarque de *beautiful* dans *a beautiful picture*, qui lui, est considéré comme étant un modifieur : *very* et *beautiful* sont tous

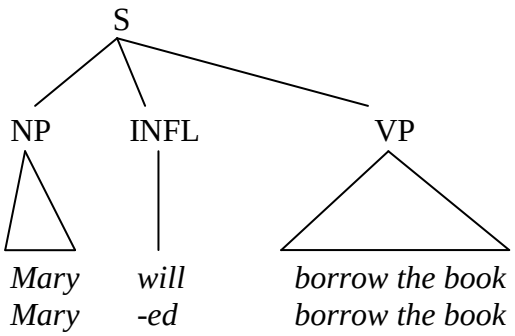
deux optionnels (non requis par la tête). Nous verrons quelle(s) réponse(s) peu(ven)t être apportée(s) en section 2.3, lorsque nous étudierons l'hypothèse avancée par Abney (1987).

2.2. L'analyse de la phrase dans la Théorie X-barre

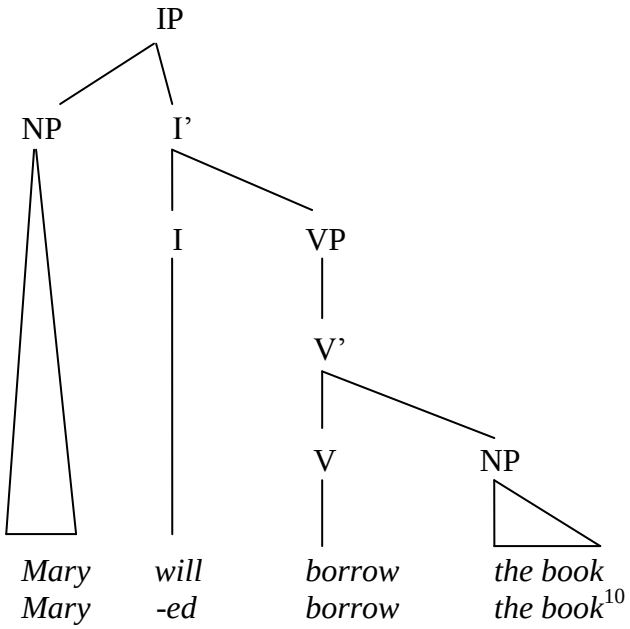
En ce qui concerne la structure de la phrase, deux problèmes se posent. Tout d'abord, la phrase est jusqu'alors considérée comme étant une structure exocentrique, c'est-à-dire composée d'un ensemble de syntagmes : « S → NP + VP ». Si la Théorie X-barre permet de rendre compte de la structure de la phrase, celle-ci se doit de s'articuler autour d'un noyau ; si nous partons d'un exemple tel que *Mary will borrow the book*, il est tentant de décider que l'auxiliaire *will* est l'élément recteur, donc la tête. En revanche, que faire d'une phrase telle que *Mary borrowed the book* ? Une manipulation simple telle qu'une extraction – *what Mary did was borrow the book* – montre que la flexion pour le temps (ici, *-ed*) ne porte plus sur le verbe mais sur l'auxiliaire (ici, *did*). Par conséquent, l'élément flexionnel pour le temps est relativement indépendant du verbe dans l'exemple *Mary borrowed the book*. Il en est conclu que la tête de la phrase n'est ni l'auxiliaire – puisque celui-ci peut être absent en sortie – ni le verbe, mais un élément plus abstrait, noté INFL :

(17) *Marry will borrow the book.*

(18) *Marry borrowed the book.*



Ce qui nous donne, sous le format X-barre :

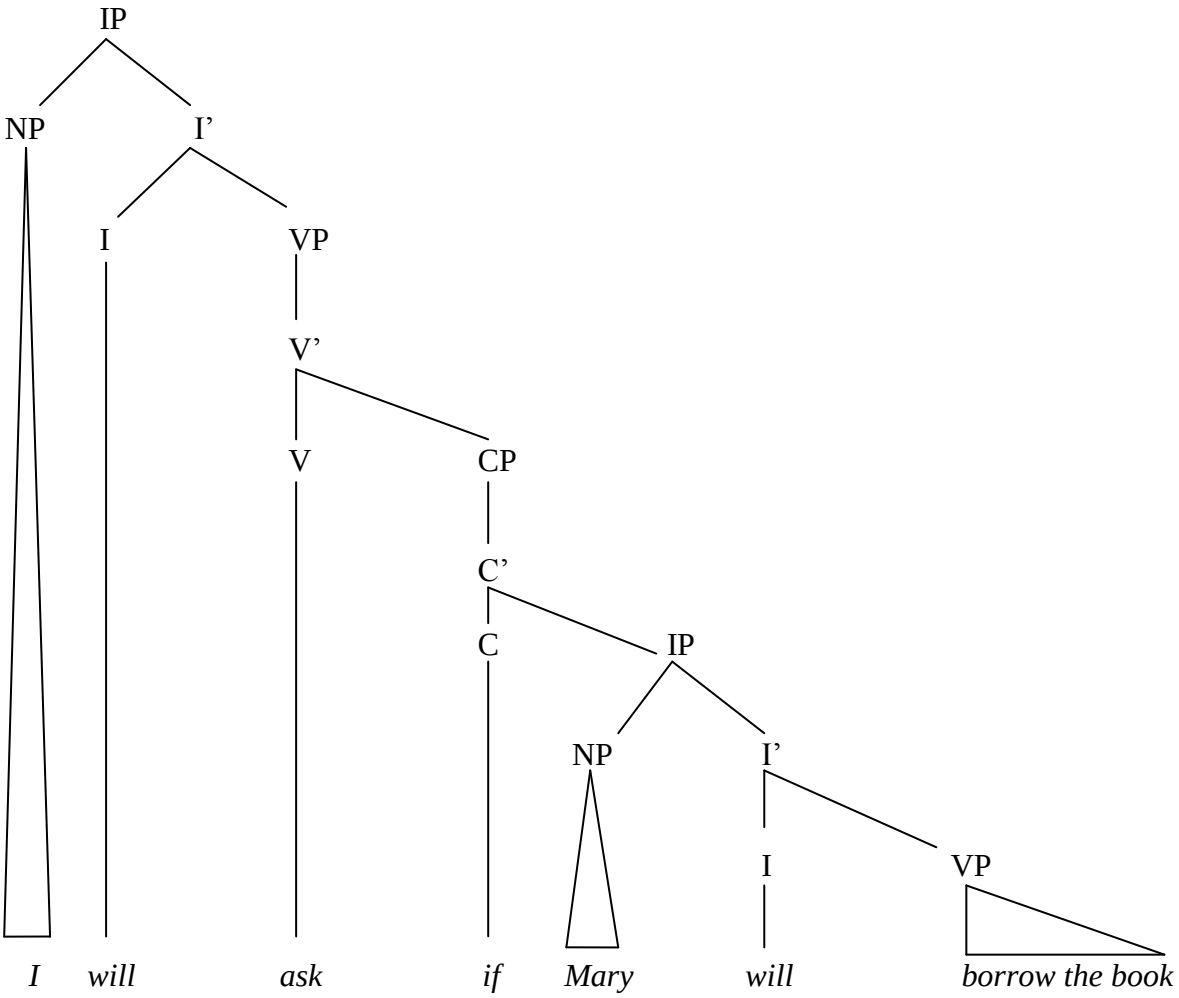


Où IP – 'Inflectional Phrase' – est la projection maximale de la tête flexionnelle. A présent, si nous étendons cette analyse aux complétives de la forme :

(19) *I will ask if Mary will borrow the book.*

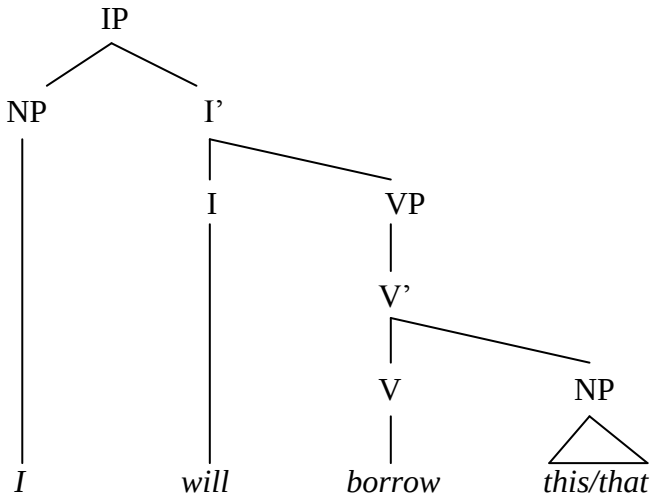
nous aurons la structure suivante, où la tête de la complétive est l'élément complémenteur C, CP étant la projection maximale de cette tête fonctionnelle :

[10] Afin d'obtenir *Mary borrowed the book*, il faut que le prédicat verbal en V se déplace en I afin de recevoir les marques morphologiques de l'accord dans un mouvement appelé 'déplacement de tête à tête'.



2.3. Le déterminant comme tête du syntagme nominal

A ce stade de la présentation de la Théorie X-barre, il serait légitime de se demander quelle est la pertinence de ce modèle pour l'analyse des démonstratifs. En effet, le SN a déjà été évoqué dans ce chapitre, présenté comme étant la projection maximale du nom-tête. Le déterminant, qu'il soit indéfini, défini, ou démonstratif, serait donc le spécifieur. *THIS* et *THAT* pronoms seraient alors considérés comme des syntagmes nominaux :



Cette analyse est envisagée comme acquise jusqu'à la thèse proposée par Abney (1987).

Le point de départ de la réflexion d'Abney est que le génitif anglais accompagné d'un gérondif – la construction 'Poss *-ing*' – partage des propriétés du SN et de la construction verbale. En effet, le génitif-gérondif apparaît en position de SN, notamment dans les constructions présentant une inversion sujet-auxiliaire, en position de sujet dans les propositions imbriquées, ainsi qu'après une préposition, là où les formes phrastiques sont prohibées :

(20) a. *John's building a spaceship*.

b. **Did [that John built a spaceship] upset you?*

Did [John] upset you?

Did [John's building a spaceship] upset you?

c. **I wondered if [that John built a spaceship] had upset you.*

I wondered if [John] had upset you.

I wondered if [John's building a spaceship] had upset you.

d. **I told you about [that John built a spaceship]*.

I told you about [John].

I told you about [John's building a spaceship]. (exemples (1) et (2a/b/c) d'Abney, 1987)

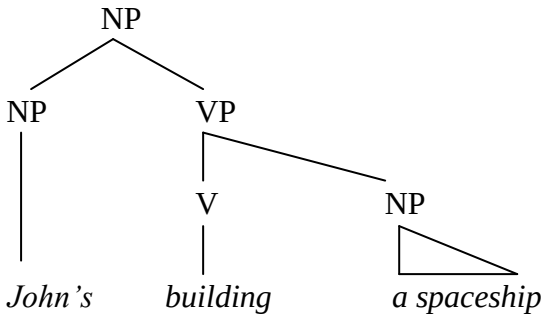
Ainsi, le nom associé au génitif – ici, *John's* – se comporte comme le possessif d'un SN, et non comme le sujet d'une phrase :

(21) a. *[John] destroyed the spaceship*.
b. *[John's] destruction of the spaceship*.
c. *[John's] destroying the spaceship*. (exemple (3) d'Abney, 1987)

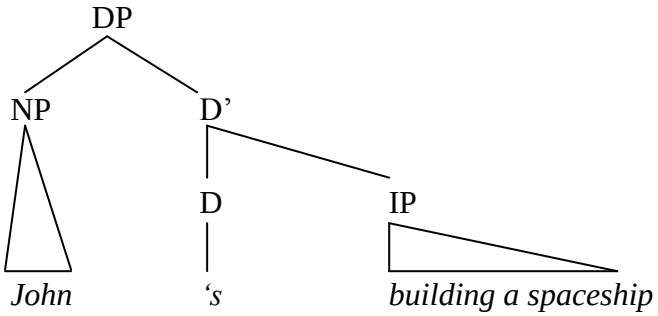
D'un autre côté, le reste du génitif-gérondif, ici, *building a spaceship*, constitue un syntagme verbal, et non un dérivé nominal en *-ing*. En effet, contrairement à un SN sujet, l'assignation du cas accusatif ne requiert pas de préposition, le sujet de surface – ici, le génitif – correspond au sujet logique de l'infinitive imbriquée, le sujet de la proposition infinitive est casuellement marqué par le verbe de la principale, le gérondif peut apparaître dans une construction à double objet, et le gérondif peut admettre une particule ; ces structures sont respectivement illustrées en (22 a, b, c, d, e) :

(22) a. **John's destruction the spaceship.*
John destroyed the spaceship.
John's destroying the spaceship.
b. **John's appearance to be dead.*
John appeared to be dead.
John's appearing to be dead.
c. **John's belief Bill to be Ceasar Augustus.*
John believed Bill to be Ceasar Augustus.
John's believing Bill to be Ceasar Augustus.
d. **John's gift/rental (of) Mary (of) a Fiat.*
John gave/rented Mary a Fiat.
John's giving/renting Mary a Fiat.
e. **John's explanation (away) of the problem (away).*
John's explained (away) the problem (away).
John's explaining (away) the problem (away). (exemples (5a/b/c/d/e) d'Abney, 1987)

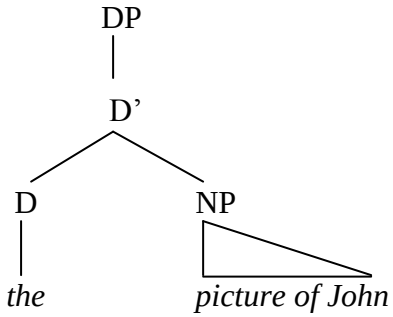
Ainsi, Abney est confronté au problème du NP au plus haut de l'arborescence n'ayant pas de tête ; en effet, VP ne peut être la tête, puisqu'il ne fait pas partie de la même catégorie que NP :



Pour résoudre ce dilemme, Abney propose de changer la structure du SN : la tête n'est plus le nom, mais un élément fonctionnel qu'il nomme D. Ainsi, la structure du NP est analogue à celle de IP, D étant le pendant nominal de INFL :



La structure du génitif-gérondif s'avérant très complexe et n'étant pas au cœur de cette thèse, nous ne nous étendons pas plus sur le sujet ; nous laissons au lecteur le soin de lire la thèse d'Abney (1987), ainsi que la vaste littérature existante sur le traitement des gérondifs au sein de ce modèle (cf. entre autres Horn, 1975, et Reuland, 1983). En revanche, l'exemple (8) – *the picture of John* – est désormais représenté de la manière suivante :

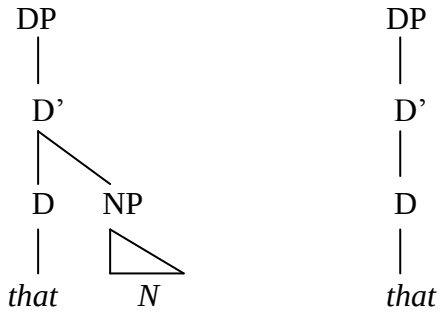


Quels sont les avantages d'une telle analyse ? Tout d'abord, le déterminant comme tête permet de rendre compte de l'impossibilité de *the* à se pronominaliser :

(23) a. *Give me the book.*
**Give me the.*
b. *Give me that book.*
Give me that.

Si le déterminant est envisagé comme tête, il permet de sélectionner ses arguments ; *the* nécessite un complément de nature nominale, tout comme *if* requiert un complément phrastique, ou comme *of* requiert un complément nominal. D'un autre côté, certains déterminants, comme *that* dans l'exemple (23b), peuvent à eux seuls former un SN ; *that*, dans ce cas précis, fonctionne tout comme un NP. Deux explications sont alors possibles : soit *that* est considéré comme un SN à lui seul, soit *that* est analysé comme étant un déterminant accompagné d'une tête lexicale vide. Nous développerons la structure interne des pronoms démonstratifs dans les chapitres VII et VIII. Quoi qu'il en soit, le choix du déterminant comme tête de syntagme pallie ce problème :

(24)



En règle générale, l'analyse des pronoms comme D-tête peut rendre compte des spécificités distributionnelles de ces derniers, notamment du fait qu'ils ne puissent accepter de spécifieur. En effet, les déterminants, possessifs, adjectifs, et quantifieurs sont bannis :

(25) a. **[the she that I talked to] was nice*.
b. **[my she] has always been good to me*.
c. **[dependable them] are hard to find*.
d. **[many they] make housecalls*.
e. **[two dozen us] signed the petition*. (exemples 304a/b/c/d/e d'Abney, 1987)

Cette spécificité distributionnelle est rendue évidente par le fait qu'un déterminant n'est jamais précédé d'un autre déterminant ; en revanche, traiter les pronoms comme N en GG

implique qu'ils puissent être précédés d'un déterminant[11]. De plus, les déterminants et les pronoms ont la caractéristique commune d'indiquer les traits grammaticaux des syntagmes nominaux, comme le genre et le nombre. En anglais, les pronoms sont les seuls éléments à être casuellement marqués – ceci n'est pas évident pour *THIS* et *THAT*, mais le demeure pour *he vs him*, par exemple. Si l'élément D contrôle les traits grammaticaux du syntagme dans son ensemble, il est légitime de penser que le déterminant est bien la tête du syntagme, ce qui permet de rendre compte des similarités entre déterminants et pronoms. Par ailleurs, il est à noter que le traitement de D comme tête a été étendu aux syntagmes adjectivaux tels que *it was [this big]*, où Deg – pour 'Degree' ; ici, *this* – est la tête du DegP *this big*.

Pour résumer, l'analyse de *THIS* et *THAT* – et des déterminants et pronoms en général – en tant que tête de syntagme les assimile à des opérateurs fonctionnels : c'est par ces items que sont gérés les phénomènes d'accord et la sélection des arguments. Ce traitement semble séduisant pour rendre compte du comportement syntaxique et distributionnel des ces items, mais qu'en est-il de la deixis et de l'anaphore ?

2.4. La Théorie du Liage

Dans la Théorie du Gouvernement et du Liage, la définition de l'anaphore est très restreinte ; celle-ci ne s'applique qu'aux pronoms réfléchis – *myself*, *yourself*, *himself*, etc. – et aux anaphoriques liés, en coréférence* avec un antécédent, compris comme un segment de texte. En d'autres termes, l'anaphore ne s'applique qu'aux éléments liés au sein du même domaine local[12] :

(26) *John$_i$ hurt **himself**$_i$ and **he**$_{i/j}$ sat down.*

En (26), *himself* est considéré comme étant un anaphorique, alors que *he* ne l'est pas. En effet, la coréférence entre *himself* et *John* est manifeste pour des raisons syntaxiques, alors que l'interprétation de *he* – « John » ou « quelqu'un d'autre » – dépend du contexte d'emploi.

[11] Nous pourrions objecter que les noms propres ne sont pas précédés d'un déterminant, alors même que les noms propres font partie des N en grammaire générative. Il existe néanmoins des exemples de cette sorte : *you no longer are **the Linda** that I used to know*. Nous évoquerons les spécificités des noms propres dans le chapitre IV, section 4.2.1 et dans le chapitre V, section 3.1.

[12] Le domaine local d'une expression α est le plus petit syntagme qui contient à la fois α et un sujet (Zribi-Hertz, 1996 : 117).

Cette définition de l'anaphore est directement reliée à la Théorie du Liage, qui correspond à l'un des six modules présentés par Fodor (1983) (cf. ce chapitre, section 2), module dont le but est de traiter la référence des syntagmes nominaux, et en particulier l'indexation des relations de coréférence entre NP. Ainsi, trois types d'expressions nominales sont distingués :

- Les anaphoriques : uniquement les réflexifs et anaphoriques liés, au sein du même domaine local.
- Les pronominaux : pronoms personnels, déterminants possessifs ; ils se doivent d'être libres dans leur domaine local, mais peuvent être liés en dehors de leur domaine local – par exemple, *he* en (26).
- Les expressions référentielles* ('r-expressions') : toutes les autres expressions nominales, y compris les pronoms démonstratifs ; elles se doivent d'être libres, à la fois au sein de leur domaine local et en dehors.

La Théorie du Liage est définie de la manière suivante :

- B est lié par A si et seulement si :
  - A c-commande B : A ne domine pas B et B ne domine pas A ; de plus, le nœud qui domine A domine également B.
  - A et B sont co-indexés.

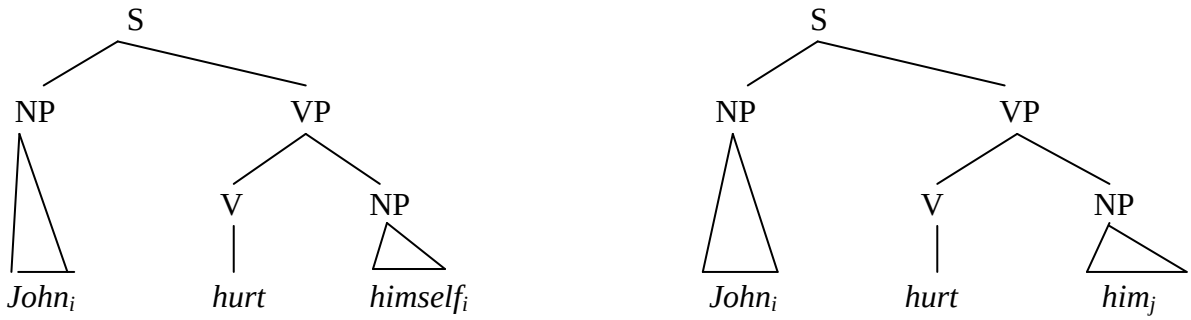
A cette définition s'ajoute une contrainte supplémentaire (Reinhart, 1983) :

- Un pronom ne peut pas c-commander son antécédent (alors que l'antécédent c-commande le pronom).
- Deux syntagmes lexicaux ne peuvent être en relation de coréférence si l'un c-commande l'autre.

Cette contrainte permet de prédire qu'une phrase telle que *John hurt himself* est acceptable, contrairement à **Himself hurt John*, puisque, dans ce dernier cas, c'est l'anaphorique qui c-commande l'antécédent. Néanmoins, seule la méthode d'indexation permet de rendre compte de la coréférence entre $John_i$ et $himself_i$ dans $John_i$ *hurt* $himself_i$, contrairement à $John_i$ et him_j dans $John_i$ *hurt* him_j, bien que ces deux phrases aient la même structure profonde. Toutefois, afin de rendre ces règles plus explicites, nous pouvons représenter les phrases *John hurt*

himself et *John hurt him* sous forme arborescente (en Théorie X-barre classique, pour ne pas complexifier la lecture) :

(27) a. *John hurt himself* b. *John hurt him*



En GG, *THIS* et *THAT* ne sont donc pas considérés comme des éléments anaphoriques, mais cette non-appartenance est due à des considérations terminologiques. *THIS* et *THAT* peuvent effectuer une reprise[13], la co-indexation libre étant permise pour tous les cas de coréférence non induite par la syntaxe. Aussi, *THIS* et *THAT* peuvent référer sans dépendre d'une expression référentielle précédente, puisque les démonstratifs sont eux-mêmes affiliés aux expressions référentielles ('r-expressions'), ce qui les rapproche des syntagmes nominaux. En revanche, l'analyse générative ne permet pas d'expliquer certains phénomènes plus complexes, tels que celui suscité par l'exemple suivant (exemple issu du corpus PAC Manchester, code locuteur emangt1, conversation libre[14]) :

(28) I : *And the sister was totally well integrated with the village, and the brother, they lived just outside, was actually, he lived there, but he didn't* <GT: ***That****'s men and women a bit isn't it, we're a bit like* ***that****.> interact at all.*

D'une part, les deux occurrences de *that* en (28) ne sont pas co-indexables avec une expression référentielle antérieure de type SN. A la rigueur, nous pourrions envisager une co-indexation entre *that* et l'ensemble de la phrase *And the sister was totally well integrated with the village, and the brother, they lived just outside, was actually, he lived there, but he didn't interact at all*, mais les antécédents propositionnels ne font pas partie des expressions

[13] Le phénomène de 'reprise' effectué par les démonstratifs, notamment la description de l'élément 'repris', sera développé dans les chapitres IV, VII et VIII. Nous gardons ici le terme vague de 'reprise' – et non pas d''anaphore', par exemple – pour ne pas faire acte de préconçu théorique à ce stade de la thèse.

[14] Le corpus PAC, ainsi que les normes d'anonymisation des locuteurs, seront présentés dans le chapitre VI.

référentielles entrant en jeu dans la co-indexation libre, en GG. De plus, et surtout, la référence opérée par chacun des deux *that* dans cet exemple dépend essentiellement de facteurs contextuels, extérieurs à la syntaxe et à la sémantique. Nous reviendrons sur ces points dans les chapitres VII et VIII. Cependant, l'analyse générative et transformationnelle apporte un éclairage nouveau sur le fonctionnement des expressions nominales ; en effet, les SN ne sont plus analysés comme une suite de constituants concaténés, mais comme ayant une structure hiérarchique, avec gouverneur et dépendant(s). Aussi, réduire les théories génératives à la simple étude de la syntaxe serait sans nul doute superficiel, puisque, depuis Jackendoff, les traits sémantiques – assignation du cas et informations lexicales – sont répercutés dans la structure profonde. Par conséquent, il en résulte une conception cognitive de la référence, la structure profonde étant la représentation formelle des règles fondamentales d'une langue donnée, ces dernières résultant d'une spécialisation de la compétence langagière selon l'environnement linguistique de l'enfant en phase de développement. Pour ce qui est des démonstratifs, l'hypothèse du déterminant comme tête de SN les assimile à des opérateurs fonctionnels, à la manière du trait flexionnel INFL. Nous verrons dans le chapitre III, section 2.4, que le mode de détermination particulier des démonstratifs est décisif dans la représentation de la structure interne du DP.

3. Le traitement des démonstratifs anglais dans les théories énonciatives

La linguistique énonciative est un espace théorique qui s'oppose aux analyses syntaxiques telles que celle proposée par la Théorie du Gouvernement et du Liage. Ce cadre épistémologique s'attache à décrire les opérations mentales qui régissent l'emploi des formes langagières. Le but n'est pas de décrire une série d'algorithmes menant à la bonne formation syntaxiques de phrases idéales, mais de décrypter le fonctionnement cognitif humain lié au langage. Deux grands courants de pensée voisins se dégagent aujourd'hui, notamment chez les linguistes français anglicistes : la linguistique énonciative héritée de la Grammaire Métaopérationnelle* d'Adamczewski et la Théorie des Opérations Prédicatives et Enonciatives* de Culioli.

3.1. La Grammaire Métaopérationnelle d'Adamczewski : un bref rappel théorique

Le modèle théorique développé par Adamczewski désire rendre compte des opérations nécessaires à la production d'un énoncé par un locuteur donné. Si la langue est considérée comme étant formée d'un ensemble fermé d'unités systémiques, l'énoncé est perçu comme

étant un résultat idiosyncrasique produit par un locuteur, dans une situation d'énonciation particulière. Ainsi, ce résultat constitue la trace visible des opérations sous-jacentes à sa production, effectuées par le locuteur à partir des éléments de la langue. Le but ultime de cette Grammaire Métaopérationnelle est de définir les traits fonctionnels invariants qui seraient propres à chaque élément de la langue. Si Adamczewski reconnaît que les structures SVO (Sujet-Verbe-Objet) sont résolument ternaires[15], il émet l'hypothèse de la structure systématique des langues comme due à un système binaire unique, qui permettrait de générer les particularités grammaticales de chaque langue. Sa réflexion part de l'observation de '*BE* + *-ing*' en anglais, traditionnellement analysé de la manière suivante :

(29) a. *I leave tomorrow*
b. *I am leaving tomorrow*

Cette structure ne permettrait pas de rendre compte de la portée de *-ing*, qui, selon Adamczewski, s'étend sur le prédicat dans son ensemble, et non sur le seul verbe. Dès lors, il propose la ré-analyse suivante :

(30) a. *I leave tomorrow*
b. *I am leaving tomorrow*

Il apparaît ainsi qu'en (30b), *tomorrow* n'a pas le même statut qu'en (30a) : dans *I leave tomorrow*, *tomorrow* appartient à ce qu'Adamczewski nomme un 'paradigme ouvert' – par exemple *on Monday, in two weeks, in a month* auraient pu se substituer à *tomorrow* – tandis que, dans *I am leaving tomorrow*, *tomorrow* relèverait d'un 'paradigme fermé' – il ne serait donc pas interchangeable – en raison de la valeur de *-ing* portant sur *leave tomorrow* et non sur *leave* seul. Si le paradigme est fermé en (30b), il en découle que « the fact that the speaker leaves tomorrow » est présupposé. La relation binaire ici extraite est celle entre *I* et *leave tomorrow*, membres reliés par la copule *BE*. Cette restructuration permettrait de rendre compte de la validité de *I was leaving tomorrow but now I won't* et de l'impossibilité de **I left tomorrow*. De plus, cette relation binaire est orientée : en (30b), la cible du prédicat en *-ing* serait « I » – expliquant ainsi que cet énoncé puisse être perçu comme une excuse provenant du locuteur – tandis qu'en (30a), la cible du prédicat serait le moment fixé par le

[15] La structure ternaire SVO se base sur la notion de fonction syntaxique et non pas de classe.

choix paradigmatique, ici, « tomorrow ». La relation en (30a) apparaît donc comme orientée à droite et est appelée 'phase 1' ou 'phase rhématique', la relation en (30b) serait orientée à gauche et appelée 'phase 2' ou 'phase thématique'. Dans sa Théorie des Phases, Adamczewski précise que thème et rhème n'ont ici rien de commun avec les définitions pragmatiques de ces termes – connus également sous le nom de 'topique'* et 'focus'* – mais qu'ils se réfèrent à des opérations de structuration de l'énoncé. La phase 2, contrairement à la phase 1, serait restreinte car opérée sur un présupposé, une anaphore, une situation particulière. Cette dichotomie ouvert/fermé, rhème/thème, phase 1/phase 2, serait applicable à l'ensemble des unités grammaticales de la langue – *a/the* ; *this/that* ; *to/-ing* ; *till/until* ; *ø, -s, -ed/do, does, did* ; *shall/will* ; *nearly/almost* ; *some/any* –, ce qu'Adamczewski nomme le Principe de Cyclicité (Adamczewski, 1995), principe à la base des analyses de *THIS* et *THAT* proposées par les linguistes français Lapaire et Rotgé.

3.2. La linguistique énonciative de Lapaire & Rotgé

3.2.1. Cadre théorique

Le cadre énonciatif dans lequel travaillent les linguistes Lapaire et Rotgé s'inspire largement des fondements de la Grammaire Métaopérationnelle d'Adamczewski, mais aussi de la Psychomécanique* de Guillaume, que nous décrirons dans le chapitre III. Ainsi, ni l'étude de la langue *per se*, ni même celle de son fonctionnement, ne sont au cœur du débat. Les 'lexèmes' (le terme de 'mot' est rejeté) sont conçus comme des symboles servant à conceptualiser le travail cognitif de l''énonciateur' (terme préféré à 'locuteur'[16] ou 'émetteur'). De là, l'objet de leurs recherches est le décryptage des opérations cognitives qui amènent l'énonciateur à utiliser, de manière systématique et normée, ces symboles langagiers :

[16] L'énonciateur est le sujet parlant, mais n'est pas forcément assimilable au locuteur ; cette différence transparaît dans le discours rapporté : ***Je*** (énonciateur) *suis sûr que Marie a dit : « **Je*** (locuteur) *n'aime pas ça ».*

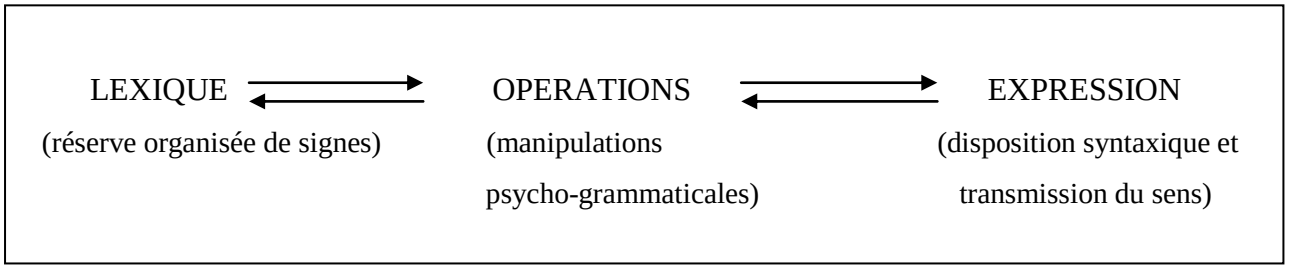


FIGURE 1 : « PAR OPERATION, NOUS ENTENDONS TOUT TRAVAIL PSYCHIQUE ET LINGUISTIQUE […] ENTREPRIS SUR LES SIGNES TELS QUE LE LEXIQUE NOUS LES PROPOSE, A DES FINS D'EXPRESSION » (LAPAIRE & ROTGE, 1992 : 32)

Cette conception de la langue – et du langage – est séduisante, en ce qu'elle est comprise comme un système de signes dont l'énonciateur dispose pour communiquer un message au co-énonciateur. Ce dernier ne sera en mesure de comprendre ce message que par l'actualisation préalable des signes lexicaux en contexte*. La langue sans contextualisation n'a pas de pouvoir communicatif, car son sens est latent. Cependant, quelques confusions possibles, à la fois terminologiques et conceptuelles (cf. ce chapitre, section 3.2.3 *infra*), pourraient hélas nuire à ce modèle, résolument cognitif et interactionniste.

3.2.2. *THIS* et *THAT* : une remise en question terminologique

Dans cette approche, l'examen critique de *THIS* et *THAT* débute par une remise en question de l'appellation catégorielle de ces items. Comme le remarquent très justement Lapaire & Rotgé, *THIS* et *THAT* ne peuvent être assimilés à des adjectifs, lorsque ceux-ci accompagnent un nom : si les adjectifs sont invariables en genre et en nombre en anglais, il n'en va pas de même pour *this* et *that* qui se déclinent en *these* et *those* au pluriel. Le terme de 'déterminant' est donc préféré à celui d''adjectif'. L'étiquette de 'pronom' est également soumise à réflexion ; en effet, 'pro-nom' sous-entendrait que *THIS* et *THAT* ne puissent se substituer qu'à un seul item lexical, bien que bon nombre d'exemples montrent que l'antécédent puisse être de forme phrastique (problème soulevé par Lyons, 1977), comme l'illustre l'exemple suivant :

(31) *Well, thank God, your poor father has been spared this disgrace.* ***That****'s all I can say*. (exemple (17) de Lapaire & Rotgé, 1992 : 50)

THIS et *THAT* pronoms sont alors renommés 'proformes'.

Afin de caractériser le fonctionnement de *THIS* et *THAT*, Lapaire et Rotgé s'appuient sur un constat terminologique, notamment à partir des grammaires de l'anglais utilisées pour l'enseignement de la langue seconde : *THIS* et *THAT* sont appelés 'démonstratifs' car ils sont utiles pour désigner un « objet » réel, souvent grâce à l'action combinée d'un geste ostensif :

(32) ***This*** *is Miss Archer*. (exemple non numéroté issu de Lapaire & Rotgé, 1992 : 53)

Bien que cette définition semble plutôt s'accorder avec celle de deixis (cf. section 1 *supra*, chapitre IV *infra*), l'étiquette même de 'démonstratif' est ici remise en question ; toutefois, les auteurs ne proposent pas d'appellation alternative. Cette fonction 'physique' de désignation du réel peut alors être transposée au texte, les démonstratifs signalant à l'auditeur/lecteur que celui-ci doit porter son attention vers un « segment de discours » (Lapaire & Rotgé, 1992 : 47), ou « fragment de texte »[17] (*op. cit.* : 53) :

(33) *I simply asked him if he was making any money. Is* ***that*** *a criticism?* (exemple (6) de Lapaire & Rotgé, 1992 : 47)

(34) *It was* ***this****: she was the most beautiful person he had ever seen.* (exemple non numéroté issu de Lapaire & Rotgé, 1992 : 53)

Cet emploi endophorique (terme emprunté à Halliday & Hasan, 1976) est qualifié d''anaphore' – exemple (33) – ou de 'cataphore'* – exemple (34) – suivant la position antérieure ou postérieure de l'antécédent textuel par rapport au démonstratif. L'emploi exophorique (terme également emprunté à Halliday & Hasan, 1976) de *THIS* et *THAT* – la faculté de référer au monde réel – serait donc un cas particulier de la monstration effectuée par les démonstratifs.

3.2.3. *THIS* et *THAT* : des éléments anaphoriques

Comme nous venons de l'évoquer, l'emploi déictique de *THIS* et *THAT*, dans ce cadre théorique, ne serait qu'un cas particulier de leurs emplois possibles. La pertinence de l'exophore ici transparaîtrait dans le fait qu'une telle faculté de monstration implique un

[17] L'alternance des termes 'segment de texte' et 'segment de discours' laisse entrevoir la confusion entre 'texte' et 'discours' ici. Nous donnerons notre définition de ces termes et argumenterons la nécessité de séparer ces notions dans le chapitre IV.

repérage par rapport à l'énonciateur (cf. section 1 *supra*). Par là même, les notions de proximité et de distance traditionnellement attribuées à *THIS* et *THAT* seraient des valeurs tout à fait relatives, puisque calculées par rapport à l'énonciateur, qui se trouve être au cœur de toutes les opérations pré-langagières décrites par la linguistique énonciative.

Néanmoins, *THIS* et *THAT* sont ici considérés comme étant des anaphoriques par essence – contrairement, par exemple, au point de vue de Lyons (1977, mais cf. également section 1 *supra*), qui argumente que l'anaphore est un procédé dérivé de la deixis. L'anaphore étant comprise comme une opération cognitive, celle-ci devrait donc se manifester par une trace précise, ancrée dans le matériel linguistique. C'est à ce stade de la réflexion qu'apparaît la notion de 'submorphème'*.

3.2.3.1. Le concept de 'submorphème' en linguistique énonciative

Selon Crystal, dans son *Dictionary of Linguistics & Phonetics* (2003), les 'submorphèmes' peuvent se définir ainsi : « a part of a morpheme that has recurrent form and meaning, such as the *sl-* beginning of *slimy, slug,* etc. » (Crystal, 2003 : 301). Si l'on s'en tient à cette définition, les 'submorphèmes' seraient des unités de sens plus petites que les morphèmes. En effet, outre *sl-*, que l'on retrouve dans *slack, slime, slop, slouch, slow* et *slug*, d'autres éléments seraient extraits de morphèmes lexicaux et pourvus d'une invariance sémantique : *sn-* (« oro-nasalité ») dans *sneeze, snore, sniff,* etc., *gl-* (« lumière ») dans *gleam, glow, gloss,* etc., *st-* (« fixité ») dans *steady, stout, sturdy,* etc...Ces 'submorphèmes' auraient donc une motivation plus iconique qu'arbitraire, ce qui remet en question, de toute évidence, l'arbitraire du signe et le rapport entre signifiant et signifié établis par Saussure. Néanmoins, si nous nous en tenons à la propriété supposée invariante de ces 'submorphèmes', il serait légitime de comprendre qu'elle s'applique à chaque item lexical ayant l'un de ces 'submorphèmes' à l'initiale. Peut-on dire que les morphèmes *snake, snail* ou *snatch* renvoient à « l'oro-nasalité » ? Que penser de la « fixité » prétendument véhiculée par les morphèmes *stab, stag* ou *stair* ?

Néanmoins, Philps (2002) nuance la définition de Crystal, en précisant que ces marqueurs sub-lexicaux n'ont pas une invariance d'ordre sémantique, mais sémiologique et notionnelle, c'est-à-dire renvoyant à un ensemble de concepts convergents. En d'autres termes, le morphème reste l'unité indivisible de sens commun. Par contre, la valeur 'notionnelle' de ces 'submorphèmes' serait attestée et invariante, à la condition que ces groupements

consonantiques soient initiaux (comme dans *sneeze, snore, gleam, glow*...), et non pas en position finale (comme *-eeze* de *sneeze*, *-ore* de *snore*, *-eam* de *gleam*...), car cette partie du mot, parfois appelée 'rime', est variable. Les 'submorphèmes' seraient des éléments systémiques dont la valeur ne pourrait être effective que lors d'une actualisation en discours, au cours d'un acte de langage effectué dans une situation de communication donnée. Il s'agirait donc d'une distinction entre sémantique et pragmatique. Pour ce faire, ces 'submorphèmes' auraient besoin d'une valeur particularisante, qui, elle, est variable : par exemple, dans *sneeze*, *sn-* est l'élément formateur, et *-eeze* est la valeur phonique particularisante ; la combinaison de ces éléments permettrait alors de former un lexème à part entière.

L'étude de la submorphémie ne porte pas que sur les lexèmes, l'existence de ce découpage submorphémique est également supposée dans les items grammaticaux, notamment dans *THIS* et *THAT*.

3.2.3.2. *THIS* et *THAT* : un microsystème

Comme nous l'évoquions dans l'introduction de cette section (3.2.3), *THIS* et *THAT* sont ici considérés comme étant intrinsèquement anaphoriques. Le point de vue adopté est que, sauf cas exceptionnel, un énonciateur donné ne peut faire usage de la monstration qu'en ayant déjà perçu l'« objet » – réel ou langagier – désigné. Pour ne pas que cette remarque puisse être qualifiée d'évidente, force est de constater que la démarche d'analyse choisie est centrée sur l'énonciateur, non pas sur le co-énonciateur ; nous pourrions objecter que le cadre énonciatif perd ici la volonté de construire un modèle interactionniste. *THIS* et *THAT* signaleraient donc, invariablement, un « déjà perçu, repéré, entendu, pensé » (Lapaire & Rotgé, 1992 : 56), codifié par le TH- initial de *THIS* et *THAT* – mais aussi des autres items grammaticaux que sont *there*, *though*, *thus*, *the*, *then*. Par là même, 'anaphorique' perd son sens textuel – distinction entre 'endophore' et 'exophore', cf. ce chapitre, section 1 – et gagne une dimension mentale par « un retour de l'esprit vers son propre passé » (Lapaire & Rotgé, 1992 : 57, d'après une citation de Guillaume, 1919 : 16). La définition cognitive de l'anaphore semble pertinente, car nous verrons dans le chapitre IV que celle-ci peut très bien se réaliser sans présence d'antécédent textuel ; néanmoins, une confusion cruciale apparaît dans cette analyse, si tant est que l'on accepte l'existence des 'submorphèmes', dotés d'une valeur de sens invariante : le préconçu cognitif est assigné au TH- initial des items grammaticaux. Si l'on suit cette analyse, la valeur notionnelle est propre à un (sub-)élément

graphique ; or, que distingue le TH- initial des mots grammaticaux du TH- initial des mots lexicaux, tels *think* ou *thick*, par exemple ? Si invariant notionnel il y a, il devrait plutôt être assigné à la consonne d'attaque /ð/ propre aux grammèmes en TH-, et non à leur particularité orthographique[18]. Cette remarque est également valable pour -IS et -AT, qui codifieraient respectivement un 'mode de pensée non clôturant' et un 'mode de pensée clôturant' (Lapaire & Rotgé, 1992 : 63), faisant écho à la distinction entre 'thème' ('phase 1') et 'rhème' ('phase 2') d'Adamczewski. En d'autres termes, -IS dans *THIS* ferait appel à un ajout d'éléments informationnels, tandis que -AT dans *THAT* marquerait un point d'orgue à ce que vient de communiquer l'énonciateur. Cette conception 'd'ouverture' *vs* 'fermeture informationnelle', volonté véhiculée par l'énonciateur grâce au recours à *THIS* ou *THAT*, est préférée à la distance relative de proximité ou d'éloignement, jugée – à juste titre par ailleurs – trop reliée à la réalité physique du monde ; la 'non-clôture' *vs* 'clôture' expliquerait en revanche la propension de *THIS* à être employé dans des cas de cataphore, contrairement à *THAT* – cf. exemples (32) et (33). Le microsystème *THIS/THAT* serait alors représenté de la manière suivante (Lapaire & Rotgé, 1992 : 63) :

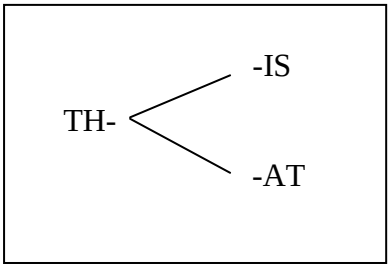


FIGURE 2 : LE MICROSYSTEME *THIS/THAT* (LAPAIRE & ROTGE, 1992 : 63)

La démarche énonciative est séduisante par plusieurs aspects, notamment la prise en compte du sujet énonciateur dans l'analyse des formes langagières et l'idée d'un système linguistique qui ne prend effet que lors d'un ancrage contextuel. En revanche, une distinction entre 'texte'* – compris comme une séquence de signes verbaux et non verbaux – et 'discours' – une représentation conceptuelle de la référence – pourrait apporter de sérieux avantages à ce modèle théorique. Nous développerons cette perspective dans le chapitre IV.

[18] A l'origine, les submorphèmes portaient le nom de 'phonesthèmes', terme introduit par Firth (1930). Selon nous, cette appellation serait plus judicieuse quant à la potentielle iconicité d'éléments récurrents dans le lexique.

De plus, la définition d'invariant notionnel propre à quelque marqueur linguistique devrait très certainement être abordée avec la plus grande précaution (cf. chapitre III, section 3.1.3).

3.3. La Théorie des Opérations Prédicatives et Enonciatives de Culioli

3.3.1. Rappels théoriques

A travers sa Théorie des Opérations Prédicatives et Enonciatives (TOPE), Culioli cherche à développer un modèle théorique de l'énonciation, prenant en considération l'activité de langage à travers la diversité des langues, des textes, et des situations. L'énoncé n'est pas ici conçu comme le résultat d'un acte de langage idiosyncrasique, produit par un quelconque énonciateur dans une situation d'énonciation donnée, mais plutôt comme une construction du sens. Cet agencement est effectué grâce à trois niveaux de représentation logique :

- Le niveau ou domaine notionnel (niveau 1) : il s'agit d'un ensemble constitué de concepts – ou 'notions'*, notées entre barres obliques – dotés de particularités physiques, psychiques, mais aussi culturelles, qui sont d'ordre cognitif, et antérieurs à toute catégorisation linguistique. Ces concepts sont mis en relation suivant un axe à trois places (π, C_0, C_1), nommé 'lexis'*. Les représentations mentales à ce niveau ne sont pas directement accessibles, puisque psychiques, et ne peuvent être perçues qu'au travers des activités humaines, dont l'activité langagière.
- Le niveau prédicatif (niveau 2) : les notions sont organisées – ou 'orientées' – afin de former une 'relation prédicative'* : un rôle d'actant est assigné à C_0 et à C_1, par rapport au noyau prédicatif π qui relie les deux arguments obtenus. Les représentations sont ici d'ordre textuel, le texte étant compris comme une représentation des représentations de niveau 1. Il s'agit donc ici de représentations de représentations, que le linguiste peut analyser car elles constituent des données observables.
- Le niveau énonciatif (niveau 3) : l'ancrage spatio-temporel T_0 et l'énonciateur S_0 considérés permettent l'actualisation de la relation prédicative obtenue, grâce à l'action combinée du sujet syntaxique S_2 et de la situation spatio-temporelle T_2 à laquelle fait référence l'énoncé – T_2 et S_2 peuvent être différents de T_0 et S_0. Les représentations de niveau 3 sont d'ordre métalinguistique, construites par le linguiste afin de décrypter les opérations prédicatives et énonciatives.

Ainsi, les données observables constituent la trace des opérations mentales antérieures à la formulation langagière. Le but de ce modèle théorique est alors de décoder et modéliser ces opérations, grâce à un système de représentation métalinguistique.

3.3.2. La forme schématique* dans la TOPE

L'un des points centraux de la TOPE est la conception de forme schématique (FS). La FS est une solution apportée aux diverses critiques reçues par l'idée d'un invariant de sens tel qu'il est défini par les théories énonciatives – cf. les théories décrites en sections 3.1 et 3.2 *supra*. La FS serait propre à chaque terme linguistique – notons ici que certains linguistes, comme De Vogüé, se refusent à allouer une FS aux items grammaticaux – et se définirait par un schéma immuable de relations que le terme considéré entretient avec son co-texte* – 'contexte'[19] chez Culioli – et un schéma de variation interne correspondant aux différentes facettes de l'item. Il n'est plus question d'invariance notionnelle, mais d'un mode de fonctionnement propre à chaque terme. Nous pourrions cependant objecter que ni l'invariant de sens ni la FS ne semblent applicables à l'analyse diachronique ; nous reviendrons sur ce point dans le chapitre III, section 3.

La forme schématique est centrée sur la notion de 'repérage'* : tout terme est 'repéré' par rapport à un autre terme, qui lui sert de 'repère'. Trois relations peuvent unir les deux expressions considérées :

- L'identification (notée =), où le terme repéré est considéré identique au repère.
- La différenciation (notée ≠), où les deux termes considérés sont distincts.
- La rupture (notée ω), où le terme repéré n'est ni identifiable au repère, ni différenciable de celui-ci.

Le repérage est particulièrement manifeste dans la construction de la temporalité, où *Wait a minute, he is working* exemplifie l'identification, *Damn ! I have lost my keys* illustre la différenciation, et *Mary left at 5 last Monday* relève de la rupture.

[19] 'Contexte' compris comme 'co-texte' chez Culioli rejoint l'assimilation entre 'texte' et 'discours' effectuée par Lapaire et Rotgé (cf. note 17).

Pour des raisons de clarté et de concision, nous n'entrerons pas plus en détail dans la description des FS potentielles. A présent, nous allons nous focaliser sur le traitement des démonstratifs dans la TOPE.

3.3.3. La forme schématique des démonstratifs

Trois opérations de repérage sont pertinentes dans l'analyse des démonstratifs, au sein de la TOPE :

- 'L'extraction'*, opérée de manière prototypique par l'article indéfini *a/an*, et que nous pourrions définir par la sélection d'une 'occurrence'*[20] parmi une 'classe d'occurrences'.
- Le 'fléchage'* : généralement considéré comme étant une opération seconde, car suivant l'extraction, le fléchage pourrait se définir par le renvoi d'une occurrence à une autre occurrence. Culioli ajoute que le fléchage implique co-identité entre les deux occurrences considérées (Culioli, 1990 : 182), mais cette définition ne semblerait s'appliquer qu'au fléchage contextuel que l'on pourrait assimiler à l'endophore (cf. ce chapitre, section 1). L'opération de fléchage est divisée en deux catégories, le 'fléchage contextuel' et le 'fléchage situationnel', elles-mêmes définies au sens étroit et au sens large :

    - Le 'fléchage contextuel étroit' manifeste une anaphore avec reprise lexicale : *I met **a woman*** (opération d'extraction) *and her daughter yesterday.**The woman*** (opération de fléchage) *told me she was looking for a good place to eat.*
    - Le 'fléchage contextuel large' définit une anaphore par inférence, où le fléchage passe par un 'préconstruit', c'est-à-dire un ensemble de valeurs

[20] Occurrence : Exemplaire, représentant, incarnation d'une notion. Les occurrences sont à la fois d'ordre référentiel et d'ordre linguistique. "Croquis", "dessin", "image", "diagramme", etc., mais aussi "dessiner", "croquer", "crayonner", "reproduire", etc., de même que "dessinateur", "illustrateur", "caricaturiste", etc. sont des occurrences linguistiques de la même notion /dessin/. Dans *il y a une mouche dans la soupe*, on pose en outre l'existence d'une occurrence référentielle de /mouche/ et d'une occurrence référentielle de /soupe/. L'occurrence peut alors être considérée comme un événement énonciatif qui opère une double délimitation sur la notion, une délimitation quantitative qui concerne son ancrage spatio-temporel, son existence, et une délimitation qualitative qui a trait à ses propriétés, sa nature. Les occurrences associées à une même notion constituent une classe d'occurrences qui correspond à l'extension de la notion en question. (*The SIL French/English Linguistic Glossary* : http://www.sil.org/linguistics/glossary_fe/defs/TOEFr.asp#Occurrence)

référentielles que le co-énonciateur doit déduire car celles-ci ne sont pas textuellement exprimées : *I visited **a small village** a few days ago. **The church** was beautiful*. (le co-énonciateur est à même d'identifier « l'église » car il est admis que tout village est construit autour d'une église).

- Le 'fléchage situationnel étroit' – ou 'fléchage déictique' – est situé par rapport à la situation d'énonciation directe, c'est-à-dire celle où se trouve l'énonciateur : *can you pass me **the salt** please ?* (« le sel qui est sur la table » au moment où l'énoncé est produit). Selon cette définition, l'occurrence construirait son propre repère par rapport à elle-même.
- Le 'fléchage situationnel large' est situé par rapport à la situation d'énonciation appartenant à l'univers socio-culturel communément admis par l'énonciateur et le co-énonciateur – savoir encyclopédique et savoir partagé[21] : ***The Prime Minister** had a televised talk yesterday*.

Selon toute vraisemblance, le 'fléchage contextuel' présenterait un rapport de dépendance référentielle entre deux occurrences matérialisées dans le texte, tandis que le 'fléchage situationnel' serait une opération de construction du sens à partir de l'extralinguistique. Ainsi, le fléchage contextuel pourrait s'apparenter à l'endophore et le fléchage situationnel à l'exophore (cf. ce chapitre, section 1). Néanmoins, l'exophore étant traditionnellement liée à une action ostensive, il conviendrait de nuancer l'appartenance du fléchage situationnel large à cette catégorie. Concernant le mode de fonctionnement des démonstratifs, les avis semblent partagés ; en effet, selon Groussier, Groussier & Chantefort (1975), les démonstratifs seraient réservés au fléchage déictique, tandis que le fléchage situationnel large ne pourrait être indiqué que par le déterminant zéro ou le défini *the*. Le fléchage contextuel, étroit ou large, pourrait être indifféremment marqué par *the*, *THIS* ou *THAT*. En d'autres termes, *THIS* et *THAT* seraient des déictiques, leur sens n'étant accessible que par la situation de communication directe, mais pourraient également être anaphoriques. Le défini *the* permettrait au co-énonciateur d'accéder au sens à partir d'informations que ce dernier a en mémoire*, sans passer par le support linguistique au préalable, mais pourrait aussi être anaphorique. D'après un article de Giancarli (2004 : 3), certains linguistes excluent les démonstratifs des opérateurs de fléchage et les cantonneraient à la deixis ; d'autres, tels que

[21] Notons ici que le recours au savoir encyclopédique ou au savoir partagé n'est pas nécessaire au seul fléchage situationnel large, puisqu'il intervient également dans l'opération de fléchage contextuel large.

Danon-Boileau (1983, 1984), seraient plus mesurés et conféreraient à *that* une marque de fléchage, contrairement à *this* uniquement déictique. La réflexion de Danon-Boileau se dégage d'exemples tels que (Danon-Boileau, 1984 : 34) :

(35) *Give me **that** ashtray on the table there.*

où *that* serait textuellement repéré par la propriété <*be on the table there*>. Naturellement, nous pourrions objecter que *that* peut construire un référent par simple ostension :

(36) *Give me **that** ashtray please.*

En (36), le co-énonciateur ne dispose que de la situation d'énonciation pour identifier « le cendrier quémandé par l'énonciateur », auquel cas *that* n'aurait rien d'anaphorique ici. Nous pouvons entrevoir les problèmes qu'est susceptible de poser l'attribution d'une forme schématique invariante aux items langagiers : cette forme doit pouvoir rendre compte de tous les emplois possibles observés en discours, tout en étant la plus épurée possible. Or, il pourrait paraître difficile d'assigner un mode de fonctionnement unique à *THIS*, à *THAT* et à *the* si leurs emplois en discours se recoupent. De plus, l'opération de fléchage semblerait regrouper sous une seule et même appellation des phénomènes aussi divers que l'anaphore avec coréférence, l'anaphore avec inférence, l'ostension directe et le recours au savoir encyclopédique ou partagé. Enfin, il convient de garder à l'esprit que le fléchage est décrit comme étant une opération seconde, basée sur une extraction préalable. Si le fléchage opère sur un déjà opéré, le fléchage est par définition anaphorique. L'appartenance du fléchage déictique à cette catégorie paraît donc bien discutable ; qui plus est, l'appellation de 'fléchage déictique' est alors un oxymore. Une réponse apportée à l'une de ces observations, à savoir celle de *that* pouvant être anaphorique ou déictique, est que le fléchage ne serait pas une opération seconde reliant deux termes linguistiques, et donc appartenant au niveau 2 (Giancarli, 2004 : 7), mais une opération double : l'extraction serait réalisée mentalement par l'énonciateur, sans que celle-ci n'apparaisse sous forme linguistique. Reste toutefois le problème de *THIS*, considéré comme étant invariablement déictique, malgré des exemples tels que *No I've been back to college, <F: Been back to college.> I've got my, my qualifications for the job that I'm doing and I think I'll do **this job** now I've, because I do like it* (exemple

issu du corpus PAC Lancashire, code locuteur elalb1, conversation guidée[22]), qui ne relève clairement pas de la deixis.

Le chapitre II offre un panorama, certes succinct, des analyses des démonstratifs anglais suivant trois courants théoriques majeurs. Même si les traitements proposés peuvent s'avérer insuffisants, ils permettent de dégager certaines propriétés distributionnelles, syntaxiques et référentielles des marqueurs indexicaux que sont *THIS* et *THAT*. Néanmoins, et ce malgré la grande diversité théorique parcourue, nous pensons que le problème commun à ces modèles est de ne pas séparer les niveaux de représentation – ce qui fait partie du système syntactico-sémantique de la langue et ce qui résulte d'une actualisation en discours –, problème posé par l'indexation post-analyse syntaxique en structure profonde en GG, par l'intégration du sens à la forme dans certaines écoles énonciatives. Or, nous pensons et arguerons dans le chapitre IV, notamment, que, bien qu'interdépendants, les 'niveaux' linguistiques ne doivent pas être amalgamés. Le chapitre suivant s'attache à décrire, suivant le même schéma que le chapitre II, les démonstratifs en français.

[22] Cf. note 14.

Chapitre III. Les démonstratifs en français : analyses structurale, générative et énonciatives

Selon l'inventaire proposé par les grammaires de référence, telles que la huitième édition de la Grammaire de l'Académie Française (1932), les pronoms et 'adjectifs'[23] démonstratifs au XX^{ème} siècle se composent chacun de formes simples et de formes dites 'composées'.

Concernant les pronoms démonstratifs, les formes simples à l'écrit sont *celui*, *celle*, *ceux*, *celles* et *ce* invariable. Ces items ne sont pas autonomes, mais suivis d'un 'pronom relatif'[24] – *qui*, *que* ou *dont* – ou d'un complément nominal précédé de la préposition *de* ; en revanche, l'une de ces occurrences suivie d'un adjectif qualificatif ou d'un participe est jugée comme étant une incorrection. Ces formes pronominales simples servent à « désigner sans la nommer la personne ou la chose dont il est question » (Grammaire de l'Académie Française, 1932 : 53-54). Les formes composées – *celui-ci/là, celle-ci/là, ceux-ci/là, celles-ci/là* et *ceci/cela* – sont complémentaires des formes simples, en ce qu'elles apparaissent dans toute autre position. Trois particularités corollaires de ces items sont proposées (1932 : 54-55) :

- Un pronom démonstratif composé « désigne une personne ou une chose plus fortement que ne le ferait le simple pronom personnel *le, la, les, lui* ». Il suffirait de comparer *J'ai choisi* ***celui-là*** *pour confident* et *Je* ***l'****ai pris pour confident*.
- Un pronom démonstratif composé « peut servir d'antécédent au pronom relatif quand on veut attirer fortement l'attention sur la personne ou la chose en question », mais la norme voudrait « que la proposition relative soit rejetée après la proposition principale » : ***Celui-là*** *mérite la mort qui trahit son pays*.
- *Celui-ci* et *celui-là*, ainsi que leurs dérivés féminins et pluriels, employés de manière connexe, sont utiles pour « distinguer, à opposer nettement, deux êtres ou deux choses,

[23] Dans cette huitième édition de la Grammaire de l'Académie Française, les 'adjectifs' sont séparés en 'adjectifs déterminatifs' et 'adjectifs qualificatifs' : « les adjectifs déterminatifs désignent le nom d'une façon plus ou moins précise. [Ils sont] appelés quelque fois pronominaux, parce que beaucoup d'entre eux ressemblent à des pronoms et peuvent être employés comme pronoms » (1932 : 67). Chaque occurrence des termes 'adjectif démonstratif' dans cette thèse est en référence aux grammaires traditionnelles ou aux ouvrages qui utilisent cette terminologie.

[24] Il serait peut-être plus juste de spécifier que *celui, celle(s)* et *ceux* peuvent être suivis d'une proposition relative ; nous dressons ici un inventaire tel qu'il peut apparaître dans les grammaires de référence et, par conséquent, ne modifions pas la terminologie employée.

ou deux groupes d'êtres ou deux groupes de choses » : *De ces deux tableaux, **celui-ci** est médiocre, **celui-là** est parfait*. De plus, ce contraste peut se traduire par une opposition spatiale ou temporelle, *celui-ci* désignant par exemple « le dernier objet dont on vient de parler ou le plus voisin, *celui-là* désignant l'objet dont on a parlé en premier lieu ou le plus éloigné » : *Les deux règles ont le même sens ; puisque vous n'admettez pas celle-là, écoutez celle-ci*.

Ce, quant à lui, « désigne une chose dont il vient d'être question ou une proposition déjà exprimée ». Cette forme simple ne peut qu'être suivie d'un « pronom relatif, de la conjonction *que*, ou d'une des formes du verbe *être* ». *Cela*, devenu *ça* en registre familier, est le complémentaire de *ce*, en ce qu'il apparaît dans une toute autre position ; néanmoins, *ce* sera conservé dans certaines tournures anciennes : *pour ce faire, ce me semble, sur ce*. Il est également noté que « *ce* sert souvent de sujet apparent » : *cette fuite, c'était une feinte* (1932 : 55). Les formes composées *ceci* et *cela* en emploi connexe jouent le même rôle contrastif que *celui-ci* et *celui-là* ; *ceci* « désigne plus particulièrement ce qui va être dit ou ce qui va advenir », tandis que *cela* « désigne plus souvent ce qui est déjà arrivé ou ce qui a été dit » (1932 : 56).

En ce qui concerne les 'adjectifs' démonstratifs, leur rôle est de « détermine[r] le nom en désignant, comme s'il[s] le montrai[en]t, la personne ou la chose en question » (1932 : 75). Les 'adjectifs' démonstratifs simples sont *ce*, devenu *cet* devant une voyelle ou un « h » muet, *cette* et *ces*. Les formes composées sont « *ce/cet + N-ci* », « *ce/cet + N-là* », « *cette + N-ci* », « *cette + N-là* », « *ces + N-ci* », « *ces + N-là* ». Les 'adverbes *ci* et *là*' sont toujours séparés de l''adjectif' démonstratif par le nom qu'il détermine, et un trait d'union est requis entre le nom et l'adverbe. A la manière de *celui-ci/celui-là* et de *ceci/cela*, deux formes composées conjointement employées peuvent créer un contraste entre deux personnes ou deux choses ; de plus, il est noté que l''adjectif' démonstratif peut produire un effet emphatique ou péjoratif (1932 : 75-76).

Cet inventaire normatif proposé par la Grammaire de l'Académie Française esquisse les démonstratifs comme ayant la faculté de 'montrer' et de 'sélectionner' un référent parmi plusieurs potentiels. Par ailleurs, nous verrons dans la section 1 de ce chapitre que deux des grammaires majeures du français – celle de Chevalier, Blanche-Benveniste, Arrivé & Peytard (1964) et celle de Dubois (1965) – reprennent et développent les idées amorcées dans la

Grammaire de 1932. Les sections 2 et 3 illustrent le traitement des démonstratifs proposé par les théories génératives tout d'abord, énonciatives ensuite.

Comme nous allons le constater dans la section 1, les deux ouvrages-phares de l'analyse grammaticale du français publiés dans les années soixante, la *Grammaire Larousse du français contemporain* de Chevalier *et al.* (1964) et la *Grammaire structurale du français* de Dubois (1965) reprennent et approfondissent la doxa présentée par l'Académie Française en 1932. Les deux classes de démonstratifs, déterminants et pronoms, sont analysées suivant leurs propriétés distributionnelles et quelques propriétés sémantiques sont esquissées. En particulier, il est implicitement reconnu que les pronoms composés en *-ci* et *-là* réfèrent à des entités discrètes, alors que *ceci*, *cela* et *ça* ont tendance à désigner des événements ou des contenus propositionnels.

1. L'analyse structurale des démonstratifs français

D'aucuns ne sauraient aborder le système des démonstratifs français sans s'intéresser à l'analyse structurale qu'en fait Dubois (1965), qui s'appuie fondamentalement sur la notion de marque. Selon lui, les démonstratifs constituent l'ensemble marqué correspondant à celui non marqué formé par les éléments *le*, *la*, *les*, auxquels s'ajoutent les pronoms de troisième personne ; en effet, un démonstratif pourrait toujours, en principe, se substituer aux articles et aux pronoms personnels. L'inverse n'étant pas avéré, ce test de remplacement justifierait la qualification de marqué *vs* non marqué.

1.1. Le rôle syntagmatique des démonstratifs

La particularité de cet ensemble marqué serait d'avoir un double statut syntagmatique : selon les termes de Dubois, les démonstratifs peuvent être 'constituants de syntagmes' et 'constituants de phrases', appellations qui désignent sans nul doute les déterminants – qu'il nomme 'adjectifs'[25] – et pronoms démonstratifs. Ainsi, *CELUI-ci* et *CELUI-là* ont la même distribution que les pronoms de troisième personne, sont marqués en genre et en nombre, et

[25] Le terme d''adjectif démonstratif' est largement employé dans les grammaires encyclopédiques du français, mais aussi de l'anglais ; par ailleurs, 'adjectif' est également usité chez certains linguistes contemporains, tels que Kleiber. Cependant, les déterminants démonstratifs ne font jamais l'objet d'une analyse assimilée à celle des adjectifs en tant que modifieurs de groupe ou syntagme nominal, puisque leur rôle est décrit comme étant déterminatif. Bien que la terminologie soit la même, les éléments que nous appelons 'déterminants' d'une part, et 'adjectifs' d'autre part, appartiennent à des catégories distinctes.

« se réfèrent à la situation ou au syntagme nominal le plus proche » (1965 : 160) ; en d'autres termes, *CELUI-ci* et *CELUI-là* pourraient être employés en première mention, ou bien en tant que rappel d'une expression nominale précédente, cas illustrés par les exemples suivants :

(1) *C'est* ***celui-ci*** *le plus grand.*
(2) *Julie s'immobilisa, un plat à la main, déposa* ***celui-ci****.* (exemples (1) et (2) issus de Chevalier *et al.*, 1964 : 242)
(3) *J'ai visité plusieurs appartements ;* ***celui-ci*** *me paraît le plus agréable.*
(4) *De tous ces romans,* ***celui-là*** *est le plus intéressant.* (exemples (3) et (4) issus de Dubois, 1965 : 160)

Par déduction, nous pouvons supposer que « se référe[r] à la situation » et « [se référer] au syntagme nominal le plus proche » correspondent peu ou prou à l'emploi déictique et anaphorique des démonstratifs. Dans leur grammaire publiée presque au même moment que l'ouvrage de Dubois, Chevalier *et al.* choisissent respectivement les appellations de 'pronoms représentants' et 'pronoms nominaux' (1964 : 244)[26]. De plus, Chevalier *et al.*, à la différence de Dubois, ajoutent que les pronoms *celui-ci* et *celui-là,* tels qu'employés dans les exemples (3) et (4), ont un mode de désignation particulier : « [la détermination] est fournie à la fois par cette présence – *i.e.* la présence antérieure d'un « substantif » – et par les circonstances extralinguistiques. Dans ce cas, substantif évoqué et pronom évocateur s'appliquent chacun à une quantité différente de la notion désignée » (Chevalier *et al.*, 1964 : 242). Plusieurs caractéristiques majeures des démonstratifs sont ici esquissées, même si elles se doivent de rester distinctes :

- la référence des démonstratifs peut dépendre à la fois du co-texte – la matière langagière, que Chevalier *et al.* nomment 'contexte' – et du contexte situationnel – ici appelé 'extralinguistique' ;
- les démonstratifs sont à même de désigner une entité particulière au sein d'une classe de référents ayant les mêmes propriétés ;

[26] La *Grammaire Larousse du français contemporain* de Chevalier *et al.* (1964) occupe une place importante dans ce retour historique. Cet ouvrage, qui a eu une très grande influence dans les années soixante, est l'un des premiers manuels de référence à s'inspirer « des méthodes distributionnelles et génératives » (Préface, 1964 : 5). Nous entrelaçons notre présentation de références à leur traitement qu'on peut considérer comme complémentaire à celui de Dubois (1965).

- le SN en première mention et le démonstratif employé dans un phénomène de reprise ne sont pas nécessairement coréférentiels ;
- les démonstratifs pourraient être à la fois anaphoriques – puisque la référence du pronom se base sur la présence antérieure d'une expression lexicale – et déictiques – puisqu'il y désignation d'un nouvel objet discursif[27].

Ces définitions sont évidemment trop simplistes, mais nous pouvons attribuer à ces grammaires le mérite d'avancer que les démonstratifs ont plus d'un emploi discursif. Rappelons cependant qu'un pronom ne peut « référer à un syntagme nominal », mais plutôt au référent produit par ce SN. Nous aurons bien évidemment l'occasion de développer ces différents niveaux d'analyse dans la suite de cette thèse. Dubois ajoute que *ce* et *ça*, dans les exemples (5) et (6), sont des 'constituants de phrases' :

(5) ***Ça*** *me paraît juste, mais il faudrait nuancer cette affirmation.*
(6) ***Ce*** *que vous dites me semble tout à fait justifié.*

Ce et *ça* sont dits substituables à des 'groupes de segments' ; nous pouvons supposer que, par ces termes, Dubois suggère que *ce* et *ça* réfèrent moins à des objets discrets qu'à des événements ou des faits. Chevalier *et al.* se rallieraient sans doute à cette hypothèse car ils avancent que *ceci*, *cela* et *ça* « désignent n'importe quelle notion ou fait » (souligné par nous) et « représentent le contenu notionnel de la subordonnée ou de la phrase qui précède » (1964 : 242-243).

En ce qui concerne les démonstratifs 'constituants de syntagmes', Dubois note que ces derniers ont la possibilité de « référe[r] à un segment précédent ou anticipe[r] sur un segment suivant », et « que la référence se fait souvent à la situation ». Une fois de plus, la terminologie est discutable – un démonstratif ne 'réfère' pas à un segment de texte, et la référence, de toute évidence, se fait grâce à l'emploi du syntagme dans son ensemble – mais nous devinons ici l'emploi anaphorique ou déictique des syntagmes démonstratifs. A l'inverse de Dubois qui attribue la référence du syntagme au seul démonstratif, Chevalier *et al.* restreignent le rôle des déterminants démonstratifs à celui de « présenter un substantif, ou l'équivalent d'un substantif, déterminé soit par les circonstances extralinguistiques, soit par le

[27] Si l'on accepte la conception qu'a Lyons (1968, 1977) de la deixis, sur laquelle nous reviendrons à plusieurs reprises dans cette thèse.

contexte » (1964 : 241). En d'autres termes, la référence se ferait grâce au seul item lexical tour à tour associé à un geste désignatif, une expression langagière précédente, une proposition relative ou un participe ; le rôle désignatif des déterminants démonstratifs reste donc bien incertain. Pour en revenir à Dubois, ce dernier – qui accorde une place importante à l'oral – note que, bien que distincts d'un point de vue catégoriel, les déterminants et pronoms démonstratifs ont en commun l'élément *ce* [sə]. Pour résumer, voici la classification des démonstratifs qu'il propose (1965 : 160) :

Adjectifs (constituants de syntagmes)		
Ce *Cet*	*Cette*	*Ces*
Pronoms (constituants de phrases)		
Ce *C'* *Ça* *Ceci* *Cela*	*Celui* *Celui-là*	*Celle* *Celle-là*
Indices		
Ci *Là*	*Ici* *Là-bas* *Là-dedans* *Là-dessus* *Là-haut*	*Voici* *Voilà*

TABLEAU 1 : LA CLASSIFICATION DES DEMONSTRATIFS SELON DUBOIS (1965 : 160)

1.2. La structure morphologique des démonstratifs

Par cette taxinomie, Dubois laisse transparaître une perspective de son analyse, originale et à notre avis très pertinente : la prise en compte de toutes les formes orales possibles, y compris l'apparition du [t] final lorsque *ce* est suivi d'un nom commençant par un phonème vocalique et la forme contractée *c'* de *ce*/*ça*. Ainsi, Dubois note que la structure morphologique des démonstratifs est construite à partir de celle des éléments non marqués auxquels ils correspondent, puisque la morphologie des démonstratifs se base sur la forme disjointe du personnel à laquelle est ajouté l''indice [s]', et éventuellement la marque [si] ou [la] (1965 : 162) :

(7) *[səlɥi]/[lɥi]* *[səlɥisi]/[səlɥila]* *[sɛl]/[ɛl]* *[sɛlsi]/[sɛlla]*

Celui/lui *celui-ci/celui-là* *celle/elle* *celle-ci/celle-là*

[sø]/[ø] *[søsi]/[søla]*

Ceux/eux *ceux-ci/ceux-là*

Si Dubois note 'indice [s]' ce qui pourrait être 'indice [sə]', c'est pour rendre compte de la variation phonologique que nous pouvons constater à l'oral, qui correspond à une élision du schwa* :

(8) *[slɥi]/[slɥisi]/[slɥila]*

De même, Dubois mentionne que les formes *ce*, *ceci* et *cela* ne sont constituées que des 'indices' [s], [si] et [la], ce dernier indice étant réduit au son vocalique [a] dans *ça* – « [s] + [a] ». Les déterminants – 'adjectifs' dans sa terminologie – sont eux, analogues aux articles :

(9) Singulier : *[lə] le* *[sə] ce*
Pluriel : *[lɛ] les* *[sɛ] ces*

Par ailleurs, Dubois mentionne également l'emploi marqué de la forme [sɛt] au lieu de [sə] lorsque le mot suivant commence par un son vocalique. Par conséquent, le point commun entre déterminants et pronoms démonstratifs est l''indice' [s]/[sə]. Même si Dubois ne l'énonce pas explicitement, il laisse penser que cet 'indice' commun serait la 'trace', pour reprendre le terme énonciativiste évoqué au chapitre II, de la monstration. Cette découpe n'est d'ailleurs pas sans rappeler la submorphémie opérée par les linguistes travaillant dans ce cadre théorique, à la différence que l''indice' est une conception phonétique, contrairement au submorphème qui, dans certains travaux, semble parfois plus graphique que phonique. Les 'indices' [si] et [la] sont quant à eux optionnels, à la fois pour les déterminants et les pronoms ; [si] et [la] – ce dernier étant réduit à [a] pour *ça* – ne sont pas tant décrits en termes de proximité ou de distance que d'emploi marqué, [la] étant la valeur par défaut de la paire [si]/[la][28].

[28] Dubois se base sur une étude quantitative de *ci* – 23 occurrences – et de *là* – 1468 occurrences – pour affirmer que *ci* est la forme marquée de la paire *ci/là* ; néanmoins, il ne mentionne aucune source quant à l'origine de son corpus « du français fondamental » (1965 : 161). Chevalier *et al.* (1964 : 243) adoptent le même point de vue, bien que la notion de forme marquée ne soit pas explicite, sans toutefois abandonner le critère de la distance pour distinguer *ci* et *là*.

1.3. L'opposition articles/pronoms personnels et démonstratifs

Selon Dubois, les démonstratifs sont les éléments marqués correspondant aux articles et pronoms personnels, eux non marqués. En effet, comme nous l'avons évoqué en introduction de cette section 1, les démonstratifs peuvent en principe se substituer à toutes les occurrences d'articles et pronoms personnels, mais le remplacement inverse n'est pas immuable. Cette non-réciprocité commutative aurait une raison sémantico-pragmatique : les démonstratifs permettent un supplément d'information que les éléments non marqués sont incapables de produire. Cette propriété autoriserait les démonstratifs à modifier la structure informationnelle de l'énoncé par la sélection référentielle qu'ils opèrent (exemple (10)), mais aussi à véhiculer une modalisation affective (exemples (11a, b, c)) :

(10) *Le fils de mon voisin est rentré de vacances ; **il** est très content*/***celui-ci** est très content.*

(11) a. ***Il** fait le fier*/***ça** fait le fier*.
b. ***Il** est en retard*/***celui-là**, encore en retard !*
c. ***Cet** enfant est insupportable*. (exemples issus de Dubois, 1965 : 162)

Ainsi, en (10), *il* pourrait référer au « fils du voisin » ou au « voisin », alors que *celui-ci* sélectionne « le voisin » comme référent. En (11a, b, c), l'emploi d'un démonstratif accompagne la référence d'un jugement de valeur, péjoratif dans les cas présents. Ces observations de Dubois sont tout à fait pertinentes, même s'il ne développe pas plus son idée ; nous aurons l'occasion d'étudier en détail la référence opérée par les démonstratifs – *ça*, en particulier – et l'incidence qu'a celle-ci sur l'articulation du discours dans le chapitre VIII.

1.4. La distribution des formes pronominales

Comme il est mentionné dans la Grammaire de l'Académie Française (1932), les démonstratifs comprennent des formes simples et des formes composées. Les propriétés distributionnelles de ces formes sont ici observées, analyse dont se dégage une complémentarité de ces items.

1.4.1. *Celui, celle(s)* et *ceux*

Les formes *celui, celle(s)* et *ceux* ne sont pas autonomes, mais requièrent une extension de types variables. L'apport sémantique de l'expansion considérée est indispensable à l'identification du référent désigné (exemples suivants issus de Dubois, 1965 : 164) :

- Une extension de type syntagme prépositionnel : *nous prendrons le train de cinq heures, **celui de six heures** nous retarderait trop ; la concierge est très aimable, mais **celle d'avant** était insupportable ; ses sentiments ne sont pas **ceux d'un ingrat**.*
- Une extension de type proposition relative : ***celui qui parle** sera puni ; sa passion pour la chasse égale **celle qu'il a pour le jeu**.*
- Une extension de type syntagme adjectival : *je n'ai pas lu son dernier roman, mais **celui écrit par l'auteur l'année dernière** ; cette inondation n'est pas aussi grave que **celle provoquée l'année dernière par le Rhône**.*

L'inventaire des diverses expansions demandées par *celui, celle(s)* et *ceux* mène à une conclusion évidente : ces éléments correspondent à ce que la grammaire traditionnelle nomme des compléments du nom. Par cette constatation, il est légitime de questionner le statut catégoriel de *celui, celle(s)* et *ceux* : sont-ce des pronoms ? Des noms ? Des syntagmes nominaux à tête lexicale vide ? De toute évidence, ils ne peuvent être assimilés à *CELUI-ci* et *CELUI-là*, qui, eux, fonctionnent de manière autonome ; de même, ils ne peuvent se confondre avec des déterminants, de par l'absence de tête lexicale. Une solution envisageable serait de les traiter comme des éléments spécifiques, à mi-chemin entre pronoms et déterminants démonstratifs, ayant un mode de fonctionnement similaire à *the one* en anglais – par exemple, *celui que j'ai vu* pourrait se traduire par 'the one who/that/ø I saw'. Par ailleurs, et même par conséquent, *celui, celle(s)* et *ceux* sont en distribution complémentaire avec *CELUI-ci* et *CELUI-là* ; ce que Dubois nomme l''indice' [si] et l''indice' [la] pourrait donc non seulement changer les propriétés distributionnelles de *CELUI*, mais également son statut catégoriel. Nous reviendrons sur *celui, celle(s)* et *ceux* dans les chapitres VIII et IX.

1.4.2. *Ce*

Les distributions de *ce* ne recouvrent pas celles des éléments que nous venons de présenter, mais *ce* partage avec *celui*, *celle(s)* et *ceux* la possibilité d'accepter une proposition relative en guise de complément (exemples suivants issus de Dubois, 1965 : 165) :

(12) *Il n'a jamais su* ***ce dont il est accusé***.
Je vous renvoie à ***ce que j'ai déjà écrit sur le sujet***.
Ce qui a été dit *est exact*.

En revanche, et contrairement à *celui*, *celle(s)* et *ceux*, Dubois note que le référent désigné par *ce* et son complément dépend entièrement du sémantisme de la proposition relative ; *ce* ne sert que d'amorce, il « anticipe sur l'expansion introduite par le relatif » (1965 : 165). De plus, il est remarqué que *ce* suivi de *que* ou *qui* est une suite non seulement très fréquente, mais surtout que ces éléments tendent à fusionner, d'un point de vue morpho-phonologique, en une locution conjonctive [skə] ou [ski] ; par ailleurs, la forme [skə] ne se cantonne pas aux propositions relatives, puisque l'ensemble *ce que* [skə] peut également introduire des phrases[29] exclamatives (exemple (13a)), causales (exemple (13b)), ainsi que des compléments du verbe (exemple (13c)) :

(13) a. ***Ce que*** *tu peux être bête !*
b. ***Ce que*** *c'est que de conduire si vite, l'accident n'est pas loin*.
c. *Je me réjouis de* ***ce que*** *vous ayez accepté*. (exemples issus de Dubois, 1965 : 165)

L'absorption de *ce* par l'élément complémenteur pourrait soutenir l'hypothèse d'un sémantisme réduit pour *ce* dans ce cas précis, pour ne lui attribuer qu'un rôle de mot-outil.

Il est un autre ensemble cohésif dont *ce* fait intégralement partie, la suite « *ce + est* », *c'est,* tour à tour analysée comme un dispositif présentatif, comme introducteur d'une clivée ou comme une structure copulative. La distinction entre l'une ou l'autre de ces constructions

[29] Nous utilisons ici le terme 'phrase' en référence aux grammaires traditionnelles, et non en faveur d'une quelconque prise de position dans les débats actuels entre les notions d''énoncé', de 'clause' et de 'phrase' – cette dernière ayant à juste titre un statut définitoire très discuté depuis l'éclosion de la linguistique de l'oral. Nous discuterons ce point au chapitre VI.

pourrait se faire sur la base d'un test de remplacement par *voici*/*voilà* ou *il y a*, en ce qui concerne le *c'est* présentatif :

(14) [on sonne à la porte] *Tiens ! **C'est** mon frère = Tiens ! **Voilà** mon frère*. (*c'est* présentatif)

(15) *Je t'assure, je n'ai rien dit à Marie, **c'est** Pierre qui a tout dit ≠ Je t'assure, je n'ai rien dit, **voilà** Pierre qui a tout dit*. (*c'est* non présentatif ; ici, introducteur d'une clivée)

(16) *Jacques, **c'est** un linguiste*[30] *≠ Jacques, **il y a** un linguiste*. (*c'est* non présentatif ; ici, structure copulative)

Si le *c'est* présentatif en (14) joue un rôle essentiellement expositoire, il n'en va pas de même pour le *c'est* introducteur de clivée, qui sert à mettre un référent particulier en exergue – (15) pourrait se paraphraser *c'est PIERRE qui a tout dit, et non pas moi*. Notons par ailleurs que le *c'est* présentatif sera préféré dans les phrases thétiques, où chaque élément de l'énoncé produit est nouveau au sein du discours, tandis que la partie subordonnée des clivées est thématique et présupposée – en d'autres termes, elle véhicule une information connue. Dans les deux cas cependant, le rôle de *c'est* semble être celui d'un tremplin, qu'il soit introductif ou contrastif. Cette observation pourrait aller dans le sens de Dubois qui analyse *ce* comme un mot-outil dans la suite *ce que*. En ce qui concerne le *c'est* des structures copulatives, l'analyse pourrait être quelque peu différente, *est* étant la copule et donc au cœur de la construction. *C'* n'en perd pas pour autant son rôle de mot-outil, puisqu'il est utile pour reprendre un référent qui, selon ce qu'il en est prédiqué, se verra attribuer une propriété ou une valeur identificatoire via la copule. Notons cependant que nous ne pourrions pas avoir de construction semblable à celle de l'exemple (16) avec un adjectif en position prédicative :

(17) a. **Jacques, **c'est** intelligent*.
b. *Jacques, **il** est intelligent*.

Il semblerait donc qu'il existe des contraintes à l'emploi de *c'* dans les structures copulatives. Nous ne rentrerons pas plus en détail dans la description de ces structures, mais laissons au lecteur le soin de se tourner vers Higgins (1979) et Pollock (1983), en ce qui concerne la lecture prédicationnelle ou identificationnelle des copulatives ; pour une analyse du choix du

[30] Exemple (16) de Reboul & Moeschler (1994 : 132).

pronom sujet *c'* ou *il* dans les structures copulatives en fonction des contraintes sémantiques liées à ces pronoms, voir Kleiber (1994) ; enfin, on pourra néanmoins consulter Reboul & Moeschler (1994), pour une critique de Kleiber.

Enfin, comme le constatent très justement Dubois (1965 : 167) et Chevalier *et al.* (1964 : 245), *ce* apparaît dans certaines structures figées, telles que *ce faisant, pour ce* et *pour ce faire*, qui relèvent plutôt d'un registre formel, et *sur ce* ou *c'est-à-dire* que l'on pourrait trouver dans des genres plus relâchés. Ce figement pourrait une fois de plus arguer en faveur d'une 'instrumentalisation' de *ce* au détriment de sa valeur référentielle.

1.4.3. *Ceci*, *cela* et *ça*

Les pronoms démonstratifs *ceci*, *cela* et *ça* sont les complémentaires distributionnels de *ce*, à la manière de *CELUI-ci* et *CELUI-là* par rapport à *CELUI*. Deux possibilités s'offrent alors :

- *Ceci*, *cela* et *ça* ne peuvent se substituer à *ce* : par exemple, aucun de ces trois éléments ne sera suivi d'une relative restrictive, c'est-à-dire indispensable à l'identification du référent visé (nous reprenons et modifions ici les exemples en (12)) :

(18) **Il n'a jamais su* ***ceci/cela/ça*** *dont il est accusé.*
Je vous renvoie à* ***ceci/cela/ça* *que j'ai déjà écrit sur le sujet.*
****Ceci/cela/ça*** *qui a été dit est exact.*

- *Ceci* et *cela* sont les formes marquées de *ce* devant le verbe *être*, lorsque celui-ci commence par une voyelle :

(19) ***Cela*** *est honteux* / ***Ceci*** *est honteux.*
Cela *est vrai* / ***Ceci*** *est vrai*. (exemples issus de Dubois, 1965 : 167)

En revanche, *ça* ne fait pas partie de cet ensemble marqué dans ce cas précis, puisque nous n'aurons jamais la suite suivante :

(20) ****Ça*** *est vrai*.

Nous pourrions avancer que l'énoncé produit par l'exemple (20) a peu de chance de se réaliser en raison du hiatus entre [sa] et [ɛ] ; cependant, cette remarque serait également

valable pour *cela est* [səlaɛ]/[slaɛ] et *ceci est* [səsiɛ]. Nous préférons donc amorcer ici l'idée de *ça* ayant un fonctionnement qui lui est propre, et ne pas l'assimiler à la simple forme orale de *cela*, bien que cette conception soit largement répandue, adoptée par les grammaires prescriptives et énoncée par Dubois (1965 : 167). Nous développerons ce point dans le chapitre VIII. En revanche, nous soutenons Dubois et son hypothèse de *ceci* comme forme marquée de *cela* – tout comme *THIS* est la forme marquée de *THAT*. En effet, l'emploi de *ceci* est à la fois bien moins fréquent et plus contraint que celui de *cela* ; par exemple, il suffit de regarder l'exemple suivant pour constater l'impossibilité de *ceci* dans des tournures qui se rapprochent de l'impersonnel :

(21) a. ***Cela*** *m'étonnerait qu'il vienne*.
b. ***Ça*** *m'étonnerait qu'il vienne*.
c. *****Ceci*** *m'étonnerait qu'il vienne*.

Nous pouvons également observer la restriction sémantique de *ceci* par rapport à *cela*. En (22), *cela* et *ceci* sont tous deux acceptables en référence à la dénotation* d'un prédicat – « avoir beaucoup d'argent » ; en revanche, *cela* en (23a) peut référer à « l'argent », qui est une entité discrète, contrairement à *ceci* :

(22) a. *Je voudrais avoir beaucoup d'argent*. ***Cela*** *est bien agréable*.
b. *Je voudrais avoir beaucoup d'argent*. ***Ceci*** *est bien agréable*.

(23) a. *L'argent ne fait pas le bonheur, mais* ***cela*** *aide à faire les courses*.
b.***L'argent ne fait pas le bonheur, mais* ***ceci*** *aide à faire les courses*. (exemples (22) et (23) issus de Corblin, 1987 : 91)

L'éventail des types de référents possibles pour *ceci* serait donc plus réduit que pour *cela*. Dans cette optique, Corblin (1987) argumente que *ceci*, contrairement à *cela*, ne peut être en référence générique, comme illustré dans l'exemple suivant :

(24) a. *Un chien,* ***ça*** *ne vit pas très vieux*.
b. *Un chien,* ***cela*** *ne vit pas très vieux*.
c. **Un chien,* ***ceci*** *ne vit pas très vieux*. (exemples issus de Corblin, 1987 : 91)

Si *ça* en référence générique dans des structures à dislocation est tout à fait acceptable, il n'en va évidemment pas de même pour *ceci* ; néanmoins, comme *cela* nous semble également très peu naturel dans une telle tournure, nous ne tablerions pas sur cet argument pour justifier du statut marqué de *ceci* par rapport à *cela*. Au contraire, la référence générique dans des structures à dislocation serait plutôt un critère de démarcation de *ça* par rapport à *cela*. Nous reviendrons sur ces structures lors du chapitre VIII.

Pour résumer, l'analyse structurale des démonstratifs français présente *celui, celle(s)* et *ceux* comme étant les complémentaires distributionnels des pronoms *CELUI-ci* et *CELUI-là.* De manière similaire, *ce* est en distribution complémentaire avec les pronoms *ceci, cela* et *ça,* mais partage certaines propriétés paradigmatiques avec *celui, celle(s)* et *ceux*. A présent, nous pouvons exposer l'analyse proposée des démonstratifs français dans des approches théoriques, en commençant par diverses versions de la grammaire GG.

2. L'analyse des démonstratifs français dans un cadre génératif

La présente section s'attache à décrire les analyses de type génératif des démonstratifs français. Ce panorama débute avec la proposition de Gross (1977, section 2.1), disciple de Harris, se poursuit avec l'amorce d'un traitement transformationnel 'à la Chomsky' effectué par Dubois & Dubois-Charlier (1970, section 2.2), avant de présenter la prise en compte des démonstratif au sein du Programme Minimaliste de Chomsky (1995, section 2.3), version la plus récente de sa théorie.

2.1. L'analyse générative harrissienne : Gross (1977)

L'analyse de Gross (1977) s'inspire de la théorie distributionnaliste et générative de Harris. Bien que l'ouvrage-support de cette présentation date de la fin des années soixante-dix, la théorie de Gross remonte à 1968. Il s'agit en premier lieu de bien différencier les conceptions de la langue et du langage qu'ont Gross et Chomsky. Si Chomsky postule une faculté biologique de langage propre à l'être humain (cf. chapitre II, section 2), il n'en va pas de même pour Gross et Harris. Pour ces derniers, nul métalangage antérieur à la production langagière – la Grammaire Universelle chez Chomsky – et par conséquent, nulle distinction entre 'compétence' – le savoir-faire linguistique inné chez l'Homme – et 'performance' – la production langagière effectuée par l'Homme à partir de cette même 'compétence'. Du point de vue technique, les transformations chez Gross et Harris ne sont pas des opérations qui

amènent une 'structure profonde' vers une 'structure de surface', mais des relations entre structures phrastiques. Selon Gross, les déterminants opèrent par inclusion suivant la dérivation « GN → Det N de GN », règle qui traduit une opération d'extraction : « *un ami* → *un ami de mes amis* ». Néanmoins, son postulat concernant les démonstratifs n'est pas sans annoncer celui de Dubois & Dubois-Charlier (cf. section 2.2 *infra*), où l'idée d'un défini sous-jacent au démonstratif apparaît. En effet, Gross considère les pronoms c*elui, celle, ceux* et *celles* comme étant composés du déterminant *ce* associé au pronom personnel *LUI*. Si *LUI* est dérivé d'un nom selon la règle « N → LUI[31] » et si cette découpe morphologique est exacte, il reste à justifier que *ce* inclus dans un pronom démonstratif reste invariable, contrairement à *ce* déterminant qui s'accorde en genre et en nombre avec le nom qu'il précède – *ce**s** hommes*, *ce**tte** femme* – et qui accepte la forme *cet* [sɛt] lorsque celle-ci est suivie d'un son vocalique – *ce**t** homme* [sɛtɔm]. L'argumentation de Gross s'expose alors de la manière suivante : un SN démonstratif s'analyse toujours en termes d'inclusion sémantique, suivant la formule « *Ce* $N_{g,n}$ = *Ce* $N_{g,n}$ de $D_{déf}$ $N_{g,plur}$ » (1986 : 129). Pour que la relation d'inclusion « $\{N_{g,n}\} \subset \{N_{g,plur}\}$ » soit respectée, il est indiqué que les deux N en présence se doivent d'être identiques :

(25) ***Cet ami** que tu as vu hier est mort = **Cet ami de nos amis** que tu as vu hier est mort.*

« *Ce* $N_{g,n}$ Modif = *Ce* $N_{g,n}$ de GN Modif où GN = $D_{déf}$ $N_{g,plur}$ et Modif = prop relative »

Si l'on pose la règle de substitution « N → LUI », nous avons :

(26) ***Cet ami** que tu as vu hier est mort = **celui** que tu as vu hier est mort.*

« *Ce* $N_{g,n}$ Modif = *Ce* $N_{g,n}$ de $D_{déf}$ $N_{g,plur}$ Modif = *ceLUI* de $D_{déf}$ $N_{g,plur}$ Modif »

La pronominalisation de *LUI* par rapport au $N_{g,plur}$ inclusif et la règle énoncée en (25) conduiraient à l'accord entre ces deux éléments, *ce* restant non marqué en genre et en nombre, et produisant ainsi les pronoms *celle*, *celles* et *ceux*. De plus, l'effacement potentiel de « de $D_{déf}$ $N_{g,plur}$ » rendrait possibles les suites telles que *celui que tu as vu*, tandis que les adverbes *ci* et *là* en instanciation de Modif génèreraient les formes pronominales *ceLUI-ci* et *ceLUI-là*.

[31] « N → LUI » traduit qu'un pronom de troisième personne est substituable à un nom, et conserve ainsi ses propriétés morphosyntaxiques.

Il est à noter cependant que la théorie grossienne d'une structure pronominale tripartite « *ce LUI* Modif » ne justifie point des différences syntaxiques et sémantiques entre les formes simples – *celui, celle, ceux, celles* – et les formes composées – *ceLUI-ci, ceLUI-là*. Néanmoins, Gross soulève nombre de remarques pertinentes, reprises et développées ensuite dans les travaux de Corblin (1987), Kleiber (1994) et Schnedecker (2006) ; en particulier, Gross observe la référence générique potentiellement réalisée par les syntagmes nominaux définis, dissocie les notions de pronominalisation et de coréférence, attribue aux SN le pouvoir d'effectuer une « référence lexicale » (1986 : 122) – SN qui désignent une notion sémantique et non un « objet » identitaire – et admet la capacité des syntagmes démonstratifs à accepter un modifieur de type adjectival qui ne modifie pas l'entité désignée[32].

2.2. L'analyse de Dubois & Dubois-Charlier (1970)

L'une des premières analyses transformationnelles du français est celle proposée par Dubois & Dubois-Charlier (1970), dans le cadre défini par *Aspects of the Theory of Syntax* (Chomsky, 1965). Cet ouvrage date de la période pré-Abney (1987), le SN est donc traité comme régi par le nom-tête, qui sélectionne le genre et le nombre du déterminant. Dubois & Dubois-Charlier proposent la règle de réécriture suivante pour le déterminant (1970 : 38) :

(27) D → (préArt) + (Dém) + Art + (postArt)

où, par exemple, « préArt = *tout, tous* », « Dém = *ce, cet, ces* », « Art = *un, le* », « postArt = *deux, mille* ». Selon cette règle, l'élément Art est le seul obligatoire, tandis que tous les autres sont optionnels. Nous pouvons déjà entrevoir ici que la catégorie 'démonstratif' est facultative, et potentiellement ajoutée à la catégorie de l'article, elle obligatoire. Le déterminant serait donc constitué *a minima* d'un seul constituant, de quatre au maximum.

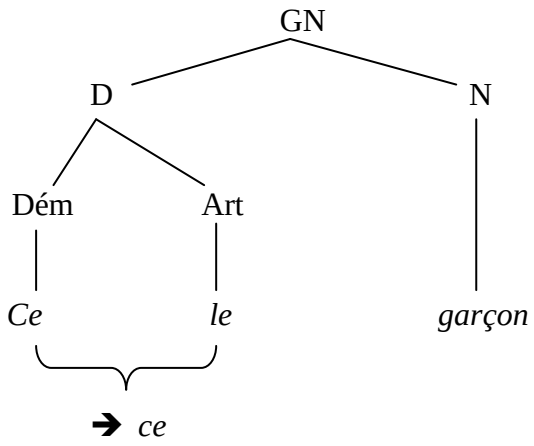
Avant d'exposer le point de vue de Dubois & Dubois-Charlier sur la structure interne des démonstratifs proprement dits, il convient de mentionner la règle de réécriture proposée pour l'article :

[32] Ces remarques seront développées dans le chapitre IV, lors de la présentation des théories 'modernes' qui traitent des démonstratifs et de la référence en général.

(28) Art → { Déf / nDéf }

Cette règle est dite « de sous-catégorisation » car elle représente un choix entre deux sous-catégories – Déf pour 'défini' et nDéf pour 'non défini' – appartenant à la même catégorie Art[33]. Ainsi, l'article peut se réécrire sous la forme de la sous-catégorie Déf ou nDéf. En ce qui concerne la catégorie des 'démonstratifs' – *ce, ces, cette*, etc – Dubois & Dubois-Charlier considèrent que ces éléments sont constitués de la composante optionnelle Dém associée à celle obligatoire Art. Ainsi, le SN *ce garçon* serait représenté de la manière suivante, selon l'arbre syntaxique en (29) :

(29)



Il est nécessaire de préciser ici que la composante Dém instanciée par *ce* en (29) est un élément abstrait, à différencier du démonstratif *ce* qui est le résultat catégoriel de l'association « Dém + Art ». De plus, il est posé que cette réalisation combinatoire ne peut s'effectuer qu'en présence d'un Art sous-catégorisé Déf. Ainsi, l'article défini est considéré comme un élément sous-jacent à la catégorie des démonstratifs, point de vue que Dubois & Dubois-Charlier argumentent par le fait que le défini est philologiquement issu des démonstratifs. Par ailleurs, il est avancé que démonstratifs et articles ne peuvent être traités comme appartenant à

[33] Le terme de 'sous-catégorisation' est également utilisé dans les traitements dérivés de Chomsky (1965) pour spécifier la valence d'un élément recteur. Par exemple, un verbe transitif exigeant un complément du type SN possède un cadre de sous-catégorisation du type suivant : [+ --- SN]. Ces deux sens sont donc à bien distinguer.

des catégories de même niveau ; en effet, une telle analyse donnerait la règle de réécriture suivante pour le déterminant (1970 : 49) :

(30) $D \rightarrow (\text{préArt}) + \left\{ \begin{array}{l} \text{Art} \\ \text{Dém} \end{array} \right\} + (\text{postArt})$

Cette règle induirait que la composante obligatoire à la structure interne d'un déterminant soit un article ou un démonstratif, ce qui poserait problème lors de l'analyse des pronoms *celui, celle* et *ceux*, qui sont considérés comme étant composés de « *ce + lui* », « *ce + elle* » et « *ce + eux* » respectivement. Ces items comporteraient par conséquent deux composantes obligatoires. Dubois & Dubois-Charlier offrent de fait une analyse transformationnelle où une forme de départ à deux éléments (cf. exemple (29)) est fusionnée en un seul élément au niveau superficiel.

2.3. Les analyses transformationnelles modernes

La présente section fait écho à celle concernant le traitement de *THIS* et *THAT* dans la théorie X-barre, en ce que l'analyse des démonstratifs anglais et français est similaire. Ainsi, nous renvoyons le lecteur au chapitre II, section 2, qui couvre le courant génératif chomskyen et l'avènement du déterminant comme tête de syntagme (Abney, 1987). Nous passons alors à l'analyse proposée des démonstratifs au sein du Programme Minimaliste* (1995), dernière version de la théorie chomskyenne.

2.4. L'analyse transformationnelle dans le cadre du Programme Minimaliste (Chomsky, 1995)

A la suite du travail d'Abney, deux questions majeures se sont posées : l'articulation de la structure fonctionnelle du DP et le mouvement[34] du nom au sein du DP. Cette dernière question concerne notamment l'ordre des mots à l'intérieur du DP – en français, l'adjectif peut précéder ou suivre le nom, mais non sans conséquences sémantiques – et l'assignation de

[34] Déplacement que subit un constituant lors des transformations syntaxiques afin d'occuper une certaine place en structure de surface, autre que celle qu'il tient en structure profonde.

la valeur référentielle d'un syntagme. Cette section s'attache à exposer les solutions proposées pour les déterminants et pronoms démonstratifs français, dans le cadre du Programme Minimaliste.

Le Programme Minimaliste est la version de la grammaire générative chomskyenne développée en 1995 et toujours en vigueur. L'architecture du Programme Minimaliste se veut optimale ; ainsi, le système computationnel qui prend en charge les entrées provenant du lexique se résume à deux opérations : 'Merge' ('fusion') et 'Move' ('déplacement'). Merge est une opération binaire qui, à partir de deux entités α et β, crée une troisième entité γ, dont α et β sont les constituants. Cette fusion est soumise aux contraintes définies par le schéma X-barre (cf. représentation au chapitre II, exemple (12)). Move est une opération de déplacement des constituants qui permet d'obtenir l'agencement observé en sortie. Les opérations Merge et Move sont à leur tour assujetties à un processus de vérification afin d'obtenir une bonne formation en sortie. Chaque item provenant du lexique est doté de traits formels – par exemple, le genre et le nombre ; ce sont ces traits qui déclenchent les dérivations syntaxiques, notamment la nécessité ou non d'effectuer un déplacement. Enfin, chaque élément en sortie est associé à deux niveaux de représentation : une 'Forme Phonologique' (PF) en relation avec les aspects physiques de la langue et une 'Forme Logique' (LF) concernant l'interface entre la syntaxe et le but communicatif.

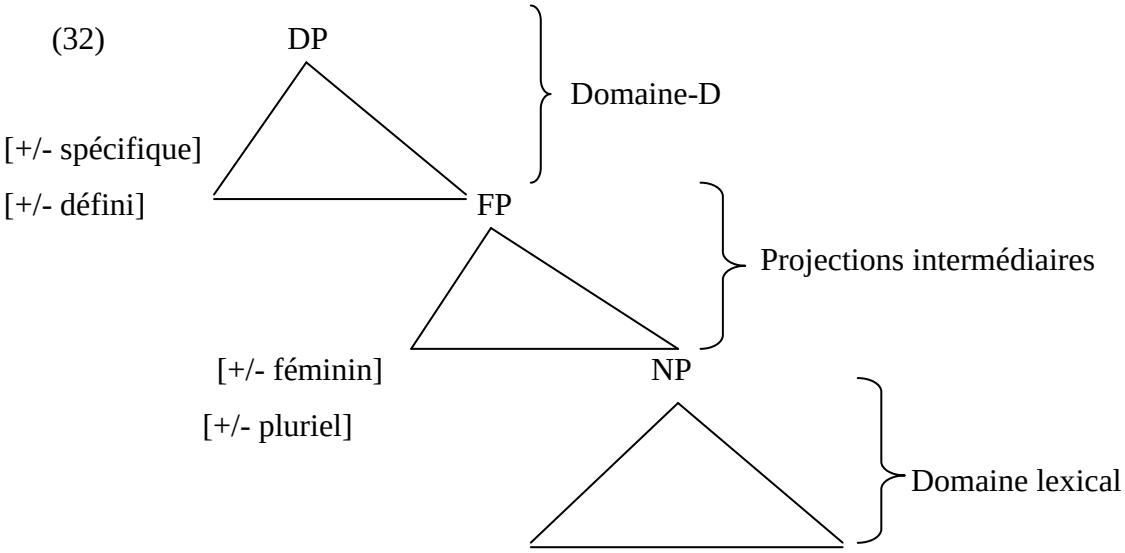
Pour ce qui est des démonstratifs, il est considéré que ces derniers n'appartiennent pas à la même catégorie que les définis. En effet, il est remarqué que, dans certaines langues comme l'espagnol, démonstratifs et définis peuvent apparaître simultanément : *el* (défini) *libro este* (démonstratif). De plus, le démonstratif peut faire office de constituant autonome, contrairement au défini qui nécessite invariablement la présence d'un élément lexical. Afin de cerner l'analyse qu'il est fait des démonstratifs dans un cadre génératif moderne, il convient tout d'abord d'expliciter la structure générale du DP en français, qui s'analyse selon la structure suivante :

(31) $[_{DP}\ D\ [_{FP}\ F\ [_{NP}\ N\ ]]]$

La représentation linéaire en (31) stipule que le déterminant D est la tête initiale du syntagme DP ; à l'extrême droite de cette formule est notée la tête N qui projette le syntagme NP du bas vers le haut, base lexicale à l'origine des dérivations ; entre les projections DP et NP sont

identifiées plusieurs projections fonctionnelles, regroupées ici sous le nom de FP. Cette représentation sous forme d'arbre syntaxique équivaut à la figure suivante :

(32)

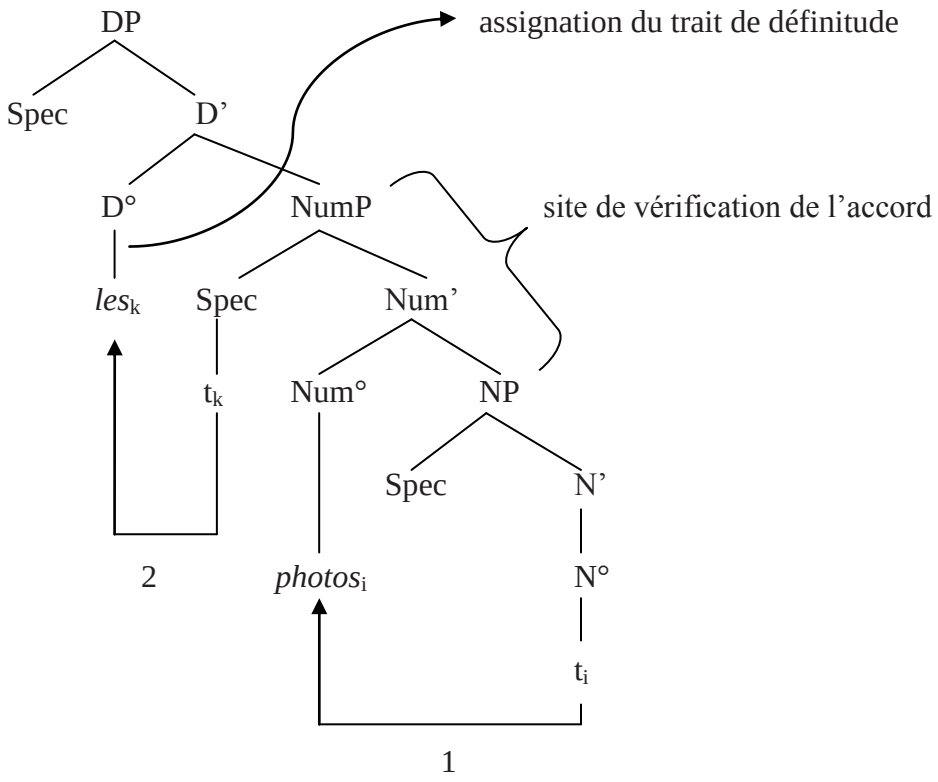


Ainsi, la tête lexicale à la base de la dérivation se déplace en FP afin de recevoir sa morphologie flexionnelle ; c'est à ce niveau que sont vérifiés les phénomènes d'accord en genre et en nombre. L'étape ultime de la dérivation est le domaine-D, qui correspond au locus de référentialité, et qui est donc en relation avec l'ancrage discursif du syntagme ; de manière prototypique, les déterminants sont les items qui couvrent ces fonctions au sein du syntagme. Afin de mieux cerner le mécanisme interne du DP et avant de nous focaliser sur le comportement des démonstratifs au sein de ce modèle, voyons ce qu'il est établi de l'article défini[35].

L'article défini en français est prénominal, obligatoire lorsque le SN se trouve en position argumentale, et fait partie de la classe des clitiques. L'élément lexical étant le point de départ de la dérivation selon le schéma X-barre, celui-ci se déplace en une position fonctionnelle – NumP, selon Valois (1991 : 87) – afin de recevoir sa morphologie flexionnelle (mouvement 1 sur la représentation en (37)). Le déterminant est alors généré en position de spécifieur de NumP avant de monter en position D° – position de la tête du DP – où il reçoit le trait de définitude (mouvement 2) :

[35] Cette section faisant partie d'un chapitre d'exposition concernant l'analyse proposée des démonstratifs dans certaines théories fondamentales, mais ne constituant pas néanmoins le cadre épistémologique choisi pour l'étude au cœur de cette thèse, les tenants et aboutissants ne sont pas explicités par souci de clarté.

(33) *Les photos*

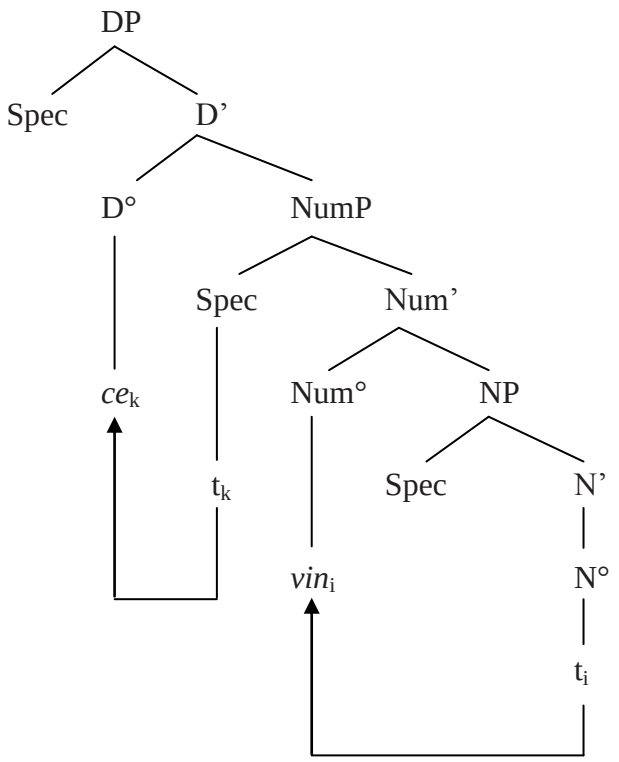


Toujours selon Valois (1991 : 87), D° est un site d'accueil du déterminant, et non une position de génération, car celui-ci s'accorde avec le nom ; or, l'accord en genre et en nombre s'effectue en NumP, qui correspond à une dérivation antérieure. La montée du déterminant en D° s'explique par le trait de définitude que celui-ci doit recevoir – cette montée n'aurait pas lieu dans un emploi générique du SN défini, comme dans ***les photos****, ça rappelle des souvenirs*. Par ailleurs, nous pouvons constater qu'une trace *t* co-indexée est laissée en position originelle de l'élément considéré.

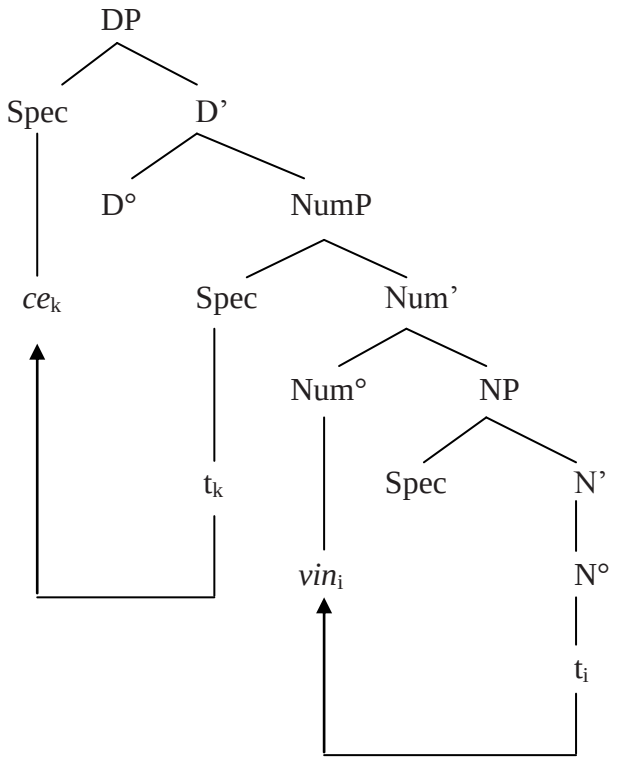
Le traitement des démonstratifs est comparable à celui des articles tout en étant plus complexe ; en effet, les démonstratifs peuvent apparaître en position de SN sans la présence d'une tête lexicale – *CELUI-ci*, *CELUI-là*, mais aussi *celui*, *celle, ceux* et *celles* accompagnés d'un complément – et leur analyse doit rendre compte du possible emploi non spécifique des démonstratifs – comme dans *J'aime ce vin*, où *ce vin*, tout en désignant potentiellement celui présent sur la table, réfère à la classe à laquelle il appartient. Selon Bernstein (1997 : 96), l'emploi non spécifique des démonstratifs est majoritaire dans les langues germaniques, mais néanmoins possible dans les langues romanes. Certaines variétés du français, le picard notamment, usent même de démonstratifs comme articles définis (Bernstein, 1997 : 94).

Rowlett (2007 : 50) argumente cette possible interprétation du SN démonstratif en français par la traduction en anglais, qui passe alors par l'article défini 'the'. Toutefois, quelle que soit la référence effectuée lors de son emploi, les déterminants et pronoms démonstratifs seraient générés en spécifieur de NumP – tout comme les articles – puisqu'ils dépendent de la forme flexionnelle du nom. Si la valeur référentielle du SN est non spécifique, le déterminant démonstratif est analysé à la manière d'un défini et se déplace en D°. Par conséquent, nous pouvons même avancer que les démonstratifs sont ici assimilés à des articles définis : ils sont dépourvus de leur faculté de pronominalisation – puisque amalgamés à des clitiques – et devraient également perdre la sélection référentielle qu'ils sont supposés effectuer (cf. Introduction de ce chapitre). En revanche, si l'entité désignée est spécifique, le site d'accueil du démonstratif est en spécifieur de DP (Valois, 1991 ; Bernstein, 1997) :

(34) *J'aime* ***ce vin*** (valeur non spécifique, compris comme « ce genre de vin »)



(35) *J'aime* ***ce vin*** (valeur spécifique)

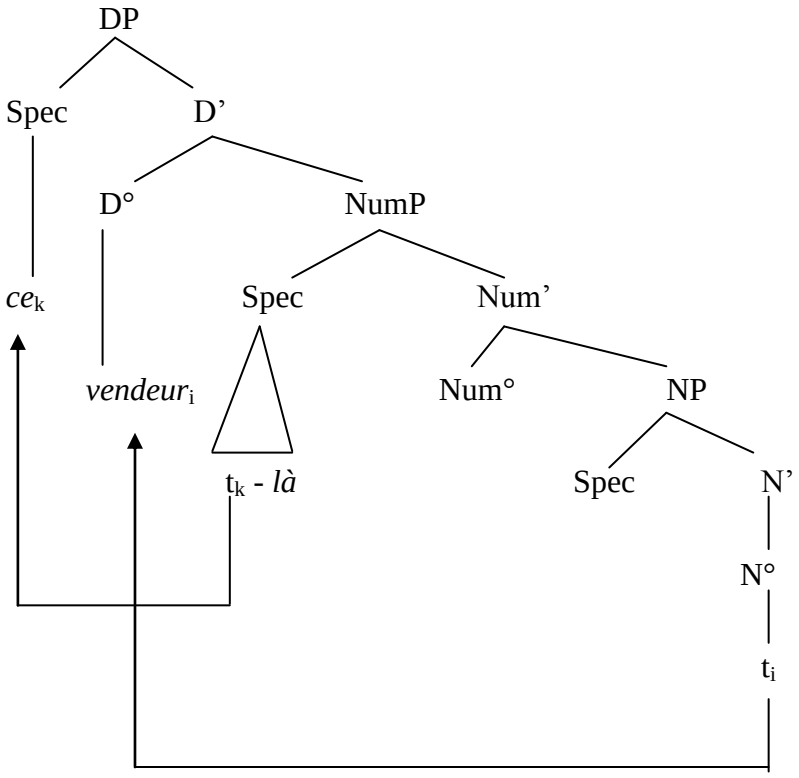


A partir de là s'est posé le problème du traitement des particules adverbiales *ci* et *là* : où sont-elles générées ? Où se déplacent-elles, si tant est qu'il y ait un mouvement effectué ? En préambule de sa démonstration, Bernstein (1997) note que *ci* et *là* dépendent de la présence d'un démonstratif, mais que cette relation est univoque – un démonstratif n'appelle pas nécessairement la présence de l'une ou l'autre de ces particules adverbiales. De plus, elle remarque que *ci* et *là* ne sont employés que lors d'une référence spécifique du syntagme ou pronom démonstratif[36] et que le groupe lexical éventuel peut se présenter sous la forme d'un nom composé – *ce vendeur de voitures-là* (cf. également Rowlett (2007 : 51) à ce sujet). Ces propriétés mènent Bernstein à la conclusion que l'ensemble « démonstratif + groupe lexical éventuel + *ci/là* éventuel » est généré sous forme d'élément complexe en NumP, *ci/là* étant la tête représentée en Num° – tête vide si *ci/là* n'est pas instancié – et l'ensemble « démonstratif + groupe lexical éventuel » étant le spécifieur en SpecNumP. Le déplacement du démonstratif en SpecDP lui fait gagner ses propriétés fonctionnelles, tandis que *ci/là* reste *in situ* ; si le

[36] Cette remarque pourrait s'avérer discutable, notamment lors de certains emplois au pluriel : *J'aime ces vins-là* peut très bien s'accommoder d'une référence à la classe à laquelle le *demonstratum* appartient.

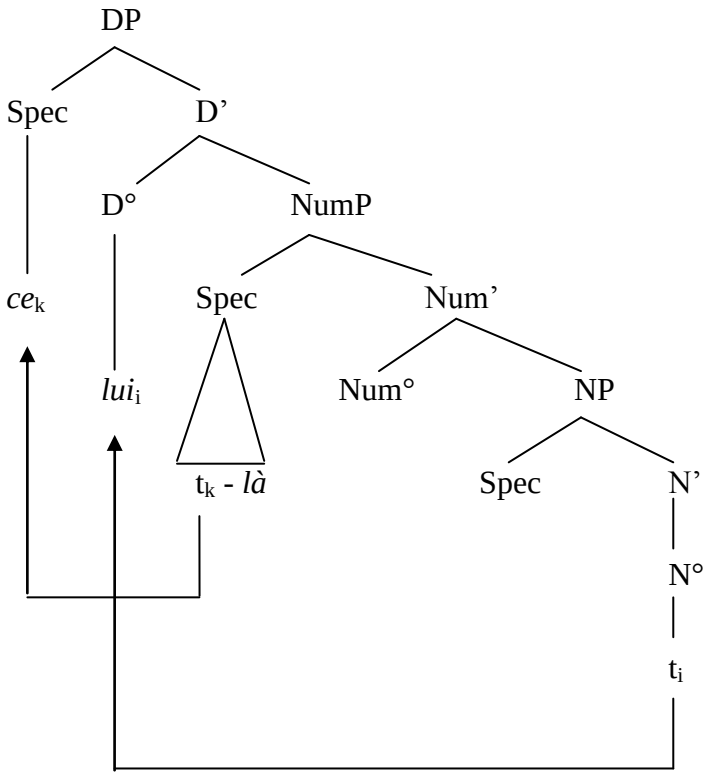
démonstratif employé est un déterminant, le nom ou groupe nominal monte en D°[37]. Il est à noter que Bernstein qualifie le mouvement du démonstratif comme un mouvement de syntagme et non de tête, ce qui mène à penser que le pronom démonstratif est considéré comme étant compositionnel – analysé en « *ce* + *LUI* », à la manière des théories d'ores et déjà présentées dans ce chapitre :

(36) *Ce vendeur-là*



[37] Le déplacement de l'élément lexical en D° trouve sa logique si nous considérons qu'un SN démonstratif en référence spécifique est par défaut défini car identifiable et identifié par le destinataire. Nous reviendrons sur cette remarque lors de l'analyse proposée au cœur de cette thèse.

(37) *Celui-là*



A la lumière de ces analyses, nous pouvons conclure que les théories génératives assimilent les démonstratifs à des définis, si le syntagme est à référence non spécifique – bien que les démonstratifs et les définis soient considérés comme appartenant à des catégories distinctes. En revanche, le démonstratif à référence spécifique est compris comme un ensemble syntagmatique pouvant être augmenté de l'une ou l'autre des marques adverbiales *ci* et *là*. Par conséquent, la dérivation obtenue sera différente selon l'emploi discursif de l'expression référentielle. Précisément, la faille principale d'un tel modèle pourrait résider dans l'assimilation entre propriétés systémiques des éléments linguistiques et fonctions activées lors de leur emploi discursif.

3. Les analyses énonciatives des démonstratifs français

3.1. Guillaume et la Psychosystématique du langage (1919)

3.1.1. Présentation générale

La Psychomécanique – ou Psychosystématique – du langage de Guillaume (1883, 1960) assure une transition parfaite entre la théorie générative transformationnelle de Chomsky et les linguistiques de l'énonciation. En effet, même si leurs préoccupations

scientifiques sont différentes – Guillaume s'attache à décrire des opérations mentales, tandis que l'objet d'étude de Chomsky est la syntaxe – il demeure quand bien même certaines similitudes entre les préceptes des deux linguistes. Ainsi, Guillaume et Chomsky remarquent que la faculté de langage est propre à l'homme, que le développement langagier s'effectue de manière toute naturelle et avec une facilité déconcertante par les enfants en phase d'acquisition – contrairement à l'adulte qui éprouvera toujours plus de difficultés à apprendre une langue seconde –, ce qui les mène à penser qu'il existe chez l'être humain un savoir-faire linguistique nécessaire à la communication verbale. La distinction compétence/performance de Chomsky fait sans nul doute écho à ces considérations, qui sont la base même du modèle théorique élaboré par Guillaume. Aussi, Guillaume est un pionnier dans la problématisation linguistique sous forme dynamique, décrivant des opérations qui permettent de générer – osons employer ce terme – les éléments d'un système. La métalangue utilisée en Psychosystématique ne laisse planer aucun doute sur ce point : la 'chronogénèse'* est l'opération qui génère les formes verbales, la 'topogénèse'* (introduite par Lafont en 1967) l'actualisation du nom par les déterminants. Si cette dernière opération est en relation avec le traitement des démonstratifs par les linguistes guillaumiens, il convient toutefois de présenter, ne serait-ce que de manière succincte, la Psychosystématique du langage.

Pour Guillaume, le langage est un phénomène dont l'étude ne peut s'effectuer que dans la perspective du sujet parlant : nul phénomène ne peut exister s'il n'est pas produit. 'Langage' et 'locuteur' sont donc indissociables, comme le sont 'langage' et 'pensée' – l'activité langagière a cours dans les pensées du sujet – ainsi que 'langage' et 'temps' – tout phénomène est créé dans un laps de temps, si minime soit-il. Si parler équivaut à un 'faire', cette action nécessite un 'savoir-faire', qui correspond à la langue d'une communauté donnée. Pour Guillaume, la langue est un système constant exploitable par le sujet parlant afin de produire un discours, résultat éphémère car temporellement situé, dont l'unité minimale est la phrase. Ainsi, à l'instar de Saussure – dont il ne prit connaissance des idées fondatrices qu'après la publication de son premier ouvrage, *Problème de l'Article et sa solution dans la langue française* (1919) –, Guillaume distingue les concepts de 'langue' et 'discours' – 'parole' chez Saussure. Guillaume définit trois états d'existence du langage :

- *In posse* : correspond à la langue en tant que système potentiellement exploitable.
- *In furie* : correspond à l'acte de langage en cours d'existence.
- *In esse* : correspond au discours en tant que matérialisation linguistique.

Il est à noter que Guillaume inscrit cette vision tripartite de l'existence du langage dans une observation plus large de tout élément dans l'univers : toute chose qui existe répond aux facteurs qui ont permis cette existence, puis cette chose existe, avant de s'éteindre. Pour illustration, Guillaume développe les formes verbales impersonnelles du français qui recouvrent ces trois états : l'infinitif *in posse*, le participe présent *in furie*, le participe passé *in esse*. Par ailleurs, Guillaume indique que la production langagière répond elle-même à certaines conditions : il ne suffit pas de dire quelque chose, il faut non seulement avoir quelque chose à dire – ce qu'il nomme la 'visée du discours'* – mais aussi vouloir le dire – ce qu'il appelle la 'visée phrastique'*. Les éléments systémiques de la langue, ou 'unités de puissance'*, n'étant par définition pas observables – puisque présentes dans la pensée du sujet parlant –, il est par conséquent nécessaire de s'appuyer sur des données physiques afin de rendre compte de l'activité langagière. Ces données sont fournies par les éléments du discours, ou 'unités d'effet'*. La volonté de Guillaume est de remonter des unités d'effets vers les unités de puissance. Pour ce faire, il développe un système représentatif et interprétatif des mécanismes psychiques de la pensée sous forme de schème binaire, qu'il nomme 'tenseur binaire radical'*. Ce schème bi-tensif représente le mode de fonctionnement de la pensée qui, selon Guillaume, se caractérise par un mouvement binaire : l'un allant du large à l'étroit – du général au particulier – l'autre de l'étroit au large – du particulier au général (*Leçon du 28 Novembre 1947*, Série C) :

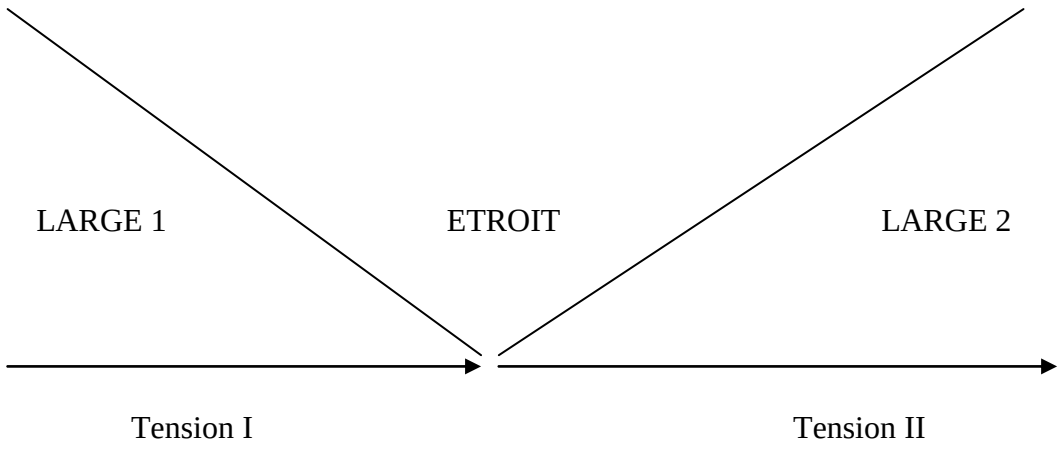


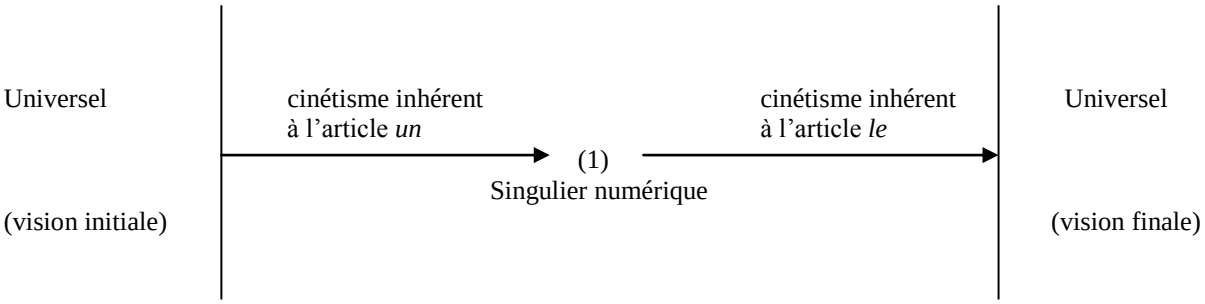
FIGURE 3 : LE SCHEME BI-TENSIF EN PSYCHOMECANIQUE

Ce schème serait applicable à la construction de l'ensemble des structures de la langue par l'adaptation conceptuelle des pôles extrêmes de ces deux mouvements miroirs. Ainsi, les

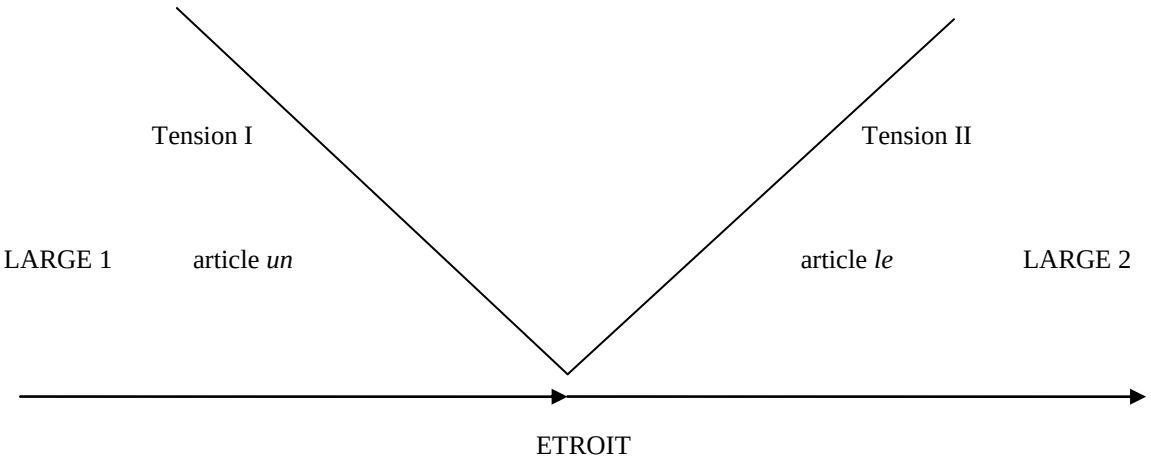
limites pour le nombre seraient le pluriel pour le pôle large, le singulier pour le pôle étroit ; en ce qui concerne la détermination, et en particulier l'article, l'universel[38] instancierait le pôle large, le distinctif le pôle étroit. Ce va-et-vient continu de la pensée entre le large et l'étroit serait susceptible de s'arrêter à un instant *t*, permettant ainsi d'obtenir un effet de langue qui soit plus ou moins étroit selon la distance relative de la structure langagière observée par rapport à l'un ou l'autre des deux pôles.

3.1.2. La détermination en Psychomécanique

C'est dans *Problème de l'article et sa solution dans la langue française* (1919) que Guillaume pose les jalons de sa théorie et effectue une analyse très fine de la détermination nominale en français. Ainsi, comme nous l'avons évoqué dans le paragraphe précédent, Guillaume développe l'idée que l'article, quel qu'il soit, permet « de réduire l'idée générale du nom en une idée étroitement particulière et momentanée » (1919 : 16). En d'autres termes, l'article assure la transition entre nom en puissance et nom en effet en effectuant une coupure dans la signification toute notionnelle dénotée par le nom – par exemple, le concept « homme » – afin d'obtenir une image unique – « l'homme dont je t'ai parlé hier » ; « j'ai croisé un homme bizarre hier ». L'article indéfini, de par l'extraction qu'il opère, est caractérisé par un mouvement particularisant allant de l'universel à l'unique, mouvement qualifié de 'cinétisme* anti-extensif' (*Leçon du 09 Janvier 1948*, Série C), tandis que l'article défini signale un mouvement miroir généralisant qualifié de 'cinétisme extensif' (*op. cit.*), allant du particulier à l'universel :



[38] Chez Guillaume, 'universel', 'large' et 'général' sont des termes synonymes : 'universel' n'a pas ici le même sens qu'en grammaire générative – *i.e.* « commun à toutes les langues du monde ».



FIGURES 4 & 5 : LE CINETISME ANTI-EXTENSIF DE L'ARTICLE INDEFINI ET LE CINETISME EXTENSIF DE L'ARTICLE DEFINI EN PSYCHOMECANIQUE

Concernant *le* plus particulièrement, une interruption du mouvement extensif proche du pôle étroit – 'précoce' en termes guillaumiens – induirait alors une référence spécifique : *l'homme entra, s'assit, et attendit* ; à l'inverse, une interruption du mouvement proche du pôle large – interruption 'tardive' – entraînerait une référence générique : *l'homme est un loup pour l'homme*.

Dans *Problème de l'article et sa solution dans la langue française*, Guillaume compare également les déterminants – 'articles' dans sa terminologie – *ce* et *le*, basant son analyse sur des arguments diachroniques, le défini étant historiquement issu du démonstratif. Selon Guillaume, l'article défini est apparu afin de pouvoir distinguer 'montrer le réel' – fonction qui incombe au démonstratif – et 'montrer virtuellement dans la langue' – fonction assurée par le défini. Par cette approche, nous discernons l'opposition entre 'exophore' et 'endophore', entrevue au chapitre II, section 1. Guillaume va même jusqu'à qualifier le défini de 'démonstratif anaphorique' (1919 : 16). En ses propres termes, Guillaume énonce les prémices de ce qui est établi aujourd'hui des propriétés du démonstratif et du défini. En effet, il distingue ces déterminants en fonction de la sélection référentielle opérée par le démonstratif, rendue possible par le caractère essentiellement déictique de cet item – même s'il n'emploie pas ce terme – *a contrario* du défini qui a tendance à continuer la référence d'une entité déjà identifiée par l'allocutaire (1919 : 18). Pareillement, Guillaume souligne la capacité du démonstratif à désigner « une chose déterminée parmi des choses possibles en pensée » et celle du défini à référer au générique : le démonstratif sélectionne « un point »

dans « l'étendue » désignée par le défini (1919 : 44-45, note 2), ce qui mène Guillaume à la conclusion d'un *ce* non extensible – à la manière de l'indéfini – et d'un *le* extensible (1919 : 227). Ces caractéristiques observables dans le discours sont appuyées par l'effet produit par le démonstratif et par le défini dans les expressions temporelles, où *ce* permet de situer la référence dans un espace-temps en marge du moment de l'énonciation – Guillaume parle d''actuel relatif' (1919 : 207) – tandis que *le* ancre la référence dans l''actuel absolu' ; cette différence émane des cadres temporels exprimés par *à cet instant* et *à ce moment* qui sont déplacés par rapport au *hic* et *nunc* signalé par *à l'instant* et *pour le moment*.

3.1.3. Les pronoms démonstratifs en Psychomécanique

Qu'en est-il des pronoms démonstratifs dans une perspective guillaumienne ? Pour Guillaume, un pronom, quel qu'il soit, manifeste un phénomène de 'subduction'*, qui pourrait se gloser comme une dématérialisation du sens en faveur d'un mouvement de rappel de ce qui a été dit ou d'un mouvement d'appel de ce qui va être dit. Nous distinguons ici une définition possible de l'anaphore et de la cataphore. Les pronoms démonstratifs ne sont pas tant développés par Guillaume que par son plus fervent disciple, Moignet (1974 ; 1981), qui entrevoit les pronoms *celui* et *celle* de manière compositionnelle : *ce [qui est] lui*, *ce [qui est] elle*. De plus, il est à noter que Moignet regroupe les démonstratifs et les possessifs sous l'étiquette de 'pronoms dépersonnels' (1981 : 169), en distinction des pronoms personnels que sont *je/tu/il*. L'analyse guillaumienne de *ce* – en comparaison avec le *il* explétif –, mais aussi de *cela* et *ça* – *ça* étant considéré comme la forme orale de *cela* –, part d'une observation distributionnelle : *il* explétif ne s'emploie qu'en sujet de verbes dits 'impersonnels' (***il** pleut ; **il** faut que tu viennes*) ; *ce* peut être support de 'phrase nominalisée'[39] (***ce que tu me dis** est très intéressant*) ou sujet du verbe *être* (***c'**est mon frère qui est venu hier*) ; *cela*/*ça* n'est soumis à aucune règle restrictionnelle, sauf celle de ne pas être sujet du verbe *être* – *cela*/*ça* est donc perçu comme complémentaire de *ce*. Les propriétés syntactico-sémantiques de ces trois formes pronominales sont définies par le rapport qu'elles entretiennent avec 'la personne d'univers'[40] – ou encore 'personne cardinale' ou 'personne logique' – ainsi qu'en fonction de

[39] En Psychomécanique, une 'phrase nominalisée' est une proposition – en règle générale, relative ou conjonctive – à laquelle est substituable un syntagme nominal, la 'nominalisation' étant l'opération qui transforme une phrase en nom de discours.

[40] Aussi nommée 'personne cardinale' ou 'personne logique', la 'personne d'univers' désigne « la personne dont il est parlé », présente sous toutes les 'personnes ordinales' (*je, tu, il*). Guillaume dissocie 'personne ordinale' et 'personne d'univers' parallèlement à la distinction Moi/Hors-Moi, termes qui définissent

l'apport sémantique dont elles ont besoin pour référer. *Il* explétif est 'impersonnel' ou encore 'unipersonnel' car le verbe en présence n'appelle pas d'incarnation du sujet parlant – en termes guillaumiens, le verbe impose un 'refus d'incidence'* à la personne humaine, *il* explétif est alors uniquement lié à la personne d'univers. *Ce*, contrairement à *il*, n'est pas unipersonnel mais 'pluripersonnel', car sa référence sort des frontières imposées par le verbe – en termes guillaumiens, *ce* appelle une 'sémantèse'* extérieure à la sémantèse verbale – ce qui permet à *ce* d'apparaître dans des structures disloquées, telles que *c'est utile, la réflexion*, alors que nous ne pourrions avoir **il est utile, la réflexion* (exemples issus de Guenette, 1997). Aussi, il est remarqué que la référence de *ce* dépend irrémédiablement d'un complément quelconque, qu'il s'agisse d'une relative – *ce qui t'arrive est fantastique* – ou d'une conjonctive – *nous devons nous préparer à ce qu'il soit de mauvaise humeur* –, ou d'une structure copulative – *je ne pourrai pas vous accompagner ce weekend, c'est bien dommage*. Ainsi, *ce in posse* est décrit comme un élément sémantiquement vide en attente de remplissage[41], le contenu additionnel pouvant être un élément statique – ***cet homme**, c'est son père* – ou événementiel – ***conduire la nuit**, c'est un calvaire pour moi*. *Ceci*, *cela* et *ça* sont analysés comme étant composés de *ce* auquel est adjoint l'une des particules locatives *ci* et *là*. Si *ce* est compris comme un espace vide en attente de contenu, *ceci/cela/ça in posse* sont décrits comme des formes pleines, preuve en est qu'elles ne requièrent pas la présence obligatoire d'un complément. Le contenu sémantique de l'une ou l'autre de ces formes est supposé se déterminer en discours, que ce contenu provienne de la situation d'énonciation – auquel cas nous aurions affaire à une référence exophorique – ou d'un élément du discours antérieur – référence endophorique.

La théorie et les analyses de Guillaume sont motivées par la recherche d'un invariant : tout compte fait, les propriétés intrinsèques de chaque élément de la langue sont les conditions minimales et suffisantes à tout emploi observé dans le discours. Comme le note Guillaume lui-même, « La vraie réalité d'une forme, ce ne sont pas les effets de sens multiples et fugaces qui résultent de son emploi, mais l'opération de pensée, toujours la même, qui préside à sa définition dans l'esprit » (Guillaume, 1929). Or, cet invariant est intimement lié à l'idée d'un

respectivement « la personne pensante et parlante » et « la personne qui est en permanence au fond de la pensée ». Exemple : *je* instancie le sujet parlant (première personne ordinale) qui parle de lui-même (personne d'univers délocutée).

[41] La qualification 'vide en attente de remplissage' est reprise par Lapaire & Rotgé (1992) en référence à la valeur invariante des éléments en WH- de l'anglais, tels *which*, *who* et *whose.*

contenu sémantique, comme en témoigne le terme même de 'sémantèse'. Attribuer un invariant de sens aux unités langagières n'est pas sans poser problème, surtout lorsque l'on sort des balises de l'étude synchronique pour s'étendre à la diachronie – et ce, même si les analyses guillaumiennes s'attachent à argumenter leurs démonstrations à la lumière des changements historiques. En effet, comment justifier d'un invariant de sens, inscrit dans les 'gènes' du mot[42], si ce dernier ne cesse de varier au cours du temps ? De plus, les fondements de la différence postulée entre signifié de puissance et signifié d'effet, à savoir valeur unique et permanence *vs* valeurs multiples et fugacité, expriment un rejet explicite de la polysémie : la valeur de puissance en langue étant considéré par Guillaume comme « d'une autre espèce » (Guillaume, 1945[1970] : 12) que les valeurs en discours, nul ne saurait dans ce cas concevoir une polysémie du mot[43]. Le postulat d'un invariant comme contenu sémantique propre à chaque item langagier pourrait donc sembler hautement discutable. Néanmoins, l'approche guillaumienne de la langue en tant que système dynamique, actualisé par l'emploi qu'en fait le sujet parlant dans un cadre énonciatif donné, est un principe absolument essentiel pour quiconque s'attache à décrire le fonctionnement langagier. Ce postulat constitue l'un des axiomes majeurs de l'étude des démonstratifs que nous développerons dans la suite de cette thèse.

3.2. Les pronoms chez Benveniste (1956)

Benveniste est considéré comme le père de la linguistique de l'énonciation en France. Même si nous avons déjà employé le terme d''énonciateur' (cf. chapitre II, section 3.2.1, note 16), c'est Benveniste qui introduit ce terme en 1956 en l'opposant à 'locuteur'. Dans *La nature des pronoms* (1956[1966]), Benveniste décrit les pronoms comme appartenant à la syntaxe et au discours, puisque ces éléments langagiers sont des instances de discours : la référence qu'ils effectuent résulte d'une actualisation idiosyncrasique. Selon lui, *je* et *tu*, indicateurs de personne, ne sont que des réalités de discours : l'identification du référent n'est possible que par un retour à la situation d'énonciation, à chaque fois que l'occurrence apparaît. En revanche, la troisième personne *IL*, mais aussi *ce*, *ceci*, *cela* – les pronoms qui ne

[42] Si la notion d'un pôle d'invariance propre à chaque élément de la langue est un concept présent chez Adamczewski, Culioli, Guillaume et Lapaire & Rotgé, la découpe submorphémique est principalement adoptée par les linguistes anglicistes.

[43] Guillaume souhaite se démarquer de la polysémie développée par les diachroniciens localistes, qui établissent une hiérarchie des signifiés d'un même terme observés en discours, où les 'sens figurés' sont dérivés d'un 'sens propre', si possible à connotation spatiale.

sont pas *je* et *tu* –, instancient la « non-personne » (1956[1966]) : 255) puisque ces pronoms sont non réflexifs et renvoient à une entité objective[44]. D'après Benveniste, la caractéristique principale de ces items est de servir de « substituts abréviatifs de matériel de l'énoncé » (1956[1966] : 256) : pour des raisons d'économie, l'énonciateur choisit d'utiliser l'un ou l'autre de ces pronoms en représentation d'une « partie du discours » (*op. cit.*)[45] qui peut s'étendre du nom à la proposition. Par cette description, Benveniste fait sans nul doute allusion à la référence anaphorique que peuvent opérer les pronoms de troisième personne, d'autant plus qu'il ajoute que, dans ce cas précis, les occurrences de *IL*, *ce*, *ceci* et *cela* ne font pas partie du même paradigme que les déictiques *ici* et *maintenant*. Le paramètre qui détermine la complémentarité distributionnelle des pronoms non réflexifs et des déictiques est ici l'ancrage au *hic* et *nunc* énonciatif. Néanmoins, Benveniste adjoint les démonstratifs aux déictiques tels que *ici* et *maintenant*, lorsque « l'identification de l'objet [se fait] par un indicateur d'ostension concomitant à la présente instance de discours contenant l'indicateur de personne » (1956[1966] : 253). Cette définition n'est pas sans rappeler le repère déictique tel qu'il est décrit par Bühler (1934[2009], cf. chapitre II, section 1). Toutefois, *La nature des pronoms* n'expose pas de point de vue tranché entre vision dichotomique et conception dérivée de l'anaphore et de la deixis.

3.3. Les démonstratifs français dans la Théorie des Opérations Prédicatives et Enonciatives

3.3.1. Rappels théoriques

La Théorie des Opérations Prédicatives et Enonciatives (TOPE) est une théorie du langage développée par Culioli. Si la TOPE se classe parmi les linguistiques de l'énonciation, elle s'en démarque par l'approche choisie, radicalement différente des théories voisines. Ici, l'acte de langage n'est pas vu comme l'actualisation d'un système langagier par un énonciateur quelconque, mais comme une construction du sens. Ainsi, l'espace cognitif

[44] Dans la vaste littérature existante sur les déictiques, *je* et *tu* sont souvent référencés comme 'déictiques purs' ou 'directs', tandis que *il* et les démonstratifs sont appelés 'déictiques opaques'.

[45] L'emploi anaphorique des pronoms chez Benveniste est décrit comme suit : « […] les formes telles que *il*, *le*, *cela*, etc. ne servent qu'en qualité de substituts abréviatifs. […] ils remplacent ou relaient l'un ou l'autre des éléments matériels de l'énoncé. […] c'est une fonction de « représentation » syntaxique qui s'étend ainsi à des termes pris aux différentes « parties du discours » (1956[1966] : 256, souligné par nous). Par cette citation, nous devinons que 'discours' est ici amalgamé à 'texte' (cf. glossaire pour ces termes) ; nous pouvons donc supposer que Benveniste postule une définition endophorique de l'anaphore.

humain abrite un ensemble de notions – concepts dotés de particularités physico-culturelles – antérieures à toute lexicalisation. Ces notions sont ensuite organisées selon un schéma à trois places afin de former des lexèmes (niveau 1), qui gagnent à leur tour un rôle d'actant afin de former la structure syntactico-sémantique de l'énoncé à venir (niveau 2), qui ne sera effectif que lors d'un repérage énonciatif dont l'origine est l'énonciateur (niveau 3). Les opérations décrites par la TOPE sont complexes et variées mais sont toutes centrées sur la notion de repérage : tout terme est repéré par rapport à un terme repère. Le repérage est une opération essentielle au niveau des constructions énonciatives en ce qu'elle est à la base de la détermination et donc de la construction d'un énoncé par rapport au point origine du repère énonciatif. Les opérations de repérage cruciales dans le traitement de la détermination chez Culioli sont l'extraction et le fléchage. L'extraction est une opération de construction d'occurrence à partir d'une notion qui sélectionne une occurrence parmi une classe d'occurrences, prototypiquement effectuée par l'article indéfini – ***un appartement en ville*** *est bien plus pratique qu'****une maison à la campagne***. Le fléchage est une opération de construction d'occurrence qui se base sur l'extraction. Sont distingués le fléchage situationnel étroit et large – qui a recours à la situation d'énonciation – ainsi que le fléchage contextuel étroit et large – qui nécessite la présence d'un élément linguistique. Le fléchage est prototypiquement réalisé par l'article défini – *J'ai acheté une veste et une paire de chaussures ;* ***la veste*** *est en cuir*.

3.3.2. *Ce, cette, ces* dans la Théorie des Opérations Prédicatives et Enonciatives

Les questions que pourrait soulever l'appartenance ou non des démonstratifs aux marqueurs de fléchage ont été posées au chapitre II, section 3.3.3, lors de la présentation de *THIS* et *THAT* tels qu'ils sont analysés dans la TOPE. En effet, la valeur tantôt anaphorique, tantôt déictique que peuvent prendre les démonstratifs en discours – niveau 3 chez Culioli – remet en question la définition du fléchage comme étant une opération seconde à l'extraction, certains linguistes préférant le terme d'opération double afin d'éviter l'assignation d'une vertu immuablement anaphorique à l'opération de fléchage. Les mêmes interrogations sont suscitées par le traitement des démonstratifs français, le fonctionnement opératif de *ce, cette* et *ces* étant décrit de manière analogue à celui de *THIS* et *THAT*. Ainsi, *ce, cette* et *ces* sont affiliés aux déictiques, puisqu'ils sont utiles pour repérer des « objets » dans l'espace-temps énonciatif, par rapport à l'énonciateur – 'fléchage situationnel étroit' en termes culioliens (cf. chapitre II, section 3.3.3). Les particules adverbiales *ci* et *là* seraient propices à renforcer ce

repérage par l'indication locative qu'elles fournissent, *ci* désignant un « objet » situé dans la sphère de l'énonciateur, *là* signalant un « objet » hors de cette sphère (nous reviendrons sur cette idée de sphère locative dans les chapitres VII et VIII, lors de notre étude des démonstratifs anglais et français). L'appartenance de *ce, cette* et *ces* à la famille des déictiques les différencie de l'article défini, marqueur de fléchage, en ce que l'« objet » désigné serait nouveau. Néanmoins, l'ajout de la particule adverbiale *là* permettrait à ces déictiques d'acquérir une valeur anaphorique – ***cet étudiant-là**, dont on parlait il y a quelques minutes, est passé me voir hier* (Tchagbalé, s.d.). L'ajout de *ci*, quant à lui, indiquerait un contraste systématique avec un « *ce* + N-*là* » – *je veux **cette robe-ci**, pas cette robe-là* (Giancarli, 2004). Notons par ailleurs que les démonstratifs peuvent opérer le fléchage contextuel strict, après une opération d'extraction – *j'ai interpelé un homme hier. **Cet homme** avait apparemment beaucoup de choses à se reprocher*. Par conséquent, l'analyse des démonstratifs au sein de la TOPE rend compte des différentes fonctions discursives réalisées par ces items lors d'un ancrage énonciatif, par les types de fléchage opérés. L'analyse que nous proposerons dans les chapitres VII et VIII rejoint, par certains côtés, celle de Culioli – notamment par la notion d'opérateur propre aux démonstratifs ; néanmoins, notre description du fonctionnement indexical donne une importance majeure aux contextes et à la mémoire (cf. chapitre IV pour l'explication de ces termes et la terminologie que nous choisissons d'employer).

A la manière du chapitre II, le chapitre III parcourt un siècle de traitement des démonstratifs français, de l'analyse structurale à celles des linguistiques de l'énonciation. Le système des démonstratifs français s'avère complexe : d'une part, déterminants et pronoms ne partagent pas les mêmes formes ; d'autre part, il existe une complémentarité distributionnelle au sein même de la classe des pronoms démonstratifs. De plus, les analyses présentées dans ce chapitre semblent s'accorder sur le mode de détermination opéré par les démonstratifs français, qui peut se rapprocher ou se distinguer de celui du défini. Enfin, il ressort de ce panorama un accord plus ou moins explicite sur 'montrer le réel' – 'deixis', 'exophore' – et sur 'montrer dans le texte' – 'anaphore', 'endophore'. Voici donc venue l'heure du bilan entre démonstratifs anglais et français, à la lumière de cet état de l'art.

Bilan des chapitres II et III. Les ressemblances et les divergences entre démonstratifs anglais et français

Le panorama offert par les chapitres II et III permet de dégager certaines propriétés des démonstratifs anglais et français mais soulève également un certain nombre de questions.

Tout d'abord, déterminants et pronoms partagent les mêmes formes en anglais, ce qui n'est pas le cas pour le français. Notre étude étant consacrée aux pronoms démonstratifs, voici la classification que nous pouvons retenir :

<table>
<tr><th></th><th>Anglais</th><th>Français</th></tr>
<tr><td rowspan="2">Marqués en genre et nombre (en anglais, seulement pour le nombre)</td><td>This one, these ones, that one, those ones

(problème : ici, le pronom est one, pas le démonstratif qui est ici déterminant)</td><td rowspan="2">Celui-ci (prox)/ceux-ci, etc.

Celui-là (distal)/ceux-là, etc.</td></tr>
<tr><td>This (prox), that (distal), these, those</td></tr>
<tr><td>Neutres</td><td>It[46]

problème : it n'est pas un démonstratif</td><td>Ceci, cela (distance marquée)

Ce, c'

ça</td></tr>
</table>

En anglais, le système binaire des pronoms démonstratifs passe au ternaire avec l'ajout de *it*, ce dernier étant complémentaire des deux autres dès que l'on étudie la référence neutre. En français, la situation semble plus complexe. Dans les deux langues, l'idée d'un défini sous-jacent[47] jalonne la littérature, ce qui revient à poser le problème de la détermination effectuée par les démonstratifs : si ces derniers comportent un défini, ils doivent cependant comprendre une caractéristique supplémentaire. Ainsi, tandis qu'un SN défini désigne un « objet »

[46] Nous expliquerons l'intégration de *it* au sein de notre système de pronoms démonstratifs neutres dans le chapitre VII.

[47] L'idée d'un défini sous-jacent aux démonstratifs trouve sa source dans l'évolution de ces unités grammaticales, le défini étant issu des démonstratifs (cf. chapitre IX).

correspondant à ce qui est dénoté par la tête lexicale – ce que Corblin (1987) nomme la 'référence virtuelle' –, les démonstratifs autorisent un ajout d'information via la tête lexicale du SN (cf. Schnedecker, 2006), ce qui transparaît dans les exemples suivants :

(1) a. *Un chasseur est arrivé hier ;* ***cet Allemand*** *a manqué tous ses tirs*. (exemple (24) de Milner, 1982)

b. **Un chasseur est arrivé hier ;* ***l'Allemand*** *a manqué tous ses tirs*.

En (1a), *un chasseur* et *cet Allemand* sont deux expressions co-référentielles, ce qui n'est pas le cas en (1b). De même, les SN démonstratifs permettent de véhiculer un jugement du locuteur (cf. les remarques de Dubois, 1965, présentées au chapitre III, section 1.3).

D'autre part, en français, il est remarqué que *CELUI* se classe difficilement parmi les pronoms démonstratifs, au même titre que le seraient *CELUI-ci* et *CELUI-là* (cf. chapitre III, section 1.4.1). En effet, *celui*, *celle(s)* et *ceux* ne fonctionnent pas de manière autonome mais requièrent une extension de type syntagme prépositionnel, proposition relative ou syntagme adjectival, à la manière de *the one* en anglais. En revanche, l'ajout de la marque distale *là* ou proximale *ci* permettrait à ce 'semi-pronom' d'accéder au statut de pronom. Par conséquent, nous pourrions intuitivement déduire les correspondances suivantes :

Anglais	Français
The one	*Celui*
This one	*Celui-ci*
That one	*Celui-là*
This	*Ceci ?*
That	*Cela ?*
It	?
?	*Ça*

Il serait tentant de passer par la traduction afin de compléter ce tableau, mais nous pouvons aisément constater la non-correspondance unaire entre les éléments de chacune de ces deux langues : qui n'a pas été confronté au problème du défini anglais *the* traduit par le démonstratif français *ce* et *vice-versa* ? Il est donc nécessaire de décrire le fonctionnement référentiel de ces pronoms pour rendre compte de leurs similitudes et de leurs différences.

Enfin, et ce pour l'anglais comme pour le français, se pose le problème de la référence 'externe' et de la référence 'interne' – nous avons tour à tour employé les termes de 'deixis' et d''exophore' pour le premier type de référence et d''anaphore' et d''endophore' pour le second. En effet, les démonstratifs sont des déictiques en ce qu'ils peuvent être interprétés grâce à un ancrage spatio-temporel, mais ils peuvent également être anaphoriques, c'est-à-dire dépendre de la référence d'un autre élément langagier afin de pouvoir être interprétés. La pomme de discorde est ici la relation qui existe entre dimension extralinguistique et dimension langagière : doit-on séparer le réel et le langage ? Comment s'effectue le pont entre les deux ? Ces deux questions fondamentales, ainsi que bien d'autres encore, sont au cœur du débat proposé au chapitre IV.

Chapitre IV. La référence et l'indexicalité*

Un élément central au langage et à la communication, dont nous n'avons pas encore évoqué le nom, est la cognition. Si la cognition est le fonctionnement de l'esprit et que ces mécanismes de la pensée sont propres à l'Homme, toute théorie du langage pourrait en principe se réclamer de la linguistique cognitive. Ce qui différencie les différentes écoles linguistiques, outre leur origine géographique et leur filiation scientifique, est le point de vue qu'elles portent sur le langage. Ainsi, les grammaires génératives conçoivent la faculté de langage comme génétique, les grammaires cognitives considèrent les mécanismes psychiques du langage comme analogues à ceux intervenant dans d'autres activités cognitives et les grammaires fonctionnelles envisagent le langage comme un instrument grâce auquel les individus peuvent communiquer. Par conséquent, les grammaires génératives s'allient au concept d'innéité tandis que les deux autres courants sont unis par la notion d'émergence. Au niveau méthodologique, cette différence se retrouve dans le caractère déductif ou inductif de leur recherche. Nous serions encline à nous inscrire dans la lignée des linguistiques fonctionnelles, en ce que le langage comme instrument de communication induit une analyse essentiellement pragmatique, interactionnelle et prenant en considération les choix du locuteur de l'emploi de telle ou telle forme langagière. D'un autre côté, nous ne rejetons pas certains mécanismes décrits par les linguistiques cognitives. Ainsi, nous pourrions qualifier notre approche de 'cognitivo-fonctionnaliste'. Toutefois, et ce quel que soit le positionnement théorique choisi, l'étude de la référence et de l'indexicalité – n'oublions pas que notre thèse porte sur les pronoms démonstratifs – ne peut s'effectuer sans définir ce qu'est la réalité et la place que celle-ci occupe dans les phénomènes langagiers. En effet, c'est à partir d'une stabilité conceptuelle du réel que peut s'envisager la description de la référence et des phénomènes qui lui sont liés, comme la deixis et l'anaphore. Le chapitre IV s'attache à définir cet ensemble de notions : dans un premier temps, nous spécifions quelle conception nous avons de la réalité (section 1), ce qui permet de préciser la terminologie que nous adoptons (section 2) pour notre étude ; bien que nous choisissions de ne pas nous inscrire dans un modèle particulier de formalisation du langage, nous nous inspirons de certaines théories qu'il est nécessaire de présenter (section 3) afin de caractériser les phénomènes de deixis et d'anaphore (section 4).

1. La référence, entre langage et philosophie

‘Se référer à’, ‘en référence à’, ‘x réfère à y’… Ces termes peuvent sembler triviaux, d’usage tout à fait banal et courant. Preuve en est que quiconque est à même de les prononcer au quotidien, sans avoir à recourir à l’usage d’un dictionnaire, mais néanmoins dans des contextes relativement différents. Qui plus est, il serait peut-être plus délicat de s’essayer à définir ces termes : ‘se rapporter à’ ? ‘vouloir dire’ ? ‘parler de’ ? ‘désigner’ ?... Ce problème n’est en aucun cas destiné au seul commun des mortels, ni même au seul linguiste : la référence est l’objet de débats houleux entre philosophes de tous bords depuis la nuit des temps. Tout du moins depuis l’Antiquité. En effet, si Frege (1892) apporte précisions et distinctions sémantiques – nous sommes bien confrontée au(x) problème(s) du sens –, l’étude de la référence ne peut s’effectuer sans la prise en compte de la connaissance, du réel et du rapport entre connaissance et réel, préoccupations au cœur des évolutions épistémologiques depuis près d’un demi-millénaire avant notre ère. Sans vouloir dresser un inventaire encyclopédique de tous les courants de pensée qui jalonnent et façonnent notre façon de concevoir le monde depuis l’Antiquité, ni remettre en cause des préceptes qui alimentent encore de nos jours les débats philosophiques et scientifiques, il est néanmoins nécessaire de présenter deux courants antagonistes concernant la connaissance et la réalité : l’objectivisme et le constructivisme.

1.1. L’approche objectiviste

L’objectivisme trouve sa source dans la philosophie aristotélicienne, qui pourrait se résumer en deux préceptes corollaires :

- Il existe une coupure nette entre l’entendement humain et le monde physique ;
- Le monde réel est objectif, en ce que les objets – au sens large – qui le composent existent indépendamment de la pensée humaine.

Ainsi, la table de bois noir sur laquelle nous sommes en train d’écrire ce chapitre existe même si notre œil n’est pas là pour la percevoir ; le fait que cet objet soit « une table », qu’elle soit « en bois » et de « couleur noire » constitue certaines de ses propriétés intrinsèques. Si nous transposons cette dualité entre esprit et matière et l’existence immuable de la réalité, nous pouvons conclure sans peine que les expressions linguistiques désignent des objets réels. C’est dans cette optique référentialiste que Frege (1892) pose plus tard les bases de la

sémantique vériconditionnelle, en distinguant 'Sinn' – que nous traduirons par 'sens'* – et 'Bedeutung' – que l'on traduit habituellement par 'référence'. Pour Frege, la référence d'une expression linguistique est ce qu'elle représente dans le monde, tandis que le sens est le mode de présentation de la référence, en d'autres termes, la façon dont une expression linguistique représente ce qu'elle représente. Dans ce système, le sens d'une expression est accessible à quiconque sait juger si cette expression est vraie ou fausse, à la condition de connaître les circonstances auxquelles s'applique cette expression – autrement dit, l'état du monde : nous pouvons accéder au sens du mot *table* car nous sommes capables de reconnaître si un objet du monde est une table ou non ; nous savons que l'objet « table » existe, même si nous ne percevons aucune table au moment où nous prononçons ou entendons ce mot. La référence d'une expression linguistique est l'objet du monde désigné par cette expression ; en langage formel, la valeur de vérité des circonstances à la fois nécessaires et suffisantes à sa désignation, dans un domaine de définition particulier, est vraie. Ainsi, le mot *table* trouve sa référence si l'objet du monde que nous percevons vérifie les propriétés intrinsèques de cet objet. Cependant, Frege souhaite distinguer le processus référentiel des expressions désignatives de celui des expressions prédicatives – *le ciel vs bleu* dans *le ciel est bleu*, par exemple, mais aussi *le Président de la République Française* et *un homme de gauche* dans *le Président de la République Française est un homme de gauche* – et sépare alors les termes singuliers (désignateurs) des termes généraux (prédicatifs). Notons ici que, toutes choses égales par ailleurs, *le Président de la République Française* et *François Hollande* réfèrent à la même portion de réalité, au moment où nous écrivons cette thèse ; de même, nous pourrions renvoyer à « François Hollande » par la description *l'ancien Premier Secrétaire du Parti Socialiste français* ; toutefois, l'accès à la référence s'effectue de manière différente, puisque les expressions linguistiques *le Président de la République Française* et *l'ancien Premier Secrétaire du Parti Socialiste français* ne décrivent pas la même propriété ; en d'autres termes, la référence est ici identique, mais le sens est différent. Revenons alors à la distinction entre termes singuliers et termes généraux : les termes singuliers dénotent – c'est-à-dire qu'ils renvoient à un aspect de la réalité – et désignent, tandis que les termes généraux dénotent seulement. De là, Frege ajoute un troisième niveau au couple sens/référence, celui de l'extension : l'extension d'une expression linguistique renvoie aux objets qui ont la propriété dénotée par cette expression – par exemple, « le ciel » est l'extension de « bleu ». Pour résumer, les termes singuliers réfèrent, ou dénotent, (à) des « objets », tandis que les termes généraux réfèrent, ou dénotent, (à) des propriétés ; l'extension renvoyant aux « objets »

porteurs de la « propriété » dénotée par une expression linguistique, l'extension ne s'applique alors qu'aux termes généraux :

	Sens	Référence	Extension
Termes singuliers	Propriété	Objet	
Termes généraux		Propriété	Objets

TABLEAU 2 : LE SYSTEME DE FREGE (1892)[48]

Comme nous pouvons le constater, la sémantique frégéenne s'inscrit dans le courant de pensée objectiviste : le sens est constitutif de la langue, qui fait elle-même partie des connaissances humaines. En cela, la langue représente la réalité objective, les objets du monde qui existent indépendamment de nous. De cette conception découle un lien indéfectible entre 'dénotation', ou 'référence', et 'existence' : pour qu'une expression linguistique puisse référer, il faut qu'elle soit en accord avec l'objet de la réalité qu'elle est censée désigner. En d'autres termes, il devrait être impossible de référer à un objet imaginaire. Que faire alors d'un terme général comme *licorne* dans *cet animal est une licorne* ? Tout un chacun conviendrait qu'une « licorne » n'existe pas ; néanmoins, nous serions capables de décrire les propriétés qui s'appliquent à une « licorne ». Que penser alors de l'expression singulière *le roi de France* dans l*e roi de France est chauve* ? A l'heure actuelle, il est vrai que « la France » n'a pas de « roi ». Ce problème est justement résolu par le système à trois niveaux de Frege, puisqu'il distingue l'extension des termes généraux de leur référence : un terme général comme *licorne* réfère à une ou des propriétés, tandis qu'un terme singulier comme *le roi de France* n'a pas de référence, dans les circonstances imposées par la configuration du monde actuelle. Toutefois, le problème posé par ces termes vides constitue

[48] Notons qu'à la suite de Frege, Carnap (1947) propose une simplification du système dissymétrique frégéen, en introduisant les concepts d''intension' et d''extension' : l'intension d'un terme est sa ou ses propriétés dénotées, tandis que l'extension renvoie à l'objet ou aux objets de la réalité qui possèdent cette ou ces propriétés. Mill (1896), quant à lui, contraste l'extension, qu'il nomme 'dénotation', avec la 'connotation' d'un terme : la connotation désigne la ou les propriétés qu'un objet doit posséder pour que ces propriétés soient dénotées par l'expression. De plus, Mill s'attache à décrire la référence directe manifestée par les noms propres, qui, selon lui, dénotent des objets sans connoter de propriété. Cette approche rejoint celle de Kripke (1972), qui décrit la rigidité des noms propres comme étant une relation unaire entre un nom propre et son extension : quelle que soit la manière dont l'extension est définie, elle reste la même ; en revanche, ceci n'est pas le cas pour les expressions définies, puisque l'extension varie selon l'état du monde à l'instant t_0. Nous reviendrons sur la référence des noms propres en section 4.2.1 *infra*.

l'une des objections majeures à la sémantique référentialiste et à la remise en cause du paradigme stipulant que le sens est une représentation abstraite de la réalité objective.

1.2. L'approche constructiviste

Le constructivisme[49] s'oppose à l'objectivisme, en ce qu'il nie l'existence d'une réalité indépendante de la perception. En d'autres termes, le monde tel que nous le percevons est la représentation cognitive, construite par les êtres humains, d'un monde parmi un ensemble de mondes possibles. Les fondements du paradigme constructiviste sont repris par Piaget (1937) pour expliquer le développement de la connaissance qui se construit progressivement lorsque l'individu, notamment pendant l'enfance, entre en interaction avec le monde qui l'entoure. Les préceptes constructivistes sont l'apanage des théories de la physique quantique – le fameux chat de Schrödinger, à la fois vivant et mort tant que l'œil humain n'a pas décidé de son état, ainsi que la perception des couleurs qui dépend crucialement de l'observateur – mais se retrouvent également dans la psychologie gestaltique, la philosophie de Schopenhauer et les jeux de langage de Wittgenstein. Si nous nous centrons sur les théories cognitives du langage contemporaines, nous pouvons citer, notamment, la linguistique incarnée de Lakoff (1980), la Grammaire Cognitive de Langacker (1987) ainsi que la Théorie des Espaces Mentaux de Fauconnier (1984). Il est un autre courant de théories cognitives du langage, les théories dites fonctionnelles, que nous développerons en section 2.4, qui héritent plus ou moins fortement des hypothèses constructivistes.

L'abandon de l'objectivité du monde ainsi que l'irruption du sujet parlant mènent à une nouvelle conception de la sémantique. Le sens n'est plus descriptif, puisqu'il n'existe plus de monde objectif, mais résulte d'une interaction entre le sujet et l'environnement physico-social. Par conséquent, la catégorisation due à la validation de conditions nécessaires et suffisantes, héritée d'Aristote, est supplantée par une théorie des prototypes : chaque prototype est le meilleur représentant de sa catégorie. De plus, l'hypothèse du 'tout cognitif' ainsi que la cognition incarnée de Lakoff laissent libre place à la métaphore (Lakoff & Johnson, 1980), aux tropes, à l'imaginaire – par exemple, l'expression *boire un verre* qui signifie, bien entendu, boire « le contenu du verre », et non pas « le verre lui-même », par métonymie. Ainsi, les expressions linguistiques n'ont plus de lien avec la réalité stable mais

[49] Le constructivisme est avant tout une théorie de l'acquisition. Toutefois, les concepts qui en découlent s'appliquent à la description théorique que nous faisons des problèmes soulevés par la référence.

renvoient à des entités mentales, nommées ‘référents’, ‘objets-de-discours’ ou encore ‘objets du discours’ dans les théories cognitives du langage. Comme nous pouvons nous en douter, une position aussi radicale du ‘tout est subjectif’ entraîne des conséquences tout aussi extrêmes qui peuvent sembler absurdes ou, tout du moins, contredire ce point de vue. Nous pensons ici au pessimisme de Quine (1977, 1992) sur l’accès au sens : si le monde est construit par le sujet psychologique, comment ce dernier peut-il envisager connaître et représenter ce monde par des expressions langagières ? Si le sens et la référence sont construits par le sujet parlant, comment l’interlocuteur peut-il partager le même sens et la même référence pour une expression linguistique donnée ? Ces questions, subsumées par un constructivisme radical doublé de relativisme, mènent à l’inévitable conclusion que la communication n’aura des fins qu’asymptotiques.

Bien entendu, il est des arguments en faveur de l’existence d’une réalité objective et de la conception frégéenne de la sémantique : « la table sur laquelle nous écrivons », que nous nommons *table*, est palpable, visible, définissable en fonction de critères qui semblent bel et bien inhérents à cet « objet ». De même, il est des raisons de croire à la subjectivité de l’univers : l’adjectif *rouge* renvoie à une certaine longueur d’onde captée par notre œil, sans lequel cette « couleur » n’est qu’une longueur d’onde. Cependant, l’une et l’autre de ces positions épistémologiques mènent à des aberrations scientifiques, notamment dès lors qu’on s’attache à l’étude du sens : l’approche véric onditionnelle exclut les fonctions imagées du langage – or, *boire un verre* signifie « boire le contenu » et non « le contenant » – tandis que le ‘tout cognitif’ peut conclure à une approximation perpétuelle de la communication. Comment alors ne pas tomber dans ces écueils ? Comment réconcilier réalité et subjectivité ? La section suivante s’attache à présenter certaines propositions de réponses.

1.3. L’objectivité en termes d’intersubjectivité

Putnam (1970, 1975) rejette la conception frégéenne du sens et de la référence, notamment les deux préceptes corollaires et conjoints selon lesquels (i) le sens – valable pour n’importe quelle configuration du monde – détermine la référence – valable selon un certain état du monde – d’une expression linguistique et (ii) l’accès au sens, et, par conséquent, à la référence, dépend des connaissances du monde du sujet (cf. ce chapitre, section 1.1). La réfutation de ces principes s’illustre chez Putnam par les termes linguistiques qui dénotent la classe des espèces et substances naturelles. En effet, le mot *chat* dénote invariablement « la classe des chats », même si le sujet parlant ne sait pas définir ce qu’est « un chat » ; or, si la

dénotation est déterminée par le sens, et que ce sens n'est accessible qu'à la condition que le sujet parlant soit capable de reconnaître ce qu'est « un chat », le mot *chat* ne devrait pas avoir de dénotation – ou de référent – si le sujet ne sait pas ce qu'est « un chat ». Cette contradiction apparente mène Putnam à modifier l'assertion (ii) en ne basant plus le sens d'une expression linguistique sur les seules connaissances du sujet parlant, mais sur le principe d'intersubjectivité. Ainsi, Putnam conserve la vision d'un monde objectif qui existe indépendamment de l'esprit humain, mais le sens d'un terme linguistique censé représenter un « objet » de ce monde n'est plus viable grâce à la seule aptitude du sujet parlant à reconnaître cet « objet », car ce dernier a la possibilité de recourir à un expert pour connaître le sens de ce terme. Si nous reprenons l'exemple du mot *chat*, ce n'est pas parce qu'un sujet parlant donné ne sait pas reconnaître ce qu'est « un chat » que « la classe des chats » n'existe pas ; il suffit au sujet de chercher dans un dictionnaire, ou, mieux encore, de demander à un scientifique, et le sens du mot *chat* lui apparaîtra. L'objectivité de la représentation est donc assurée par la ratification collective.

La notion d'intersubjectivité dans la définition du sens et de la référence est développée par Kleiber (2001), notamment dans l'acception du réel qu'il propose. Selon Kleiber, dont nous partageons le point de vue, que la réalité soit objective ou construite par la pensée humaine ne change pas grand-chose. Le point crucial est que cette représentation soit commune à l'ensemble de la communauté. Par là même, si personne n'a accès au réel vraiment réel, mais que chaque être humain partage la même conception du réel, il devient superflu de préciser *ad vitam aeternam* que ce que nous croyons être la réalité objective ne l'est pas. Si le monde perçu par l'ensemble de la communauté est le seul accessible, nous pouvons l'assimiler au réel. Bien entendu, nous pourrions trouver des contre-arguments à cette hypothèse, comme la dénomination des couleurs, qui comme chacun sait, est loin d'être uniforme d'un point de vue interlinguistique – le pirahã[50], par exemple, paraphrase l'adjectif *rouge* en 'ça ressemble à la couleur du sang' – ou comme la perception non visuelle de l'environnement physique qu'ont les non-voyants. Cependant, ce n'est pas parce qu'un item unique ne désigne pas la couleur que cette dernière n'est pas perçue et les non-voyants, malgré leur cécité, ont une appréhension du monde physique relativement similaire à la nôtre. De manière générale, la représentation cognitive du monde est semblable d'un individu à l'autre, ce qui permet tout d'abord de confondre cette conception avec une réalité objective,

[50] Le pirahã est une langue parlée par un peuple vivant au Brésil, le long d'un des affluents de l'Amazone.

puisque collective, et, surtout, de considérer cette représentation comme une norme. L'idée d'une réalité intersubjective comme point de repère permet d'éviter les écueils constructivistes radicaux – notamment le pessimisme de Quine où toute véritable communication est impossible – tout en permettant de concevoir des mondes imaginaires, d'appréhender les métaphores, les métonymies, et autres figures rhétoriques propres au langage. En définitive, nous pouvons distinguer ce qui est réel de ce qui ne l'est pas, ce qui est vrai de ce qui est faux.

Comme le montre Kleiber (2001), les avantages qu'offre la conception d'une réalité intersubjective au sens et à la référence sont nombreux. Tout d'abord, elle permet de définir la dénomination – *x est le nom de X* – et la désignation – *x renvoie à X* – réalisables par les expressions linguistiques. Que l'on ait affaire à la dénomination – ***un canif*** *est le nom d'un petit couteau de poche, composé d'une ou de plusieurs lames pliables* ; *un petit couteau de poche, composé d'une ou de plusieurs lames pliables, s'appelle* ***un canif*** – ou à la désignation – ***un canif*** *réfère à/renvoie à un petit couteau de poche, composé d'une ou de plusieurs lames pliables* – nous sommes face à une relation sémiotique du type « x → X », où x représente l'entité linguistique et X l'« objet » du réel. Cependant, nous ne pouvons avoir de relation dénominative si une expression descriptive remplace l'item dénominatif :

(1) a. **L'outil avec lequel les graveurs sur bois creusent différentes parties de leurs planches est le nom du petit couteau de poche, composé d'une ou de plusieurs lames pliables.*

b. **Le petit couteau de poche, composé d'une ou de plusieurs lames pliables, s'appelle l'outil avec lequel les graveurs sur bois creusent différentes parties de leurs planches.*

tandis qu'un énoncé désignatif est tout à fait acceptable :

(2) *L'outil avec lequel les graveurs sur bois creusent différentes parties de leurs planches réfère au/renvoie au petit couteau de poche, composé d'une ou de plusieurs lames pliables.*

La différence entre la dénomination et la désignation est que la dénomination se base sur une présupposition d'existence : je ne peux appeler une « chose » par son nom que si celle-ci a

déjà été nommée auparavant. Cette présupposition d'existence réside justement dans le caractère intersubjectif de ce que nous concevons des « choses » de ce monde. Ce phénomène est particulièrement visible avec les noms qui renvoient à des « catégories de choses ». Supposons que A indique son adresse à une personne B :

(3) A : *J'habite en face de la pharmacie.*

B est tout à fait à même de reconnaître ce qu'est « la pharmacie » car il relie ce terme à une « catégorie de choses » dotée de concepts généraux et communément admis – « un magasin où l'on vend des médicaments, en général avec une insigne verte en forme de croix, etc ». Notons ici que la présupposition d'existence n'est pas uniquement valable pour les noms. Si nous reprenons l'exemple bien connu de Russell (1905) *Le roi de France est chauve*, *Le roi de France* présuppose l'existence d'« un individu x qui soit roi de France », mais *chauve* présuppose également qu'il existe « la propriété d'être chauve ». La présupposition d'existence est cruciale dans le fonctionnement dénominatif des expressions langagières, comme nous pouvons le constater avec le test des mots inventés : si nous rencontrons au sein d'un syntagme référentiel le mot *podillon*, dont nous ne connaissons pas la définition, nous en déduirons que ce terme renvoie à une entité quelconque qui appartient au monde physique. Qui plus est, grâce à l'apport informatif des éléments qui l'entourent – par exemple, *je vais manger un podillon au bord de la mer* –, nous aurons le réflexe de l'assigner à une catégorie de choses qui pourraient lui ressembler –, en l'occurrence, à « quelque chose de comestible, peut-être à base de poisson ». Nous voyons bien ici la primauté du réel par rapport aux mondes imaginaires : d'instinct, nous nous accommodons d'un item lexical dont nous ignorons le sens en le rapprochant d'un objet ou d'un concept que nous connaissons et qui fait partie de notre environnement. Cette primauté du réel est effacée dans les théories constructivistes radicales, puisque le 'tout cognitif' place la conception du monde construite par la pensée humaine au même niveau que les autres mondes possibles. Il y a donc nécessité d'une norme, qui se trouve être celle du monde 'extralinguistique' tel que l'ensemble des individus le conçoit, à partir de l'expérience qu'il en a. Par conséquent, nous pouvons déduire que, si un item lexical renvoie à une catégorie, il y a présupposition d'existence et, donc, préconstruit ; le sens est ici codé, conventionnel, intersubjectif. A l'inverse, il n'y a pas de préconstruit conceptuel à l'emploi d'une expression à plusieurs termes – par exemple, *le petit couteau de poche, composé d'une ou de plusieurs lames pliables* –, d'où la possibilité d'avoir un énoncé désignatif en (2). Dans ce cas-ci, le sens s'élabore par construction : à la manière

de la compositionnalité gestaltique, le sens global de l'expression se compose de la somme des sens des termes qui la forment. La différence entre sens préconstruit et sens construit se ressent par ailleurs intuitivement : « un canif » n'est pas tout à fait la même chose qu'« un petit couteau de poche, composé d'une ou de plusieurs lames pliables », « détaler » est sensiblement différent de « courir vite », « ingurgiter » n'est pas la même action que « manger vite ». Les termes dénominatifs d'une catégorie – *canif*, *détaler*, *ingurgiter* – sont des unités formées d'une composante descriptive – « un petit couteau de poche, composé d'une ou de plusieurs lames pliables », « courir vite », « manger vite » –, d'un statut présupposé et d'une valeur indicative d'assignation à une catégorie de « choses ». Comme le dit Kleiber (2001), et comme le montre très justement Culioli avec sa forme schématique, une unité lexicale est formée d'une identification formelle fixe – la dénomination d'une catégorie – et d'une composante variable, descriptive, représentationnelle. D'une part, cette valeur ajoutée d'appartenance ontologique est rendue possible par la conception commune que nous avons de l'univers ; d'autre part, cette catégorisation est capitale car c'est elle qui permet à une simple information sémantique d'accéder au statut de référent. Nous reviendrons sur ce point lors de notre étude des pronoms démonstratifs en anglais et en français.

Plusieurs points en guise d'(inter-)conclusion : tout d'abord, afin d'éviter les écueils d'une vision rigoriste de la réalité comme existence objective ou, au contraire, d'un monde construit par la perception humaine au même titre que l'ensemble des mondes possibles, il convient d'attribuer à la perception que nous avons de notre environnement physique un caractère intersubjectif ; cette conception commune du réel permet la mise en place d'une norme, celle de notre réalité accessible, grâce à laquelle il est possible d'envisager différents espaces imaginaires et secondaires. Ce réel est composé de « choses », de « concepts », de « propriétés », que notre faculté de langage permet de désigner. Ainsi, les items lexicaux, en particulier les items ontologiques, renvoient à ces « catégories » préconstruites dotées de notions générales et communément admises. Par là même, il convient de distinguer très nettement le matériel langagier dont nous disposons et ces « choses » que nous percevons et désignons. La section suivante s'attache à définir les termes et notions indispensables à l'analyse des pronoms démonstratifs que nous effectuerons dans les chapitres VII et VIII ; en particulier, il est crucial de distinguer et relier 'référence' et 'indexicalité'.

2. Précisions conceptuelles et terminologiques

La première étape dans la description de notre modèle sémantique consiste à affiner ces « choses », « concepts » et « propriétés » qui appartiennent à notre réel intersubjectif. Pour cela, nous nous basons sur la typologie des entités établie par Lyons (1977) et complétée par Dik (1997).

2.1. La classification ontologique des entités*

L'entreprise de Lyons (1977) de classifier les entités référentielles ne part pas d'un simple motif philosophique teinté d'une quelconque prise de position sur la nature du réel – prise de position que nous pouvons cependant lire entre les lignes. La volonté de Lyons est de catégoriser les « parties du discours » (ce qu'il nomme « parts of speech », 1977 : 440 et suivantes), que nous choisissons d'appeler 'catégories morphosyntaxiques'[51], par la sémantique plutôt que par la syntaxe. Le postulat de base est donc que les catégories morphosyntaxiques – noms, adjectifs, verbes, adverbes, etc. – sont, du moins en partie, découpées par leur valeur sémantique ; cette hypothèse rend possible la description de langues, comme le nootka[52], censées ne pas faire d'opposition lexicale entre nom et verbe (Swadesh, 1938 : 78). Par conséquent, la classification de Lyons se situe à l'interface entre syntaxe et sémantique plus qu'au niveau sémantique *per se*, choix que nous discuterons ultérieurement.

La classification sémantico-syntaxique[53] de Lyons s'inscrit dans une articulation entre 'langue', 'sens' et 'dénotation' différente de celle établie par Frege (cf. ce chapitre, section 1.1). Si la définition classique de ces termes propose le sens d'une expression linguistique à la fois comme indépendant de la configuration du monde et déterminant pour la dénotation de cette expression dans un état du monde particulier, il n'en va pas de même pour Lyons qui

[51] Nous choisissons de ne pas employer 'parties du discours' ici afin de bien distinguer le 'discours', qui, comme nous le verrons ultérieurement, est une représentation cognitive produite lors d'un acte communicatif, des catégories morphosyntaxiques – ou encore, grammaticales – qui sont plus intimement liées à la langue, et, donc, au 'texte', dans notre acception.

[52] Le nootka est une langue de la famille des langues amérindiennes, parlée sur l'île de Vancouver.

[53] Nous choisissons 'sémantico-syntaxique' au lieu de 'syntactico-sémantique', comme nous le trouvons la plupart du temps, afin d'appuyer la démarche de Lyons qui souhaite décrire la syntaxe à partir de la sémantique, et non l'inverse.

oppose nettement ‘sens’ et ‘dénotation’, couple terminologique auquel il ajoute ‘référence’ – ‘dénotation’ et ‘référence’ ne sont donc plus interchangeables. Pour Lyons, le sens d’un terme linguistique autorise la délimitation des mots, par les connaissances que nous avons des propriétés associées à ce terme : le mot *chat* est une unité lexicale qui se différencie des termes *chien*, *table* et *ciel*, par exemple, mais qui entretient des relations plus ou moins proches avec les termes *animal*, *félin* et *siamois*, en faveur de ce que l’on croit être un « chat » : « animal de compagnie, mammifère, félin, de caractère indépendant, souvent ingrat, voire sournois ». Toujours selon Lyons, la ‘dénotation’ d’un terme renvoie à l’image mentale suscitée par ce terme ; ainsi, la ‘dénotation’ serait comparable à une photo dans le lexique mental[54]. Enfin, la ‘référence’ est la désignation d’un signifié par un signifiant, dans un ancrage énonciatif particulier : ***ce tableau*** *est immense* n’aura certainement pas le même référent si l’énoncé est prononcé dans une salle de classe ou dans un musée. Notons ici que Lyons introduit le paramètre crucial qu’est le contexte dans la description de la référence[55]. Ainsi, la classification de Lyons propose une hiérarchie des ‘entités’, et non pas des ‘référents’ – comme nous pouvons parfois le trouver dans la littérature existante –, pour appuyer le fait que sa typologie prend en compte les signifiés indépendants ou dépendants du contexte.

Cette mise au point terminologique effectuée, nous pouvons entrer en détail dans la hiérarchie des entités proposée par Lyons. Pour rappel, cette classification vise à dégager les catégories morphosyntaxiques par la sémantique, dans une volonté de description ‘universaliste’ – mais pas ‘universelle’, ne confondons pas élargissement typologique et Grammaire Universelle chomskyenne – des langues. Dans son approche, Lyons rejoint la théorie des rangs* de Jespersen (cf. chapitre II, section 1, note 5) ainsi que la logique des prédicats : il existe une corrélation entre le rang d’une catégorie morphosyntaxique (ce que le terme est à même de modifier) et sa portée sémantique (ce que le terme dénote) ; si l’on transpose ceci à la logique des prédicats, une expression de rang 1 dénote une entité de premier ordre, comprise comme « un individu » ou « une chose » (prototypiquement, les noms), les expressions de rang 2 dénotent des propriétés de premier ordre (on peut les

[54] Bien entendu, nous pourrions nous demander ce qu’il advient de la dénotation des termes « abstraits » : peut-on associer une image mentale au mot *courage*, par exemple ? La solution se trouverait peut-être dans le maintien de l’extension frégéenne des termes généraux (cf. ce chapitre, section 1.1 *supra*).

[55] Le contexte – ou les contextes – est un paramètre crucial dans l’étude de la référence et sera traité en section 2.3 *infra*.

attribuer à des entités de premier ordre ; il s'agit prototypiquement des verbes et des adjectifs), les expressions de rang 3 dénotent des propriétés de deuxième ordre ou plus (on peut les attribuer à des propriétés de premier ordre ; il s'agit prototypiquement des adverbes). Néanmoins, Lyons critique le principe véricondi tionnel de cette ontologie qui prédit une relation unaire entre la catégorie morphosyntaxique et l'ordre – ou le rang – de l'entité désignée. Par exemple, un nom serait indissociable d'une entité de premier ordre, or ce n'est non seulement pas le cas pour toutes les langues, mais il serait aussi difficilement concevable d'attribuer à *amour* la dénotation d'une « chose ». Lyons précise alors à bon escient qu'une ontologie des entités n'est valable que dans un but de catégorisation sémantique de sous-ensembles prototypiques, et non pas de la classe entière de chaque catégorie morphosyntaxique[56]. La hiérarchie de Lyons classe trois catégories d'entités selon leur niveau d'abstraction :

- Les entités de premier ordre sont définies dans l'espace, dotées de caractéristiques stables ; dans notre réalité intersubjective, elles ont une matérialité physique, elles 'existent'. De manière générale, les « individus », les « animaux » et les « choses » sont des entités de premier ordre. Les catégories morphosyntaxiques les plus à même de désigner ces entités sont certains noms – *chat*, *chien*, *eau*, *table*, où nous retrouvons les fameux noms à dénotation de classes et substances naturelles de Putnam –, les noms propres, les pronoms ainsi que les syntagmes nominaux ;
- Les entités de deuxième ordre sont définies dans l'espace et le temps, ce sont des « événements », des « procès », des « états de choses » ; nous pouvons dire des entités de deuxième ordre qu'elles 'arrivent'. Les structures langagières les plus à même de désigner ces entités sont certains noms – *mort*, *naissance*, *arrivée*, *départ*, etc. – ainsi que les verbes qui leur correspondent. Ainsi, nous pouvons constater que les substantifs prototypiques désignateurs d'entités de deuxième ordre résultent d'une nominalisation, processus particulièrement visible en anglais : *set*/*setting*, *leave*/*leaving*, etc. ;
- Les entités de troisième ordre sont de nature abstraite, en dehors de l'espace et du temps, ce sont des « concepts », des « sentiments » ; de manière générale, nous pouvons soumettre ces entités à la réfutabilité, à la croyance ou au doute. Les

[56] Même si Lyons dit se baser sur le réalisme naïf (1977 : 442) pour classifier les entités, il n'en reste pas moins que le postulat de catégories sémantiques prototypiques exclut son approche du cadre objectiviste.

structures propositionnelles sont les plus à même de renvoyer à des entités de troisième ordre, mais certains noms, comme *pensée*, *sentiment* ou *raison*, désignent également ce type d'entités. Tout comme ceux désignant des entités de deuxième ordre, nous pouvons constater que ces noms résultent d'une nominalisation : *penser*/*pensée* ('think/thought'), *sentir*/*sentiment* ('feel/feeling'), *raisonner*/*raison* ('reason (V)/reason (N)').

Plusieurs remarques concernant cette classification : tout d'abord, comme le soutient Lyons, le dégagement de catégories sémantiques prototypiques ne permet pas une découpe franche des catégories morphosyntaxiques : nous voyons bien que les noms sont à même de dénoter des entités de premier, deuxième et troisième ordres ; *exit* donc la correspondance unaire « catégorie → entité » présagée par les principes vériconditionnels. Par ailleurs, la frontière entre ordres d'entités est loin d'être étanche dans certains cas. Si la distinction entre entités de premier ordre et entités de deuxième ordre semble somme toute pouvoir se faire sans trop de difficultés – et encore, que dire, par exemple, du nom *repas* ? Dénote-t-il l'entité de premier ordre qu'est « la nourriture composée d'un ou plusieurs mets prise à certains moments de la journée » ou bien celle de deuxième ordre qu'est « l'acte de manger un ou plusieurs plats » ? –, il paraît plus délicat, voire nettement difficile, de séparer les niveaux d'entités, dès que l'on quitte le domaine de la dénotation pour celui de la référence. Prenons comme exemple un slogan publicitaire vantant les mérites d'une crème amincissante :

(4) *Cosmatoline Cosmétique, **ça** fonctionne !*[57]

Le pronom *ça* renvoie-t-il à « la crème Cosmatoline », auquel cas nous aurions affaire à une entité de premier ordre, ou bien à « l'application en massages de cette crème », qui ferait partie des entités de deuxième ordre ? Que dire également de l'emploi en usage et en mention de certains mots ? :

(5) a. *Le **rhododendron** est une fleur de montagne, cultivée comme plante ornementale*. (emploi en usage)

b. ***Rhododendron** est un nom commun composé de douze lettres*. (emploi en mention)

[57] Le pronom démonstratif neutre *ça* sera au cœur de l'étude dans le chapitre VIII.

Rhododendron en (5a) réfère à « l'objet physique, concret, végétal » qu'est le « rhododendron » dans notre réalité collective et qui est une entité de premier ordre. En revanche, *rhododendron* en (5b) réfère au « mot » lui-même, au « matériel linguistique » ; un « mot » est-il une réalité concrète ? Après tout, il est doté de caractéristiques orthographiques et phoniques stables… En même temps, un « mot » n'est-il pas un concept, et donc une entité de troisième ordre ? Bien entendu, répondre à une telle question renvoie aux débats linguistico-philosophiques sur la découpe du monde par le langage, sur l'arbitraire du signe et sur les conventions orthographiques. De manière générale, le passage de la dénotation à la référence est rendu possible par la contextualisation de l'énoncé considéré, mais également par la prédication dans laquelle s'inscrit l'emploi de telle ou telle expression référentielle – nous approfondirons ces notions dans les chapitres VII et VIII. Attention cependant, nous ne prétendons pas ici nous faire l'avocat du diable : la porosité entre les niveaux d'entités et les difficultés de catégorisation des référents ont déjà été remarquées, notamment par Apothéloz (1995), et par Lyons lui-même lors de la présentation de son ontologie (1977). Il faut donc bien garder à l'esprit qu'il s'agit d'un essai de classification de prototypes sémantiques et du rapport que ceux-ci entretiennent avec des catégories morphosyntaxiques également prototypiques. En revanche, Lyons soutient que les entités de deuxième et troisième ordres sont dérivées de celles de premier ordre. En effet, il remarque, à juste titre d'ailleurs, que la simple appartenance des « événements », des « procès », des « états de faits » et des « concepts » à la classe des « entités » les associe indéniablement à la notion d'existence (cf. ce chapitre, section 1.1), ce qui les ramène, dans une sorte de mouvement centripète, vers la catégorie des entités de premier ordre. Cette tendance à concrétiser un phénomène abstrait se nomme 'hypostatisation'* et est un argument supplémentaire pour l'interprétation cognitive du réel comme étant vraiment réel : la pensée humaine a besoin d'une réalité normative, constituée de « choses » et de « propriétés » préconstruites. L'hypostatisation est d'ailleurs une réponse possible au questionnement soulevé par Lyons sur les adjectifs dénotant une propriété de premier ordre – *rouge*, *rond*, *métallique*, etc. En effet, comment classer ces propriétés selon la hiérarchie proposée des entités ? Si certains logiciens se refusent à distinguer 'classe' et 'propriété', ces deux catégories sont tout de même distinguables dans notre vision du réel : nous différencions bien le « fruit » qu'est la « pomme », sa « couleur » qui est « rouge », même si la « pomme rouge » ne forme qu'un seul et même « objet ». Or, il semble évident que lorsque nous évoquons la couleur « rouge », nous l'associons instinctivement à l'« objet » auquel elle se rapporte. Le phénomène d'hypostatisation prend ici toute sa dimension, en ce que nous avons besoin du concret pour visualiser l'abstrait. Deux

solutions émergent alors pour la classification des propriétés de premier ordre dénotées par les adjectifs. Soit nous suivons Lyons, qui remarque que les adjectifs prototypiques se situent entre les noms prototypiques et les verbes prototypiques, à la manière de l'analyse proposée par la Grammaire Notionnelle Dépendancielle* de Anderson (1991, 1997, 2007, 2011) et, dans un certain sens, dans la mouvance des Grammaires Catégorielles* (cf. entre autres Morrill, 2010), auquel cas « rouge » sera associé aux entités de premier ordre et « rapide » aux entités de deuxième ordre, par exemple ; soit nous augmentons la hiérarchie de Lyons de deux catégories d'entités supplémentaires, comme le propose Dik (1997) : les entités d'ordre zéro sont alors des dénotations de prédicats :

(6) *Cette fille est intelligente, ma sœur* ***l'****est aussi*. (*l'* → « être intelligent(e) »)

Les entités de quatrième ordre sont des illocutions :

(7) *« Je ne croyais pas avoir épousé une aventurière. » Elle parut très froissée et me répondit : « Ce n'est pas une aventure, ça dure depuis deux ans. » Elle me fit* ***cette déclaration****, je dois le dire, avec une certaine dignité*. (Flers & Cavaillet, *Monsieur Brotonneau*, 1914)

Si les entités d'ordre zéro nous semblent judicieuses – une telle entreprise permet à la fois de rendre évident le fait que la dénotation d'un prédicat puisse accéder au statut de référent, notamment par la pronominalisation (cf. ce chapitre, exemple (6)), mais aussi de distinguer 'classe' et 'propriété' – les entités de quatrième ordre paraissent moins pertinentes : après tout, bien que cette déclaration indique un point de vue de l'énonciateur sur l'acte illocutoire de sa compagne, il réfère avant tout, à notre avis, au « contenu des paroles prononcées par celle-ci » ; auquel cas nous aurions 'simplement' affaire à une entité de troisième ordre.

Afin de faire le point sur la hiérarchie des entités, il est utile de la réinscrire dans le questionnement philosophique et sémantique dont nous avons débattu en section 1, mais aussi dans le cadre de l'articulation entre langage, structure de la langue et sémantique. Lyons nous dit se fonder sur un réalisme naïf pour le dégagement de ses ordres d'entités, tout en précisant qu'il tient compte de prototypes sémantiques. Nous pouvons sans peine et sans trop de conséquences fâcheuses transposer ce réalisme prétendument naïf vers une réalité intersubjective ; l'idée même de prototypes, ainsi que le phénomène d'hypostatisation qui ramène les entités de deuxième et troisième ordres à des « objets » concrets, semblent être des

arguments tout en faveur d'une réalité collective. Cependant, nous devons garder à l'esprit que la volonté première de Lyons est d'effectuer un pont entre catégories sémantiques et morphosyntaxe ; or, nous voyons bien les difficultés que cela entraîne : non seulement les catégories sémantiques sont perméables – nous avons constaté la porosité entre les différents ordres d'entités – mais la relation qu'elles entretiennent avec les catégories morphosyntaxiques est loin d'être unaire ; en outre, la classification de Lyons ne propose pas de lien logique entre ordre d'entité et catégorie morphosyntaxique : la découpe sémantique est instinctive et ne rend pas compte de la possible structure interne des éléments du lexique qui pourrait justifier le renvoi vers tel ou tel ordre d'entité. De plus, l'assignation d'une entité à l'un ou l'autre des ordres proposés est d'autant plus ambiguë lors de la prise en compte des référents et non des dénotations, en raison de la dépendance contextuelle de l'énoncé et de la prédication dans laquelle s'inscrit l'emploi de telle ou telle expression référentielle. Il conviendrait donc très certainement de séparer les niveaux de représentation, même si, au niveau du fonctionnement langagier, il est selon nous évident que syntaxe et sémantique sont liées. Attention cependant à ne pas jeter le bébé avec l'eau du bain : Lyons lui-même se défend d'interpréter cette hiérarchie comme un carcan catégoriel, car elle prétend être et est d'une grande aide à quiconque s'attache à décrire la référence, en se basant sur des cas prototypiques ; une fois de plus, nous revenons à ce recours à la norme. Cela va sans dire, cette classification référentielle constitue un cadre-clef de l'analyse des pronoms démonstratifs que nous effectuerons dans les chapitres VII et VIII ; dès lors que nous évoquerons les ordres d'entités, nous nous rapporterons à la classification ontologique de Lyons.

2.2. La 'référence', l''indexicalité', le 'texte' et le 'discours'

Nous suivons Lyons et sa distinction entre 'sens', 'dénotation' et 'référence', et insistons tout particulièrement sur la contextualisation, indispensable à l'accès à la référence – contrairement à la dénotation qui est comparable au rapport entre signifiant et signifié établi par Saussure. Le cœur de notre étude étant consacré aux pronoms démonstratifs neutres – *ça* en français, *this*, *that* et *it*[58] en anglais –, l'objet de notre débat est par conséquent la référence, mais aussi l'indexicalité. Si la référence est la désignation d'un « objet » sémantique par une expression linguistique – que cette expression soit un nom, un SN, une proposition, une phrase, etc. – contextuellement située, l'indexicalité pourrait se gloser par le besoin

[58] Nous verrons dans le chapitre VII les raisons pour lesquelles nous considérons *it* dans notre étude.

qu'éprouve une expression langagière de pointer vers un élément quelconque afin de pouvoir référer. Prenons un exemple tout simple :

(8) *Le petit chat de la voisine est intenable ! Il s'est encore enfui ce matin*.

Le petit chat de la voisine est une expression référentielle, tandis que *il* est un élément indexical : dans ce cas-ci, nous ne pourrions pas déduire que *il* renvoie au « petit chat de la voisine » si celui-ci n'avait pas été mentionné auparavant. De manière générale, les pronoms sont les éléments indexicaux prototypiques. Néanmoins, il ne faudrait pas se fourvoyer en prétendant qu'une expression à tête lexicale est forcément référentielle, sans être indexicale :

(9) *Prendre quatre pommes et une dizaine de fraises. Peler les pommes, couper les fraises et mélanger le tout dans un bol*.

(10) *Il était une fois un homme qui vivait avec ses trois fils. Cet homme était bon et aimé de tous*.

Dans l'exemple (9), *les pommes* et *les fraises* sont interprétables grâce à la mention antérieure de « quatre pommes et une dizaine de fraises ». Il en va de même pour la référence effectuée par *cet homme* en (10). Ainsi, nous voyons bien qu'un syntagme à tête lexicale peut dépendre d'un élément[59] pour pouvoir référer. La distinction entre 'référence' et 'indexicalité' n'est pas toujours très claire si l'on parcourt la littérature, certains linguistes prennent le parti d'opposer ces deux notions (cf. Chastain, 1975 ; Corblin, 2005). Nous choisissons d'adopter le point de vue de l'indexicalité comme étant un mode d'accès à la référence particulier, puisqu'engageant un processus de pointage, mais nous n'adhérons pas à la thèse qui exclut les éléments indexicaux des expressions référentielles[60]. Deux points concernant la référence et l'indexicalité : d'une part, ces deux processus sont effectuables grâce à un ancrage contextuel ; d'autre part, l'indexicalité dépend d'un élément quelconque puisque celle-ci

[59] Nous employons le terme 'élément' ici de façon tout à fait volontaire, afin de ne pas présager de la nature linguistique, sémantique ou autre de cet élément ; nous discuterons ce point en section 2.5 *infra* et lors des chapitres VII et VIII.

[60] Kaplan (1989) va jusqu'à classer les indexicaux parmi les expressions directement référentielles, au même titre que les noms propres. Selon lui, les indexicaux dénotent une propriété – par exemple, *je* dénote « la propriété d'être le locuteur » – qui permet l'identification référentielle en fonction de la situation d'énonciation. C'est l'appariement entre l'expression indexicale et l'« objet » visé qui justifie, selon Kaplan, cette classification.

appelle un pointage. Le questionnement directement déductible de cette assertion concerne bien entendu la spécification du contexte ainsi que la nature de cet élément quelconque.

Afin de pouvoir répondre aux questions ci-dessus, il convient de poser clairement les jalons de notre modèle sémantique, à partir des discussions que nous avons tenues jusqu'alors. La conception de la réalité que nous choisissons d'adopter est celle d'un réel intersubjectif, dont l'objectivité se trouve dans la norme collective. Comme nous l'avons déjà souligné, une telle vision de la réalité permet d'établir une primauté du réel par rapport aux autres mondes possibles. Ainsi, nous percevons des « choses », des « concepts », des « propriétés », qui forment notre environnement physico-culturel. En tant qu'individus parlants, nous disposons d'un matériel langagier pour renvoyer à ce monde ; il est important de noter ici que le matériel langagier n'est pas seulement constitué de la langue, mais également de l'ensemble de signes corollaires que nous pouvons employer, de manière consciente ou non : signes de la tête, du doigt, du regard, etc. Nous nommons ces signes verbaux et non verbaux le 'texte'. En tant qu'individus psychologiques, nous pouvons envisager des mondes alternatifs à la réalité normative que nous connaissons, nous pouvons entrevoir des espaces imaginaires secondaires ; qui plus est, nous pouvons les mêler à notre réalité : l'Intégration Conceptuelle – 'Conceptual Blending'* – est la théorie de la cognition développée par Fauconnier et Turner (Fauconnier & Turner, 2002), à la suite des Espaces Mentaux de Fauconnier (1984). Les seuls termes 'espaces imaginaires' mènent à une évidence, qu'il convient toutefois de souligner : les « choses », « concepts », « propriétés » appartenant à ces espaces sont cognitifs, des représentations mentales non palpables et uniquement situées dans l'esprit de l'individu. Pour l'instant, nous pourrions donc penser que nous avons les éléments physiques et culturels de notre réalité tangible d'un côté, les entités mentales des espaces secondaires de l'autre ; par exemple, un locuteur A ordonnant *passe-moi le beurre !* à un individu B serait utile pour ce que ce dernier comprenne que la requête de A concerne « le beurre qui est sur la table », ou à la rigueur, « le beurre qui est dans le frigidaire », et non pas un objet fantasque de couleur improbable que l'on pourrait ingérer sans prendre un gramme parce que, dans ce monde onirique, le beurre serait sans matière grasse. A l'inverse, la lecture du *Cycle des Dieux* de Werber (2004-2007) conduit le lecteur à conceptualiser l'île d'Aeden où se déroule l'éducation des élèves-dieux par les Dieux de la mythologie grecque en personne. Cependant, si nous revenons à l'usage commun du texte et si nous inscrivons la référence dans un cadre plus large que le simple énoncé, produit par une phrase canonique contextuellement interprétée – nous pensons par exemple à une conversation en face à face, téléphonique, mais

aussi à la lecture, à l'écoute d'une conférence, etc. –, cette dimension cognitive devient indispensable à l'interprétation du message ; en effet, nous ne pouvons comprendre ce qui est dit/lu/communiqué à l'instant t_0 qu'en faisant le lien avec ce qui a été dit/lu/communiqué auparavant. Certes, la lecture permet d'effectuer des retours en arrière, puisqu'il reste une trace écrite du texte ; en revanche, la communication conversationnelle et l'écoute d'un monologue n'admettent pas ce mouvement rétrogressif : « les paroles s'envolent, les écrits restent ». Par conséquent, la trace de ce qui a été dit/communiqué – nous pouvons même ajouter 'lu', car il est peu probable d'effectuer des retours en arrière incessants lors de la lecture – réside dans l'esprit de l'individu, est cognitive. Ces traces sont une représentation mentale d'« objets » évoqués, de « concepts », auxquels ont été attribuées des « propriétés », en somme des « objets » dont on a parlé. Ces traces sont en perpétuel développement lors de l'acte communicatif, que ce dernier soit interactif, comme dans une conversation, ou monologique, comme lors d'une conférence. Nous nommons ces traces le 'discours'[61]. Ainsi, nous pouvons effectuer une corrélation avec la définition des 'référents' proposée par le paradigme constructiviste (cf. ce chapitre, section 1.2), encore appelés 'objets-de-discours' ou 'objets du discours', comme étant des entités mentales auxquelles renvoient les expressions linguistiques, sans lien avec la réalité stable ; néanmoins, de par notre conception intersubjective de la réalité, nous choisissons de ne pas exclure le réel des paramètres permettant l'accès à la référence. La problématique réside dans la manière d'intégrer cette composante.

Le paramètre que nous avons évoqué à maintes reprises dans les processus que sont la référence et l'indexicalité est le contexte. Terme largement usité dans la littérature, il n'en reste pas moins défini de manière inadéquate ou insuffisante, selon nous : pour certains, le 'contexte' est synonyme de 'co-texte' (chez Culioli par exemple, cf. chapitre II, section 3.3.2). Selon notre définition, le contexte est une composante complexe, intégrant l'extralinguistique, la situation d'énonciation ainsi que d'autres éléments, à l'exclusion du texte. Afin d'affiner la description du contexte et de tendre vers une définition plus détaillée de cette structure, nous nous basons sur l'étude de Connolly (2007).

[61] Cette définition du discours n'est pas la plus répandue ; en effet, 'discours' est très souvent synonyme de 'texte' – 'texte' selon notre définition du terme – comme chez Lapaire & Rotgé (1992), par exemple (cf. chapitre II, section 3.2.2). Néanmoins, cette définition du 'discours' se retrouve dans la 'mémoire discursive' de Berrendonner (1990).

2.3. Le ‘contexte’

Selon Connolly, dont nous partageons le point de vue, le contexte est un élément indispensable à toute communication. Par conséquent, il est nécessaire de l’intégrer à tout modèle représentatif et analytique, au même titre que les autres niveaux d’étude linguistiques – phonologie, morphologie, syntaxe, sémantique –, et de ne pas le laisser ‘au-dehors’ – n’oublions pas que, jusqu’aux travaux de Grice (1957), la pragmatique ne faisait pas partie de l’étude du langage. Afin de définir ce qu’est le contexte, il suffit de partir du terme lui-même : « contexte = con + texte = qui accompagne le texte ». Ainsi, le contexte est tout ce qui est pertinent à la production et à l’interprétation de ce qui est dit/lu/communiqué à l’instant t_0. Nous envisageons le contexte comme un élément essentiellement dynamique. Prenons quelques exemples pour illustrer cette conception. Imaginons la scène suivante : nous sommes dans notre chambre, plongée dans la lecture d’un ouvrage sur la philosophie du langage, lorsque notre voisin se met à augmenter le niveau de sa chaîne hi-fi, au point que le rythme régulier et assourdissant des basses devient insupportable. Nous décidons alors de frapper à sa porte et lui disons d’un ton agacé :

(11) *Tu peux baisser le son de ta chaîne s’il-te-plaît ? Y en a qui travaillent !*

Le ton de notre énoncé, c’est-à-dire l’acte perlocutoire que nous avons produit, est déterminé par l’augmentation de la musique, qui a eu pour effet de nous énerver. Nous n’aurions pas produit un tel énoncé si l’élément contextuel qu’est « le son des basses parvenant jusqu’à notre chambre » n’avait pas eu lieu. De la même manière, notre production langagière subit un net changement si nous discutons posément lors d’un petit-déjeuner en famille et que notre directeur nous appelle afin de débattre de nos récentes avancées de thèse. Néanmoins, il est des aspects du contexte qui ne varient pas : dans notre dernier exemple, le lieu et l’ancrage temporel restent identiques. Si nous revenons à la définition du contexte – « qui accompagne le texte » –, il serait incohérent dans notre modèle d’inclure le texte, à savoir l’ensemble de signes verbaux et non verbaux dont nous disposons pour communiquer. Comme nous allons le remarquer, l’analyse de Connolly (2007) inclut ces signes dans son modèle du contexte. Voici la structure en quatre dichotomies générales proposée par Connolly :

(i) Contexte discursif *vs* contexte situationnel

(ii) Contexte physique *vs* contexte socioculturel

(iii) Contexte étroit *vs* contexte large

(iv) Contexte mental *vs* contexte extra-mental

L'opposition (i) sépare ce qui fait partie du domaine langagier de ce qui fait partie du monde extérieur. Ainsi, chez Connolly, le contexte discursif est l'ensemble de signes verbaux et non verbaux que nous appelons 'texte'. Le point (ii) est une subdivision du contexte situationnel, celui-ci étant composé d'une part des « choses » qui nous entourent, des « propriétés » que nous attribuons à ces « choses », inscrites à la fois dans le temps et dans l'espace – les entités d'ordre zéro, un et deux –, et d'autre part des codes relevant des domaines social, de pensée et d'attitude auxquels nous sommes soumis. L'opposition (iii) s'applique aux contextes discursif et situationnel, chacun étant divisé en contextes étroit et large ; ainsi, le contexte discursif étroit s'assimile au co-texte – au sein du même texte –, le contexte discursif large à l'intertexte – au sein de textes différents ; de même, le contexte situationnel étroit concerne la situation d'énonciation immédiate – ce que l'approche modulaire de Roulet (1991) nomme les 'scénarios' –, tandis que le contexte situationnel large englobe les implications que peut entraîner un scénario : par exemple, si nous achetons des billets de train au guichet SNCF, la gare et le guichetier appartiennent au contexte situationnel étroit alors que notre destination fait partie du contexte situationnel large. La dernière dichotomie (iv) différencie ce qui appartient à l'esprit des individus parlants de ce qui en est externe ; ainsi, chez Connolly, le contexte mental inclut le savoir partagé entre les participants à un acte de communication, l'état psychologique du locuteur et celui de l'interlocuteur.

Comme nous l'avons mentionné, Connolly inclut le texte dans son modèle, puisque le contexte discursif est comparable à ce qui pourrait être notre contexte textuel. Etant donné que nous choisissons d'exclure le texte des composantes contextuelles, nous gardons la terminologie de Connolly mais modifions le contenu : le contexte discursif ne concerne plus le matériel langagier, linguistique et gestuel, mais ces traces dynamiques que nous nommons 'discours'. Le contexte situationnel reste identique à celui décrit ci-dessus, comme étant composé de notre monde intersubjectif et des codes socioculturels auxquels nous sommes attachés. En revanche, et ce afin d'entrevoir un modèle interactionnel, nous ajoutons le calcul qu'effectue le locuteur au sujet de l'état psychologique de son allocutaire à la composante 'contexte mental'. Pourquoi séparer 'texte' et 'discours' ? Pourquoi ne pas intégrer le texte

dans la description du contexte ? Pour différencier système langagier et usage et ne pas mélanger les niveaux d'analyse. Selon nous, le discours est la production conjointe du texte et du contexte ; cette représentation cognitive est en perpétuel développement lors de l'acte communicatif. Puisque ce discours influence à son tour la production langagière à l'instant t_0 – gardons en mémoire que nous ne pouvons comprendre[62] ce qui est dit/lu/communiqué à l'instant t_0 que par le souvenir et la compréhension de ce qui a été dit/lu/communiqué auparavant –, nous pouvons admettre que le discours fait partie intégrante du contexte :

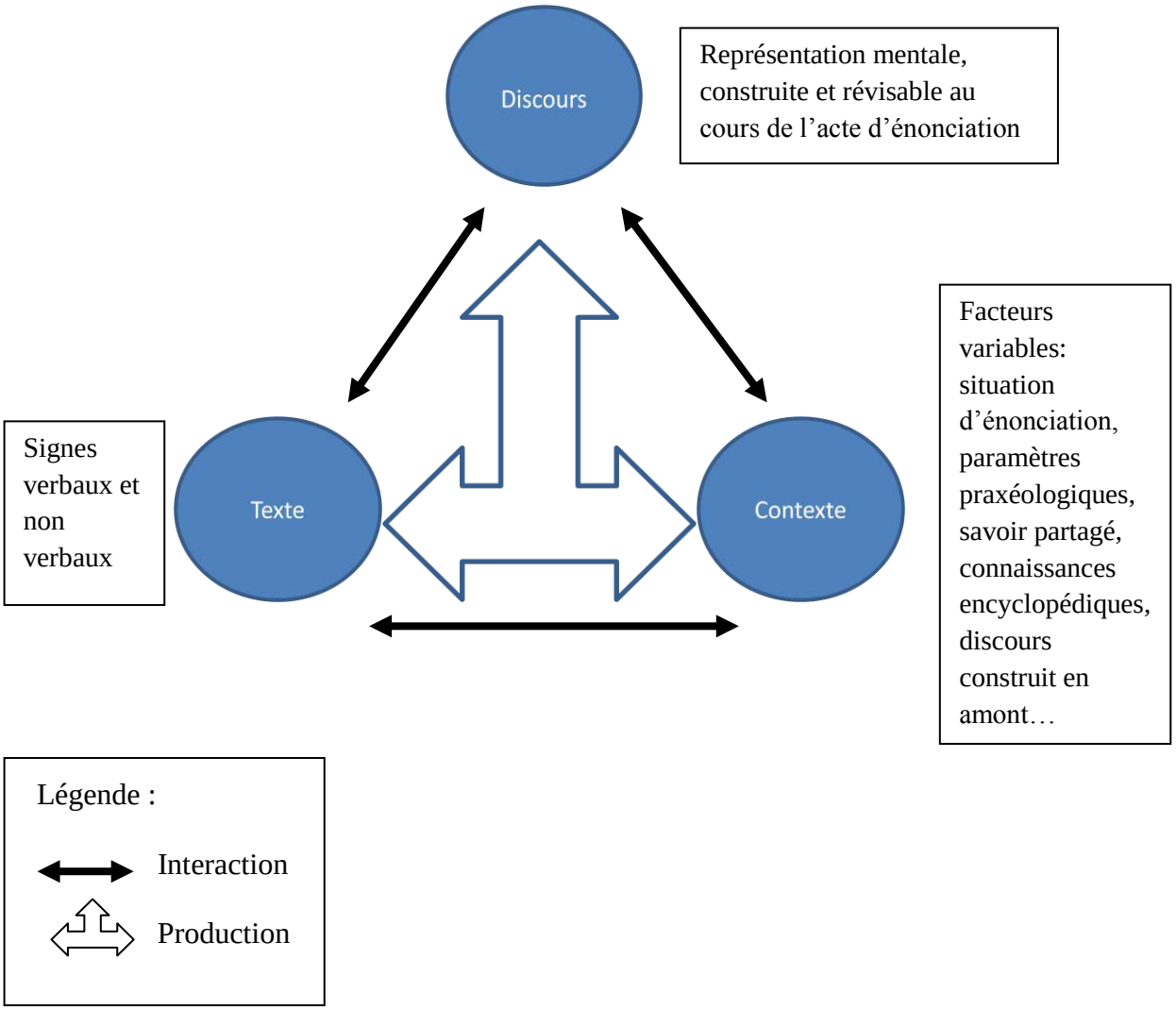


FIGURE 6 : 'TEXTE', 'CONTEXTE' ET 'DISCOURS' : UN SYSTEME INTERDEPENDANT

[62] Notons ici que nous pouvons comprendre, mais aussi intervenir de manière adéquate. Nous développerons ces principes de cohérence ultérieurement.

Contexte					
Situationnel		Savoir		Discursif	
Etroit	Large	Etroit	Large	Etroit	Large
Environnement physico-social immédiat (Quelles limites ?)	Environnement physico-social étendu (Quelles limites ?)	Savoir partagé entre les participants à l'acte communicatif + état psychologique de chaque participant + calcul de l'un et l'autre des participants sur l'état psychologique de l'un et l'autre des participants	Connaissances du monde	Discours construit en amont de la production langagière en t_0 (Quelles limites ? L'énoncé ?)	Discours construit en amont (Quelles limites ? Au-delà du simple énoncé ?)

TABLEAU 3 : LA STRUCTURE DU CONTEXTE (D'APRES CONNOLLY, 2007)

A la lecture du tableau ci-dessus, nous remarquons que certaines zones de cette composante complexe qu'est le contexte restent floues, notamment les frontières qui pourraient délimiter contextes étroits et larges. Nous apporterons des éléments de réponse en observant les données de notre corpus.

Le contexte tel qu'il est défini par Connolly (2007) s'inscrit dans un modèle néo-fonctionnaliste, la FDG (Functional Discourse Grammar*), dont les préceptes théoriques dérivent partiellement du courant structuraliste et des conceptions fonctionnelles de Saussure. Les linguistiques fonctionnelles considèrent le langage comme un instrument dont disposent les individus afin de communiquer à bon escient. La section 2.4 présente deux modèles-phares de la linguistique fonctionnelle : la FDG et la SFG (Systemic Functional Grammar*).

2.4. Quelques théories discursivo-fonctionnelles

Les notions de 'texte', de 'contexte' et de 'discours' telles que nous les définissons pourraient s'inscrire dans le cadre général fourni par les grammaires fonctionnelles comme la FDG (Functional Discourse Grammar) de Hengeveld & Mackenzie (2008), elle-même issue

de la FG (Functional Grammar*) de Dik (1997), ainsi que de la SFG (Systemic Functional Grammar) de Halliday (1985, théorie développée dans les années soixante). Le point commun à ces cadres théoriques est la conception de la langue et du langage comme un réseau de systèmes qui interagissent pour faire sens – idée initialement développée par Firth dans les années cinquante. Ainsi, la langue est composée d'ensembles finis, comme la syntaxe, et non finis, comme le lexique, et n'est pas porteuse de sens ; elle en est le potentiel – « meaning potential » (Halliday, 1978 : 39). Les grammaires fonctionnelles ont en commun la structure descendante de leurs modèles, au contraire de formalisations ascendantes comme celle proposée par la GG, afin de refléter le caractère multidimensionnel de l'expérience humaine et les relations interpersonnelles qu'engage un acte communicatif – volonté partagée par les approches connexionnistes et modulaires. Sans vouloir adopter l'un ou l'autre de ces cadres pour notre analyse à venir, nous nous accordons néanmoins sur la fonction communicative au cœur de tout modèle représentatif du langage et présentons, de manière succincte, les deux théories-phares que sont la FDG et la SFG.

2.4.1. La FDG (Functional Discourse Grammar) de Hengeveld & Mackenzie (2008)

La FDG est une expansion de la FG (Functional Grammar) développée par Dik en 1997. Si la FG est avant tout un modèle d'analyse des énoncés produits par des phrases, la FDG désire rendre compte de phénomènes ayant trait à des unités plus larges que la phrase, mais aussi de ceux concernant des unités plus petites. En effet, la phrase est une division normative dont le représentant canonique – unité formée autour d'un verbe, commençant par une majuscule et se terminant par un point – se trouve principalement dans les productions relevant du registre de l'écrit. Dès lors que nous passons à des données authentiques, notamment en ce qui concerne l'oral spontané, il est très difficile de s'accorder sur ce qui relève du concept de 'phrase' ; nous discuterons ce point dans le chapitre VI. Ainsi, la FDG ne choisit pas la phrase comme unité minimale de production d'énoncé – 'acte discursif' dans leur terminologie. Le modèle d'analyse de la FDG est une structure descendante et modulaire : les configurations choisies en niveaux supérieurs déterminent celles en niveaux inférieurs, mais chaque niveau est indispensable au fonctionnement général de l'acte communicatif. La FDG décrit quatre composantes majeures :

- La composante conceptuelle concerne les intentions de communication des participants à l'acte langagier ainsi que les représentations cognitives qui leur

correspondent. C'est en fonction de cette composante qu'est décidée la configuration de la composante grammaticale ;

- La composante grammaticale porte sur la formulation langagière déterminée par les intentions de communication. Cette composante grammaticale constitue le cœur de l'architecture de la FDG et se divise en quatre niveaux d'analyse :
    - Le niveau interpersonnel, que l'on pourrait assimiler à la pragmatique ; c'est à ce niveau que sont pris en compte les participants à l'acte communicatif et les actes perlocutoires ;
    - Le niveau représentationnel concerne la sémantique ; c'est à ce niveau que sont co-indexés les expressions référentielles et les anaphoriques correspondants, par exemple, mais aussi la catégorisation des entités selon leur ordre ontologique, la marque distale ou proximale des démonstratifs si l'on décrit l'anglais ;
    - Le niveau morphosyntaxique analyse la structure syntaxique et morphologique de la séquence considérée ;
    - Le niveau phonologique décrit les unités phonologiques et le contour intonatif de l'énoncé.
- La composante contextuelle est principalement consacrée au stockage des informations accessibles aux participants à tout point de l'acte communicatif – ce que nous appelons le 'discours'. Néanmoins, la définition du contexte par Connolly au sein de ce modèle affine la description de la composante contextuelle (cf. ce chapitre, section 2.3) ;
- La composante de sortie ('output') ressortit aux signes phoniques, graphiques et sémiotiques employés par les participants à l'acte langagier.

L'architecture générale de la FDG se représente alors de la manière suivante :

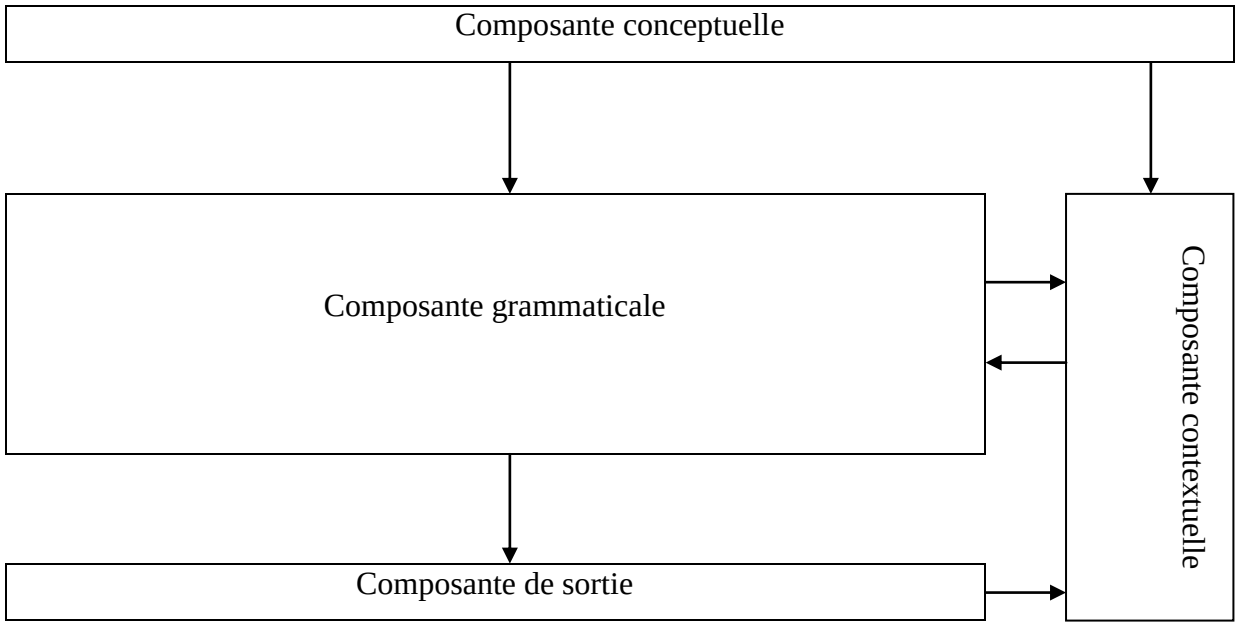


FIGURE 7 : L'ARCHITECTURE GENERALE DE LA FDG (HENGEVELD & MACKENZIE, 2008 : 13)

Nous pouvons constater que ce modèle décrit une macrostructure principale, la composante grammaticale, bordée par les composantes subalternes, mais néanmoins indispensables, que sont les composantes conceptuelle, contextuelle et de sortie. Le problème, selon nous, d'une telle formalisation de la communication est que les mécanismes régissant la structure systémique de la langue appartiennent à la même composante que les aspects qui relèvent de l'usage de cette même langue ; or, il semble indispensable que système et usage soient clairement séparés. Sans cette distinction fondamentale, il serait impossible de rendre compte de la différence entre 'dénotation' et 'référence', par exemple (cf. ce chapitre, section 2.1). Il en va de même pour l'interprétation de cet avertissement, visible sur les autoroutes de France :

(12) *Attention aux hommes en jaune !*

Ici, *les hommes en jaune* désigne « les employés de Vinci Autoroutes® dont la tenue de travail est jaune », et non pas « certains coureurs cyclistes vainqueurs du Tour de France » par exemple. Comment le destinataire de ce message peut-il saisir la référence des *hommes en jaune* si ce n'est par la contextualisation de l'énoncé ? *Les hommes en jaune* ne renvoie pas de manière intrinsèque à des « employés de Vinci Autoroute® » ; l'interprétation de ce SN est rendue possible par l'actualisation de l'expression linguistique en contexte. Par conséquent, il

est de notre avis qu'un modèle descendant et modulaire convient à la représentation de la communication, à la condition que système et usage soient séparés.

2.4.2. La SFG (Systemic Functional Grammar) de Halliday (1985)

La SFG concerne principalement les choix de langue dont disposent les individus parlants afin de communiquer à bon escient. La SFG comporte trois niveaux d'analyse, à savoir la sémantique, la phonologie et la morphosyntaxe – 'lexicogrammar' selon la terminologie de Halliday – régies par des bases fonctionnelles appelées 'métafonctions' : les métafonctions idéationnelle, interpersonnelle et textuelle :

- La métafonction idéationnelle est en relation avec le domaine d'application du texte – ce que Halliday nomme 'field', parfois traduit par 'champ' – et que l'on pourrait décrire comme le genre du texte considéré, le thème majeur discuté, le scénario (cf. ce chapitre, section 2.3). Ce domaine d'application détermine les choix lexicaux et l'agencement des différentes catégories morphosyntaxiques ;
- La métafonction interpersonnelle concerne les relations entre les divers participants à l'acte communicatif ainsi que les conséquences que celles-ci pourraient avoir sur les choix perlocutoires et lexicaux ;
- La métafonction textuelle prend en compte la cohérence* et la cohésion* internes d'un texte – relations anaphoriques, enchaînements, ellipses, etc.

Comme en FDG (cf. ce chapitre, section 2.4.1), la SFG ne distingue pas la langue en tant que système et la langue actualisée en discours. Cependant, hormis la description de son modèle fonctionnel et systémique, Halliday développe le rapport qui existe entre référence et cohérence discursive, notamment dans son ouvrage de 1976, co-écrit avec Hasan (Halliday & Hasan, 1976). Dans leurs termes, la cohésion est une notion sémantique qui se définit par les liens grammaticaux et lexicaux qui unissent un ensemble de segments textuels. Ainsi, le principe de cohésion fait qu'une suite d'énoncés forme un tout intelligible que le destinataire peut percevoir comme un message cohérent. De cette définition ressort d'ores et déjà une assimilation entre 'texte' et 'discours' d'une part et entre 'système' et 'usage' d'autre part, les liens grammaticaux et lexicaux ainsi que les segments de texte relevant du domaine du système langagier et la cohésion étant un concept sémantique. La cohésion est sous-divisée en cinq procédés : la référence, la substitution, l'ellipse, la conjonction et la cohésion lexicale. Les deux catégories qui nous intéressent tout particulièrement sont la référence et la

substitution. Chez Halliday, la référence s'applique à des termes spécifiques qui ne peuvent être sémantiquement interprétés par eux-mêmes. Deux conjonctures sont alors envisagées : soit l'item en question dépend de l'extralinguistique pour référer, auquel cas nous avons affaire à de l''exophore'[63], soit cet item dépend d'un autre élément textuel pour être interprété, auquel cas nous avons affaire à de l''endophore' :

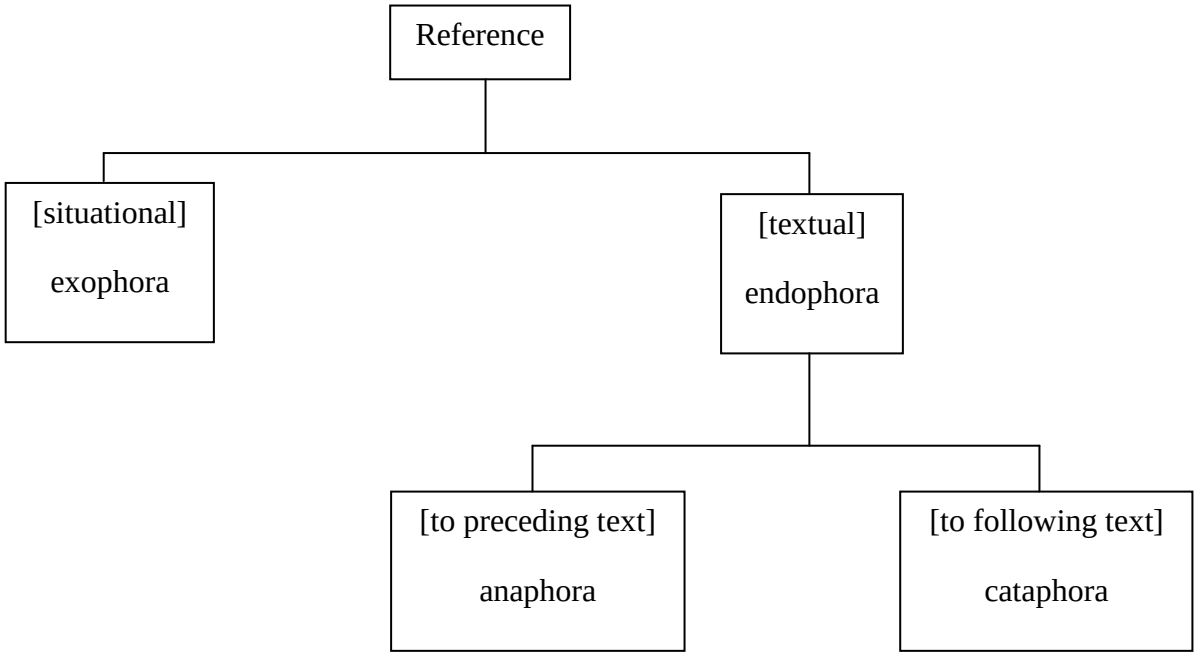


FIGURE 8 : LES TYPES DE REFERENCE EN SFG (HALLIDAY & HASAN, 1976 : 33)

Halliday et Hasan précisent qu'il existe trois types de référence, qui pourraient *a priori* se combiner avec une référence de type exophorique ou endophorique : les références personnelle, démonstrative et comparative, correspondant de manière respective aux catégories morphosyntaxiques que sont les pronoms personnels et possessifs, les démonstratifs proximaux – *this*, *these*, *here* – et distaux – *that*, *those*, *there* – et les comparatifs. La substitution, quant à elle, est définie comme une relation linguistique – au niveau morphosyntaxique, 'lexicogrammar' chez Halliday – grâce à laquelle un terme est remplacé par un autre : en anglais, *one*, *ones* et *same* sont des termes substitutifs, contrairement aux autres pronoms, personnels et démonstratifs, qui relèvent du principe

[63] La définition de l''exophore' est similaire à celle de 'deixis' (cf. chapitre II, section 1), cette dernière étant caractérisée dans certaines théories par la référence via l'ancrage dans la réalité spatio-temporelle. Chez Halliday, cependant, le terme déictique renvoie au mode de détermination effectué par les déterminants, à valeur spécifique ('specific deictics') ou non spécifique ('non-specific deictics').

cohésif de référence. Notre intuition semble se confirmer : il n'y a pas de distinction entre 'texte' et 'discours' en FSG – nommés respectivement 'surface text' et 'world text' – ni de séparation entre système langagier et usage de la langue. Preuve en est la correspondance directe entre catégories morphosyntaxiques et mode de référence ainsi que la caractérisation d'un concept sémantique – la 'cohésion' – par des relations entre items linguistiques[64]. Un élément langagier est dit référer à une portion de la réalité ou à un segment de texte, mais les paramètres qui rendent possible cette référence – notamment l'intervention du contexte – ne sont pas pris en compte. La conséquence immédiate de ces confusions est l'opposition entre 'exophore' et 'endophore' et, par là même, entre réel et fonctionnement langagier. Les répercussions de cette conception ne sont pas des moindres ; prenons pour exemple le pronom personnel clitique *il*, en français. Supposons que nous soyons dans une chambre d'hôpital ; notre ami est au chevet de son père, endormi et visiblement inconscient ; nous nous adressons alors à cet ami, à propos de son père :

(13) *Tu penses qu'il peut m'entendre ?*

Dans le cas présent, *il* serait donc à référence exophorique, puisque ce pronom dépend de la situation extralinguistique pour référer. D'un autre côté, si nous discutons entre amis et que l'un d'eux prononce l'énoncé produit par l'exemple suivant :

(14) *J'ai parlé à Jean, il m'a dit qu'il ne viendrait pas ce soir.*

Ici, *il* serait à référence endophorique, puisque ce pronom fait appel à *Jean* pour référer. Devrions-nous déduire qu'il existe deux *il* ? Un *il* à référence exophorique et un *il* à référence endophorique ? En tout cas, une telle conclusion découle directement du modèle prôné par la FSG. Comme nous l'avons déjà soulevé à plusieurs reprises, il est crucial de distinguer langue en tant que système et langue en usage, de ne pas confondre 'texte' et 'discours' et de séparer les niveaux de représentation dans tout modèle de la communication langagière.

[64] Il est évident que la cohérence discursive est assurée par l'assemblage cohésif d'items linguistiques, mais il convient de séparer les niveaux en définissant quels paramètres agissent afin que ces items, actualisés en discours, puissent rendre un discours cohérent.

2.5. L'indexicalité : la deixis et l'anaphore

Dans la section précédente, nous avons exprimé notre désaccord envers l'opposition entre exophore et endophore, qui rejoint celle entre deixis et anaphore. Il est une autre vision de ces deux notions, déjà entrevue dans le chapitre II, section 1, qui considère l'anaphore comme dérivée de la deixis. Ce point de vue est défendu par de nombreux linguistes – entre autres Bühler, Cornish, Kleiber, Lyons et de Mulder – et soutenu notamment par des arguments d'ordre diachronique : les démonstratifs, essentiellement déictiques, sont à l'origine des définis et pronoms personnels plus à même de réaliser l'anaphore. De plus, cette conception scalaire allant de la deixis à l'anaphore s'accorde avec notre position sur la réalité intersubjective : la deixis canonique étant centrée sur un ancrage spatio-temporel dans l'instant t_0, elle permet ainsi de rendre compte de la primauté du réel par rapport aux autres mondes possibles. Dès lors, nous pouvons situer les formes de deixis secondaire : la deixis *am phantasma** (Bühler, 1934[2009]) qui déplace le centre déictique dans un monde imaginaire, la deixis textuelle* (Lyons, 1977 : 667) où l'usage d'un pronom est à référence mentionnelle, ou encore ce que Lyons nomme 'impure textual deixis'* (1977 : 668) et définit comme la référence d'un pronom à une entité de troisième ordre précédemment évoquée au sein du discours. S'ensuit une définition résolument cognitive de la deixis et de l'anaphore, allant jusqu'à négliger le rapport entre deixis et réalité. Ce point de vue est notamment défendu par Cornish (1996, 1999, 2007, 2010), qui définit la deixis comme l'insertion, par un locuteur/scripteur donné, d'un référent nouveau pour l'allocutaire et l'anaphore comme la continuation référentielle d'une entité considérée comme connue et centrale dans l'esprit de l'interlocuteur/lecteur. Le modèle proposé par Cornish distingue les expressions langagières qui relèvent du 'texte' et les entités désignées qui, elles, appartiennent au 'discours' ; par conséquent, Cornish réfute la relation traditionnelle entre un 'anaphorique' et un 'antécédent' textuel en proposant les termes de 'déclencheur d'antécédent'* – *i.e.* le signe textuel, linguistique ou sémiotique –, d''antécédent' – le référent premièrement évoqué – et d''anaphorique' – l'expression dont la référence dépend de cet antécédent. Entre la deixis et l'anaphore est alors définie une série de « procédures discursives » (Cornish, 1996, 1999, 2007, 2010) alliant deixis et anaphore, ce que Cornish appelle, à la suite de Ehlich (1982), l''anadeixis'*. Ainsi, la deixis et l'anaphore sont conçues comme les pôles extrêmes d'une échelle, chaque procédure discursive étant réalisée en discours par un marqueur langagier prototypique :

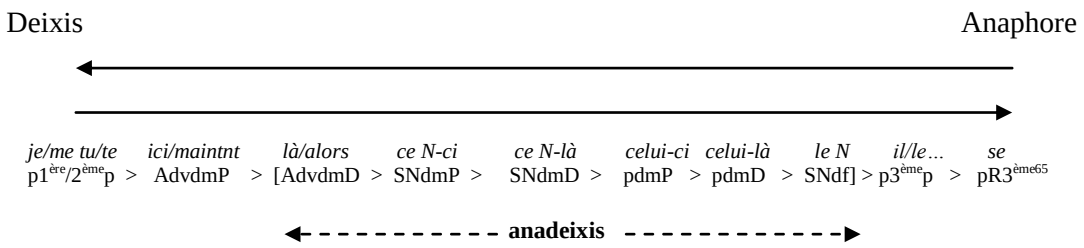


FIGURE 9 : L'ECHELLE DE PHORICITE CODEE PAR CERTAINES CATEGORIES D'EXPRESSIONS INDEXICALES (CORNISH, 2007 : 149)

Cette échelle fait écho aux théories sur l'accessibilité des référents de Ariel ('Accessibility Scale'*, 1988) et Gundel, Hedberg & Zacharski ('Givenness Hierarchy'*, 1993), qui se proposent de hiérarchiser certaines expressions indexicales en fonction du statut cognitif de leur référent. L'échelle d'Ariel (1988 : 84) classe ces expressions langagières selon trois facteurs : la représentation mentale plus ou moins saillante du référent dans l'esprit des interlocuteurs, le nombre de référents potentiels pour l'expression considérée et la distance relative entre l'expression anaphorique et son 'antécédent' textuel (cf. chapitre II, section 1) :

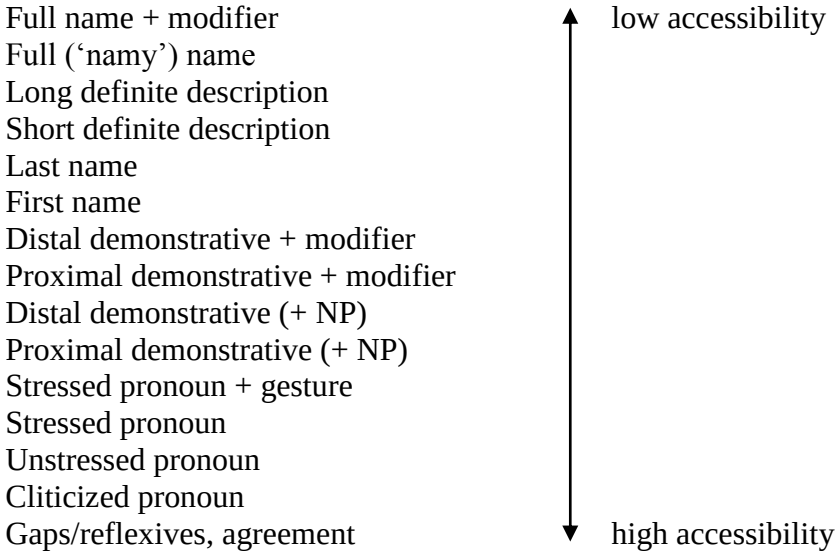


FIGURE 10 : L''ECHELLE D'ACCESSIBILITE DES REFERENTS' (ACCESSIBILITY SCALE) D'ARIEL (1988 : 84)

Ainsi, les pronoms clitiques, par exemple, seraient plus à même de référer à une entité discursive saillante au moment de leur emploi, tandis qu'un nom propre serait plus propice à

[65] Abréviations employées dans la Figure 9 : 'p1ère/2ème/3èmep' : 'pronom de première/deuxième/troisième personne' ; 'P' : 'proximal' ; 'D' : 'distal' ; 'dm' : 'démonstratif' ; 'adv' : 'adverbe' ; 'SN' : 'syntagme nominal' ; 'p': 'pronom' ; 'df': 'défini' ; 'R': 'réfléchi'.

une première désignation. Néanmoins, il est remarqué que la distance n'est pas un facteur valable pour la description référentielle des démonstratifs – cf. la distance entre *un homme qui vivait avec ses trois fils* et *cet homme* dans ce chapitre, exemple (10) – et nous avons évoqué des cas d'expressions anaphoriques sans 'antécédent' textuel – cf. ce chapitre, exemple (13). L'échelle de Gundel *et al*. (1993 : 275), quant à elle, se base uniquement sur le statut cognitif des référents, qui sont hiérarchisés et imbriqués du plus au moins saillant :

in focus	> activated	> familiar	> uniquely identifiable	> referential	> type identifiable	
{*it*}	{*this*, *that*, *this* N}	{*that* N}	{*the* N}	{indefinite *this* N}	{*a* N}	

FIGURE 11 : L''ECHELLE DES STATUTS COGNITIFS DES REFERENTS, CODES PAR CERTAINES EXPRESSIONS LINGUISTIQUES' (GIVENNESS HIERARCHY) DE GUNDEL *ET AL*. (1993 : 275)

Ainsi, le clitique *it* serait typiquement anaphorique puisque continuant la référence d'une entité saillante, les pronoms démonstratifs ainsi que le SN démonstratif *this N* référeraient à une entité activée – mais moins saillante que celle visée par *it* – le SN démonstratif *that N* coderait un référent connu des interlocuteurs, un SN défini serait prototypique de la référence spécifique, contrairement au SN démonstratif *this N* présentatif et au SN indéfini, ce dernier introduisant un référent nouveau et donc étant prototypiquement déictique. Plusieurs inadéquations ressortent de cette échelle : d'une part, elle présuppose une correspondance directe entre expression référentielle et statut cognitif du référent visé, problème qui se trouverait résolu par l'insertion des paramètres rendant possible cette alliance au sein du modèle ; d'autre part, la double occurrence de *this N* dans cette représentation implique l'existence de deux SN démonstratifs *this N*, l'un par défaut, l'autre présentatif – 'indefinite *this N*'. Cette dernière remarque est un argument supplémentaire à la nécessité d'une séparation entre système et usage : selon nous, il existe une seule expression *this N*, qui peut être à référence spécifique ou introductive, dirons-nous, selon son emploi en contexte. Néanmoins, les échelles de Cornish, Ariel et Gundel *et al*. permettent une mise en relation entre expressions indexicales, deixis, anaphore et cognition, ce qui laisse présager d'une étude possible de la fonction structurale de ces expressions au sein du discours. Nous pouvons toutefois regretter que le facteur « réalité » ait disparu de ces représentations.

Il est de notre avis que l'ensemble des préceptes décrits dans ce chapitre présentent des avantages : il ne s'agit pas ici de défaire des décennies de recherche sur les démonstratifs, la

deixis et l'anaphore et de proposer une théorie révolutionnaire et universelle. Notre désir n'est pas non plus de réinventer la roue en proposant une énième analyse des démonstratifs suivant l'un ou l'autre des modèles proposés. Tout d'abord, cela n'aurait aucun intérêt. Ensuite, il s'agit de trouver un compromis permettant la prise en compte du réel, l'intégration du contexte dans le fonctionnement langagier, tout en s'accordant sur le fait que le produit d'un acte communicatif est avant tout cognitif.

3. Un candidat possible au chaînon manquant entre réalité et cognition : la mémoire

Avant d'exposer notre théorie sur la référence et l'indexicalité, dressons un bilan à ce point de la réflexion. Tout d'abord, nous considérons notre environnement physico-social comme réel puisque reconnu comme tel par l'ensemble des individus. Le langage est l'un des moyens dont dispose l'être psychologique, parlant et écoutant afin de renvoyer à ce monde, assimilable à une norme composée d'« objets », de « concepts » et de « propriétés ». Dès lors que nous sommes confrontés à la lecture ou engagés dans une conversation, ces « objets », « concepts » et « propriétés » ne sont pas physiquement situés de manière immuable, mais gagnent un statut cognitif, ce sont des représentations mentales localisées dans notre esprit qui évoluent au fur et à mesure que se déroule l'acte communicatif : ces représentations mentales sont des référents et elles appartiennent au monde du discours. Si nous choisissons de distinguer sens, dénotation et référence – et ceci est la position que nous adoptons –, le sens est lié aux termes linguistiques et par là même appartient au système de la langue, tandis que la dénotation renvoie aux représentations mentales prototypiques évoquées par ces termes. La référence, quant à elle, désigne le résultat de l'activation en contexte énonciatif de la dénotation. Il existe donc un pont 'dénotatif' entre les expressions langagières, verbales et non verbales, et ces référents. Dans les théories qui opposent exophore et endophore, ce pont est supprimé lors de la référence à l'extralinguistique au profit d'un appariement direct entre langue et réalité nommé deixis. Dans les cadres théoriques cognitifs, c'est le facteur réalité qui disparaît dans la définition de la deixis, qui désigne alors la procédure effectuée par un locuteur L afin d'insérer au sein du discours une entité nouvelle pour l'allocutaire A. Selon nous, si un écueil consiste à séparer de façon drastique réel et langage, un autre est de délaisser la réalité. Comment allier ces définitions ? Comment réconcilier prise en compte de la réalité intersubjective et considération cognitive de la deixis et de l'anaphore ? Nous nous proposons de passer par la mémoire.

La mémoire est un élément qui se retrouve dans certaines théories ; par exemple, Berrendonner (1990) l'utilise pour décrire sa 'mémoire discursive', qui s'apparente à notre conception du discours – au sein, toutefois, d'un modèle essentiellement cognitif –, et Ariel (1988) a recours à la mémoire encyclopédique et épisodique pour construire son échelle d'accessibilité. Néanmoins, si le facteur mémoriel apparaît parfois, il n'est selon nous pas toujours intégré de manière adéquate ou suffisante aux modèles proposés. Si nous accordions à la mémoire un rôle fondamental, elle permettrait de faire le lien entre réalité et cognition : l'« objet » du réel n'acquerrait de statut cognitif que lors de son passage en mémoire ; pour ce faire, l'individu psychologique devrait s'approprier cet « objet » et le conceptualiser :

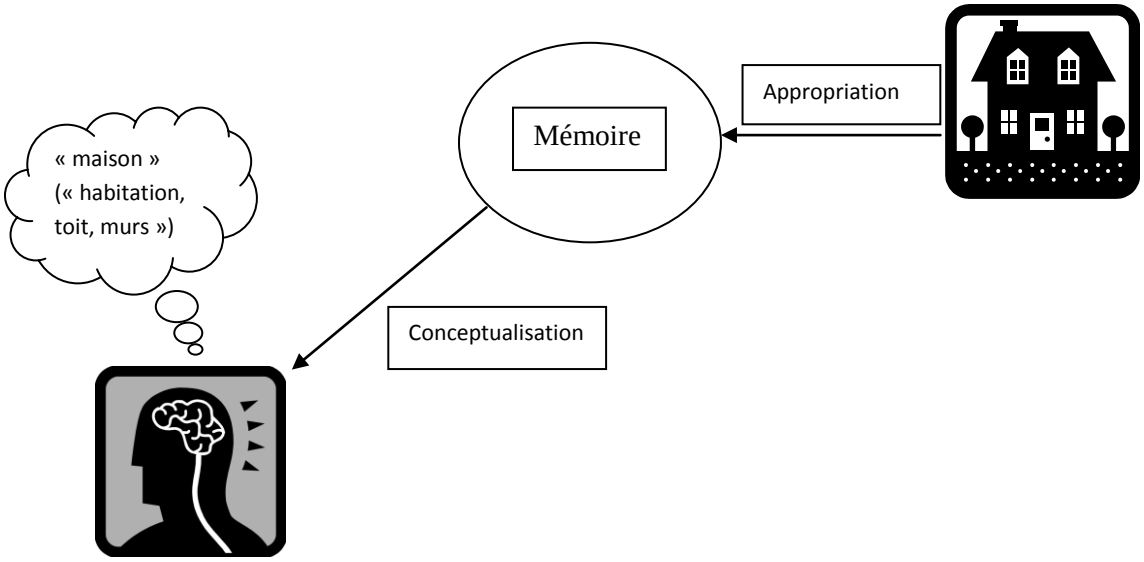


FIGURE 12 : LA MEMOIRE COMME TRANSITION ENTRE REALITE ET COGNITION

Nous pouvons dès lors restituer au réel sa primauté : si chaque individu n'avait pas de perception normative de la réalité, il ne pourrait pas conceptualiser un « objet » quelconque. De plus, nous (ré-)intégrons ici le plan de la dénotation, qui comporte l'ensemble d'images prototypiques que nous avons en mémoire à partir de notre perception du monde : les « objets concrets » mais aussi les « propriétés » que nous sommes en mesure de leur assigner. De par l'intervention de la mémoire, il serait également possible à la même expression langagière de référer par deixis et par anaphore. En effet, un élément textuel n'est plus déictique ou anaphorique : le texte n'est qu'un signal d'accès à la référence via l'activation de la mémoire, selon un procédé déictique ou anaphorique. Par conséquent, le problème posé par la correspondance directe entre expression langagière et référent, statut cognitif du référent ou

procédure indexicale ne se pose plus : c'est la mémoire qui intervient. Bien entendu, il serait erroné de penser que la mémoire agit seule : le contexte est un paramètre crucial à ne pas omettre et à intégrer dans notre modèle. Dès lors, nous proposons une description de la mémoire suivant trois couches concentriques :

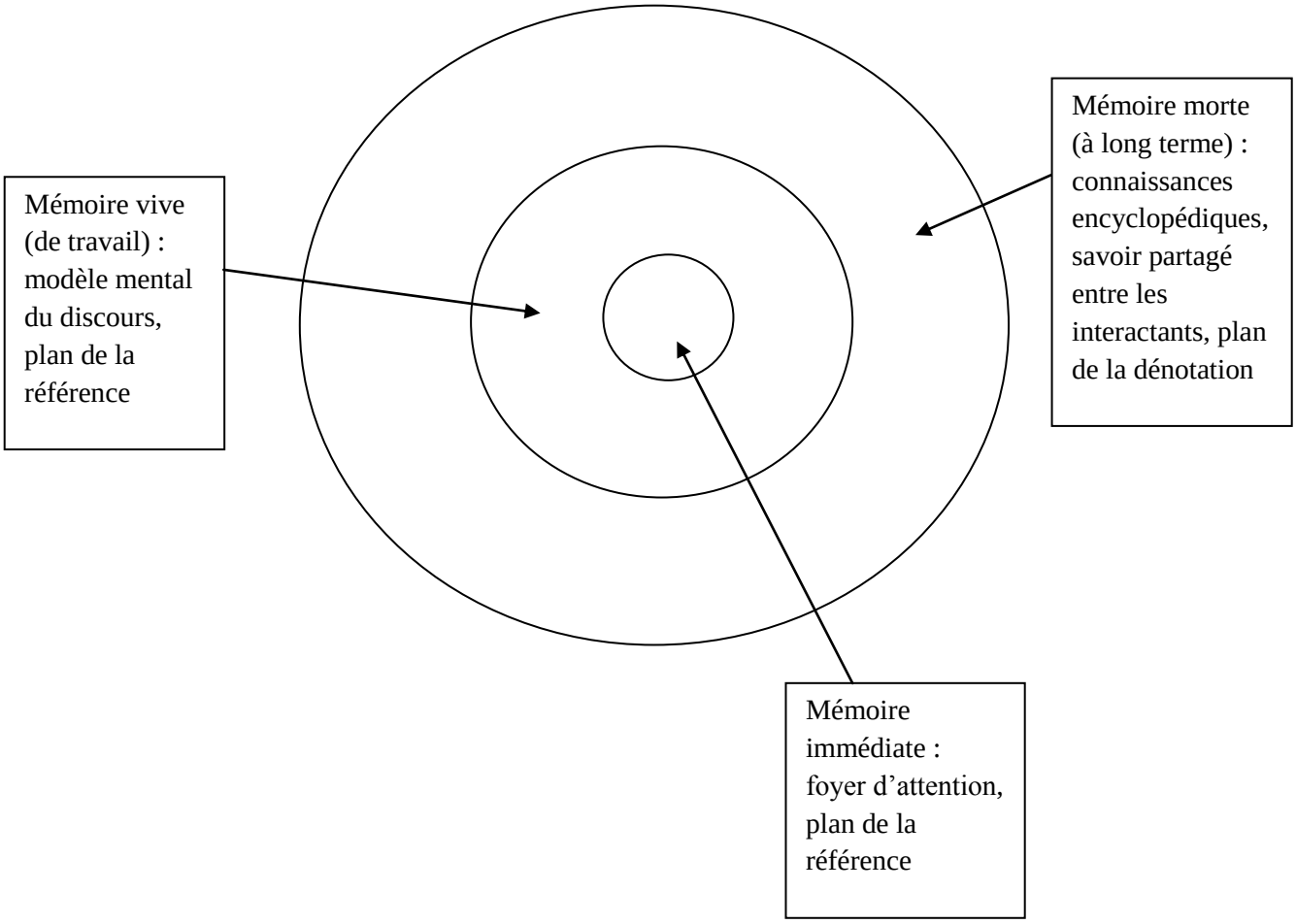


FIGURE 13 : LA MEMOIRE SELON TROIS CERCLES CONCENTRIQUES

Ainsi, un peu à la manière d'un ordinateur, la mémoire serait divisible en trois strates : la mémoire morte[66] serait constituée de notre arrière-plan de connaissances, à savoir notre expérience construite à partir de l'interaction avec le monde physico-social ainsi que le savoir partagé entre les interactants potentiels à un acte communicatif ; la mémoire vive serait

[66] La métaphore 'mémoire morte' est utilisée en informatique, raison pour laquelle nous employons ces termes. Dans la littérature existante évoquant la mémoire comme paramètre de fonctionnement référentiel, les appellations 'mémoire morte', 'mémoire inactive' ou encore 'mémoire à long terme' sont adoptées.

essentiellement fondée sur le discours en construction ; la mémoire immédiate serait celle qui représente le plus haut foyer d'attention. Nous pouvons d'ores et déjà constater le lien étroit entre contexte(s) et mémoire(s) (cf. ce chapitre, section 2.3 *supra* pour une description des contextes). Bien entendu, la problématique à résoudre est de combiner contextes et types de mémoire afin de rendre compte des phénomènes d'anaphore et de deixis, ce qui, nous l'espérons, pourra préciser le fonctionnement référentiel des pronoms démonstratifs ainsi que le rôle joué par ces derniers dans l'articulation du discours.

4. Une théorie de la référence des indexicaux : un seul mode de fonctionnement référentiel pour plusieurs accès possibles

4.1. Les éléments déictiques, les éléments indexicaux, la deixis et l'anaphore

Afin de définir notre positionnement théorique de manière claire, il convient de revenir aux fondamentaux. Notre étude étant centrée sur les pronoms démonstratifs, il est nécessaire de considérer l'ensemble des items traditionnellement appelés 'déictiques', mais aussi les expressions référentielles comme les noms propres et les syntagmes nominaux définis.

Les déictiques sont généralement divisés en deux sous-catégories. D'un côté, nous avons les déictiques purs comme *ici*, *maintenant*, *je* et *tu*, qui sont dits 'token-reflexive' dans le sens où ces items renvoient directement à l'instanciation qui leur correspond au moment de l'énonciation. D'un autre côté, nous disposons des déictiques « opaques » (Kleiber, 1983), tels que les expressions démonstratives, dont la référence repose sur le repère énonciatif sans en dépendre totalement. En effet, les déictiques opaques comportent une composante descriptive – information lexicale, genre, nombre – qui implique une construction référentielle en deux étapes, que sont le repérage énonciatif et le besoin de trouver un « objet » qui corresponde à la description fournie. Supposons que deux amies soient dans une boutique et que la locutrice L prononce *Je vais choisir cette robe pour la soirée* ; l'allocutaire A associera *cette robe* à un objet physique présent en rayon qui satisfasse l'image stéréotypique qu'elle a en mémoire. Cette double construction référentielle est également valable pour les pronoms : *celle-ci* est porteur des traits [+singulier ; +féminin] et repose sur un ancrage énonciatif pour acquérir un référent. Si nous poursuivons notre analyse en ce sens, cette remarque est également valable pour *elle*[67], pronom pourtant traditionnellement classé parmi les anaphoriques. Pourquoi alors

[67] L'emploi d'un pronom personnel dit 'anaphorique' en référence à l'extralinguistique est présenté par l'exemple (13) de ce chapitre ; les théories cognitives de la deixis et de l'anaphore énoncent que, dans ce cas, le

opposer éléments déictiques et éléments anaphoriques, puisque leur fonctionnement référentiel est identique ? C'est là que nous faisons intervenir la mémoire et que nous décidons de faire table rase des appellations 'éléments déictiques opaques' et 'éléments anaphoriques' en leur préférant le terme général d''indexicaux'. Ainsi, les individus communiquent grâce à des éléments langagiers, dont les indexicaux, qui leur permettent d'accomplir des actes de référence. Lors d'un acte communicatif existe une influence constante du réel, monde physico-social dans lequel sont immergés les individus, d'une part, et du discours qui se construit au fur et à mesure du développement de l'acte langagier, d'autre part, réel et discours faisant partie des composantes contextuelles. Nous postulons alors que c'est l'activation en contexte des éléments systémiques de la langue qui détermine l'emploi déictique ou anaphorique des indexicaux. Dès lors, il n'est plus de correspondance directe entre items langagiers et usage en discours :

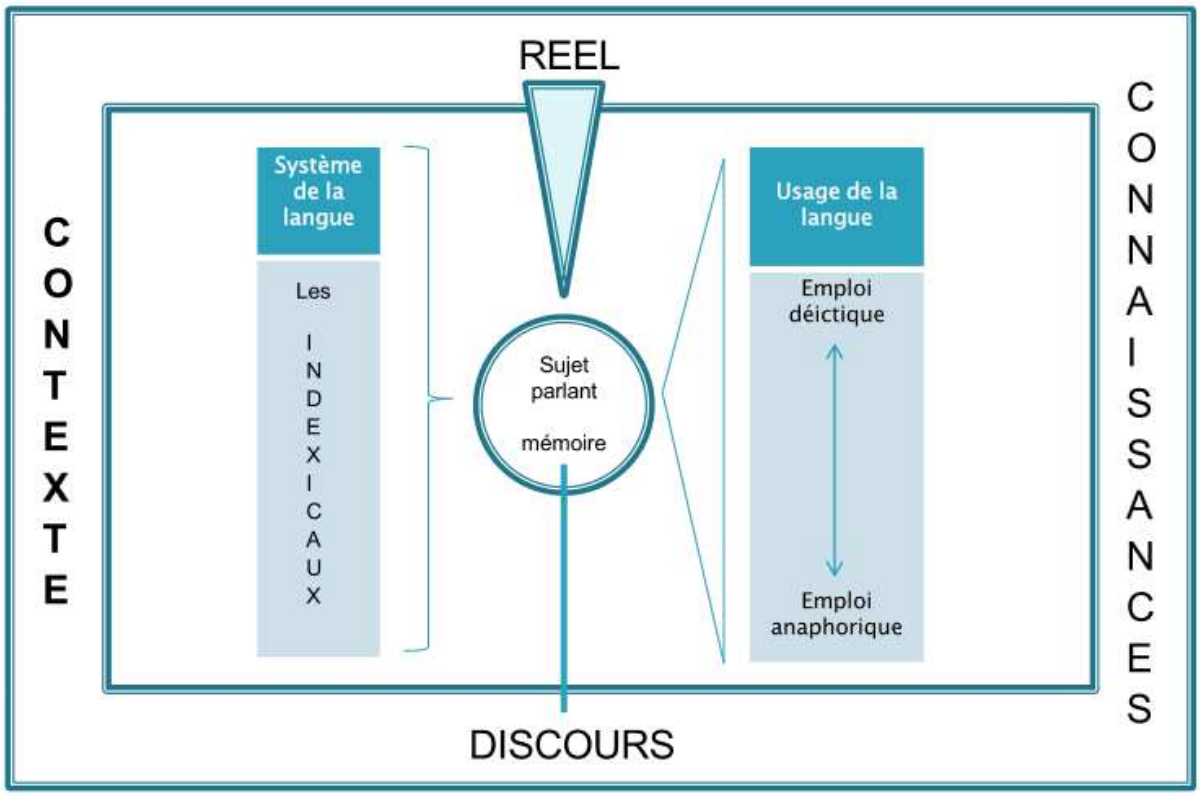


FIGURE 14 : L'ACTIVATION DES INDEXICAUX EN DISCOURS

pronom est anaphorique en raison du statut cognitif hautement saillant du référent au moment de l'occurrence.

Toutefois, si le terme ‘indexicaux’ remplace ‘éléments déictiques opaques’ et ‘éléments anaphoriques’, nous conservons les ‘déictiques purs’ – que nous appelons simplement ‘déictiques’ puisqu’ils ne sont plus en opposition avec les ‘déictiques opaques’. En effet, étant ‘token-reflexive’, ces éléments nous permettent de définir la deixis comme opérant un retour sur le repère énonciatif pour trouver un référent. Pourquoi un retour ? Parce qu’entre l’élément déictique et le référent se trouve la mémoire ; le référent n’étant pas présent au sein de la mémoire, un mouvement de rétroaction est nécessaire pour que l’élément langagier soit activé :

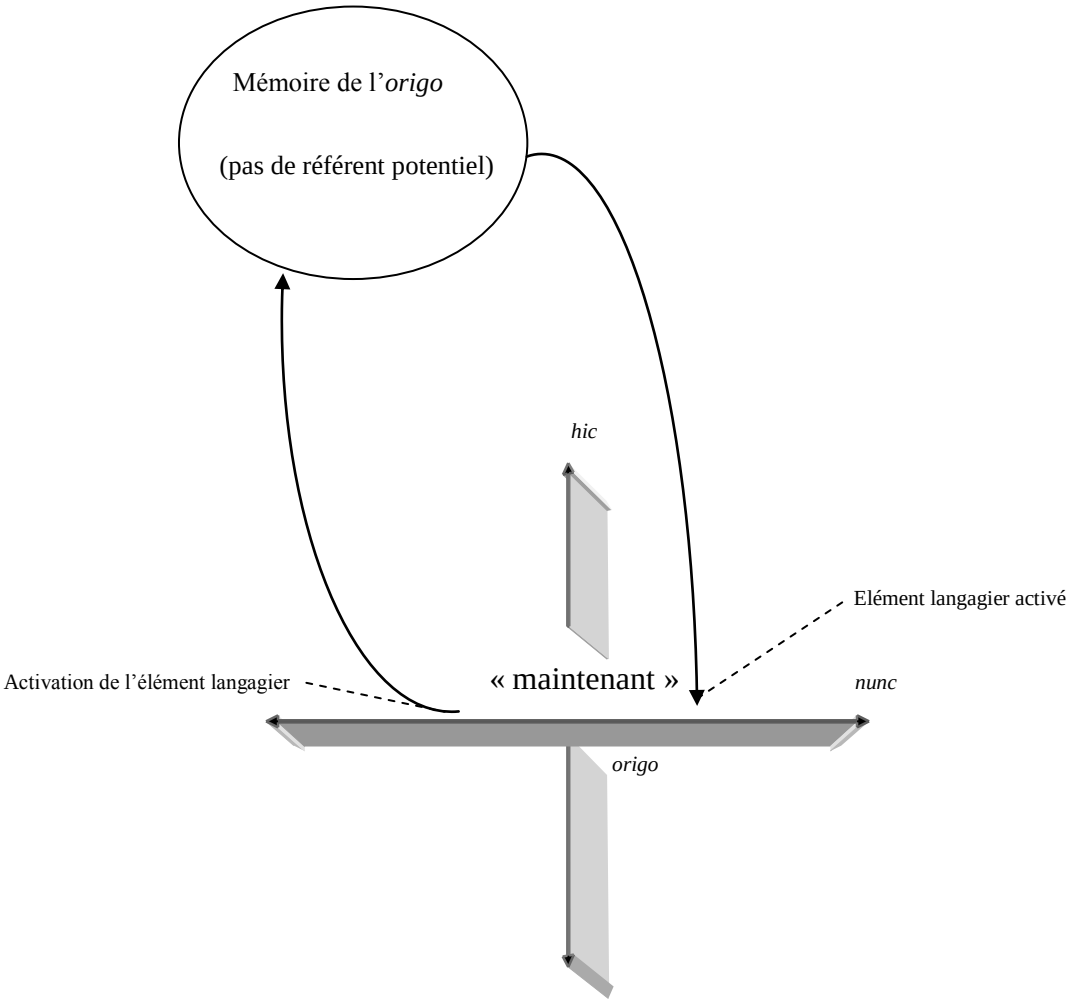


FIGURE 15 : LA DEIXIS COMME UN MOUVEMENT RETROACTIF VERS LE REPERE ENONCIATIF

En revanche, l’anaphore se réalise lorsque le référent voulu est déjà présent en mémoire focus :

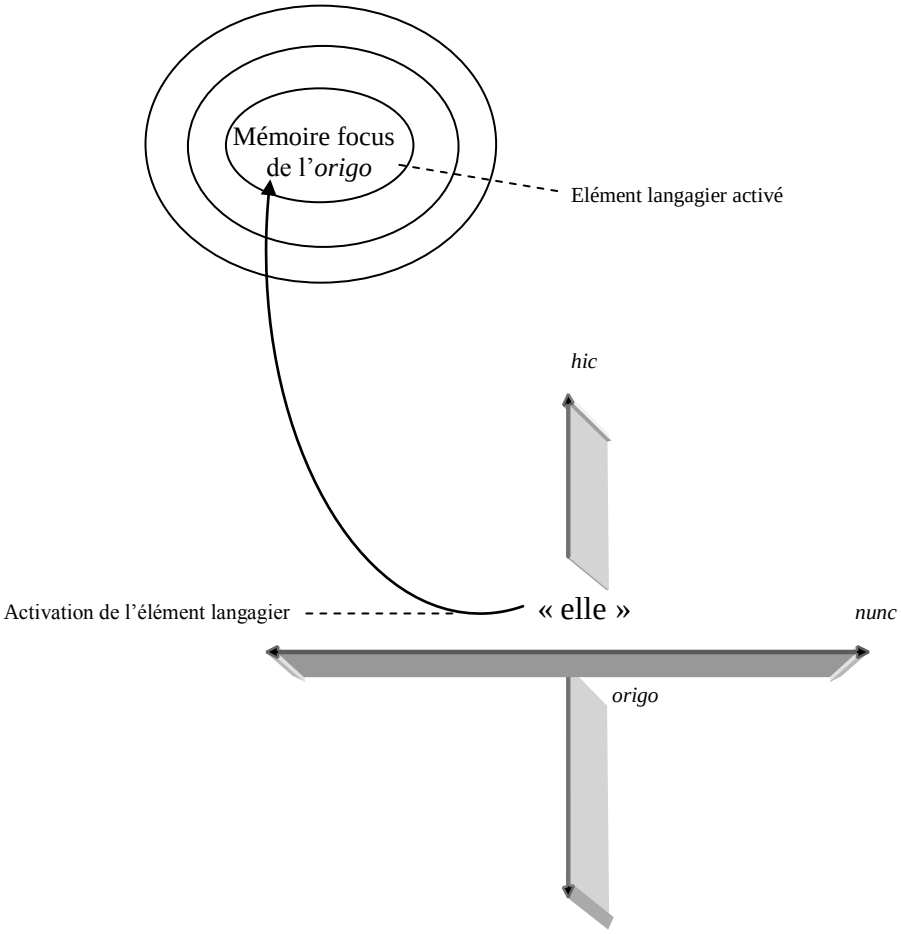


FIGURE 16 : L'ANAPHORE COMME UN ENVOI REFERENTIEL UNIQUE VERS LA MEMOIRE IMMEDIATE

Notons ici que ces schémas représentent l'énonciateur et le récepteur comme étant confondus ; ce cas pourrait se retrouver dans l'acte de lecture par exemple car, même si le véritable énonciateur est l'auteur du texte, il n'en reste pas moins que les éléments langagiers sont activés par le lecteur. Dans des cas de communication prototypiques, qui mettent en jeu au moins deux participants, l'activation de l'élément langagier est effectuée par l'énonciateur – terme que nous alternons avec celui de 'locuteur' lorsque la distinction entre les deux n'est pas pertinente – tandis que l'élément activé, c'est-à-dire le référent, est accessible à l'ensemble des allocutaires. La deixis et l'anaphore prennent alors une dimension inter-communicative basée sur un calcul de l'énonciateur sur l'état mémoriel de son destinataire.

4.2. Les conséquences de la deixis comme opérant un retour sur le repère énonciatif

Quelles sont les implications d'une définition aussi stricte de la deixis ? Selon notre théorie, seuls les déictiques 'purs' seraient à même de réaliser la deixis. Nous verrons dans les sections suivantes ainsi que lors de l'étude des pronoms démonstratifs que cette conclusion est hâtive. Toutefois, nous pouvons préciser la relation graduelle entre la deixis et l'anaphore en faisant intervenir les trois types de mémoire (cf. ce chapitre, section 3). En effet, l'anaphore pure ne s'opère que lorsque le référent est en mémoire immédiate ; si l'entité désignée est présente en mémoire vive ou en mémoire morte, la procédure employée sera moins anaphorique, et se rapprochera ainsi de la deixis. En préambule aux chapitres VII et VIII, qui, nous le rappelons, sont centrés sur l'étude des pronoms démonstratifs à l'oral en anglais et en français, il est bienvenu de recadrer ces indexicaux au sein de la famille à laquelle ils appartiennent, celle des expressions référentielles. Esquisser le fonctionnement référentiel, dans notre modèle, des noms propres et des expressions définies permet ainsi de mieux cerner la particularité des démonstratifs.

4.2.1. Le fonctionnement référentiel des noms propres

Il semble intuitif d'avancer qu'il existe une relation bi-univoque entre un nom propre et son référent : *François Hollande* renvoie au « Président de la République Française » – sachant qu'au moment où nous écrivons, François Hollande dirige notre pays – et quiconque entendra ce nom propre saura identifier de qui il s'agit. En d'autres termes, quel que soit le sens – *i.e.* le mode de présentation (cf. ce chapitre, section 1.1) –, le référent sera le même : *François Hollande, le Président de la République Française, le vingt-quatrième Président de la France* renvoient au même référent. Dans un cadre communautaire plus restreint, l'évocation du nom *Jacques Durand* renverra sans nul doute au « directeur de thèse de Laurie Buscail ». Bien que « François Hollande » fasse partie des connaissances encyclopédiques et « Jacques Durand » des connaissances partagées, il demeure que c'est la mémoire morte qui est activée. Néanmoins, si notre allocutaire ne possède pas l'information « Jacques Durand » en mémoire, il se verra contraint de questionner l'identité de la personne qui correspond au nom *Jacques Durand*. Nous pouvons alors avancer que ce dernier cas relève de la deixis : le référent n'étant pas présent en mémoire, un mouvement rétroactif vers le repère énonciatif est opéré, afin de trouver un référent adéquat au moment où l'individu détiendra assez d'informations pour se construire une image satisfaisante alors associée au nom *Jacques Durand*. Revenons à présent sur cet appariement (apparemment) unaire. Il est vrai que

l'évocation d'un nom propre complet semble renvoyer à un référent unique. Si tel est le cas, cela implique que le plan de la dénotation et le plan de la référence sont confondus. En effet, pourquoi passer par une étape intermédiaire si l'activation en discours aboutit au même résultat immuable ? Cependant, nous pourrions avoir dans notre base de données mémorielle deux instances possibles pour le même nom propre – problème qui contraint certains artistes à utiliser un pseudonyme. Dans le cas majoritaire où un nom propre renvoie à un seul référent possible, il n'en reste pas moins que l'image prototypique correspondante à ce nom, c'est-à-dire la dénotation, reste en mémoire, à l'état latent ; le référent, qui est la valeur activée de cette image, n'est rendue possible et accessible que lors d'une instanciation en contexte énonciatif. Ainsi, nous choisissons de conserver la distinction entre sens, dénotation et référence, même pour les noms propres.

4.2.2. Le fonctionnement référentiel des expressions définies

Nous posons que l'article *the, le, la* ou *les* dans les expressions définies signale à l'allocutaire d'accéder à un référent qui soit directement identifiable – contrairement à l'article indéfini qui suppose une extraction. L'accès à la référence se fait via la dénotation du noyau lexical, qui doit être suffisante pour l'identification de l'entité voulue lors de son activation en discours ; moins le référent est accessible au destinataire, plus l'information lexicale fournie se doit d'être riche : *le professeur d'anglais de l'université de Perpignan* désigne un référent moins accessible que *le professeur d'anglais* qui désigne un référent moins accessible que *le professeur* qui désigne un référent moins accessible que *il*. Cette relation d'identification implique que l'entité visée soit présente en mémoire, plus ou moins centrale selon le degré d'accessibilité du référent. Dès lors, nous pouvons mettre en relation la quantité d'information véhiculée par le noyau lexical, le type de mémoire activée et la procédure déictique ou anaphorique : plus l'information lexicale est riche, plus le référent visé est en mémoire morte, plus la procédure est déictique ; plus l'information lexicale est pauvre, plus le référent visé est en mémoire centrale, plus la procédure est anaphorique. La possibilité d'activer les trois strates de la mémoire permet d'expliquer les expressions définies en tant qu''anaphores associatives' – *J'ai visité un petit village,* ***l'église*** *était très jolie* – où l'identification de « l'église » repose à la fois sur les connaissances du monde, et, donc, la mémoire morte – il est un fait que chaque village comporte une église –, et sur la mémoire vive, grâce à l'évocation préalable du « village ». De même, nous pouvons considérer le fait qu'une expression définie puisse être déictique ; si l'allocutaire A est confronté à l'énoncé *Le*

chien de mon voisin a hurlé toute la nuit sans que « le chien de mon voisin » ne fasse partie de sa banque de données mémorielle, A s'accommodera de ce référent : *le chien de mon voisin* ne trouvant pas de référent correspondant en mémoire, un mouvement rétroactif sera effectué pour que A construise *hic et nunc* une entité qui corresponde « au fait que j'ai un voisin et que celui-ci a un chien ».

Dans le cas extrême des pronoms personnels, l'information lexicale est à son minimum, réduite à la marque du genre et du nombre. Plusieurs conclusions découlent de cette analyse : tout d'abord, les pronoms personnels activent la mémoire la plus centrale, c'est-à-dire la mémoire focus, pour rendre possible l'accès à la référence ; la conséquence directe est que, par l'activation de la mémoire focus, les pronoms personnels réalisent l'anaphore ; enfin, même minime, un indice lexical est fourni, ce qui implique que les pronoms soient composés d'un déterminant et d'une catégorie nominale sous-jacente.

4.2.3. Le fonctionnement référentiel des expressions démonstratives

Nous posons que le déterminant *this*, *that*, *ce*, *cette* ou *ces* dans les expressions démonstratives signale à l'allocutaire une identification du référent en deux étapes. Cette double construction référentielle peut s'effectuer de diverses manières, la plus évidente étant celle en fonction de la situation directe ; si nous reprenons l'exemple des deux amies achetant des vêtements (cf. ce chapitre, section 4.1), l'énoncé *Je vais choisir cette robe pour la soirée* contraint l'allocutaire A à effectuer un retour sur le repère *hic et nunc* pour trouver un référent correspondant à *cette robe*. L'emploi de cette robe relève ici tout simplement de la deixis, me direz-vous – si tant est que l'on considère que « la robe en question » n'a pas été évoquée auparavant. Quelle est alors la différence avec la deixis effectuée par *le chien de mon voisin* ? Cette différence repose sur le processus d'accommodation. Si *cette robe* entraîne un retour sur le repère énonciatif, *le chien de mon voisin* s'accommode d'un référent en mémoire morte, comme si « le chien en question » faisait déjà partie de la base de données mémorielle :

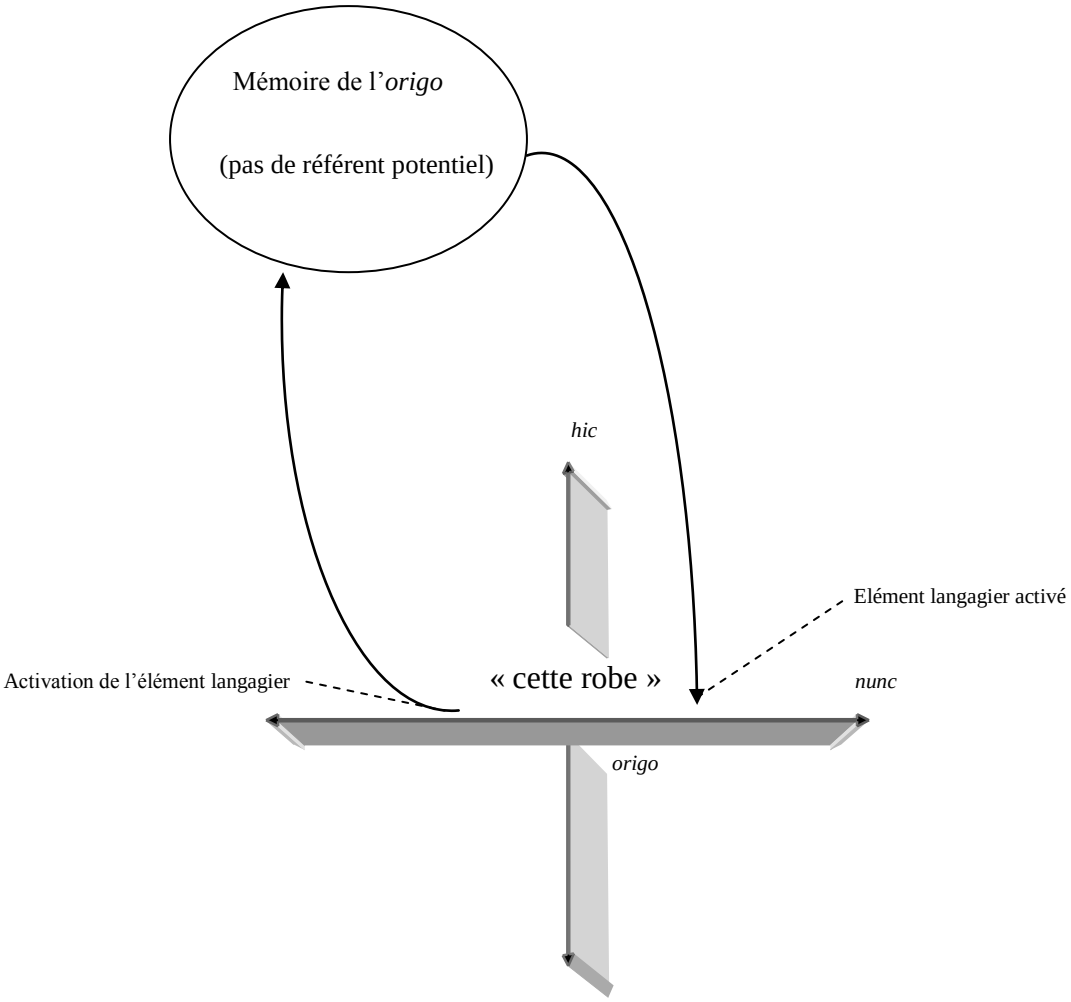


FIGURE 17 : UNE CONSTRUCTION REFERENTIELLE EN DEUX ETAPES POUR LES EXPRESSIONS DEMONSTRATIVES

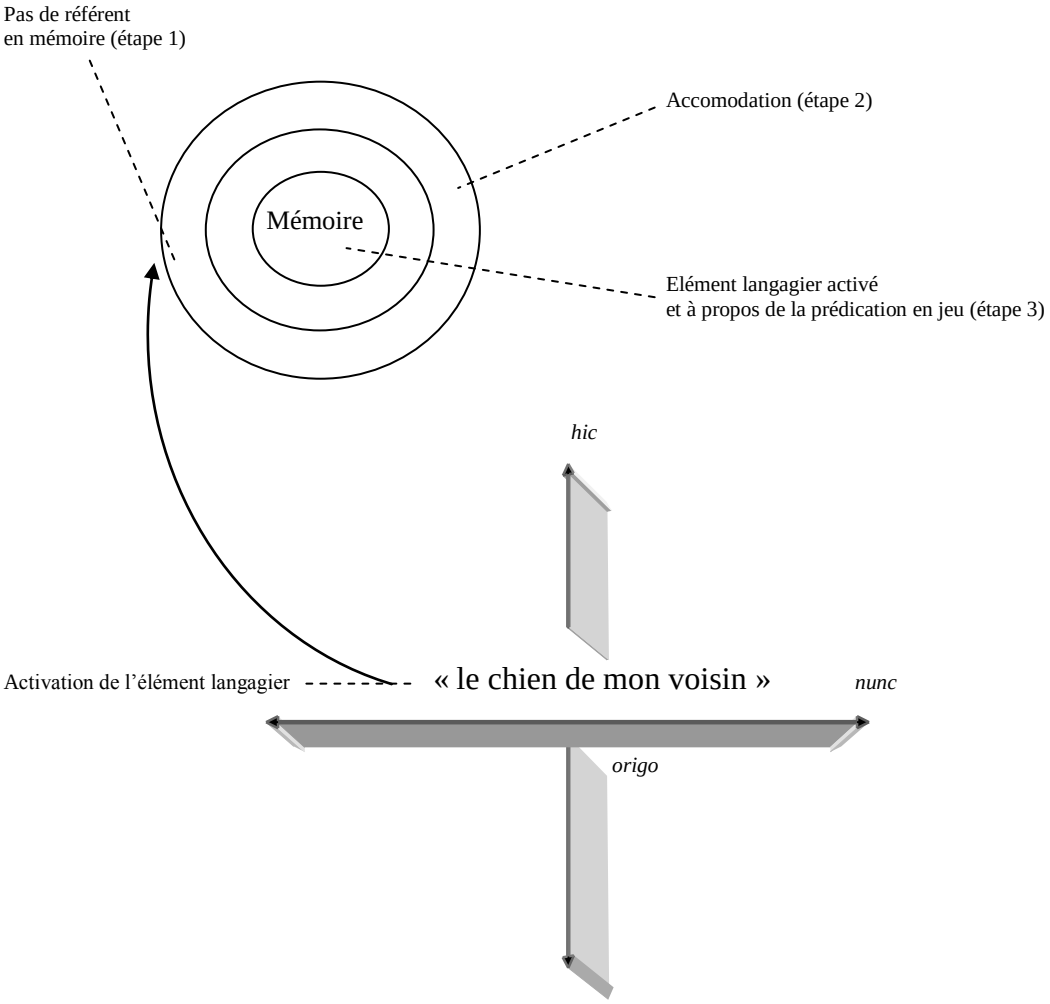


FIGURE 18 : L'ACCOMMODATION D'UNE EXPRESSION DEFINIE S'EFFECTUE DANS LA MEMOIRE

Une fois « le chien de mon voisin » acquis comme s'il faisait partie de la banque de données, le référent pourra migrer vers la mémoire focus afin d'être inséré dans une prédication dont il est l''à propos'[68].

Revenons alors à l'exemple du shopping entre amies. Supposons cette fois-ci que « la robe en question » ait déjà été évoquée : la locutrice L a essayé plusieurs robes, dont « celle-

[68] L''à propos' est une notion définie par rapport à celle d''aboutness' de Lambrecht (1994), qui définit la relation entre le topique et l'information véhiculée par l'énoncé : « A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent ; *i.e.* as expressiong information which is relevant to and which increases the addressee's knowledge of this referent. » (Lambrecht, 1994 : 131)

ci », et prononce *Je vais choisir cette robe pour la soirée*. Dans ce cas, l'identification en deux temps s'explique par la sélection que l'allocutaire A effectue afin d'identifier « la robe choisie parmi l'ensemble des robes essayées ». Cette double construction référentielle permet d'expliquer les emplois métaphoriques des expressions démonstratives, comme dans l'exemple suivant :

(15) *Les ouvriers considérèrent le plancher du wagon et décidèrent que pour une nuit, on pouvait se satisfaire de* ***ce lit***. (exemple (27) de Corblin, 1987 : 212)

La double construction référentielle permet au destinataire du message de ré-identifier « le plancher du wagon » comme étant comparé à « un lit ». Notons la différence avec l'emploi d'un SN défini :

(16) *Les ouvriers considérèrent le plancher du wagon et décidèrent que pour une nuit, on pouvait se satisfaire* ***du lit***.

La construction référentielle produite par le SN défini n'étant pas double mais directe, *le lit* ici mène le destinataire à chercher un référent qui s'accorde à la dénotation du noyau lexical *lit* ; bien entendu, cette interprétation entraîne une incompréhension du message, « le plancher du wagon » et « le lit » étant alors deux entités distinctes. Dans le cas des expressions démonstratives métaphoriques, l'identification du référent ne passe pas par la dénotation du noyau lexical *per se*, mais par la réinterprétation de cette dénotation. Par ce fonctionnement référentiel, nous pouvons également expliquer les cas d'expressions démonstratives dont l'activation en discours passe par les connaissances partagées des allocutaires. Considérons à nouveau l'exemple *le chien de mon voisin a hurlé toute la nuit*, mais en modifiant le contexte : l'allocutaire A connaît l'existence du « chien du voisin » et sait que celui-ci hurle souvent. Nous pourrions alors être confrontée à l'exemple suivant, sans mention antérieure du « chien » :

(17) *Ce chien a hurlé toute la nuit*.

La double construction référentielle produite par le SN démonstratif permet à l'allocutaire de faire appel à sa mémoire morte pour déduire que *ce chien* réfère au « chien du voisin ». L'information lexicale fournie ici étant pauvre, un SN défini serait insuffisant, et nécessiterait

par conséquent l'ajout du complément du nom *de mon voisin* afin d'identifier le référent voulu :

(18) L : *Le chien a hurlé dans la nuit.*
A : *De quel chien parles-tu ?*

(19) L : *Le chien de mon voisin a hurlé toute la nuit.*
A : *Encore ??*

Comment pourrions-nous alors expliquer la différence de fonctionnement référentiel entre les expressions définies et les expressions démonstratives ? Nous pourrions dire que, si la construction référentielle des expressions définies passe directement par la dénotation du noyau lexical, celle des expressions définies s'effectue par cette dénotation à laquelle s'ajoute un recours au(x) contexte(s), situationnel immédiat ou mémoriel. Ainsi, nous classons les expressions définies et les pronoms personnels parmi les indexicaux faibles, tandis que les expressions démonstratives sont des indexicaux forts. L'enjeu des chapitres VII et VIII sera d'étudier ce processus référentiel lors de l'emploi des pronoms démonstratifs.

Chapitre V. Représentation des expressions référentielles en Grammaire Notionnelle Dépendancielle

Pour mener à bien l'étude des démonstratifs, nous sommes confrontée à un double problème : offrir une description du fonctionnement référentiel de ces items et proposer un système de notation explicite qui puisse rendre compte des éléments langagiers, de leur distribution et de leurs valeurs potentielles en discours. La Grammaire Notionnelle Dépendancielle (désormais GND) de John Anderson (1991, 1997, 2007, 2011)[69] est une réponse à la question de la formalisation. Toutefois, si ce procédé représentationnel est utile pour schématiser la structure interne des éléments du lexique et les projections que ceux-ci peuvent obtenir, il n'engage en rien sur la réalité psychologique des phénomènes discursifs.

1. La Grammaire Notionnelle Dépendancielle : une introduction

Nul ne saurait contredire l'idée que les mots renvoient à des concepts mentaux et que nos facultés langagières nous mènent à organiser ces mots selon un ordre précis. De plus, ces mots, en théorie comptables à l'infini, se rangent parmi un ensemble fermé de catégories. Il s'agit donc de répondre à la double exigence des éléments langagiers, qui sont l'appartenance à un système formel et la faculté de susciter des phénomènes cognitifs, ce problème étant à moitié soulevé par le fameux *Colourless green ideas sleep furiously* de Chomsky dans *Syntactic Structures* (1957). La GND permet de rendre compte à la fois des contraintes syntaxiques et des contraintes sémantiques des énoncés. En effet, la GND rejette la théorie 'autonomiste' qui propose une syntaxe semblable à un ensemble d'algorithmes libre de toute 'substance'. L'hypothèse ici est que les catégories syntaxiques sont déterminées par l'image mentale qu'elles suscitent et représentées par les signes linguistiques, qui sont eux-mêmes liés aux catégories phonétiques représentant la perception humaine de la substance phonique :

[69] Les travaux de John Anderson dans la GND s'enracinent dans la grammaire des cas qu'il a contribué à développer dans les années soixante et soixante-dix (cf., par exemple, Anderson, 1971), et pour laquelle il a collaboré avec Roger Böhm (cf., par exemple, Böhm, 1982).

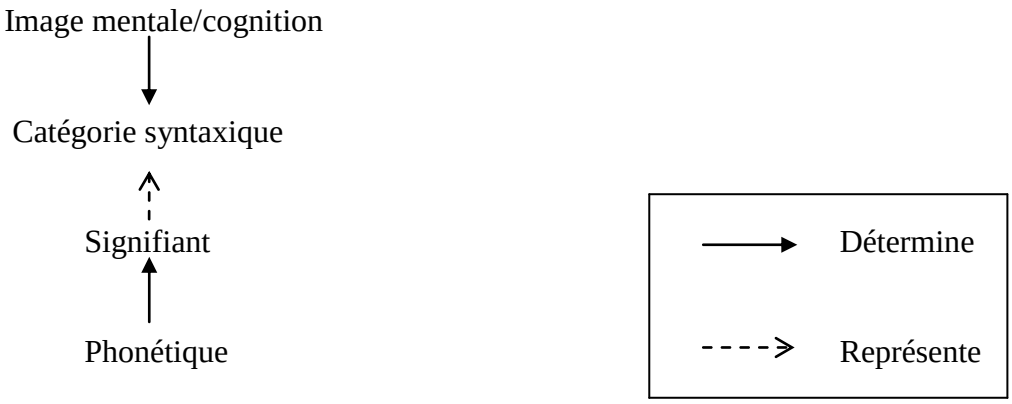


FIGURE 19 : LA STRATIFICATION SCHEMATIQUE DE LA GND

Ainsi, le lexique précède la phonologie, mais aussi la syntaxe, qui sert à son tour la phonologie, en particulier la structure prosodique des énoncés :

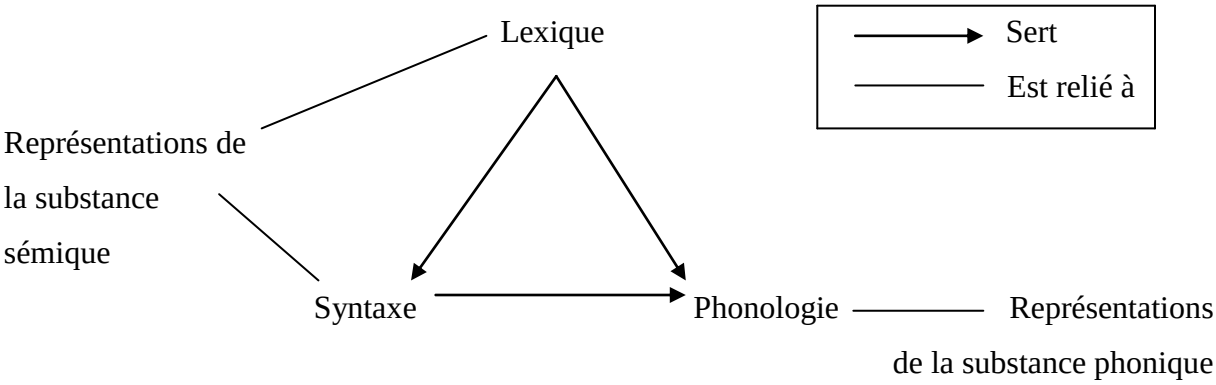


FIGURE 20 : LA REPRESENTATION SCHEMATIQUE DE LA GND (ANDERSON, 2011 : 110)

La GND se veut être une grammaire explicite, descriptive, par une combinaison de symboles qui représentent les éléments atomiques que sont les mots, regroupés en catégories syntactico-sémantiques. Ces catégories forment à leur tour des constituants syntagmatiques à structure asymétrique dont les dépendants sont régis par la tête. Les relations de complémentation (Figure 21) et de modification (Figure 22) sont alors représentées de la manière suivante (Anderson, 2011 : 115-116) :

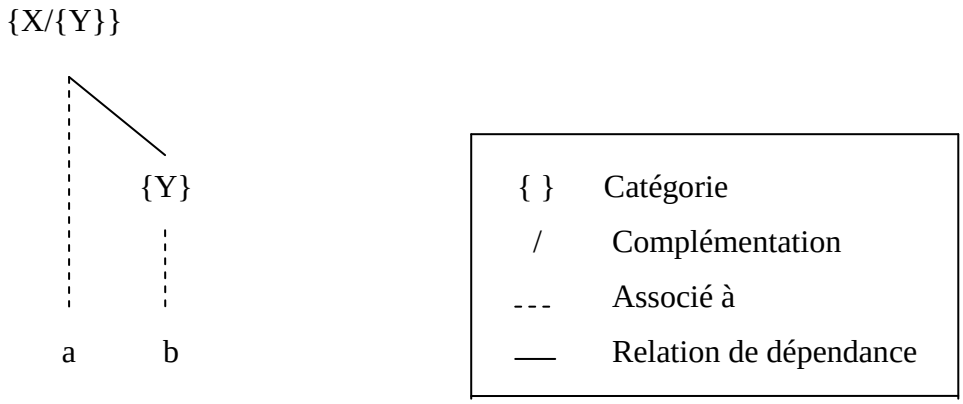


FIGURE 21 : LA RELATION DE COMPLEMENTATION EN GND (ANDERSON, 2011 : 115)

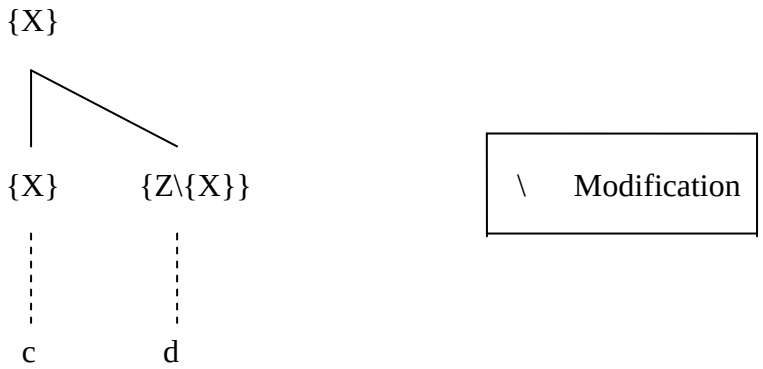


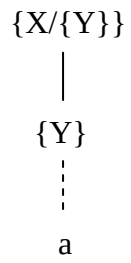
FIGURE 22 : LA RELATION DE MODIFICATION EN GND (ANDERSON, 2011 : 116)

Sur ces schémas, les accolades { } entourent le symbole représentatif de la catégorie concernée, X pouvant être instancié, dirons-nous, par un nom ou par un verbe. Ce symbole peut apparaître seul ou accompagné de celui de la catégorie dépendante, indiqué à droite d'une barre oblique marquant la relation de complémentation (/) ou à gauche d'une barre oblique inversée marquant la relation de modification (\)[70]. Ainsi, la représentation catégorielle au sommet de l'arbre en Figure 21 explicite le fait que la catégorie X prend Y comme complément. {X/{Y}} décrit alors la structure interne de l'élément 'a' du lexique considéré, {X/{Y}} et 'a' étant associés par une ligne en pointillés. Notons ici que la Figure 21 représente une relation de complémentation où 'a' et 'b' – 'b' étant associé à {Y} – sont

[70] L'ordre des catégories dans les notations {X/{Y}} et {Z\{X}} ne reflète pas nécessairement des relations de précédence linéaire ; dans nos schémas, cependant, {Y} apparaissant à droite de {X} et {Z} à gauche de {X} suit la position prototypique des compléments à droite de la tête et des modifieurs à gauche de celle-ci, comme c'est le cas pour l'anglais.

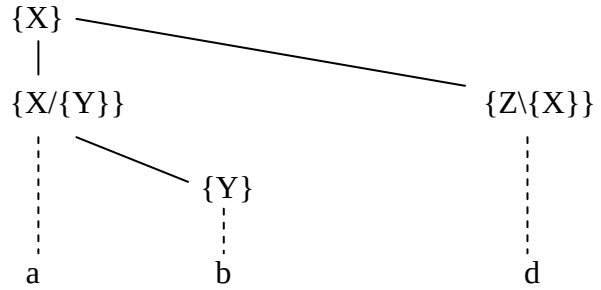
distincts ; toutefois, la structure interne {X/{Y}} peut se manifester au sein d'un seul et même item :

(1)



L'arbre en Figure 22, quant à lui, schématise un membre 'c' de la catégorie {X} comme étant accompagné d'un élément 'd' de catégorie {Z} en relation de modification. Contrairement aux compléments, les modifieurs ne sont pas requis par la structure interne de l'élément considéré. Cependant, les modifieurs pouvant aussi être compléments, il est nécessaire de pouvoir représenter cette double relation potentielle. De plus, une tête est susceptible de régir à la fois des compléments et des modifieurs. Ainsi, {Z\{X}} n'est pas directement relié à {X} mais à une projection de {X}, ce qui permet d'avoir des représentations telles que :

(2)



où 'b' et 'd' dépendent de 'a', le premier étant en relation de complémentation avec la tête rectrice et le second en relation de modification.

La structure de la GND rappelle sans nul doute certains aspects de la théorie X-barre chomskyenne ; en effet, ces deux grammaires partagent le même idéal d'explicitation par une représentation symbolique et l'idée d'une génération de constituants à partir du lexique, dont les éléments atomiques sont régis par le principe de compositionnalité. Néanmoins, la différence fondamentale entre la GND et les préceptes générativistes se joue dans la conception et la combinaison des traits distinctifs qui déterminent les catégories lexicales primaires.

2. Les catégories syntactico-sémantiques en Grammaire Notionnelle Dépendancielle

L'identification des catégories syntactico-sémantiques en GND se fait par l'interaction de deux traits notionnels distinctifs qui sont N (trait référentiel) et P (trait prédicatif), que l'on pourrait rapprocher des traits [± N, ± V] chomskyens. Toutefois, si les préceptes générativistes considèrent N et V comme des traits binaires parallèles, la GND conçoit N et P comme des traits unaires ou monovalents constituant deux pôles prototypiques de nature très différente : N manifeste un pouvoir référentiel tandis que P exprime le caractère essentiellement relationnel des prédicats. Ainsi, les extrémités N et P complexifient le continuum catégoriel et leurs combinaisons déterminent les classes lexicales et fonctionnelles des éléments langagiers.

2.1. Les catégories lexicales en Grammaire Notionnelle Dépendancielle

Le nom et le verbe sont les catégories syntactico-sémantiques basiques et intuitives, car elles semblent être en distribution complémentaire dans toutes les langues – même si certaines ne distinguent pas noms et verbes sur le plan lexical : le nom prototypique (*boy*, *table*, *computer*, *chaise*, *arbre*) désigne une « entité » tandis que le verbe prototypique (*go*, *walk*, *give*, *construire*, *parler*) décrit un « événement », « entité » et « événement » étant des concepts cognitifs. Ainsi, les « entités » canoniques sont stables et discrètes et les « événements » prototypiques sont essentiellement dynamiques et relationnels ; en effet, les verbes dénotent un ensemble de concepts délimités dans le temps qui sont appliqués aux concepts spatialement circonscrits dénotés par les noms. Cette intuition concernant le caractère relationnel inhérent au verbe disparaît dans la description des catégories lexicales chomskyennes, de par le parallélisme effectué entre les traits syntaxiques N et V :

(3) Noms [+N –V]
Verbes [-N +V]
Adjectifs [+N +V]
Prépositions [-N –V]

En revanche, le trait notionnel P en GND souligne le caractère relationnel des prédicats, et par un jeu de combinaisons potentielles entre N et P, il est désormais possible d'expliquer, par exemple, que les adjectifs – une catégorie qui est loin d'être universelle – manifestent des propriétés à mi-chemin entre noms et verbes. Les prépositions sortent du système lexical car

elles appartiennent aux catégories fonctionnelles. Enfin, la GND traite comme trompeur le parallélisme établi en syntaxe X-barre entre noms et verbes sur le plan de la complémentation à partir d'exemples comme ceux en (4) et (5) :

(4) *His fear of darkness.*
He fears darkness.

(5) *He is a student of French.*
He studies French.

En effet, des noms comme *fear* ou *student* sont dérivationnellement liés à des prédicats dont ils héritent une partie de la complémentation. A la différence de ces noms dérivés, les membres centraux de la catégorie des substantifs – *boy*, *table*, *computer*, *chaise*, *arbre* – n'ont précisément pas de compléments. La GND distingue cinq catégories lexicales primaires (les prépositions en sont exclues, cf. section 2.2 *infra*) associées aux combinaisons pondérées de N et de P :

(6)

	N	N;P	N:P	P;N	P
	noms propres	noms communs	adjectifs	verbes	auxiliaires
N/P =	4/0	3/1	2/2	1/3	0/4

Dans cette représentation, ';' signifie 'prépondère' et ':' indique une égale distribution de N et de P ; ainsi, les noms propres constituent la catégorie lexicale au pouvoir référentiel maximal et les auxiliaires la catégorie lexicale purement relationnelle. Les noms communs sont composés d'une prépondération de N par rapport à P, les verbes d'une prépondération de P par rapport à N, et les adjectifs, à mi-chemin entre les noms et les verbes, comportent autant de N que de P.

Les combinaisons de N et de P décrivent la structure primaire des catégories lexicales, à laquelle peuvent s'ajouter un ou plusieurs traits secondaires, notés entre accolades. Ces traits peuvent ou non être morphologiquement marqués, comme le temps et l'aspect des verbes ou le genre et le nombre du nom. Ainsi, nous pourrions avoir :

(7) *boy*
{N;P{sg}}

Nous aurons l'occasion de parcourir un certain nombre de ces traits secondaires dans les sections 3.1, 3.2 et 3.3 *infra*.

2.2. Les catégories fonctionnelles en Grammaire Notionnelle Dépendancielle

Contrairement aux catégories lexicales, les catégories fonctionnelles primaires ne sont pas définies par un jeu de combinaisons entre N et P. Anderson (1997 : 21, 2011 : 125) distingue quatre catégories fonctionnelles, '.' signifiant une coprésence de N et de P :

(8) Finiteness ('finitude') {P}
Determinative ('déterminatif') {N}
Comparator ('comparateur') {P.N}
Functor ('foncteur') { }

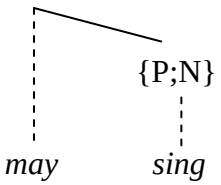
La structure notionnelle de ces catégories met en jeu de simples combinaisons de N et de P, y compris la combinaison nulle. A chacune des trois premières catégories de la liste en (9) correspond un représentant lexical prototypique, suivant la prépondérance de N ou de P de ce dernier :

(9)

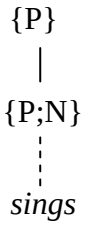
Catégories fonctionnelles		Catégories lexicales	
Finiteness	{P}	Verb	{P;N}
Determinative	{N}	Noun	{N;P}
Comparator	{P.N}	Adjective	{N:P}

La finitude maximise le caractère relationnel et produit des structures qui sous-tendent l'accomplissement d'actes de parole, la détermination est liée au pouvoir référentiel, et le trait 'comparateur' ajoute une dimension possible de comparaison manifeste aux adjectifs graduels, comme c'est le cas dans *more comfortable*. Ces catégories fonctionnelles étiquettent la tête syntagmatique et peuvent être exprimées de diverses manières dans le système langagier, que ce soit par un mot distinct – comme *more* dans *more beautiful*, ou encore *may* dans *may sing* – ou par une marque flexionnelle – comme dans *sing**s*** :

(10)

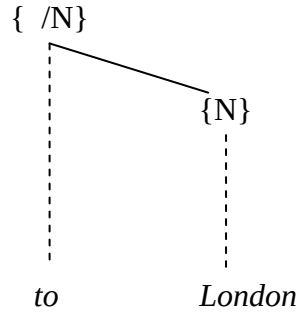


(12)

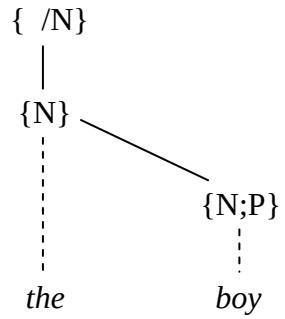


La catégorie fonctionnelle 'foncteur', marquée { }, définit un ensemble de termes langagiers purement relationnels, indépendamment de leur contenu sémantique. Parmi cet ensemble d'éléments, nous pouvons noter les prépositions – *at, to, for*, qui entrent dans la liste des catégories lexicales dans la théorie X-barre – mais aussi les déterminants, les auxiliaires et les comparatifs ; toutefois, ces trois derniers reçoivent une étiquette catégorielle lexicale, à la différence des prépositions. De plus, le foncteur peut apparaître sous forme de flexion (exemple (15)) :

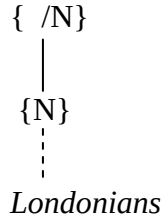
(13)



(14)



(15)



{ /N} dans les exemples (13), (14) et (15) signifie que l'élément fonctionnel requiert un complément ('/') nominal ('N'). Une caractéristique fondamentale des catégories

fonctionnelles est qu'elles sont toutes rectrices d'un complément, ce qui justifie leur étiquette fonctionnelle. Ainsi, nous pouvons modifier la notation des catégories fonctionnelles listées en (9) de la manière suivante :

(16) Finiteness {P/}
Determinative {N/}
Comparator {P.N/}
Fonctor { /}

Ces éléments fonctionnels, têtes rectrices de syntagmes et définis par la relation qui leur incombe au sein de l'énoncé, ne sont pas sans rappeler les opérateurs des théories énonciatives (cf. chapitres II et III).

3. Les expressions référentielles en Grammaire Notionnelle Dépendancielle

Dans cette section, nous souhaitons expliciter le fonctionnement des noms propres et des syntagmes nominaux définis et démonstratifs par la représentation de syntagmes simples en GND. Ces trois types d'expressions référentielles appartiennent à la catégorie fonctionnelle des 'déterminatifs', puisqu'elles réfèrent à une entité – ce qui correspondrait à un DP en termes X-barre. Cette première formalisation sert à la fois d'introduction et de base théorique à l'étude syntactico-sémantique des pronoms démonstratifs que nous effectuerons dans les chapitres VII et VIII.

3.1. Les noms propres en Grammaire Notionnelle Dépendancielle

Les noms propres montrent un certain nombre de caractéristiques prototypiques qui les différencient des noms communs[71]. En effet, même si l'on accorde une dénotation aux noms propres (cf. chapitre IV, section 4.2.1), l'accès à la référence, c'est-à-dire à l' « objet »

[71] Il est un phénomène particulièrement remarquable que nous devons souligner et qui alimente l'idée d'un fonctionnement des noms propres différent de celui des noms communs : certaines langues ne différencient pas les noms communs des adjectifs ou des verbes ; la valeur nominale ou prédicative est acquise en discours, lors de l'ancrage énonciatif ; en revanche, ces langues comportent une catégorie 'nom propre'. Ainsi, si la différence entre nominal et verbal ne s'obtient que par l'acte de parole, les noms propres sont d'ores et déjà distingués dans le lexique (cf. Anderson, 1997 : 23-29 et Lyons, 1977 : 433 à ce sujet). Cette observation va également dans le sens de la théorie de Culioli (cf. chapitres II et III, section 3.3) qui ne marque pas le nominal et le verbal au niveau notionnel (niveau 0) ; ces propriétés s'acquièrent grâce à la lexis et à la mise en œuvre d'opérations du type 'extraction', 'fléchage' ou 'parcours'.

désigné en contexte énonciatif, se fait de manière directe : l'image dénotée et l'image référentielle sont la plupart du temps confondues, ce qui implique que ni la mémoire focus, ni la mémoire vive, n'interviennent dans l'activation d'un nom propre en discours ; seule la mémoire morte agit. En d'autres termes, les participants à l'acte communicatif sont à même d'identifier le référent voulu sans l'intervention d'un contexte particulier. D'autre part, les noms propres ne sont pas sensibles à la distinction comptable *vs* non-comptable, et leur emploi en tant que vocatifs les rapproche des déictiques purs :

(17) a. *Mary!*
b. *Mary! Come and see that!*
c. *You! Come and see that!*

De plus, notons que, dans un premier temps, les noms propres sont plutôt réfractaires à la modification restrictive, ce qui les rapproche des pronoms. Par conséquent, un nom propre est catégorisé en GND comme un élément référentiel pur :

(18) {N}
|
London

Toutefois, il est des traits notionnels secondaires signalés par les noms propres, qui sont certes minimes en ce qui concerne leur pouvoir descriptif, mais à ne pas omettre pour autant. Ainsi, *London* et *François Hollande* renvoient à une entité unique :

(19) {N{sg}}
|
London

(20) {N{sg}}
|
François Hollande

Néanmoins, cette caractérisation s'avère trop simple pour la description du fonctionnement des noms propres. Nous rappelons que, bien que certaines théories défendent l'idée de noms propres dénués de sens, nous considérons cette prise de position comme extrême et accordons aux noms propres un sens minimal. En effet, prenons comme exemple le prénom *Mary* ; dans le monde tel qu'il est aujourd'hui, et toutes choses égales par ailleurs, *Mary* renvoie à du féminin. Bien entendu, d'aucuns pourraient rétorquer que l'énonciateur est toujours à même de briser cette convention ; néanmoins, *Mary* renvoyant à du féminin est un usage pour le moins normatif et régulier. Toutefois, il est important de souligner que les noms propres se

différencient des noms communs et des pronoms, en ce qu'ils ne renvoient pas à des propriétés mais ont en discours un référent unique. Nous traduisons cela par l'introduction d'un indice dénotationnel R dans la dénotation des noms propres : cet indice lie, dans la pensée de chaque énonciateur, un nom propre donné à une référence unique, par rapport à un certain contexte d'utilisation. En effet, le contexte d'emploi est déterminant, puisque chaque forme peut désigner plusieurs référents : *London* peut renvoyer à « London, United Kingdom » et à « London, Ontario ». Ainsi, chaque nom propre comporte un certain nombre de propriétés : par exemple, *London* signale une entité unique, un lieu, auxquels il convient d'ajouter la forme phonologique /ˈlʌndn/. Pour rappel, l'échelle présentée en (6) *supra* caractérise les noms propres comme des éléments référentiels purs ; en d'autres termes, ils peuvent se caractériser de façon négative et sont alors considérés comme non fonctionnels. Nous proposons donc, en accord avec les représentations des catégories lexicales et fonctionnelles énoncées en (6) et en (9), la formalisation suivante pour le nom propre *London* :

(21) $\{R_{133}\{sg,place,/ˈlʌndn/\}\}$

London

Le numéro 133 indicé est ici arbitraire et indicatif : 133 lie, dans l'espace mental, la forme *London* à « London, United Kingdom » comme 134 pourrait lier *London* à « London, Ontario ». L'indice dénotationnel R est le paramètre nécessaire pour que les caractéristiques sémantiques profondes du nom propre projettent un {N} – équivalent du DP en GG – et se comporte alors comme un nom commun ou un pronom au niveau syntagmatique et référentiel, puisque ces trois catégories sont en principe interchangeables et susceptibles d'être en coréférence ; c'est également ce R qui différencie les noms propres des noms communs, puisque ces derniers dénotent des classes :

(22) $\{N\{def\}\}$

$\{R_{133}\{sg,place, /ˈlʌndn/\}\}$

London

(23) $\{N\{def,pl\}\}$

$\{N;P\{female, adult, etc.\}\}$

women *(are free to go)*

Notons que nous sommes en mesure de représenter les propriétés syntactico-sémantiques de *women* en (23) uniquement parce que, d'une part, ce nom est ici au pluriel, et fonctionne ainsi en anglais en tant que catégorie syntagmatique et, d'autre part, parce que *women* s'inscrit ici dans un énoncé donné. En effet, *woman* seul ne pourrait être directement représenté, puisque ce nom nécessite un élément de détermination pour fonctionner comme argument : nous aurions donc une manifestation de la linéarité langagière où la tête déterminative requiert un complément nominal. Cette remarque nous mène à l'observation de certaines langues, comme le grec, qui requièrent un article défini placé devant le nom propre pour que celui-ci fonctionne comme un {N} référentiel – dans des énoncés du type *John came* et non pas *They named him John*. Ainsi, de telles langues imposent une projection de ce R qui soit séparée dans la séquence, puisque l'article défini est exigé. Ceci n'est pas le cas pour l'anglais, par exemple, puisque ce R ne projette pas de sérialisation.

Le nom propre prototypique est donc porteur du seul trait catégoriel lexical {N} en GND. Néanmoins, il est des exemples d'emplois de noms propres certes marginaux, mais à ne pas omettre pour autant. Lors de ces emplois, le nom propre est dévié de sa catégorie lexicale {N} pure, en quelque sorte, et se rapproche alors d'un nom commun (exemple (23)) ou d'un verbe (exemple (24)), par exemple :

(23) a. *C'est le Paris que j'aime*.
b. *She is no longer the Susan I used to love*.

(24) a. *We are 'dextering' tonight*. (*Dexter* est le nom d'une série télévisée américaine)
b. *'To argue' c'est se disputer et 'to fight' c'est 'se trintignaniser' c'est ça ?* (communication personnelle avec un ami)

Ici, *Paris* et *Susan* sont alors traités comme des noms communs {N;P} auxquels il faudrait ajouter le trait secondaire { /{prt}}, pour 'partitif' ; en effet, un nom propre dans cet emploi-ci est utile pour faire ressortir un aspect particulier du concept dénoté dans son ensemble. Notons par ailleurs que l'opposition partie/tout qui en découle se manifeste par la présence obligatoire d'une modification restrictive – *c'est le Paris* ***que j'aime*** *; she is no longer the Susan* ***I used to love***. '*Dextering*' et '*se trintignaniser*' sont alors traités comme {P;N{prog}} et {P;N{inf}}, respectivement. Ces exemples-là, au fond, ne posent pas de réel problème : le

nom propre est décatégorisé et transite de {N} à {N;P} ou {P;N}, et analysé comme tel. En revanche, les exemples (25) et (26) présentent deux usages particuliers de noms propres qui restent {N} :

(25) *Cette femme, c'est François Hollande au féminin*.

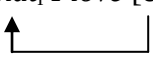
(26) *Magalie Dupré, c'est un joli nom*.

Comment appréhender l'emploi de ces noms propres ici ? En (25), la résistance à la modification restrictive est annihilée, par la présence de *au féminin* ; en (26), *Magalie Dupré* renvoie au patronyme et non pas à l'image de la personne physique qui répond à ce nom. Ces cas précis sont bien entendus spéciaux et ne remettent pas en question les caractéristiques prototypiques des noms propres. Néanmoins, ils mettent en exergue un ensemble de traits syntactico-sémantiques qui ne s'acquièrent que lors de l'activation en contexte de l'élément langagier et qui doivent quand même apparaître au sein de la représentation notationnelle. Ainsi, *François Hollande* en (25) doit être catégorisé comme nécessitant une complémentation – {N/ } – tandis que *Magalie Dupré* est interprété comme étant en référence au métalinguistique grâce au contexte énonciatif. L'enjeu serait de pouvoir faire apparaître ces caractéristiques acquises en discours sur le schéma représentationnel sans pour autant les amalgamer avec les traits syntactico-sémantiques inhérents à l'élément langagier considéré.

3.2. Les syntagmes nominaux définis en Grammaire Notionnelle Dépendancielle

Comme nous l'évoquions en fin d'introduction de ce chapitre, la représentation des syntagmes nominaux – et des syntagmes en général – en GND partage certains points communs avec la représentation X-barre en GG. En effet, les syntagmes ont une architecture asymétrique construite autour d'une tête rectrice, élément à la fois fonctionnel et porteur de traits lexicaux. En GG, ces traits lexicaux sont conservés lors des diverses transformations entre structure profonde et structure de surface ; les transformations opérées permettent de justifier le déplacement des constituants observable en surface et ont pour conséquence directe le postulat des catégories vides :

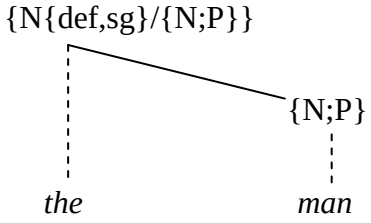
(27) *The Paris that*$_i$ *I love* [e_i] (GG)



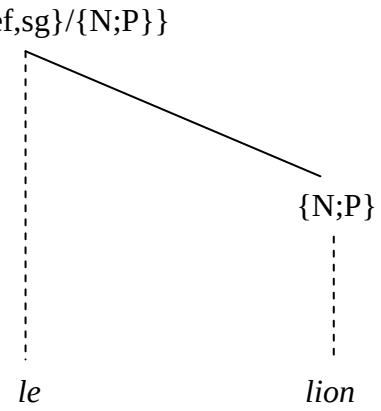
L'élément [e] est ici la trace de l'argument en position objet laissée en structure profonde. Il n'existe pas de telle distinction entre structure de surface et structure profonde en GND, ce qui élimine l'hypothèse des transformations et exclut les catégories vides. Concernant les syntagmes nominaux, et ce comme en GG, l'élément-tête est le déterminant, auquel est assigné l'ensemble des traits syntactico-sémantiques de l'ensemble du syntagme, non pas par transformations successives mais par la sous-jonction de l'élément sous-catégorisé.

Le noyau nominal des syntagmes définis et démonstratifs est un nom commun, dont le comportement syntaxique et sémantique diffère de celui des noms propres (cf. ce chapitre, section 3.1) : la distinction entre dénotation et référence est bien plus marquée concernant les noms communs, ces derniers sont sensibles à la différence entre comptable et non comptable et n'opposent aucune résistance à la modification restrictive. En anglais, *the* n'est jamais pronominalisé et nécessite par conséquent une complémentation nominale toujours exprimée linguistiquement ; ainsi, la représentation catégorielle primaire de l'article défini est notée {N/{N;P}} – nous conservons cette notation catégorielle pour le déterminant *le* en français également : 'N' à gauche de la barre oblique signifie que l'expression considérée est référentielle, la barre '/' signale une relation de complémentation entre le gouverneur à gauche et le dépendant à droite, et {N;P} signifie que le complément est à base nominale. A cette étiquette catégorielle primaire s'ajoutent celles de traits secondaires. Ainsi, les expressions définies signalant à l'allocutaire d'accéder à un référent directement identifiable, le trait {def} est inhérent à la notation syntactico-sémantique de *the*, tandis que le nombre est indiqué par le nom ; concernant l'article *le* en français, le nombre est morphologiquement explicite :

(28)

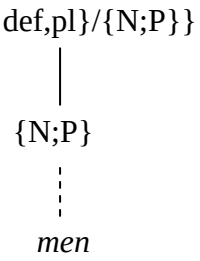


(29)



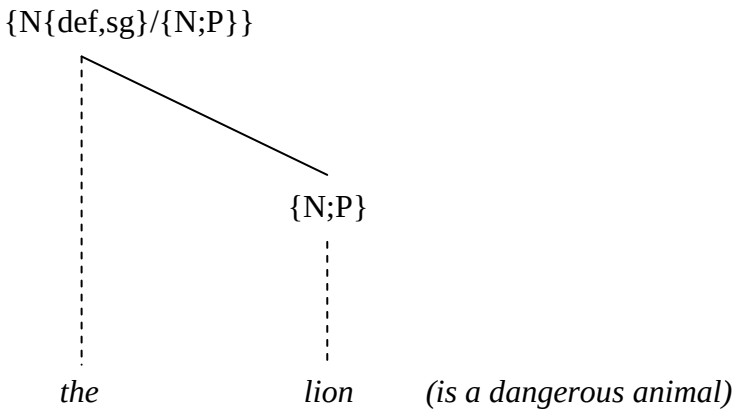
Notons qu'en anglais, le déterminant défini n'est pas forcément exprimé – *men, women, lions, love,* etc. Comme le représente la structure générale en (1), la relation de complémentation {N/{N;P}} est alors manifeste au sein d'un seul item, où l'élément sous-catégorisé est absorbé par la tête :

(30)

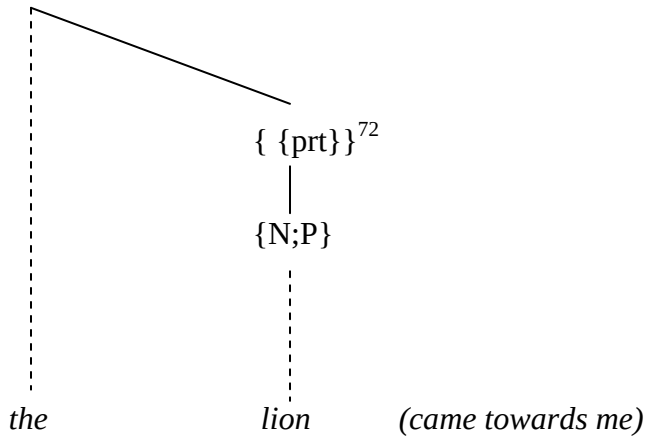


Les traits {def} et {sg} – ou {pl} – appartiennent au système langagier, morphologiquement marqués dans le cas des articles définis en français, le nombre étant signalé par le nom en anglais. En d'autres termes, ces caractéristiques sémantiques sont identifiables hors contexte. Toutefois, il est des propriétés discursives qui ne s'acquièrent et ne sont reconnaissables que lors d'une activation en contexte, comme la lecture spécifique ou générique du SN défini, mais aussi la procédure déictique ou anaphorique réalisée (cf. chapitre IV, section 4.2.2). Comment inscrire ces traits au sein de la représentation notationnelle, sans pour autant les assimiler aux traits notionnels inhérents à la langue ? En ce qui concerne la lecture spécifique ou générique du SN, la GND propose d'attribuer ou non au nom-noyau l'étiquette {prt} pour 'partitif' :

(31) SN à lecture générique



(32) SN à lecture spécifique

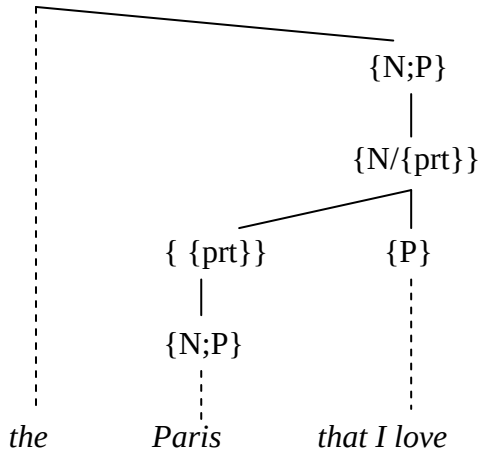


En effet, attribuer le trait 'partitif' au nom-tête permet de rendre compte de l'extraction opérée lors d'une lecture spécifique : *the lion came towards me* renvoie à « this particular lion amongst all the lions », ce qui rejoint l'analyse de Gross (cf. chapitre II, section 2.1) concernant la détermination, toujours considérée comme étant basée sur une extraction. De plus, notons que {prt} serait une étiquette utile pour caractériser un nom propre modifié par une relative restrictive, comme c'est le cas pour *the Paris that I love – Paris* étant ici traité comme un nom commun ; la présence de la relative restrictive permet d'extraire certaines

[72] Nous notons ici que l'espace entre '{' et '{prt}' dans '{ {prt}}' signifie que nous avons affaire à une catégorie fonctionnelle.

caractéristiques de « Paris » parmi l'ensemble des caractéristiques qui pourraient s'appliquer à cette ville :

(33) {N{def,sg}/{N;P}}



Qu'en est-il de la représentation des procédures indexicales potentiellement réalisées ? Puisque nous réfutons la deixis et l'anaphore comme deux phénomènes opposés et les concevons comme les pôles extrêmes d'un continuum discursif, les traits {dei} (pour deixis) et {ana} (pour anaphore) ne sont pas notés {± dei} et {± ana} mais sont des traits discursifs combinatoires, à la manière de N et de P :

(34) Procédures discursives

déictique	déictique et anaphorique	anaphorique
{dei}	{dei : ana}	{ana}

Ainsi, nous pourrions imaginer des cas où la procédure effectuée serait plus déictique qu'anaphorique et vice-versa :

(35)

plus déictique qu'anaphorique	déictique et anaphorique	plus anaphorique que déictique
{dei ; ana}	{dei : ana}	{ana ; dei}

Ainsi, nous pouvons étiqueter les pronoms de troisième personne prototypiques de la manière suivante :

(36)

Notons ici que les pronoms *he* et *she* ne sont pas analysés en GND comme incorporant un {N;P}, et ce pour deux raisons : tout d'abord, *he* et *she* sont tout à fait à même de se substituer à un nom propre ; les noms propres ayant un fonctionnement syntactico-sémantique différent de celui des noms communs, il fallait donc transcrire la possibilité qu'ont *he* et *she* de co-référer avec noms propres et noms communs, d'où le choix de la notation 'neutre' {X} ; ensuite, les traits {masc} et {fem} sont attribués à ce {X} et non pas au {N{def,sg}} – qui serait noté DP en formalisation X-barre – pour des raisons sémantiques. En effet, nous pourrions imaginer des énoncés du type *I saw the doctor yesterday ;* ***she*** *told me that I had the flu / J'ai vu le médecin hier ;* ***elle*** *m'a dit que j'avais la grippe*, où le genre signalé par le pronom choisi dépend du référent visé et non pas du déclencheur d'antécédent *the doctor / le médecin*. Par conséquent, de manière générale, le genre est attribué à un élément 'X', et non pas inscrit en tant que trait catégoriel secondaire de la catégorie primaire (cf. Anderson, 2007 : 256).

3.3. Les syntagmes nominaux démonstratifs en Grammaire Notionnelle Dépendancielle

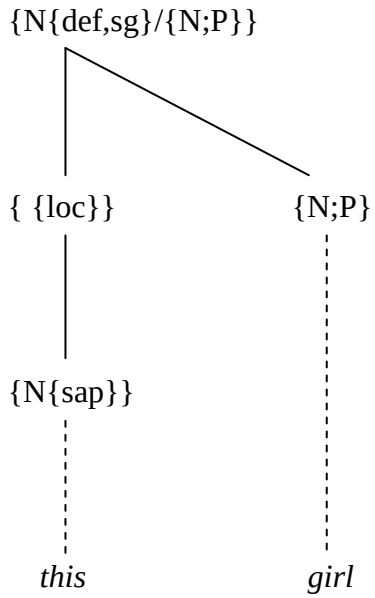
Le traitement des démonstratifs en GND prend comme base celui des déictiques purs *I* et *you*. *I* et *you* ont pour particularité de référer à la source et au destinataire de l'énoncé, respectivement. Ainsi, {ego} est assigné à *I* et {tu} à *you* :

(37)

En GND, *this* et *that* sont essentiellement compris comme des éléments de pointage utiles pour désigner une entité présente dans le contexte situationnel immédiat. Même si la deixis ne se limite pas à la seule référence extralinguistique, le système de notation que nous allons présenter pourra être adapté en fonction des analyses et résultats obtenus lors de notre étude dans les chapitres VII et VIII. En GND, *this* et *that* sont donc analysés comme des indices de

localisation d'un « objet » quelconque par rapport à l'énonciateur ou à son destinataire, ces derniers étant étiquetés {sap}, pour 'speech act participants'. *This* est supposé situer un objet proche du 'sap', tandis que *that* signale une certaine distance. Ainsi, le SN *this/that girl* est représenté de la manière suivante :

(38)



Ce schéma indique que la localisation – opération notée { {loc}}[73] – par rapport à l'un ou l'autre des participants à l'acte énonciatif est interne à *this* et *that*. Ces participants étant des entités que l'on pourrait nommer *I* et *you*, {sap} est assigné à un {N} référentiel. Nous verrons au chapitre VII que la proximité/distance signalée par *this* et *that* est loin de se réduire à la géographie spatiale physique ; cette dernière n'est en fait qu'une conséquence dérivée du fonctionnement des démonstratifs anglais. Néanmoins, nous conservons ce système de notation de base en vue de le préciser en aval de nos analyses.

Ce chapitre s'est attaché à présenter les grandes lignes de la formalisation des éléments langagiers en Grammaire Notionnelle Dépendancielle, système représentatif et descriptif que nous utiliserons aux chapitres VII et VIII pour notre analyse des pronoms démonstratifs anglais et français. Selon nous, la force de la GND est de décrire le fonctionnement des éléments du lexique en fonction de leurs propriétés syntactico-sémantiques, sans toutefois

[73] Nous rappelons que l'espace entre les deux accolades ouvrantes dans '{ {loc}}' signifie que nous avons affaire à une catégorie fonctionnelle.

amalgamer les traits systémiques des mots et les diverses propriétés acquises en discours. En effet, si la formalisation en GND se rapproche de la syntaxe X-barre par le désir de représentation et d'explicitation des énoncés, elle s'en distancie par la réinstanciation du caractère relationnel inhérent aux prédicats et par la complexification du continuum entre 'référentiel' et 'prédicatif'. De plus, la GND se rapproche aussi des théories énonciatives, par certains côtés, puisque la détermination démonstrative s'effectue par rapport à un énonciateur – S_0 en TOPE culiolienne, {ego} ou {sap} en GND – et puisque les têtes syntagmatiques sont avant tout définies par les opérations qu'elles effectuent au sein de l'énoncé. Le chapitre suivant s'attache à décrire les corpus oraux qui servent de base empirique à notre étude des pronoms démonstratifs anglais (chapitre VII) et français (chapitre VIII).

Chapitre VI. Les corpus oraux en linguistique : avantages, limites et choix des données

L'avènement des grands corpus en linguistique remet en question les recherches s'appuyant exclusivement sur des exemples pris à la volée ou construits de toute pièce afin d'illustrer des hypothèses. Désormais, le volume des données authentiques disponible pour le chercheur permet à la fois une conception nouvelle de l'objet scientifique et de ses théories. Toutefois, il convient de distinguer deux approches méthodologiques sur ces corpus : la linguistique de corpus et la linguistique avec (ou sur) corpus. De plus, il convient de garder à l'esprit que ces bases de données certes présentent de nombreux avantages mais ont aussi leurs limites. Les occurrences analysées dans cette thèse (cf. chapitres VII et VIII) proviennent de conversations orales spontanées, ce qui pose en amont une réflexion concernant les différences entre parlé et écrit. Le présent chapitre s'attache à définir ce qu'est le corpus en linguistique et à discuter la place qui peut lui être accordée (section 1) ; l'oral et l'écrit sont également cernés selon leurs ressemblances et leurs divergences (sections 2 et 3) ; enfin, les sections 4 et 5 présentent les projets PFC[74] (Phonologie du Français Contemporain) et PAC[75] (Phonologie de l'Anglais Contemporain), d'où sont extraites les données analysées dans les chapitres VII et VIII, et offrent des perspectives pour l'annotation micro- et macro-syntaxique des textes oraux en contexte de conversations spontanées.

1. Quelle place accorder aux corpus en linguistique ?

1.1. Définition du corpus

Un corpus tel que nous l'entendons est un ensemble de textes authentiques, oraux ou écrits, présélectionnés selon l'usage auquel ils sont destinés et représentatifs d'une ou plusieurs langues. Un corpus n'est pas un simple amas de données en ce qu'il est structuré, ne serait-ce que par le choix des textes qui le composent. Afin d'être utile à une communauté de chercheurs et non pas au seul individu qui l'a monté, un corpus est idéalement annoté à des fins spécifiques.

[74] Cf. Durand, Laks & Lyche (2002, 2009).

[75] Cf. Carr, Durand & Pukli (2004).

1.2. Les enjeux épistémologiques de l'utilisation des corpus en linguistique

Longtemps, une conception trop radicale du dualisme saussurien – interprété à partir du fameux *Cours* posthume – a fait de la linguistique une science de la langue et non de la parole. Ainsi, les énoncés idiosyncrasiques produits par le sujet parlant ne faisaient pas partie du domaine d'étude, mais relevaient de sciences hybrides et connexes telles que la psycholinguistique ou la sociolinguistique. Les analyses linguistiques s'effectuaient alors hors corpus, car les 'bruits' qui jalonnent les données authentiques – hésitations, reprises, amorces, syntaxe non canonique, etc. – gênaient le théoricien. Le représentant de la linguistique hors corpus est bien entendu Chomsky et sa GG, dont les postulats et les règles trop puissantes s'accordent difficilement à l'étude de la performance langagière. Aujourd'hui, la situation est tout autre : il existe de moins en moins de linguistes travaillant sans corpus. Nous passons donc de l'étude 'sans corpus' *vs* 'avec corpus', à la linguistique 'sur corpus' *vs* 'de corpus'. Cette différence relève de deux conceptions du corpus et de la place que celui-ci occupe au sein de la méthodologie employée. En effet, s'il est indéniable qu'un corpus est un observable, il peut être envisagé comme le reflet d'une théorie *a priori* ou comme un observé dynamique qui permet de décrire dans un premier temps puis d'élaborer des modèles *a posteriori*. Ces deux pôles correspondent aux méthodes déductive et inductive de la formation des hypothèses et de la construction des modèles, dualité que Tognini-Bonnelli (2001) exprime par 'corpus-based' *vs* 'corpus-driven'. Toutefois, cette distinction ne peut être aussi tranchée qu'elle le paraît, selon nous ; si la déduction radicale est à proscrire – au risque de tomber dans l'écueil de chercher à tout prix les exemples qui confortent la théorie *a priori* – l'induction pure semble aussi difficilement concevable que ne l'est le vide absolu : l'intuition[76] est inévitable et préexiste à la recherche de tout phénomène.

1.3. Les avantages et les limites de l'utilisation des corpus en linguistique

Nous pouvons reconnaître de nombreux avantages à l'utilisation des corpus. Tout d'abord, les corpus permettent d'analyser des données réelles ; ce point pourrait paraître évident, mais la description d'énoncés authentiques contraint le chercheur à considérer ces fameux 'bruits' (cf. ce chapitre, section 1.2), soigneusement évités si les études se basent sur

[76] L'intuition peut être définie comme étant une hypothèse qui se présente à l'esprit avec la clarté d'une évidence, qui servira alors de fondement au raisonnement scientifique. L'intuition chez les chomskyens correspond plutôt à des jugements (grammaticalié, synonymie, etc.).

des exemples sélectionnés, voire forgés. Force est alors de constater que la langue n'est pas un carcan prescriptif mais un observable dynamique[77]. Dès lors, la finalité n'est pas de s'obstiner à prouver la véracité d'une théorie possiblement mal adaptée à la langue telle que nous l'employons au quotidien, mais de dégager des normes parmi la variation observée. Ces normes, justement, font le lien entre langue et parole, entre système et usage, entre règles et variation. Enfin, les corpus autorisent une réflexivité de la recherche : c'est en confrontant nos intuitions théoriques aux données authentiques que nous pouvons confirmer ou infirmer nos hypothèses.

N'en déplaise à Leibniz – Pangloss pour les férus de Voltaire – « tout [n']est [pas] pour le mieux dans le meilleur des mondes possibles » : les corpus ont leurs limites. D'un côté, il est peu probable de trouver toutes les occurrences rencontrées 'dans la vraie vie' au sein d'un seul et même corpus, voire dans plusieurs : par exemple, le pronom anglais *them* en tant que déterminant démonstratif – comme dans *Them two had better watch it!* – n'apparaît pas dans notre corpus d'étude ; néanmoins, cet emploi est bel et bien attesté dans certains registres ou certaines variétés de l'anglais qui restent à circonscrire. A l'inverse, dure serait la tâche du linguiste désirant rendre compte de chaque particularité langagière observée. La difficulté est de trouver la juste mesure entre le désir de confirmer nos intuitions par la confrontation aux corpus et la nécessité de ne pas exclure les occurrences qui pourraient contredire nos hypothèses. Ce compromis rejoint celui à atteindre entre approche quantitative et approche qualitative. En effet, il est légitime de penser que l'utilisation des corpus tendrait vers des études quantitatives, au risque de négliger le qualitatif : avec de grands corpus, certains exemples particulièrement intéressants à considérer pourraient être noyés. Une issue serait peut-être de privilégier le qualitatif dans un premier temps, avant d'étendre les recherches sur des corpus plus larges afin de vérifier la tendance.

2. L'oral et l'écrit

2.1. L'oral et l'écrit, deux facettes d'une même pièce

L'oral et l'écrit peuvent sembler bien différents, tant dans leur matérialisation linguistique que dans leur conception. En effet, s'il est évident que le médium employé les oppose – l'oral est régi par les sons, l'écrit par le signe graphique –, il en va de même pour la

[77] Cf. ce chapitre, section 2 *infra*.

portée communicative de l'énoncé. Faut-il pour autant séparer ces deux modes de production ? A l'origine de ce questionnement réside sans nul doute la dichotomie entre langue et parole introduite par Saussure, mais aussi la relation qu'il envisageait entre oral et écrit :

> « Langue et écriture sont deux systèmes de signes distincts : l'unique raison d'être du second est de représenter le premier ; l'objet linguistique n'est pas défini par la combinaison du mot écrit et du mot parlé ; ce dernier constitue à lui seul son objet. Mais le mot écrit se mêle si intimement au mot parlé dont il est l'image, qu'il finit par usurper son rôle principal ; on en vient à donner autant et plus d'importance à la représentation du signe vocal qu'à ce signe lui-même. C'est comme si l'on croyait que pour connaître quelqu'un, il vaut mieux regarder sa photographie que son visage. » (Saussure, 1916[1972] : 45).

Ce point de vue est partagé par Bloomfield, qui l'expose de manière plus radicale encore :

> « Writing is not language but merely a way of recording language by means of visible marks. » (1933: 21).

Cette primauté de l'oral sur l'écrit s'explique par nombre d'arguments phylogénétiques et ontogénétiques (cf. Durand, 2000) qui confèrent à l'oral une vertu 'naturelle', contrairement à celle, artificielle, de l'écrit. Cependant, la linguistique moderne naquit à contre-courant de la tradition normative pré-structuralisme, qui ne considérait que les *belles lettres* comme dignes d'intérêt. Il faut dire que les données écrites étaient facilement collectées et étudiées, à l'ère où le matériel d'enregistrement n'avait pas encore vu le jour.

Néanmoins, bien que l'oral soit premier, du point de vue du développement à l'échelle humaine et individuelle, il n'en souffre pas moins d'une certaine dévalorisation par rapport à l'écrit. En effet, contrairement à ce dernier, qui peut être réfléchi, révisé, corrigé, l'oral a la particularité d'être ineffaçable : « ce qui est dit est dit ». Ainsi, l'oral spontané est ponctué d'hésitations, de reprises, d'inachèvements. Le concept même de phrase peut sembler peu pertinent, tant il est difficile de trouver, au sein d'une conversation, la suite canonique « syntagme nominal + syntagme verbal » (cf. Rossi-Gensane, 2010 à ce sujet). Il en résulte que la phrase est de plus en plus délaissée par les spécialistes de l'oral, au profit d'unités (micro-)syntaxiques* minimales nommées 'noyau'* (Blanche-Benveniste *et al.*, 1990) ou

‘clause’* (Berrendonner, 1990). Ces noyaux ou ces clauses peuvent alors se combiner pour former des unités plus larges, appelées ‘regroupement’* ou ‘période’* ; ces unités dépassent le seuil de rection d’un seul verbe et relèvent alors de la macro-syntaxe*. Toujours est-il que l’oral n’en est pas pour autant moins structuré que l’écrit. Halliday (1985), pour ne citer que lui, souligne que l’écrit et l’oral sont deux systèmes à part entière, régis par des règles, avec une complexité qui leur est propre. Seulement, un tel jugement, s’il n’est pas nuancé, sous-entend que ces deux modes de production sont diamétralement opposés. De plus, cela supposerait qu’il n’existe qu’un seul oral et qu’un seul écrit. Or, il n’en est rien.

2.2. Le genre et la visée communicative

Selon nous, l’écrit et l’oral ne peuvent être appréhendés comme deux systèmes dichotomiques. En effet, qu’entendons-nous par ces termes ? Une œuvre de littérature classique pour le premier ? Une conversation libre pour le second ? Que faire alors des articles de presse, des lettres entre amis, d’un discours politique ou bien du cours monologique d’un professeur ? Et, de nos jours, comment classer les emails, textos et autres twitts ? De toute évidence, le rapport entre écrit et oral semble plus complexe qu’une simple division manichéenne.

Une des différences majeures entre le texte écrit et le texte oral, qui pourrait sembler évidente, est la distance, ou, au contraire, l’adéquation spatio-temporelle entre le destinateur et le destinataire. En effet, seule la représentation d’un échange conversationnel au sein d’un texte écrit serait en mesure de manifester une concordance spatio-temporelle entre les actants. Un autre critère qui pourrait sembler discriminatoire entre l’écrit et l’oral est celui de la réciprocité : si les allocutaires d’une conversation participent l’un et l’autre à l’échange, il n’en est pas de même pour le scripteur et le lecteur, ce dernier n’étant que le bénéficiaire, en quelque sorte, du discours représenté. Seulement, tous les textes oraux ne sont pas des conversations, et la réciprocité est probablement bien moindre entre un professeur et ses élèves lors d’un cours académique, et même inexistante lors d’une allocution politique. L’œuvre majeure et pionnière qui sonne le glas de cette vision bipartite de l’écrit et de l’oral est sans nul doute l’étude de corpus effectuée par Biber (1988 ; 1995), ayant pour objectif de définir, de manière inductive, une typologie des textes selon différents traits linguistiques et communicatifs qui les caractérisent. Voici, à titre d’exemple, les quatre principaux registres que Biber a sélectionnés pour le *Longman Student Grammar of Spoken and Written English* (2002 : 13) :

	Conversations	Fiction	News	Academics
Mode	Spoken	Written	Written	Written
Interactiveness and real-time production	Yes	Restricted to fictional dialog	No	No
Shared situation	Yes	No	No	No
Main communicative purpose/content	Personal communication	Pleasure reading	Information or evaluation	Argumentation or explanation
Audience	Individual	Wide-public	Wide-public	Specialist

TABLEAU 4 : LES QUATRE PRINCIPAUX REGISTRES SELON BIBER (2002 : 13)

Il en résulte – entre autres choses, bien sûr – que de la variation, notamment morphosyntaxique, existe en fonction du genre pris en compte, mais est également notoire au sein d'un même genre. Ainsi, si l'oral spontané peut être limité sur le plan lexical, privilégie une structure paratactique et nécessite un grand nombre de calculs inférentiels de la part de l'allocutaire, il n'en est pas de même pour l'écrit publié très normatif, qui se caractérise, notamment, par une densité lexicale accrue et des relations de cohérence morphologiquement marquées. Cependant, le monologue oratoire d'un politicien partage sans doute certaines caractéristiques morphosyntaxiques de l'écrit, tandis qu'un e-mail envoyé à un ami rappelle probablement, par de nombreux aspects, les propriétés de l'oral spontané. De même, la proximité communicative est différente selon le genre : élevée dans une conversation familière entre proches, moindre lors du cours d'un professeur à ses élèves, nulle ou quasi nulle lors d'un discours oratoire. Cappeau (2001), Deulofeu (2000) et Gadet (2003), par exemple, perpétuent cette quête des genres, en ce qui concerne le français. De cette grande variété découle une conception graduelle de l'écrit et de l'oral (Blanche-Benveniste, 1997 ; McCarthy, 2001), sans doute préférable à une opposition binaire entre ces deux modes de production :

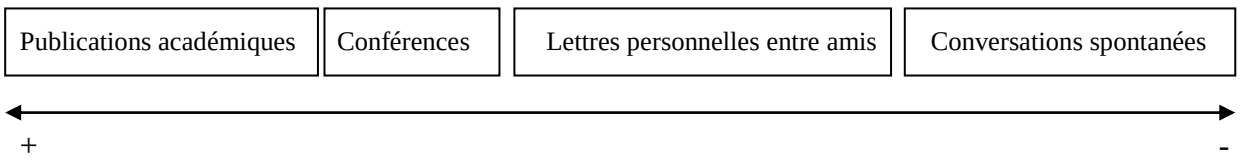


FIGURE 23 : LA COMPLEXITE GRAMMATICALE SELON LE GENRE : UN CONTINUUM

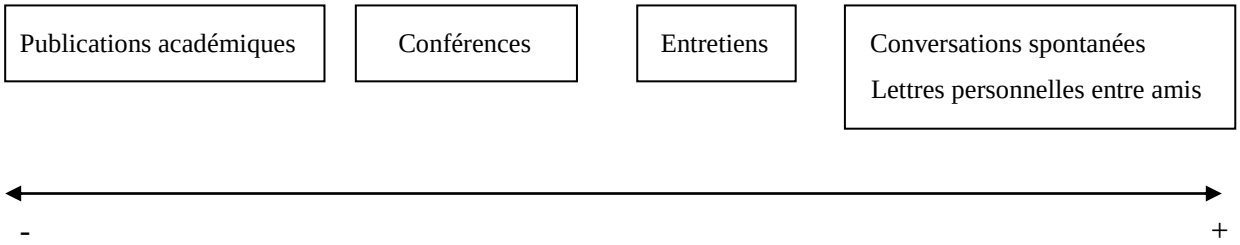


FIGURE 24 : L'IMPLICATION INTERPERSONNELLE SELON LE GENRE : UN CONTINUUM

Nous voyons alors l'importance d'un modèle interactionnel de la communication : que les participants soient en adéquation spatio-temporelle ou non, il y a toujours un émetteur et (au moins) un bénéficiaire ; ainsi, nous devons prendre en compte l'énonciateur, le destinataire, mais aussi le calcul que chacun d'eux effectue afin d'être compris par l'autre. La différence selon le genre se manifeste sans nul doute dans les expressions langagières employées : il nous semble évident que, lors d'une conversation spontanée en face à face, le contexte, notamment situationnel, joue un rôle plus important que lors de la lecture d'un texte écrit quelconque ; nous pouvons donc supposer que le facteur contextuel comblerait le déficit lexical dans le fonctionnement référentiel de l'oral conversationnel. Ce dernier point nous amène à une autre réflexion et nous fait comprendre pourquoi 'contexte' et 'co-texte' sont parfois amalgamés (cf. chapitre IV, section 2.3). Ce qui pourrait engendrer cette confusion est qu'il est impossible à l'oral de revenir sur le texte, à cause de sa non-persistance[78]. Ainsi, dès lors que nous souhaitons revenir sur ce qui a été dit, il faut effectuer un retour mental dans l'univers du discours, et non pas dans le texte lui-même. Bien entendu, cette non-distinction entre 'contexte' et 'co-texte' résulte d'un paramètre supplémentaire : l'étude de l'oral se base sur des transcriptions. Par conséquent, le linguiste est forcément influencé par le médium écrit. La solution serait de se détacher de ce support, autant que faire se peut ; ne pas se concentrer sur la matière linguistique mais sur les objets sémantiques et les phénomènes pragmatiques induits par le choix des structures langagières.

[78] Notons ici qu'il est possible de renvoyer au texte lui-même, comme l'illustrent les exemples suivants : *Comment tu écris **ça/ce mot/rhododendron** ? Que veut dire **ce regard** ?* ; néanmoins, cette référence au texte résulte de l'emploi en mention des expressions langagières considérées.

3. Structure du texte, structure du discours

Selon nous, comme nous venons de l'évoquer, il est important de considérer la structure du discours en fonction de la matière textuelle, et non pas la structure du texte *per se*. L'articulation du discours est l'objet de bon nombre d'études (cf. entre autres Adam, 2005 ; Brown & Yule, 1983 ; Roulet, Filliettaz, Grobet & Burger, 2001), mais nous insistons sur la distinction entre 'discours' et 'texte', une fois de plus, en vue de bien séparer les niveaux, les mécanismes régissant chacun d'eux et les paramètres qui relient ces deux notions. Afin de cerner ce qu'est la structure discursive, il suffit d'inscrire les notions abordées au chapitre IV – deixis et anaphore notamment – dans une optique d'organisation cognitive. Ainsi, la deixis serait un moyen disponible pour l'énonciateur de diriger l'attention de son allocutaire vers un nouvel objet-de-discours tandis que l'anaphore permettrait de continuer la référence d'un objet sémantique quelconque hautement saillant dans la conscience de l'allocutaire. Cet objet hautement saillant est appelé 'topique'* et se caractérise de la manière suivante :

> « Le topique se définit comme une information (un référent ou un prédicat) identifiable et présente à la conscience des interlocuteurs, qui constitue, pour chaque acte[79], l'information la plus immédiatement pertinente liée par une relation d'à propos avec l'information activée par cet acte. » (Grobet, 2002 : 96)

Bien entendu, l'enjeu de notre étude est de cerner les paramètres contextuels qui interviennent dans ces procédures : contextes situationnel, discursif, savoirs encyclopédique et partagé. Schématiquement parlant, nous serions encline à prédire que la deixis se trouve à l'initiale d'un ensemble cohésif de discours, au contraire de l'anaphore qui tisserait les liens internes à l'ensemble considéré : la deixis serait aux articulations du squelette anaphorique. La relation entre deixis, anaphore, structure discursive et structures langagières fait écho aux échelles d'Ariel et Gundel *et al*. (cf. chapitre IV, section 2.5) : si les expressions définies codent un référent à faible accessibilité (Ariel, 1988) ou uniquement identifiable (Gundel *et al*., 1993),

[79] L'acte est caractérisé de la manière suivante : « L'acte est défini comme la plus petite unité délimitée de part et d'autre par un passage en mémoire discursive, dans le sens de Berrendonner (1993). […] Le passage en mémoire discursive est marqué en particulier par la possibilité d'utiliser indifféremment comme anaphore un pronom ou une expression définie pour indiquer la coréférence. » (Roulet, 1999 : 210 - 211). Défini ainsi, l'acte correspond à ce que Berrendonner appelle 'énonciation'.

elles auraient tendance à instancier un point d'articulation du discours, alors que les pronoms de troisième personne seraient le support du squelette anaphorique. Une hiérarchie possible des expressions langagières dans l'organisation du discours est proposée par Dik (1997) et notamment reprise par Cornish (1998) ; seulement, cette structure s'applique aux textes écrits mais s'avère insuffisante pour l'oral conversationnel, en raison de la variation lexicale entre les deux genres et de l'intervention prépondérante du contexte dans la communication verbale. Dik (1997 : 215-228) propose un modèle à *n*-maillons pour structurer le discours :

- Le premier maillon est la position remplie par le déclencheur d'antécédent (cf. chapitre IV, section 2.5). Ce premier maillon introduit un « Topique Nouveau »[80] ou réinstalle un « Topique Ancien » sous la forme d'un nom propre, d'un SN indéfini, voire d'un SN défini, mais jamais sous la forme d'un clitique ou d'une ellipse ;
- Le deuxième maillon confirme le « Topique Nouveau », nommé alors « Topique Repris », qui sera par la suite un « Topique Donné ». Ce deuxième maillon est instancié par les démonstratifs sous toutes leurs formes, parfois par des pronoms personnels ;
- Le troisième maillon occupe toutes les positions après la seconde. Les *n*-maillons sont instanciés par des clitiques ou des ellipses, tant que le référent évoqué est un « Topique Donné ».

Pour illustration, inventons une suite calquée sur un début de conte pour enfants classique, comme l'exemple (10) du chapitre IV :

(1) *Il était une fois* ***un homme*** *qui vivait avec ses trois fils.* ***Cet homme*** *était bon et aimé de tous.* ***Il*** *sacrifiait sa vie pour que ses enfants ne manquent de rien.*

Selon le modèle de Dik, le SN indéfini *un homme* serait en position de premier maillon et introduirait un référent nouveau, susceptible par la suite de devenir topique ; le SN démonstratif *Cet homme* instancierait le deuxième maillon et serait utile pour conférer à « l'homme en question » un statut cognitif suffisamment saillant pour pouvoir être continué par l'emploi du clitique *Il*, en position de *n*-maillon. Comme nous pouvons nous en douter,

[80] Le topique étant défini comme « une information identifiable et présente à la conscience des interlocuteurs » (Grobet, 2002 : 96), l'association des termes 'topique' et 'nouveau' ne nous semble pas judicieuse ; 'topique à venir' serait une appellation plus appropriée.

l'oral conversationnel bouleverse les codes formatés du genre textuel d'où pourrait être extrait cet exemple, et la chaîne de Dik s'y trouve difficilement applicable :

(2) CM: *Donc euh, tu fais quel euh, rôle au théâtre ?*
LH: *Je fais euh, encore un rôle de, de folle, Consuela (xxx).*
CM: *Et, c'est quoi ?*
LH: *C'est, en fait, tu, tu connais l'histoire non ?*
CM: *Non pas du tout.*
LH: *Oh c'est, c'est hyper compliqué en fait, ça s'appelle 'Treize à table', et euh, et en fait, tout le long de la pièce, euh, enfin c'est une bourgeoise qui invite des gens, mais elle va s'apercevoir qu'ils sont treize, et tout le long de la pièce, ils vont essayer d'inviter des gens, mais ils vont, pourront pas et ça va être tout le long de la pièce, elle saura qu'ils seront treize de toute façon. Ça va la tenir jusqu'au bout. Et il y a plein de personnages qui vont arriver, repartir, et tout ça.* (exemple issu du corpus PFC Domfront, code locuteur 61acm1, conversation libre)

Ce simple extrait de conversation spontanée révèle les difficultés relatives à l'analyse discursive en termes de chaînes référentielles, notamment par l'emploi massif de pronoms dits 'neutres' – *c', ça*. Par ailleurs, le même constat s'effectue en anglais, où le clitique *it* et le démonstratif *that* supplantent l'emploi d'expressions lexicales :

(3) MD: *Oh no I'm just reading a book about er, by Ben Elton, have you, heard of it, 'Dead Famous'?*
LB: *I've, I've heard of Ben Elton, yeah.* <MD1: *Right.*>
MD: *It's, and it's like about, you know er, Big Brother.*
LB: *Yeah.*
MD: *It's like about that, I mean, he's like, s/ some of the ideas are from, Big Brother and it's, but there's like a murder in there.* (laughs) *But it's a bit, I was just thinking, it's, weird, because I'm watching Celebrity Big Brother at the moment,* <LB1: *Oh yeah.> and so like I'm thinking like, have you seen it?*
LB: *That's boring.*
MD: *Don't like it.*
LB: (telephone ringing) *I do, I watch it because it's compulsive viewing but, I do think it's awfully boring because there's, they're not letting themselves go at all* <MD1: *No, no.> it's not like the ordinary Big Brother, where, you don't know these people that are*

in there and you get to know the personalities as, as the program goes on. (exemple issu du corpus PAC Lancashire, code locuteur elalb1, conversation libre)

L'utilisation massive de pronoms neutres à l'oral conversationnel se comprend aisément : les interlocuteurs produisant l'acte communicatif au même endroit et au même moment, ils partagent par conséquent le même contexte situationnel immédiat ; de plus, la construction du discours s'effectue de manière conjointe, les participants agissant en interaction. Ainsi, les locuteurs n'ont pas à produire l'effort de choisir l'expression lexicale qui convienne à leur volonté communicative : un pronom neutre est un moyen économique de renvoyer à des objets-de-discours que le destinataire peut identifier grâce à l'intervention des différents paramètres contextuels. Reste à savoir comment ces paramètres agissent et quelle est leur contribution à la structure du discours.

4. Les projets PFC et PAC

Les données orales analysées dans les chapitres VII et VIII proviennent des projets PFC (Phonologie du Français Contemporain) et PAC (Phonologie de l'Anglais Contemporain)[81], respectivement codirigés, d'une part, par Marie-Hélène Côté (Université d'Ottawa), Jacques Durand (ERSS, Université Toulouse II-Le Mirail), Bernard Laks (MoDyCo, Université Paris X) et Chantal Lyche (Universités d'Oslo et de Tromsø), et, d'autre part, par Philip Carr (EMMA, Université Montpellier III), Jacques Durand et Anne Przewozny (ERSS, Université Toulouse II-Le Mirail). Bien entendu, ces projets ne sont pas les seuls à collecter des données orales. Citons par exemple, pour le français, le corpus du GARS (Groupe Aixois de Recherche en Syntaxe), initialement dirigé par Claire Blanche-Benveniste, et dont la partie numérisée est appelée CorpAix, mais aussi les différents corpus organisés autour de l'équipe DELIC (Description Linguistique Informatisée sur Corpus) pilotée par Jean Veronis, appartenant désormais à l'équipe TALEP (Traitement Automatique du Langage Ecrit et Parlé, du Laboratoire d'Informatique Fondamentale de Marseille) : le CRFP (Corpus de Référence du Français Parlé, monté par l'équipe DELIC et financé par la Délégration Générale à la Langue Française et aux Langues de France (DGLFLF)) et le C-ORAL-ROM (Corpus de référence pour les langues ROManes, corpus multilingue monté par

[81] Le projet PAC est né à la suite du projet PFC, lancé en 1999. Pour plus de renseignements concernant les participants aux projets, mais aussi les points d'enquêtes et les locuteurs, nous renvoyons le lecteur aux sites internet http://projet-pfc.net et http://w3.pac.univ-tlse2.fr.

l'équipe DELIC, le laboratoire LABLITA (Université de Florence, Italie), le centre de linguistique de l'Université de Lisbonne et l'Université Autonome de Madrid). En ce qui concerne l'anglais, il convient de signaler le Brown Corpus (anglais américain, collecté dans les années soixante) et son équivalent britannique, le Lancaster Corpus, fondé dans les années quatre-vingt ; citons également le TRAINS Corpus, créé dans les années quatre-vingt-dix et centré sur des dialogues orientés par la tâche, ainsi que le Survey of English Usage (1959) qui s'est poursuivi par le London-Lund Corpus et le projet ICE (cf. ce chapitre, section 5.2.1 *infra*). Toutefois, la volonté originelle des deux projets-frères PFC et PAC est d'étudier la variation et l'hétérogénéité phonologiques des espaces francophone et anglophone par la constitution et l'analyse de vastes corpus oraux. Dans la plus pure tradition labovienne (1966), le protocole méthodologique de ces projets est basé sur des enregistrements de locuteurs d'âge et de provenance sociale variés, confrontés à des productions orales contrôlées, comme la lecture, et relâchées, comme la conversation spontanée.

4.1. La méthodologie des projets PFC et PAC

4.1.1. Les points d'enquêtes et le choix des locuteurs

A ce jour, le projet PFC rassemble plus de soixante points d'enquêtes à travers l'espace francophone, tandis que le projet PAC en compte une trentaine. Idéalement, chaque enquête réalisée comprend une douzaine de locuteurs enregistrés, si possible six hommes et six femmes, représentatifs de trois tranches d'âges : junior, de trente à cinquante ans, senior.

4.1.2. Les tâches effectuées

Après signature d'une feuille de consentement, le locuteur doit s'acquitter de quatre tâches : la lecture du texte PFC ou du texte PAC, la lecture d'une liste de mots spécifiques (deux listes pour le projet PAC), une conversation guidée par l'enquêteur et une conversation libre entre le locuteur et l'un de ses proches. L'ensemble de ces tâches permet de couvrir un panel de registres, du plus formel – la lecture – au plus relâché – la conversation spontanée. Les projets PFC et PAC étant initialement destinés aux phonologues, les textes et les listes de mots du protocole d'enquête sont bien entendu voués à refléter la variation phonologique et phonétique qui existe au sein des espaces francophone et anglophone. Ainsi, un système de codage alphanumérique est mis en place afin d'étudier la (non-)présence du schwa dans certaines variétés du français, du /r/ de liaison dans diverses variétés de l'anglais et de la liaison dans les deux langues.

4.1.3. La méthode d'annotation

Les fichiers sonores enregistrés au format WAV sont traités grâce au logiciel Praat (cf. Boersma & Weenick (2009) pour une description de la version 5.1.05 du logiciel) qui permet d'aligner son et transcription afin de réaliser des analyses acoustiques fines. Le niveau zéro du traitement des fichiers sonores est la transcription orthographique, qui permettra alors de servir de base aux annotations phonologiques concernant le schwa et la liaison pour le français et le /r/ de liaison pour l'anglais. Par ailleurs, notons ici que les composantes phonologiques telles que la courbe prosodique ou la qualité de la voix ne sont pas annotées ; chaque chercheur ou équipe de recherche peut en revanche décrire les phénomènes linguistiques souhaités grâce au logiciel Praat.

4.1.3.1. Les conventions de transcription

La volonté des projets PFC et PAC est que les transcriptions soient le plus fidèlement représentatives des données orales sans présager d'analyses acoustiques fines. Il s'agit donc de trouver un compromis entre une transcription phonétique qui serait au plus proche de l'oralité et une transcription 'littéraire' répondant aux codes de l'écrit normatif. Ainsi, la transcription choisie est de type 'standard', où les lemmes sont écrits dans leur intégralité mais où les mots non prononcés ne sont pas reproduits. Par exemple, [i s r a p a l a] sera transcrit *il sera pas là* et non pas *il ne sera pas là* ou encore *i s'ra pas là*. La ponctuation permise se limite au point et à la virgule, suivant les pauses plus ou moins longues effectuées par le locuteur, ainsi qu'au point d'interrogation. Concernant les tours de parole et les troncations de mots, les conventions de transcription s'inspirent de celles établies par les équipes du GARS (Groupe Aixois de Recherche en Syntaxe, Université d'Aix-Marseille) et de VALIBEL (Université catholique de Louvain, Belgique)[82]. Ainsi, chaque locuteur est anonymisé et représenté par ses initiales ; ces lettres sont précédées du code du département où l'enquête s'est déroulée – par exemple, 61 pour le corpus PFC Domfront, ela pour le corpus PAC Lancashire – et suivies d'un chiffre ordinal selon le nombre de locuteurs ayant les mêmes initiales – par exemple, 13brp2 signifie que deux locuteurs différents portent les mêmes initiales RP au sein de l'enquête à Marseille ; la lettre 'b' indique que cette enquête est

[82] Pour une description en détail des conventions de transcription des projets PFC et PAC, voir le Bulletin PFC n°1 disponible à l'adresse web http://projet-pfc.net/bulletins-et-colloques/cat_view/918-bulletins-pfc/919-bulletin-pfc-nd1.html et le descriptif PAC fourni à l'adresse web http://w3.pac.univ-tlse2.fr/conventions.htm.

la deuxième réalisée dans cette région. L'enquêteur est représenté par la lettre 'E' en français, 'I' en anglais – 'I' pour 'interviewer'. Les chevauchements sont encadrés de chevrons, comme l'illustre l'exemple suivant :

(4) BM: *Un, un petit peu plus oui* <CM: *Cinquante-trois ?*>, *cinquante-deux.* (exemple issu du corpus PFC Domfront, code locuteur 61abm1, conversation libre)

Les mots tronqués sont signalés par une barre oblique – *il m'a pro/ pro/ mis de venir* – tandis que les répétitions sont toutes comptabilisées et séparées par une virgule – *il, il, il est venu*. Les bruits sont signalés par des parenthèses – *(rire)* ; *(silence)* ; *(background voices)* – et des croix remplacent les syllabes non décryptées – *Je fais euh, encore un rôle de, de folle, Consuela (XXX).* Enfin, le discours rapporté est indiqué par l'encadrement de guillemets simples – *il m'a dit 'tu es complètement idiot d'avoir accepté ce boulot', et je crois bien qu'au fond il avait raison*.

4.1.3.2. Le codage phonologique

La transcription orthographique sert de base aux codages du schwa et de la liaison pour le français et du /r/ de liaison pour l'anglais. Ces codages sont alphanumériques, sous forme de champs instanciés par une valeur[83] :

(5) Exemple de codage du schwa : *Mon père0411 était arrivé*, où le premier chiffre indique la présence ou l'absence du schwa, le deuxième chiffre la position initiale, finale ou interne de ce schwa, le troisième chiffre le contexte phonologique gauche et le quatrième chiffre le contexte phonologique droit ;

(6) Exemple de codage pour la liaison : *Un11n an ou deux*, où le premier chiffre indique le caractère monosyllabique ou polysyllabique du mot liaisonnant, le deuxième chiffre la présence ou l'absence de la liaison et le troisième champ par une lettre – ici 'n' – qualifie la consonne réalisée ;

(7) Exemple de codage du /r/ de liaison : *Before021 our nine a.m. meeting*, où le premier chiffre indique la présence ou l'absence du /r/, le deuxième chiffre le caractère

[83] Les questions phonologiques n'étant pas au cœur de notre étude, nous ne décrivons pas en détail les différents codages appliqués au sein des projets PFC et PAC. Néanmoins, l'esquisse de ces codages numériques est nécessaire à la compréhension des difficultés rencontrées lors de la constitution de notre corpus, qui seront développées en section 5.1 *infra*.

monosyllabique ou polysyllabique du mot liaisonnant et le troisième chiffre le caractère monosyllabique ou polysyllabique du mot suivant.

Un tel codage numérique se prête à l'analyse intuitive, sans fournir de préconçu théorique ni s'aventurer dans une pré-analyse phonétique. Une fois codées, les données traitées sont rendues accessibles à l'utilisateur par la plateforme Dolmen (développée par Julien Eychenne (Hankuk University of Foreign Studies, Corée du Sud) et disponible librement à l'adresse http://julieneychenne.info/dolmen), spécialement créée à cet effet. Ainsi, l'ensemble des données recueillies, transcrites puis annotées, permet ensuite à une large communauté de phonologues d'effectuer des analyses fines et d'observer la variation au sein des espaces francophone et anglophone.

5. Les projets PFC et PAC 'au-delà de la phonologie'

Bien qu'initialement destinés aux phonologues, les projets PFC et PAC touchent des linguistes dont les intérêts de recherche ne se limitent pas à la phonologie[84]. Une des raisons principales pour laquelle nombre de chercheurs s'intéressent à ces corpus est que ces derniers sont disponibles en ligne et exploitables. Le problème qui se pose est bien sûr celui du traitement automatique de ces données : si des phénomènes phonologiques comme le schwa et la liaison sont annotés, il n'en va pas de même pour la syntaxe, la sémantique ou la structure du discours[85] ; à l'heure actuelle, la seule entreprise d'automatisation à ces différents niveaux est un découpage morpho-syntaxique par TreeTagger (cf. Schmid, 1995), mais le moteur de recherche du site PFC ne permet pas d'effectuer des requêtes directement à partir des catégories obtenues. Nous-même avons été confrontée à ce problème lors de la constitution de notre corpus de démonstratifs. En effet, la seule solution était de relever manuellement chaque occurrence au sein des conversations transcrites, ce qui est une méthode extrêmement coûteuse. Afin d'espérer gagner du temps et rendre la recherche d'occurrences automatique, nous nous sommes essayée à la mise au point d'un codage spécifique pour les démonstratifs.

[84] Citons ici, par exemple, Sichel-Bazin, Buthke & Meisenburg (2012) sur la prosodie des questions, Likhacheva-Philippe (2008) sur les constructions détachées à gauche, ou encore Boutin (2009) sur les procédés de catégorisation dans des variétés de français de l'Afrique de l'Ouest.

[85] Il existe certains analyseurs de l'anglais et du français écrit et oral, pour la plupart en cours de développement (cf. ce chapitre, section 5.2.1 *infra*).

5.1. Comment construire un corpus voué à l'étude des démonstratifs à partir de corpus destinés aux phonologues ?

5.1.1. Un essai d'annotation chiffrée 'à la PFC/PAC'

Notre première idée fut de nous inspirer du système de codage adopté au sein des projets PFC et PAC. Ainsi, nous avons inséré des symboles alphanumériques dans la transcription orthographique après chaque occurrence de pronom démonstratif afin d'automatiser, à des fins descriptives, la quantification et la qualification des phénomènes linguistiques associés. Dans un premier temps, nous avons déterminé trois champs pour le codage du démonstratif neutre *ça* et quatre pour celui des pronoms *this* et *that*. Le codage se présentait alors de la manière suivante :

- Codage du démonstratif français *ça* :
  - 1er chiffre pour la position syntaxique de l'occurrence :
    - 0 sujet
    - 1 objet
    - 2 objet second
  - 2ème chiffre pour l'ordre du référent visé :
    - 0 entité d'ordre zéro
    - 1 entité de premier ordre
    - 2 entité de deuxième ordre
    - 3 entité de troisième ordre
    - 4 entité de quatrième ordre
  - 3ème chiffre pour la procédure indexicale réalisée :
    - 0 deixis
    - 1 anaphore
    - 2 anadeixis

Ce qui pourrait être illustré par :

(8) AL : *Il faut trois fois changer de métier dans sa vie* <E : *C'est, qui c'est qui a dit ça132 ?> Les Jésuites*. (Corpus PFC Douzens, code locuteur 11aal1, conversation guidée)

- Codage des pronoms anglais *this* et *that* :
    - 1er chiffre pour distinguer les deux pronoms :
        - 1 *this/these*
        - 2 *that/those*
    - 2ème chiffre pour la position syntaxique de l'occurrence :
        - 0 sujet
        - 1 objet
        - 2 objet second
    - 3ème chiffre pour l'ordre du référent visé :
        - 0 entité d'ordre zéro
        - 1 entité de premier ordre
        - 2 entité de deuxième ordre
        - 3 entité de troisième ordre
        - 4 entité de quatrième ordre
    - 4ème chiffre pour la procédure indexicale réalisée :
        - 0 deixis
        - 1 anaphore
        - 2 anadeixis

Ce qui pourrait être illustré par :

(9) MO : *There's quite a few repeats, but you needn't look at those*. <LC: *Oh, don't worry, don't worry*.> (silence) *Right, this1010 is your, great grand-son?* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

Le premier problème auquel nous avons été confrontée est la description de certains phénomènes que nous souhaitions faire apparaître dans le codage lui-même. Ainsi, le code alphanumérique pour qualifier *this* et *that* s'est vu augmenté d'un champ supplémentaire en vue de noter la (non-)accentuation de l'occurrence :

- Codage des pronoms anglais *this* et *that* :
    - 1er chiffre pour distinguer les deux pronoms :
        - 1 *this/these*
        - 2 *that/those*
    - 2ème chiffre pour la position syntaxique de l'occurrence :

        - 0 sujet
        - 1 objet
        - 2 objet second
    - 3ème chiffre pour l'accentuation de l'item :
        - 0 non accentué
        - 1 accent nucléaire
        - 2 accent contrastif
    - 4ème chiffre pour l'ordre du référent visé :
        - 0 entité d'ordre zéro
        - 1 entité de premier ordre
        - 2 entité de deuxième ordre
        - 3 entité de troisième ordre
        - 4 entité de quatrième ordre
    - 5ème chiffre pour la procédure indexicale réalisée :
        - 0 deixis
        - 1 anaphore
        - 2 anadeixis

Ce qui pourrait être illustré par :

(10) MO : *There's quite a few repeats, but you needn't look at those*. <LC: *Oh, don't worry, don't worry*.> (silence) *Right, this10110 is your, great grand-son?* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

De même, le premier champ du codage choisi pour le démonstratif neutre *ça* a dû être accru de plusieurs valeurs potentielles, notamment pour faire face à certains emplois particuliers de *ça* :

- Codage du démonstratif français *ça* :
    - 1er chiffre pour la position syntaxique de l'occurrence :
        - 0 sujet
        - 1 objet
        - 2 objet second
        - 3 attribut
        - 4 complément

        - 5 adjoint
        - 6 dislocation ou clivée
        - 7 post-adverbe + lettre (ex : ‘t’ pour *tout ça* ; ‘c’ pour *comme ça*)
        - 8 marqueur discursif (ex : *Ah ça !* ; *Alors ça !*)
    - 2ème chiffre pour l’ordre du référent visé :
        - 0 entité d’ordre zéro
        - 1 entité de premier ordre
        - 2 entité de deuxième ordre
        - 3 entité de troisième ordre
        - 4 entité de quatrième ordre
    - 3ème chiffre pour la procédure indexicale réalisée :
        - 0 deixis
        - 1 anaphore
        - 2 anadeixis

Ce qui pourrait être illustré par :

(11) AL : *On nous avait donné des primes d’arrachage tout ça7t12 et j’ai arraché, j’ai tout arraché.* (Corpus PFC Douzens, code locuteur 11aal1, conversation guidée)

Par conséquent, la liste des valeurs potentielles pour chaque champ, mais aussi le nombre de champs codés, risquaient d’être continuellement enrichis au fur et à mesure de l’observation des phénomènes. Par ailleurs, il a fallu nous rendre à l’évidence que certains paramètres, comme l’ordre de l’entité désignée ou la procédure indexicale réalisée, s’accordent mal avec un système de codage chiffré ; en effet, ces phénomènes d’ordre sémantique forment plus un continuum qu’un ensemble de valeurs clairement délimitées (cf. chapitre IV). De plus, l’annotation systématique du fonctionnement des démonstratifs ne peut s’effectuer dans la même optique que celle du schwa ou de la liaison car elle ne peut se priver de pré-analyse. Enfin, il est une autre caractéristique à faire apparaître et dont il est difficile de rendre compte à l’aide d’un codage alphanumérique : la récursivité de la langue. Cette dernière remarque nous a menée vers un essai d’annotation de type attribut-valeur.

5.1.2. Un essai de système de codage à attributs-valeurs

Un système d'annotation de type attribut-valeur ne modifie pas les champs codés, puisque la valeur chiffrée est remplacée par le nom que l'on attribue au phénomène décrit auquel est associée une valeur ; ainsi, les problèmes rencontrés lors d'un codage alphanumérique – listes des champs et des valeurs potentielles croissantes, continuum des phénomènes observés et pré-analyse – demeurent. Néanmoins, ce système pourrait permettre de résoudre celui de la récursivité de la langue. En effet, le point fort de cette méthode d'annotation est de pouvoir représenter les enchâssements, comme pourrait l'illustrer l'exemple suivant :

(12) TG : <P> *Non* <ench1 fct = 'cause'> *parce que ça* {stx = S ; ordre_ref = 1 ; proc_idx = ana} *ressemblait pas au patois que euh, parlait ma grand-mère* </ench1></P>. (Corpus PFC Douzens, code locuteur 11atg1, conversation guidée)

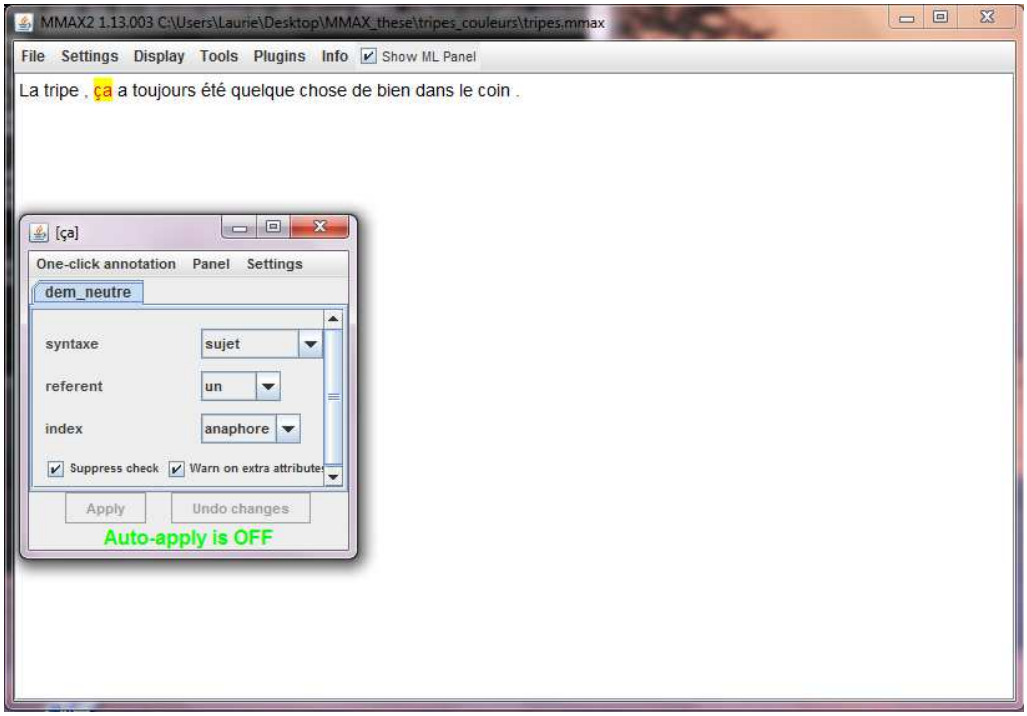
où <P> et </P> indiquent le début et la fin de la phrase considérée et où <ench1> et </ench1> encadrent la circonstancielle de cause. Hélas, un tel système d'annotation requiert une plateforme capable de décrypter le codage afin de rendre les données traitées accessibles ; or, la plateforme Dolmen n'est pas prévue à cet effet. Il a donc fallu trouver un logiciel capable de lire ces codes, si possible un programme à interface graphique, de manière à rendre les données accessibles au novice en informatique.

5.1.3. Une plateforme open-source pour l'annotation linguistique multi-niveaux : MMAX2

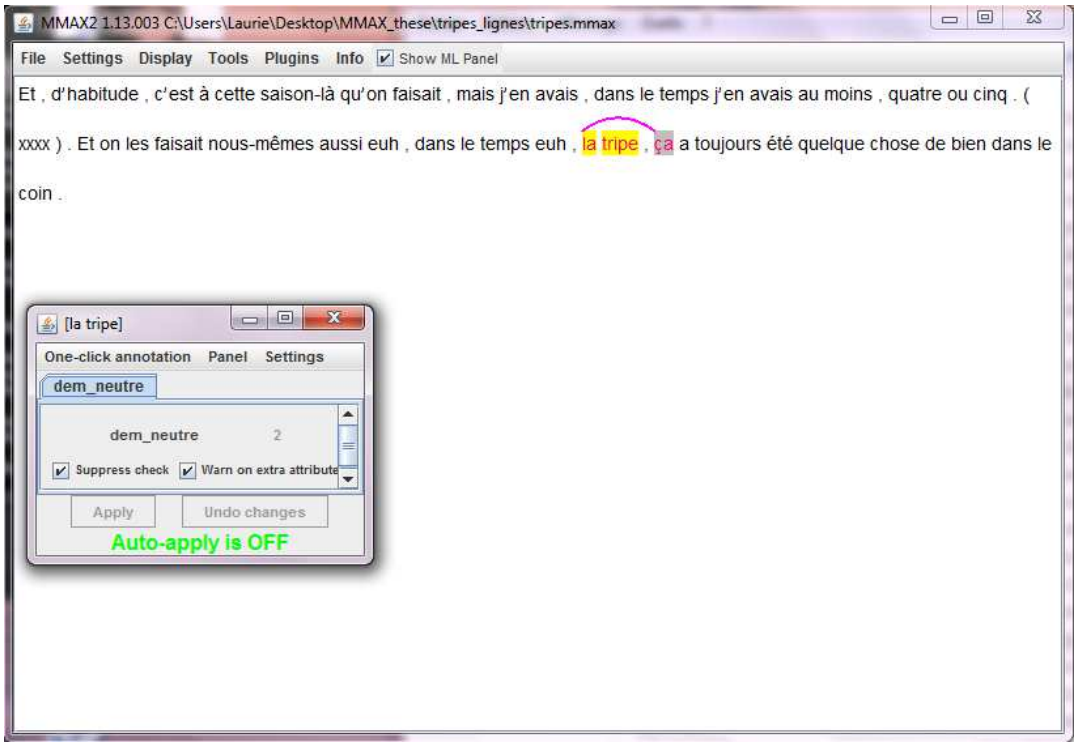
MMAX2 (Müller & Strube, 2006) est un outil d'annotation à usages multiples et basé sur le format XML. Cette plateforme permet notamment d'étiqueter des mots ou des groupes de mots selon les critères souhaités par l'utilisateur. Il suffit pour cela d'entrer le texte à annoter et de paramétrer les attributs et les valeurs possibles. Supposons que nous voulions annoter le démonstratif neutre de l'exemple suivant :

(13) JB : *La tripe, **ça** a toujours été quelque chose de bien dans le coin*. (Corpus PFC Domfront, code locuteur 61ajb1, conversation libre)

Après le réglage des paramètres, la plateforme MMAX2 présente sous forme d'interface graphique les attributs et valeurs associés à cette occurrence :



Ça est en surbrillance et les paramètres autrefois représentés par un codage chiffré sont affichés dans la fenêtre d'annotation correspondante : *ça* est ici sujet, réfère à une entité de premier ordre et est en emploi anaphorique. Il est également possible d'associer deux ou plusieurs unités linguistiques en coréférence :



Ici, *ça* est relié à *la tripe*. Notons en passant que MMAX2 autorise l'envoi de requêtes sous forme SQL[86], utiles si l'on souhaite par exemple recenser toutes les occurrences qui réfèrent au même objet-de-discours, ou tous les pronoms en position sujet.

Cette plateforme présente de nombreux avantages, que ce soit du point de vue des annotations ou de l'interface ; néanmoins, elle n'autorise le chargement que d'un fichier à la fois – il est donc difficile d'effectuer des comparaisons entre locuteurs, par exemple – et ne propose pas d'automatisation des annotations. En d'autres termes, l'utilisateur doit lui-même paramétrer les attributs et les valeurs, ce qui n'occasionne aucun gain de temps. Toutefois, MMAX2 s'approche indéniablement d'un outil qui serait capable de traiter des données en vue de recherches sur la syntaxe de l'oral, voire sur les phénomènes sémantiques associés.

5.2. Les projets PFC et PAC et l'annotation syntaxique : le projet SAXO (Buscail & Eychenne)

Les difficultés rencontrées lors de la constitution de notre corpus de démonstratifs a généré la volonté d'étendre cette réflexion sur les corpus et l'annotation micro- et macro-syntaxique à l'ensemble des bases PFC et PAC. C'est ainsi qu'est née la collaboration avec Julien Eychenne, déjà à l'origine de la plateforme Dolmen (cf. ce chapitre, section 4.1.3.2), afin de créer un analyseur du français parlé : SAXO (SyntAXe de l'Oral). Si cet outil sera dans un premier temps opérationnel seulement pour le français, nous souhaitons l'adapter ensuite à l'anglais.

5.2.1. Quelques méthodes d'annotation existantes

La recherche en traitement automatique du langage offre nombre d'outils d'annotation syntaxique et discursive, plus ou moins aboutis. Hélas, ces outils sont majoritairement voués à l'anglais et difficilement accessibles au public. Quelques annotateurs sont destinés au français mais sont généralement tournés vers l'automatisation de textes écrits, comme le logiciel Syntex co-développé à Toulouse par le laboratoire CLLE-ERSS (UMR 5263) et par la société Synomia. Toutefois, un petit historique de ces différents projets permet de situer notre futur SAXO parmi l'ensemble d'outils existants et fournit des indices sur la méthode que nous serons amenée à employer pour développer cette plateforme.

[86] SQL (Structured Query Language) est un langage informatique utile pour effectuer des requêtes sur des bases de données relationnelles, qui sont des stockages d'informations organisées dans des matrices.

Parmi les systèmes mis au point en vue d'une annotation syntaxique 'et au-delà' des textes, nous pouvons signaler MapTask (Anderson *et al.*, 1991), lancé dans les années quatre-vingt-dix, qui enregistre, transcrit et code des dialogues entre locuteurs souhaitant s'orienter suivant les cartes géographiques et les indications qu'ils ont en main. Notons également les travaux autour des conférences MUC (Message Understanding Conferences, cf. Grishman & Sundheim (1996) pour un historique de ces conférences) qui ont pour volonté de développer des méthodes d'extraction d'information à partir de rapports militaires et journalistiques ; ces conférences ont permis de générer des systèmes d'annotation de la coréférence, d'abord de manière 'inline', où le codage est inséré aux données elles-mêmes, puis 'standoff', où le fichier contenant les codages est annexé aux données. L'avantage de l'annotation 'standoff' est qu'elle permet de marquer certains phénomènes référentiels comme la deixis de discours* (Cornish, 1999, 2007, et cf. chapitre IV, section 2.5) où le déclencheur d'antécédent n'est pas sous forme de SN. Dans la lignée des travaux de MUC, il convient de relever le projet MATE (Multilingual Annotation Tools Engineering, cf. McKelvie *et al.*, 2001) dont le but est de développer des outils pour l'annotation de dialogues. De MATE est né un sous-projet, GNOME (cf. Poesio, 2004), qui se focalise sur les facteurs influençant l'apparition et la forme des expressions référentielles au sein des textes. Le projet ARRAU (Université de Glasgow et Université d'Essex), issu de GNOME, s'intéresse à la sous-spécification référentielle qui se manifeste dans les cas de reprise anaphorique d'une entité méréologique – par exemple, *The engineer hooked up the engine and the boxcar and sent **it** to London*, où « engine » et « boxcar » font partie du même ensemble « train » (exemple issu de Filik *et al.*, poster présenté à la ECEM 2005). L'ensemble de ces projets et des développements qui en sont issus – MapTask, MUC, MATE, GNOME et ARRAU – est orienté selon des fins spécifiques, que ce soit pour des raisons financières – MUC est lancé par la DARPA (Defense Advanced Research Projects Agency), qui est en charge de la création de nouvelles technologies à usage militaire aux Etats-Unis – ou scientifiques – MATE, GNOME et ARRAU s'intéressent quasi exclusivement aux phénomènes de référence et d'anaphore. Ainsi, les systèmes d'annotation qui y sont développés ne sont pas foncièrement destinés à l'automatisation de l'anglais parlé contemporain. L'entreprise qui rejoindrait peut-être les desseins du projet PAC et de notre futur SAXO serait le projet ICE (International Corpus of English, cf. le numéro consacré de *World Englishes* 23 (2)), dont la composante britannique est nommée ICE-GB. Le corpus ICE-GB est issu du Survey of English Usage, pour la première fois compilé en 1959 par Greenbaum, Quirk, Leech et Svartvik. La version computationnelle de ce corpus est le célèbre London-Lund Corpus (LLC), qui a généré ICE en 1990 et ICE-GB en 1998. Le projet ICE est

à l'origine d'un logiciel nommé ICECUP, qui permet d'effectuer diverses requêtes lexicales et syntaxiques sur le corpus ICE-GB. La méthode d'annotation du corpus ICE-GB se base sur une découpe morphosyntaxique puis sur un étiquetage des constituants, eux-mêmes analysés en tant que gouverneurs et dépendants syntaxiques. En ce qui concerne l'analyse semi-automatique du français parlé, notons l'équipe du projet Rhapsodie (officiellement terminé depuis juin 2012, initialement porté par les laboratoires MoDyCo (Université Paris Ouest La Défense), IRCAM (Centre Pompidou), LATTICE (Université Sorbonne Nouvelle), ERSS (Toulouse II-Le Mirail) et LPL (Université d'Aix-en-Provence)) qui s'essaie au développement d'un annotateur intono-syntaxique ; toutefois, ce projet s'intéresse principalement à l'analyse prosodique et demeure éloigné de nos objectifs concernant SAXO. Il existe néanmoins certains analyseurs du français écrit et parlé, en cours de développement, et tous plus ou moins centrés sur une grammaire générative à base de représentations linéaires ou d'arbres : l'analyseur Fips (du laboratoire LATL, Laboratoire d'Analyse et de Technologie du Langage, à l'université de Genève), le FRMG (de l'institut INRIA, Institut de Recherche en Informatique et en Automatique), le JSynATS (du laboratoire LORIA, Laboratoire Lorrain de Recherche en Informatique et ses Applications), ainsi que l'analyseur LEOPAR (LORIA et INRIA). Même si SAXO ne se base pas sur un annotateur existant quelconque, il est indéniable que ces divers analyseurs offrent des pistes non négligeables pour la mise au point de notre propre outil.

5.2.2. La plateforme SAXO : où en sommes-nous ?

Dans l'idéal, SAXO est une plateforme d'analyse du français parlé tout d'abord, de l'anglais parlé dans un second temps. A terme, cet analyseur sera en mesure de traiter des données syntaxiques mais aussi discursives. L'enjeu est donc de structurer et annoter des transcriptions orthographiques qui reflètent au mieux l'oralité caractéristique d'une conversation.

5.2.2.1. Quelques spécificités du français parlé difficiles à traiter de manière computationnelle

Parmi les caractéristiques du français – et de l'anglais – parlé délicates à implémenter se trouvent les dysfluences qui jalonnent nos conversations, comme les hésitations et les remplisseurs, ainsi que les mots tronqués :

(14) CM : *Donc **euh**, tu fais quel **euh**, rôle au théâtre ?*

(15) CM : ***Tu vois**, **tu vois**, ils sont pourris mes profs parce que, **je veux dire**, toi euh, le seul truc positif c'est 'personne très gentille et très polie'.*

(16) CM : *Non, mais hier **s/** en plus hier attends, je retrouve euh, l'enveloppe. CM: Hier regarde ce que j'ai reçu. **C/** non il est pas là.* (exemples (14), (15) et (16) issus du corpus PFC Domfront, code locuteur 61acm1, conversation libre)

Ces disfluences sont d'autant plus complexes à traiter par un analyseur que ce dernier devrait pouvoir les distinguer de répétitions volontaires :

(17) ML: *il est **très, très** attentionné ce petit, la fille aussi n'en parlons pas.* <E: *Et Chloé, aussi.> Elle est **très, très** bien, elle est **très, très** bien Chloé, oh oui, oui, oui*. (Corpus PFC Douzens, code locuteur 11aml1, conversation libre)

Au niveau syntaxique, les abandons de structures, avec ou sans reprise différée, les dislocations et les constituants non canoniques représentent une gageure certaine :

(18) PL : ***On me demandait euh si** euh qu'est-ce que j'étais **si j'étais basque ou si j'étais française.*** (Corpus PFC Pays Basque, code locuteur 64apl1, conversation guidée)

(19) RB: *Et puis là euh, **Anne-Laure et Christian, ils** sont montés au chalet euh, samedi, et puis il y avait plus le panneau à vendre*. (Corpus PFC Nyon (Suisse), code locuteur svarb1, conversation libre)

(20) *Tu as choisi celle aux anchois, **de pizza** ?* (exemple fictif)

Enfin, les tours de parole ainsi que les chevauchements sont deux caractéristiques de l'oral conversationnel que notre annotateur devra maîtriser. Au stade actuel du développement de SAXO, le problème posé par les troncations est résolu par l'adaptation de TreeTagger aux données PFC (cf. ce chapitre, section 5.2.2.2.2 *infra*).

5.2.2.2. Les étapes du développement de SAXO

A terme, SAXO sera un analyseur multi-niveaux. Ainsi, son développement s'effectue par couches successives, chacune prenant en charge un module spécifique : mots, catégories morphosyntaxiques, constituants, micro-syntaxe* et macro-syntaxe.

5.2.2.2.1. D'un sous-corpus-test à l'ensemble de la base de données PFC

La base de données PFC étant composée d'une soixantaine de points d'enquêtes et représentant environ 200 giga-octets de données à ce jour, il serait impossible de songer à développer un annotateur à partir de cet ensemble. Nous avons donc choisi de nous concentrer sur une seule enquête dans un premier temps afin de pouvoir construire un système robuste qui, nous l'espérons, pourra s'élargir à toute la base PFC. L'enquête sélectionnée se devait de répondre aux critères imposés par le protocole PFC, à savoir douze locuteurs répartis de manière égale en trois tranches d'âge (cf. ce chapitre, section 4.1.1). De plus, les transcriptions se devaient d'avoir été vérifiées par une personne autre que le transcripteur/codeur lui-même. C'est ainsi que l'enquête PFC Domfront (61, Normandie) a été retenue pour servir de première base à la création de SAXO.

5.2.2.2.2. La segmentation par TreeTagger et l'adaptation de ce logiciel au corpus PFC

TreeTagger est un outil de segmentation des textes en mots qui indique la base lemmatique et la catégorie morphosyntaxique de chaque occurrence isolée. Pour le français, la liste des catégories morphosyntaxiques disponibles compte trente-trois items :

ABR	Abréviation
ADJ	Adjectif
ADV	Adverbe
DET:ART	Article
DET:POS	Déterminant possessif
INT	Interjection
KON	Conjonction[87]
NAM	Nom propre
NOM	Nom
NUM	Numéral
PRO	Pronom

[87] Notons ici que, pour l'instant tout du moins, TreeTagger inclut les subordonnants et les coordonnants dans sa catégories des conjonctions.

PRO:DEM	Pronom démonstratif
PRO:IND	Pronom indéfini
PRO:PER	Pronom personnel
PRO:POS	Pronom possessif
PRO:REL	Pronom relatif
PRP	Préposition
PRP:det	Préposition + article (*au, du, aux, des*)
PUN	Ponctuation
PUN:cit	Ponctuation citation
SENT	Phrase
SYM	Symbole
VER:cond	Verbe au conditionnel
VER:futu	Verbe au futur
VER:impe	Verbe à l'impératif
VER:impf	Verbe à l'imparfait
VER:infi	Verbe à l'infinitif
VER:pper	Participe passé
VER:ppre	Participe présent
VER:pres	Verbe au présent
VER:simp	Verbe au passé simple
VER:subi	Verbe au subjonctif imparfait
VER:subp	Verbe au subjonctif présent

TABLEAU 5 : LES CATEGORIES MORPHOSYNTAXIQUES DU FRANÇAIS PROPOSEES PAR TREETAGGER

Ainsi, un fichier texte (*.txt) en entrée de TreeTagger aura la structure suivante en sortie :

Occurrence	POS[88]	Lemme
mon	DET:POS	mon
père	NOM	père
était	VER:impf	être
arrivé	VER:pper	arriver
avant	PRP	avant
que	KON	que
de	PRP	de
,	PUN	,
se	PRO:PER	se
marier	VER:infi	marier
.	SENT	.

TABLEAU 6 : UN EXTRAIT DE CONVERSATION PFC ETIQUETE PAR TREETAGGER

La simple 'tokenisation' – *i.e.* la segmentation en occurrences de mots – ci-dessus offre un avant-goût des difficultés auxquelles est confronté TreeTagger : *avant que* est scindé en deux, *avant* étant étiqueté comme une préposition et *que* comme une conjonction ; or, il conviendrait de traiter *avant que* comme un ensemble cohésif. Outre les mots composés, la tokenisation des textes par TreeTagger engendre des erreurs fréquentes comme la non-reconnaissance de certains mots lexicaux, le mauvais étiquetage morphosyntaxique des marqueurs discursifs – *ben*, *euh*, *quoi*… – mais aussi de certaines formes de verbes – confusion entre présent et participe passé, par exemple – et l'absence de gestion des auxiliaires. Concernant les problèmes au niveau lexical, il est possible de contrôler TreeTagger dans une certaine mesure, puisqu'il est admis de définir un lexique complémentaire. Un fichier supplémentaire a donc été créé pour que des termes particuliers tels que *trailleuse* – terme issu du Corpus PFC Domfront, code locuteur 61abm1, conversation guidée, et désignant un appareil utile pour traire les vaches – soient reconnus par le logiciel. Les locutions conjonctives ont été reliées par un caractère de soulignement – par exemple, *avant_que* – pour que l'ensemble soit reconnu comme une conjonction. De même, les catégories DISC et AUX ont été ajoutées à la liste initiale de TreeTagger pour les marqueurs discursifs et les auxiliaires ; notons par ailleurs que l'étiquetage correct des auxiliaires réduit

[88] POS signifie 'part of speech', c'est-à-dire la catégorie morphosyntaxique.

considérablement le nombre d'erreurs d'étiquetage des participes passés. De plus, l'ajout des étiquettes <TRUNC-X> et <GUESS-X> permet de résoudre le problème posé par les troncations (cf. ce chapitre, section 5.2.2.1) : un mot tronqué est étiqueté <TRUNC-X> (où X est le mot tronqué). Nous appliquons alors une heuristique simple : nous cherchons un mot qui commence par la troncation dans les cinq mots qui la suivent – par exemple, *i/ les autres ils se chargent*. Si nous trouvons un mot cible, la troncation est alors étiquetée <GUESS-X> et nous lui attribuons la catégorie et le lemme du mot cible :

Occurrence	POS	Lemme
oui	ADV	oui
nous	PRO:PER	nous
avons	AUX:pres	avoir
vécu	VER:pper	vivre
<GUESS-av>	PRP	avec
j'	PRO:PER	je
ai	AUX:pres	avoir
vécu	VER:pper	vivre
avec	PRP	avec
donc	ADV	donc
mes	DET:POS	mon
deux	NUM	deux
frères	NOM	frère

TABLEAU 7 : UNE ILLUSTRATION DES ETIQUETTES <TRUNC-X> ET <GUESS-X>

Nous avons repéré dans le corpus domfrontais 429 mots tronqués ; après application de cette méthode, il n'en reste que 205. Après modification, la liste des étiquettes 'version PFC' est la suivante :

ABR	Abréviation
ADJ	Adjectif
ADV	Adverbe
AUX:cond	Auxiliaire au conditionnel
AUX:futu	Auxiliaire au futur
AUX:impe	Auxiliaire à l'impératif
AUX:impf	Auxiliaire à l'imparfait
AUX:infi	Auxiliaire à l'infinitif

AUX:pper	Auxiliaire au participe passé
AUX:ppre	Auxiliaire au participe présent
AUX:pres	Auxiliaire au présent
AUX:simp	Auxiliaire au passé simple
AUX:subi	Auxiliaire au subjonctif imparfait
AUX:subp	Auxiliaire au subjonctif présent
DET:ART	Article
DET:POS	Déterminant possessif
DISC	Marqueur discursif
INT	Interjection
KON	Conjonction
NAM	Nom propre
NOM	Nom
NUM	Numéral
PRO	Pronom
PRO:DEM	Pronom démonstratif
PRO:IND	Pronom indéfini
PRO:PER	Pronom personnel
PRO:POS	Pronom possessif
PRO:REL	Pronom relatif
PRP	Préposition
PRP:det	Préposition + article (*au, du, aux, des*)
PUN	Ponctuation
PUN:cit	Ponctuation citation
SENT	Phrase
SYM	Symbole
VER:cond	Verbe au conditionnel
VER:futu	Verbe au futur
VER:impe	Verbe à l'impératif
VER:impf	Verbe à l'imparfait
VER:infi	Verbe à l'infinitif
VER:pper	Participe passé

VER:ppre	Participe présent
VER:pres	Verbe au présent
VER:simp	Verbe au passé simple
VER:subi	Verbe au subjonctif imparfait
VER:subp	Verbe au subjonctif présent

TABLEAU 8 : LES CATEGORIES MORPHOSYNTAXIQUES DU FRANÇAIS PROPOSEES PAR TREETAGGER 'VERSION PFC'

A titre illustratif, voici le découpage proposé par la nouvelle version de TreeTagger :

Occurrence	POS	Lemme
eh_ben	DISC	eh_ben
c'est-à-dire	ADV	c'est-à-dire
que	KON	que
,	PUN	,
ben	DISC	ben
moi	PRO:PER	moi
je	PRO:PER	je
suis	AUX:pres	être
arrêté	VER:pper	arrêter
au	PRP:det	au
certificat	NOM	certificat
d'	PRP	de
étude	NOM	étude

TABLEAU 9 : UN EXTRAIT DE CONVERSATION PFC ETIQUETE PAR TREETAGGER 'VERSION PFC'

Notons que cette liste n'est pas définitive et que nous désirons ajouter d'autres étiquettes – celle de DET:DEM pour 'déterminant démonstratif', par exemple. Aussi, nous avons momentanément mis de côté les chevauchements effectués par les différents locuteurs en séparant chacun d'eux afin de ne traiter qu'un tour de parole à la fois. Toutefois, les résistances éprouvées par TreeTagger face aux difficultés tout juste recensées sont facilement corrigées lors de cette phase de premier étiquetage. En revanche, il est des questions bien plus épineuses qui ne peuvent se passer d'une analyse linguistique et se résolvent par conséquent de manière *ad hoc*. Nous pensons ici notamment aux occurrences de *que* qui se classeraient parmi les adverbes – *que* au sens de *uniquement*, comme dans *je ne veux* ***que*** *celui-ci* – ou

bien à *après* qui est tantôt adverbe, tantôt préposition[89]. L'attribution d'un item à l'une ou l'autre des catégories morphosyntaxiques proposées sera donc traitée ultérieurement, lors de la construction de l'analyseur lui-même.

5.2.2.2.3. La gestion de la structure syntaxique des énoncés

La phase de développement de SAXO au niveau micro-syntaxique n'a pas encore été effectuée, mais nous pouvons d'ores et déjà avancer quelles seront nos stratégies pour gérer la syntaxe du français parlé. A la manière de Syntex (cf. ce chapitre, section 5.2.1) et de la méthode adoptée pour les corpus ICE-GB, nous choisirons une structure d'analyse définie selon la paire gouverneur-dépendant(s). Ainsi, les relations entre les différents constituants seront explicitées par des triplets de la forme (gouverneur, relation, dépendant(s)), par exemple :

(21) *Donc tu as grandi ici.*
($Donc_1$ tu_2 as_3 $grandi_4$ ici_5)
(Aux, 4, 3)
(Sujet, 3, 2)
(Circ, 4, 5)
(Connect, 4, 1)

Le découpage des textes sera effectué en clauses (cf. ce chapitre, section 2.1, ainsi que le glossaire pour une définition de ce terme), en supposant qu'il y ait autant de clauses que de mots non gouvernés. De plus, et ce dans la perspective d'obtenir un analyseur syntaxique mais aussi discursif, nous opterons pour une grammaire de type notionnel (cf. chapitre IV, section 2.1, et chapitre V, ainsi que le glossaire) ; par ce biais, les informations lexicales sont intégrées à la syntaxe et pourront donner lieu à des analyses au niveau sémantique.

Le chapitre V offre aux corpus une place similaire à celle qu'ils occupent au sein de la recherche en linguistique : centrale, postérieure aux intuitions théoriques et antérieure aux analyses. Comme nous venons de le présenter, les données orales d'où sont extraites les

[89] Notons qu'en GG, *après* est toujours considéré comme une préposition, orpheline lorsqu'elle apparaît sans complément.

occurrences au cœur de l'étude des chapitres VII et VIII proviennent des projets PFC et PAC, dont les corpus sont initialement collectés par et pour des phonologues. Si la volonté d'une annotation automatique des démonstratifs n'est pas aboutie à l'heure actuelle, elle s'est néanmoins concrétisée en la création d'un futur analyseur du français et de l'anglais parlés, la plateforme SAXO. De fait, le défi scientifique ultime sera de confronter nos analyses théoriques à l'épreuve de l'annotation automatique.

Chapitre VII. Retour sur les démonstratifs anglais : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle

1. Introduction : rappels et problématisation

Les chapitres II et III ont permis de dégager un certain nombre de propriétés des démonstratifs (cf. Bilan des chapitres II et III pour un récapitulatif), parmi lesquelles celle d'être des indices de détermination et de monstration, en lien étroit avec l'anaphore et la deixis. Nous choisissons de concentrer notre étude sur les pronoms démonstratifs à l'oral, ce qui pose le problème du manque d'information lexicale (non)véhiculée par le noyau nominal : en l'absence d'élément dénotatif, l'auditeur se voit contraint de faire appel à un ou plusieurs paramètres subsidiaires afin d'accéder à la référence. Reste à savoir quels sont ces paramètres et comment les interlocuteurs y ont recours.

Nous rappelons que le cadre choisi pour notre étude définit les notions de texte, de contexte et de discours comme étant trois composantes d'un ensemble dynamique : dans ce modèle, le texte est l'ensemble de signaux verbaux et non verbaux – texte écrit, texte oral, mais aussi signes de la tête, du regard, etc. – en fonction duquel le discours est co-construit par les participants à l'acte de communication ; ce discours est le produit de la séquence hiérarchisée et contextuellement située d'actes énonciatifs, indexicaux, propositionnels et illocutoires effectués dans la poursuite d'un but communicatif quelconque et intégrés dans un contexte donné ; ce contexte, lui-même divisé en trois sous-parties que sont le contexte situationnel, le contexte discursif et les connaissances des interlocuteurs, est assujetti à un processus continu de construction et de révision au fur et à mesure que le discours se déroule. C'est à travers l'invocation d'un contexte approprié que l'auditeur peut convertir en discours la séquence connexe d'indices textuels qu'est le texte. De plus, nous supposons que les trois sous-domaines contextuels énumérés ci-dessus (cf. chapitre IV, section 2.3, pour plus de précisions) peuvent être mis en relation avec différents types de mémoire, de la plus saillante à la plus latente : la mémoire focus, la mémoire vive et la mémoire morte (cf. chapitre IV, section 3). Ainsi, la deixis et l'anaphore ne se définissent plus exclusivement en termes de présence ou d'absence co-textuelle (cf. chapitre II, section 1), mais comme étant des stratégies discursives employées par les interlocuteurs afin d'accomplir des actes de référence : la deixis s'accorde d'un retour vers l'*origo* et permet alors l'insertion d'un « objet » quelconque en mémoire focus, tandis que l'anaphore poursuit la référence d'une entité déjà présente en

mémoire focus et donc précédemment introduite au sein du discours. La deixis et l'anaphore représentent alors les pôles extrêmes d'un continuum entre ces deux notions, reliant un ensemble de procédures discursives combinant la deixis et l'anaphore à des degrés divers (cf. chapitre IV, sections 2.5 et 4.1, pour une description approfondie). Le point central ici est que ces notions ne sont plus comprises comme des propriétés inhérentes aux éléments langagiers ; nous choisissons alors d'abandonner les termes d''éléments déictiques' et d''éléments anaphoriques' au profit d''éléments indexicaux', afin de ne pas confondre l'item langagier en tant que membre d'un système et l'actualisation que ce dernier acquiert en discours. En d'autres termes, les individus parlants et écoutants ont à leur disposition des éléments indexicaux en vue d'accomplir des actes de référence ; lors d'un échange communicatif, l'influence du réel – *i.e*. la situation d'énonciation immédiate – est constante, tout comme l'est celle du discours construit en amont, réel et discours étant des composantes contextuelles. C'est l'activation en discours de ces éléments par le sujet parlant qui détermine alors l'emploi déictique ou anaphorique des indexicaux. Ainsi, il n'y a plus de correspondance directe entre système et usage (cf. chapitre IV). Un autre point crucial, selon nous, est de ne pas omettre que l'acte communicatif est avant tout un échange : il nous semble primordial de prendre en compte l'interaction dans toute analyse discursive, qui plus est lors d'analyses d'occurrences extraites de conversations spontanées. Dès lors, notre problématique générale est de préciser l'intervention des paramètres contextuels dans l'accès à la référence indexicale, mais également de mettre en relation les contextes, les types de mémoire et les stratégies discursives que sont la deixis et l'anaphore. De plus, nous souhaitons mettre à l'épreuve les valeurs notionnelles traditionnellement attribuées aux démonstratifs – proximité de *this* et de *ceci*, éloignement de *that* et de *cela* – en proposant une conception alternative de cette distance relative qui soit un emploi volontaire de l'un ou l'autre des indexicaux par le locuteur afin de s'engager dans un acte communicatif d'où résulte une construction conjointe du discours. Ainsi, l'étude combinée de l'ensemble de ces facteurs permet de mettre en exergue le fonctionnement interne de chacun des items considérés, par la distinction effectuée entre propriétés systémiques et phénomènes observés lors de l'activation en discours ; la mise en évidence de ces propriétés systémiques peut alors conduire à une représentation syntactico-sémantique des démonstratifs anglais (chapitre VII, section 5) et français (chapitre VIII, sections 4 et 5).

Nous avons vu au chapitre VI que nos corpus d'étude sont extraits de conversations spontanées enregistrées dans le cadre des projets PAC et PFC. Si la recherche de la totalité

des occurrences de *this*, *that*, *these* et *those* en anglais s'avère fructueuse – 164 *this*, 981 *that*, 40 *these* et 24 *those*[90] sur un total de 52608 mots, mais ne pouvant hélas pas distinguer pronoms et déterminants en raison de l'absence d'annotation morphosyntaxique des données de l'anglais – il apparaît un constat frappant et sans appel au regard de la distribution des pronoms démonstratifs en français : 28 *ceci*, 35 *cela*, 6 *celui-ci*, 32 *celui-là*[91], contre 10 450 occurrences de *ça*, sur un total de 831042 mots. Si nous réduisons la recherche en français au corpus de Domfront, afin d'avoir 12 locuteurs à la fois pour l'anglais et pour le français, il résulte zéro occurrence de *ceci*, de *cela* et de *celui-ci*, 9 de *celui-là* et 1 039 de *ça* sur un total de 73772 mots. Cette prépondérance de *ça* au sein du corpus PFC nous amène à considérer les indexicaux dits 'neutres' – même si en anglais, 'neutre' prend un sens particulier en raison de l'absence générale de marque du genre – et à traiter *this*, *that* et *it* pour l'anglais, *ça* et *ce* pour le français. Nous décidons d'inclure *it* dans le système des indexicaux neutres de l'anglais et d'exclure partiellement *ce* de notre corpus d'étude pour des raisons de correspondance ; en effet, *ça* peut se traduire par *this*, *that* ou *it*, tandis que *ce* présente des propriétés distributionnelles différentes du triptyque anglais : n'oublions pas que notre analyse s'effectue dans une perspective comparative. Néanmoins, la confrontation des systèmes de l'anglais et du français ne sera opérée que lors du chapitre VIII. En effet, l'anglais présentant trois unités et le français une seule, il y a fort à parier qu'une première analyse propre à chaque système soit la solution la plus rigoureuse à adopter en vue de confronter le fonctionnement de ces deux langues.

Le chapitre présent est consacré à l'étude du fonctionnement référentiel et discursif de *this*, *that* et *it*. Cependant, sont exclues de l'analyse les occurrences employées en tant que déterminants et adverbes de degré – par exemple, *it was* ***that*** *big*. Ne seront pas non plus considérées les occurrences possibles de *this*, *that* et *it* précédées d'une préposition – par exemple, *like* ***this***, *like* ***that***. Enfin, les cas de structures ébauchées et abandonnées en cours de production – PK: *Oh yes I love sewing that's what I wanted you to know all the time er,* ***this*** *was just erm, when I was little I used to make, clothes for dolls you know, so*. (corpus PAC Lancashire, code locuteur elapk1, conversation guidée) – sont également mises à part, puisque l'analyse de la référence ne peut s'effectuer que dans l'ancrage d'un énoncé complet.

[90] Notons ici que nous trouvons surprenant le résultat quantitatif obtenu : 40 *these* et 24 *those* vont *a priori* à l'encontre du statut de *this* comme étant le membre marqué de la paire *this/that*.

[91] *Ceci* et *cela*, ainsi que *celui-ci* et *celui-là*, apparaissent majoritairement dans des structures de mise en contraste, telles que *ceci explique cela* (Corpus PFC Ile-de-France, code locuteur 91asl1, conversation libre).

Dans un premier temps, nous nous attacherons à décrire le fonctionnement référentiel de *this*, de *that* et de *it* en fonction du contexte invoqué (situationnel immédiat, discursif, savoirs). Dans un deuxième temps, nous reviendrons sur les notions de distances proximale et distale attribuées à *this* et à *that* afin de leur donner une dimension à la fois cognitive et interactionnelle. Dans un troisième temps, nous considèrerons *this*, *that* et *it* comme des signaux d'appel à la co-construction du discours par les participants à l'acte langagier. Enfin, nous proposerons une formalisation de ces pronoms dans le cadre de la GND.

2. La référence indexicale par le recours aux contextes et l'activation de certains types de mémoire

Cette section s'attache à décrire le fonctionnement indexical de *this*, de *that* et de *it* par la caractérisation des différents contextes nécessaires à l'accès à la référence de ces items et par la mise en relation de ces contextes avec l'activation de certains types de mémoire. Cette volonté n'est pas sans rappeler les échelles d'accessibilité de Gundel *et al*. ou de Cornish (cf. chapitre IV, section 2.5) qui associent *this*, *that* et *it* à la saillance cognitive du référent visé. Ainsi, il est établi que *it*, typiquement anaphorique, réfère à une entité hautement saillante, tandis que les démonstratifs *this* et *that* désignent un « objet » qui n'est pas en focus d'attention. Néanmoins, la correspondance directe effectuée entre l'unité indexicale et le statut cognitif de l'entité visée nous semble inadéquate en raison de l'unicité qui est impliquée entre un élément systémique quelconque et la valeur mémorielle d'un « objet » sémantique. Voici donc l'analyse que nous proposons de la référence indexicale de *this* (section 2.1), de *that* (section 2.2) et de *it* (section 2.3).

2.1. *This*

Comme nous allons le constater par l'analyse d'exemples caractéristiques de l'oral spontané, *this* semble être en mesure de référer par le truchement de deux catégories contextuelles parmi les trois proposées. Les occurrences issues de notre corpus d'étude permettent en outre de mettre en relation les contextes invoqués et l'activation de certains types de mémoire. Enfin, nous pouvons intégrer à cette description indexicale les procédures discursives que sont la deixis et l'anaphore, telles que nous les avons définies au chapitre IV, section 4.

2.1.1. La référence indexicale de *this* par l'invocation du contexte situationnel immédiat

La visée référentielle effectuée par *this* peut s'effectuer via l'intervention du contexte situationnel immédiat, comme l'illustrent les exemples (1) à (5). *This* dans les exemples (1), (2), (3) et (4) réfère par le contexte spatial, du plus concret au plus cognitif, tandis que *this* dans l'exemple (5) passe par un ancrage temporel :

(1) [MO et LC regardent et commentent des photos de famille appartenant à MO]

MO: *There's quite a few repeats, but you needn't look at those* <LC: *Oh don't worry, don't worry.*>. (silence) LC: *So* ***this*** *is your great grandson* <MO: *That's him.*>, *isn't that lovely? And how old is he now?* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

En (1), *this* renvoie à l'une des photos que regardent les deux locutrices MO et LC ; l'objet-de-discours ainsi visé est une image mentale de premier ordre (cf. chapitre IV, section 2.1), à ne pas confondre avec l'objet réel, physique, palpable, qu'est la photo représentant l'arrière-petit-fils de MO. L'accès à la référence se fait par l'intervention de l'environnement extralinguistique, dont le lien entre l'élément indexical et la situation immédiate classe souvent, de manière réductrice selon nous, les démonstratifs parmi l'ensemble des déictiques. Notre description du fonctionnement référentiel effectué par *this* ici est quelque peu différente. Par l'emploi de ce démonstratif, LC signale à MO qu'un retour sur le *hic* du repère énonciatif est nécessaire pour identifier le référent voulu. Si nous nous plaçons du côté du destinataire MO, l'occurrence de *this* provoque un envoi référentiel vers la mémoire de cette dernière ; le référent désigné n'ayant pas été introduit auparavant au sein du discours, celui-ci n'est pas disponible en mémoire. Par conséquent, un second mouvement référentiel s'opère, cette fois-ci vers le repère énonciatif, afin de trouver le référent correspondant à *this*. Cette description s'accorde à notre définition de la deixis (cf. chapitre IV, section 4.1), caractérisée par l'entrée en mémoire focus d'un objet-de-discours par un mouvement rétroactif vers le repère *origo-hic-nunc*. Notons ici que ce repère peut être déplacé, comme l'illustre l'exemple (2), issu de la même conversation que l'exemple (1) :

(2) LC: *Oh, he's beautiful*. <MO: *So Mary, my daughter-in-law, says 'I'll bathe him in the sink',> so* (laughter) *she puts a towel, on the draining board, down and then up again*

and, and he went and his eyes, you know, (XX), ‘*this isn't* <LC: *Yes, oh.*> *my bath’, and all at once he went*.

En (2), l’*origo* du repère énonciatif n’est ni LC ni MO, mais le petit-fils de cette dernière. Cette procédure déictique correspond à la *deixis am phantasma* de Bühler (cf. chapitre IV, section 2.5).

(3) [I est l’enquêtrice et DK le locuteur enregistré ; DK suit ses études à l’université d’Edinbourg, en Ecosse, mais l’enquête a lieu à Bury, dans le comté urbain du Grand Manchester]

I: *And er, what about motor cycling? You said* <DK: *Yeah, yeah.*> *it's one of your, one of the things you do during your leisure?* DK: *Yeah, well that was sort of being at uni and don't get the chance cause I leave the bike here, er, i/ in Manchester. Er, when I'm at uni, cause it's, Edinburgh is just freezing, cold and rainy all the time so you just wouldn't get the opportunity to ride it,* <I: *Really?*> *yeah, oh it's horrid, yeah, yeah,* (laughter) ***this*** *is warm, out there is warm, compared to Edinburgh* (laughter). *It's nice to come home and get a bit of sun.* (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

En (3), *this* réfère grâce à l’invocation d’un contexte situationnel plus large que l’est celui de l’exemple (1), puisqu’il s’agit ici de la ville, et même du comté, où l’enregistrement est effectué. Ici encore, le fonctionnement référentiel de *this* est déictique, caractérisé par un retour sur le *hic* du repère énonciatif : le référent visé pourrait se définir comme étant « l’atmosphère ou la température moyenne qui règne dans le Grand Manchester ». Néanmoins, la mémoire du destinataire I n’est pas entièrement dépourvue de traits notionnels associés à ce référent lors de l’occurrence du démonstratif, puisque « Manchester » est évoqué peu auparavant. Par conséquent, nous pouvons dire qu’il existe un soupçon de procédure anaphorique à la procédure référentielle opérée. Toutefois, « Manchester » n’étant pas au cœur du débat lors de l’occurrence de *this* – DK parle d’ « Edinbourg et de ses températures glaciales » – les indices référentiels qui lui sont associés font plus partie de la mémoire périphérique du discours que de la mémoire focus. Ainsi, la référence de *this* ici s’effectue principalement par l’ancrage déictique.

(4) [Après avoir parcouru les différents métiers qu'elle a exercés par le passé, LB parle de sa profession actuelle]

LB: *And, from then on, I've been to college for two years doing my N.V.Q. level one and two. I've been on courses galore, erm, and* ***this*** *is where I'm now, I run my own kitchen,* <I: *So you*.> *which I love.* (Corpus PAC Lancashire, code locuteur elalb1, conversation guidée)

A la manière de *this* dans les exemples (1) et (3), *this* dans l'exemple (4) réfère par le recours au contexte situationnel spatial, à la différence que ce contexte perd sa dimension physique au profit d'une caractérisation cognitive. En quelque sorte, *this* signale ici un retour sur le repère énonciatif dont l'*origo* est LB et dont le *hic* est interne à cette même locutrice. Ainsi, le destinataire I doit pénétrer l'espace personnel de LB afin d'identifier le référent de *this* comme étant un point d'ancrage marquant la situation professionnelle de LB à l'instant t_0. Nous remarquons par ailleurs que le référent visé est difficile à caractériser. S'agit-il de « la situation professionnelle de LB à l'instant t_0 » ? Cette solution n'est vraiment pas convaincante, et pour cause, puisque « la situation professionnelle de LB à l'instant t_0 » n'est pas généré par *this*, mais bien par l'ensemble de l'énoncé *this is where I'm now, I run my own kitchen* contextuellement activé. Le rôle discursif de *this* ici se résume à l'ancrage déictique, spatial et cognitif, utile pour situer l'ensemble de la prédication.

(5) [MC fait part de son désir de vendre sa maison]

MC: *er, you know, because it'll take about twelve months for me to move I think, it may (X) take less, I don't know, but* ***this*** *is the time to sell the house and I want to, I really don't need a house this size now and I want to bring my overheads down, Marianne, you know so*. (Corpus PAC Lancashire, code locuteur elamc1, conversation libre)

Contrairement aux exemples (1) à (4), *this* en (5) ne réfère pas via l'invocation du contexte spatial mais par le contexte temporel. Cependant, l'emploi de *this* partage avec celui de l'exemple (4) le rôle de point d'ancrage déictique, ici temporel, utile pour situer l'ensemble de la prédication.

Par l'observation des exemples (1) à (5), nous comprenons l'amalgame (trop) souvent effectué entre les démonstratifs, alors classés parmi les déictiques, et la référence à l'extralinguistique ; il s'avère en effet que la référence de *this* peut s'effectuer par le recours au contexte situationnel immédiat. Néanmoins, nous caractérisons la référence déictique opérée par *this* dans ces cas-ci par un retour sur le repère énonciatif, à défaut de référent correspondant disponible dans la mémoire du destinataire. Cependant, l'emploi de *this* le plus frappant, selon nous, est celui où ce démonstratif est dépourvu de référence et ne sert que d'ancrage déictique ; de plus, à l'exception peut-être de l'exemple (1), substituer *that* à *this* dans ces énoncés ne semble pas naturel – *#that is warm* (sous-entendu, *here in Manchester*) ; *#that is where I'm now, I run my own kitchen* ; *#that is the time to sell the house*. Ce rôle discursif de localisation spatio-temporelle semble indéniablement rapprocher *this* de *here* (exemples (3) et (4)) et de *now* (exemple (5)).

2.1.2. La référence indexicale de *this* par l'invocation du contexte discursif

Afin d'accomplir sa référence, *this* peut dépendre du discours suivant (section 2.1.2.1) et précédant (2.1.2.2.) son occurrence. Les deux sections suivantes sont donc consacrées à la description du fonctionnement indexical de *this* dans ces deux cas de figure.

2.1.2.1. *This* en position de référence cataphorique

Les exemples (6) et (7) illustrent un emploi propre à *this*, qui est celui d'apparaître en position antéposée à celle du déclencheur d'antécédent. En d'autres termes, *this* annonce que la construction du référent visé s'effectue grâce au discours à venir :

(6) [I questionne GT sur son environnement linguistique]

GT: *no they don't.* <I: *Right.*> I: *And, I don't know* ***this*** *is a bit embarrassing,* <GT: *Not at all.*> *is there a significant other that, you speak, interact linguistically with, or? N/ hmm*. (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

(7) [GT étudie la linguistique à Salford ; I, qui a longtemps fréquenté cette université, déduit que la linguistique en tant que matière principale est un programme récemment instauré]

I: *And you do, do you do French and another language, I'm sorry.* <GT: *I don't, I don't know languages (XX)/ no, no, no, just at all, just linguistics*.> *You just do French. No you're just doing linguistics.* <GT: *Last year I did some French*.> *Just linguistics sorry I.* <GT: *Yeah, linguistics and that's all*.> *I got it. I knew you were doing linguistics but I thought erm.* GT: *Yes I mean there's all sorts of programs* <I: *But, **this** is new isn't it in Salford, you can do linguistics, er as a single, single (XX)*.> *on that. Yes it is, yes, yes, yeah.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

En (6), le déclencheur d'antécédent, c'est-à-dire la portion de texte qui sert de support référentiel à l'identification de l'objet-de-discours visé par *this*, est *is there a significant other that, you speak, interact linguistically with, or?* ; en (7), ce déclencheur est *you can do linguistics, er as a single, single (XX)*. Dans les deux cas, le déclencheur d'antécédent est subséquent à l'occurrence de *this* ; ce phénomène est traditionnellement connu sous le nom de 'cataphore'. Comment pouvons-nous décrire ce fonctionnement référentiel ? Si nous nous plaçons du côté du destinataire, l'occurrence de *this* provoque un envoi référentiel vers sa mémoire ; nul référent potentiel n'y étant présent, un mouvement rétroactif vers le repère énonciatif est opéré. A la différence des retours référentiels décrits dans les exemples (1) à (5), celui qui se manifeste dans les emplois de *this* en (6) et en (7) ne revient pas sur le repère spatio-temporel mais sur le repère discursif. En quelque sorte, le signal amorcé par *this* pourrait se traduire par « je vise un référent qui n'est pas présent dans ta mémoire – *i.e.* celle du destinataire – il va donc falloir que tu effectues un retour sur le discours immédiat afin d'identifier le référent voulu ». Grâce à ce signal, l'énonciateur annonce à son destinataire qu'il s'apprête à introduire un référent dans sa mémoire focus. Ainsi, nous pouvons qualifier cette procédure référentielle de déictique.

2.1.2.2. *This* puise sa référence dans le discours en amont

Deux cas possibles émergent des occurrences de *this* pour lesquelles l'accès à la référence passe par le contexte discursif : soit *this*, grâce à sa position d'argument, permet la topicalisation d'un référent discret à partir d'informations diffuses présentes en mémoire discursive (exemple (8)), soit *this* maintient la référence d'un objet-de-discours précédemment évoqué (exemple (9)) :

(8) [La conversation entre I et RH tourne autour du football, en particulier de la manière qu'ont les enfants de grandir en pratiquant ce sport dans la rue]

I: *Is there still a lot of that, five-a-side or, you know that's.* <RH: *Yeah.*> RH: *there is where, there is where I am, y/ you get kids, sort of like, they're playing this, they're playing this street (XX) and they're all on their way to school kicking the ball on the road and stuff it's, yeah, I think, I think, especially I think, they grow up playing from, I think they're just, they're told to play, if you know what I mean, I think* ***this*** *is something they grow up with.* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(9) [GT est atterrée par le prix exorbitant de certains vêtements et accessoires, uniquement justifié par la marque de fabrique de ces objets]

GT: *But, you know bu/ I was looking in erm, in Selfridges, and, some of the stuff are (XXX). And, it had all threads, hanging from the hand, like it was unfinished.* <RH: *Yeah.*> *And* ***this*** is, *(XXX) four hundred and seventy-five pounds.* (Corpus PAC Manchester, code locuteur emangt1, conversation libre)

En (8), le déclencheur d'antécédent est l'ensemble propositionnel *y/ you get kids, sort of like, they're playing this, they're playing this street (XX) and they're all on their way to school kicking the ball on the road and stuff it's, yeah, I think, I think, especially I think, they grow up playing from, I think they're just, they're told to play.* Cet empan de texte permet le passage en mémoire discursive d'un agglomérat informationnel, que l'on pourrait qualifier de topique discursif – *i.e.* ce de quoi l'on parle. *This*, de par sa nature référentielle et sa position argumentale au sein de l'énoncé, permet de circonscrire cet ensemble d'informations en un objet-de-discours que l'on pourrait paraphraser par « playing football ; 'jouer au football' », et le place au rang de thème de l'énoncé. Le fonctionnement référentiel de *this* ici se caractérise par un envoi référentiel vers la mémoire discursive du destinataire où est situé l'agglomérat informationnel premièrement évoqué ; le retour sur le repère discursif au moment t_0 de l'énonciation permet à *this* de trouver son référent, par l'ancrage au sein de la prédication en jeu. Ainsi, nous pourrions dire que la procédure référentielle effectuée par *this* dans ce cas-ci est à la fois anaphorique – par la source référentielle présente en mémoire discursive – et déictique – par la circonscription du référent concomitante à son identification. En (9), *this* est utile pour maintenir la référence d'une entité de premier ordre, d'ores et déjà présente en mémoire focus puisqu'auparavant référée par le pronom *it* ; cette entité est le topique local de l'énoncé. Ici, le fonctionnement référentiel de *this* s'effectue par un envoi référentiel vers la mémoire focus du destinataire où se trouve le référent voulu. Selon notre définition, la

procédure référentielle réalisée est anaphorique – déduction confirmée par la possibilité de *it* à la place de *this* dans cet exemple. A la différence de *it* néanmoins, *this* semble mettre en évidence le commentaire négatif effectué par GT à propos de ce vêtement. En effet, nous pourrions aisément substituer un SN du genre *this ugly thing* au pronom ici. En d'autres termes, *this* entraîne une recatégorisation sémantique du référent en fonction de ce que GT en prédique, ce qui conforte le retour sur le repère énonciatif, dont l'origine est l'*origo*, opéré par *this*.

Les exemples (8) et (9) mettent en évidence la possibilité de *this* de réaliser l'anaphore lors de son activation en discours. Ce constat est une preuve supplémentaire qu'un appariement direct entre démonstratifs et deixis est inadéquate. Les deux occurrences de *this* ici réfèrent grâce à l'invocation du contexte discursif : en (8), *this* permet la délimitation d'un référent discret à partir d'un ensemble d'informations connues du destinataire, tandis qu'en (9), *this* maintient la référence d'une entité présente en mémoire focus. Si l'on ajoute les observations extraites des exemples (1) à (7), *this* réfère via le recours au contexte situationnel, en plus de celui au contexte discursif ; de plus, *this* est le seul pronom démonstratif à apparaître en position antéposée au déclencheur d'antécédent dans notre corpus. A présent, nous allons présenter nos résultats en faveur des contextes invoqués par *that* pour référer.

2.2. *That*

2.2.1. La référence indexicale de *that* par l'invocation des contextes, d'une part, situationnel immédiat et, d'autre part, discursif

Le fonctionnement indexical de *that* couvre les recours aux contextes effectués par *this*, à savoir le contexte spatial immédiat (exemple (10)) et le contexte discursif (exemples (11) et (12) :

(10) [MO et LC regardent et commentent des photos de famille appartenant à MO]

LC: *Who is* ***that***? MO: ***That****'s daddy.* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(11) [LC explique qu'à l'âge de vingt et un ans, elle devait obtenir des diplômes scolaires élémentaires afin de pouvoir suivre une formation dans un 'Training College']

LC: *And because I was then only about twenty-one, er, they said that I ought to go to, two-year training college or three-year training college that was just coming up. Er, but* ***that*** *meant that I had to get O levels*. (Corpus PAC Lancashire, code locuteur elalc1, conversation guidée)

(12) [PK a joué à la loterie et demande à MD ce qu'elle désirerait si le ticket s'avérait gagnant]

PK: *So I got that and got my lottery ticket*. PK: *And er*. <MD: *Are you gonna er win this week?*> PK: *Yes. So are you putting, you want some money do you?* <MD: *Yeah er.> If I win?* MD: *What do I want? You could pay for my M.O.T. if* ***that*** *goes wrong.* (Corpus PAC Lancashire, code locuteur elapk1, conversation libre)

Dans ces exemples, la description du fonctionnement référentiel de *that* est analogue à celui de *this* dans les exemples (1), (3) et (4) : un retour déictique sur le repère énonciatif s'effectue en (10), la création d'un objet-de-discours circonscrit à partir d'informations discursives est opérée en (11), grâce à la position argumentale de *that* et à un retour effectué sur le repère discursif, tandis que *that* en (12) permet le maintien anaphorique d'une entité discursive présente en mémoire focus. Notons ici qu'à la manière de *this* (cf. exemple (5)), *that* autorise l'ancrage temporel de l'énoncé ; toutefois, cet ancrage ne se fait pas au moment t_0 de l'énonciation, mais à un moment décalé :

(13) [I demande à RH s'il s'est déjà rendu à Edimbourg]

I: *Have you been to Edinburgh, have you?* RH: *No I've not, I've been, I've been to erm, Glasgow once,* <I: *Right.> but* ***that****'s,* ***that****'s the only time I've really been to Scotland, yeah, never been, I've been up, I've been up to sort of the Newcastle way*. (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

That permet ici de situer l'énoncé dans un temps passé ; le retour référentiel opéré par *that* ne se produit pas sur le repère *origo-here-now*, mais sur le repère *origo-there-then*. Par ailleurs, *this* ici ne serait pas pertinent – *#I've been to Glasgow once, but this is, this is the only time*

I've really been to Scotland. Nous pourrions d'ores et déjà conclure que *this* est associé à *here* et *now*, et *that* à *there* et *then*. Tournons-nous à présent vers l'exemple (14) :

(14) [La conversation entre GT et RH tourne autour d'un concert des Red Hot Chili Peppers qui a eu lieu à The Arena, quatre années auparavant, et de la propension qu'ont les groupes de rock à se séparer puis se reformer]

<RH: *Oh is it? I saw,> (XXXXX).* RH: *erm, Red Hot Chili Peppers there*. GT: *Oh did you,* <RH: *Yeah.> was it good?* RH: *Yeah*. GT: *Yeah I'd like to see them.* RH: *Which it u/ yeah it was pretty good. It was entertaining, it was a show more than anything else.* GT: *Yeah, yeah, are they, are they (XXXX) now or,* <RH: *I think.> or some doing s/* RH: *They're all doing different things* <GT: *Are they?> now but,* RH: *but there's, there's too much money for them, as a band so they, they won't, I don't think they'll split*. GT: *Hmm*. RH: *They j/ they do it all the time* <GT: *That's true.> no don't they? The bands, they go off and do their own* <GT: *True.> little thing,* <GT: *Yeah, yeah.> and they all come back because there's a, there's a bigger need for them then.* GT: *Yeah, that's true enough.* RH: *Everyone wants them to,* <GT: *That's true enough.> come back (XX).* GT: *Yeah, yeah.* RH: ***That** was, **that** was f/ four years ago now.* GT: *Oh was it?* RH*: It was four years ago*. GT: *Gosh*. (Corpus PAC Manchester, code locuteur emanrh1, conversation libre)

That ici réfère au « concert des Red Hot Chili Peppers » qui s'est donc produit quatre ans auparavant. Ce « concert » a déjà été introduit au sein du discours et a même été anaphoriquement maintenu par *it – was it good?* ; *it was pretty good* ; *it was a show more than anything else* – ce qui nous amène à conclure, d'une part, que *that* invoque ici le contexte discursif pour référer, mais aussi que *that* renvoie à une entité discrète, précédemment évoquée au sein du discours en tant que telle. A la différence de *this*, qui ne peut continuer la référence d'un « objet » discret qu'en procédure anaphorique, *that* peut récupérer un « objet » circonscrit si celui-ci ne fait plus partie de la mémoire focus. En effet, « le concert des Red Hot Chili Peppers » n'est plus au point culminant de l'attention lors de l'occurrence du démonstratif, puisque le topique discursif est « les raisons pour lesquelles les groupes de rock se séparent et se reforment ». La procédure indexicale réalisée par *that* ici est donc à la fois anaphorique – puisque le référent est présent en mémoire discursive – et déictique – puisque ce même référent est replacé en mémoire focus. Notons ici que ni *this* ni *it* ne sont judicieux dans cet exemple – *#this was, this was four years ago now* ; *#it was, it was four years ago*

now : *this* n'a pas cette capacité de récupération d'un référent discret pour le remettre en mémoire focus, tandis que *it* est lui utile pour le simple maintien référentiel de ce référent.

2.2.2. La référence indexicale de *that* par l'invocation des savoirs partagés et encyclopédiques

Dans cette section, nous allons constater que l'étendue contextuelle de *that* est supérieure à celle de *this* – ce qui confirme *this* comme membre marqué de la paire *this/that*. En effet, les exemples (15) et (16) montrent que *that* à la possibilité de recourir aux contextes alliant savoirs partagés et connaissances encyclopédiques :

(15) [Tout d'abord, MO explique à son amie LC les problèmes de dents qu'elle subit et qui la mènent à consulter un dentiste. La conversation part ensuite sur le voyage que LC s'apprête à effectuer]

MO: *No I'm looking forward to going to the dentist, I went, a, fortnight ago, it's this one, it's, it is a lot of filling in but he said it wants (XXXX) so, I'll have to see what he can do, cause I haven't many left.* (laughter) *I've only these left, but er, I did have a denture at the bottom* <LC: *Yes.*> *but it started to rub, I think that's what rotted these,* <LC: *Oh, probably, yes.*> *you know, (XX),* <LC: *Yes.*> *so, out they came. So I'm a bit of a bunny rabbit now.* LC: *Well you know I'm going away on the thirteenth, don't you?* <MO: *Is* ***that*** *still going?*> *Yes. She's coming back after the half hour. Er, I'm going away on the thirteenth, and.* <MO: *You are going then?*> *Oh yes, we made the decision, I rang Steven, last week I sent him an email with all the pros and cons and said 'But in the decision we'll be guided by you'. And the thing that clinched it, when I was humming and haring about whether we should go he said 'well it's, much the same as it was after September the eleventh fear', I said 'Yes, but I'm frightened about global war and being stuck out with you for a few years'. So he said 'Oh I don't think that will happen'. He said 'Sean goes to the British school'. So he said 'Well, two of Sean's classmates' grandmas are coming over from England'. So I said 'Well that clinches it', I said 'We can't let Sean's grandma be a wimp can we? So. And I thought well really if we postpone it, we're both getting older and you don't know if you'll be well and enough or if there will be a global war, so I feel you've got to take the opportunities while you can. But, because there's a lot of traffic here at Christmas time, in the close, the other*

> *day, I was mentioning that I was looking for a car park that wasn't very often used, where I could park my car while I'm away*. (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

En (15), *that* renvoie au « voyage que LC s'apprête à effectuer », qui devient le topique discursif de l'échange conversationnel subséquent. Le déclencheur d'antécédent se limite à la proposition *Well you know I'm going away on the thirteenth, don't you?*, qui n'apporte que peu d'information, en somme : « LC part quelque part, le treize du mois ». Toutefois, ces seuls éléments permettent à MO de déduire que LC réfère à ce fameux « voyage » ; en retour, LC comprend que *that* renvoie à ce même référent. Comment pourrions-nous décrire le fonctionnement indexical de *that* ici ? *That* invoque certes le discours tout juste construit – sans « le fait que LC parte le treize », *that* ne pourrait accomplir sa référence – mais surtout le savoir partagé entre les deux interlocutrices. C'est grâce aux connaissances communes que MO peut inférer que le déclencheur d'antécédent renvoie à ce « voyage », mais aussi que LC déduit sans peine que MO a bien compris ce dont il est question ; d'ailleurs, *you know* et *don't you* sont des marqueurs textuels indiquant que LC et MO partagent les mêmes connaissances à ce sujet. Ainsi, le déclencheur d'antécédent active la mémoire morte de MO, qui peut alors instancier le référent voulu par *that* ; à son tour, *that* effectue un retour indexical sur le repère énonciatif *origo*$_{\text{LC}}$*-hic-nunc* afin d'introduire ce référent au moment t_0 de l'énonciation. La particularité ici est que l'*origo* transite de MO à LC lors du mouvement référentiel rétroactif opéré par *that* ; la construction du référent visé est conjointe :

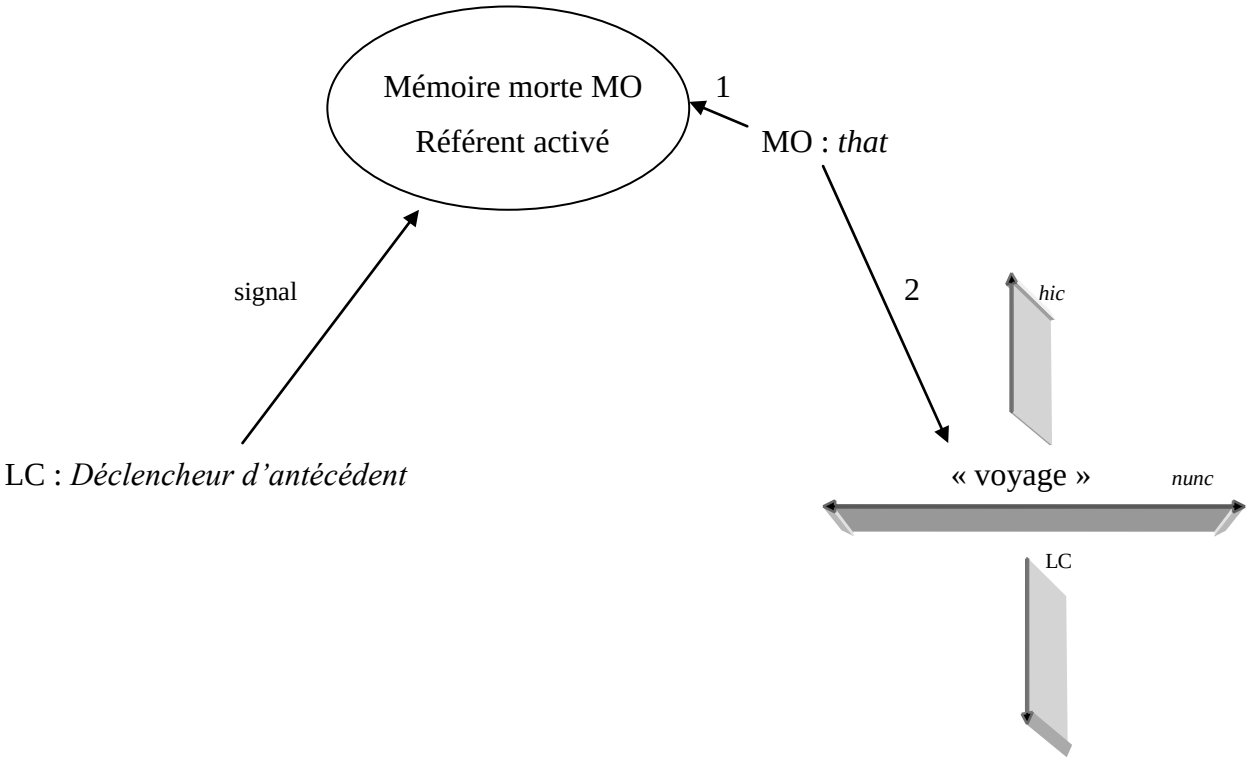


FIGURE 25 : LE FONCTIONNEMENT INDEXICAL DE *THAT* LORS DE L'ACTIVATION DES SAVOIRS ENCYCLOPEDIQUES OU PARTAGES

Le fonctionnement référentiel de *that* en (16) est analogue, à la différence que la mémoire encyclopédique est activée :

(16) [Il fut un temps où GT était intervenante de musique dans les écoles ; I, lui-même professeur des universités, approuve l'enseignement par la musique]

I: *So apparently,* <GT: *It's great.*> *you know, I'm/ I have to tell you I'm not musical and I rather regret it* <GT: *Yeah.*> *I'm very sad about that, but I know a lot of my friends teach languages through music,* <GT: *Yes, oh yes I can see how* ***that*** *works.*> *and it's apparently quite effective.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

En (16), le déclencheur d'antécédent *a lot of my friends teach languages through music* active la mémoire morte de GT, qui peut alors instancier par *that* le référent voulu, que l'on pourrait paraphraser « l'enseignement par la musique » ; *that*, à son tour, effectue un retour sur le repère discursif à l'instant t_0 afin que le destinataire I puisse accéder au référent. Comme en

(15), la construction du référent visé par *that* est conjointement effectuée par les interlocuteurs. Nous reviendrons sur cette co-construction en section 4 *infra*.

Par l'analyse des exemples (10) à (16), nous avons pu constater l'étendue contextuelle à laquelle *that* peut recourir afin de référer ; si *that* est en mesure d'invoquer l'ensemble des contextes possibles – situationnel, discursif, connaissances partagées et encyclopédiques – et, par conséquent, d'activer les trois types de mémoire, *this* se limite au recours aux contextes situationnel immédiat et discursif. Il est un paramètre supplémentaire pour lequel *that* s'avère plus extensif : le pouvoir référentiel.

2.2.3. *That* et la génération d'un nouvel objet-de-discours

Selon les données de notre corpus, lors d'un phénomène de reprise (cf. exemples (8) et (9) pour illustration), *this* est à même de renvoyer à la même entité que celle évoquée par le déclencheur d'antécédent ou bien de topicaliser un ensemble d'informations discursives, par sa nature référentielle et sa position d'argument. *That*, quant à lui, manifeste des possibilités de référence supplémentaires, que l'on pourrait qualifier de 'génératrices'. Tournons-nous alors vers les exemples (17) et (18) :

(17) [LC et MO parlent de Michael Caine, un acteur célèbre]

> LC: *He* [*i.e.* Michael Caine] *said, you know 'It's alright for people who come from money families'. But he said 'When you come from a working-class family, you've got to give it away'. So he* <MO: *You don't know what to do with.*> *said, 'I did this film'. He said, 'and* ***that*** *bought my mother's house', and I think he bought several in the family, a house with his winnings from the various films.* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(18) [I et GT débattent de la prononciation du mot 'clerk' en anglais]

> GT: (laughs) <I: *A clerk* [klɑːk] *or a clerk* [klɜːk]*?*> *Clerk* [klɑːk]*, you would say, clerk* [klɑːk] *(XXX).* I: *(X) you're posh.* GT: *Hmmm, yeah,* <I: *Rob said clerk* [klɜːk]*, for his mum, yeah.*> *is that right, did he really?* I: *She was a clerk* [klɜːk]. GT: *Ah* ***that*** (a) *surprises me I would've thought* ***that*** (b) *was American English*. (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

En (17), le déclencheur d'antécédent est *I did this film*, qui vise une entité de deuxième ordre, car spatio-temporellement située. Afin de référer, *that* s'appuie donc sur le contexte discursif et active ainsi la mémoire vive. Toutefois, le référent correspondant à *that* ici n'est pas « le film qu'a tourné Michael Caine » mais « l'argent gagné grâce à ce film ». *That*, grâce à la prédication de l'énoncé dans lequel il s'inscrit, permet de créer son propre référent. En d'autres termes, *that* signale qu'un effort déductif est nécessaire afin d'apporter des informations supplémentaires, essentielles à l'identification du référent voulu :

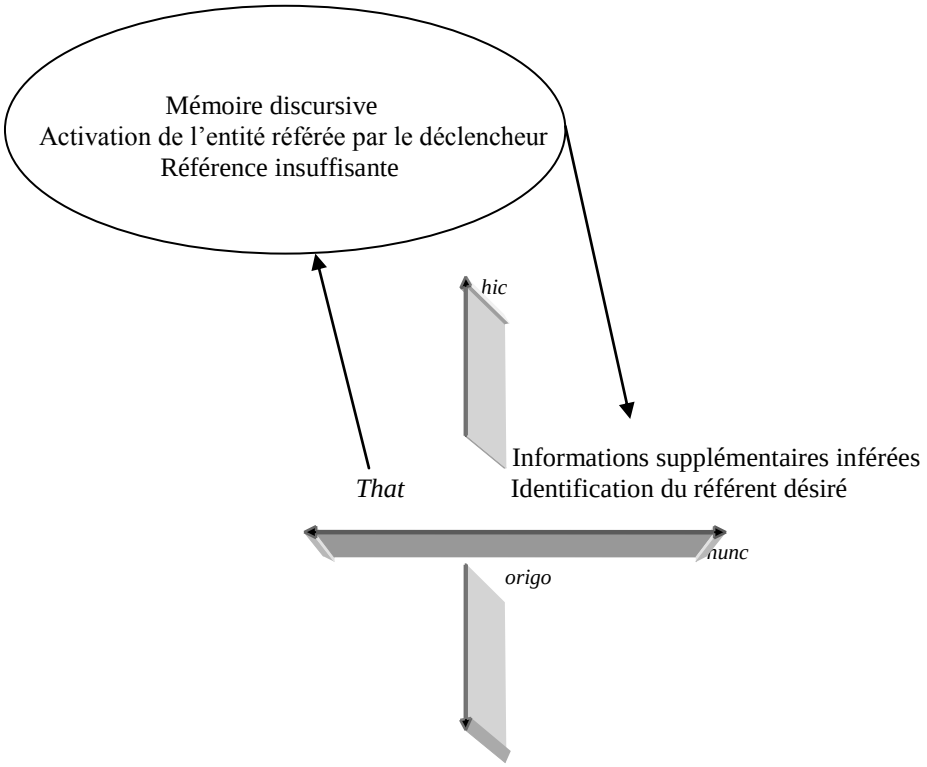


FIGURE 26 : LE FONCTIONNEMENT INDEXICAL DE *THAT* 'GENERATEUR DE REFERENCE'

En (18), deux occurrences de *that* nous intéressent. Concernant *that* (a), le déclencheur d'antécédent est *Rob said clerk* (prononcé [klɜːk]), *for his mum*. Le référent visé par *that* (a), c'est-à-dire « ce qui surprend GT », est « le fait que Rob prononce 'clerk' [klɜːk] ». Dans ce cas-ci, rien que nous n'ayons déjà vu : *that* topicalise une information discursive, à la manière de *this* dans l'exemple (8). *That* (b), en revanche, est plus intéressant. En effet, le référent visé ici n'est pas celui désigné par *that* (a) ; « ce que GT pensait être de l'anglais américain » est « [klɜːk] ». Par l'invocation du contexte discursif – « la prononciation [klɑːk] ou [klɜːk] de 'clerk' » étant le topique discursif – et par la prédication de l'énoncé, *that* signale qu'un

apport informationnel est nécessaire à l'identification du référent voulu, ce dernier relevant ici de l'ordre du métalinguistique. Les deux exemples (17) et (18) sont révélateurs de la structure interne de *that*. En effet, ces deux occurrences sont remplaçables par les SN *cet argent* et *cette prononciation*, alors que l'élément lexical de chaque SN n'apparaît pas dans le texte ; *argent* et *prononciation* sont déductibles du contexte et de la prédication de l'énoncé, et c'est le démonstratif *that* qui permet l'introduction du référent correspondant ; ainsi, nous pouvons dire que *that* est à même d'effectuer une nominalisation. Dans le cas présent, plus que pronom, *that* est créateur de SN, ce qui nous mène à considérer les pronoms démonstratifs comme étant composés d'un élément lexical sous-jacent.

Un bilan provisoire s'impose à ce stade de l'étude de *this*, *that* et *it*. En effet, les sections 2.1 et 2.2 présentent l'éventail des contextes invoqués par *this* et *that* nécessaires à la référence de ces démonstratifs, une fois activés par l'ancrage énonciatif :

	Contextes et types de mémoire correspondants						
	Environnement	Mémoire vive				Mémoire morte	
	Situationnel	Discours en amont			Discours en aval	Savoirs	
		Informations actives (topique discursif)	Topique local (référent en focus)	Référent activé (mais pas en mémoire focus)		Partagé	Encyclopédique
This	x	x	x		x		
That	x	x	x	x		x	x

TABLEAU 10 : LES CONTEXTES INVOQUES ET LES TYPES DE MEMOIRE ACTIVES PAR *THIS* ET *THAT*

Tout d'abord, l'accès à la référence de *that* peut s'effectuer via tous les contextes possibles, activant ainsi potentiellement toutes les strates de la mémoire ; *this*, en revanche, ne semble pas pouvoir activer la mémoire morte. D'un autre côté, *this* est le seul démonstratif à pouvoir apparaître en position de cataphore ; ainsi, nous déduisons que le lien étroit entre *this* et le repère énonciatif *origo-here-now* est à l'origine de l'annonce référentielle opérée (cf. exemples (6) et (7)). Ensuite, nous avons pu constater des capacités 'génératrices' de référent

de *that* (cf. exemple (17)), lorsque ce dernier procède à ce que nous avons nommé une 'nominalisation'. Enfin, nous pouvons sans appel conclure que *this* et *that* ne sont pas exclusivement déictiques : les procédures discursives réalisées par *this* et *that* dans chacun des exemples *supra* vont de la deixis à l'anaphore.

2.3. *It*

Comme il est annoncé dans la section introductive de ce chapitre, nous choisissons d'inclure *it* dans le système *this/that/it*. En effet, *it* est un élément indexical neutre ; cependant, à la différence de *this* et *that*, *it* n'est pas un démonstratif mais un clitique. Un clitique étant à mi-chemin entre le mot indépendant et le morphème lié, il y a fort à parier que la référence indexicale de *it* soit différente de celle de *this* et de *that*. C'est donc à la description du fonctionnement référentiel de *it* qu'est consacrée la présente section.

2.3.1. La référence indexicale de *it* par l'invocation du contexte situationnel immédiat

Selon les données de notre corpus, la référence de *it* via le contexte situationnel immédiat n'est pas le mode de fonctionnement principal de ce pronom, mais il existe toutefois :

(19) [Prélude à la conversation guidée entre I et RH. L'enquêteur I propose de se rapprocher du magnétophone pour que l'enregistrement soit de bonne qualité]

I: *I'll move you a little bit forward*, <RH: *Yeah*.> *if I may, because I think, I think we're not near enough the/ it's a bit difficult to*, <RH: *Yeah*.> *plan it exactly*. (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

En (19), *it* renvoie au magnétophone utile pour réaliser l'enregistrement de la conversation, dans le cadre du projet PAC. Ici, *it* n'a pas de déclencheur d'antécédent textuel ; c'est grâce à la présence du « magnétophone » dans l'environnement physique immédiat que l'allocutaire est à même d'identifier le référent voulu. Quelle est alors la différence entre l'indexicalité de *it* ici et celle opérée par *this* et *that* dans la référence aux « photos » des exemples (1) et (10), reproduits ici pour l'aisance de la lecture ?

(1) [MO et LC regardent et commentent des photos de famille appartenant à MO]

MO: *There's quite a few repeats, but you needn't look at those* <LC: *Oh don't worry, don't worry*.>. (silence) LC: *So **this** is your great grandson* <MO: *That's him*.>, *isn't that lovely? And how old is he now?* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(10) [MO et LC regardent et commentent des photos de famille appartenant à MO]

LC: *Who is **that***? MO: ***That**'s daddy.* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

Si le locuteur I n'emploie pas de démonstratif, c'est qu'il ne vise pas à montrer. Ceci se traduit par le calcul qu'effectue I sur la saillance cognitive qu'a le référent dans l'esprit de l'allocutaire. En effet, I considère que « le magnétophone » se situe dans la mémoire focus de RH, et pour cause, puisque « le magnétophone » est physiquement saillant – devant les interlocuteurs, sur la table – mais est aussi au cœur de l'événement en cours, puisqu'il s'agit d'un enregistrement formel dans le cadre d'un projet scientifique. Ainsi, à la différence de *this* et de *that*, *it* ne produit pas de retour référentiel sur le repère énonciatif ; un seul envoi référentiel vers la mémoire focus est nécessaire à l'identification du référent. Selon notre définition, la procédure indexicale réalisée par *it* est anaphorique, mais l'absence de déclencheur d'antécédent textuel associée à la dépendance au contexte physique immédiat démontre, une fois de plus, qu'il ne peut exister de correspondance immuable entre texte et anaphore.

2.3.2. La référence indexicale de *it* par l'invocation du contexte discursif

Le discours est le contexte de prédilection par lequel *it* peut référer. En effet, si l'on considère que la procédure indexicale réalisée par *it* est l'anaphore, il semble logique que la mémoire morte ne puisse être directement activée par ce pronom. Trois possibilités discursivo-référentielles émergent de l'analyse de notre corpus.

2.3.2.1. *It* et le maintien référentiel en mémoire immédiate

Comme le prédisent le modèle de Dik concernant les chaînes topicales (cf. chapitre VI, section 3) et les différentes échelles d'accessibilité (cf. chapitre IV, section 2.5), *it* permet de maintenir la référence d'un objet-de-discours en haut focus d'attention :

(20) [L'enquêteur I et la locutrice GT parlent des inconvénients et des avantages des ordinateurs]

GT: *Yeah the computer, you know, has* ***its*** *down sides but* ***it****'s also got, you know,* ***it*** *has a lot of good things*. (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

(21) [L'enquêteur I et le locuteur RH discutent cinéma]

I: *Have you seen Invictus, the film?* RH: *No I haven't but I'd like to, well, I've not seen* ***it*** *yet.* I: ***It****'s a good film. I've seen* ***it****,* <RH: *Yeah.*> *yeah,* ***it****'s good. You don't have to like rugby, I mean* ***it****'s rugby but,* <RH: *Yes* ***it****'s historical isn't it, yeah,* ***it****'s political, yeah.*> *yeah, well, Nelson Mandela really. Yeah*. RH: ***It*** *looks interesting, yeah.* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(22) [La conversation entre l'enquêtrice I et la locutrice JM tourne autour de l'ancien métier de cette dernière, qui était serveuse]

I: *Er, what other job have you done, apart from?* JM: *Apart from the personnel?* <I: *Your current job.*> *Er. while I was at school, and sixth form at university I was a waitress. Just part-time work it was just to get a bit of extra money, but I did that I've done that for years, I think my first job was when I was about, about seventeen, and, I just, had a part-time job until I went to university, and then, oh yeah I worked in a sandwich-shop as well, in, while I (XX), while I was in Newcastle, er, and then, when was it, just wh/ when I finished my master's, I worked as a waitress again, until I got my, job at the hotel, in personnel*. I: *Did you like it*? <JM: *Waitressing?*> *yeah.* <JM: *Yeah.*> JM: *I'd nev/ yeah, I'd never want to do as a career or anything, but, er,* ***it****/ I don't know, i/* ***it****'s, er,* ***it*** *makes you more confident, and,* ***it*** *makes you cope under pressure because, people think, have you, have you already done* ***it****?* (Corpus PAC Lancashire, code locuteur elajm1, conversation guidée)

Comme il est conforme à la définition de l'anaphore, *it* signale à l'allocutaire d'associer l'occurrence au référent correspondant, présent en mémoire focus. Si la tendance majoritaire est au renvoi à une entité de premier ordre (exemples (20), en référence à « l'ordinateur », et

(21), en référence à « *Invictus*, le film »), il est également possible pour *it* de maintenir la référence d'une entité de deuxième ordre (exemple (22)). Par ailleurs, cet exemple (22) est intéressant, par l'occurrence de *it* dans *Did you like it?* Ici, l'enquêtrice I souhaite faire correspondre *it* et « le métier de serveuse », mais le référent correspondant n'étant pas assez saillant, et, par conséquent, n'étant pas en mémoire focus, l'allocutaire RH souhaite confirmer le référent visé par sa question *Waitressing?* ; il y a fort à parier que RH n'aurait pas souhaité clarifier les propos de I si cette dernière avait utilisé *that* – I: *Did you like that?* RH : *Yeah* – ce qui témoigne du retour référentiel opéré par le démonstratif et du simple envoi référentiel vers la mémoire focus effectué par *it*.

2.3.2.2. *It* en tant que support de prédication

Le seul titre '*it* en tant que support de prédication' renvoie probablement le lecteur aux cas prototypiques de *it* explétif, comme dans *it's raining, how is it going?*, etc. Effectivement, le *it* explétif – ou 'dummy *it*' en anglais – est caractéristique de ce pronom, comme l'illustrent les exemples (23) et (24) :

(23) [LC explique l'histoire de Michael Caine, un acteur connu, qui n'a connu l'existence de son demi-frère handicapé qu'une fois adulte ; leur mère s'est occupée de lui toute sa vie sans révéler son existence à Michael Caine]

LC: *And **it** seems that his mother, every Monday afternoon, she used to say 'I'm going shopping', and he said she'd come back and she'd put one or two things in the fridge.* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(24) [RH est atterré par le prix exorbitant de certains vêtements, uniquement justifié par la marque du créateur]

RH: *My cousin, when he went for, for Christmas, erm, and got one of them, erm, from, a shop, in town, I can't remember which shop it was, **it** was like, **it** was seventy pounds for the tee-shirt.* (Corpus PAC Manchester, code locuteur emangt1, conversation libre)

En (23) et (24), la fonction de *it* est purement grammaticale : *it* ne renvoie à aucun objet-de-discours, sa présence est justifiée par le besoin de satisfaire la valence argumentale du verbe.

Toutefois, il est des exemples de *it* (apparemment) explétif qui semblent fonctionner de manière quelque peu différente :

(25) [L'enquêteur I et le locuteur RH parlent de la violence qui peut régner lors de matches de football]

I: *And, is there any er, violence there in football, as* ***it*** *used to be a number of years ago or is it, has it, calmed down a little bit?* RH: *Erm, you don't see it, erm, you don't see violence so much, sort of, week to week, you do, when there's erm, when there're some big games like, when erm, (XXXX) played, erm, about a month ago, in the Cup. And there was, violence there, and that was like sort of in the ground and outside the ground and,* <I: *Right.*> *even like, like throwing (X) things (X) to the (XX). But from, week to week* ***it*** *doesn't seem too bad, the, there always does seem to be, violence or when English teams go abroad and (XX),* <I: *Yeah.*> *that's when* ***it*** *tends to really flare up is it so.* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(26) [L'enquêteur I et la locutrice GT parlent des relations entre voisins]

I: *I've had a case of, two people I was interviewing in a village in France,* <GT: *Hmm.*> *and the sister was totally well integrated with the village,* <GT: *Yeah.*> *and the brother, they lived just outside, was actually, he lived there, but he didn't* <GT: *That's, men and women a bit isn't it, we're a bit like that.*> *interact at all. Yeah.* GT: *Er, I mean, w/ with the actual neighbours in my road,* ***it*** *was more just you know 'good morning', you know, 'good evening' type of thing, but I have lived, where I live for many many years, since I was six, so, I know a lot of people.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

Les occurrences de *it* en (25) et (26) ressemblent à des cas de *it* explétif ; en effet, il est difficile d'accorder une référence à *it* dans *it doesn't seem too bad, it tends to really flare up* et *it was more just you know 'good morning', 'good evening' type of thing*. Néanmoins, force est de reconnaître qu'il existe une différence sémantique entre un *it* purement grammatical, comme dans *it's raining*, et les exemples que nous venons d'énumérer. Selon nous, cette différence provient du fait que *it*, dans ces derniers cas, n'est pas totalement dénué de référence. Bien que celle-ci soit extrêmement vague, elle semble tendre vers ce que l'on

pourrait nommer « la situation » : *is there any er, violence there in football, as* ***it*** (= « the situation in football games ») *used to be a number of years ago,* ***it*** (= « the situation in football games ») *doesn't seem too bad,* ***it*** (= « the situation between GT and her neighbours ») *was more just you know 'good morning', you know, 'good evening' type of thing.* Par conséquent, *it* ici n'est pas seulement support syntaxique, il est aussi support discursif : il permet à la prédication en jeu d'être assignée à « la situation dont il est question ». Notons d'ailleurs que cette conclusion est en accord avec la reprise effectuée par RH en (25), où *there* se supplée à *it* – ***it*** *doesn't seem too bad, the,* ***there*** *always does seem to be, violence.* De plus, nous pouvons remarquer que *that* ne peut se substituer à *it* dans ces exemples : *#is there any er, violence there in football, as* ***that*** *used to be a number of years ago ; #But from, week to week* ***that*** *doesn't seem too bad, the, there always does seem to be, violence or when English teams go abroad and (XX),* <I: *Yeah.> that's when* ***that*** *tends to really flare up is it so ; #with the actual neighbours in my road,* ***that*** *was more just you know 'good morning', you know, 'good evening' type of thing.* Ceci s'explique par le trop fort pouvoir référentiel de *that* (cf. ce chapitre, section 2.2.3), qui mène le destinataire à associer *that* à un objet-de-discours, quitte à le créer par inférence. En revanche, *this* semble plus probable, à condition que le verbe soit au présent : *But from, week to week this doesn't seem too bad* ; *that's when this tends to really flare up*. Ceci s'explique par la capacité de *this* à ne servir que de point d'ancrage spatio-temporel (cf. ce chapitre, section 2.1.1) ; toutefois, si *this* renvoie à la situation *hic et nunc*, *it* ne fait qu'inscrire la prédication en jeu au sein d'un espace-temps beaucoup plus diffus.

2.3.2.3. *It* et la référence au topique discursif

A la manière de *this* et de *that* (cf. ce chapitre, sections 2.1.2.2 et 2.2.1), *it* est à même de topicaliser un ensemble d'informations discursives, grâce à sa fonction d'argument :

(27) [LC narre à MO l'histoire de Michael Caine]

> LC: *She* [*i.e.* la mère de Michael Caine]*'d had a, an illegitimate child before she met his father, and in those days, it was a big shame. So, he'd been put, I think he wasn't, er I think he was a little bit, mentally brain-damaged, I can't think what caused it, oh I know, epilepsy, he had epilepsy, and they used to put him in the cellar, and he used to lie there banging his head on the floor, and he caused himself brain damage so he's actually in a mental institution, and Michael Caine knew not one thing about* ***it*** *until she died. And it seems that his mother, every*

Monday afternoon, she used to say 'I'm going shopping', and he said she'd come back and she'd put one or two things in the fridge. And every Monday afternoon, for fifty years, she had visited that son, in the mental home. She never said a word to her husband, and never said a word to Michael Caine. And how they found out was, his mother had told him that Michael Caine was his half brother, and she put a photograph of Michael, on the wall of his room, in one of his films. And then there were others, who spoke better, than the half brother did, and, then, er, (silence) ***it*** *came out after she died, they said*. (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(28) [JM est hôtesse événementielle ; elle explique à son amie MD en quoi ce métier consiste]

JM: *I'm gonna be, working basically, we have to work, when we're there because we have to hand out vouchers for drinks, we have to make sure everyone's okay, make sure there's no trouble and things like* <MD: *Oh charming.*> *that. I know.* (laughter) MD: *Are you getting paid?* JM: *No. But,* ***it*** *kind of goes with the job really, I don't mind really.* (Corpus PAC Lancashire, code locuteur elajm1, conversation libre)

En (27), le référent visé par *it* pourrait se définir par « l'existence du demi-frère de Michael Caine » et, en (28), « les à-côtés du métier d'hôtesse événementielle », ce qui correspond dans les deux cas au topique discursif de la conversation, « ce de quoi l'on parle ». Cette fonction référentielle étant aussi assumée par *this* et *that*, quelle est alors la différence avec *it* ? :

(27') […] *and Michael Caine knew not one thing about* ***this/that*** *until she died.* […] *And then there were others, who spoke better, than the half brother did, and, then, er,* (silence) ***this/that*** *came out after she died, they said*.

(28') […] *No. But,* ***this/that*** *kind of goes with the job really, I don't mind really.*

La différence ressentie est qu'avec le démonstratif, un référent particulier se crée : par le retour référentiel opéré par le démonstratif, l'allocutaire identifie un objet-de-discours correspondant qu'il serait certainement capable de désigner par un SN ; avec l'emploi de *it*, la référence est comme inachevée, car l'entité correspondante n'est pas circonscrite. En d'autres

termes, si le démonstratif attire l'attention de l'allocutaire sur l'argument, *it* la porte sur la prédication. Ceci s'explique par le fonctionnement référentiel de *it* : puisque ce dernier est un signal de correspondance référentielle en mémoire focus, l'entité désignée est considérée comme acquise ; cependant, elle ne l'est pas, puisqu'elle n'a jamais été introduite en tant que telle au sein du discours. L'agglomérat informationnel désigné par *it* reste en l'état, diffus, sous-spécifié ; la fonction discursive de *it* est alors de servir de relais afin que la prédication en jeu puisse être appliquée à ces informations.

En conclusion de cette étude référentielle de *it*, nous pouvons établir le fait que ce pronom est à même de référer via le contexte situationnel immédiat et via le discours. *It*, étant clitique, possède un faible pouvoir référentiel : soit ce pronom maintient la référence d'une entité déjà établie (exemples (20), (21) et (22)), soit le référent n'est pas créé, ce qui se traduit par une absence de référence (exemples (23) et (24)) ou par une sous-spécification référentielle (exemples (25) à (28)). De plus, la seule procédure indexicale qu'il soit capable de réaliser est l'anaphore.

2.4. La référence indexicale de *this*, *that* et *it* : bilan

Au cours de cette section 2, nous avons observé plusieurs paramètres : les contextes par lesquels *this*, *that* et *it* permettent de référer, l'activation des différentes strates de la mémoire selon le contexte invoqué, mais aussi le parachèvement référentiel de chaque item et les procédures discursives, déictique, anaphorique ou combinatoire, que ces éléments indexicaux sont à même de réaliser. *That* est l'élément le plus complet : il permet de recourir à tous les contextes possibles, son pouvoir référentiel est tel qu'il est à même de créer un référent par inférence, et ses emplois discursifs vont de la deixis à l'anaphore. *This*, quant à lui, est plus marqué : il ne peut invoquer les savoirs partagés et encyclopédiques, mais il apparaît en position de cataphore ; contrairement à *that*, il ne semble pas propice à la création référentielle, mais couvre également le continuum entre deixis et anaphore. Enfin, *it*, pronom clitique et anaphorique, est doté d'un faible pouvoir indexical, qui, la plupart du temps, le limite au rôle de point d'ancrage fonctionnel. Nous proposons de résumer ces caractéristiques sous forme de tableaux :

	Contextes et types de mémoire correspondants						
	Environnement	Mémoire vive				Mémoire morte	
	Situationnel	Discours en amont			Discours en aval	Savoirs	
		Informations actives (topique discursif)	Topique local (référent en focus)	Référent activé (mais pas en mémoire focus)		Partagé	Encyclopédique
This	x	x	x		x		
That	x	x	x	x		x	x
It	x	x	x				

TABLEAU 11 : LES CONTEXTES INVOQUES ET LES TYPES DE MEMOIRE ACTIVES PAR *THIS*, *THAT* ET *IT*

	Procédures discursives		
	Deixis [dei]	Combinatoire [dei : ana]	Anaphore [ana]
This	x	x	x
That	x	x	x
It			x

TABLEAU 12 : LES PROCEDURES DISCURSIVES REALISEES PAR *THIS*, *THAT* ET *IT*

A présent, nous souhaitons reconsidérer les notions de distances proximale et distale attribuées à *this* et à *that*. Selon les théories, cette notion est à caractère d'abord spatial ; selon d'autres courants, la proximité et la distance prennent une dimension cognitive. D'après nous, la valeur spatiale de cette distance est dérivée de la valeur cognitive, mais pas seulement : si l'on considère l'ensemble de la structure discursive des conversations spontanées, nous remarquons que chaque indexical joue un rôle particulier dans la co-construction du discours. Voici l'analyse que nous proposons.

3. *This, that* et la notion de distance

Ceci n'est pas un fait nouveau : *this* est traditionnellement associé à la notion de proximité, *that* à celle d'éloignement. Selon les théories, cette distance est définie comme spatiale ou cognitive (cf. chapitre II, section 1). Il est vrai que la proximité de *this* et l'éloignement de *that* s'expriment à ces deux niveaux, mais nous allons voir qu'il est un troisième paramètre auquel cette distance relative s'applique. Nous en concluons par conséquent que les traits syntactico-sémantiques incorporés à *this* et à *that* doivent alors traduire toutes ces manifestations potentielles.

3.1. *This* et la proximité

3.1.1. *This* et la proximité spatio-temporelle

Nous comprenons pourquoi la notion de proximité propre à *this* – et celle d'éloignement concernant *that* – est traditionnellement associée à l'idée d'une distance spatiale, et même quasi-géographique. En effet, les démonstratifs étant alors affiliés aux déictiques et la deixis étant comprise comme une référence à l'extralinguistique, il en découle un fort lien entre les démonstratifs et l'environnement physique. Les exemples (1) et (3), que nous recopions ici pour l'aisance de la lecture, peuvent analyser *this* comme manifestant une proximité spatiale :

(1) MO: *There's quite a few repeats, but you needn't look at those* <LC: *Oh don't worry, don't worry*.>. (silence) LC: *So* ***this*** *is your great grandson* <MO: *That's him*.>, *isn't that lovely? And how old is he now?* (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(3) I: *And er, what about motor cycling? You said* <DK: *Yeah, yeah*.> *it's one of your, one of the things you do during your leisure?* DK: *Yeah, well that was sort of being at uni and don't get the chance cause I leave the bike here, er, i/ in Manchester. Er, when I'm at uni, cause it's, Edinburgh is just freezing, cold and rainy all the time so you just wouldn't get the opportunity to ride it,* <I: *Really?*> *yeah, oh it's horrid, yeah, yeah,* (laughter) ***this*** *is warm, out there is warm, compared to Edinburgh* (laughter). *It's nice to come home and get a bit of sun.* (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

En (1), *this* renvoie à la photo représentant l'arrière-petit-fils de MO, que nous pouvons imaginer être dans la main d'une des interlocutrices, ou posée sur une table devant elles, en tout cas dans une localisation proche de MO et LC. En (3), *this* est un point d'ancrage au Grand Manchester, comté dans lequel résident les deux interlocuteurs. Si nous considérons, dans la lignée des thèses localistes, que le temps est dérivé du lieu, nous pouvons ajouter les exemples (4) et (5) à la manifestation de proximité spatio-temporelle de *this* :

(4) LB: *And, from then on, I've been to college for two years doing my N.V.Q. level one and two. I've been on courses galore, erm, and* ***this*** *is where I'm now, I run my own kitchen,* <I: *So you.*> *which I love.* (Corpus PAC Lancashire, code locuteur elalb1, conversation guidée)

(5) MC: *Er, you know, because it'll take about twelve months for me to move I think, it may (X) take less, I don't know, but* ***this*** *is the time to sell the house and I want to, I really don't need a house this size now and I want to bring my overheads down Marianne, you know so*. (Corpus PAC Lancashire, code locuteur elamc1, conversation libre)

Ici, la proximité de *this* s'exprime par le *here and now* dans lequel s'inscrivent les énoncés.

Qu'en est-il alors de la conception cognitive de cette proximité ?

3.1.2. *This* et la proximité cognitive

Les thèses cognitives au sujet de la proximité de *this* se défendent par le caractère affectif que revêt la prédication de l'énoncé : l'énonciateur emploie *this* pour référer à un « objet » qu'il aime ou qui lui est proche, comme l'illustre l'exemple (29) :

(29) [LB explique à l'enquêtrice I le cheminement qu'elle a dû suivre pour occuper son poste actuel]

LB: *And then I used to be at college two till ten. On Wednesdays. So I had to get a friend in to watch the children while I went to work.* I: *It must have been difficult times going back to school then* <LB: *Yeah, it was.*> *at the same time, the children and, the job.* LB: *It was. It was very, very tiring. But, it was worth it, to me, it was worth it. I'd do it again.* I: *Because otherwise you didn't have the qualifications to be, to be a cook right?* LB: *I would have done I would have still gotten where I*

went/ but, the commercial services now do it in-house, they didn't at the time. And to actually get me NVQ level one and two in catering I had to go to college, and so, I found a job that I enjoyed doing. I found a job that fits in with my family. So I thought 'Well, this is it, I w/ I'll take ***this*** *further.* <I: *It was worth it?*> LB: *It was worth it and I'm quite happy now*. (Corpus PAC Lancashire, code locuteur elalb1, conversation guidée)

Ici, *this* renvoie au métier actuel de LB, qu'elle aime et qui correspond à son mode de vie – *a job that fits in with my family, I'm quite happy now* – ce qui conforte la dimension cognitive que peut couvrir la proximité de *this*. De même, nous pouvons considérer les exemples (4) et (5) ci-dessus comme présentant un *this* proximal au niveau cognitif : en (4), LB précise que gérer sa propre entreprise culinaire est quelque chose qu'elle adore – *which I love* – tandis qu'en (5), MC argumente vouloir vendre sa propriété par son refus d'avoir une maison si grande, ce qui prouve que « vendre sa maison » lui tient vraiment à cœur. Par convention, nous choisissons de nommer ce type d'énoncés des 'énoncés à polarité prédicative positive'. Sur l'ensemble du corpus, la manifestation spatiale de la proximité de *this* est un épiphénomène et l'exemple (5) montre que certaines analyses spatio-temporelles peuvent se ramener à la dimension cognitive ; dès lors, nous pouvons envisager la proximité physique de *this* comme étant dérivée de la prise en compte cognitive.

De plus, il est de nombreux exemples où *this* ne semble manifester ni de proximité spatio-temporelle ni d'affection particulière portée par l'énonciateur au référent voulu. Néanmoins, ces occurrences peuvent s'analyser comme indiquant un fort lien entre le référent et l'énonciateur. Nous pensons tout d'abord aux cas de *this* en position de cataphore, dont nous redonnons les exemples présentés en section 2.1.2.1 *supra* :

(6) GT: *no they don't.* <I: *Right.*> I: *And, I don't know* ***this*** *is a bit embarrassing,* <GT: *Not at all.*> *is there a significant (XX) that, you speak, interact linguistically with or? N/ hmm.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

(7) I: *And you do, do you do French and another language, I'm sorry.* <GT: *I don't, I don't know languages (XX)/ no, no, no, just at all, just linguistics.*> *You just do French. No you're just doing linguistics.* <GT: *Last year I did some French.*> *Just linguistics sorry I.* <GT: *Yeah, linguistics and that's all.*> *I got it. I knew you were doing linguistics but I thought erm.* GT: *Yes I mean there's all sorts of*

programmes <I: *But, **this** is new isn't it in Salford, you can do linguistics, er as a single, single (XX).> on that. Yes it is, yes, yes, yeah.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

Quel pourrait être le lien entre l'énonciateur et la référence de *this* dans ces cas-ci ? Si nous élargissons la définition cognitive de la proximité de *this*, celle-ci ne se résume plus à l'affection que peut porter l'énonciateur à l'objet-de-discours visé. Dès lors, cette dimension cognitive inclut tout ce qui pourrait concerner l'énonciateur, pas seulement au niveau des sentiments. Les cas de *this* en position de cataphore supposent que le référent visé ne soit accessible qu'au seul locuteur, au moment de l'occurrence (cf. section 2.1.2.1 *supra*), puisque ce *this* déictique signale à l'allocutaire que le référent n'est pas encore disponible dans sa mémoire. Nous comprenons alors aisément la proximité cognitive qui lie ici *this* et l'énonciateur. Tournons-nous à présent vers les exemples (30) (qui reproduit l'exemple (8)) et (31) :

(30) I: *Is there still a lot of that, five-a-side or, you know that's.* <RH: *Yeah*.> RH: *there is where, there is where I am, y/ you get kids, sort of like, they're playing this, they're playing this street (XX) and they're all on their way to school kicking the ball on the road and stuff it's, yeah, I think, I think, especially I think, they grow up playing from, I think they're just, they're told to play, if you know what I mean, I think **this** is something they grow up with.* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(31) [ST évoque ses souvenirs de lycée]

I: *Any teachers who influec/ who influenced you, who impressed you, a lot?* ST: *My English teacher, my English teacher, Miss Wo. She did. I remember her, erm yes, yes erm, I* <I: *So it was in primary school?> often think about. **This** was at high school.* <I: *High school*.> *I often think about her now when I'm teaching you know*. (Corpus PAC Lancashire, code locuteur elast1, conversation guidée)

Au niveau du fonctionnement référentiel, ces deux occurrences de *this* manifestent le même mode d'indexicalité : *this* topicalise un ensemble d'informations discursives déclenché par un

groupe propositionnel – *y/ you get kids, sort of like, they're playing this, they're playing this street (XX) and they're all on their way to school kicking the ball on the road and stuff it's, yeah, I think, I think, especially I think, they grow up playing from, I think they're just, they're told to play* en (30) et *My English teacher, my English teacher, Miss Wo. She did. I remember her* en (31). Comment pourrions-nous déceler l'expression de la proximité de *this* ici ? A l'évidence, nous ne pouvons recourir à la géographie spatiale, ni même à la contiguïté cognitive induite par un référent connu du seul énonciateur au moment de l'occurrence du démonstratif. D'un autre côté, l'ensemble des énoncés en (30) et (31) ne montre pas d'affection particulière entre l'énonciateur et l'« objet » visé par *this*. Quelques indices textuels laissent cependant penser qu'il existe un chemin à suivre pour mettre au jour la proximité de *this* ; nous songeons à *there is where **I** am, **I think**, **especially I think*** et ***if you know what I mean*** en (30) et à ***My** English teacher, **I remember** her* et ***I** often **think** about her now when **I**'m teaching* en (31). Ces petites marques, en particulier *I, I think* et *I remember*, indiquent que l'énonciateur, dans chaque cas, parle de quelque chose qu'il expérimente, qu'il pense ou qu'il a connu dans sa propre vie ; de même pour *if you know what I mean*, qui indique clairement que l'énonciateur développe des propos en fonction de ce que lui seul a en tête. Dans ce prolongement, nous pouvons désormais expliquer la différence de fonctionnement de *this* par rapport à *it* dans *it had all threads, hanging from the hand, like it was unfinished.* <RH: *Yeah.*> *And **this**/**it** is, (XXX) four hundred and seventy-five pounds* (exemple (9)) : *this* permet de recatégoriser le référent, au niveau sémantique, selon l'appréciation du seul énonciateur.

Si nous regardons à nouveau les exemples présentés en cette section 3.1.2, nous aboutissons à un constat frappant : aucune occurrence de *this* ne peut être remplacée par *that* sans modifier la valeur de l'énoncé – *#I'll take that further* (exemple (29)) ; *#that is a bit embarrassing* (exemple (6)) ; *#that is new isn't it in Salford* (exemple (7)) ; *#I think that is something they grow up with* (exemple (30) ; *#that was at high school* (exemple (31)). A priori, ceci ne provient pas du mode de fonctionnement référentiel, puisque nous avons vu en section 2.4 *supra* que *that* couvre à la fois les contextes possiblement invoqués par *this* et les procédures discursives réalisées par ce dernier. Il est donc probable que cette impossibilité émane de ce fameux trait proximal attribué à *this*. Notons par ailleurs que, si *that* est exclu, *it* est possible : *I'll take it further ; it is a bit embarrassing ; it is new isn't it in Salford ; it is something they grow up with ; it was at high school*. Cette observation suggère une complémentarité des traits proximal et distal ; de plus, elle permet d'expliquer pourquoi *that*

ne peut se substituer à *this* dans les cas de référence au *hic et nunc* (cf. exemples (1) à (5)) et dans les cas de *this* en position cataphorique (exemples (7) et (8)) : cette impossibilité de substitution n'est pas due à la deixis, ni au retour référentiel caractéristique de cet emploi discursif, mais à la capacité de *this* à se rapprocher de l'*origo*. Enfin, ceci indiquerait une certaine neutralité de *it* concernant la notion de distance relative. Ainsi, *this* serait un marqueur de référence interne à la sphère personnelle de l'énonciateur.

3.2. *That* et l'éloignement

3.2.1. *That* et l'éloignement spatio-temporel

Les données de notre corpus d'étude ne fournissent hélas pas de support vidéo ; il est donc très difficile de juger la proximité ou l'éloignement de l'objet servant de support référentiel à *this* ou *that*, dans un cas de référence via le contexte spatial immédiat. Néanmoins, l'exemple (10), reproduit ici, présente manifestement un tel emploi de *that* :

(10) [MO et LC regardent ensemble des photos de famille]

LC: *Who is* ***that****?* MO: ***That****'s daddy*. (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

Cet exemple est extrait de la même conversation dont est issu l'exemple (1). Pendant leur entretien, LC et MO regardent et commentent des photos de famille ; ces clichés appartenant à MO, c'est LC qui questionne MO au sujet des différentes personnes représentées, et elle emploie tour à tour *this* et *that* pour désigner les portraits en question. Comme nous le mentionnions, nous ne sommes pas en mesure de témoigner de la distance relative entre les photos et la locutrice, puisque nous manquons de support vidéo. Néanmoins, nous pouvons supposer que le jeu entre *this* et *that* dépend de la localisation plus ou moins proche des photos par rapport au centre énonciateur. En revanche, nous disposons d'exemples où *that* ne réfère pas via le contexte spatial, mais via le contexte temporel, comme l'illustre l'exemple (13), déjà présenté en section 2.2.1 *supra* :

(13) [I demande à RH s'il s'est déjà rendu à Edinbourg]

I: *Have you been to Edinburgh, have you?* RH: *No I've not, I've been, I've been to erm, Glasgow once,* <I: *Right.*> *but* ***that****'s,* ***that****'s the only time I've really been to*

Scotland, yeah, never been, I've been up, I've been up to sort of the Newcastle way. (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

Par le trait distal de *that*, nous pouvons confirmer nos intuitions soumises en section 2.2.1, lors de l'analyse du fonctionnement indexical de *that* dans cet exemple : si *this* est associé à *now*, *that* est associé à *then*, foncièrement éloigné du repère *origo-hic-nunc*.

3.2.2. *That* et l'éloignement cognitif

Dans les théories qui supportent la définition cognitive du trait distal de *that*, il est avancé que ce démonstratif réfère à un objet-de-discours qui implique une distanciation, voire un rejet, de l'énonciateur. En effet, nous disposons de nombreuses occurrences de *that* qui répondent à ce critère :

(32) [RH est étudiant en linguistique ; l'enquêteur I, lui-même linguiste, demande à RH s'il a eu l'occasion d'étudier sa propre variation de l'anglais]

I: *Right. And have you studied your own, speech in fact through, through doing linguistics have you discovered features of your own, variety of English that surprised you?* RH: *We have the, what we've been doing in, in dialects is, erm, a sort of (X), not so much in the lectures but in the seminars, we've been given sort of like wordlists,* <I: *Yeah*.> *and, because, there's people in the class that are from, from all over so,* <I: *Sure*.> *we've all been just sort of like reading at different times and stuff, not so much planned but er naturally.* I: *Yes*. RH: *Erm, we've been comparing the differences there,* <I: *Right*.> *but* ***that, that****, really* ***that****'s the farthest we've gone.* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(33) [DK explique à l'enquêtrice I les raisons pour lesquelles il n'aime pas les Ecossais]

DK: *I hate a f/ small,* <I: *Oh really?*> *group of Scottish people yeah, but. Yeah, the majority of them are okay, I do have Scottish friends up there and, they're quite like, chilled out and decent people but, there's the odd few that like, tar everybody else. You walk into the local pub, where I live, er, I ordered a pint in English, accent, didn't try and change my accent in any way and it just fell silent, the pub,*

is just completely dead, everybody turns around, 'What are you doing in our pub?' type of thing, it's like 'Okay, I'll leave', (laughter) <I: *Oh, right.> so, yeah, so. It's strange to be on the receiving end of* ***that****, because especially in England it's Asians that get, that side of like, if you've got pale skin then you/ you're fine here but, this is really weird, to receive* ***that*** *in Scotland, but.* (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

En (32), *that* renvoie à « ce que RH a étudié de la variation de l'anglais, à l'Université ». Par la prédication dans laquelle s'inscrit la référence de *that*, nous pouvons déduire que RH émet un avis plutôt négatif sur son avancée dans le domaine. En soi, l'énoncé produit par *that's the farthest we've gone* ne véhicule pas de point de vue foncièrement péjoratif, mais l'ajout de *really* ainsi que la contextualisation de la conversation – un entretien guidé au sein de projet PAC, mené par un Professeur des Universités, lui-même linguiste – laisse deviner que RH en serait presque à s'excuser de ne pas avoir plus de connaissances en la matière et de ne pouvoir répondre à la requête de I. Par convention, nous choisissons de nommer ce type d'énoncés des 'énoncés à polarité prédicative négative'. Ainsi, nous pouvons faire le lien entre la distance cognitive indiquée par *that* et la distanciation de l'énonciateur. De même en (33), où *that* renvoie à « l'œil accusateur des autochtones envers les étrangers », qui fait ressentir à DK qu'il n'est pas écossais ; ici, la distance cognitive entre l'énonciateur et le référent visé ne fait aucun doute. Ainsi, nous pouvons associer à *that* la référence à la sphère non personnelle de l'énonciateur.

Toutefois, l'observation des occurrences de *that* au sein de notre corpus révèle un fait remarquable : *that*, et *that* seulement, apparaît dans des commentaires effectués par l'interlocuteur – que nous étiquetterons B – lors du tour de parole du locuteur principal A. Voici quelques exemples, pour illustration :

(34) RH: *and he's a new (X) at nine,* <GT: *Oh no.> and it's all because he spends so much work, money <GT: Yeah, the expensive.> on the Vivienne Westwood,* <GT: *lifestyle.> that he's got to, that he's got to work to pay for it because,* <GT: ***That's crazy isn't it?****> obviously, he doesn't want to get into debt and stuff.* (Corpus PAC Manchester, code locuteur emanrh1, conversation libre)

(35) RH: *They j/ they do it all the time* <GT: ***That's true****.> no don't they?* (Corpus PAC Manchester, code locuteur emanrh1, conversation libre)

(36) RH: *and, I can tell immediately whether they're, they're from, sort of, an area around where I am. If you know what I mean,* <I: ***That's amazing***.> *yeah.* I: *Could it be expre/ ju/ i/ is it the accent or could it be just words or expressions or? I mean are there typical, Salford expressions you know things like, erm, words that might be more typical of Salford than Manchester or?* RH: *Erm. It's ha/ it's hard, I th/ I think there is a, maybe not the accent so much but maybe the, the pronunciation in certain words, that, th/ there's a, there's a distinct difference, in, in,* <I: *Right*.> *some things, definitely yeah.* I: ***That's interesting.*** (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(37) I: *I need to ask you,* <GT: *But I love it, yes sure, no problem, no problem*.> *we'll just go through the questions then talk. Y/ your, Yo/ your parents, it's, it's kind of personal,* <GT: ***Yeah, that's fine, that's fine***.> *they, they are from this area?* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

(38) I: *I've had a case of, two people I was interviewing in a village in France,* <GT: *Hmm*.> *and the sister was totally well integrated with the village,* <GT: *Yeah*.> *and the brother, they lived just outside, was actually, he lived there, but he didn't* <GT: ***That's, men and women a bit isn't it****, we're a bit like that*.> *interact at all. Yeah.* GT: *Er, I mean, w/ with the actual neighbours in my road, it was more just you know 'good morning', you know, 'good evening' type of thing, but I have lived, where I live for many many years, since I was six, so, I know a lot of people.* (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

Dans ces exemples, le trait distal de *that* ne peut être analysé comme manifestant une distanciation ou un rejet de la part de l'énonciateur : hormis peut-être *that's crazy*, aucun des intervalles en gras n'indique des propos à connotation négative. En revanche, au niveau de l'interaction, ces commentaires sont utiles à B, qui peut ainsi participer à l'intervention de A. Grâce à *that*, B peut récupérer des objets du discours générés par les paroles de A, en quelque sorte se les approprier, et produire un énoncé qui le fasse participer à l'échange conversationnel. Ainsi, la sphère non personnelle de *that* serait mieux caractérisée comme étant une sphère interpersonnelle. Cette intuition nous mène donc à l'étude plus approfondie de *this* et de *that* en tant que signaux interactionnels du discours.

3.3. La distance exprimée par *this* et *that* au niveau de l'(inter)communication

Comme nous venons de le présenter, la position majoritaire de *that* dans des commentaires brefs de l'interlocuteur B ponctuant le tour de parole du locuteur principal A (55% de nos occurrences) nous conduit à l'étude du rôle joué par *this* et *that* dans l'échange conversationnel. En effet, cette spécificité suggère que *that* permet à B de récupérer un objet du discours introduit par A afin de se l'approprier ; ainsi, B pénètre l'espace discursif du locuteur A pour le faire sien. Notons en passant que, souvent, *isn't it* clôt l'intervention de B – *that's crazy isn't it?* ; *that's men and women isn't it?* ; cet indice textuel, qui indique clairement la volonté de B de s'unir à l'espace discursif de A, contraste radicalement avec ceux entourant les occurrences de *this – I think, I remember, if you know what I mean* – qui, eux, ont pour effet de cantonner l'énoncé produit à la sphère interne du locuteur. Comment pouvons-nous alors appréhender, dans cette optique, les occurrences de *that* n'apparaissant pas dans des ponctuations discursives de B ? Tournons-nous vers les exemples (39), (40) et (41) :

(39) [RH est étudiant en linguistique, l'enquêteur I est lui-même linguiste et enseigne à l'université ; RH explique qu'il doit passer un examen de phonologie et qu'afin de se préparer, il utilise un ouvrage accompagné d'un CD-rom]

RH: *Yeah,* <I: *Yes.*> *so, I've got that due in, at the end of the month so I've been listening to my erm, <u>like a CD-rom</u>, over and over and over (XXX) down (XX).* <I: *Is* ***that****.*> I: *the Peter Roach's course or* <RH: *Yeah, yeah.*>*?* (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(40) [RH et l'enquêteur I discutent football, en particulier des fervents supporters de certaines équipes]

I: *What do they call it in er, in Liverpool, i/ is, is there k/ is there not a name for the?* RH: *Anfield*. I: *Anfield.* <RH: *Yeah.*> RH: *Erm, you get, the, the (XX) fans are in like a small corner,* <I: *Right.*> *so it's really intimidating because you've got like, the, the rest of the ground around you, and, and always there is real tension there,* <I: *Yeah.*> *yeah*. I: *<u>I meant is, in Anfield isn't there the, the s/ the s/ Liverpool supporters</u> is there* <RH: *Oh <u>the Kop</u>* (a).> *a name for the?* I: <RH: *Yeah, yeah, yeah.*> *Yeah,* ***that****'s* **(a)** *what I mean, the Kop*. RH: ***That*** **(b)** *is,* ***that*** *is*

intimidating because, (laughter) *that yeah. Yeah it is*. (Corpus PAC Manchester, code locuteur emanrh1, conversation guidée)

(41) RH: *I went to J.Lo. But there's plenty of peo/ there's plenty of people going to that, there's a/ there's all, there's all girls (XXX).* GT: *Little girls as well.* RH: *Yeah. But when er, when it (X), when we bought the tickets and stuff (XX), our mates were like 'you're not going', like 'yeah but, only because you got dragged into it'. It was my idea to go really but.* <GT: *Was it?*> GT: *So was* ***that*** *with the girlfriend that y/* <RH: *Yeah, erm.*> *yeah*. (Corpus PAC Manchester, code locuteur emanrh1, conversation libre)

Dans chacun de ces exemples, *that* n'apparaît pas dans un énoncé produit par B, ponctuant le tour de parole de A ; au contraire, *that* fait partie d'un énoncé propre à B, doté d'une réelle intention communicative qui élabore le discours en construction. Toutefois, la récupération référentielle opérée par *that* reste effective. En effet, si nous regardons de plus près chaque exemple, le déclencheur d'antécédent de *that* est irrémédiablement prononcé par A – déclencheur que nous avons souligné dans chaque exemple. En (39), *that* renvoie au « CD-rom que RH écoute en boucle pour réviser sa phonologie », que I suspecte d'être « le cours de Peter Roach » ; *that* est prononcé par I, le déclencheur *a CD-rom* par RH. En (40), le jeu des différentes occurrences de *that* est subtil : *that* (a), prononcé par I, réfère à « The Kop, le nom », fourni par RH ; *that* (b), prononcé par RH, renvoie à « The Kop, l'équipe de supporters », référent introduit par I. En (41), *that*, prononcé par GT, renvoie au « concert de J.Lo auquel RH a assisté », objet-de-discours produit par RH. Cette saisie référentielle effectuée par *that* est en parfaite complémentarité avec *this*. En effet, dans l'ensemble de notre corpus, tous les référents visés par *this* trouvent leur source dans un objet du discours préalablement évoqué par le même énonciateur. Ainsi, au niveau de la construction du discours par l'interaction conversationnelle, nous pourrions comparer *this* et *that* à des élastiques : si *this* ne peut récupérer qu'un objet-de-discours introduit par le même énonciateur, il serait un élastique tendu ; si *that* a la capacité de récupérer un objet du discours généré par les paroles de l'interlocuteur, il serait un élastique relâché. Dès lors, nous comprenons pourquoi *that* est le seul indexical à pouvoir remettre en focus un objet-de-discours non topical au moment de l'occurrence (exemple (14)), mais aussi sa capacité à recourir à la mémoire morte pour référer (exemples (15) et (16)) : son élasticité référentielle étant supérieure à celle de *this*, il peut s'étendre aux zones les plus excentrées de la mémoire.

Néanmoins, il est des exemples dans lesquels le déclencheur d'antécédent et *that* sont prononcés par le même locuteur :

(42) [GT et RH parlent de différents groupes de rock]

GT: *Yeah I love Radiohead. I really do because I find them miserable*. RH: (laughs) *Like Coldplay*. <GT: *Oh yeah*.> ***That****'s the most miserable band.* (Corpus PAC Manchester, code locuteur emangt1, conversation libre)

(43) [L'enquêtrice I questionne le locuteur DK sur les métiers qu'il a exercés]

I: *Er, you wrote that you were, a sound engineer,* <DK: *Yes*.> *how long did you do that for?* DK: *It was er, I did it as work experience from school, and then, I took it as a qualification when I left school, er as a diploma, and it was just bits and pieces really, sort of in and out of recording studios but, it's rare and there were very few jobs so, this that's why I ended up not doing it as a career. Which was unfortunate but.* (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

En (42), *that* renvoie à « Coldplay », un groupe de rock de la même famille musicale que Radiohead, les deux étant connus pour leurs chansons mélancoliques ; en (43), *that* réfère au « fait que les postes en tant qu'ingénieur du son sont rares », ce qui fait que DK n'a pas suivi ce chemin de carrière. Dans les deux cas, *that* et le déclencheur d'antécédent, que nous avons souligné dans chaque exemple, sont prononcés par le même locuteur. Ici, donc, la marque distale de *that* ne s'applique pas à quelque récupération référentielle. Toutefois, nous pouvons remarquer que chaque énoncé dans lequel s'inscrit *that* indique une distanciation appréciative de l'énonciateur : en (42), RH pense que « Coldplay est le groupe le plus misérable qui existe » – sachant que précédemment, RH critique Radiohead à cause de leur style musical qui tend vers la mélancolie ; en (43), DK explique que « le peu de débouchés en tant qu'ingénieur du son est la raison pour laquelle il a dû abandonner l'idée d'en faire son métier ». Ainsi, si la marque distale de *that* ne permet pas de récupération référentielle, elle semble marquer la distanciation cognitive de l'énonciateur. Par conséquent, nous pouvons ajouter les exemples (32) et (33) à cette analyse.

Pour résumer, la marque proximale de *this* et la marque distale de *that* peuvent se ramener aux notions de sphère personnelle et de sphère interpersonnelle élaborées par ces démonstratifs. Cette distance relative peut se traduire par une proximité/distance spatiale, temporelle, cognitive et référentielle. Dans ce dernier cas, il faut non seulement prendre en compte le type de mémoire activé mais aussi l'interaction communicative :

Marque proximale de *this*	Proximité spatiale	Sphère personnelle
	Proximité temporelle	
	Proximité cognitive	
Marque distale de *that*	Distance spatiale	Sphère interpersonnelle
	Décalage temporel	
	Distanciation cognitive	
	Récupération référentielle ➢ Par l'interlocution ➢ Par les zones excentrées de la mémoire	

TABLEAU 13 : LA SPHERE PERSONNELLE DE *THIS* ET LA SPHERE INTERPERSONNELLE DE *THAT*

Ainsi, *this* est associé au repère *origo-here-now* grâce auquel se dessine une sphère personnelle ; *that*, quant à lui, est associé au repère *sap*[92]*-there-then* grâce auquel est esquissée une sphère interpersonnelle. Comment s'imbrique alors *it* dans ce système ?

4. *This*, *that* et *it* en tant qu'indices de co-construction du discours

Jusqu'à présent, nous avons mis en évidence le fonctionnement indexical de *this*, de *that* et de *it* ainsi que la sphère personnelle et la sphère interpersonnelle générées par *this* et *that*. En ce qui concerne le fonctionnement indexical, *this* est la forme marquée de *that* : ce dernier couvre les modes d'accès à la référence du premier et peut entraîner la génération d'un nouvel objet-de-discours. *It*, lui, est typique de la référence anaphorique ou sous-spécifiée et ne sert souvent que d'outil syntaxique et discursif. Concernant les espaces discursifs de *this* et de *that*, nous avons vu que ces deux démonstratifs sont complémentaires : *this* génère une

[92] 'sap' signifie 'speech act participants'.

sphère personnelle dont le repère est *origo-here-now*, tandis que *that* élabore une sphère interpersonnelle dont le repère est *sap-there-then*.

Qu'advient-il de *it* dans le système des espaces discursifs ? La simple observation du fonctionnement indexical de ce pronom nous donne des éléments de réponse. En effet, hormis les cas où *it* est explétif, *it* renvoie à ce que nous avons nommé « la situation » (cf. exemples (25) et (26)) ; comme nous l'avons vu en section 2.3.2 *supra*, cette référence vague est due au faible pouvoir référentiel de *it*, qui, lors d'une référence au topique discursif, produit une sous-spécification sémantique et porte l'attention de l'allocutaire sur la prédication de l'énoncé, non pas sur l'argument (cf. exemples (27) et (28)). Si l'on associe ces propriétés référentielles aux traits syntactico-sémantiques de *it* – clitique, inanimé – nous pouvons envisager un espace discursif neutre pour ce pronom. Regardons plus en détail les exemples suivants :

(44) [Reproduction de l'exemple (28)]

[JM est hôtesse événementielle ; elle explique à son amie MD en quoi ce métier consiste]

JM: *I'm gonna be, working basically, we have to work, when we're there because we have to hand out vouchers for drinks, we have to make sure everyone's okay, make sure there's no trouble and things like* <MD: *Oh charming.*> *that. I know.* (laughter) MD: *Are you getting paid?* JM: *No. But,* ***it*** *kind of goes with the job really, I don't mind really.* (Corpus PAC Lancashire, code locuteur elajm1, conversation libre)

(45) [L'enquêtrice I et le locuteur DK évoquent les différentes prononciations du mot *loch*]

DK: *You wouldn't say 'loch'* [lɔχ] *in Scotland you would say 'loch* [lɔχ]'. I: *Yes some people said,* <DK: *Well, yeah.*> *'loch'* [lɔχ]. DK: *But it's yeah,* ***it*** *depends where you come from*. (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

(46) [MD a eu une amende pour s'être garée sans payer ; en fait, elle avait oublié de mettre son ticket en évidence sur le pare-brise]

MD: *So I thought 'Oh, I'm not putting any more money in' so I stuck it back on the windscreen and walked off,* (laughter) *because otherwise, what's the point?* <SC: *Yeah,> I've already got a* <SC: *for being so early, yeah> fine. Well, they'd already given me the ticket so* ***it*** *made no difference so I just stuck it back as if I hadn't been back.* (Corpus PAC Lancashire, code locuteur elasc1, conversation libre)

(47) [MD explique ses choix pour le 'A-level' (équivalent du baccalauréat) en fonction des matières qu'elle voulait étudier à l'université]

MD: *And sociology, when I looked in the prospectuses for universities, it was more social policy, and, and that's a bit,* ***it*** *didn't really interest me*. (Corpus PAC Lancashire, code locuteur elamd1, conversation guidée)

(48) [ST donne la raison pour laquelle elle a abandonné son métier à la naissance de son fils]

ST: *And then I wa/-s with them for er you know er, a couple of years and er, eventually left because you know erm, I had my little boy, and my husband didn't want me carrying on, erm, doing it, because* ***it*** *would have meant getting a nanny, for our son*. (Corpus PAC Lancashire, code locuteur elast1, conversation guidée)

Les exemples (44) à (47) confirment la neutralité discursive de *it* par la prédication de l'énoncé en jeu. En effet, la 'polarité prédicative' – pour reprendre les termes employés en sections 3.1.2. et 3.2.2 *supra* – de chaque énoncé n'est ni positive, ni négative, mais intermédiaire : ***it*** *kind of goes with the job really, I don't mind really* (exemple (44)), ***it*** *depends where you come from* (exemple (45)), ***it*** *made no difference* (exemple (46)) et ***it*** *didn't really interest me* (exemple (47)) traduisent une indifférence certaine de l'énonciateur envers le référent visé par *it*. Par convention, nous choisissons de nommer ce type d'énoncés des 'énoncés à polarité prédicative neutre'. L'exemple (48), quant à lui, exprime une non-actualisation de l'événement, par l'usage du conditionnel dans *it would have meant getting a nanny* ; de plus, cet énoncé ne concerne pas la locutrice ST, mais son mari, puisque « le fait qu'il eût fallu employer une nourrice » est la raison pour laquelle ce dernier souhaitait que ST arrête de travailler. Ainsi, nous ne pouvons inscrire cet énoncé dans une proximité ou une distanciation cognitive de la locutrice pour deux raisons : d'une part, l'événement décrit ne

s'est pas réalisé et, d'autre part, elle n'y est pas engagée de manière directe. Les exemples (49) et (50) témoignent également d'une non-implication du locuteur dans l'énoncé produit :

(49) [Reproduction de l'exemple (27)]

[LC narre à MO l'histoire de Michael Caine]

LC: *She* [*i.e.* la mère de Michael Caine]*'d had a, an illegitimate child before she met his father, and in those days, it was a big shame. So, he'd been put, I think he wasn't, er I think he was a little bit, mentally brain-damaged, I can't think what caused it, oh I know, epilepsy, he had epilepsy, and they used to put him in the cellar, and he used to lie there banging his head on the floor, and he caused himself brain damage so he's actually in a mental institution, and Michael Caine knew not one about* ***it*** *until she died. And it seems that his_mother, every Monday afternoon, she used to say 'I'm going shopping', and he said she'd come back and she'd put one or two things in the fridge. And every Monday afternoon, for fifty years, she had visited that son, in the mental home. She never said a word to her husband, and never said a word to Michael Caine. And how they found out was, his mother had told him that Michael Caine was his half brother, and she put a photograph of Michael, on the wall of his room, in one of his films. And then there were others, who spoke better, than the half brother did, and, then, er,* (silence) ***it*** *came out after she died, they said*. (Corpus PAC Lancashire, code locuteur elamo1, conversation libre)

(50) [GT explique la profession qu'exerçaient ses parents]

GT: *My father worked for the Ministry of Defense. So I'm not terribly sure of what he did.* I: *you can't,* <GT: *It's, i/> he can't tell you I presume.* <GT: *If he, if he t/>* GT*: told us they do have to have er, you know, they do have to have killed us,* <I: *Right, and.> so I don't know, and my mum worked at er, Manchester Poly, she was a career adviser. So* ***it*** *was in the days* <I: *Poly.> of the Poly*. I: *Career adviser*. GT: *Yeah*. (Corpus PAC Manchester, code locuteur emangt1, conversation guidée)

En (49), *it* renvoie au topique discursif, « l'histoire de Michael Caine » ; en (50), *it* situe « la période où la mère de GT travaillait comme conseillère de carrière » à une certaine époque

révolue. Dans les deux cas, la locutrice n'est pas actrice des événements relatés ; elle en est seulement le rapporteur. Constatons la différence avec *this* et *that* dans ces cas-là :

> (49') #[...] *and he caused himself brain damage so he's actually in a mental institution, and Michael Caine knew not one about* ***this/that*** *until she died.* […] *And then there were others, who spoke better, than the half brother did, and, then, er,* (silence) ***this/that*** *came out after she died, they said.*

> (50') #[…] *so I don't know, and my mum worked at er, Manchester Poly, she was a career adviser. So* ***this/that*** *was in the days* <I: *Poly.*> *of the Poly.*

This et *that* auraient impliqué l'énonciateur dans chaque énoncé produit, que ce soit par rapprochement ou par distanciation. Ainsi, l'inconvenance de l'un ou l'autre des démonstratifs dans ces exemples n'est pas justifiée par des raisons sémantiques – le référent n'est pas affecté par la substitution – mais par des motivations pragmatiques. *This*, *that* et *it* sont des indices d'implication du locuteur dans l'énoncé produit. Dès lors, nous pouvons associer à *it* la génération d'une sphère discursive externe.

Lorsque nous nous sommes attachée à définir les espaces discursifs suscités par *this* et par *that*, nous avons mis en exergue les indices textuels entourant l'occurrence du démonstratif ; ainsi, nous avons noté *this* souvent associé à *I think*, par exemple, et *that* à *isn't it*. En ce qui concerne *it*, l'indice textuel qui lui est adjoint est *just* : *it's just a mixed dinner* (elajm1g), *it's just not the problem* (emangt1i), *it's just the way he is* (emanrh1i), *it's just to put a little bit of thought* (elalc1i), *it's just all too much* (emangt1i), *it was just a family thing* (elamo1i), *it was just the way it worked* (elalc1g), *it just became hard work* (emangt1g), *it just made it much harder* (elamd1g), etc. Cette liste n'est pas exhaustive mais illustre parfaitement la propension à associer *it* et *just* : sur la totalité de notre corpus, nous trouvons 72 occurrences de « *it* + *just* », 9 seulement de « *that* + *just* » et 1 de « *this* + *just* ». Comment pouvons-nous interpréter cet indice textuel ? Par son sémantisme réducteur, dirons-nous, *just* atténue la valeur prédicative jusqu'à un point zéro, comme si l'énonciateur désirait tempérer l'information véhiculée par ses propos. Dès lors, nous pouvons constater la différence entre le rôle joué par les indices textuels accompagnant *this* et *that*, qui est d'appuyer la séparation ou la fusion des espaces discursifs dont le centre est l'*origo*, et celui joué par *just*, qui a pour effet de restreindre la valeur prédicative de l'énoncé. D'une part, ce constat semble confirmer la sphère discursive externe générée par *it*, puisque les interlocuteurs ne sont pas concernés par

cet espace ; d'autre part, ces réflexions sont conformes à l'emploi typiquement anaphorique de *it* : si *it* est à référence anaphorique, cela veut dire que le référent visé est hautement saillant dans la mémoire des allocutaires, ce qui implique alors que les participants à l'acte conversationnel partagent le même espace discursif. Ainsi, nous pourrions décrire la sphère discursive générée par *it* de la manière suivante :

Marque neutre de *it*	Neutralité spatiale	Sphère externe
	Neutralité temporelle	
	Espace cognitif externe	

TABLEAU 14 : LA SPHERE EXTERNE DE *IT*

Les sphères propres à chaque indexical étudié sont alors récapitulées dans le tableau suivant :

Marque proximale de *this*	Proximité spatiale	Sphère personnelle
	Proximité temporelle	
	Proximité cognitive	
Marque distale de *that*	Distance spatiale	Sphère interpersonnelle
	Décalage temporel	
	Distanciation cognitive	
	Récupération référentielle ➢ Par l'interlocution ➢ Par les zones excentrées de la mémoire	
Marque neutre de *it*	Neutralité spatiale	Sphère externe
	Neutralité temporelle	
	Espace cognitif externe	

TABLEAU 15 : LES SPHERES DISCURSIVES DE *THIS*, *THAT* ET *IT*

Quel est alors le rôle joué par les sphères propres à *this*, *that* et *it* dans la structure du discours ? Tout d'abord, il est nécessaire d'insister ici sur le fait que la structure discursive est le résultat d'une interaction communicative : le discours n'est pas construit, mais co-construit par les participants à l'acte conversationnel. De plus, nous rappelons que, dans notre définition, le discours est une représentation mentale créée par les actes de langage : le

discours est un co-construit cognitif. Dans cette optique, *this*, *that* et *it*, par les sphères qu'ils suscitent, sont des signes de maillage du discours dont les énonciateurs sont doublement à l'origine : à l'origine du repère énonciatif grâce auquel l'espace discursif s'oriente – *origo-here-now* pour *this*, *sap-there-then* pour *that*, externe-neutre pour *it* – mais aussi à l'origine créatrice. En effet, l'emploi de l'un ou l'autre de ces indexicaux est un choix effectué par l'énonciateur en fonction de sa volonté interactionnelle – impliquer ou non l'interlocuteur dans sa sphère, s'unir ou non à la sphère de l'allocutaire, exprimer une polarité positive, négative ou neutre, indiquer une proximité ou une distance spatiale, etc. Associer à *this*, *that* et *it* une sphère communicative présente de nombreux avantages :

- La sphère communicative dissocie système et usage : la faculté de générer un espace discursif orienté est propre à chaque élément indexical en tant que membre systémique ; que l'usage de cet espace s'exprime par le contexte situationnel spatial, temporel, par la cognition relève de l'usage ;
- La sphère communicative réconcilie les thèses 'textuelles' et les thèses 'cognitives' de la deixis et de l'anaphore, puisque l'extralinguistique, le texte et la saillance cognitive sont tous des paramètres pris en compte dans la description du fonctionnement indexical ;
- La deixis, l'anaphore et les procédures indexicales mixtes relèvent de l'usage, ce qui explique qu'un seul et même item puisse couvrir le continuum ; de plus, la procédure en jeu, déictique, anaphorique ou combinatoire, ne dépend pas du contexte invoqué : si tel était le cas, *it*, typiquement anaphorique, ne pourrait invoquer qu'un seul et même contexte. La procédure discursive est en fonction du calcul de l'énonciateur effectué sur l'état mémoriel de l'allocutaire. Ainsi, si la mémoire focus est activée, la procédure en jeu est anaphorique ; à l'inverse, si le référent visé n'est pas présent en mémoire, l'indexical est en emploi déictique. Entre ces deux cas extrêmes se positionne un ensemble de procédures indexicales combinant la deixis et l'anaphore.
- La sphère communicative apporte de la souplesse au modèle de Dik concernant les chaînes topicales (cf. chapitre VI, section 3), plutôt applicable à l'écrit, en raison de ses règles normatives. Dans notre modèle, ce n'est pas l'expression langagière qui est à la base du fonctionnement discursif, mais le statut du référent dans la mémoire des interlocuteurs – même si les éléments langagiers entrent en ligne de compte, bien entendu ;

- La sphère communicative permet de considérer l'interaction propre à l'échange conversationnel, hélas souvent oubliée des théoriciens.

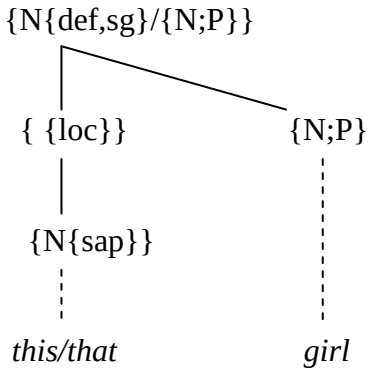
A présent, nous souhaitons proposer une représentation possible de *this*, *that* et *it* en GND (cf. chapitre V). Pour ce faire, il est nécessaire d'établir au préalable les caractéristiques minimales attribuables à chaque indexical.

5. La structure interne de *this, that* et *it* et une représentation possible des pronoms démonstratifs anglais en GND

5.1. La représentation de *this* et *that* en GND : rappels

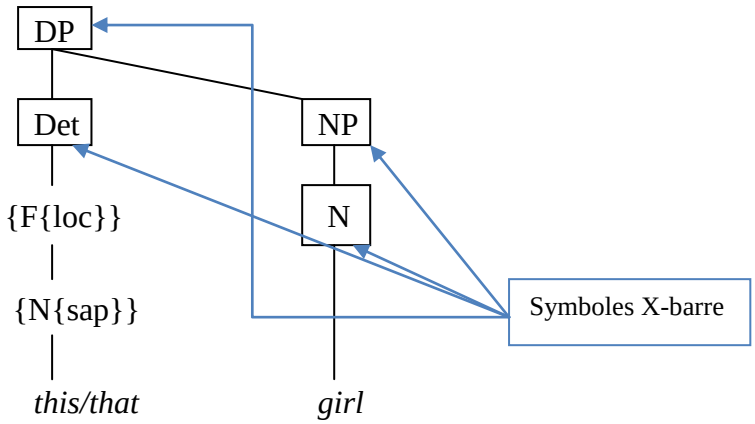
Pour rappel (cf. chapitre V, section 3.3), voici une représentation possible au sein de la GND pour la formalisation de *this* et de *that* :

(51)



Afin de faciliter la lecture de ce schéma, ramenons les notations typiques de la GND à des symbôles plus classiques, empruntés aux représentations X-barre :

(52)

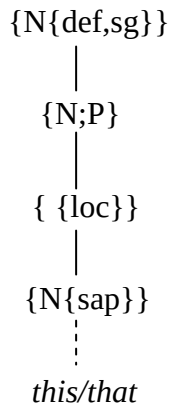


This/that girl est donc un DP ('Determiner Phrase', cf. chapitre II, section 2.3) composé du déterminant-tête et du complément nominal requis. { {loc}}, noté avec un espace entre le

crochet ouvrant '{' et '{loc}' sur le schéma en (51), signifie que nous avons affaire à une catégorie fonctionnelle spécifiée comme un type de localisation (abstraite ou concrète). On pourrait aussi l'écrire {F{loc}} en GND. En (51), {N{def,sg}/{N;P}} remplace la notation DP mais renvoie à une analyse linguistique similaire : {N{def,sg}/{N;P}} signifie qu'un SN démonstratif est une expression référentielle – 'N' signifie 'référentiel' – dont la tête syntagmatique est le déterminant. Ainsi, c'est la forme du démonstratif qui régit le nombre {sg} ou {pl} du syntagme dans son ensemble. De plus, il est indiqué par {def} qu'un démonstratif entraîne l'identification d'un référent présupposé connu du destinataire. Ensuite, { /{N;P}} traduit qu'un complément nominal est requis par la tête déterminative ; la pondération entre N et P – ici, ' ;' signifiant « N > P » – indique que le complément nominal peut-être le résultat de la nominalisation d'une proposition : le nom obtenu est à majorité 'N' puisque référentiel, mais conserve des propriétés prédicatives 'P' de par l'appareil verbal dont il est dérivé. Enfin, { {loc}} renvoie à une préposition abstraite, primitive de la sémantique, et {N{sap}} indique une localisation par rapport à l'un ou l'autre des participants à l'acte énonciatif.

Sans rentrer dans tous les détails de la notation possible des démonstratifs *per se*, proposons à ce point de la réflexion une représentation des pronoms démonstratifs :

(53)



Dans ce schéma, { /{N;P}} a disparu, puisque le complément nominal à droite de la tête déterminative n'est plus exprimé linéairement. Néanmoins, nous avons vu que les démonstratifs incorporent un élément nominal (cf. ce chapitre, section 2.2.3). Ainsi, cet élément n'est plus linéairement associé mais sous-joint, raison pour laquelle nous notons un {N;P} sous l'élément-tête {N{def,sg}} ; remarquons ici qu'un tel système élimine les catégories vides telles qu'elles existent en X-barre. A présent, nous allons affiner la description des pronoms *this*, *that* et *it* en GND.

5.2. Une représentation possible de *this, that* et *it* en GND d'après leur structure interne

Afin de préciser la représentation de *this, that* et *it* en GND, il est nécessaire de distinguer ce qui fait partie des caractéristiques inhérentes à ces indexicaux de ce qui relève de leur usage en discours. Ainsi, selon les conclusions de nos analyses, nous pouvons associer aux traits fondamentaux de *this, that* et *it* les paramètres suivants :

- Au niveau du fonctionnement indexical : *this* et *that* opèrent par une opération référentielle double. En effet, l'occurrence du démonstratif provoque un premier envoi référentiel vers la mémoire de l'allocutaire puis un second vers le repère *origo-here-now* pour *this*, *sap-there-then* pour *that* (*sap* pouvant être l'*origo*). Le repère associé à chaque démonstratif va de pair avec la sphère personnelle ou interpersonnelle considérée. *It* non explétif, quant à lui, procède par un envoi référentiel unique, dont l'objet-de-discours correspondant est présent en mémoire focus ; ce référent est souvent sous-spécifié.
- *This* est intrinsèquement relié à l'*origo*, tandis que *that* est en rapport avec les participants à l'acte communicatif (*origo* étant inclus). *It*, quant à lui, est externe à ces points d'origine.
- *This* et *that* couvrent le continuum des procédures indexicales, tandis que *it* non explétif n'est qu'anaphorique. Par conséquent, *this* et *that* peuvent être déictiques mais ne le sont pas foncièrement. Cette observation nous mène à considérer la structure interne de *this* et de *that* ; selon nous, les démonstratifs ont une composante à valeur déictique – ce qui fait que ces items sont des démonstratifs, justement – plus ou moins exprimée lors de l'activation en discours. Si cette valeur est pleinement activée, le démonstratif est en emploi déictique ; si cette valeur est moindre, la procédure en jeu est l'anaphore. Toutefois, la valeur en discours de cette composante déictique n'est jamais nulle. En effet, *a minima*, *this* indique un repérage déictique *origo-here-now* et *that* un repérage déictique *sap-there-then*. Dès lors, nous pouvons affiner les correspondances possibles entre éléments indexicaux et procédures discursives de la manière suivante :

	Procédure indexicale associée à l'acte référentiel					Référence nulle ➔ absence de procédure indexicale
	Deixis	Combinaison de deixis et d'anaphore			Anaphore	
This	{dei}[93]	{dei ; ana}[94]	{dei : ana}[95]	{ana ; dei}[96]		
That	{dei}[97]	{dei ; ana}[98]	{dei : ana}[99]	{ana ; dei}[100]		
It non explétif					{ana}	
It explétif						{ }

TABLEAU 16 : LES PROCEDURES INDEXICALES REALISABLES PAR *THIS*, *THAT* ET *IT* EN DISCOURS

- *This*, *that* et *it* étant à même de topicaliser un ensemble d'informations discursives, un élément {N;P} leur est incorporé.
- *This*, *that* et *it* ne pouvant renvoyer à des objets-de-discours dotés du trait [+ humain], mais pouvant toutefois référer à des êtres vivants, nous assignons au {N;P} en question le trait {neuter}[101].

Voici alors la représentation que nous proposons pour chacun de ces indexicaux :

[93] Par exemple, cf. énoncés en (1), (2), (4), (5), (6) et (7).

[94] Par exemple, cf. énoncé en (3).

[95] Par exemple, cf. énoncé en (8).

[96] Par exemple, cf. énoncé en (9).

[97] Par exemple, cf. énoncés en (10) et (12).

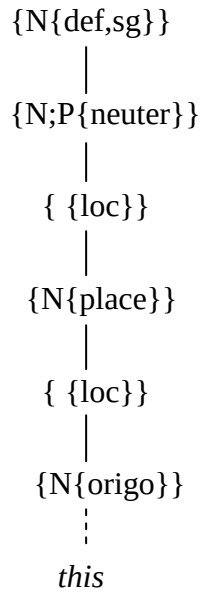
[98] Par exemple, cf. énoncé en (15).

[99] Par exemple, cf. énoncé en (14).

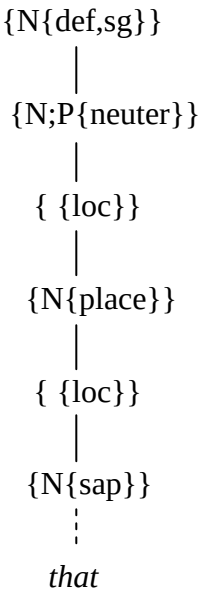
[100] Par exemple, cf. énoncé en (12).

[101] Les notations utilisées dans nos formalisations en GND sont en anglais lors d'une analyse de l'anglais, en français lors d'une analyse du français. Ainsi, nous notons {neuter} pour caractériser *this*, *that* et *it*, nous noterons {neutre} au chapitre VIII pour caractériser les démonstratifs neutres en français ; de même, nous employons ici 'sap' pour 'speech act participants' et utiliserons 'pap' pour 'participants à l'acte de parole' au chapitre VIII.

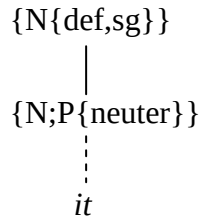
(54)



(55)



(56)



Sur ces schémas, { {loc}} {N{place}} et { {loc}} {N{origo/sap}} représentent la double opération référentielle opérée par *this* et *that* : { {loc}} {N{place}} indique le premier pointage référentiel vers une localisation mémorielle ; { {loc}} {N{origo/sap}} indique le retour référentiel sur le repère d'origine *origo* pour *this*, *sap* pour *that*, nécessaire à l'identification du référent. Nous remarquons alors que *it* ne comporte aucune marque de localisation. En effet, le référent voulu est soit en mémoire focus – le pointage référentiel est à son minimum, puisque l'objet-de-discours visé est hautement saillant – soit sous-spécifié – pour les cas de renvoi vers un ensemble d'informations discursives – soit nul – pour les cas de *it* explétif. La représentation de *it* se rapproche alors de celles de *he* et de *she* (cf. chapitre V, section 3.2), à la différence que le trait {neuter} est alors assigné à {N;P}.

Les représentations en (54), (55) et (56) correspondent aux propriétés syntactico-sémantiques de *this*, *that* et *it* en tant que membres systémiques de la langue. Que faire alors de la procédure indexicale acquise par l'un ou l'autre de ces éléments en discours ? Est-il possible de représenter la deixis et l'anaphore sur ce type de représentation ? La question n'apporte certainement pas de réponse définitive, mais, selon nous, il serait incorrect de noter {dei}, {ana}, {ana ; dei}, {dei ; ana} ou {dei : ana} sur ces schémas. Pourquoi ? Parce que la

procédure indexicale en jeu résulte d'un mécanisme très complexe, mêlant tout un ensemble de paramètres *ad hoc*, tels que le contexte énonciatif, les valeurs perlocutoires et illocutoires des énoncés, les paramètres praxéologiques, etc. Pour que la valeur plus ou moins déictique des indexicaux soit possiblement représentée, il faudrait mettre en place un système d'annotation qui prenne en compte l'ensemble de ces facteurs. Or, la GND, tout comme les autres systèmes représentationnels, est une formalisation des énoncés-types et non pas des énoncés idiosyncrasiques. Nous choisissons donc de reconnaître que la GND a ses limites, puisque tout ne peut pas être représenté dans ce cadre, tout en présentant des avantages notoires à l'analyse des propriétés syntaxiques, sémantiques et fonctionnelles des éléments langagiers (cf. chapitre V pour une présentation détaillée de la GND et de ses apports à l'analyse linguistique).

Ce chapitre explore la référence indexicale de *this*, de *that* et de *it*, notamment par la description fine du mécanisme opéré par ces indexicaux lors d'un accès à la référence via les différents contextes, mais aussi par la mise en exergue des sphères cognitives et discursives propres à ces items. Ainsi, *this*, *that* et *it* sont non seulement des dispositifs référentiels, mais aussi des indices de co-construction du discours grâce auxquels les interlocuteurs tissent un maillage référentiel par la prise en compte de chaque participant à l'acte de parole. Enfin, une représentation de chacun de ces éléments au sein de la GND permet de rendre compte des propriétés syntactico-sémantiques des pronoms démonstratifs en anglais. Le chapitre suivant s'attache à décrire le fonctionnement indexical des pronoms démonstratifs en français.

Chapitre VIII. Retour sur les démonstratifs français : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle

Ce chapitre est consacré à l'étude de la référence indexicale des pronoms démonstratifs en français. *Ça* étant majoritairement employé à l'oral (à 99% selon les données de notre corpus), notre analyse se concentre sur le fonctionnement de ce pronom neutre, bien que nous l'inscrivions au sein du système des pronoms démonstratifs dans son ensemble. Les particularités référentielles de *ça* font l'objet de bon nombre d'études (cf. par exemple Cadiot (1988), Carlier (1996), Guérin (2007) et Maillard (1987)), notamment en ce qui concerne la diffusion sémantique dont il semble faire preuve lors de reprises anaphoriques. Notre dessein est alors de mettre en rapport l'ensemble des phénomènes observés en discours afin de ramener ces manifestations discursives à deux traits minimaux propres à *ça* : la double opération référentielle caractéristique des démonstratifs et la neutralité.

Dans un premier temps, nous rappelons et précisons les propriétés des pronoms démonstratifs présentées dans le chapitre III ; dans un deuxième temps, nous mettons en évidence la variété des contextes grâce auxquels *ça* accomplit sa référence et proposons une analyse du fonctionnement référentiel de *ça* à partir de données orales issues du corpus PFC ; dans un troisième temps, nous offrons une étude pragmatique de *ça*, en fonction du rôle joué par ce démonstratif neutre dans la co-construction du discours ; enfin, nous dégageons la structure syntactico-sémantique interne de *ça* avant de proposer une formalisation dans le cadre de la GND.

1. Les pronoms démonstratifs français : rappels et précisions

La présente section fait écho au chapitre III, en ce qu'elle rappelle et affine les propriétés de *celui-ci* et de *celui-là*, de *ceci* et de *cela*, de *ce* et de *ça*. En effet, il est selon nous indispensable de (ré-)inscrire le pronom neutre *ça* au sein du système indexical démonstratif dans son ensemble afin de mieux cerner les particularité de cet item.

1.1. *Celui-ci* et *celui-là, ceci* et *cela, ça*

Celui-ci, celui-là, ceci, cela et *ça* sont tous des pronoms démonstratifs qui fonctionnent seuls en tant qu'argument. Au chapitre III, section 1.1, nous avions déduit, à

partir de travaux remontant à Dubois (1965) et Chevalier *et al*. (1964), que *celui-ci* et *celui-là* tendaient à référer à des « objets » stables, de premier ordre, tandis que *ceci*, *cela* et *ça* pouvaient renvoyer à des entités événementielles (deuxième ordre) ou factuelles (troisième ordre). Cette conclusion est l'arbre qui cache la forêt : en effet, *celui-ci* et *celui-là* sont à même de désigner des entités de deuxième (exemple (1)) et troisième ordres (exemple (2) ; (1) et (2) forgés) :

(1) *De tous ces événements, c'est* ***celui-ci*** *qui m'a le plus marqué.*
(2) *L'amour, la haine et la passion sont des sentiments extrêmement proches ; néanmoins, l'amour est selon moi le plus complexe. Quel poison violent* ***celui-là*** *!*

Toutefois, il reste vrai que le peu d'occurrences de *celui-ci* et de *celui-là* qui ressortissent de notre corpus sont en référence à des entités de premier ordre :

(3) [DP et l'enquêtrice regardent des photos de famille de DP]

E : *Eh ben c'est des beaux enfants. Ils ont quel âge maintenant ?* DP : *Ben dites,* ***celle-là****, a soixante ans aujourd'hui.* <E : *Oh la la.*> *Gérard, je, Norbert, a cinquante-trois ans,* ***celui-là*** *cinquante, non,* ***celui-là*** *cinquante-cinq, cinquante-trois, et cinquante.* E : *Ah oui donc il est né en c/ en quarante-neuf comme moi.* DP : *En quarante-neuf.* <DP : *En quarante-neuf.*> *Hum, d'accord. Et lesquels habitent à Domfront ?* <DP : *Hein ?*> E : *Qui c'est qui habite à Domfront ?* <DP : *Annie.*> DP : *qu/ Annie, bon,* ***celui-là*** *habite à Domfront, qui travaille à la ville, qui est comme mécanicien à la ville, et puis Annie qui est euh coiffeuse dans la, à Domfront dans la rue, dans la grande rue*. (Corpus PFC Domfront, code locuteur 61adp1, conversation guidée)

(4) [LH évoque les bars et cafés de Domfront]

LH : *Euh non* <E : *Non.*>. *Euh non, non, le Bar des Sports c'est euh, à côté du parking en bas, quatre-vingt places là, enfin, dans la rue du Maréchal Foch. C'est pas terrible non plus hein, c'est ambiance euh foot et compagnie, c'est pas top. Voilà. Sinon il y a l'Etrier euh, le, le bar euh du samedi soir, moi j'y, j'y vais jamais, j'y suis jamais allée, à ce qu'il paraît* <E : *C'est où ça l'Etrier ?*> *c'est très bon. C'est à côté du lycée là.* E : *C'est vrai que c'est* ***celui-là*** *qui est près, (xxx).*

<LH : *A côté de la s*/> LH : *En face de la salle multisports*. (Corpus PFC Domfront, code locuteur 61alh1, conversation guidée)

En (3), *celui-là* et sa variante *celle-là* désignent tour à tour « les différents enfants de la locutrice », tandis qu'en (4), *celui-là* renvoie au bar « l'Etrier ». Nous notons par ailleurs que *celui-ci* peut être à référence non spécifique :

(5) [CG se rappelle d'un échange de timbres effectué dans son enfance]

CG : *Et puis euh, tu te rappelles on était au, on avait été au château et puis, un jour où il faisait beau, et on, on avait étalé tous les timbres*. <ML : *Ah oui*.> *Tu te rappelles pas ? On avait étalé tous les timbres et puis on s'était, parce qu'on faisait toutes la collection* <ML : *Ouais*.> *à peu près, et on se les était euh* <ML : *Echangés oui*.>, *échangés euh, 'toi tu prends* ***celui-là****, moi je prends* ***celui-là****', chacun notre tour on en prenait un*. (Corpus PFC Domfront, code locuteur 61acg1, conversation libre)

(6) [JH exprime son désaccord envers l'existence de classes de perfectionnement pour les élèves en difficulté]

JH : *Mais, j'ai d/ j'ai bien observé depuis, Lucie à la maternelle, j'ai suivi, cas par cas, et euh bon si vous voulez c'était pas mon affaire de, de regarder euh, l'enfant de* ***celle-ci*** *l'enfant de* ***celui-là****, mais, j'ai commencé à faire euh, ça a fait entonnoir avec ce que j'avais vécu*. (Corpus PFC Domfront, code locuteur 61ajh1, conversation guidée)

En (5), *celui-là* désigne « un timbre quelconque parmi l'ensemble de timbres » et, en (6), *celle-ci* et *celui-là* renvoient à « un parent d'élève quelconque parmi l'ensemble de parents d'élève ». Toutes ces occurrences révèlent que *celui-ci* implique la distinction de l'entité visée par rapport à un ensemble d'entités de même nature : « un événement particulier parmi un ensemble d'événements » en (1), « l'amour parmi l'ensemble des sentiments humains » en (2), « chaque enfant de la locutrice » en (3), « un bar parmi les bars de Domfront » en (4), « un timbre parmi les timbres » en (5) et « un parent parmi l'ensemble des parents d'élève » en (6). De plus, le point commun à ces occurrences est que le référent spécifique désigné par *celui-ci* est déjà 'nommé' par un élément lexical antérieur : *ces événements* en (1), *l'amour* en

(2), *mes enfants* en (3) et *l'Etrier* en (4). Toutefois, *celui-ci* pouvant référer grâce au seul contexte situationnel – [une caissière A à sa cliente, au moment de payer la robe choisie] A : *Vous prendrez celle-ci ?* – mais pouvant également être à référence non spécifique – comme dans les exemples (5) et (6) –, nous dirons que le référent est 'circonscrit'. L'ensemble de ces remarques corrobore l'analyse de Dubois & Dubois-Charlier (1970 ; cf. chapitre III, section 2.2) qui incorpore à *celui-ci* un élément défini ; si défini il y a, ceci implique que le référent visé est identifiable et, par conséquent, circonscrit. Dès lors, nous comprenons l'inclination de *celui-ci* à référer à des entités de premier ordre : dans notre monde réel, les objets circonscrits sont en majorité assimilés à des choses concrètes, puisque palpables. De plus, nous expliquons alors l'impossibilité qu'a *celui-ci* de renvoyer à un objet-de-discours dont le déclencheur d'antécédent n'est pas nominalisé :

(7) **Aimer quelqu'un est le corollaire de détester quelqu'un ; néanmoins, selon moi, aimer quelqu'un est très complexe. Quel poison violent celui-là !*

Aimer quelqu'un et *amour* renvoient sans nul doute à des concepts extrêmement similaires, pour ne pas dire identiques ; néanmoins, la forme prédicative *aimer quelqu'un* fait barrage à l'emploi de *celui-ci* car, ainsi, l'objet-de-discours généré est à l'état d'information et non de référent ; en termes de GND et d'autres cadres, nous pourrions dire qu'*aimer quelqu'un* est fondamentalement une prédication et donc à dominance {P}) ; en revanche, le terme *amour* fonctionne comme un substantif à dominance {N} – même s'il est déverbal et inclut une prédication sous-jacente. *Celui-ci* ne peut renvoyer qu'à des entités dont le déclencheur textuel est à dominante {N}.

En ce qui concerne *ceci* et *cela*, la contrainte du déclencheur d'antécédent nominal disparaît (le déclencheur est souligné) :

(8) TH : *Comment il se comporte devant l'empereur des Gourmantchés ? Comment il se comporte ?* TI : <u>*Il rampe, et même, il est courbé sur le sol*</u>. <BTH : *Il rampe.*> <AB : *C'est pas possible.*> TH : *Hein ?* <TI : *Oui.*> <AB : ***Cela** est impossible.*> (Corpus PFC Burkina Faso, code locuteur bfath1, conversation libre)

(9) AS : <u>*Donc par conséquent, à l'inverse de moi qui par exemple suis bien, ou sur la plage, parce que farniente, c'est tout à fait dans mes cordes, ou sous l'eau, parce que, plongée, (XXX) et cetera*</u>, <E : *Ah sous l'eau ça me plait aussi hein*> <u>*vous, vous êtes bien sous l'eau mais vous êtes bien aussi sur l'eau* <E : *Sur l'eau, oui*</u>

> *c'est ça>, et etcetera.* E : *Je suis un poisson moi hein.* AS : *Voilà.* E : *D'ailleurs* <AS : *Donc> je suis du mois de mars hein vous le savez* <AS : *oui* (rire)> AS : (en riant) *Donc par conséquent vous êtes bien. Bon* ***ceci*** *mis à part, en, matière de, de quoi parlerons-nous ?* (Corpus PFC Aix-Marseille, code locuteur 13adg1, conversation libre)

En (8), *cela* renvoie à « la manière dont il se comporte devant l'empereur des Gourmantchés », tandis qu'en (9), *ceci* topicalise l'ensemble du sujet sur lequel porte l'échange conversationnel précédant l'occurrence du démonstratif. Dans les deux cas, le déclencheur d'antécédent est sous forme propositionnelle. Toutefois, cette constatation ne signifie pas que *ceci* et *cela* soient incapables de référer à une entité circonscrite ; preuve en est la possibilité qu'ont ces démonstratifs de désigner un « objet » stable via le contexte situationnel immédiat. En revanche, nous constatons l'impossibilité de *ceci* et de *cela* en lieu et place des occurrences de *celui-ci* :

(1') *#De tous ces événements, c'est* ***ceci/cela*** *qui m'a le plus marqué.*

(2') **L'amour, la haine et la passion sont des sentiments extrêmement proches ; néanmoins, l'amour est selon moi le plus complexe. Quel poison violent* ***ceci/cela*** *!*

(3') **E : Eh ben c'est des beaux enfants. Ils ont quel âge maintenant ?* DP : *Ben dites,* ***ceci/cela****, a soixante ans aujourd'hui.* <E : *Oh la la.> Gérard, je, Norbert, a cinquante-trois ans,* ***ceci/cela*** *cinquante, non,* ***ceci/cela*** *cinquante-cinq, cinquante-trois, et cinquante.*

(4') **LH : <E : C'est où ça l'Etrier ?> c'est très bon. C'est à côté du lycée là.* E : *C'est vrai que c'est* ***ceci/cela*** *qui est près, (xxx).*

(5') **CG : On avait étalé tous les timbres et puis on s'était, parce qu'on faisait toutes la collection* <ML : *Ouais.> à peu près, et on se les était euh* <ML : *Echangés oui.>, échangés euh, 'toi tu prends* ***ceci/cela****, moi je prends* ***ceci/cela'****, chacun notre tour on en prenait un.*

(6') **JH : Mais, j'ai d/ j'ai bien observé depuis, Lucie à la maternelle, j'ai suivi, cas par cas, et euh bon si vous voulez c'était pas mon affaire de, de regarder euh, l'enfant de* ***ceci/cela*** *l'enfant de* ***ceci/cela****, mais, j'ai commencé à faire euh, ça a fait entonnoir avec ce que j'avais vécu.*

Selon nous, l'impossibilité de *ceci/cela* dans ces exemples est moins due à la circonscription du référent évoqué qu'à la sélection référentielle uniquement opérée par *celui-ci* ; preuve en est la simple inadéquation de l'un ou l'autre des pronoms neutres en (1'). Par ailleurs, nous pouvons remarquer que *ceci* et *cela* ne renvoient pas à des entités dotées du trait [+ humain] (cf. exemples (3') et (6')). Cette dernière remarque pourrait paraître évidente, *ceci* et *cela* étant des pronoms neutres ; néanmoins, il est utile de le souligner en vue de la formalisation en GND que nous proposerons en section 5 *infra*. Nous ne nous étendons pas plus sur *celui-ci, ceci* et *cela* car l'analyse de Dubois (1965), présentée au chapitre III, suffit à nos desseins.

Qu'en est-il de *ça* dans ce système ? Bien évidemment, tout l'enjeu de ce chapitre est de le découvrir. Toutefois, nous pouvons introduire notre étude par l'observation des exemples ci-dessus modifiés de la manière suivante :

(1'') *De tous ces événements c'est* ***ça*** *qui m'a le plus marqué.*

(2'') *L'amour, la haine et la passion sont des sentiments extrêmement proches ; néanmoins, l'amour est selon moi le plus complexe. Quel poison violent* ***ça*** *!*

(3'') *E : Eh ben c'est des beaux enfants. Ils ont quel âge maintenant ?* DP : *Ben dites,* ***ça****, a soixante ans aujourd'hui.* <E : *Oh la la.> Gérard, je, Norbert, a cinquante-trois ans,* ***ceci/cela*** *cinquante, non,* ***ça*** *cinquante-cinq, cinquante-trois, et cinquante.*

(4'') LH : <E : *C'est où ça l'Etrier ?> c'est très bon. C'est à côté du lycée là.* E : *C'est vrai que c'est* ***ça*** *qui est près, (xxx).*

(5'') CG : *On avait étalé tous les timbres et puis on s'était, parce qu'on faisait toutes la collection* <ML : *Ouais.> à peu près, et on se les était euh* <ML : *Echangés oui.>, échangés euh, 'toi tu prends* ***ça****, moi je prends* ***ça'****, chacun notre tour on en prenait un.*

(6'') *JH : Mais, j'ai d/ j'ai bien observé depuis, Lucie à la maternelle, j'ai suivi, cas par cas, et euh bon si vous voulez c'était pas mon affaire de, de regarder euh, l'enfant de* ***ça*** *l'enfant de* ***ça****, mais, j'ai commencé à faire euh, ça a fait entonnoir avec ce que j'avais vécu.*

Les seuls cas où *ça* est impossible est lors d'une référence à une entité dotée du trait [+ humain] en dehors d'exemples que nous considérons plus loin[102] ; les autres exemples modifiés sont tout à fait acceptables, et nous paraissent naturels à l'oral. Pour cette raison, *ça* est traditionnellement considéré comme étant la forme orale, voire familière, de *cela*. Cet héritage prescriptif fait qu'un phénomène d'hypercorrection peut apparaître chez les locuteurs/scripteurs, lors d'entretiens formels ou d'exercices d'écriture ; ainsi, *cela* peut remplacer *ça* de manière (quasi)systématique, sans tenir compte des propriétés inhérentes de chacun. Cependant, il serait erroné de ne différencier *cela* et *ça* que par le registre de leur emploi. En effet, leur distribution n'est pas identique : l'expression *cela dit* n'alternera jamais avec **ça dit* et, à l'inverse, *ça va* ne deviendra jamais **cela va*. De plus, lors d'un phénomène de reprise, le rapport que les référents de *cela* et *ça* entretiennent avec l'entité introduite au préalable diffère. Considérons pour cela les exemples suivants :

(10) [Interview de Manolo Preciado, entraîneur sportif du Sporting Gijon, qui doit affronter le Real Madrid]

E : *Considérez-vous le Real comme l'équipe la plus en forme du moment ?*
MP : *Beaucoup de gens me disent qu'on affronte l'équipe la plus en forme du monde et ils ont raison. Ce sont les leaders de la Liga, ils survolent la Ligue des Champions, mais on est en novembre… **Ça** ne veut rien dire*. (www.sofoot.com, 17/01/2012)

(10') E : *Considérez-vous le Real comme l'équipe la plus en forme du moment ?*
MP : *Beaucoup de gens me disent qu'on affronte l'équipe la plus en forme du monde et ils ont raison. Ce sont les leaders de la Liga, ils survolent la Ligue des Champions, mais on est en novembre… **Cela** ne veut rien dire.*

En (10), *ça ne veut rien dire* sous-entend « ce n'est pas parce que le Real est la meilleure équipe du moment que l'on ne peut pas gagner ». En revanche, en (10'), *cela ne veut rien dire* signifierait plutôt « les paroles qui viennent d'être prononcées n'ont aucun sens ». En d'autres termes, *ça* semble élargir le référent premièrement introduit par le déclencheur d'antécédent, tandis que *cela* paraît strictement se rapporter à l'énoncé qui précède. Bien entendu, nous

[102] En effet, nous verrons en section 2.2.2 *infra* que *ça* peut renvoyer à une entité dotée du trait [+humain], notamment dans les constructions à double sujet.

reviendrons sur ce point lors de l'étude approfondie de *ça*, à partir de données issues du corpus PFC.

1.2. *Ça* et *ce*

Comme nous l'avons déjà souligné dans le chapitre III, *ce* est en distribution complémentaire avec les pronoms *ceci*, *cela* et *ça*, mais partage certaines propriétés paradigmatiques avec *celui*, *celle(s)* et *ceux*. En effet, *ceci*, *cela* et *ça* ne peuvent se substituer à *ce* : par exemple, aucun de ces trois éléments ne sera suivi d'une relative restrictive, c'est-à-dire indispensable à l'identification du référent visé (nous reprenons ici les exemples du chapitre III, section 1.4.3) :

(11) **Il n'a jamais su* ***ceci/cela/ça*** *dont il est accusé*.

Je vous renvoie à* ***ceci/cela/ça* *que j'ai déjà écrit sur le sujet*.

****Ceci/cela/ça*** *qui a été dit est exact*.

De plus, *ceci* et *cela* sont les formes marquées de *ce* devant le verbe *être*, lorsque celui-ci commence par une voyelle :

(12) ***Cela*** *est honteux* / ***Ceci*** *est honteux*.

Cela *est vrai* / ***Ceci*** *est vrai*.

En revanche, *ça* ne fait pas partie de cet ensemble marqué dans ce cas précis, puisque nous n'aurons jamais la suite suivante :

(13) ****Ça*** *est vrai*.

Nous pouvons ajouter à ces observations la suite de transformations suivantes, présentant des substitutions entre *ça* et *ce* dans des cas de dislocation :

(14) *Un chien, ça reste fidèle* / **Un chien, ce reste fidèle*.
(15) *Un chien, c'est fidèle* / **Un chien, ça est fidèle*.
(16) *Une femme, ça reste fidèle* / **Une femme, ce reste fidèle*.
(17) *Une femme, c'est fidèle* / **Une femme, ça est fidèle*.

(18) Le *gosse de mon voisin, c'est un lève-tôt* / **Le gosse de mon voisin, ça est un lève-tôt.*

(19) *Le gosse de mon voisin, ça reste un lève-tôt* / **Le gosse de mon voisin, ce reste un lève-tôt.*

(20) *Une femme, ça parle tout le temps* / **Une femme, ce parle tout le temps*.

(21) *Les femmes, ça parle tout le temps* / **Les femmes, ce parlent tout le temps*.

(22) **La femme de mon voisin, ça/ce parle tout le temps*.

Les exemples (14) à (19) confirment que *ça* et *ce* sont complémentaires face aux verbes d'état : *ce* et *est* sont indissociables – la copule permet d'ailleurs à *ce* de fonctionner seul comme argument ; à défaut, *ce* requiert un complément – tandis que *rester* appelle *ça* comme sujet. Nous notons par ailleurs que la nature adjectivale ou substantive de l'attribut n'a pas d'importance : dans les deux cas, *fidèle* et *un lève-tôt* sont des termes classifiants et l'ensemble de l'énoncé permet de catégoriser le référent de *ce* ou *ça* parmi « la classe des fidèles » ou parmi « la classe des lève-tôt ». De la même manière, nous constatons que le SN disloqué peut être à référence spécifique (exemples (18) et (19)) ou non (exemples (14) à (17)) ; de plus, contrairement à ce que pouvaient suggérer les exemples (3'') et (6''), *ça* peut renvoyer à de l'animé humain. Le seul paramètre qui semble donc décider de l'emploi de *ce* ou de *ça* est le verbe en jeu. Les exemples (20) à (22) ne présentent pas de verbe d'état et soulèvent des contraintes à l'emploi de *ça* dans les constructions à double sujet. En effet, dans ce cas, le SN disloqué doit être à référence générique. *Ça* serait donc incapable de reprendre une entité définie, identifiable et circonscrite. Ce phénomène est largement constaté dans la littérature existante, notamment par Corblin (1987, 1995), Kleiber (1987, 1994) et Maillard (1987), qui caractérisent *ça* comme évoquant un référent non catégorisé, non délimité, indistinct. Toutefois, la difficulté qui demeure est de déterminer comment la structure interne de *ça* permet de rendre compte d'un tel phénomène.

1.3. *Ça* : un indexical singulier au sein du système des démonstratifs français

Comme nous pourrons le constater tout au long de ce chapitre, *ça* est un élément avant tout polyvalent ; ne serait-ce qu'au niveau de sa distribution, puisque *ça* peut fonctionner dans toute position de SN :

(23) ***Ça** ne fait rien*.

(24) *Je vois **ça***.

(25) *Je lui ai pris* ***ça***.

(26) *Le colis, je l'ai posé sur* ***ça***.

(27) *Je ne l'ai pas fait pour* ***ça***.

(28) *En dépit de* ***ça****, je continuerai à lui faire confiance*.

Les exemples (23) à (28) laissent d'ailleurs entrevoir que *ça* n'impose pas de restriction à l'ordre de l'entité évoquée. De plus, *ça* entre en compétition avec les pronoms de troisième personne (exemple (29)), avec le *il* explétif (exemple (30)), mais apparaît également dans des expressions qui lui sont propres (exemples (31) et (32)) :

(29) *Les étudiantes,* ***elles*** *préfèrent le russe* / *Les étudiantes,* ***ça*** *préfère le russe*.

(30) ***Il*** *pleut* / ***Ça*** *pleut*.

(31) * ***Il*** *bosse ici !* / ***Ça*** *bosse ici !*

(32) *C'est son truc à lui* ***ça***.

En (29), *ça préfère le russe* insinue un commentaire dépréciatif, nuance qui ne se manifeste pas lors de l'emploi du pronom *elles* ; en (30), *ça pleut* fait songer à une pluie le plus souvent torrentielle, contrairement à *il pleut* qui ne semble évoquer qu'une pluie ordinaire. L'effet de *ça* par rapport au pronom de troisième personne produit donc un phénomène sémantique supplémentaire. Nous pouvons analyser l'exemple (31) de la même manière : *il* n'étant pas ici dans une construction (quasi)figée, comme le sont les expressions météorologiques, il est référentiel ; n'étant par conséquent pas explétif, *il* se doit de renvoyer à une entité identifiable, évidemment inexistante dans cet énoncé ; *ça* permet alors d'inscrire la prédication en jeu dans ce que nous pourrions nommer « l'atmosphère », en étendant les propriétés du prédicat considéré, caractéristiques d'un agent animé, à la situation d'énonciation. Enfin, *ça* en (32) apparaît en fin d'énoncé, position caractéristique que Morel (2007) nomme 'postrhème'[103]. Pour finir ce rapide inventaire syntagmatique de *ça*, nous citerons Guérin (2007), qui décrit très finement les spécificités sémantiques et pragmatiques de *ça* en fonction de l'analyse de constructions présentant un phénomène de dislocation, telles que :

(33) *Moi je comprends pas et ça ça m'intrigue.* (titre de l'article de Guérin, 2007)

[103] Nous étudierons *ça* en tant que postrhème en section 3.2 *infra*.

Cet exemple montre que *ça* peut occuper « les deux positions (désormais P1 et P2), *a priori* concurrentes, dévolues à deux éléments qui semblent évoquer le sujet d'un même verbe » (Guérin, 2007 : 27). La possibilité pour *ça* d'apparaître en P1 – position disloquée – le démarque d'un simple clitique ; de plus, sa capacité à être sujet grammatical ne peut l'assimiler à un pronom disjoint. Pour Guérin, *ça* se rapprocherait plutôt d'un SN, à la différence que ces derniers ne peuvent occuper la position P2 au sein de structures disloquées (*Mon frère, il est venu* mais pas **Il, mon frère est venu*). P1 étant destiné à poser le thème de l'énoncé, et P2 à servir de relais, *ça* serait donc, selon elle, à la fois référentiel et anaphorique (Guérin, 2007 : 6). C'est cette dernière remarque qui nous a menée à nous interroger sur la possible structure interne de *ça*.

A présent, nous nous apprêtons à inventorier les contextes grâce auxquels *ça* accomplit sa référence ; dès lors, nous nous attacherons à mettre en relation les phénomènes sémantiques observés en discours afin de dégager les propriétés internes et minimales de *ça*.

2. La référence indexicale de *ça*

Une description fine du fonctionnement des indexicaux selon les contextes invoqués ayant été effectuée lors du chapitre VII, nous présentons sans plus de précisions la dépendance contextuelle de *ça*. En revanche, la section 2.2 analyse en profondeur le fonctionnement référentiel de ce pronom neutre à partir de nos données orales, en vue d'extraire les caractéristiques syntactico-sémantiques minimales de *ça*.

2.1. Les contextes invoqués et les types de mémoire activés par *ça*

Au vu des fréquences dans notre corpus, *ça* est quasiment le seul pronom démonstratif employé à l'oral (10 450 occurrences de *ça* contre 28 de *ceci*, 35 de *cela*, 6 de *celui-ci* et 32 de *celui-là*). Il paraît donc logique de déduire que *ça* est à même de recourir à tous les contextes possibles pour référer, intuition que confirment nos analyses. Nous choisissons de ne pas développer de manière extensive le mécanisme opéré par *ça* pour recourir aux différents contextes, mais nous présentons toutefois quelques exemples caractéristiques qui montrent l'étendue contextuelle et mémorielle par laquelle *ça* accomplit sa référence.

2.1.1. La référence indexicale de *ça* par le recours au contexte situationnel immédiat

Tout comme *this* et *that* en anglais, *ça* est à même d'introduire un référent dans la mémoire focus de l'allocutaire via le contexte situationnel spatial immédiat ; nous avons en (34) un cas typique de référence indexicale de ce type, où *ça* est par ailleurs en emploi déictique :

(34) [DP et l'enquêtrice E regardent des photos ; E demande à DP si ces clichés représentent ses enfants]

E : ***Ça*** *c'est vos enfants*. DP : ***Ça*** *c'est mes,* ***ça*** *c'est mes enfants*. E : *Ah, oui*. (Corpus PFC Domfront, code locuteur 61adp1, conversation guidée)

Toutefois, l'accès à la référence via le contexte spatial immédiat n'entraîne pas de manière systématique de procédure déictique :

(35) [LD demande à l'enquêtrice E si le magnétophone est en marche]

LD : ***Ça*** *tourne là* <E : *Oui, oui,* ***ça*** *tourne*> (rires) ? (Corpus PFC Douzens, code locuteur 11ald1, conversation guidée)

En (35), *ça* renvoie au « magnétophone », élément central – au sens propre comme au figuré – d'une conversation guidée dans le cadre du projet PFC. Par conséquent, « le magnétophone » est déjà en mémoire focus du destinataire, ce qui fait que *ça* est en procédure anaphorique. Un autre indice, syntaxique cette fois-ci, de la procédure anaphorique réalisée ici par *ça* est le fait que le pronom soit sujet d'une construction figée (*ça tourne*) ; nous pourrions même avancer que *ça* est ici un simple clitique. *Ça* en (35) se rapproche indéniablement de *it* en anglais (cf., par exemple, exemple (19) du chapitre VII). Néanmoins, nous proposons en (36) un fonctionnement indexical de *ça* plus complexe :

(36) [L'aboiement du chien coupe la conversation entre LD et l'enquêtrice]

LD : *Le port de chasse, le fusil, les pièges à lapins euh, à moineau, il y en a même un de gros qui je pense* (aboiement) *doit être un pièg/ un piège à lapins.* (aboiement) ***Ça***

c'est quelqu'un qui arrive <DP : *D'accord*>. (sonnerie) (Corpus PFC Douzens, code locuteur 11adp1, conversation libre)

En (36), *ça* est en position P1 de construction à double sujet (cf. ce chapitre, exemple (33)) ; comme le dit Guérin (2007), P1 est une position typique de SN, à vocation référentielle : c'est grâce à P1 qu'un référent est premièrement posé pour être ensuite repris par l'élément en P2 – ici, *c'* – et ainsi établi comme topique de l'énoncé. Par conséquent, *ça* a ici valeur de SN, ce que semble confirmer la substitution possible de *ça* par *cet aboiement*. Ainsi, nous pouvons déduire que *ça* a ici un pouvoir de nominalisation, ce qui le rapproche du fonctionnement de *that* (cf. chapitre VII, section 2.2.3). Toutefois, la référence indexicale de *ça* ici ne s'arrête pas à la nominalisation ; regardons les transformations suivantes :

(36') *Ça c'est quelqu'un qui arrive ≠ cet aboiement, c'est quelqu'un qui arrive*

Ça c'est quelqu'un qui arrive = cet aboiement veut dire que quelqu'un arrive

Ça ne se contente pas d'avoir valeur de SN, il instaure également une relation entre l'entité instanciée en P1 et le reste de l'énoncé produit – ici, une valeur de signification. *Ça* a donc ici pouvoir de nominalisation et de mise en relation de deux unités discursives. De plus, *ça* visant un référent absent de la mémoire de l'allocutaire au moment de son occurrence, nous pouvons inférer que *ça* est ici en emploi déictique. Par conséquent, la seule observation de trois occurrences de *ça* en référence via le contexte situationnel immédiat soulève de nombreuses pistes liées à son fonctionnement interne : *ça* peut être déictique, anaphorique, et semble non seulement incorporer un élément lexical sous-jacent – ce qui lui permet d'être pronom – mais aussi 'produire du sens'.

2.1.2. La référence indexicale de *ça* par le recours au contexte discursif

Ça en référence indexicale via le contexte discursif couvre les fonctionnements possibles de *this*, *that* et *it*. En effet, tout comme *this*, *ça* peut apparaître en position de cataphore (dans chaque exemple, *ça* est en gras et le déclencheur d'antécédent est souligné) :

(37) [JB et E discutent de la chasse comme elle se pratiquait quelques décennies auparavant ; E se demande alors si les déjeuners aux tripes sont une tradition toujours en vigueur]

E : *Et euh, **ça**, est-ce que **ça** s'est fait des fois euh, on en/ on a entendu, beaucoup parler des, déjeuners aux tripes ?* JB : <E : *Des euh, vous en faisiez des déjeuners aux tripes ?> Ah ben oui, oui oui, oui oui, oui ben, on, on f/ oui, oui oui, oui oui, j'en ai encore fait un cette année*. (Corpus PFC Domfront, code locuteur 61ajb1, conversation libre)

Néanmoins, nous notons que *ça* peut, en quelque sorte, être en position d''ana-cataphore' : il réfère à une entité discursive déjà introduite tout en dépendant d'une entité similaire dont le déclencheur est subséquent à l'occurrence de *ça* :

(38) E : *Il y a pas mal de problèmes de drogue au lycée ou au collège, ou bien ou,* <LH : *Euh, non, non, non, c'est*.> *comment **ça** commence ces choses-là?* (Corpus PFC Domfront, code locuteur 61alh1, conversation guidée)

(39) E : Est-*ce qu'il fallait les brosser, les laver ?* <GR : *Oui oui*.> GR : *les brosser, les faire boire, les brosser, tous les jours, tous les jours, c'était l'ouvrage à papa **ça** de brosser ses chevaux, ses juments*. (Corpus PFC Domfront, code locuteur 61agr1, conversation guidée)

Ça semble ici créer une relation entre deux entités, à l'image de ce que nous avons observé en (36) ; il est intéressant de noter ici qu'une structure telle que celle présentée en (39) se traduit par *it* suivi de *that* en anglais : *I find it quite incredible that he got away with it for so long* / **I find quite incredible that he got away with it for so long* (Huddleston & Pullum, 2002 : 1408).

Tout comme *this*, *that* et *it*, *ça* peut renvoyer au topique discursif (exemple (40)) ; à la manière de *that*, *ça* peut nominaliser un référent-antécédent déclenché par une proposition (exemple (41)) mais aussi récupérer un objet-de-discours pour le remettre en mémoire focus (exemple (42)) ; ainsi, *ça* est à même d'activer les zones centrales de la mémoire :

(40) [La conversation entre JB et E tourne autour de la chasse]

JB : *Oui mais là euh, je crois pas que tu aies vu tellement de monde avec nous à la chasse on y allait tous les trois*. E : *J'ai des vagues souvenirs, mais je me rappelle pas tellement*. JB : *Oui **ça** démarrait là oui*. (Corpus PFC Domfront, code locuteur 61ajb1, conversation guidée)

(41) [AL narre son parcours professionnel]

AL : *J'ai eu trois reconversions dans ma vie, vous savez qu'il faut en avoir trois dans sa vie de reconversions ?* <E : (inaudible)> ***Ça*** *vous le savez pas*. (Corpus PFC Douzens, code locuteur 11 aal1, conversation guidée)

(42) [A la demande de E, JB explique la tradition des déjeuners aux tripes]

E : *Des euh, vous en faisiez des déjeuners aux tripes ?* JB : *Ah ben oui, oui oui, oui oui, oui ben, on, on f/ oui, oui oui, oui oui, j'en ai encore fait un cette année, euh, c'était une entreprise où que, que je travaille av/ tu as bien connu ca (xx) toi ? Oui, et, il est encore en vie hein, il a quatre-vingt* <E : *Oui ?*>, *sept, quatre-vingt-sept ans, c'est mon premier patron, et, on s'est retrouvé ce, avec mon nouveau patron euh, on a mangé des tripes à Lonley là. Un dimanche matin là. Et, c'est encore bien, bien servi encore à l'ancienne là si tu veux, on te change les assiettes à chaque fois euh.* E : *Chaque fois que…* JB : *Et ben c'est-à-dire que, tu manges une euh, première euh, assiettée de, tripes, et, tu sais la tripe euh, c'est* <E : *Un peu, collant, oui c'est un peu.*>, *ç/ ça fige un peu, oui c'est un peu collant, et, c'est un, c'est, plutôt dommage de remettre des nouvelles tripes, dans, ton assiette, et en principe là, c'est ce qu'ils faisaient chez Fourmentin où qu'était Suzanne, et on, on te change l'assiette à chaque fois, jusqu'à temps que tu sois rassasié. Et puis, après, fromage euh, euh enfin, le vin adapté euh,* <E : *Le vin blanc ?*> *un petit blanc, un petit blanc avec oui, un petit, muscadet, et puis euh, le fromage avec un petit coup de rouge quoi, hein, et puis* <E : *Et tout ça, à dix heures, à di/ neuf dix heures le matin, non ?*>, *le, le café. Oui, oui, oh ben, on finit il est onze heures hein.* (rires) *Ouais,* ***ça*** *cale hein*. (Corpus PFC Domfront, code locuteur 61ajb1, conversation libre)

Enfin, tout comme *that*, *ça* est à même de faire appel à la mémoire partagée des interlocuteurs pour parfaire sa référence :

(43) [E et le locuteur AL sont amis ; E s'inquiète de l'exploitation de pommes gérée par AL, bien que le sujet n'ait pas été abordé jusqu'à présent]

E: *Et là, les pommes* ***ça*** *en est où alors ?* (Corpus PFC Douzens, code locuteur 11aal1, conversation libre)

En (43), la seule évocation des « pommes » permet à AL d'activer sa mémoire morte et d'inférer que E ne parle pas de « pommes en tant que telles », mais bien de son « exploitation fruitière ».

Les sections 2.1.1 et 2.1.2 présentent de manière succincte mais néanmoins exhaustive l'inventaire des contextes grâce auxquels *ça* peut accomplir sa référence. Il est tout à fait volontaire de notre part d'offrir un panorama aussi concis ; en effet, l'observation et la description du fonctionnement référentiel de *ça* sont l'objet de la section 2.2. Néanmoins, nous pouvons d'ores et déjà comparer *this, that, it* et *ça* en ce qui concerne l'invocation des différents contextes et l'activation des types de mémoire :

	Contextes et types de mémoire correspondants						
	Environnement	Mémoire vive				Mémoire morte	
	Situationnel	Discours en amont			Discours en aval	Savoirs	
		Informations actives (topique discursif)	Topique local (référent en focus)	Référent activé (mais pas en mémoire focus)		Partagé	Encyclopédique
This	x	x	x		x		
That	x	x	x	x		x	x
It	x	x	x				
Ça	x	x	x	x	x	x	x

TABLEAU 17 : LES CONTEXTES INVOQUES ET LES TYPES DE MEMOIRE ACTIVES PAR *THIS*, *THAT*, *IT* ET *ÇA*

2.2. Le rendu référentiel de *ça* : une production sémantique

La section 2.2 propose un inventaire des emplois syntaxiques et discursifs où *ça* semble entraîner une certaine production référentielle ; ainsi, *ça* ne serait jamais explétif et serait donc source de valeur sémantique par lui-même. Nous désirons alors expliquer quels sont les facteurs internes de ce pronom qui autorisent cette référence indexicale si singulière.

2.2.1. *Ça* : un indexical relationnel

Ça, en tant qu'indexical, ne se contente pas de pointer vers une entité. Au contraire, nous allons voir que ce pronom neutre a la capacité de mettre en relation deux éléments. Cette caractéristique est particulièrement ostensible dans les expressions à caractère temporel (44), météorologique (45) et 'atmosphérique' (46), dans les collocations *et tout ça* (47) et *comme ça* (48), ainsi que dans les 'ana-cataphores' (49) :

(44) E : *Oui donc **ça** fait déjà un bout de temps hein.* GR : *Treize ans* <E : *Oui.*>, *treize ans enfin l'année prochaine*. (Corpus PFC Domfront, code locuteur 61agr1, conversation libre)

(45) ***Ça** pleut.*

(46) ***Ça** réfléchit ici !*

(47) E : *Comment ça se passe l'internat de nos jours? Parce que bon, à, à mon, il y a trente ans ça devait être autre chose que maintenant hein*. <LH : *Ben.*> LH : *l'internat euh, bon euh, c'est, c'est franchement c'est, c'est, c'est c'est bien sympathique c'est, il y a, il y a rien à dire c'est vrai que la période de s/ s'adapter, vivre en communauté **et tout ça** c'est vrai qu'il y a des, en seconde on est encore un petit peu des, des petits gamins alors euh, la, les filles qui se mettent des bâtons dans les roues, **et tout ça** c'est, c'est, des fois il y a des sales ambiances*. (Corpus PFC Domfront, code locuteur 61alh1, conversation guidée)

(48) LH : *Et euh, ouais ben oui. Et en plus, tout le long je dois avoir l'accent, je dois parler **comme ça** tout le long de la pièce*. (Corpus PFC Domfront, code locuteur 61alh1, conversation libre)

(49) CG : *Ben euh, si les attentats aux Etats-Unis. On en a, on en a énormément parlé. On a fait un travail aussi euh là-dessus. Euh on a, ils nous ont consacré une semaine pour euh vraiment éplucher la presse et euh réaliser des, des dossiers. Euh. Pff.* E : *Qu'est-ce que vous en pensez, vous ?* CG : *De quoi ?* <E : *De.*> E : *ce*

qui s'est passé. CG : *Aux Etats-Unis ?* <E : *Oui.*> *Ben je trouve* ***ça*** *euh horrible quoi, d'en, d'en arriver euh, jusque là pour, pour de la religion.* (Corpus PFC Domfront, code locuteur 61acg1, conversation guidée)

Dans les expressions temporelles (44), *ça* est utile pour délimiter une durée, et donc pour mettre en relation deux dates précises – à la différence de *this* qui situe l'énoncé à « maintenant », et *that* qui inscrit l'énoncé dans un espace temporel en décalage avec la situation d'énonciation. Dans les expressions météorologiques (45), *ça* produit une amplification de la prédication en jeu – nous avons déjà souligné en section 1.3, exemple (30), que *ça pleut* renvoie le plus souvent à une « pluie torrentielle », par rapport à *il pleut* qui n'évoque qu'une « pluie ordinaire » ; de même en (46), où les propriétés indiquées par le prédicat employé semblent s'étendre à la situation d'énonciation dans son ensemble (cf. section 1.3, exemple (31)). En ce qui concerne *et tout ça* (47) et *comme ça* (48), le premier permet de compléter la référence amorcée par le déclencheur qui précède la suite *et tout ça* – ici, *vivre en communauté* et *les filles qui se mettent des bâtons dans les roues* – afin que le destinataire construise l'image mentale d'un ensemble d'éléments similaires à l'entité produite par ce déclencheur ; le second permet de relier deux éléments par une relation de comparaison. Enfin, la position d''ana-cataphore' (49) autorise, comme nous l'avons souligné avec les exemples (38) et (39), la mise en relation de deux entités discursives semblables, déclenchées par deux éléments textuels qui entourent l'occurrence de *ça*. *Ça* est donc à même de créer un lien, que ce soit entre deux périodes temporelles, entre la prédication en jeu et « l'atmosphère », entre un « objet » et un ensemble d'« objets » ayant les mêmes caractéristiques, ou entre deux entités discursives.

2.2.2. *Ça* : un indexical à élargissement référentiel

Le caractère relationnel de *ça* semble impliquer que ce pronom neutre ne puisse reprendre une entité définie et circonscrite, du moins lorsqu'il s'agit de lui assigner une propriété. Ce constat est particulièrement visible dans les cas de constructions à double sujet :

(50) *<u>Une femme</u>,* ***ça*** *parle tout le temps.*

(51) *<u>Les femmes</u>,* ***ça*** *parle tout le temps.*

(52) **<u>La femme de mon voisin</u>,* ***ça*** *parle tout le temps.*

(53) JB : *Et je me souviens une fois euh, il fallait qu'on passe la rivière hein, et, je me souviens une fois, on a tué un lièvre, je crois que c'est mon frère, c'est Georges*

*qui a tué le lièvre, en partant, et ben c'est moi qui l'a porté tout le temps dis-donc euh, un lièvre de six ou sept livres dans le dos, **ça** tient chaud mais **ça** pèse hein*. (Corpus Domfront, code locuteur 61ajb1, conversation libre)

(54) JB : *Et ben c'est-à-dire que, tu manges une euh, première euh, assiettée de, tripes, et, tu sais la tripe euh, c'est* <E : *Un peu, collant, oui c'est un peu.*>, ***ç/ ça** fige un peu, oui c'est un peu collant.* (Corpus Domfront, code locuteur 61ajb1, conversation guidée)

(55) [Copie de l'exemple (43)]

E: *Et là, les pommes **ça** en est où alors ?* (Corpus PFC Douzens, code locuteur 11aal1, conversation libre)

Que l'entité introduite par l'expression disloquée soit dotée du trait [+ humain] (exemples (50) et (51)) ou pas (exemples (53) et (54)), elle résulte forcément d'une lecture générique du SN, que ce dernier soit défini ou indéfini. Preuve en est l'impossibilité de **la femme de mon voisin, ça parle tout le temps*, où le complément *de mon voisin* induit nécessairement une lecture spécifique du SN. Comment pourrions-nous expliquer ce phénomène ? Si nous considérons que *ça* engendre une relation quelconque (cf. ce chapitre, section 2.2.1 *supra*), il ne peut se contenter de reprendre une entité pour l'inscrire dans un énoncé ; quelque part, il faut qu'un lien soit créé. Si l'expression disloquée peut être en lecture spécifique ou générique, c'est cette dernière interprétation qui sera sélectionnée par *ça* ; ainsi, *ça* élargit l'entité considérée à la classe d'entités à laquelle elle appartient, pour l'inscrire dans l'énoncé en jeu. Si l'expression disloquée ne peut être qu'à lecture spécifique, comme c'est le cas dans l'exemple (52), il en résulte une impossibilité. Que pouvons-nous alors dire de l'exemple (55) ? Ici, compte tenu de ce qui en est prédiqué, « les pommes » ne peuvent résulter d'une lecture générique du SN : nous ne sommes pas face à un énoncé du type *les pommes, ça reste le fruit préféré des Français*, qui permet de mettre en relation « la classe des pommes » et « être le fruit préféré des Français ». Pour que l'énoncé soit acceptable, il faut donc que *ça* puisse créer un lien entre « les pommes particulières dont parle le locuteur » et quelque chose d'autre. Ce lien est celui qui part de l'entité « les pommes » vers la mémoire morte de l'interlocuteur, afin que ce dernier comprenne que le locuteur ne réfère pas aux « pommes en tant que telles » mais à « son exploitation fruitière ». Ainsi, en quelque sorte, le SN disloqué déclenche un référent qui sert de point de départ à la genèse du référent évoqué par *ça*, complétée par l'apport d'informations contextuelles issues de la mémoire à long terme.

Que faire alors de la différence ressentie entre *cela* et *ça* lors d'une reprise anaphorique ? Nous pensons ici à l'exemple (10), que nous reproduisons pour faciliter la lecture :

(10) E : *Considérez-vous le Real comme l'équipe la plus en forme du moment ?*
MP : *Beaucoup de gens me disent qu'on affronte l'équipe la plus en forme du monde et ils ont raison. Ce sont les leaders de la Liga, ils survolent la Ligue des Champions, mais on est en novembre… **Ça** ne veut rien dire*. (www.sofoot.com, 17/01/2012)

(10') E : *Considérez-vous le Real comme l'équipe la plus en forme du moment ?*
MP : *Beaucoup de gens me disent qu'on affronte l'équipe la plus en forme du monde et ils ont raison. Ce sont les leaders de la Liga, ils survolent la Ligue des Champions, mais on est en novembre… **Cela** ne veut rien dire.*

Nous avons remarqué, lors de la première mention de cet exemple, que *ça ne veut rien dire* sous-entend « ce n'est pas parce que le Real est la meilleure équipe du moment que l'on ne peut pas gagner ». En revanche, en (10'), *cela ne veut rien dire* signifierait plutôt « les paroles qui viennent d'être prononcées n'ont aucun sens ». En d'autres termes, *cela* est en coréférence avec le déclencheur d'antécédent, ce qui n'est pas le cas pour *ça*. Guérin (2007) analyse des exemples similaires et propose une structure interne hybride pour *ça*, qui serait formée d'une composante anaphorique et d'une composante référentielle (cf. ce chapitre, section 1.3). Nous n'adhérons pas à cette proposition, dans la mesure où anaphorique et référentiel ne sont pas des propriétés de même niveau : en termes de GND, une expression référentielle est un élément textuel dont la catégorie lexicale est {N} – {N} et {P} étant les pôles extrêmes du continuum catégoriel entre {N} et {P} (cf. chapitre V, section 2.1) – tandis qu'anaphorique désigne une procédure indexicale réalisée par une expression référentielle – SN ou pronom – lors de son activation en discours. Ainsi, une expression référentielle peut être anaphorique ou pas, selon l'emploi énonciatif. Selon nous, le décrochage référentiel observé provient de la catégorie indexicale à laquelle appartient *ça*, celle des démonstratifs. En effet, même si les démonstratifs peuvent être déictiques ou anaphoriques en discours, et que, par conséquent, démonstratifs et deixis ne sont pas à confondre, il n'en reste pas moins que la deixis est la procédure indexicale prototypique des démonstratifs – puisque l'anaphore est dérivée de la deixis. Ainsi, nous en revenons à ce fonctionnement référentiel double, caractérisé par un retour sur le repère énonciatif. Certes, mais *this* et *that* ne manifestent-ils pas le même

fonctionnement indexical ? Pourquoi donc les démonstratifs anglais ne révèlent-ils pas ce genre de décrochage sémantique ? Selon nous, le décalage sémantique produit par *ça* provient également de son caractère neutre. *This* et *that*, eux, indiquent *a minima* la localisation spatio-mémorielle de l'entité visée, puisque chacun de ces démonstratifs définit une sphère en rapport avec l'*origo* et les participants à l'acte langagier. *Ça*, lui, est totalement neutre, puisqu'il est le seul élément de son système ; ainsi, en parallèle du mouvement rétroactif caractéristique des démonstratifs s'opère une dé-sémantisation du référent visé, que l'allocutaire réinterprète alors, lors de l'identification référentielle, en fonction de la prédication de l'énoncé. Dès lors, nous pouvons mieux expliquer les raisons pour lesquelles *ça* bloque la lecture spécifique d'un SN disloqué (cf. exemple (52)), mais également l'ensemble des phénomènes observés dans les exemples (50) à (55) : le mouvement référentiel rétroactif associé à la dé-sémantisation du référent premièrement évoqué entraîne une réinterprétation de ce même référent, réinterprétation nécessaire pour que la prédication de l'énoncé puisse s'appliquer ; les traits sémantiques liés à *ça* étant neutres, le référent reconfiguré est lui aussi neutre, ici, donc, non spécifique.

2.2.3. *Ça* : nominalisation et dynamisme référentiel

A présent, nous allons nous intéresser au rendu référentiel de *ça* en fonction de sa position syntaxique. Avant d'amorcer toute explication anticipée, tournons-nous vers les exemples (56) et (57) :

(56) [AL explique les circonstances de la mort de son père]

> AL : *Donc euh, mon père a eu une mort extraordinaire*. <E : *Ah oui ?*> *Ouais. Mon père avait un cancer, bon, alors, il a,* (bruit) *il avait mal aux reins,* (toux) *il avait mal aux reins, ma mère de suite, au mois de euh, mai juin, l'amène à Carcassonne. Et euh, on lui fait des, des analyses, enfin des radios des reins tout ça, rien. Et un jour moi je m'y trouve, à, à cinq heures, j'étais allé le voir, et le chirurgien arrive et lui dit 'demain matin, je vous enlève euh', i/, il lui presse sur le ventre à l'endroit de la, de l'appendicite, évidemment mon père a crié parce que euh,* ***ça*** *lui faisait mal, quand il, je l'ai vu faire moi, il appuyait comme une bête dessus*. (Corpus PFC Douzens, code locuteur 11aal1, conversation guidée)

En (56), *ça* est en position sujet. Ici, l'objet-de-discours visé par *ça* pourrait se ramener à « le geste qu'effectuait le médecin » ou bien « la manière dont le médecin appuyait sur l'appendice du père de AL ». Le déclencheur de cette entité est *il lui presse sur le ventre à l'endroit de la, de l'appendicite*. Quelle que soit la manière dont nous formulons l'expression qui y renvoie, cette entité est dynamique ; en d'autres termes, le référent de *ça* ici est temporellement situé, renvoie à une activité, est une entité de deuxième ordre. L'exemple (57), quant à lui, propose un *ça* en position objet :

(57) [JB explique qu'il existe un A.O.C. du poiré]

JB : *Comme il y a un A.O.C. du poiré aussi, oui. Et, là ils sont c/ euh, ou ils peuvent les avoir dans leur cave, mais ils sont contrôlés, ils savent, euh les g/ la g/ les, la régie sait combien ils ont de litres de, de, d'eau de vie chez eux, ils ont même pas le droit d'en perdre hein, si ils ont une fuite hein, il faut qu'ils, appellent aussitôt hein, parce que, sinon c'est, c'est qu'ils l'auraient vendu, enfin bon* <E : *Ah oui d'accord.*>, *hein, euh, ah oui oui oui* <E : *Ben oui ça plaisante pas.*>, *ah ben c'est suivi hein, c'est suivi, et le, donc euh, et ils redonnent* ***ça****, et c'est vendu avec tous les droits*. (Corpus PFC Domfront, code locuteur 61ajb1, conversation guidée)

Ici, *ça* renvoie au « poiré », entité de premier ordre qui conceptualise un objet stable de la réalité. Ces deux exemples alimentent une intuition que nous aurions à la simple observation d'énoncés-types comme *ça m'intéresse* et *donne-moi ça* ; dans le premier cas, *ça* renverrait plutôt à une activité, voire à un fait, tandis que, dans le second cas, *ça* désignerait sans doute une entité correspondant à un objet physique de la réalité. Ainsi, nous considérons une série d'exemples authentiques où *ça* apparaît en position sujet (exemples (58) à (61)) et d'autres où *ça* apparaît en position objet (exemple (62)) ou marquée, comme dans les structures clivées (exemples (63)) ou en position P1 de construction à double sujet (exemple (64)) :

(58) AL : *Ben ce que je pense c'est que les expériences que j'ai eues dans ma vie elles m'ont toujours été profitables quoi parce que euh, bon on fait un truc bon* ***ça*** *apprend des choses, on en fait un autre* ***ça*** *apprend des choses*. (Corpus PFC Douzens, code locuteur 11aal1, conversation guidée)

(59) CG : *Euh du cannabis bon ben, je pense que, ben i/ on a du mal aussi à connaître les effets, enfin euh actuellement il y a peu d'études qui ont été faites sur les effets du cannabis. Et euh, et c'est u/ une fois euh, peut-être dans dix ans où on*

verra vraiment la dépendance qui se crée, euh, et puis les effets euh neurologiques et psychologiques que ***ça*** *peut engendrer*. (Corpus PFC Domfront, code locuteur 61acg1, conversation guidée)

(60) E : *C'était quelle pièce, une pièce connue ?* CM : *Oui, 'Du vent dans les branches de Sassafras'*. E : *Ouais*. CM : *De René de Obaldia. C'est ça hein ouais ? Je me trompe pas ? Et puis euh, cette année, ben j'ai recommencé parce que je trouvais ça bien, finalement j/ j'ai adoré. Et cette année en plus il y avait les filles qui étaient là. Voilà. Donc euh, j'avais un, un copain Nico là, qui, qui, qui, qu'on a fait en plus, donc euh. C'est vrai que, j'ai des copains qui sont venus nous voir et ils ont adoré donc euh,* ***ça*** *a r/* ***ça*** *a apporté du monde en fait. Euh, l/ la pièce a été réussie, on peut le dire ?* (Corpus PFC Domfront, code locuteur 61acm1, conversation guidée)

(61) L2 : *Aussi il y a des inond/ ce qu'on, il faut voir comment que le monde il est bousculé quand même, toutes les inondations qu'il y a, du côté du* <HD : *Ah oui.>, Gard*. HD : *Dans le Midi ils en connaissent euh*. L2 : *Hein ? En s/ en quatre-vingt-dix-neuf nous notre étable a été, pas culbutée complètement mais toute la couverture était enveloppée, et le pillon, de l'étable, le, du bâtiment, ben je vous assure que ça faisait peut-être bien plus de potin, oh ben, ça fais/* <HD : *C'était.> c'était un ouragan*. HD : *C'était des ardoi/ des ardoises ils volaient partout hein. On avait peur que* ***ça****, défonce la* <L2 : *La, au pillon de l'autre maison oui.>, la (x) de pillons là*. (Corpus PFC Domfront, code locuteur 61ahd1, conversation libre)

(62) DP : *Elle est même enterrée à, Saint-Roch, la mère de, madame, la mère L.* <E : *Ah.>* E : *oui c'est vrai, et celle qui s'appelait Elisa L.* <DP : *L., oui, Elisa L.>, oui, oui oui*. E : *Elle était venue mourir à*. <DP : *Oui.>* DP : *je, c'ét/ c'était peut-être Elisa mais enfin c'était une femme L.* <E : *Oui, c'était Elisa.>, oui. Oui, elle est venue mourir, puis je me ra/ enfin, la mère Girard nous a raconté* ***ça*** *bien des fois, quand même*. (Corpus PFC Domfront, code locuteur 61adp1, conversation libre)

(63) CG : *Et euh, et c'est u/ une fois euh, peut-être dans dix ans où on verra vraiment la dépendance qui se créait, euh, et puis les effets euh neurologiques et psychologiques que ça peut engendrer. Euh, ben que, les jeunes comprendront euh, peut-être, et encore, même si on connaît euh les effets du tabac euh c'est pas pour* ***ça*** *qu'il y en a qui, qui arrêtent*. (Corpus PFC Domfront, code locuteur 61acg1, conversation guidée)

(64) LH : *Non, je, enfin soit je ferai du classique, soit, soit je ferai des comédies musicales. Parce que j'adore le théâtre aussi donc euh*. CM : *Oui mais, on va perdre une des grandes artistes françaises, je veux dire. Pourquoi tu vas pas à la 'Recherche d'une nouvelle star' ?* LH : *Oh, oh lala* <CM : *La pitié.*>. *Non merci pour la t/ la télé c'est* <CM : *Non mais, non mais franchement, **ça**, **ça** c'est une émission mais, c'est n'importe quoi.*> *pas mon truc*. (Corpus PFC Domfront, code locuteur 61acm1, conversation libre)

Tous les référents auxquels renvoie *ça* en position sujet (exemples (58) à (61)) sont dynamiques : en (58), *ça* évoque « l'activité professionnelle que l'on pratique à un moment donné dans notre vie » ; en (59), *ça* ne renvoie pas au « cannabis », mais à « la pratique qui consiste à fumer du cannabis de façon régulière » ; en (60), *ça* réfère à « la représentation théâtrale à laquelle a participé le locuteur », et non pas à « la pièce elle-même » ; enfin, en (61), l'image mentale suscitée par *ça* comprend « l'ouragan », bien entendu, mais surtout « les violents mouvements en spirale » qui le caractérisent. En revanche, *ça* dans les exemples (62) à (64) provoque un rendu référentiel différent : en (62), *ça* pourrait se ramener à *cette histoire*, SN renvoyant à une entité de troisième ordre mais n'ayant rien de dynamique ; pareillement en (63), où le SN *ces raisons* pourrait remplacer *ça* ; enfin, en (64), *ça* équivaut au titre de l'émission *A la recherche de la nouvelle star*. Par conséquent, nous constatons que *ça* en position sujet tend à renvoyer à « de l'activité », tandis que *ça* dans les autres positions argumentales préfère s'accorder d'une référence à « du statique ». En d'autres termes, *ça* sujet est moins facilement substituable par un SN que dans les autres positions argumentales. Toutefois, nous observons que ce dynamisme référentiel de *ça* sujet disparaît lorsque l'énoncé est à la voix passive :

(65) SL : *Quatre cents, quatre cents personnes hein, il y avait quatre cents personnes. L'usine Gaston J. c'était. Et maintenant il y a plus maintenant bon **ça*** [renvoie à « l'usine »] *a été repris mais il y a euh, il y a peut-être cinquante personnes qui travaillent c'est tout hein.* (Corpus PFC Vendée, code locuteur 85asl1, conversation libre)

(66) L2 : ***Ça*** [renvoie à « la laiterie »] *a été bâti en quelle année, comme ta maison* <DP : *Ben, je sais pas oui.*>, *à toi Denise ?* <DP : *Ben.*> DP : *il paraît que c'est en soixante-cinq je crois* <L2 : *Ah oui.*>, *je sais pas, en soixante-cinq*. (Corpus PFC Domfront, code locuteur 61adp1, conversation libre)

Le dynamisme référentiel de *ça* en position sujet est l'objet de certains travaux, notamment ceux d'Achard (2001, 2011) et de Cadiot (1988), mais, selon nous, le seul traitement de *ça* sujet consiste à ne mettre au jour que la partie visible de l'iceberg. Le contraste observé entre *ça* sujet, lorsque l'énoncé est à la voix active, et *ça* dans les autres positions argumentales possibles révèle que le fonctionnement référentiel de *ça* dépend crucialement des contraintes imposées par le verbe. Si nous reprenons les exemples (56) et (58) à (61), nous disposons des verbes ou locutions verbales *faire mal*, *apprendre*, *pouvoir engendrer*, *apporter* et *défoncer*. Nous remarquons que tous ces verbes acceptent un appareil verbal comme sujet : *appuyer sur un endroit douloureux fait mal, étudier/travailler apprend des choses, fumer du cannabis peut engendrer des dommages, jouer cette pièce de théâtre a apporté du monde, provoquer des tourbillons si violents défonce les constructions*. En revanche, *ça* dans les exemples (57) et (62) à (64) n'est pas remplaçable par un verbe. Que se passe-t-il au niveau du fonctionnement référentiel de *ça* ? Dans les exemples (56) et (58) à (61), *ça* s'appuie sur le discours en amont pour topicaliser une information connue et déjà saillante, tout comme l'effectuent *this* et *that* en référence via le contexte discursif (cf. chapitre VII, sections 2.1.2.2 et 2.2.1) ; *ça* en tant que démonstratif provoque un retour sur le repère discursif, ce qui permet à l'allocutaire de construire un référent à partir de celui-ci ; mais *ça* en tant que démonstratif neutre se contente de renvoyer à cette information sans la modifier : rien n'est prédiqué de cette information, puisque *ça* n'apporte nulle indication de genre ou de localisation spatiale ; *ça* poursuit le topique discursif sans l'altérer. Les exemples (57) et (62) soumettent un *ça* en position objet ; dans les deux cas, la contrainte syntaxique impose une nominalisation : si le sujet peut être une proposition verbale, l'objet de *redonner* et de *raconter* doit être sous forme nominale ; ainsi, *ça* en (57) reprend une entité déjà introduite par un SN (*du poiré, des litres d'eau de vie*) et même renvoyée par des pronoms clitiques (*en, l'*), tandis que *ça* en (62) nominalise un contenu propositionnel ; la différence entre *ça* et un simple pronom clitique dans ces cas-là est la faculté qu'a le pronom démonstratif de placer une entité non saillante au moment de l'occurrence en mémoire focus, ceci étant possible par le mouvement rétroactif sur le repère énonciatif opéré par *ça*. L'explication du fonctionnement de *ça* dans les exemples (63) et (64) est à la fois similaire et différente de celle soumise pour les exemples (57) et (62) : *ça* effectue également une nominalisation, mais cette dernière résulte principalement du dispositif dans lequel s'insère *ça*, à savoir une construction clivée et une dislocation, qui sont deux procédés de mise en relief ; c'est la fonction communicative de ces dispositifs associée aux propriétés du démonstratif qui permettent la nominalisation et la mise en mémoire focus du référent considéré. Dans tous les cas, *ça* en tant que démonstratif permet de (re)mettre un référent ou

une information discursive en mémoire focus ; dans tous les cas également, *ça* en tant qu'indexical neutre renvoie à une entité sans rien en prédiquer, ni au niveau du genre, ni au niveau de la localisation spatiale ; *ça* en tant que démonstratif neutre autorise une topicalisation d'« objets » sémantiques dans sa définition la plus large : les « choses » ne sont pas nommées, les contenus propositionnels conservent leur dynamisme tout en perdant la prédicativité instaurée par l'emploi initial d'une construction verbale. *Ça* est une illustration parfaite de deux phénomènes langagiers : l'hypostatisation de la langue – qui tend à ramener des faits à des choses concrètes, d'où le procédé de nominalisation – et la sous-spécification de la langue – les éléments du lexique ont des capacités référentielles limitées qui sont compensées par l'intervention des contextes situationnels et mémoriels, surtout lors de conversations spontanées où les participants sont en coexistence spatiale et temporelle.

2.2.4. *Ça* : une référence indexicale qui résulte de deux paramètres

Considérer *ça* comme ayant une référence indexicale marquée par ses propriétés de démonstratif d'une part et d'indexical neutre d'autre part permet de réexaminer les cas où ce pronom s'accompagne d'un accroissement référentiel. Pour rappel, les exemples canoniques qui illustrent ce fonctionnement particulier sont les expressions météorologiques (exemples (30) et (45)), les expressions que nous avons choisi de qualifier d''atmosphériques' (exemple (46) : *ça réfléchit ici !*) et les occurrences de *ça* en position P2 de construction à double sujet (exemples (50) à (55)). Dans tous ces cas-là, *ça* provoque un retour référentiel sur le repère énonciatif associé à une neutralisation sémantique qui amène l'allocutaire à (re)construire un référent. Dans les expressions météorologiques, cette construction référentielle se traduit par une diffusion des propriétés véhiculées par le verbe employé à l'ensemble de la situation d'énonciation, ce qui fait qu'au contraire de *il*, *ça* n'est pas explétif ; dans le cas des expressions 'atmosphériques', le prédicat employé requiert un sujet-argument agentif, la conséquence étant que la production référentielle induite par *ça* semble étendre cette propriété caractéristique d'un agent animé à l'atmosphère ambiante ; dans les cas de position P2 de structure disloquée, un décalage référentiel se manifeste en majorité par la référence à la classe. Ainsi, les caractéristiques de *ça* que nous avons proposées dans les sections 2.2.1 et 2.2.2, à savoir *ça* comme étant un indexical relationnel et *ça* comme induisant un élargissement référentiel, sont les conséquences de ce mode référentiel marqué par ses deux caractéristiques internes : le fonctionnement référentiel d'un démonstratif et la neutralité. Toutefois, il ne faut pas oublier que *ça* est propice à la nominalisation si la position

argumentale qu'il occupe appelle un SN (exemples (62), (63) et (64)), tandis que *ça* sujet tend à renvoyer à du dynamique ; ceci est dû à la référence de *ça* qui est tributaire des contraintes syntactico-sémantiques imposées par le prédicat. Enfin, nous devons garder à l'esprit que *ça* peut être en procédure anaphorique, même si tout laisse à penser désormais que ses propriétés intrinsèques produisent nécessairement une modification référentielle et, donc, une manifestation de référence déictique. Notre corpus présente des occurrences de *ça* témoignant du maintien d'une entité en mémoire focus : nous pensons ici notamment aux exemples (35) et (40). Le premier exemplifie un cas d'anaphore sans mention référentielle antérieure et le second un cas de continuation topicale. Ces deux emplois-types de *ça* semblent les plus proches d'une procédure anaphorique.

3. Les fonctions (inter-)communicatives de *ça*

Tout au long de ce chapitre, nous avons pu témoigner de l'extrême polyvalence de *ça*. En effet, cet indexical seul autorise l'accès à la référence via tous les contextes possibles et renvoie à des entités de tous ordres. De plus, et surtout, *ça* est un indexical complexe alliant propriétés des démonstratifs et neutralité sémantique. Bien entendu, l'ensemble de ces caractéristiques permet à *ça* de couvrir le continuum des procédures indexicales. Dès lors, l'omniprésence de *ça* à l'oral n'est pas surprenante : cette très grande diversité référentielle fait de ce pronom neutre un candidat idéal à l'acte de référence en contexte conversationnel. Grâce à lui, les participants à l'acte de communication renvoient à tout un éventail d'objets-de-discours tout en produisant le minimum d'effort linguistique ; *ça* permet de référer sans nommer par le simple truchement de facteurs contextuels, qu'ils proviennent de la situation spatio-temporelle directe, du discours ou de la mémoire morte. La condition *sine qua non,* toutefois, est que l'ensemble des ces paramètres contextuels soient partagés par les interlocuteurs ; sans cela, lesdits interlocuteurs ne disposeraient pas des éléments nécessaires à l'identification du référent visé par *ça*. Cet impératif est, selon nous, la raison première pour laquelle *ça* est si largement usité à l'oral et si peu à l'écrit.

A présent, nous souhaitons mettre en évidence le rôle pragmatique joué par *ça* dans le processus de co-construction du discours. Ainsi, nous espérons déceler une sphère discursive propre à ce pronom neutre, à la manière des sphères personnelle, interpersonnelle et externe décrites par *this*, *that* et *it* (cf. chapitre VII, section 4). La définition de cette sphère émerge de l'observation de trois fonctions intercommunicatives et discursives de *ça* : la fonction de résomption, la fonction de ratification et la fonction de repérage.

3.1. *Ça* et la fonction de résomption

La capacité résomptive de *ça* se détache notamment des constructions clivées, telles que celles en *c'est pour ça que (X)* et *c'est ça que (X)* :

(67) [La jeune locutrice, LH, explique qu'entre les leçons de piano, de danse, de théâtre et le lycée, elle n'a pas beaucoup de temps pour profiter de ses parents]

LH : *C'est vrai que ben le vendredi soir après le théâtre ben je sors avec mes copains, et puis bon, c'est vrai que je, je, je mets tout le vendredi soir et puis, le samedi donc je fais mes ateliers de musique et le dimanche je travaille mes leçons. C'est vrai que, pas beaucoup de repos, les vacances sont toujours bienvenues*. E : *Et maman et papa alors, tu as pas beaucoup le temps d'en profiter*. <LH : *Non*.> LH : *maman et papa c'est vrai que, ben ouais le, le s/ le samedi midi, le dimanche, voilà.* ***C'est pour ça que*** *ben, on s'arrange euh, depuis l'année dernière, on, à chaque vacances on essaie de, de beaucoup voyager euh, parce que.* E : *Et tu aimes ça ?* (Corpus PFC Domfront, code locuteur 61alh1, conversation guidée)

(68) [Le locuteur DL explique sa méthode qui consiste à replanter du bois après l'abattage des arbres]

DL : *Ben c'est lui qui m'a appris de, qu'autrefois il abattait beaucoup de bois, mais quand il abattait du bois il replantait du bois ou il r/ il ne plantait pas mais il élevait des jeunes euh, des jeunes, des jeunes scions pour euh, pour remplacer. C'est lui un peu qui m'a, appris à, et abattre, il abattait beaucoup de bois, mais il élevait beaucoup pour euh, il élevait.* VL : *Et c'est ce que tu fais encore maintenant,* ***c'est pour ça que****, tu replantes beaucoup. Oui je replante euh, oui beaucoup parce que, moi je suis pour euh, produire, mais je crois que on peut produire et garder du bois.* (Corpus PFC Domfront, code locuteur 61adl1, conversation libre)

(69) [BM évoque les difficultés à être agriculteur de nos jours]

E : *Non. Vous trouvez donc que c'est assez dur de, de vivre comme un agriculteur à temps complet aujourd'hui ?* BM : *Ah ben oui hein, c'est, c'est pas simple hein* <E : *Non ?*>. *Puis bon, je trouve qu'on soutient pas, le, la, la qualité et la bonne*

chose aujourd'hui, ***c'est ça qui*** *m'agace moi. Au niveau de l'Europe tout ça euh, enfin il y a peut-être pas que l'Europe hein, mais, la France aussi, puis il y a la finance derrière tout ça quoi. On parle tout le temps de la qualité puis on fait l'inverse quoi.* (Corpus PFC Domfront, code locuteur 61abm1, conversation guidée)

(70) [BM exprime son mécontentement envers la politique actuellement menée concernant les importations]

BM : *Non mais des moments je trouve qu'on devrait un peu leur euh, taper sur les doigts un petit peu plus parce que, encore aujourd'hui moi je vois comme dans la viande tout ça euh, pff. ils, ils font ce qu'ils veulent hein. Et puis si vous êtes pas contents ben on fait venir d'ailleurs donc ils en ont rien à foutre hein.* ***C'est ça moi qui*** *me fait peur. Que ben pff, la France elle recule souvent hein. Parce que d'abord on n'a pas les mêmes coûts de production hein, quand on compare le, la viande d'Argentine (xx) de chez nous, c'est pas pareil. Et puis c'est pas la même non plus*. (Corpus PFC Domfront, code locuteur 61abm1, conversation libre)

Ça, dans les constructions clivées *c'est pour ça que* et *c'est ça que/qui*, permet de renvoyer à une information connue – le plus souvent, le topique discursif – pour la mettre en relief. Par conséquent, *ça* est ici à la fois anaphorique – par la récupération de l'information présupposée – et déictique – par la (re)mise en mémoire focus de l'objet-de-discours visé. Nous pourrions même avancer que l'emploi de *ça* ici se rapproche des 'ana-cataphores' (cf. exemples (38), (39) et (49)), puisque le topique auquel renvoie *ça* est développé de manière subséquente à l'occurrence du démonstratif, les deux unités discursives jointes par le dispositif clivé étant en relation de cause-conséquence ou d'élaboration :

(67) [La jeune locutrice, LH, explique qu'entre les leçons de piano, de danse, de théâtre et le lycée, elle n'a pas beaucoup de temps pour profiter de ses parents]

LH : *C'est vrai que ben le vendredi soir après le théâtre ben je sors avec mes copains, et puis bon, c'est vrai que je, je, je mets tout le vendredi soir et puis, le samedi donc je fais mes ateliers de musique et le dimanche je travaille mes leçons. C'est vrai que, pas beaucoup de repos, les vacances sont toujours bienvenues*. E : *Et maman et papa alors, tu as pas beaucoup le temps d'en profiter*. <LH :

Non.> LH : *maman et papa c'est vrai que, ben ouais le, le s/ le samedi midi, le dimanche, voilà.* ***C'est pour ça*** [« manque de temps pour voir ses parents » = cause] ***que*** (*c'est pour ça que* → cause-conséquence) *ben, on s'arrange euh, depuis l'année dernière, on, à chaque vacance on essaie de, de beaucoup voyager* [« voyagent ensemble pendant les vacances » = conséquence] *euh, parce que.* E : *Et tu aimes ça ?* (Corpus PFC Domfront, code locuteur 61alh1, conversation guidée)

(69) [BM évoque les difficultés à être agriculteur de nos jours]

E : *Non. Vous trouvez donc que c'est assez dur de, de vivre comme un agriculteur à temps complet aujourd'hui ?* BM : *Ah ben oui hein, c'est, c'est pas simple hein* <E : *Non ?*>. *Puis bon, je trouve qu'on soutient pas, le, la, la qualité et la bonne chose aujourd'hui,* ***c'est ça*** [« la qualité est dénigrée »] ***qui*** (*c'est ça qui* → élaboration) *m'agace moi. Au niveau de l'Europe tout ça euh, enfin il y a peut-être pas que l'Europe hein, mais, la France aussi, puis il y a la finance derrière tout ça quoi. On parle tout le temps de la qualité puis on fait l'inverse quoi* [« l'Europe, mais aussi la France, dénigrent la qualité pour des raisons financières, contrairement à ce que les institutions promettent » = élaboration topicale]. (Corpus PFC Domfront, code locuteur 61abm1, conversation guidée)

Nous allons voir à présent que cette fonction de résomption se rapproche de celle de ratification.

3.2. *Ça* et la fonction de ratification

Il est deux types d'occurrences de *ça* qui manifestent une fonction de ratification : les holophrases résomptives et ce que Morel (2007) nomme le 'postrhème'. Les holophrases résomptives s'apparentent aux dispositifs clivés que nous avons vus en section 3.1. *supra* – *c'est pour ça que* et *c'est ça que* – à la différence majeure que la relative disparaît –, ce qui nous donne *c'est pour ça* et *c'est ça*. Le 'postrhème', quant à lui, se situe après le rhème[104]

[104] Le rhème s'apparente *grosso modo* au noyau ou à la clause (cf. chapitre VI, section 2.1 pour une définition de ces termes), mais se définit par la prise en compte du contour prosodique. Ainsi, si le rhème présente un seul verbe recteur, il est surtout considéré pour l'unité intonative qu'il forme. Souvent, le rhème-noyau est suivi d'un postrhème, et peut être précédé d'un préambule (cf. Morel, 2007).

dont il n'est jamais séparé par une pause, et est toujours en intonation basse. Les exemples (71), (72) et (73) clarifient et illustrent ces définitions :

(71) [Conversation entre GR, SG et E ; GR, qui a huit ans d'écart avec son époux, travaillait encore quand ce dernier a pris sa retraite]

E : *Ouais parce que vous avez pris votre retraite il y a, une quinzaine d'années ?* GR : <E : *Une dizaine ?> Et ben en, et ben j'avais soixante et un ans quand ils l'ont donnée la retraite combien de c'était, en dégressif. Alors ils l'ont donnée, mon mari l'a eu à soixante-cinq ans mais moi je l'ai eu à soixante et un ans parce que la loi avait passé, qu'ils donneraient la l/ la retraite à soixante ans mais, dégressivement je, alors je l'ai eu à soixante et un. Oui oui*. E : *Oui donc ça fait déjà un bout de temps hein.* GR : *Treize ans, treize ans enfin l'année prochaine. Mon mari avait la retraite et puis, on exploitait quand même. Parce qu'il avait le droit d'exploiter*. <E : *Ah il avait le droit quand même.> Il avait le droit d'exploiter. On a exploité trop longtemps.* SG : *Ben oui, vous en avez pas assez profité à vous reposer tous les deux mais bon.* E : *Mais vous aimiez ça aussi donc euh.* <GR : *Ah.*> GR : *ben oui, ben oui, **c'est pour ça** puis moi j'étais pas à la, mon mari était à la retraite mais moi j'étais pas à la retraite*. SG : *Ben il y avait huit ans de différence entre les deux donc c'est là qu'on voit un petit peu le décalage.* (Corpus PFC Domfront, code locuteur 61asg1, conversation libre)

(72) [BM parle de sa femme qui tenait un petit magasin à Saint-Front]

E : *Racontez un petit p/ moi un petit peu comment c'était à l'époque euh, racon/ parlez-moi un petit peu de, de ce qu'elle faisait, d'où elle vient*. <BM : *Et ben.*> BM : *elle tenait un petit magas/ un magasin, de, ben d'alimentation, **c'est ça**, à Saint-Front, et puis ben, ça commençait à être dur aussi hein, la grande distribution, les grands magasins se sont montés donc euh, les petits magasins ils, il a lu/ il a baissé aussi puis bon, elle a arrêté quoi.* <E : *Et vous vous, oui ?> C'est pas évident, de lutter contre les, les gros.* (Corpus PFC Domfront, code locuteur 61abm1, conversation guidée)

(73) [CM et HL discutent de ce qui pourrait arriver si deux copies du baccalauréat étaient malencontreusement interchangées]

> CM : *Mais là regarde, il faut, il faut, je me suis déjà inscrit au lycée, le mec il, le pauvre mec il s'est pas inscrit, si ça se trouve il est inscrit à la fac et tout ça*. <HL : *Ben*.> HL : *ils vont pas lui retirer le bac attends*. CM : *Ben j'en sais rien*. HL : *Ben non*. CM : *Non mais ouais ce serait* <HL : *Moi j'irais bien hein, voir*.>, *ça serait, ça serait trop marrant* ***ça****, voir mon truc de bio. C'est pas le mien, paf. Et, et dire après 'Bon ben on vous donne le Bac'. Paf cash comme ça*. (Corpus PFC Domfront, code locuteur 61acm1, conversation libre)

Les exemples (71) et (72) illustrent un cas d'holophrase résomptive, tandis que l'exemple (73) offre une occurrence de *ça* en position postrhématique. Les locutions *c'est pour ça* et *c'est ça* fonctionnent de manière autonome sans pour autant avoir de signification propre. Ainsi, *c'est pour ça* et *c'est ça* sont comparables à des 'petits mots de l'oral', sont des marqueurs discursifs. La fonction de ces éléments holophrastiques est de ponctuer le discours de l'énonciateur : par l'une ou l'autre de ces séquences, ce dernier marque une pause dans le développement topical, pause qui lui permet d'assembler un empan de discours précédent avant de reprendre la progression discursive. La fonction de *ça* ici est comparable à celle que ce pronom neutre prend dans les dispositifs clivés *c'est pour ça que* et *c'est ça que*, à la différence que l'holophrase ne créé pas de pont entre deux unités discursives ; au contraire, elle clôt la première avant que la suivante ne commence. En (71), *c'est pour ça* permet au locuteur d'exprimer son accord envers l'énoncé produit par l'interlocuteur – E : *Mais vous aimiez ça aussi donc euh.* <GR : *Ah*.> GR : *ben oui, ben oui,* ***c'est pour ça*** *puis moi j'étais pas à la, mon mari était à la retraite mais moi j'étais pas à la retraite* ; en (72), *c'est ça* donne au locuteur l'occasion de confirmer ses propres propos – BM : *elle tenait un petit magas/ un magasin, de, ben d'alimentation,* ***c'est ça****, à Saint-Front* ; dans un cas comme dans l'autre, le locuteur confirme la véracité d'un propos, d'où le terme de 'ratification' que nous avons attribué à cette fonction. L'exemple (73) offre une occurrence de *ça* en position de posthrème ; dans ce cas-ci aussi, nous pouvons témoigner de la reprise référentielle effectuée par *ça* – *ça* posthrème renvoie au même « objet » que *ça* sujet dans *ça serait trop marrant ça* –, mais également de la clôture discursive qu'il produit – la pause intonative est marquée après l'occurrence de *ça*, le posthrème faisant partie du contour intonatif du rhème *ça serait trop marrant*. *Ça* postrhème permet au locuteur d'asseoir son énoncé, d'où la fonction de 'ratification'. Par ailleurs, nous notons que les procédés holophrastiques *c'est pour ça* et *c'est ça*, ainsi que les *ça* posthrèmes, permettent à l'énonciateur de ratifier ses propres propos comme ceux du co-énonciateur.

Qu'en est-il de la fonction de repérage ?

3.3. *Ça* et la fonction de repérage

Les occurrences de *ça* qui nous intéressent ici sont celles qui apparaissent en liaison avec les mots interrogatifs ; en effet, notre corpus offre un grand nombre de *Où ça ? Quand ça ? Qui ça ?* A la manière de *ça* posthrème, *ça* ici n'est jamais séparé de l'unité intonative formée par le mot interrogatif ; à la différence du postrhème, bien entendu, l'intonation est ici ascendante. La fonction que pourrait revêtir *ça* dans ce cas le rapproche de ce qui le caractérise en tant que démonstratif, c'est-à-dire l'emploi déictique. En effet, de façon notoire, *ça* dans *Où ça ? Quand ça ?* et *Qui ça ?* effectue un repérage spatio-temporel par rapport aux participants à l'acte énonciatif. A l'évidence, cet emploi fait écho à l'ancrage spatio-temporel opéré par *this* et *that* (cf. chapitre VII, sections 2.1.1 et 2.2.1). En anglais, la paire marqué/non marqué formée par *this* et *that* autorise une distinction des repères énonciatifs pour chacun des démonstratifs, le premier étant associé à *origo-here-now* et le deuxième à *origo-there-then* – le second repère pouvant inclure le premier. Concernant *ça*, cette différence disparaît, bien entendu, puisque ce pronom neutre est le seul membre de son système. Ainsi, nous pourrions dire que le repérage énonciatif qui lui est associé localise un site S spatio-temporel de manière neutre par rapport aux participants à l'acte de communication.

3.4. Les fonctions (inter-)communicatives de *ça* : bilan

Les sections 3.1, 3.2 et 3.3 ont permis de révéler trois fonctions communicatives et discursives de *ça* : une fonction résomptive, qui, lorsque *ça* s'insère dans un dispositif clivé, permet de renvoyer à un empan de discours tout en créant un lien logique avec l'empan à venir ; une fonction de ratification, qui autorise l'énonciateur à confirmer ses propres propos ou ceux du co-énonciateur ; enfin, une fonction de repérage, qui ancre l'énoncé dans un repère spatio-temporel tourné vers les participants à l'acte langagier. Les fonctions de résomption et de ratification semblent avoir plus de connivence entre elles qu'avec la fonction de repérage, de prime abord. En effet, la résomption et la ratification manifestent un assemblage discursif : l'occurrence du démonstratif permet de topicaliser une information présupposée avant que ne se produise le développement subséquent du discours. Ainsi, *ça* s'apparente ici à un point de jonction du discours. Au niveau de l'intercommunication, cette jonction permet d'orienter l'attention des interlocuteurs sur le même emplacement discursif ; nous pourrions dès lors déduire une fonction phatique de *ça*. Une fois les participants à l'acte énonciatif en phase, le

discours peut continuer sa progression. Nous dirons que *ça* ici a pour vocation communicative de faire partager le même terrain d'entente aux interlocuteurs. En ce qui concerne la fonction de repérage, nous pourrions dire que *ça* ici inscrit l'énoncé et les participants à l'acte communicatif dans un repère orienté vers ces mêmes participants. Nous retrouvons alors la notion de 'terrain d'entente'.

Lorsque nous avons procédé à l'étude de *this* et de *that* en fonction de l'intercommunication, nous avons évoqué les notions de sphère personnelle pour *this* et de sphère interpersonnelle pour *that*. Dans le chapitre VII, nous avons choisi d'inclure *it* à ce système et avons donc attribué à ce clitique la notion de sphère externe. Comme nous l'avons déjà souligné, cette distinction s'efface dès lors que nous prenons *ça* en compte, puisque ce pronom neutre est le seul membre de son système. Toutefois, nous avons plusieurs fois remarqué au cours de notre développement que *ça* en français couvre les emplois de *this*, de *that* et de *it* en anglais. De plus, nous avons mis en exergue les propriétés relationnelles de *ça* (cf. ce chapitre, section 2.2.1), l'élargissement référentiel que ce pronom neutre peut produire (cf. ce chapitre, section 2.2.2), ainsi que les propriétés inhérentes à ce pronom neutre – celle d'être un démonstratif et celle d'être neutre – qui l'autorisent à absorber tous les rôles sémantiques. Si nous mettons en relation l'ensemble de ces caractéristiques avec la jonction discursive et intercommunicative que *ça* effectue, nous pourrions qualifier la sphère de ce pronom de 'corrélative' : *ça* met en corrélation des objets-de-discours, des empans de discours, les participants à l'acte langagier, et est en corrélation avec les contraintes argumentales imposées par le prédicat.

Marque proximale de *this*	Proximité spatiale	Sphère personnelle
	Proximité temporelle	
	Proximité cognitive	
Marque distale de *that*	Distance spatiale	Sphère interpersonnelle
	Décalage temporel	
	Distanciation cognitive	
	Récupération référentielle ➢ Par l'interlocution ➢ Par les zones excentrées de la mémoire	
Marque neutre de *it*	Neutralité spatiale	Sphère externe

	Neutralité temporelle	
	Espace cognitif externe	
Marque neutre de *ça*	Corrélation entre deux objets-de-discours	Sphère corrélative
	Corrélation entre lecture spécifique et lecture générique	
	Corrélation entre deux unités discursives	
	Corrélation entre les participants à l'acte langagier	
	Corrélation avec les contraintes verbales	

TABLEAU 18 : LES SPHERES DISCURSIVES DE *THIS*, *THAT*, *IT* ET *ÇA*

4. La structure interne de *ça* et une représentation possible de ce pronom neutre en GND

Afin de proposer une représentation de *ça* en GND, il est nécessaire d'extraire les caractéristiques inhérentes à ce pronom, d'après l'étude que nous venons d'effectuer :

- *Ça* étant un démonstratif, il présente certains traits minimaux en commun avec *this* et *that* ; ainsi, nous pouvons parler d'opération référentielle double pour *ça*, comme nous l'avons soumis pour les démonstratifs anglais : l'occurrence du démonstratif provoque un premier envoi référentiel vers la mémoire de l'allocutaire, puis un second vers un repère spatio-temporel ayant les participants à l'acte langagier à l'origine.
- Toutefois, *ça* est le seul membre de son système, contrairement au triptyque anglais *this-that-it*. Comme nous l'avons maintes fois souligné, le fonctionnement de *ça* se rapproche de celui de *this*, de *that* et de *it* selon son emploi. Voici quelques exemples caractéristiques, en relation avec la procédure plus ou moins déictique en discours :

	Procédure indexicale associée à l'acte référentiel					Référence nulle ➔ absence de procédure indexicale
	Deixis	Combinaison de deixis et d'anaphore			Anaphore	
This	{dei}	{dei ; ana}	{dei : ana}	{ana ; dei}		
That	{dei}	{dei ; ana}	{dei : ana}	{ana ; dei}		
It non explétif					{ana}	
It explétif						{ }
Ça	{dei}	{dei ; ana}	{dei : ana}	{ana ; dei}	{ana}	{ }

TABLEAU 19 : LES PROCEDURES INDEXICALES REALISABLES PAR *THIS*, *THAT*, *IT* ET *ÇA* EN DISCOURS

- *Ça* {dei}[105] se rapproche de *this* et de *that*, l'exemple caractéristique étant la référence déictique via le contexte situationnel immédiat (cf. exemple (34) : *ça c'est vos enfants*). Cependant, l'absence de marque proximale ou distale ne permet pas de trancher, pour *ça,* entre un fonctionnement comparable à celui de *this* ou un fonctionnement comparable à celui de *that* ;
- *Ça* en procédure indexicale combinant la deixis et l'anaphore se rapproche indéniablement de *that*, et ce par trois exemples caractéristiques issus de notre étude : l'exemple (41), qui présente un cas de nominalisation (***ça** vous le savez pas*, que l'on pourrait d'ailleurs traduire par *you don't/didn't know **that***), l'exemple (42), où *ça* renvoie à un référent déjà introduit en discours mais non saillant au moment de l'occurrence du démonstratif (***ça** cale*, que l'on pourrait traduire par ***that** fills you up*), et, enfin, l'exemple (43), où *ça* réfère via les connaissances partagées (*les pommes, **ça** en est où alors ?*). Nous avons vu dans le chapitre VII que la nominalisation, la récupération d'un objet-de-discours et la référence via la mémoire morte ne sont apparemment effectuées que par *that* ;

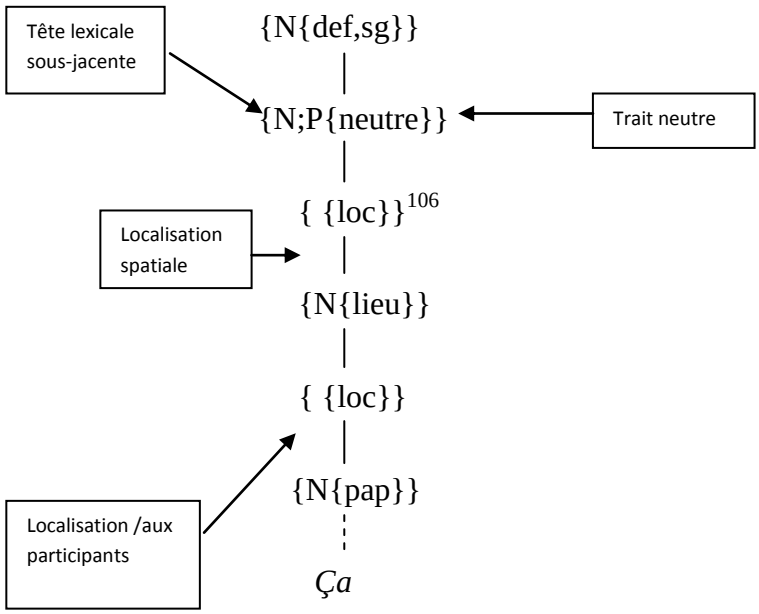
[105] *Ça* {dei}, *ça* {ana} et *ça* { } ne signifient en aucun cas qu'il y ait plusieurs *ça*. Au contraire, nous postulons un seul *ça* dont l'usage plus ou moins déictique ou anaphorique dépend de son emploi en discours. La notation *ça* {procédure} ne vise ici qu'à faciliter la compréhension.

- Enfin, *ça* {ana} et *ça* { } sont comparables à *it*. Nous pensons ici notamment aux occurrences caractéristiques de *ça* clitique, comme dans les expressions figées (exemple (35) : ***ça*** *tourne là ?*), météorologiques (exemple (45) : ***ça*** *pleut*) ou 'atmosphériques' (exemple (46) : ***ça*** *réfléchit ici !*). Cependant, nous avons noté que *ça* n'est jamais totalement explétif, d'où la différence sémantique entre *il pleut* et *ça pleut*, par exemple. Ceci est dû premièrement au fait que *ça* est un démonstratif.

- Enfin, la dernière propriété inhérente à *ça* est la neutralité : neutralité vis-à-vis du genre mais aussi vis-à-vis de la localisation. C'est ce caractère neutre qui pourvoit *ça* d'une telle polyvalence syntaxique et sémantique, mais c'est également ce caractère-là qui fait barrage à la reprise d'une entité spécifique (cf. **La femme de mon voisin, ça parle tout le temps*).

Voici donc ce qui pourrait être une représentation possible de *ça* en GND (nous nous basons sur les représentations de *this* et de *that* (cf. chapitre VII, section 5.2) pour la notation des traits minimaux propres aux démonstratifs) :

(74) *Ça*



[106] Nous rappelons que l'espace entre le crochet ouvrant '{' et '{loc}' signifie que nous avons affaire à une catégorie fonctionnelle.

A présent, et pour conclure cette étude du fonctionnement référentiel et interne de *ça*, nous souhaitons réinscrire ce pronom dans le système auquel il appartient : celui des indexicaux démonstratifs. Ainsi, nous proposons une formalisation de *celui-ci*, *celui-là*, *ceci* et *cela* dans le cadre de la GND.

5. Une représentation possible des pronoms démonstratifs en GND : bilan

Afin que l'étude de *ça* dans le cadre de la GND soit complète, il est nécessaire de proposer une représentation pour les pronoms démonstratifs que sont *celui-ci/celui-là* et *ceci/cela*. Dès la section 1 de ce chapitre, nous avons pu constater la différence entre ces différents indexicaux, en termes d'ordre d'entité visée, mais aussi de circonscription référentielle. Ainsi, nous observons une échelle descendante de référence partant du plus concret ([+ ordre 1 ; + circonscrit]) au plus abstrait ([- ordre 1 ; - circonscrit]) :

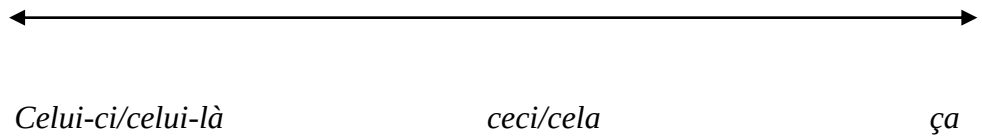
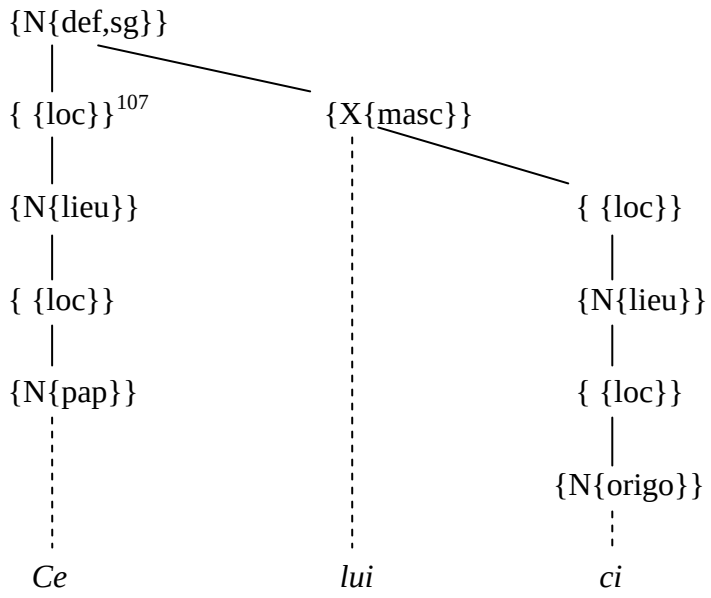


FIGURE 27 : UNE ECHELLE DESCENDANTE DE REFERENCE ENTRE *CELUI-CI/LA* ET *ÇA*

Voici quelles pourraient être les représentations des différents pronoms démonstratifs français :

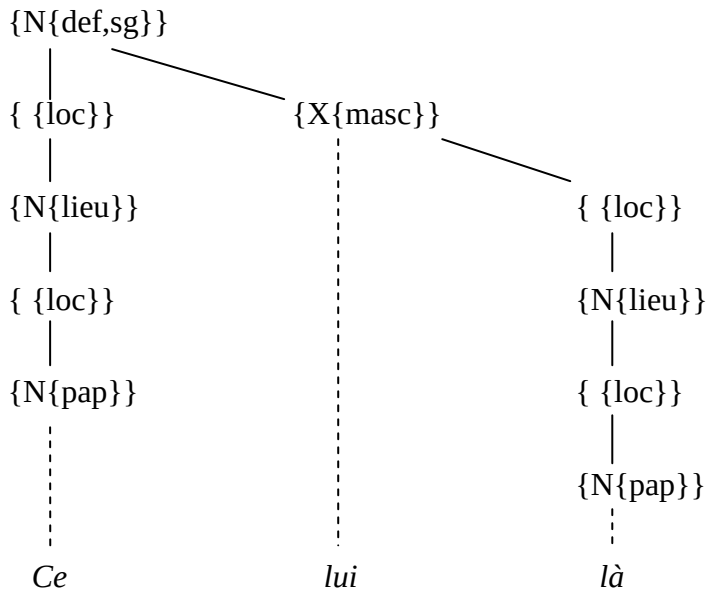
(75) *Celui-ci*



Ce est un déterminant qui pointe entitativement de façon définie et localisée par rapport à l'énonciateur ; *ce* étant un démonstratif, nous retrouvons la double construction référentielle, représentée par les deux catégories fonctionnelles { {loc}}. Dans ce contexte, *lui* est simplement une sorte de nom à référence vague puisque son seul trait grammatical est {masc}, pour 'masculin' ; néanmoins, *celui-ci* pouvant renvoyer à un référent doté du trait [+ humain] et premièrement évoqué par un nom propre – par exemple : *Ah ! Mais c'est Pierre qui a fait ça ! Il m'énerve, celui-ci !* – nous ne projetons pas *lui* comme {N;P}, ce qui signifierait que *celui-ci* ne peut se substituer qu'à un seul nom commun, mais comme {X}. *Ci* est interprété comme « *ici* → à ce lieu localisé par rapport à *ego* ».

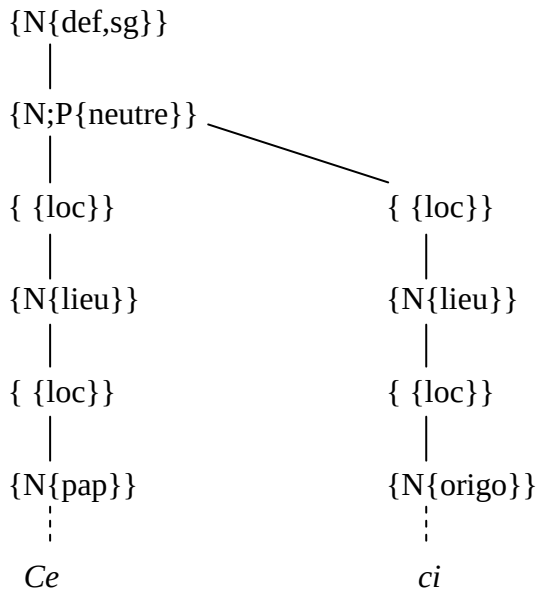
[107] Rappelons ici que { {loc}} est un foncteur de localisation, que nous pourrions aussi écrire {F{loc}}.

(76) *Celui-là*

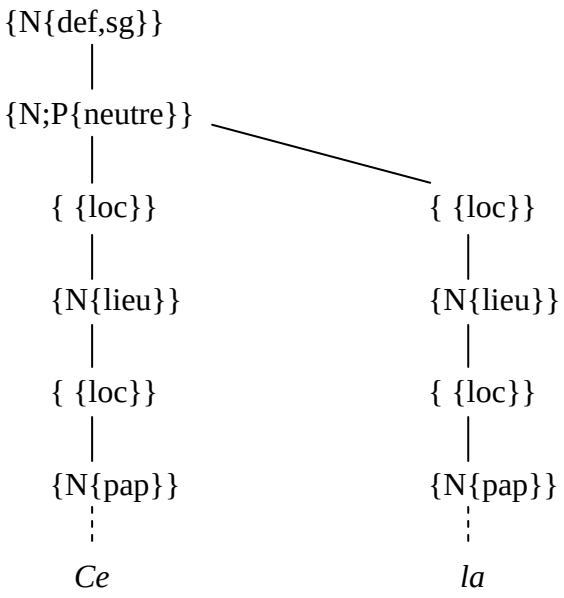


En (77), la différence avec la représentation précédente est la présence de *là* interprété comme localisé par rapport à {pap}. {pap} est plus général que {origo}, puisque {origo} est une valeur de {pap} (cf. chapitre VII, section 5.2, lors de la représentation de *this* et de *that* en GND), d'où la valeur non marquée, plus générale, de *celui-là* par rapport à *celui-ci*.

(77) *Ceci*



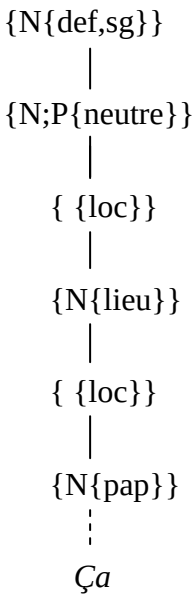
(78) *Cela*



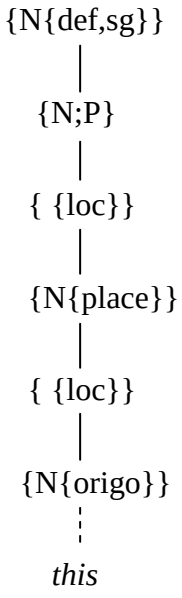
Concernant *ceci* et *cela*, la composante lexicale est sous-jacente, 'incorporée' en termes de GND. Ainsi, nous notons un {N;P} directement projeté par *ce*, et non pas adjoint de manière linéaire comme c'est le cas pour *celui-ci* et *celui-là* ; de plus, *ceci* et *cela* ne renvoyant jamais à un référent premièrement évoqué par un nom propre, nous pouvons conserver la notation {N;P}.

Nous pouvons alors réitérer la représentation de *ça* en GND, ainsi que celles des démonstratifs anglais, afin de présenter l'intégralité de nos systèmes :

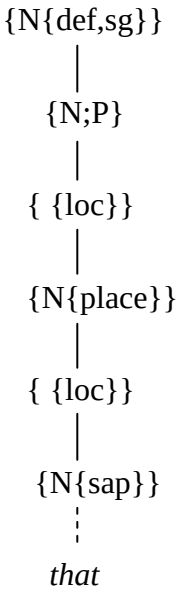
(79) *Ça*



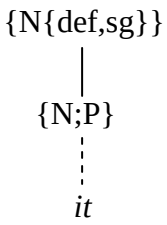
(80) *This*



(82) *That*



(83) *It*



Nous constatons que les représentations de *ça* et de *that* sont identiques : ces deux indexicaux incorporent les traits de localisation { {loc}} {N{place}} (ou {lieu}) et { {loc}} {N{sap}} (ou {pap}). La différence entre les deux réside dans un état de fait et son corollaire : *that* est l'un des membres du binôme *this/that*, tandis que *ça* est le seul élément de son système. Ainsi, le fonctionnement, et par conséquent la représentation, de *that* prend toute sa dimension lors d'une mise en contraste avec le membre marqué *this*. Ainsi, si *this*, *that* et *ça* sont tous trois des indexicaux neutres – ce qui fait qu'ils peuvent absorber tous les cas grammaticaux et renvoient difficilement à de l'animé humain –, *this* et *that* sont marqués par la complémentarité proximité/distance, tandis que *ça*, lui, est totalement neutre. De plus, le trait de neutralité est marqué en français, puisque les autres éléments de la même famille indexicale de *ça* fonctionnent par paires proximal/distal : *celui-ci* et *celui-là*, *ceci* et *cela*. Bien entendu, nous pourrions relever ici la présence de *ce*, également neutre, et complémentaire distributionnel de *ça* ; toutefois, *ce* tend à se rapprocher du défini – preuve en est son incapacité à fonctionner seul comme argument, en dehors de son association avec la copule. Dès lors, nous pourrions d'ailleurs (re)mettre en question l'incorporation de deux traits { {loc}} à *ce*.

Ainsi, il résulte un fonctionnement indexical et une représentation notionnelle de *ça* qui couvrent ceux de *that* et de *it*. De manière prototypique, *ça* et *it* sont comparables par la totale neutralité dont ils font preuve, neutralité qui leur permet d'absorber tous les rôles sémantiques et d'être des manifestations syntactico-sémantiques de l'hypostatisation de la langue et de la sous-spécification des éléments du lexique. *Ça* et *that* sont comparables par leur fonctionnement référentiel, caractéristique des démonstratifs, que nous avons décrit comme étant un mouvement rétroactif sur le repère déictique dont l'origine est l'ensemble des participants à l'acte communicatif. Dès lors, l'association des caractéristiques syntactico-sémantiques de *it* et de celles de *that* permet à *ça* de couvrir le continuum des procédures indexicales.

Chapitre IX. Perspectives de recherches

Ce chapitre se veut une ouverture, ou plutôt des ouvertures possibles, à la thèse qui vient d'être exposée. Notre étude des pronoms démonstratifs se limite aux domaines de la synchronie, de l'oral, mais aussi des deux langues que sont l'anglais et le français. Selon nous, il serait intéressant d'approfondir d'autres pistes en rapport avec les indexicaux, notamment l'évolution des démonstratifs, de leurs origines à nos jours, l'incidence que peuvent avoir les facteurs phonologiques sur la référence des démonstratifs, ainsi que l'exploration typologique de ces items. Voici ce que pourrait nous offrir la prise en compte de chacun de ces domaines.

1. Des perspectives concernant la prise en compte de la diachronie

1.1. Un bref historique du développement des démonstratifs en anglais, du V^{ème} siècle à nos jours

Tout comme aujourd'hui, le système des démonstratifs en vieil-anglais ne différenciait pas, d'un point de vue morphologique, les déterminants des pronoms. En revanche, la forme distale (*THAT*, de nos jours) et la forme proximale (*THIS*) étaient déjà marquées :

	Masculin	Neutre	Féminin	Pluriel
Nominatif	*se*	*þæt*	*sēo*	*þā*
Accusatif	*þone*		*þā*	
Génitif	*þæs*		*þāre*	*þāra, þāra*
Datif	*þām*			*þām*
Instrumental	*þȳ, þon*			

TABLEAU 20 : LE SYSTEME DES DEMONSTRATIFS 'DISTAUX' EN VIEIL-ANGLAIS (VEME-FIN XIEME)

	Masculin	Neutre	Féminin	Pluriel
Nominatif	*þes*	*þis*	*þēos*	*þās*
Accusatif	*þisne*		*þās*	
Génitif	*þisses*		*þisse, þisre*	*þissa*
Datif	*þissum*			*þissum*
Instrumental	*þȳs*			

TABLEAU 21 : LE SYSTEME DES DEMONSTRATIFS 'PROXIMAUX' EN VIEIL-ANGLAIS (VEME-FIN XIEME)

Nous pouvons remarquer que le singulier distinguait les genres masculin, féminin et neutre, et qu'il existait cinq marques casuelles. Si *se* a donné naissance au *the* actuel, *þæt* a donné *that*.

Le passage du vieil-anglais au moyen-anglais coïncide avec l'invasion de la Grande-Bretagne par Guillaume Le Conquérant, en 1066. Les changements notables qui se produisent, concernant les démonstratifs, sont la perte de la flexion marquant le genre et le cas, ainsi que l'émergence de l'article défini indéclinable *þe*, à partir de *se*. D'ailleurs, *that* persistait parfois alors en tant qu'article défini, devant un mot ayant une voyelle à l'initiale – *that one, that other*. Ainsi, au XV^{ème} siècle, le système des démonstratifs était quasi identique à celui que nous connaissons aujourd'hui :

Singulier	Pluriel
þat	*þas, þos*
þis	*þise, þese*

TABLEAU 22 : LE SYSTEME DES DEMONSTRATIFS EN MOYEN-ANGLAIS (FIN XIEME-XVEME)

It, quant à lui, proviendrait de la racine proto-indo-européenne **ki*- – cette dernière étant une variante de **ko*-, dont est issu l'ensemble des pronoms personnels en anglais – origine de *this/here*. La forme *hit* en vieil-anglais serait devenue *it* par la perte du *h*- initial au début du moyen-anglais. Voici donc un tableau récapitulatif de l'évolution de *this*, *that*, *it* et *the* :

PIE	***þa* (proto-germanique) + -s** **Hypothèse : serait aussi à l'origine du vieil-anglais *seo*, forme impérative de *see* (*behold*)**	**tod-* (base pronominale) ****þa* (proto-germanique)**	**so-* (*this, that*)	**ki-* (*this*/*here*, opposé à *that/there*)
Vieil-anglais	*þis*	*þæt*	*se* (fem. of *þæt*)	*hit*
Moyen-anglais	*þis*	*þat*	*þe*	*it*
Anglais moderne	*this*	*that*	*the*	*it*

TABLEAU 23 : L'EVOLUTION DIACHRONIQUE DE *THIS*, *THAT*, *IT* ET *THE*

Nous notons ici que les racines indo-européennes sont, bien entendu, hypothétiques. Néanmoins, nous trouvons surprenant que *it* ait une racine potentielle renvoyant à *this/here*. Si tel est le cas, ceci implique que *it* est intrinsèquement lié au repère *origo-here-now*, ce qui conforte alors notre définition de la deixis et de l'anaphore. Pour rappel, *this* est un démonstratif qui réfère par un retour sur le repère *origo-here-now*, ce qui fait que cet indexical est en emploi déictique si nul référent n'est présent dans la mémoire de l'allocutaire. *It*, quant à lui, est un pronom personnel clitique typiquement anaphorique, qui opère par un seul envoi référentiel vers la mémoire focus de l'allocutaire. *This* et *it* sont donc intimement liés au repère *origo-here-now*, la différence étant l'absence ou la présence de référent potentiel en mémoire focus, ainsi que le choix induit de l'indexical par l'énonciateur, suivant le besoin d'effectuer ou non de retour sur le repère déictique. *This* et *it* sont ainsi les représentants canoniques des pôles extrêmes reliant la deixis et l'anaphore. Par ailleurs, l'hypothèse stipulant que la racine proto-germanique de *this* est également celle du vieil-anglais *seo* – impératif de *see*, renvoyant initialement à *behold*, soit 'détenir' – conforte notre idée d'une sphère personnelle pour *this*.

1.2. Un bref historique du système des démonstratifs en français, du latin à nos jours

Contrairement au système des démonstratifs anglais, qui a somme toute évolué seulement de manière symétrique depuis le V^{ème} siècle – les paradigmes actuels *this* et *that* sont à la fois déterminants et pronoms, tout comme l'étaient *þis* et *þæt* –, il n'en va pas de même pour les démonstratifs français. En effet, le système uni-paradigmatique des démonstratifs latins a donné naissance aux démonstratifs actuels, déterminants et pronoms étant morphologiquement distincts. D'autre part, tout comme en anglais, les définis ont émergé de ces anciens démonstratifs.

1.2.1. Du latin à l'ancien français (jusqu'au IX^{ème} siècle)

Le système des démonstratifs latins est composé de quatre séries de démonstratifs à la fois pronoms et déterminants, chacune déclinable en genre – masculin, féminin, neutre – et selon cinq cas : *is* (masc.)*/ea* (fém.)*/id* (neutre), *hic/hæc/hoc*, *iste/ista/istud*, *ille/illa/illud*. Ces paradigmes pouvaient initialement être employés comme déictiques et comme anaphoriques, mais la valeur déictique s'est progressivement estompée. Selon Serbat (1980 : 95), cette tendance à l'anaphore pourrait expliquer l'émergence de l'article défini *le/la/les* et des pronoms de troisième personne *il/le/lui* à partir de la série *ille/illa/illud*. Quoi qu'il en soit, les démonstratifs se sont vus préposés de l'adverbe *ecce* – que l'on pourrait traduire *voici* ou *voilà* selon Gaffiot (1934 : 569) –, créant ainsi un nouveau système : *ecce hic* (neutre), *ecce iste* (proximal), *ecce ille* (distal).

Des quatre séries de démonstratifs latins, il n'existe nul vestige de *is/ea/id*. En revanche, le clitique *o* de la forme neutre *hoc* s'est combiné avec *ecce* pour donner naissance au pronom neutre *ço/ce/c'*. Les paradigmes *ecce iste* et *ecce ille*, quant à eux, ont amené les séries *CIST* et *CIL* de l'ancien français.

1.2.2. De l'ancien français au moyen français (IX^{ème}-XIV^{ème} siècle)

Jusqu'au XIII^{ème} siècle environ, les démonstratifs en ancien français sont composés des formes neutres *ço/ce/c'*, uniquement pronominales, et de *CIST* (proximal) et *CIL* (distal), pronoms et déterminants[108] :

CIST	Masc.sing.	Masc.plur.	Fém.sing.	Fém.plur.	Neutre
Cas sujet (nominatif)	*cist*	*cist*	*ceste*	*cestes, cez*	*cest*
Cas régime 1 (accusatif)	*cest*	*cez*	*ceste*	*cestes, cez*	*cest*
Cas régime 2[109]	*cestui*		*cesti*		

CIL	Masc.sing.	Masc.plur.	Fém.sing.	Fém.plur.	Neutre
Cas sujet (nominatif)	*cil*	*cil*	*cele*	*celes*	*cel*
Cas régime 1 (accusatif)	*cel*	*ceus*	*cele*	*celes*	*cel*
Cas régime 2	*celui*		*celi*		

TABLEAU 24 : *CIST* ET *CIL* JUSQU'AU XIIIEME SIECLE

Bien entendu, la différence sémantique entre *CIST* et *CIL* est l'objet de bon nombre d'études et de théories. Expliquée tout d'abord en termes d'opposition spatio-temporelle dont le référentiel est le locuteur (Mathews, 1907), puis en fonction du genre dialogique ou narratif du texte considéré (Yvon, 1951), c'est en 1985 que Kleiber apporte un regard nouveau à l'analyse de *CIST* et *CIL*. En effet, Kleiber s'éloigne du point de vue dualiste entre ces deux séries de démonstratifs et envisage *CIST* comme étant la forme marquée de *CIL*. Vingt ans

[108] Nous ne traitons pas des formes préfixées en *-i – icist, iceste, icil, icele*, etc.

[109] L'ancien français oppose le cas sujet (originaire du nominatif latin) et le cas régime (provenant de l'accusatif latin). Néanmoins, les démonstratifs admettent une deuxième forme de ce cas régime, issu du datif latin, notée ici « cas régime 2 ».

plus tard, la complémentarité entre *CIST* et *CIL* acquiert une définition nettement plus abstraite que celles proposées jusqu'alors, avec Marchello-Nizia (2004, 2005), qui propose une explication cognitive et pragmatique à l'emploi de *CIST* et *CIL* ; en effet, *CIST* serait préféré lorsque le locuteur réfère à quelque chose qui lui est propre, qu'il affectionne, ou même à des paroles qu'il vient juste de prononcer ; à l'inverse, *CIL* aurait tendance à être employé si le locuteur souhaite se distancier du référent évoqué. Ainsi, le caractère proximal ou distal communément prêté aux démonstratifs serait plus lié à une stratégie discursive délibérément choisie par le locuteur qu'à un trait sémantique. Cette hypothèse semble être étayée par les travaux de Guillot (2006) sur le français médiéval, qui révèlent que *CIST* est le démonstratif choisi lorsque le syntagme dans lequel il s'inscrit véhicule un jugement additionnel du locuteur. Ainsi, *CIST* est toujours employé lorsque l'élément lexical du SN est un terme métadiscursif, comme l'illustre l'exemple suivant :

(1) *« […] et quant nos avrons les vïandes recuillies, alomes devant la ville et fesons ce que Nostre Sires nos avra porveü. Quar plus seürement guerroie cil qui a la vïande que cil qui n'en a point ». A* ***cest*** *conseil s'acorderent li conte et li baron, et s'en ralerent tuit a lor nés chascuns et a ses vaissiaus*. (exemple (13) de Guillot (2010) en citation de Villehardouin, *Conqueste*, t.1 : 132)

Par l'emploi du lexème *conseil*, le locuteur exprime l'avis personnel qu'il porte sur l'acte langagier pouvant correspondre au contenu propositionnel auquel il se réfère. Par conséquent, les connotations, positives ou négatives, que pourrait véhiculer l'élément lexical considéré sont l'effet d'un choix imputable au locuteur ; dans ce cas précis, nous pourrions donc supposer que modalisation énonciative et sphère du locuteur sont deux concepts allant de pair. Ainsi, la notion de proximité attribuée à *CIST* peut très bien s'accommoder d'une définition résolument cognitive.

Aux alentours du XIII^{ème} siècle apparaît la série *ce/ces*, seulement employée en tant que déterminant. Ces morphèmes se neutralisent très vite, à la fois au niveau sémantique – la différence entre proximité et distance s'efface – et au niveau du genre, la forme masculine prenant le pas sur la forme féminine :

	Masc.sing.	Masc.plur.	Fém.sing.	Fém.plur.
Cas sujet (nominatif)				*ces*
Cas régime (accusatif)	*ce* + consonne	*ces*		*ces*

TABLEAU 25 : LA SERIE DES DETERMINANTS NEUTRES *CE/CES* AU XIIIEME SIECLE

Comme nous pouvons le constater d'après le tableau ci-dessus, *ce* est le déterminant masculin du cas régime singulier ; l'autre forme de ce déterminant est *ces*. L'origine de cette forme neutre déterminative soulève de nombreux débats, mais les travaux de Dees (1971) semblent apporter des réponses convaincantes. En effet, l'hypothèse majeure avancée par les grammaires du XIX^{ème} siècle présente *ces* – prononcé [sɛs] – comme étant issu de la forme *cez* – prononcée [sɛts] – du paradigme *CIST*. Le XIII^{ème} siècle étant la période durant laquelle les occluso-constrictives ont disparu du système consonantique du français, *ces* [sɛs] serait donc le résultat de la perte de l'affriquée [ts]. Ceci dit, les analyses discursives de *ces* semblent montrer que ce déterminant relèverait plutôt du paradigme *CIL*. Par conséquent, comment justifier l'émergence de *ces* du seul paradigme *CIST* ? L'origine d'une forme et son contexte d'emploi appartenant chacun à deux séries paradigmatiques complémentaires pourraient paraître contradictoires. Pour pallier cette contradiction, Dees (1971) propose que *ces* proviendrait non seulement du paradigme *CIST*, mais également de *CIL*. En effet, la forme *cels* de la série *CIL* aurait évolué en *ces* par vocalisation de la liquide [l], parallèlement à *dels* et *als* qui ont donné *des* et *as*. Enfin, la forme *ce* de la paire *ce/ces* résulterait d'une analogie avec *le/les*.

Vers la fin du XIII^{ème} siècle, le paradigme en *CIST* disparaît progressivement pour laisser la place à *CIL*. Ainsi, la distinction sémantique entre proximité et distance est neutralisée. Le système des démonstratifs jusqu'au début du XIV^{ème} siècle se compose alors de *CIL*, à la fois pronom et déterminant, des pronoms neutres *ço/ce/c'* – qui eux, n'ont pas évolué depuis le IX^{ème} siècle – et du déterminant neutre *ce/ces*.

1.2.3. Du moyen français au français moderne (fin XIV^{ème}-aujourd'hui)

De la même manière que *ille*, en latin vulgaire, s'est vu préposé de *ecce* afin de ré-instancier la monstration propre aux démonstratifs, *ce* déterminant peut désormais être

postposé de *ci* ou *la*, particules adverbiales respectivement proximale et distale : *ce-ci* et *ce-la*. *Ci* et *la* auraient en premier lieu été adjointes à *ceus*, du paradigme *CIL*, afin de ré-instancier l'opposition proximité/distance disparue en même temps que la série *CIST*. Par glissement, les formes déterminatives neutres *ces* et *ce* ont pu s'accommoder de *ci* et *la*. Par ailleurs, la suite *CIL* disparaît à son tour, dès le début du XV^{ème} siècle. Ainsi, la fin du moyen français présente un système de démonstratifs similaire à celui que nous connaissons, composé de *ce* déterminant neutre, de *ço* pronom neutre, et des formes *ce-ci* et *ce-la* – devenues *ceci* et *cela* par enclise de la particule initialement postposée. Au XVII^{ème} siècle, tous les pronoms démonstratifs s'analysent en *ce* forme neutre combinée à l'un des pronoms personnels toniques : par exemple, *celui* se compose de *ce* + *lui*, *celle* de *ce* + *elle*. Par la suite, tous les pronoms personnels sont potentiellement augmentés de la particule *ci* ou *là*, comme, par exemple, *celui-ci*, *celui-là*, *celle-ci*, *celle-là*.

1.3. Quelles pistes la perspective diachronique pourrait-elle nous offrir pour le traitement des pronoms démonstratifs en français ?

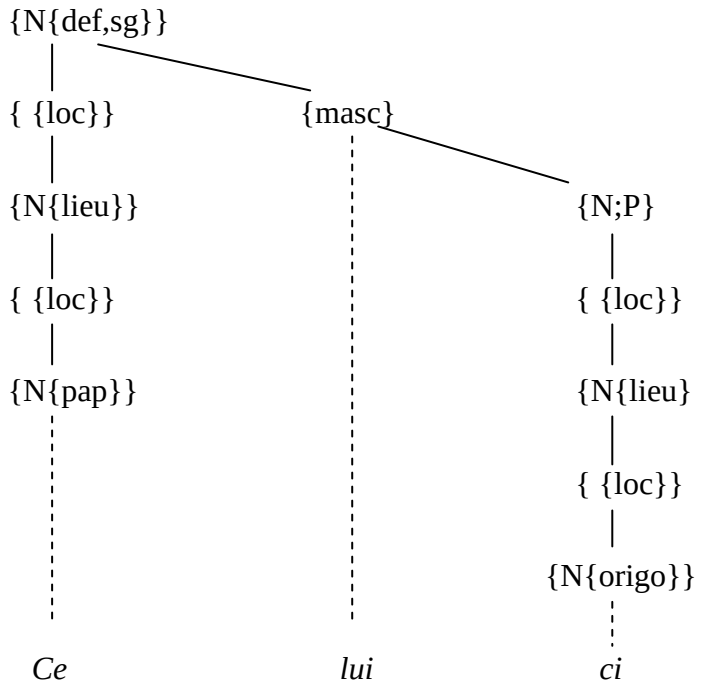
Que peut nous apporter ce parcours historique ? Tout d'abord, nous remarquons que l'évolution des éléments de détermination est cyclique en français. En effet, le phénomène de grammaticalisation des déterminants et pronoms démonstratifs se traduit par une perte de la faculté de monstration, une neutralisation de la localisation spatiale et une disparition de la pronominalisation. En d'autres termes, les démonstratifs évoluent vers le simple déterminant défini neutre. Par deux fois depuis la période latine, la dé-sémantisation des démonstratifs s'est vue comblée par la réinstauration des propriétés initiales, par l'ajout préfixé de *ecce* tout d'abord, de *ci* et *là* suffixés ensuite. Ainsi, de nos jours, et ce depuis le XVII^{ème} siècle, tous les pronoms démonstratifs s'analysent en *ce* forme neutre + pronom personnel tonique pour *celui*, *celle(s)* et *ceux*, en *ce* forme neutre + pronom personnel + *ci/là* pour *celui-ci/là*, *celle(s)-ci/là* et *ceux-ci/là*, en *ce* forme neutre + *ci/là* pour *ceci* et *cela*. Nous avons déjà constaté que les formes ne comportant pas *ci* ou *là* (*ce* et *celui/celle(s)/ceux*) ne fonctionnent pas seules comme argument mais nécessitent un complément :

(2) *Ce que tu me dis est absurde* / **ce est absurde*
(3) *Celui/celle de mon frère est bien mieux* / **celui/celle est bien mieux*

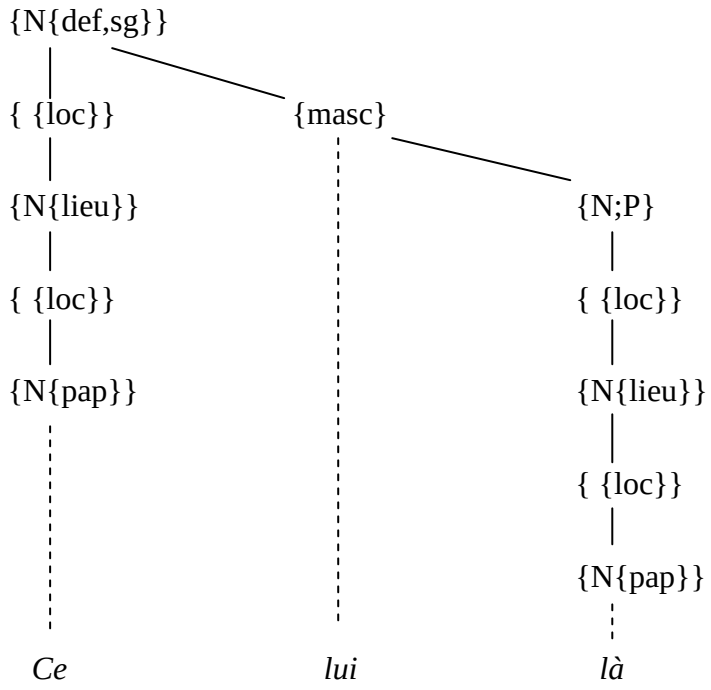
Il faut donc expliquer pourquoi *ci/là* semblent être responsables de la pronominalisation des indexicaux démonstratifs. De deux choses l'une : soit nous attribuons la tête nominale à *ci/là*,

soit nous l'incorporons à *ce*. Dans le premier cas, il faudrait considérer *ci/là* comme suffixés au déterminant neutre, ce qui nous mènerait à supposer la tête lexicale sous-jacente à *ci* et *là* et ce qui donnerait alors une formalisation en GND de ce type :

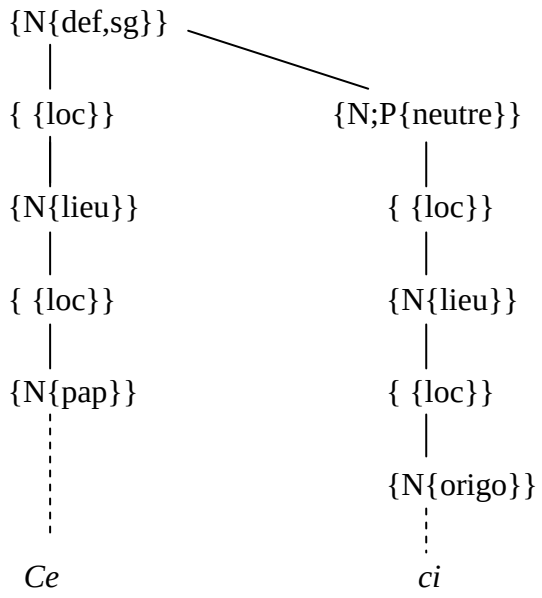
(4) *Celui-ci*



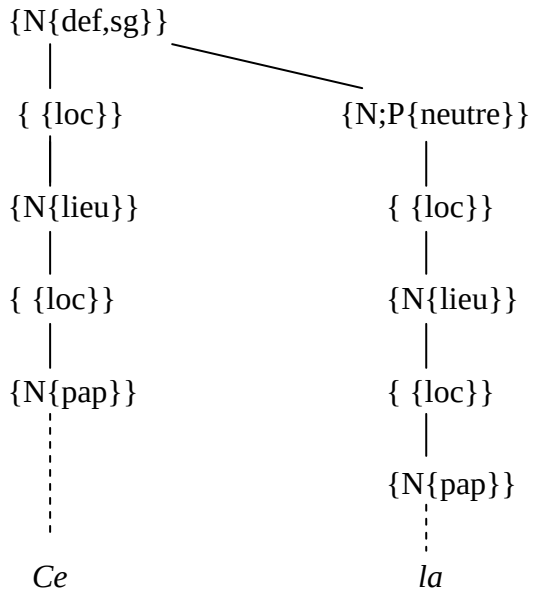
(5) *Celui-là*



(6) *Ceci*



(7) *Cela*



Bien entendu, une telle hypothèse est contre-intuitive pour la raison simple que *ici* et *là*, d'où proviennent directement *ci* et *là* suffixés, sont des adverbes alors que l'étiquette {N;P} les assimile à des substantifs. En revanche, *ici* et *là*, ainsi que l'ensemble des marqueurs de localisation, fonctionnent non seulement comme adverbiaux mais aussi comme pointeurs entitatifs :

(8) A : *Où voudrais-tu qu'on se retrouve ?* B : *Près de l'université me conviendrait*.

Cette particularité explique sans doute pourquoi *ci* et *là* suffixés suffisent à rendre aux démonstratifs leur faculté de pronominalisation. Dès lors, nous conservons l'hypothèse de la tête lexicale sous-jacente incorporée à *ce* (dans la lignée des analyses de Price, 1971) : *ceci* et *cela* sont neutres, en ce que la marque du genre et du nombre est effacée ; en revanche, *celui-ci/là*, *celle(s)-ci/là* et *ceux-ci/là* sont caractérisés par les traits syntactico-sémantiques du genre et du nombre, ce qui traduit la plus grande restriction référentielle qu'ils manifestent. De même, *ce* et *celui/celle/ceux* nécessitent un complément de par l'absence de marque de localisation. Par conséquent, *ce* et *celui/celle/ceux* manifestent un stade plus avancé de grammaticalisation que leurs homologues suffixés en *ci/là* ; cette hypothèse se vérifie par la faiblesse phonologique de ces formes, accentuée par la chute quasi systématique du schwa à l'oral, et par leur inclination à tendre vers le défini. Si nous considérons *ça* dans cette analyse, nous expliquons par la lourdeur phonologique et par la double localisation incorporée, spatiale et par rapport aux participants à l'acte communicatif – { {loc}} {N{lieu}} et { {loc}} {N{pap}} en GND –, la capacité qu'a ce pronom neutre d'agir seul comme argument. Enfin, cette dernière remarque conforte notre analyse des démonstratifs comme présentant un procédé référentiel double, représenté en GND par les deux traits { {loc}} {N{lieu}} et { {loc}} {N{pap}}.

2. Des perspectives liées à la prise en compte de la prosodie

Que syntaxe et prosodie, syntaxe et sémantique, sémantique et prosodie soient liées est indiscutable à nos yeux. Néanmoins, nous souhaitons soulever quelques questions liées à l'incidence que pourrait avoir la phonologie suprasegmentale sur l'acte de référence effectué par les démonstratifs. Pour cela, nous présentons de manière succincte l'annotation nécessaire à ce type d'étude, puis nous offrons quelques résultats préliminaires.

2.1. ToBI : un projet de conventions pour l'analyse des facteurs prosodiques

ToBI (Tones and Break Indices) est un projet d'annotation destiné à l'étude de l'intonation et des différents facteurs prosodiques pour un panel de langues défini. Etant donné que chaque langue possède sa propre structure mélodique, il existe un système ToBI pour chacune d'elles. En particulier, Pierrehumbert (1980) et Beckman & Pierrehumbert (1986) développent un ensemble de conventions pour l'étiquetage de l'intonation de l'anglais, qui se décrit comme une combinaison de tons hauts ('high tones', H) et de tons bas ('low tones', L). L'ensemble de ce système comprend deux niveaux d'unités tonales : l'unité

intonative et l'unité intermédiaire. L'unité intonative se compose d'une ou plusieurs unités intermédiaires et se termine par deux tons finaux : un ton qui marque la dernière syllabe du groupe intonatif et un autre qui s'applique à la syllabe suivant la dernière syllabe accentuée du même groupe. L'unité intermédiaire, quant à elle, ne présente qu'un seul ton final mais contient un certain nombre de syllabes accentuées, potentiellement réalisées selon six accents mélodiques possibles : H*, L*, L+H*, L*+H, H*+L et H+L*. Notons ici que le français ne fait pas partie des langues initialement décrites par le système ToBI ; toutefois, de nombreux travaux récents sur l'analyse prosodique du français utilisent ce système d'annotation (cf., par exemple, Mertens, 2004). En ce qui concerne nos observations préliminaires, nous utilisons seulement les symboles représentatifs des accents mélodiques, à la fois pour l'anglais et pour le français.

2.2. Les accents mélodiques dans le système ToBI

Dans le système ToBI, les accents mélodiques sont annotés par les symboles H et L – représentant, respectivement, un ton haut et un ton bas – ou par une combinaison de H et de L. L'astérisque * qui suit parfois l'un ou l'autre de ces symboles signale que la syllabe considérée est accentuée. Pour nos observations préliminaires, les seuls accents mélodiques H*, L* et L+H* sont utilisés ; voici donc un descriptif succinct de ces annotations :

- H* correspond à l'accent primaire en anglais et à l'accent lexical en français, H* étant l'accent par défaut. H* signale une proéminence mélodique qui est en réalité une montée de la fréquence fondamentale (F0) si H* se trouve au début de la syllabe proéminente, et une descente de la F0 si H* est situé à la fin de la syllabe proéminente :

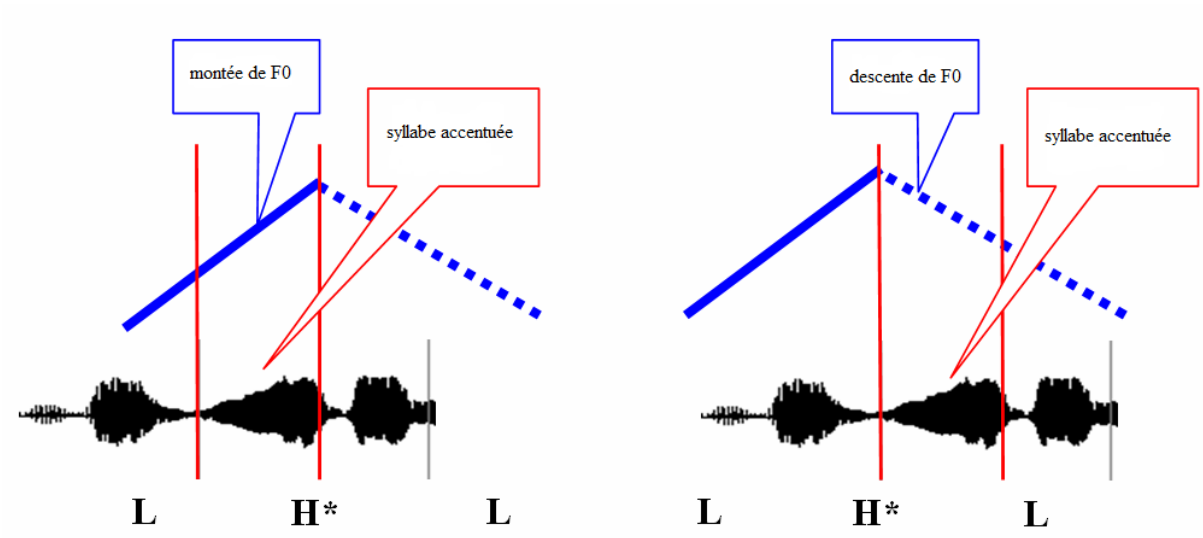


FIGURE 28 : UN TON H* EST LIE A UNE MONTEE MELODIQUE S'IL EST PLACE AU DEBUT DE LA SYLLABE PROEMINENTE, ET A UNE DESCENTE S'IL SE TROUVE A LA FIN (D'APRES MARTIN, 2008 : 3)

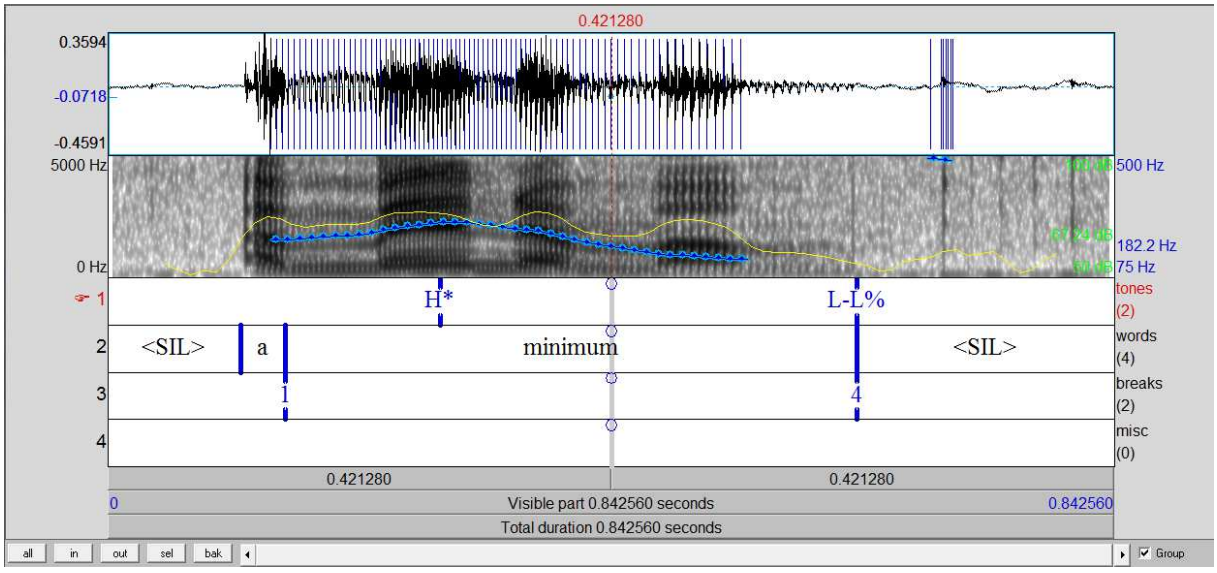


FIGURE 29 : L'ACCENT MELODIQUE H* PORTE PAR LA SYLLABE PROEMINENTE DE *MINIMUM* (CAPTURE D'ECRAN DU LOGICIEL PRAAT (CF. BOERSMA & WEENICK (2009) POUR UNE DESCRIPTION DE LA VERSION 5.1.05 DU LOGICIEL))

- L* est le contraire de H*, en ce qu'il signale une chute de la F0 si L* est placé en début de syllabe proéminente, et une montée de la F0 si L* est en fin de syllabe proéminente :

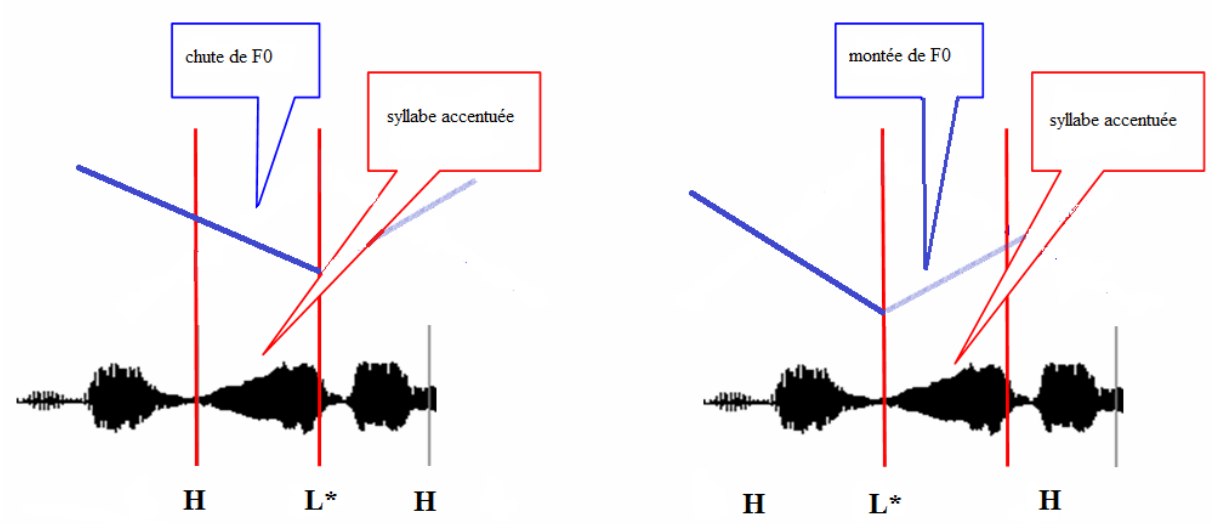


FIGURE 30 : UN TON L* EST LIE A UNE CHUTE MELODIQUE S'IL EST PLACE AU DEBUT DE LA SYLLABE PROEMINENTE, ET A UNE MONTEE S'IL SE TROUVE A LA FIN (D'APRES MARTIN, 2008 : 3)

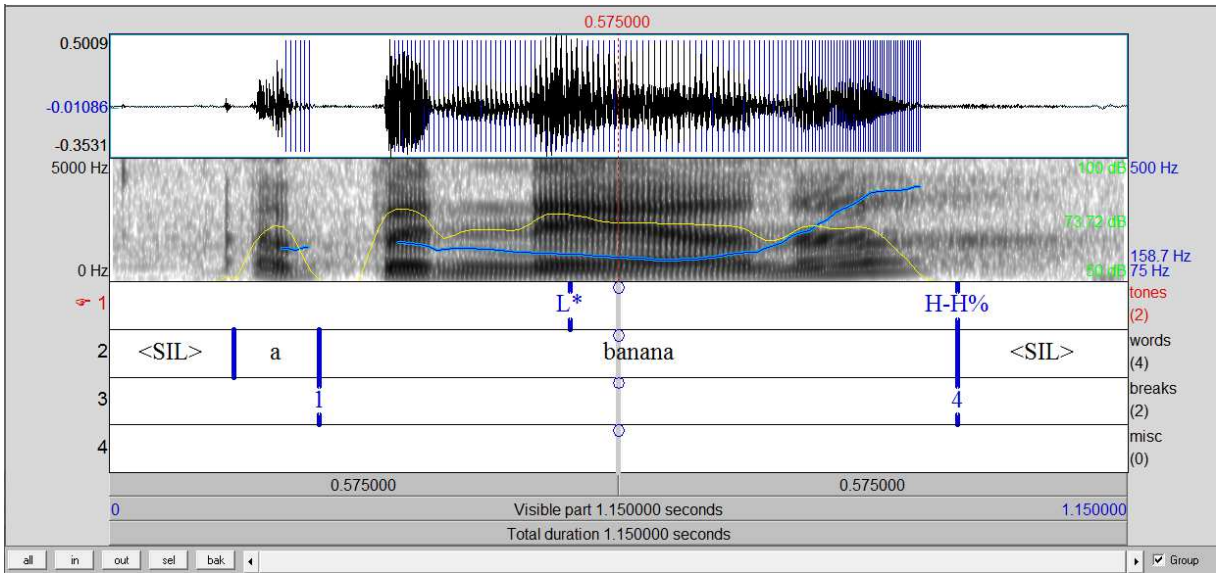


FIGURE 31 : L'ACCENT MELODIQUE L* PORTE PAR LA SYLLABE PROEMINENTE DE *BANANA* (CAPTURE D'ECRAN DU LOGICIEL PRAAT)

- Enfin, L+H* caractérise une syllabe proéminente (H*) immédiatement précédée d'une montée brusque à partir d'un L*. L+H* est typique de l'anglais et parfois nommé 'accent contrastif' :

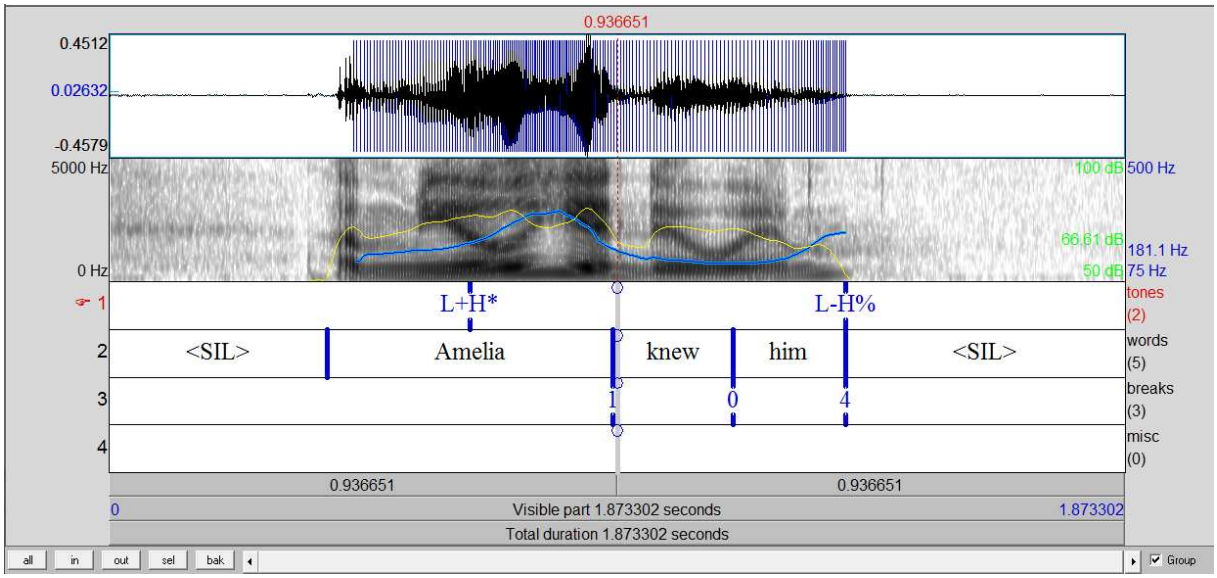


FIGURE 32 : L'ACCENT MELODIQUE L+H* PORTE PAR LA SYLLABE PROEMINENTE DE *AMELIA* (CAPTURE D'ECRAN DU LOGICIEL PRAAT)

2.3. Des corrélations possibles entre accentuation mélodique et accès à la référence

La recherche autour de la structure informationnelle conduit à décrire le lien qui pourrait exister entre cette dernière et la phonologie suprasegmentale (cf., entre autres, Jackendoff (1972) et Lambrecht (1994)). En particulier, il est noté que l'accent nucléaire – H* en notation ToBI – tombe sur la part de l'énoncé qui véhicule une information nouvelle. En ce qui concerne l'anglais, Kameyama (1999) propose un mode d'interprétation des pronoms accentués qu'il nomme 'Complementary Preference Hypothesis' (que l'on pourrait traduire par 'l'Hypothèse du choix référentiel complémentaire') :

> Complementary Preference Hypothesis (CPH) : « A focused pronoun takes the complementary preference of the unstressed counterpart. » (Kameyama, 1999 : 315, que l'on pourrait traduire par 'un pronom accentué renvoie à une entité contraire de celle à laquelle renverrait le même pronom inaccentué')

En d'autres termes, il existerait un ensemble de référents potentiels, ordonnés selon leur probabilité d'être visés par un pronom inaccentué, dans un énoncé donné ; l'accent nucléaire ou contrastif porté par ce pronom déclencherait une référence alternative à celle qu'aurait le pronom inaccentué :

(8) *Paul called Jim a Republican. Then he inSULted him*. (Paul insulted Jim)
H*L

Paul called Jim a Republican. Then HE insulted HIM. (Jim insulted Paul)
L+H* L+H*
(exemples (2) et (3) de Kameyama, 1999 : 307)

En (8), l'accent nucléaire porté par le prédicat induit que *he* sujet et *him* objet sont en coréférence avec *Paul* sujet et *Jim* objet ; à l'inverse, *he* et *him* accentués entraînent le renvoi de *he* à « Jim » et de *him* à « Paul ». Le référent correspondant à *he* et *him* accentués est donc alternatif à celui visé par *he* et *him* non accentués, si l'on considère que seuls deux référents potentiels sont présents. Une étude préliminaire menée en 2011 (Cornish & Buscail, 2011) tend à affiner cette relation entre accentuation et référence. En effet, l'analyse d'exemples oraux montre que l'identification référentielle d'un item accentué ne s'effectue pas irrémédiablement suivant l'entité désignée par son homologue inaccentué. Plus précisément, l'accent contrastif porté par un pronom entraînerait une mise en focus de l'argument et, en parallèle, l'identification d'un référent selon la relation de cohérence induite par cette accentuation marquée :

(9) [Extrait de l'émission « Any Questions ? », sur *BBC Radio 4*, 25.08.07. Question : Should the country apologise for the Slave Trade?]

Louise Bagshaw : /…I think it's just cheap point-scoring and party-political posturing to go back and say we're sorry for something that our ancestors did; it was for THEM to apologise, it was for THEM to look themselves in the mirror and
L+H* L+H*
be ashamed of what they did.

En (9), il apparaît que la locutrice souhaite exprimer un contraste entre deux groupes de personnes susceptibles de s'excuser pour la traite des esclaves : d'un côté, « les ancêtres des habitants actuels de la Grande-Bretagne, contemporains de la traite des esclaves » et, d'un autre côté, « les descendants de ces ancêtres, vivant de nos jours ». L'interprétation des pronoms *them* accentués sous-entend la suite contrastive *... and not US*, bien que ce pronom de première personne ne soit pas prononcé par la locutrice – *we* et *our* sont mentionnés mais inaccentués. En d'autres termes, le référent en contraste est inféré. Notons en passant que l'accent nucléaire porté par le prédicat – *it was for them to aPOlogise, it was for them to*

LOOK themselves in the mirror and be aSHAMEd of what they did – serait incohérent, en ce que l'information véhiculée est présupposée car déjà mentionnée par *we're sorry* ; or, l'accent nucléaire tend à correspondre à de l'information nouvelle. En (9), donc, le *them* accentué renvoie « aux ancêtres des habitants de Grande-Bretagne », qu'il met en contraste avec le référent implicite « les habitants actuels de Grande-Bretagne » ; il n'existe en aucun cas d'ensemble de référents alternatifs correspondant à *them* accentué et à *them* inaccentué.

Dès que l'on s'attache à prendre en compte l'accentuation mélodique portée par les pronoms démonstratifs et l'impact de celle-ci sur l'interprétation référentielle, les résultats sont plus nuancés. En effet, notre corpus présente une seule occurrence de *that* contrastif (L+H*), qui ne semble pas véhiculer de réel contraste :

(10) [Exemple (11) du chapitre VII]

> LC: *And because I was then only about twenty-one, er, they said that I ought to go to, two-year training college or three-year training college that was just coming up. Er, but* ***that*** **(L+H*)** *meant I had to get O levels*. (Corpus PAC Lancashire, code locuteur elalc1, conversation libre)

L'accent contrastif en (10) semble plutôt induire une seule mise en relief. Toutefois, nous notons que la majorité des *that* de notre corpus, et tous les *that* à fonction de récupération référentielle (cf. chapitre VII, section 3.3), portent l'accent nucléaire H*, qui est l'accent par défaut en anglais. Cette particularité est à mettre en relation avec *it*, qui, lui, n'est jamais accentué : *that* et *it* sont souvent interchangeables, l'un ou l'autre de ces indexicaux n'altérant pas le référent visé, mais la propension de *that* à porter l'accent nucléaire tend à mettre en relief l'argument et non le prédicat. De plus, cela soutiendrait *that* comme étant le membre marqué de la paire *that*/*it* : *it* ne pouvant être accentué, c'est *that* qui est choisi s'il y a volonté de mettre en focus le référent visé. Néanmoins, l'observation de *this* mène à une réflexion intéressante ; en effet, nous avons vu dans le chapitre VII que *this* décrit une sphère personnelle qui s'accompagne souvent de ce que nous avons nommé une « polarité prédicative positive » (cf. chapitre VII, section 3.1.2). Tous les énoncés relevant de cet emploi présentent un *this* accentué H* ; en revanche, nous notons deux occurrences de *this* L* qui s'accompagnent d'une polarité prédicative négative. *This* L* serait donc marqué :

(11) [Exemple (9) du chapitre VII]

GT: *But, you know bu/ I was looking in erm, in Selfridges, and, some of the stuff are (bits of scurves). And, it had all threads, hanging from the hand, like it was unfinished.* <RH: *Yeah.*> *And* ***this*** **(L*)** is, *(XXX) four hundred and seventy-five pounds.* (Corpus PAC Manchester, code locuteur emangt1, conversation libre)

(12) [Exemple (33) du chapitre VII]

DK: *I hate a f/ small,* <I: *Oh really?*> *group of Scottish people yeah, but. Yeah, the majority of them are okay, I do have Scottish friends up there and, they're quite like, chilled out and decent people but, there's the odd few that like, tar everybody else. You walk into the local pub, where I live, er, ordered a pint in English, accent, didn't try and change my accent in any way and it just fell silent, the pub, is just completely dead, everybody turns around, 'What are you doing in our pub?' type of thing, it's like 'Okay, I'll leave',* (laughter) <I: *Oh, right.*> *so, yeah, so. It's strange to be on the receiving end of that, because especially in England it's Asians that get, that side of like, if you've got pale skin then you/ you're fine here but,* ***this*** **(L*)** *is really weird, to receive that in Scotland, but.* (Corpus PAC Lancashire, code locuteur eladk1, conversation guidée)

L'étude de la corrélation entre l'accent mélodique porté par le démonstratif et les conséquences référentielles ou communicatives associées mérite d'être approfondie. Néanmoins, nous pourrions proposer l'hypothèse suivante :

L*	H*	L+H*
This	accent primaire	*that*
That		contraste
Prédication à	*this*	
caractère négatif	*that*	
ou polarité négative	contraste pour *this*	

FIGURE 33 : DES HYPOTHESES SUR LE LIEN ENTRE ACCENTUATION ET REFERENCE DES DEMONSTRATIFS EN ANGLAIS

Il y aurait une échelle accentuelle allant de L* à L+H*. L* sur un démonstratif entraînerait un énoncé à polarité prédicative négative. H* et L+H* formeraient une paire dont le membre marqué serait L+H*, H* étant l'accent nucléaire par défaut. Ainsi, L+H* ne pourrait s'accorder d'un *this*, *this* étant lui-même déjà marqué. En revanche, L+H* exprimerait un contraste pour *that* – de manière prototypique, bien entendu –, tandis que H* serait suffisant à produire un contraste à partir de *this*. Enfin, *that* H* serait à la fois la valeur par défaut de *that* et la forme marquée de *it* inaccentué. Cela va sans dire, toutes ces suppositions nécessitent des analyses approfondies.

En ce qui concerne le français, l'accentuation lexicale semble plus sensible à la position syntaxique que cela ne l'est pour l'anglais. Un exemple notoire est la construction à double sujet, où l'argument en P1 reçoit un accent primaire tandis que le sujet clitique en P2 est inaccentué. Bien entendu, cette observation est valable pour *ça*, ce qui entraîne un *ça* inaccentué en position P2 et un *ça* accentué en P1. De manière générale, nous notons dans notre corpus que *ça* est inaccentué lorsque son emploi discursif tend vers l'anaphore et que *ça* est accentué lorsqu'un phénomène de création référentielle s'opère. En d'autres termes, l'absence ou la présence d'un accent lexical sur *ça* irait de pair avec la procédure anaphorique ou déictique en jeu.

3. Et si nous nous intéressions aux autres langues du monde ?

Comme nous l'avons souligné dans le chapitre II, certaines langues ne différencient pas les déterminants des pronoms démonstratifs (143 langues dans le monde) ; d'autres, en revanche, possèdent des formes distinctes, qu'elles correspondent à des morphèmes différents (37 langues) ou à des variations flexionnelles (21 langues). En ce qui concerne la marque de distance, il est intéressant de noter que certaines langues qui différencient déterminants et pronoms démonstratifs les démarquent aussi en fonction de cette distance. Par exemple, le tongan (langue polynésienne) possède le système suivant :

Pronoms démonstratifs		Déterminants démonstratifs	
Neutre	*e*		
Près du locuteur	*eni*	Proximal	*ni*
Près de l'allocutaire	*ena*	Distal	*na*
Loin du locuteur et de l'allocutaire	*ia*		

TABLEAU 26 : LE SYSTEME DES DEMONSTRATIFS EN TONGAN (CHURCHWARD, 1953 : 150-152)

Cette caractéristique est également valable pour le français : les déterminants démonstratifs *ce*, *cette* et *ces* sont neutres, contrairement aux pronoms *CELUI-ci/là*, *ceci* et *cela*.

De manière générale, la grande majorité des langues du monde présente un contraste entre deux ou trois marques de distance différentes (215 langues au total sur 234, soit 92%). Voici une classification plus détaillée :

Pas de marque de distance	7	3%
Contraste binaire	127	54.3%
Contraste ternaire	88	37.6%
Quatre marques de distance	8	3.4%
Cinq marques de distance ou plus	4	1.7%
	234	100%

TABLEAU 27 : LES MARQUES DE DISTANCE DES DETERMINANTS DEMONSTRATIFS DANS LES DIFFERENTES LANGUES DU MONDE (DIESSEL, 2011)

Le contraste entre quatre, cinq, voire plus de marques de distance est donc un épiphénomène. Regardons de plus près le fonctionnement de trois langues particulières, à savoir le hunzib (parlé en Russie), le japonais et le hausa (parlé au Niger) :

	Déterminants démonstratifs en hunzib
Proximal	*b ǝd*
Médian	*b ǝl*
Distal	*ǝg*

	Déterminants démonstratifs en japonais
Près du locuteur	*kono*
Près de l'allocutaire	*sono*
Loin du locuteur et de l'allocutaire	*ano*

	Déterminants démonstratifs en hausa
Près du locuteur	*nân*
Près de l'allocutaire	*nan*
Loin du locuteur et de l'allocutaire	*cân*
Très loin du locuteur et de l'allocutaire	*can*

TABLEAU 28 : LE SYSTEME DES DETERMINANTS DEMONSTRATIFS EN HUNZIB, EN JAPONAIS ET EN HAUSA (RESPECTIVEMENT, VAN DEN BERG, 1995 : 61 ; KUNO, 1973 : 27 ; WOLFF, 1993 : 119-120)

En termes de distance, nous constatons que le système ternaire impose deux ou trois marques possibles : proximal/médian/distal (hunzib) ou proximal/distal (japonais). Le hunzib se décline en quatre démonstratifs pour trois marques de distance : proximal/distal/très distal – à la manière de *this*, *that*, *yon* pour certaines régions du nord de l'Angleterre et de l'Ecosse. Les marques de localisation alimentent une prise de position qui découle également de l'étude de nos démonstratifs : plutôt que de considérer les traits 'proximal' et 'distal' comme des valeurs parallèles, il serait peut-être plus juste de les concevoir comme les deux pôles extrêmes d'un continuum. Ainsi, la sphère personnelle de *this* est marquée par rapport à la sphère inter-personnelle de *that*, mais celle de *that* comprend celle de *this* ; la distance 'médian' en hunzib est à mi-chemin du continuum entre les valeurs 'proximal' et 'distal' ; *can* en hausa se situe à l'extrêmité 'distal'. En termes de traits syntactico-sémantiques, cette observation reviendrait à établir le fait que 'distal' est une primitive localisatrice ; ceci expliquerait également pourquoi

l'évolution cyclique des démonstratifs en français manifeste la perte de la marque proximale au profit de la marque distale, qui de ce fait, se dé-sémantise en trait neutre. Nous notons ici que le trait de localisation des démonstratifs n'est pas le seul à révéler ce phénomène d'inclusion, puisque ce qui est particulièrement notable à travers les langues du monde est le centre déictique par rapport auquel se situent ces distances relatives : le locuteur, l'allocutaire, mais aussi le locuteur et l'allocutaire. Il faudrait donc considérer trois origines de repère énonciatif : *origo* ou *ego, tu* et l'ensemble des participants à l'acte langagier – {origo}, {tu} et {pap} en termes de GND, où {pap} incluerait les deux autres valeurs. De plus, le fait que les démonstratifs forment un système de localisation spatiale en rapport avec différents centres déictiques conforte notre analyse de ces indexicaux comme incorporant deux traits localisateurs, l'un étant { {loc}} {N{place}}, l'autre étant { {loc}} {N{origo}} ou { {loc}} {N{sap}}, auxquels il conviendrait donc d'ajouter { {loc}} {N{tu}}.

Ce court chapitre propose trois pistes de recherches futures pour l'analyse théorique des démonstratifs. La prise en compte de l'évolution diachronique permet sans nul doute de mieux comprendre certains faits observables en synchronie ; le crédit apporté aux facteurs phonologiques ouvre de nouvelles portes pour les études à l'interface entre prosodie, sémantique et pragmatique, puisqu'il semble possible d'établir un lien entre l'accentuation mélodique et la valeur communicative des énoncés ; enfin, l'observation du système des indexicaux au sein des autres langues du monde offre un panorama instructif sur les marques de localisation des démonstratifs en rapport avec les différents centres déictiques possibles.

Chapitre X. Conclusion générale

Les recherches présentées dans cette thèse ont exploré le fonctionnement indexical des pronoms démonstratifs anglais *this, that* et *it* d'une part, et des pronoms démonstratifs du français, en particulier *ça*, d'autre part, en vue d'une comparaison entre ces deux systèmes. Les résultats obtenus ont donné lieu à une proposition de représentation syntactico-sémantique des indexicaux démonstratifs anglais et français, dans le cadre de la Grammaire Notionnelle Dépendancielle de John Anderson. Les occurrences de *this, that, it* et *ça* étant issues de conversations spontanées, notre étude apporte un questionnement sur les avantages et les limites des grands corpus oraux pour les recherches en linguistique théorique.

Les analyses structurales des démonstratifs pointent les propriétés syntagmatiques et distributionnelles de ces éléments, grâce à l'émergence du syntagme nominal dont le nom-noyau appelle un ou plusieurs constituant(s), parmi lesquels, potentiellement, les démonstratifs. Les analyses génératives modernes caractérisent les démonstratifs comme des éléments fonctionnels et têtes de syntagme ; si le déterminant démonstratif nécessite la présence matérielle d'un nom, le pronom démonstratif s'accompagne d'une catégorie nominale vide. Les analyses énonciatives laissent une grande place au rapport qu'entretiennent les démonstratifs avec l'énonciateur : les démonstratifs étant des déictiques, ils sont intimement liés à l'acte énonciatif pour pouvoir être référentiellement interprétés. Cependant, la capacité qu'ont les démonstratifs d'agir en tant qu'anaphoriques, couplée aux multiples définitions de la deixis et de l'anaphore, mène à plusieurs interprétations qui catégorisent ces éléments tantôt parmi les déictiques, tantôt parmi les anaphoriques.

Nos travaux partent d'une volonté d'homogénéisation des différentes analyses proposées pour les démonstratifs ; notre étude étant contrastive entre l'anglais oral et le français oral, et *ça* étant majoritairement employé dans les conversations spontanées de notre corpus (à 99% selon nos données), nous nous sommes concentrée sur les pronoms neutres *this*, *that, it* et *ça*. Pour ce faire, il était nécessaire de (re)définir les concepts de référence, de coréférence, de deixis et d'anaphore. Le cadre théorique dans lequel s'inscrit notre thèse définit le texte, le contexte et le discours comme trois composantes d'un ensemble dynamique : le texte se compose des signes verbaux et non verbaux dont disposent les locuteurs afin de communiquer ; le contexte est un ensemble de facteurs variables – situation d'énonciation, savoir partagé entre les co-énonciateurs, connaissances du monde, discours construit en amont… – assujetti à un processus continu de construction et de révision au fur et

à mesure que l'acte communicatif se poursuit ; le discours est une représentation mentale co-construite par les interlocuteurs, produit de la séquence hiérarchisée et contextuellement située d'actes énonciatifs, indexicaux, propositionnels et illocutoires effectués dans la poursuite d'un but communicatif quelconque et intégrés dans un contexte donné. Ainsi, les phénomènes discursifs tels que la deixis et l'anaphore sont conçus comme des stratégies indexicales employées par le locuteur afin d'effectuer des actes de référence, en fonction du statut cognitif plus ou moins saillant du référent visé dans la mémoire de l'interlocuteur. La deixis et l'anaphore ne sont alors plus définies selon l'absence ou la présence d'un antécédent textuel à l'occurrence de l'item indexical – 'indexical' étant préféré aux termes de 'déictique' et d''anaphorique', afin de ne pas amalgamer système de la langue et usage en discours – mais comme étant le résultat de l'activation d'un élément langagier en discours. Plus précisément, la deixis est décrite comme un processus de retour référentiel sur le repère énonciatif ; l'anaphore est, quant à elle, définie comme produisant un envoi référentiel unique vers la mémoire immédiate des interlocuteurs. Par conséquent, la référence effectuée par un élément indexical dépend de facteurs divers tels que ses propriétés intrinsèques, la prédication en jeu dans l'énoncé considéré, certains paramètres contextuels et l'invocation de la mémoire plus ou moins active des interlocuteurs. Dès lors, nous avons exploré la référence indexicale de *this, that* et *it* dans un premier temps, de *ça* dans un second temps, notamment par la description fine du mécanisme opéré par ces indexicaux lors d'un accès à la référence via les différents contextes – situationnel immédiat, discursif, savoirs partagés et encyclopédiques –, mais aussi par la mise en exergue des sphères cognitives et discursives propres à ces items – sphère personnelle pour *this*, interpersonnelle pour *that*, externe pour *it* et corrélative pour *ça*. Ainsi, *this*, *that*, *it* et *ça* sont non seulement des dispositifs référentiels, mais aussi des indices de co-construction du discours grâce auxquels les interlocuteurs tissent un maillage référentiel par la prise en compte de chaque participant à l'acte de parole. Les différents phénomènes observables en discours produits par *this, that, it* et *ça* ont alors été ramenés à deux paramètres minimaux intrinsèques à chacun de ces indexicaux : le fonctionnement syntactico-sémantique propre aux démonstratifs et la sphère cognitive et discursive définie par chacun d'eux. De là, nous avons proposé une représentation de la structure interne des pronoms démonstratifs anglais et français en Grammaire Notionnelle Dépendancielle. Ce cadre n'a pas été choisi dans un but de formalisation *per se*, mais parce que les postulats conceptuels offerts par la GND correspondent à nos attentes théoriques : les éléments langagiers ont une structure syntactico-sémantique qui permet de rendre compte de leur comportement distributionnel, référentiel et discursif. La représentation des différents constituants sous forme de graphes de

dépendance a alors une valeur descriptive et explicative, mais ne permet pas de tout révéler : les formalisations sont effectuées sur des énoncés-types, ne peuvent se passer de pré-analyses en vue d'extraire des tendances générales, et ne dispensent en aucun cas de l'examen précis des phénomènes idiosyncrasiques de la langue. Toutefois, notre prise en compte de *this*, *that*, *it* et *ça* dans le cadre de la GND aboutit à une représentation notionnelle de *ça* qui couvre celles de *that* et de *it*, et qui corrobore nos analyses théoriques. De manière prototypique, *ça* et *it* sont comparables par la totale neutralité dont ils font preuve, neutralité qui leur permet d'absorber tous les rôles sémantiques et d'être des manifestations syntactico-sémantiques de l'hypostatisation de la langue et de la sous-spécification des éléments du lexique. *Ça* et *that* sont comparables par leur fonctionnement référentiel, caractéristique des démonstratifs, que nous avons décrit comme étant un mouvement rétroactif sur le repère déictique dont l'origine est l'ensemble des participants à l'acte communicatif. Dès lors, l'association des caractéristiques syntactico-sémantiques de *it* et de celles de *that* permet à *ça* de couvrir le continuum des procédures indexicales. *This*, quant à lui, est un indexical marqué par la localisation par rapport à l'*origo* qu'il effectue.

En dernier lieu, cette thèse offre des pistes de recherches possibles afin d'approfondir la description des indexicaux, notamment par la prise en compte de l'évolution diachronique, de l'accentuation prosodique et de l'observation des autres langues du monde. Ces perspectives restent à explorer, mais les travaux effectués en amont de la rédaction de cette thèse ont d'ores et déjà contribué à la naissance de trois projets : l'élaboration d'un analyseur du français parlé, SAXO, en collaboration avec Julien Eychenne (Hankuk University of Foreign Studies, Corée du Sud), la formation du projet PAC-SYNTAXE, coordonné par Nathalie Rossi-Gensane (Université Lyon II & ICAR) et nous-même, et la création du projet PFC-SYNTAXE, coordonné par Julie Auger (Indiana University, Bloomington), Béatrice-Akissi Boutin (CLLE-ERSS), Nathalie Rossi-Gensane (Université Lyon II & ICAR) et Paul Rowlett (University of Salford, Royaume-Uni).

Bibliographie

Abney, S. (1987). *The English Noun Phrase in Its Sentential Aspect*. Thèse de doctorat non publiée. MIT.

Académie Française (1932). *Grammaire de l'Académie Française*. Huitième édition. Paris : Firmin-Didot.

Achard, M. (2001). "French *Ça* and the Dynamics of Reference". *The LACUS Forum, Speaking and Comprehending* 27 : 49-62.

Adam, J.-M. (2005). *Introduction à l'analyse textuelle des discours*. Paris : Armand Colin.

Adamczewski, H. (1976). BE+ING *dans la grammaire de l'anglais contemporain*. Thèse d'Etat. Lille : Atelier de reproduction de l'Université de Lille et Paris : Librairie Champion.

Adamczewski, H. (1982[1993]). *Grammaire linguistique de l'anglais*. Paris : Armand Colin.

Adamczewski, H. (1995). *Caroline grammairienne en herbe, ou comment les enfants inventent leur langue maternelle.* Paris : Presses de la Sorbonne Nouvelle.

Alexiadou, A., L. Haegeman & M. Stavrou (2007). *Noun Phrase in the Generative Perspective*. Berlin, New York : Mouton de Gruyter.

Anderson, A.H., M. Bader, E.G. Bard, E. Boyle, G. Doherty, S. Garrod, S. Isard, J. Kowto, J. McAllister, J. Miller, C. Sotillo, H. Thompson & R. Weinert (1991). "The HCRC Map Task Corpus". *Language and Speech* 34 (4) : 351-366.

Anderson, J.M. (1971). *The Grammar of Case: Towards a Localist Theory*. Cambridge : Cambridge University Press.

Anderson, J.M. (1991). "Notional Grammar and the Redundancy of Syntax". *Studies in Language* 15 (2) : 301-333.

Anderson, J.M. (1997). *A Notional Theory of Syntactic Categories*. Cambridge : Cambridge University Press.

Anderson, J.M. (2006). "Structural Analogy and Universal Grammar." *Lingua* 116 (5) : 601-633.

Anderson, J.M. (2007). *The Grammar of Names*. Oxford : Oxford University Press.

Anderson, J.M. (2011). *The Substance of Language*. 3 vols. Oxford : Oxford University Press.

Anderson, S.R. & E. Keenan (1985). "Deixis". In T. Shopen (ed.). *Language Typology and Syntactic Description. Vol. III, Grammatical Categories and the Lexicon*. Cambridge : Cambridge University Press.

Apothéloz, D. (1995). "Nominalisations, référents clandestins et anaphores atypiques". *TRANEL* 23 : 143-173.

Apothéloz, D. & M.-J. Reichler-Béguelin (1999). "Interpretations and Functions of Demonstrative NPs in Indirect Anaphora". *Journal of Pragmatics* 31 (3) : 363-397.

Ariel, M. (1988). "Referring and Accessibility". *Journal of Linguistics* 24 (1) : 65-87.

Atlani, F., A. Grésillon, J.-L. Lebrave, L. Danon-Boileau & J. Simonin (1984). *La langue au ras du texte*. Lille : Presses Universitaires du Septentrion.

Barber, C. (2000). *The English Language: a Historical Introduction*. Cambridge : Cambridge University Press.

Beckman, M.E. & J.B. Pierrehumbert (1986). "Intonational Strucure in Japanese and English". *Phonology Yearbook* 3 : 255-309.

Benveniste, E. (1956[1966]). *La nature des pronoms*. Paris : Gallimard.

Bernstein, J. (1997). "Demonstratives and Reinforcers in Romance and Germanic Languages". *Lingua* 102 : 87-113.

Berrendonner, A. (1990). “Pour une macro-syntaxe”. *Travaux de linguistique* 21: 25-36.
Berrendonner, A. (1993). “Périodes”. In H. Parret (ed.). *Temps et discours*. Louvain : Presses universitaires. Pp. 47-61.
Berrendonner, A. (2004). “Grammaire de l’écrit *vs* grammaire de l’oral : le jeu des composantes micro- et macro-syntaxiques”. In A. Rabatel (ed.). *Interaction orales en contexte didactique : mieux (se) comprendre pour mieux (se) parler et pour mieux (s’)apprendre*. Lyon : Presses Universitaires de Lyon. Pp : 249-264.
Biber, D. (1988). *Variation across Speech and Writing*. Cambridge : Cambridge University Press.
Biber, D. (1995). *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge : Cambridge University Press.
Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan (1999). *Longman Grammar of Spoken and Written English*. Londres : Longman.
Biber, D. (2002). *Longman Student Grammar of Spoken and Written English*. Londres : Longman.
Bilger, M. (1988). “Les réalisations en ‘et tout (ça)’ à l’oral”. *Recherches sur le français parlé* 9 : 9-109.
Blanche-Benveniste, C. (1997). “La notion de variation syntaxique dans la langue parlée”. *Langue française* 115 : 19-29.
Blanche-Benveniste, C., C. Rouget, M. Bilger, K. van den Eynde & P. Mertens (1990). *Le français parlé. Etudes grammaticales*. Paris : Editions du Centre National de la Recherche Scientifique.
Bloomfield, L. (1933). *Language*. New York : H. Holt.
Boersma, P. & D. Weenink (2009). *Praat: doing phonetics by computer* (version 5.1.05) [Computer program]. http://www.praat.org.
Böhm, R. (1982). *Topics in Localist Case Grammar*. Thèse de doctorat. Université d’Ulster, Coleraine.
Botley, S. & A. McEnery (eds.)(2000). *Corpus-Based and Computational Approaches to Discourse Anaphora*. Amsterdam : John Benjamins.
Bouscaren, J., J. Chuquet & L. Danon-Boileau (1987). *Grammaire et textes anglais : guide pour l’analyse linguistique*. Paris : Ophrys.
Boutin, B.A. (2009). “Procédés d’identifications et de catégorisation en français (Afrique de l’Ouest)”. In S. Osu, G. Col, N. Garric & F. Toupin (eds.). *Construction d’identité et processus d’identification*. Bern : Peter Lang. Pp. 33-50.
Branco, A., T. McEnery & R. Mitkov (2005). *Anaphora Processing: Linguistic, Cognitive and Computational Modelling*. Amsterdam : John Benjamins.
Britain, D. (ed.)(2007). *Language in the British Isles*. Cambridge : Cambridge University Press.
Brown, G. & G. Yule (1983). *Discourse Analysis*. Cambridge : Cambridge University Press.
Brown-Schmidt, S., D.K. Byron & M.K. Tanenhaus (2005). “Beyond Salience: Interpretation of Personal and Demonstrative Pronouns”. *Journal of Memory and Language* 53 (2) : 292-313.
Brunot, F. (1934). *Histoire de la langue française des origines à 1900. Tome II. Le seizième siècle*. Troisième édition. Paris : Armand Colin.
Bühler, K. (1934[2009]). *Sprachtheorie*. Iéna : Fischer. [Traduction française 2009, *Théorie du langage*, Marseille : Agone.]
Butler, C. (sous presse). “An Ontological Approach to the Representational Lexicon in Functional Discourse Grammar”. *Language Sciences* 34 (4).
Butler, C. (2008). “Cognitive Adequacy in Structural-functional Theories of Language.” *Language Sciences* 30 (1) : 1-30.

Bybee, J. (2003). “Cognitive Processes in Grammaticalization”. In M. Tomasello (ed.). *The New Psychology of Language, Vol. II*. Mahwah, New Jersey & Londres : Erlbaum. Pp. 145-167.
Cabredo-Hofherr, P. (2004). “Les clitiques sujets du français et le paramètre du sujet nul”. *Langue française* 141 (1) : 99-109.
Cadiot, P. (1988). “De quoi *ça* parle ? A propos de la référence de *ça*, pronom-sujet”. *Français moderne* 56 (3-4) : 174-192.
Capone, A. (2011). “The Attributive/Referential Distinction, Pragmatics, Modularity of Mind and Modularization”. *Australian Journal of Linguistics* 31 (2) : 153-186.
Cappeau, P. (2001). “Faits de syntaxe et genres à l’oral”. *Le français dans le monde*. Numéro spécial : 69-77.
Carlier, A. (1996). “‘Les gosses, ça se lève tôt le matin’ : l’interprétation générique du syntagme nominal disloqué au moyen de *ce* ou *ça*”. *Journal of French Language Studies* 6 (2) : 133-162.
Carlier, A. & W. de Mulder (2006). “Du démonstratif à l’article défini : le cas de *ce* en français moderne”. *Langue française* 152 (4) : 96-113.
Carnap, R. (1947). *Meaning and Necessity*. Chicago : University of Chicago Press.
Carr, P., J. Durand & M. Pukli (2004). “The PAC Project: Principles and Methods”. *Tribune des langues vivantes* 36 : 24-35.
Chastain, C. (1975). “Reference and Context”. In *Language, Mind and Knowledge*. Minneapolis : University of Minnesota Press. Pp. 194-269.
Chen, R. (1990). “English Demonstratives: A Case of Semantic Expansion”. *Language Sciences* 12 (2-3) : 139-153.
Chevalier, J.-C., C. Blanche-Benveniste, M. Arrivé & J. Peytard (1964). *Grammaire Larousse du français contemporain*. Paris : Larousse.
Chomsky, N. (1957). *Syntactic Structures*. La Haye : Mouton.
Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Mass : MIT Press.
Chomsky, N. (1972). *Studies on Semantics in Generative Grammar*. La Haye : Mouton.
Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht : Foris.
Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass : MIT Press.
Churchward, C.M. (1953). *Tongan Grammar*. Oxford : Oxford University Press.
Clark, H. & E. Schaefer (1989). “Contributing to Discourse”. *Cognitive Science* 13 (2) : 259-294.
Clark, M. (2000). “Deixis and Anaphora and Prelinguistic Universals”. *Oceanic Linguistics Special Publications* 29 : 46-61.
Cole, P. (1974). “Indefiniteness and Anaphoricity”. *Language* 50 (4) : 665-674.
Colman, F. (à paraître). *The Grammar of Names in Anglo-Saxon England: the Linguistics and Culture of the Old English Onomasticon*. Oxford : Oxford University Press.
Condamines, A. (2003). *Sémantique et corpus spécialisés : constitution de bases de connaissances terminologiques*. Mémoire d’Habilitation à Diriger les Recherches. Université de Toulouse II-Le Mirail.
Connolly, J. (1997). *Discourse and Pragmatics in Functional Grammar*. Berlin, New York : Mouton de Gruyter.
Connolly, J. (2007). “Context in Functional Discourse Grammar”. *Alfa : revista de linguística* 51 (2) : 11-33.
Corblin, F. (1985). “Remarques sur la notion d’anaphore”. *Revue québécoise de linguistique* 15 (1) : 173-195.
Corblin, F. (1987). *Indéfini, défini et démonstratif : constructions linguistiques de la référence*. Genève : Droz.
Corblin, F. (1995). *Les formes de reprise dans le discours. Anaphores et chaînes de référence*.

Rennes : Presses Universitaires de Rennes.
Corblin, F. (2005). "Les chaînes de la conversation et les autres". In J.-M. Gouvard (ed.). *De la langue au style*. Lyon : Presses Universitaires de Lyon. Pp. 233-254.
Cornish, F. (1996). "'Antecedentless' Anaphors: Deixis, Anaphora, or What? Some Evidence from English and French". *Journal of Linguistics* 32 (1) : 19-41.
Cornish, F. (1998). "Les chaînes topicales : leur rôle dans la gestion et la structuration du discours". *Cahiers de grammaire* 23 : 19-40.
Cornish, F. (1999). *Anaphora, Discourse, and Understanding: Evidence from English and French*. Oxford : Clarendon Press.
Cornish, F. (2007). "English Demonstratives: Discourse Deixis and Anaphora. A Discourse-Pragmatic Account". In R.A. Nilson, N.A.A. Amfo & K. Borthen (eds.). *Interpreting Utterances: Pragmatics and its Interfaces*. Essays in Honour of Thorstein Fretheim. Oslo : Novus Press. Pp. 147-166.
Cornish, F. (2010). "Anaphora: Text-based or Discourse-dependent? Functionalist *vs* Formalist Accounts". *Functions of Language* 17 (2) : 207-241.
Cornish, F. & L. Buscail (2011). "When Text Creates its Own Context: Demonstratives, Accenting and Contrast in English". *Meaning and Contexte Conference*. University of the West of England. Bristol, 13-15 Avril 2011.
Cornish, F., A. Garnham, H.W. Cowles, M. Fossard & V. André (2005). "Indirect Anaphora in English and French: A Cross-linguistic Study of Pronoun Resolution". *Journal of Memory and Language* 52 (3) : 363-376.
Cotte, P. (1993). "De l'étymologie à l'énonciation ; deixis, anaphore abstraite, syntaxe génétique dans quelques mots en th- de l'anglais contemporain". *Travaux de linguistique et de philologie* 4 : 43-89.
Crystal, D. (2003). *A Dictionary of Linguistics and Phonetics*. Cinquième édition. Londres : Blackwell.
Culioli, A. (1999). *Pour une linguistique de l'énonciation: Domaine notionnel*. Paris : Ophrys.
Culioli, A., C. Fuchs & M. Pêcheux (1970). *Considérations théoriques à propos du traitement formel du langage, tentative d'application au problème des déterminants*. Paris : Dunod.
Curat, H. & L. Meney (1983). *Gustave Guillaume et la psycho-systématique du langage : bibliographie annotée*. Sainte-Foy : Presses de l'Université Laval.
Creissels, D. (2006). *Syntaxe générale. Une introduction typologique 1 : catégories et constructions*. Paris : Hermès-Lavoisier.
Danon-Boileau, L. (1983). "*This, that, which, what* et la construction de références, quelques hypothèses". *Méthodes en linguistique anglaise* 39 : 35-56.
Danon-Boileau, L. (1984). "*That* is the question". In F. Atlani, A. Grésillon, J.-L. Lebrave, L. Danon-Boileau & J. Simonin (eds.). *La langue au ras du texte*. Lille : Presses Universitaires du Septentrion. Pp. 31-56.
Danon-Boileau, L. (1992). "Ce que ça veut dire : les enseignements de l'observation clinique". In M.-A. Morel & L. Danon-Boileau (eds.). *La deixis, colloque en Sorbonne 8-9 juin 1990*. Paris : Presses Universitaires de France. Pp. 414-425.
Dees, A. (1971). *Etude sur l'évolution des démonstratifs en ancien et en moyen français*. Groningen : Wolters-Noordhoff.
Deulofeu, J. (2000). "Types d'énoncés et genres : le cas des commentaires sportifs." In M. Bilger (ed.). *Corpus – Méthodologie et applications linguistiques*. Paris : Honoré Champion. Pp. 271-295.
Diaz-Negrillo, A., N. Ballier & P. Thompson (eds.)(à paraître). *Automatic Treatment and Analysis of Learner Corpora*. Amsterdam : John Benjamins.

Diessel, H. (1999a). *Demonstratives: Form, Function, and Grammaticalization*. Amsterdam : John Benjamins.
Diessel, H. (1999b). "The Morphosyntax of Demonstratives in Synchrony and Diachrony." *Linguistic Typology* 3 (1) : 1-50.
Diessel, H. (2006). "Demonstratives, Joint Attention, and the Emergence of Grammar." *Cognitive Linguistics* 17 (4) : 463-489.
Diessel, H. (2011). "Distance Contrasts in Demonstratives". In M.S. Dryer & M. Haspelmath (eds.). *The World Atlas of Language Structures Online*. http://wals.info/chapter/41.
Dik, S. (1997). *Theory of Functional Grammar: The Structure of the Clause*. K. Hengeveld (ed.). Deuxième édition. Berlin, New York : Mouton de Gruyter.
Dubois, J. (1965). *Grammaire structurale du français*. Paris : Larousse.
Dubois, J. & F. Dubois-Charlier. (1970). *Eléments de linguistique française : syntaxe*. Paris : Larousse.
Ducrot, O. & T. Todorov (1979). *Dictionnaire encyclopédique des sciences du langage*. Paris : Seuil.
Durand, J. (1995). "Universalism in Phonology: Atoms, Structures, Derivations." In J. Durand & F. Katamba (eds.). *Frontiers of Phonology. Atoms, Structures, Derivations*. Londres : Longman. Pp. 267-288.
Durand, J. (2000). "Oral, écrit et faculté de langage". In M.-N. Guillot & M.-M. Kenning (eds.)(2000). *Changing Landscapes in Language and Language Pedagogy*. Londres : AFLS/CILT. Pp. 40-72.
Durand, J. (2011). "Quelques remarques sur les propositions de l'anglais, l'hypothèse localiste et le principe d'analogie structurale." In F. Neveu, P. Blumenthal & N. Le Querler (eds.). *Au commencement était le verbe. Syntaxe, sémantique et cognition*. Bern : Peter Lang. Pp. 93-122.
Durand, J., B. Laks & C. Lyche (2002). "La phonologie du français contemporain : usages, varieties et structure". In C. Pusch & W. Raible (eds.). *Romanistische Korpuslinguistik - Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen : Gunter Narr Verlag. Pp. 93-106.
Durand, J., B. Laks & C. Lyche (2009). "Le projet PFC : une source de données primaires structurées". In J. Durand, B. Laks & C. Lyche (eds.)(2009). *Phonologie, variation et accents du français*. Paris : Hermès. Pp. 19-61.
Ehlich, K. (1982). "Anaphora and Deixis: Same, Similar or Different?" In R.J. Jarvella & W. Klein (eds.). *Speech, Place and Action. Studies in Deixis and Related Topics*. Chichester : Wiley. Pp. 315-338.
Engel, D. & N. Rossi-Gensane (2004). "Sémantique et fréquence : étude de *c'est/il est* dans un corpus journalistique". *Revue romane* 39 (2) : 177-202.
Epstein, R. (2011). "The Distal Demonstrative as Discourse Marker in Beowulf". *English Language and Linguistics* 15 (1) : 113-135.
Ewert, A. (1943). *The French Language*. Londres : Faber & Faber.
Fauconnier, G. (1984). *Espaces mentaux : aspects de la construction du sens dans les langues naturelles.* Paris : Editions de minuit.
Fauconnier, G. & M. Turner (2002). *The Way We Think. Conceptual Blending and the Mind's Hidden Complexities*. New York : Basic Books.
Filik, R., A.J. Sanford, P. Sturt & M. Poesio (2005). "Underspecification in anaphora reference to structured entities". *13th European Conference on Eye Movements*. Bern, Suisse.
Firth, J.R. (1930). *Speech*. Londres : Ernest Benn.
Fodor, J. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Mass : MIT Press.

Fossard, M., A. Garnham & H.W. Cowles (2011). "Between Anaphora and Deixis. The Resolution of the Demonstrative Noun Phrase *That N*". *Language and Cognitive Processes Online* : 1-20.
Fox, B.A. (1996). *Studies in Anaphora*. Amsterdam : John Benjamins.
Fraser, T. & A. Joly (1979). "Le système de la deixis. Esquisse d'une théorie d'expression en anglais". *Modèles Linguistiques* 1 (2) : 97-157.
Frege, G. (1892[1971]). "Uber sinn und bedeutung". *Zeitschrift für Philosophie und philosophische Kritik* 100 : 22-50. [Traduction française, 1971. "Sens et denotation". In *Ecrits logiques et philosophiques*. Paris : Seuil. Pp : 102-126.]
Gadet, F. (2003). *La variation sociale en français*. Paris : Ophrys.
Gaffiot, F. (1934). *Dictionnaire illustré latin-français*. Paris : Hachette.
Gaillat, T. (2013). "Annotation automatique d'un corpus d'apprenants d'anglais avec un jeu d'étiquettes modifié du Penn Treebank". In E. Morin & Y. Estève (eds.). *Actes de la conférence TALN-RECITAL 2013*. Pp. 414-425.
Garrod, S., D. Freudenthal & E. Boyle (1994). "The Role of Different Types of Anaphor in the Online Resolution of Sentences in a Discourse". *Journal of Memory and Language* 33 (1) : 39-68.
Garside, R., S. Fligelstone & S. Botley (1997). "Discourse Annotation: Anaphoric Relations in Corpora." In R. Garside, G. Leech, & A. McEnery (eds.). *Corpus Annotation: Linguistic Information from Computer Text Corpora*. Londres : Longman. Pp. 66-84.
Gerner, M. (2009). "Deictic Features of Demonstratives: A Typological Survey with Special Reference to the Miao Group". *The Canadian Journal of Linguistics* 54 (1) : 43-90.
Giancarli, P. (2004). "Le fléchage (spécifique et générique) : opération seconde ou opération double ?" *Cycnos* 18 (2) (en ligne). http://revel.unice.fr/cycnos/index.html?id=44.
Green, K. (1995). *New Essays in Deixis: Discourse, Narrative, Literature*. Amsterdam : Rodopi.
Greenberg, J.H. (ed.)(1966). *Universals of Language*. Deuxième édition. Cambridge, Mass : MIT Press.
Greimas, A.J. & J. Courtès (1979). *Sémiotique : dictionnaire raisonné de la théorie du langage*. Paris : Hachette.
Grice, P. (1957). "Meaning." *The Philosophical Review* 66 : 377-388.
Grishman, R. & B. Sundheim (1996). "Message Understanding Conference – 6 : A Brief History". *Proceedings of the 16th Conference on Computational Linguistics – Vol. 1*. *COLING '96.* Stroudsburg, PA : Association for Computational Linguistics. Pp. 466-471.
Grobet, A. (2002). *L'identification des topiques dans les dialogues*. Bruxelles : De Boeck/Duculot.
Gross, M. (1977[1986]). *Grammaire transformationnelle du français : syntaxe du nom*. Paris : Cantilène.
Groussier, M.L., G. Groussier & P. Chantefort (1975). *Grammaire anglaise : thèmes construits*. Nouvelle édition refondue et augmentée. Paris : Hachette.
Guenette, L. (1997). "Les pronoms neutres *il, ce, ça* : une comparaison de leurs emplois et de leur signifié". In P. De Carvalho & O. Soutet (dirs)(1994). *Psychomécanique du langage. Problèmes et perspectives. Actes du septième colloque international de Psychomécanique du langage*. Paris : Champion. Pp. 111-124.
Guérin, E. (2007). " Moi je comprends pas et ça ça m'intrigue". *Linx* 57 : 27-36.
Guillaume, G. (1919). *Problème de l'article et sa solution dans la langue française*. Paris : Hachette.
Guillaume, G. (1929). *Temps et verbe : théorie des aspects, des modes et des temps*. Paris : Champion.

Guillaume, G. (1945[1970]). “L’architectonique du temps dans les langues classiques.” In G. Guillaume. *Temps et verbe*. Paris : Champion. Pp. 9-16.
Guillot, C. (2006). “Démonstratif et deixis discursive : analyse comparée d’un corpus écrit en français médiéval et d’un corpus oral de français contemporain”. *Langue française* 152 (4) : 56-69.
Guillot, C. (2010). “Les démonstratifs de l’ancien français : un système encore personnel ?” In F. Neveu, V. Muni Toke, J. Durand, T. Klingler, L. Mondada & S. Prévost (dirs) (2010). *Congrès mondial de linguistique française CMLF-2010. Recueil des résumés et CD-ROM des actes*. Paris : Institut de Linguistique Française et EDP Sciences.
Gundel, J., N. Hedberg & R. Zacharski (1988). “On the Generation and Interpretation of Demonstrative Expressions.” In *Proceedings of the 12th Conference on Computational Linguistics – Vol. 1. COLING ’88*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 216-221.
Gundel, J., N. Hedberg & R. Zacharski- (1993). “Cognitive Status and the Form of Referring Expressions in Discourse”. *Language* 69 (2) : 274-307.
Gundel, J., N. Hedberg & R. Zacharski (2012). “Underspecification of Cognitive Status in Reference Production: Some Empirical Predictions”. *Topics in Cognitive Science* 4 (2) : 249-268.
Haase, A. (1914). *Syntaxe française du 17ème siècle*. Nouvelle édition, traduite et remaniée par M. Obert. Paris : Delagrave.
Halliday, M.A.K. (1978). *Language as Social Semiotics : the Social Interpretation of Language and Meaning*. Maryland : University Park Press.
Halliday, M.A.K. (1985). *Introduction to Functional Grammar*. Londres : Edward Arnold.
Halliday, M.A.K. (1989). *Spoken and Written Language*. Victoria, Australie : Deakin University.
Halliday, M.A.K. & R. Hasan (1976). *Cohesion in English*. Londres : Longman.
Hanks, W. (2009). “Fieldwork on Deixis”. *Journal of Pragmatics* 41 (1) : 10-24.
Haviland, S. & H. Clark (1974). “What’s New? Acquiring New Information as a Process in Comprehension.” *Journal of Verbal Learning and Verbal Behavior* 13 (5) : 512-521.
Hedberg, N., R. Zacharski & J. Gundel (2007). *The Grammar-Pragmatics Interface: Essays in Honor of Jeanette K. Gundel*. Amsterdam : John Benjamins.
Hengeveld, K. & J.L. Mackenzie (2008). *Functional Discourse Grammar: A Typologically-Based Theory of Language Structure*. Oxford : Oxford University Press.
Higgins, F. (1979). *The Pseudo-cleft Construction in English*. New York : Garland.
Himmelmann, N. (1996). “Demonstratives in narrative discourse: a taxonomy of universal uses”. In B. Fox (ed.). *Studies in Anaphora*. Amsterdam/Philadelphie : John Benjamins. Pp. 205-254.
Hirschberg, J., C. Nakatani & B. Grosz (1995). “Discourse Structure in Spoken Language: Studies on Speech Corpora”. Communication présentée lors du *AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*. Palo Alto, California, 27-29 mars 1995.
Horn, G.M. (1975). “On the Non-Sentential Nature of the POSS-*ing* Construction”. *Linguistic Analysis* 1 : 333-388.
Huddleston, R. & G. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge : Cambridge University Press.
Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Mass : MIT.
Jackendoff, R. (1977). *X-Bar Syntax*. Cambridge, Mass : MIT Press.
Jespersen, O. (1933). *Essentials of English Grammar*. Londres : Allen & Unwin.
Jespersen, O. (1971). *La philosophie de la grammaire*. Paris : Les Editions de Minuit.

Joly, A. & T. Fraser. (1979). “Le système de la deixis : endophore et cohésion discursive en anglais”. In A. Joly (1987). *Essais de systématique énonciative*. Lille : Presses Universitaires de Lille.
Joly, A. & D. O’Kelly (1990). *Grammaire systématique de l’anglais*. Paris : Nathan.
Kameyama, M. (1999). “Stressed and Unstressed Pronouns: Complementary Preferences”. In P. Bosch & R. Van der Sandt (eds.). *Focus: Linguistic, Cognitive and Computational Perspectives*. Cambridge : Cambridge University Press. Pp. 306-321.
Kaplan, D. (1989). “Demonstratives: an Essay on the Semantics, Logic, Metaphysics and Epistemology of Demonstratives and Other Indexicals”. In J. Almong, J. Perry & H. Wettstein (dir.). *Themes from Kaplan*. New York : Oxford University Press.
Kleiber, G. (1985). “Sur la spécialisation grammaticale des démonstratifs en français ancien”. In *Mélanges H. Naïs : De la plume d’oie à l’ordinateur*. Numéro spécial de la revue *Verbum*. Nancy : Presses Universitaires de Nancy. Pp. 99-113.
Kleiber, G. (1983). “Les démonstratifs (dé)montrent-ils ? Sur le sens référentiels des adjectifs et pronoms démonstratifs”. *Le français moderne* 51 : 99-117.
Kleiber, G. (1987). “Mais à quoi sert donc le mot *chose* ?” *Langue française* 73 (1) : 109-128.
Kleiber, G. (1994). *Anaphores et pronoms*. Louvain-la-Neuve : Duculot De Boeck.
Kleiber, G. (2001). “Remarques sur la dénomination”. *Cahiers de praxématique* 36 : 21-41.
Kleiber, G. (2001). “Sur le sens du sens : objectivisme et constructivisme”. In D. Keller, J.-P. Durafour, J.-F. Bonnot & R. Sock (eds.). *Percevoir : monde et langage*. Bruxelles : Mardaga. Pp. 335-370.
Kleiber, G. (2006). “Démonstratifs : emplois à la mode et mode(s) d’emploi”. *Langue française* 152 (2) : 9-23.
Kleiber, G. (2007). “Des démonstratifs mémoriels aux démonstratifs de point de vue”. In L. Begioni & C. Muller (eds.). *Problèmes de sémantique et de syntaxe. Hommage à André Rousseau.* Villeneuve d’Ascq : CEGES, Lille 3.
Kornfilt, J. (1997). *Turkish*. Londres, New York : Routledge.
Kortmann, B. & C. Upton (eds.)(2008). *Varieties of English: The British Isles*. Berlin : Mouton de Gruyter.
Kripke, S. (1972). *Naming and Necessity*. Cambridge, Mass : Harvard University Press.
Kuno, S. (1973). *The Structure of the Japanese Language*. Cambridge, Mass. : MIT Press.
Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC : Center for Applied Linguistics.
Laczkó, K. (2010). “Demonstrative Pronouns in Spatial Deixis, Discourse Deixis, and Anaphora”. *Acta Linguistica Hungarica* 57 (1) : 99-118.
Lafont, R. (1967). *La phrase occitane : essai d’analyse systématique*. Paris : Presses Universitaires de France.
Lakoff, G. & M. Johnson (1980). *Metaphors we Live by*. Chicago : University of Chicago Press.
Lambrecht, K. (1994). *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge : Cambridge University Press.
Langacker, R. (1987). *Foundations of Cognitive Grammar*. Stanford, Californie : Stanford University Press.
Lapaire, J.-R. & W. Rotgé (1992). *Linguistique et grammaire de l’anglais*. Toulouse : Presses Universitaires du Mirail.
Le Goffic, P. (1993). *Grammaire de la phrase française*. Paris : Hachette.
Lehmann, C. (2002). “Thoughts on Grammaticalization”. Deuxième édition. *Language* 73 (9) : 670.
Likhacheva-Philippe, L. (2008). *L’articulation pragmatique du discours : le cas des*

constructions détachées à gauche. Thèse de doctorat non publiée. Université Toulouse II-Le Mirail.
Lyons, C. (1999). *Definiteness*. Cambridge : Cambridge University Press.
Lyons, J. (1968). *Introduction to Theoretical Linguistics*. Cambridge : Cambridge University Press.
Lyons, J. (1977). *Semantics, Vol. II*. Cambridge : Cambridge University Press.
Maillard, M. (1987). "'Un zizi, ça sert à faire pipi debout'. Les références génériques de *ça* en grammaire de phrase". In G. Kleiber (dir.). *Rencontre(s) avec la généricité*. Paris : Klincksieck. Pp. 158-206.
Marchello-Nizia, C. (1995). *L'évolution du français*. Paris : Armand Colin.
Marchello-Nizia, C. (2004). "La sémantique des démonstratifs en ancien français : une neutralisation en progrès ?" *Langue française* 141 : 69-84.
Marchello-Nizia, C. (2005). "Deixis and Subjectivity : the Semantics of Demonstratives in Old French (9^{th}-12^{th} century)". *Journal of Pragmatics* 37 (1) : 43-68.
Martin, P. (2008). "L'intonation du français : le vilain petit canard parmi les langues romanes ?" *Language Design* 2 : 1-13.
Mathews, C.E. (1907). *Cist and Cil: A Syntactical Study*. Baltimore, Maryland : JH Furst.
McCarthy, M. (2001). *Issues in Applied Linguistics*. Cambridge : Cambridge University Press.
McKelvie, D., A. Isard, A. Mengel, M.B. Møller, M. Grosse & M. Klein (2001). "The MATE Workbench – An Annotation Tool for XML coded speech corpora". *Speech Communication* 33 (1-2) : 97-112.
Mertens, P. (2004). "Quelques allers-retours entre la prosodie et son traitement automatique". *Le français moderne* 72 (1) : 39-57.
Mill, J.S. (1896). *Système de logique déductive et inductive. Exposé des principes de la preuve et des méthodes de recherche scientifique*. Paris : Ladrange.
Milner, J.-C. (1982). *Ordres et raisons de langue*. Paris : Seuil.
Mithun, M. (1987). "The grammatical nature and discourse power of demonstratives". *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 13 : 184–194.
Mitkov, R. (2001). "Outstanding Issues in Anaphora Resolution". In *Proceedings of the Second International Conference on Computational Linguistics and Intelligent Text Processing, CICLing'01*. Londres : Springer-Verlag. Pp. 110-125.
Moignet, G. (1974). *Etudes de psycho-systématique française*. Paris : Klincksieck.
Moignet, G. (1981). *Systématique de la langue française*. Paris : Klincksieck.
Morel, M.-A. (2007). "Le postrhème dans le dialogue oral en français". *L'Information grammaticale* 113 : 40-46.
Morrill, G. (2010). *Categorial Grammar: Logical Syntax, Semantics and Processing*. Oxford : Oxford University Press.
Mulder De, W. (1998). "Du sens des démonstratifs à la construction d'univers". *Langue française* 120 (1): 21-32.
Mulder De, W. & L. Tasmowski (1996). *Coherence and Anaphora*. Amsterdam : John Benjamins.
Müller, C. & M. Strube (2006). "Multi-Level Annotation of Linguistic Data With MMAX2". In S. Braun, K. Kohn & J. Mukherjee (eds.). *Corpus Technology and Language Pedagogy. New Ressources, New Tools, New Methods*. Francfort : Peter Lang.
Nedoluzhko, A., J. Mírovský & P. Pajas (2009). "The Coding Scheme for Annotating Extended Nominal Coreference and Bridging Anaphora in the Prague Dependency Treebank". In *Proceedings of the Third Linguistic Annotation Workshop, ACL-IJCNLP'09*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 108-111.

Nunberg, G. (1993). “Indexicality and Deixis”. *Linguistics and Philosophy* 16 (1) : 1-43.
Passonneau, R. (1989). “Getting at Discourse Referents”. In J. Hirshberg (eds.). *Proceedings of the 27th Annual Meeting of Association for Computational Linguistics, ACL’89*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 51-59.
Passonneau, R. (1991). “Some Facts About Centers, Indexicals, and Demonstratives”. In D.E. Appelt (ed.). *Proceedings of the 29th Annual Meeting of Association for Computational Linguistics, ACL’91*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 63-70.
Philps, D. (2002). “Le concept de marqueur sub-lexical et la notion d’invariant sémantique”. *Travaux de linguistique* 45 (2) : 103-123.
Piaget, J. (1937). *La construction du réel chez l’enfant*. Paris : Delachaux & Niestlé.
Pierrehumbert, J.B. (1980). *The Phonology and Phonetics of English Intonation*. Thèse de doctorat. MIT.
Piwek, P. & A. Cremers (1996). “Dutch and English Demonstratives: A Comparison”. *Language Sciences* 18 (3) : 835-51.
Poesio, M. (2004). “Discourse Annotation and Semantic Annotation in the GNOME Corpus”. In B. Webber & D.K. Byron (eds.). *Proceedings of the 2004 ACL Workshop on Discourse Annotation, DiscAnnotation’04*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 72-79.
Poesio, M. & R. Artstein (2005). “The Reliability of Anaphoric Annotation, Reconsidered: Taking Ambiguity into Account”. In *Proceedings of ACL Workshop on Frontiers in Corpus Annotation*. Madison, WI : Omnipress. Pp. 76-83.
Poesio, M. & R. Artstein (2008). “Anaphoric Annotation in the ARRAU Corpus”. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation, LREC '08*. Marrakech : European Language Resources Association. Pp. 1170-1174.
Poesio, M., U. Reyle & R. Stevenson (2007). “Justified Sloppiness in Anaphoric Reference”. In H. Bunt, R. Muskens, L. Matthewson, Y. Sharvit & T. E. Zimmermann (eds.). *Computing Meaning, Vol. III*. Berlin : Springer. Pp. 11-31.
Poesio, M., P. Sturt, R. Artstein & R. Filik (2006a). “Underspecification and Anaphora: Theoretical Issues and Preliminary Evidence”. *Discourse Processes: A Multidisciplinary Journal* 42 (2) : 157-175.
Poesio, M., P. Sturt, R. Artstein & R. Filik (2006b). “Underspecification and Anaphora: Theoretical Issues and Preliminary Evidence”. *Discourse Processes* 42 (2) : 157-175.
Pollock, J.-Y. (1983). “Sur quelques propriétés des phrases copulatives en français”. *Langue française* 58 (1) : 89-125.
Price, G. (1971). *The French Language: Present and Past*. Londres : Edward Arnold.
Putnam, H. (1970). “Is Semantics Possible?” *Metaphilosophy* 1 (3) : 187-201.
Putnam, H. (1975). “The Meaning of Meaning”. In *Mind and Language and Reality, Philosophical Papers, Vol. II*. Cambridge : Cambridge University Press. Pp. 215-271.
Quine, W.V.O. (1971). “The Inscrutability of Reference”. In D.D. Steinberg & L.A. Jakobovits (eds.). *Semantics*. Cambridge : Cambridge University Press. Pp. 142-154.
Quine, W.V.O. (1977). “Intensions Revisited”. *Midwest Studies in Philosophy* 2 : 5-11.
Quine, W.V.O. (1992). *Pursuit of Truth*. Cambridge & Londres : Harvard University Press.
Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1973). *A Grammar of Contemporary English*. Londres : Longman.
Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1985). *A Comprehensive Grammar of the English Language*. Londres : Longman.
Rauh, G. (ed.)(1983). *Essays on Deixis*. Amsterdam : John Benjamin.
Reboul, A. & J. Moeschler (1994). “Les phrases copulatives avec sujet pronominal en français

et en anglais". *Cahiers de linguistique française* 15 : 131-156.
Recanati, F. (2008). *Philosophie du langage (et de l'esprit)*. Paris : Folio Essais.
Recasens, M., E. Hovy & M.A. Martí (2011). "Identity, Non-identity, and Near-identity: Addressing the Complexity of Coreference". *Lingua* 121 (6) : 1138-1152.
Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Londres : Taylor & Francis.
Reuland, E.J. (1983). "Governing *-ing*". *Linguistic Inquiry* 14 (1) : 101-136.
Riegel, M., J.-C. Pellat & R. Rioul (2004). *Grammaire méthodique du français*. Troisième édition. Paris : Presses Universitaires de France.
Rossi-Gensane, N. (2010). "Encore quelques remarques sur la phrase". *La linguistique* 46 (2) : 69-107.
Rouayrenc, C. (2010a). *Le français oral, Vol. I, les composantes de la chaîne parlée*. Paris : Belin.
Rouayrenc, C. (2010b). *Le français oral, Vol. II, l'organisation et la réalisation de l'énoncé oral*. Paris : Belin.
Roulet, E. (1991). "Vers une approche modulaire de l'analyse du discours". *Cahiers de linguistique appliquée* 83 : 117-130.
Roulet, E. (1999). *La description de l'organisation du discours. Du dialogue au texte*. Paris : Didier.
Roulet, E., L. Filliettaz, A. Grobet & M. Burger (2001). *Un modèle et un instrument d'analyse de l'organisation du discours*. Berne : Lang.
Rowlett, P. (2007). *The Syntax of French*. Cambridge : Cambridge University Press.
Russell, B. (1905). "On denoting". *Mind* 14 : 479-493.
Sales, M.-P. (2005). "Expression compositionnelle ou locution : *ça craint* vs *ça barde* ?" *Linx* 53 (en ligne). http://linx.revues.org/304.
Saussure De, F. (1916). *Cours de linguistique générale*. Publication posthume. Paris : Payot.
Schmid, H. (1995). "Improvements in Part-of-Speech Tagging with an Application to German". *Proceedings of the ACL SIGDAT-Workshop*. Dublin, Ireland.
Schnedecker, C. (2006). "SN démonstratifs 'prédicatifs' : qu'est-ce qui limite leur apport informatif ?" *Langue française* 152 (4) : 39-55.
Schnedecker, C. & M. Charolles (1993). "Coréférence et identité : le problème des référents évolutifs". *Langages* 27 (112) : 106-126.
Schwarz-Friese, M., M. Consten & M. Knees (2007). *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference*. Amsterdam : John Benjamins.
Serbat, G. (1980). *Les structures du latin*. Paris : Picard.
Sichel-Bazin, R., C. Buthke & T. Meisenburg (2012). "La prosodie des questions en question : une interaction de facteurs ?" *Journées PFC Décembre 2012, Phonologie du français contemporain : des centres aux périphéries*. FMSH, Paris.
Strang, B. (1968). *Modern English Structure*. Deuxième édition. Londres : Edward Arnold.
Strang, B. (1970). *A History of English*. Londres : Methuen & Co.
Strauss, S. (2002). "*This, That,* and *It* in Spoken American English: A Demonstrative System of Gradient Focus". *Language Sciences* 24 (2) : 131-52.
Swadesh, M. (1938). "Nootka Internal Syntax". *International Journal of American Linguistics* 9 (2/4) : 77-102.
Szabó, Z. (2005). *Semantics* vs. *Pragmatics*. Oxford : Oxford University Press.
Tchagbalé, Z. (s.d.). *Cours d'initiation à la théorie des operations énonciatives*. Cours proposé pour l'Unité de Valeur (UV) SL 401 intitulé *Théories et courants linguistiques*, à l'université de Cocody, Abidjan.
Tellier, A. (1962). *Histoire de la langue anglaise*. Paris : Armand Colin.
Tellier, A. (1971). *Cours de grammaire anglaise*. Deuxième édition. Paris : Société d'édition d'enseignement supérieur.

Tesnière, L. (1959). *Éléments de syntaxe structurale*. Paris : Klinksieck.
Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam : John Benjamins.
Toupin, F. (1994). *Principes, outils et méthodes de la théorie métaopérationnelle*. Thèse de doctorat. Paris : Université Sorbonne Nouvelle Paris 3.
Toupin, F. (1998). "Référence et deixis : le cas de *this* et *that* adverbes". *Travaux linguistiques du CERLICO* 11 : 65-84.
Tutin, A., F. Trouilleux, C. Clouzot, E. Gaussier, A. Zaenen, S. Rayot & G. Antoniadis (2000). "Annotating a Large Corpus with Anaphoric Links". In P. Baker, A. Hardie, T. McEnery & A. Siewierska (eds.). *Proceedings of the Discourse Anaphora and Reference Resolution Conference, DAARC2000*. Lancaster University, Royaume-Uni, 2000.
Valois, D. (1991). *The Internal Syntax of DP*. Thèse de doctorat non publiée. Los Angeles : Université de Californie.
Van den Berg, H. (1995). *A Grammar of Hunzib*. München, Newcastle : Lincom Europa.
Versley, Y., S.P. Ponzetto, M. Poesio, V. Eidelman, A. Jern, J. Smith, X. Yang & A. Moschitti (2008). "BART: A Modular Toolkit for Coreference Resolution." In N. Calzolari (ed.). *Proceedings of the Sixth International Conference on Language Resources and Evaluation, LREC '08*. Marrakech : European Language Resources Association. Pp. 9-12.
Vogüé De, S. (1989). "La théorie des opérations énonciatives comme théorie de la référence." *Semen. Revue de sémio-linguistique des textes et discours* 4 (en ligne). http://semen.revues.org/6653.
Walker, M.A. (1997). "Centering, Anaphora Resolution, and Discourse Structure". In M. Walker, A. Joshi & E. Prince (eds.). *Centering Theory in Discourse.* Oxford : Oxford University Press. Pp. 401-435.
Webber, B.L. (1988). "Discourse Deixis: Reference to Discourse Segments". In J.R. Hobbs (eds.). *Proceedings of the 26th annual meeting of Association for Computational Linguistics, ACL'88*. Stroudsburg, PA : Association for Computational Linguistics. Pp. 113-122.
Webber, B.L. (1991). "Structure and Ostension in the Interpretation of Discourse Deixis". *Language and Cognitive Processes* 6 (2) : 107-35.
Webber, B.L., M. Stone, A. Joshi & A. Knott (2001). "Anaphora and Discourse Structure". *Computational Linguistics* 29 (4) : 545-587.
Widdowson, H. (2005). "Text, Context, Pretext". *International Journal of Applied Linguistics* 15 (3) : 421-424.
Wiese, H. & H. Simon (2002). "Grammatical Properties of Pronouns and Their Representation : an Exposition". In H. Simon & H. Wiese (eds.). *Pronouns : Grammar and Representation*. Amsterdam : John Benjamins. Pp. 1-21.
Wilmet, M. (1997). *Grammaire critique du français*. Louvain-la-Neuve : Duculot.
Wittgenstein, L. (1934[1996]). *Le cahier bleu et le cahier brun*. Traduction française de M. Goldberg & J. Sackur. Paris : Gallimard.
Wolff, H.E. (1993). *Referenzgrammatik des Hausa* [*Reference Grammar of Hausa*]. Munster, Hambourg : Lit Verlag.
Yvon, H. (1951). "CIL et CIST articles démonstratifs". *Romania* 72 : 145-181.
Zribi-Hertz, A. (1996). *L'anaphore et les pronoms : une introduction à la grammaire générative*. Villeneuve d'Ascq : Presses Universitaires du Septentrion.

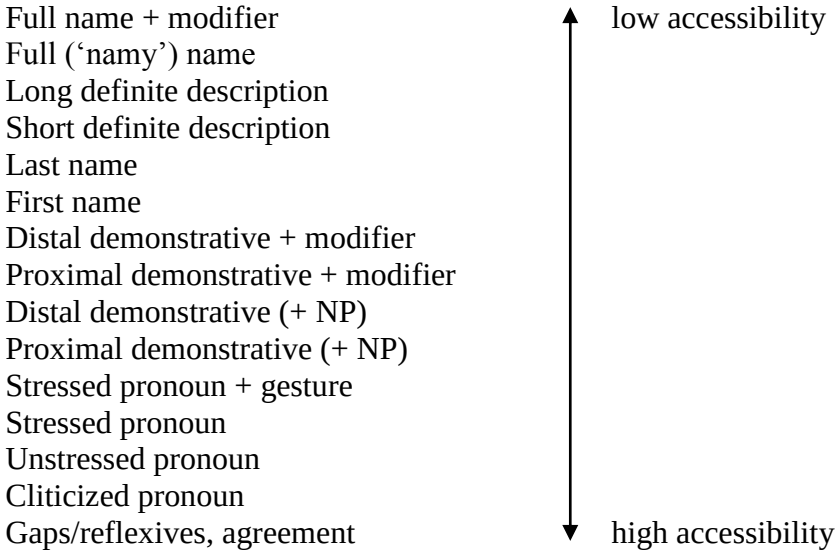
Conventions typographiques

Astérisque *	Antéposé : incorrection langagière
	Postposé : terme défini dans le glossaire (1ère occurrence seulement)
Chevrons <…>	Représentation de niveau 2 chez Culioli
	Chevauchement selon les conventions PAC/PFC
Crochets […]	En tête de définition dans le glossaire : cadre théorique ou niveau d'analyse
Guillemets simples '…'	Terme employé dans un but terminologique
	Traduction
Guillemets doubles « … »	Citation
	Définition
	Formalisation
	Référence aux entités de discours (de l'ordre de la sémantique)
Italiques	Exemples
	Emprunts linguistiques
	Référence au texte (de l'ordre du matériel linguistique)
Pap	Annotation entre crochets en GND ({pap}), 'pap' signifie 'participants à l'acte de parole' (utilisé pour les représentations du français)

Sap	Annotation entre crochets en GND ({sap}), ‘sap’ signifie ‘speech act participants’ (utilisé pour les représentations de l’anglais)

Glossaire

Accessibility Scale : Souvent traduit par ‘Echelle d’Accessibilité des Référents’ en français ; cette échelle est une classification des expressions langagières suivant la plus ou moins grande accessibilité des entités auxquelles elles renvoient. L’accessibilité est définie selon trois critères : la représentation plus ou moins saillante du référent dans l’esprit des interlocuteurs, le nombre de référents potentiels disponibles pour l’expression considérée et la distance relative entre l’expression anaphorique et son antécédent textuel. L’échelle proposée par Ariel hiérarchise quinze catégories d’expressions de la manière suivante :

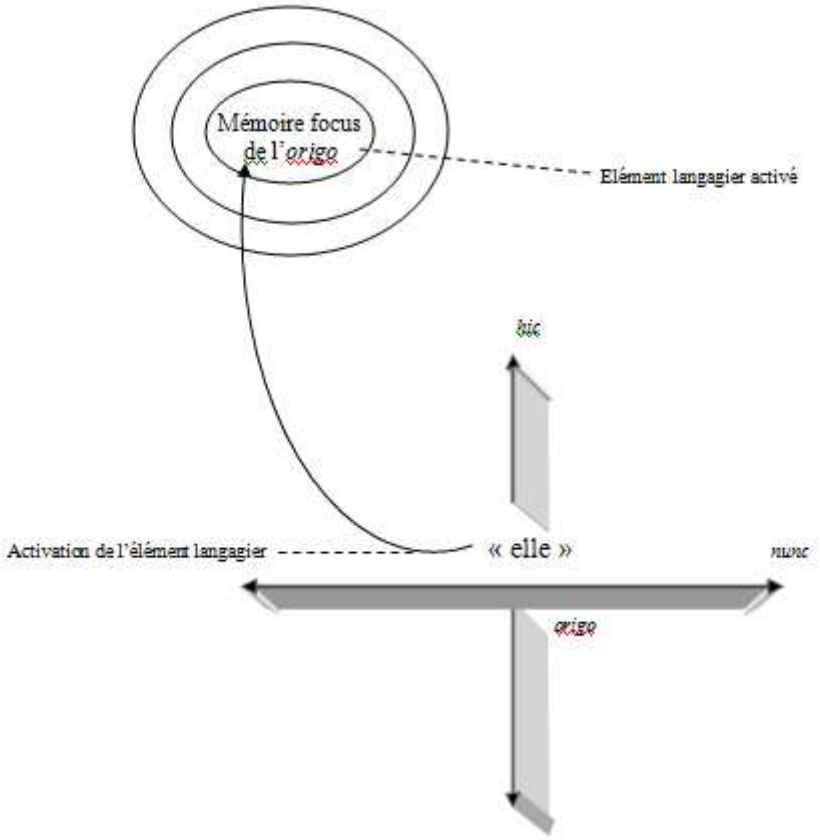


Anadeixis : Terme introduit par Ehlich (1982) afin de décrire des procédures discursives alliant anaphore et deixis.

Anaphore : [définition ‘textuelle’] Rapport de dépendance référentielle entre un élément anaphorique – pronom personnel de troisième personne, par exemple – et son antécédent – expression référentielle du type ‘syntagme nominal’. Exemple : ***Ma sœur*** (antécédent) *a dit qu’**elle*** (anaphorique) *viendrait demain* (‘**my sister** told me that **she** would come tomorrow’), où *elle* (‘she’) ne peut être compris sans la présence de *ma sœur* (‘my sister’).

[définition cognitive] Continuation référentielle d'une entité considérée comme connue et centrale dans l'esprit de l'interlocuteur/lecteur. Les pronoms personnels de troisième personne ainsi que les réflexifs sont les représentants prototypiques de l'anaphore.

[notre définition] Nous adhérons à la définition mentionnée ci-dessus, mais précisons que cette procédure discursive est caractérisée par la continuation référentielle d'un objet-de-discours déjà présent en mémoire focus (encore appelée 'mémoire immédiate') :



Anaphorique : [définition 'textuelle'] Elément référentiellement dépendant, qui sert à reprendre un antécédent – nom, syntagme nominal, proposition. Exemple : *Ces pommes sont bien rouges, mais **elles*** (anaphorique) *ne sont pas très sucrées* ('these apples are brightly coloured but **they** are not so sweet').

[définition cognitive] Elément langagier dont la référence dépend d'un antécédent, ici compris comme un objet-de-discours précédemment évoqué et cognitivement saillant dans l'esprit des allocutaires.

[notre définition] Un élément indexical est en emploi anaphorique lorsque la procédure discursive réalisée par cet élément est l'anaphore.

Antécédent : [définition 'textuelle'] Elément textuel (nom, syntagme nominal, proposition) auquel se substitue l'anaphorique. Exemple : *Le problème de **Marie*** (antécédent) *est qu'elle est persuadée d'avoir toujours raison* ('the problem is that **Mary** thinks she is always right'). Le terme 'antécédent' peut également être employé lors de la construction d'une proposition relative : ***Le garçon** qui est entré dans mon bureau hier est mon frère* ('**the boy** who entered my office yesterday is my brother').

[définition cognitive] Objet-de-discours référé de manière subséquente à sa première évocation en discours par une expression indexicale.

Cataphore : [définition textuelle] Anaphore où l'élément anaphorique est antéposé à l'antécédent. Exemple : *Depuis qu'**elle*** (anaphorique) *est partie, **ma sœur*** (antécédent) *ne donne plus de nouvelles* ('since **she** left, **my sister** has not called yet').

[définition cognitive] Cas d'anaphore où l'élément anaphorique précède le déclencheur d'antécédent. Exemple : *Depuis qu'**elle*** (anaphorique) *est partie, **ma sœur*** (déclencheur d'antécédent) *ne donne plus de nouvelles* ('since **she** left, **my sister** has not called yet').

Chronogénèse : [Psychomécanique] Opération mentale de construction de l'image-temps sur le verbe nécessaire à l'actualisation de la langue en discours. Conçu comme un mouvement de va-et-vient entre pôle étroit et pôle large, le temps opératif requis par cette opération peut être saisi à trois moments consécutifs ou chronothèses, correspondant aux modes infinitif/participe (quasi-nominal, chronothèse *in posse*), subjonctif (mouvement intercepté dans son cours, chronothèse *in furie*) et indicatif (mouvement d'actualisation intercepté à terme, chronothèse *in esse*).

Cinétisme : [Psychomécanique] Aptitude qu'a l'esprit humain de concevoir des images plus ou moins achevées d'un fait de langue donné, selon l'arrête effectué sur un

mouvement cognitif de va-et-vient entre pôle étroit et pôle large. Exemple : l'article indéfini *un* permettant d'effectuer l'extraction d'un « objet » parmi une classe d'« objets » de même nature, le cinétisme sera qualifié d''anti-extensif' ; l'article défini *le* permettant d'effectuer une référence générique, le cinétisme sera qualifié d''extensif'

Classification ontologique des entités : Initialement proposée par John Lyons (1977) et augmentée par Simon Dik (1997), cette classification vise à créer une correspondance entre expressions langagières prototypiques et catégories d'entités définies selon cinq ordres :

- Les entités d'ordre zéro (Dik) sont des dénotations de prédicats : C*ette fille est intelligente, ma sœur* ***l'****est aussi* (*l'* → « être intelligent(e) ») ;
- Les entités de premier ordre (Lyons) sont définies dans l'espace, dotées de caractéristiques stables ; dans notre réalité intersubjective, elles ont une matérialité physique, elles 'existent'. De manière générale, les « individus », les « animaux » et les « choses » sont des entités de premier ordre. Les catégories morphosyntaxiques les plus à même de désigner ces entités sont certains noms – *chat, chien, eau, table,* où nous retrouvons les fameux noms à dénotation de classes et substances naturelles de Putnam – les noms propres, les pronoms ainsi que les syntagmes nominaux ;
- Les entités de deuxième ordre (Lyons) sont définies dans l'espace et le temps, ce sont des « événements », des « procès », des « états de choses » ; nous pouvons dire des entités de deuxième ordre qu'elles 'arrivent'. Les structures langagières les plus à même de désigner ces entités sont certains noms – *mort, naissance, arrivée, départ,* etc. – ainsi que les verbes qui leur correspondent. Ainsi, nous pouvons constater que les substantifs prototypiques désignateurs d'entités de deuxième ordre résultent d'une nominalisation, processus particulièrement visible en anglais : *set*/*setting, leave*/*leaving,* etc ;
- Les entités de troisième ordre sont de nature abstraite, en dehors de l'espace et du temps, ce sont des « concepts », des « sentiments » ; de manière générale, nous pouvons soumettre ces entités à la réfutabilité, à la croyance ou au doute. Les structures propositionnelles sont les plus à même de renvoyer à des entités de troisième ordre, mais certains noms, comme *pensée, sentiment* ou *raison,*

désignent également ce type d'entités. Tout comme ceux désignant des entités de deuxième ordre, nous pouvons constater que ces noms résultent d'une nominalisation : *penser*/*pensée* ('think/thought'), *sentir*/*sentiment* ('feel/feeling'), *raisonner*/*raison* ('reason (V)/reason (N)') ;

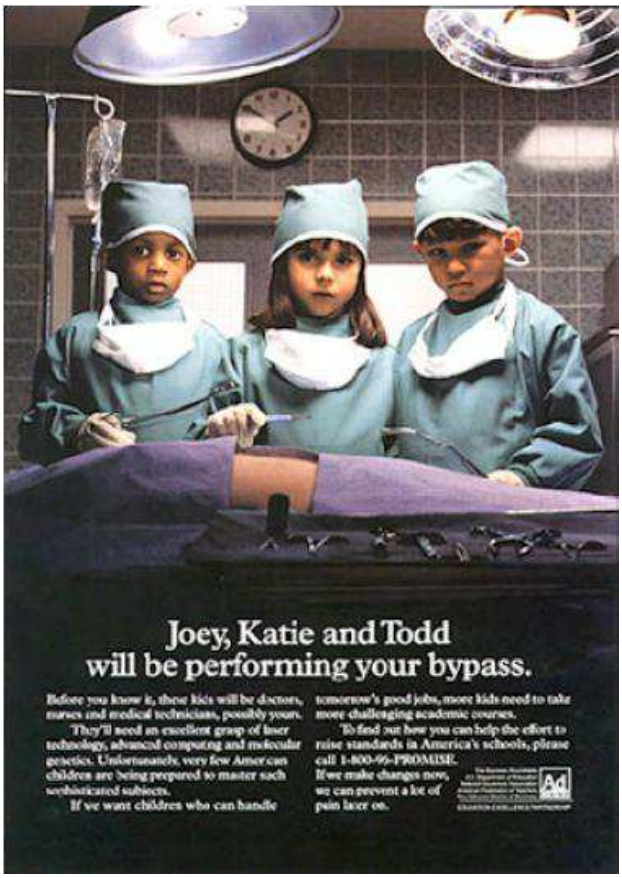
- Les entités de quatrième ordre (Dik) sont des illocutions : *« Je ne croyais pas avoir épousé une aventurière. » Elle parut très froissée et me répondit : « Ce n'est pas une aventure, ça dure depuis deux ans. » Elle me fit* ***cette déclaration****, je dois le dire, avec une certaine dignité.* (Flers & Cavailler, *Monsieur Brotonneau*, 1914)

Clause : Les clauses sont définies comme « des îlots de connexité rectionnelle », où « chaque élément (morphème, syntagme) […] est en relation de dépendance ou d'interdépendance avec au moins un autre élément, si bien que le tout se tient » (Berrendonner, 2004). Exemple : *Cette vieille voiture, ils vont arriver tard* est constitué de deux clauses, *Cette vieille voiture* et *ils vont arriver tard*, dans la mesure où *Cette vieille voiture* n'est pas régi par un élément du second segment *ils vont arriver tard* (Rossi-Gensane, 2010 : 75).

Cohérence : [SFG] La cohérence est obtenue grâce au principe de cohésion, qui fait qu'une suite d'énoncés forme un tout intelligible percevable par le destinataire.

Cohésion : [SFG] Notion sémantique qui se définit par les liens grammaticaux et lexicaux qui unissent un ensemble de segments textuels.

Conceptual Blending : [Linguistique cognitive] Théorie développée par Gilles Fauconnier et Mark Turner (2002), selon laquelle les éléments et les relations diverses qui proviennent de plusieurs scénarios sont intégrés en un seul espace cognitif, grâce à un mécanisme subconscient nommé Intégration Conceptuelle. Un exemple célèbre et très parlant est celui d'une publicité incitant le public à verser de l'argent pour les études supérieures et la recherche :



Afin de comprendre le message, le lecteur doit assimiler deux espace-temps différents : la période t_0 où Joey, Katie et Todd sont des enfants scolarisés à l'école primaire, et la période t_1 où ces enfants seront adultes, chirurgiens, et maîtres de la vie de patients (dont celle du lecteur).

Contexte : Plusieurs acceptions. Le contexte peut désigner les éléments linguistiques qui entourent une occurrence donnée ; désigne alors une part du 'co-texte', où le 'texte' dénote l'ensemble de signes verbaux et non verbaux dont l'énonciateur dispose pour effectuer un acte de communication. Peut également désigner la situation d'énonciation spatio-temporellement circonscrite qui constitue l'ancrage de l'acte de communication considéré ; désigne alors le contexte situationnel ou extralinguistique de l'acte de communication.

[notre définition] Le contexte est tout ce qui est pertinent à la production et à l'interprétation de ce qui est dit/lu/communiqué à l'instant t_0. Nous envisageons le contexte comme un élément essentiellement dynamique et composé de trois macrostructures : le contexte situationnel, le contexte discursif et les connaissances :

Contexte					
Situationnel		Savoir		Discursif	
Etroit	Large	Etroit	Large	Etroit	Large
Environnement physico-social immédiat	Environnement physico-social étendu	Savoir partagé entre les participants à l'acte communicatif + état psychologique de chaque participant + calcul de l'un et l'autre des participants sur l'état psychologique de l'un et l'autre des participants	Connaissances du monde	Discours construit en amont de la production langagière en t_0	Discours construit en amont

Coréférence : Deux expressions langagières sont en coréférence si elles se réfèrent à la même entité. Exemple : ***Le petit ami de ma sœur*** *est bien gentil, mais je trouve* ***ce garçon*** *un peu fainéant* (‘**my sister's boyfriend** seems nice but I think **the boy/he** is rather lazy’).

Co-texte : Si le ‘texte’ est compris comme l'ensemble de signes verbaux et non verbaux dont l'énonciateur dispose pour effectuer un acte de langage, le co-texte est défini comme l'ensemble de ces signes qui entourent une occurrence textuelle particulière.

Déclencheur d'antécédent : Signe textuel, linguistique ou sémiotique, renvoyant à un objet-de-discours potentiellement référé par une expression indexicale subséquente. Exemple : *Prenez* ***deux pommes bien mûres*** (déclencheur d'antécédent) *et pelez-****les*** (pronom en emploi anaphorique).

Déictique(s) : Les déictiques – terme employé ici au pluriel pour éviter toute confusion – sont des expressions langagières inséparables de l'ancrage énonciatif. Nous pouvons distinguer trois catégories de déictiques : les déictiques de personne (*je, tu*) ; les déictiques spatio-temporels (*ici, maintenant, hier*…) ; les déictiques de monstration (*ce, cette, ceci, cela, ça*…). Ces déictiques sont généralement divisés en deux sous-catégories. D'un côté, nous avons les déictiques purs comme *ici*, *maintenant*, *je* et *tu*, qui sont dits ‘token-reflexive’ dans le sens où ces items renvoient directement à l'instanciation qui leur correspond au moment de l'énonciation. D'un autre côté, nous disposons des déictiques « opaques » (Kleiber, 1983), tels que les expressions

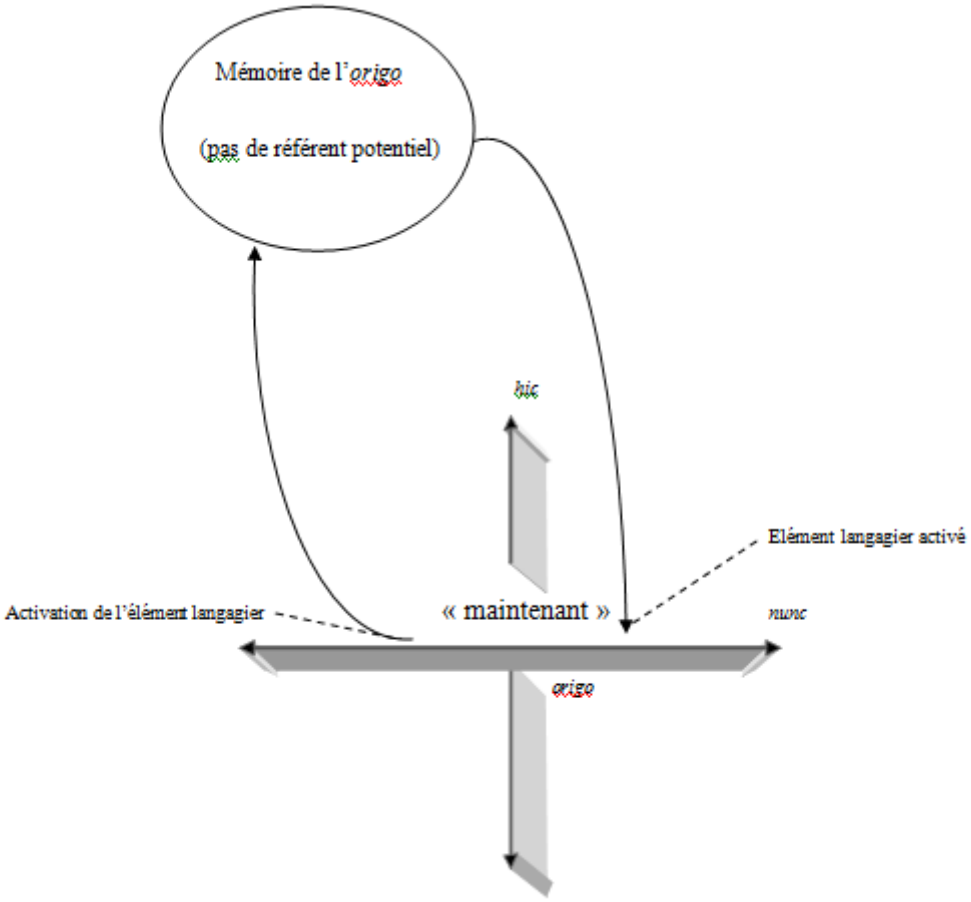
démonstratives, dont la référence repose sur le repère énonciatif sans en dépendre totalement. Aux déictiques s'opposent les anaphoriques, dont les représentants canoniques sont les pronoms de troisième personne, qui opèrent une reprise référentielle.

[notre définition] Nous limitons l'étiquette de 'déictiques' aux déictiques purs : étant 'token-reflexive', ces éléments nous permettent de définir la deixis comme opérant un retour sur le repère énonciatif pour trouver un référent. Le terme 'indexicaux' remplace alors ceux d''éléments déictiques opaques' et d''éléments anaphoriques'.

Deixis : [définition 'traditionnelle'] Mode de référence qui implique tout recours à la situation d'énonciation. Le terme 'deixis' provenant du grec *deiktikos* ('montrer'), le cas prototypique de la deixis est celle s'accompagnant d'un geste ostensif, en référence à un « objet » quelconque présent en situation, lui-même nommé *demonstratum* ('deixis *in prasentia*'): *Oh ! Lis* ***ça****, tu vas voir, c'est très drôle !* (l'énonciateur désignant un article de journal qu'il tient dans sa main). Les déictiques purs que sont *je* et *tu* sont également vecteurs de deixis, puisque ces symboles langagiers ne trouvent leur instanciation qu'en étant perpétuellement ancré dans la situation d'énonciation. Si la deixis *in prasentia* nécessite le recours à la situation directe, la deixis *in absentia* accepte une définition plus large, regroupant notamment les cas de référence indirecte, où le *demonstratum* n'est pas le référent visé, ainsi que la 'deixis mémorielle' – aussi nommée deixis *am phantasma* (Bülher, 1934[2009]), ou encore 'deixis emphatique' (Kleiber, 2007) – où la situation d'énonciation se trouve en quelque sorte transposée dans l'univers mental de l'énonciateur.

[définition cognitive] Introduction d'un nouvel objet-de-discours au sein du discours ; dans les théories alliant cognition et interaction communicative, cette définition est augmentée de la fonction qu'a la deixis d'attirer l'attention de l'allocutaire vers un nouvel objet-de-discours.

[notre définition] Nous adhérons à la définition mentionnée ci-dessus, mais précisons que cette procédure discursive est caractérisée par un retour sur le repère énonciatif ; la mémoire des interlocuteurs étant située entre l'élément déictique et le référent visé, un mouvement de rétroaction est nécessaire pour que l'élément langagier soit activé :



Deixis *am Phantasma* : La deixis *am Phantasma*, parfois traduite par 'deixis à l'imaginaire', est introduite par Bühler en 1934 et désigne un cas de deixis où l'*origo* – ou centre déictique – est déplacé. Ainsi, le repérage spatio-temporel ne s'effectue pas dans la situation d'énonciation mais dans un espace imaginaire.

Deixis de discours : Selon la terminologie employée notamment par Webber (1991), Himmelmann (1996) et Diessel (1999), la deixis de discours désigne un emploi particulier des démonstratifs en référence à des événements ou des contenus propositionnels précédemment mentionnés. Exemple : *Une deuxième piste extérieure a été créée, au total nous avons remué et tassé 124 tonnes de sable. Quarante boxes homologués ont été installés pour les chevaux et 3 ensembles de gradins montés. C'est une logistique très lourde. Toutes nos félicitations au personnel qui a effectué* ***ces lourdes manipulations*** (*l'Indépendant en ligne*, 25/09/10). Selon la terminologie employée par Cornish (1999, 2007), la deixis de discours se caractérise par l'insertion d'un référent nouveau au sein du discours à partir d'un empan de discours précédant.

Exemple : *The guest speakers will be recorded, and **these** will be available at the end of the conference* (entendu lors de la présentation d'un colloque à Edimbourg), où *these* renvoie aux « enregistrements », et non aux « conférenciers invités ».

Deixis textuelle : Terme employé par John Lyons (1977) pour désigner l'usage mentionnel – *i.e.* métalinguistique – de certains pronoms. Exemple : A : *Ceci est un rhinocéros*. B : *Un quoi ? Est-ce que tu peux me **l'épeler ?*** ('A: That's a rhinoceros. B: A what? Spell **it** for me' ; exemple issu de Lyons, 1977 : 667).

Dénotation : Dans le système sémantique de Frege (1892), 'dénotation' et 'référence' sont confondues et définissent l'objet du monde désigné par un terme singulier et la propriété à laquelle renvoie un terme général. Selon notre définition, et suivant celle de Lyons (1977), la 'dénotation' d'un terme renvoie à l'image mentale suscitée par ce terme ; ainsi, la 'dénotation' serait comparable à une photo dans le lexique mental.

Discours : Plusieurs acceptions. Selon les théories, le discours peut être assimilé au texte ou bien à l'énoncé, il couvre même parfois ces deux notions alors confondues.

[notre définition] Construction cognitive créée à partir d'un texte et du contexte à l'œuvre ; le discours est une représentation mentale d'« objets » évoqués, de « concepts », auxquels ont été attribuées des « propriétés », en somme, des « objets » dont on a parlé. Ces traces sont en perpétuel développement lors de l'acte communicatif, que ce dernier soit interactif, comme dans une conversation, monologique, comme lors d'une conférence, ou spatio-temporellement discontinu, comme l'est la lecture.

Effet : [Psychomécanique] Terme que l'on retrouve dans les appellations 'unité d'effet', 'langage effectif', '*in effe*', et qui désigne l'actualisation en discours d'un élément de la langue. La valeur d'effet ou en effet est déterminée par la saisie particulière et momentanée qu'opère l'esprit humain lors du mouvement cognitif de va-et-vient effectué entre un pôle étroit et un pôle large.

Endophore : [définition textuelle] Comparable à l'anaphore, l'endophore définit la reprise d'un élément textuel, et s'oppose ainsi à l'exophore.

Exophore : [définition textuelle] Comparable à la deixis, l'exophore signale tout recours à la situation d'énonciation, la référence ne s'opèrant pas via le texte. L'exophore s'oppose ainsi à l'endophore.

Extraction : [TOPE] Opération de construction d'occurrence à partir d'une notion, l'extraction sélectionne une occurrence parmi une classe d'occurrence. L'extraction est prototypiquement effectuée par l'article indéfini. Exemple : ***Un appartement en ville*** *est bien plus pratique qu'****une maison à la campagne*** ('A flat in the town centre is far more convenient than a house in the countryside').

Fléchage : [TOPE] Opération de construction d'occurrence qui se base sur l'extraction. Sont distingués le fléchage situationnel étroit et large – qui a recours à la situation d'énonciation – ainsi que le fléchage contextuel étroit et large – qui nécessite la présence d'un élément linguistique. Le fléchage est prototypiquement réalisé par l'article défini. Exemple : *J'ai acheté une veste et une paire de chaussures ;* ***la veste*** *est en cuir* ('I bought a jacket and a pair of shoes; **the jacket** is a leather one').

Focus : [pragmatique ; structure informationnelle] Partie de l'énoncé qui constitue l'information nouvelle, le focus est souvent paraphrasé en 'ce que l'on dit à propos de X', X étant le topique de l'énoncé. Le focus est parfois utile à des fins contrastives et est aussi appelé 'rhème'.

Forme Schématique : [TOPE] Pôle invariant de chaque item langagier qui définit à la fois les relations que celui-ci entretient avec son co-texte et l'influence de la forme schématique des autres éléments linguistiques qui pourraient interagir avec lui.

Functional Discourse Grammar : [Linguistique fonctionnelle] La FDG (Hengeveld & Mackenzie, 2008) est une expansion de la FG (Functional Grammar) développée par Simon Dik en 1997. Si la FG est avant tout un modèle d'analyse des énoncés produits par des phrases, la FDG désire rendre compte de phénomènes ayant trait à des unités plus larges que la phrase, mais aussi de ceux concernant des unités plus petites. Le modèle d'analyse de la FDG est une structure descendante et modulaire : les configurations choisies en niveaux supérieurs déterminent celles en niveaux inférieurs, mais chaque niveau est indispensable au fonctionnement général de l'acte communicatif. La FDG décrit quatre composantes majeures :

- La composante conceptuelle concerne les intentions de communication des participants à l'acte langagier ainsi que les représentations cognitives qui leur correspondent. C'est en fonction de cette composante qu'est décidée la configuration de la composante grammaticale ;
- La composante grammaticale porte sur la formulation langagière déterminée par les intentions de communication. Cette composante grammaticale constitue le cœur de l'architecture de la FDG est se divise en quatre niveaux d'analyse :
    - Le niveau interpersonnel, que l'on pourrait assimiler à la pragmatique ; c'est à ce niveau que sont pris en compte les participants à l'acte communicatif et les actes perlocutoires ;
    - Le niveau représentationnel concerne la sémantique ; c'est à ce niveau que sont co-indexés les expressions référentielles et les anaphoriques correspondants, par exemple, mais aussi la catégorisation des entités selon leur ordre ontologique, la marque distale ou proximale des démonstratifs si l'on décrit l'anglais ;
    - Le niveau morphosyntaxique analyse la structure syntaxique et morphologique de la séquence considérée ;
    - Le niveau phonologique décrit les unités phonologiques et le contour intonatif de l'énoncé.
- La composante contextuelle est principalement dédiée au stockage des informations accessibles aux participants à tout point de l'acte communicatif – ce que nous appelons le 'discours'. Néanmoins, la définition du contexte par Connolly (2007) au sein de ce modèle affine la description de la composante contextuelle (cf. ce chapitre, section 2.3) ;
- La composante de sortie ('output') ressort des signes phoniques, graphiques et sémiotiques employés par les participants à l'acte langagier.

Functional Grammar : [Linguistique fonctionnelle] La FG est un modèle initialement développé par Simon Dik dans les années soixante-dix, dont la dernière version est présentée dans son ouvrage posthume de 1997. Le but de cette théorie est d'expliquer la construction des phrases à la lumière de la volonté qu'ont les interlocuteurs de communiquer. Ainsi, la FG s'articule autour de la fonctionnalité du langage et décrit trois niveaux de fonctions : les fonctions sémantiques – agent, patient, bénéficiaire,

etc. – les fonctions syntaxiques – sujet et objet – et les fonctions pragmatiques – thème et rhème, topique et focus.

Givenness Hierarchy : Cette échelle, présentée par Gundel, Hedberg & Zacharski (1993) se propose de hiérarchiser certaines expressions indexicales en fonction du statut cognitif de leur référent. Ces statuts cognitifs sont hiérarchisés et imbriqués du plus au moins saillant :

in focus	>	activated	>	familiar	>	uniquely identifiable	>	referential	>	type identifiable
{*it*}		{*this*, *that*, *this* N}		{*that* N}		{*the* N}		{indefinite *this* N}		{*a* N}

Ainsi, le clitique *it* serait typiquement anaphorique puisque continuant la référence d'une entité saillante, les pronoms démonstratifs ainsi que le syntagme nominal démonstratif *this N* réfèreraient à une entité activée – mais moins saillante que celle visée par *it* – le SN démonstratif *that N* coderait un référent connu des interlocuteurs, un SN défini serait prototypique de la référence spécifique, contrairement au SN démonstratif *this N* présentatif et au SN indéfini, ce dernier introduisant un référent nouveau et donc étant prototypiquement déictique.

Grammaires Catégorielles : Les Grammaires Catégorielles regroupent un ensemble de représentations formalisées et d'analyses de la syntaxe des langues naturelles motivées par le principe de compositionnalité et organisées selon le principe des constituants syntaxiques agencés en relations fonctions/arguments. Une Grammaire Catégorielle représente la syntaxe par un ensemble fini de symboles et est constituée de deux paramètres : le lexique qui assigne un certain nombre de traits à chaque catégorie – et donc un certain nombre de symboles – et des règles inférentielles qui déterminent l'agencement de ces symboles. De manière prototypique, une Grammaire Catégorielle comprend deux types de fonctions, B/A et A\B : B/A résulte d'un constituant B lorsqu'il est suivi d'un constituant A, A\B résulte d'un constituant B lorsqu'il est précédé d'un constituant A ; par exemple, l'article défini *le* ('the') est représenté NP/N, car *le* ('the') suivi d'une catégorie nominale forme un syntagme nominal ('noun phrase') NP. Selon ces principes, l'organisation des constituants de chaque langue considérée est entièrement fixée par le lexique.

Grammaire Métaopérationnelle : Théorie développée par Henri Adamczewski, née d'une réflexion sur la structure « *BE + -ing* » en anglais (BE+ING *dans la grammaire de l'anglais contemporain*, 1976). Le but de cette 'grammaire' est d'une part de décoder les opérations permettant d'obtenir les énoncés produits par l'énonciateur, à partir du système langagier, d'autre part de déceler l'invariance fonctionnelle des marqueurs de surface, qui sont eux-mêmes la trace visible des opérations cognitives sous-jacentes. Les structures de la langue considérée sont dites produites par un microsystème binaire unique combinant 'rhème' ou 'phase 1' et 'thème' ou 'phase 2'. Du point de vue des choix paradigmatiques possibles, le thème caractérise un choix limité par des contraintes de type anaphorique, contrairement au rhème qui caractérise un choix ouvert. Cette dualité 'ouvert/fermé' transparaît dans ce qu'Adamczewski nomme 'le Principe de cyclicité' : *un/le, a/the* ; *ce N-ci/-là, this/that.*

Grammaire Notionnelle Dépendancielle : Développée par John Anderson (1991, 1997, 2007, 2011) et issue de la tradition de Tesnière (1959), la Grammaire Notionnelle Dépendancielle (GND) est une grammaire explicite où la syntaxe est guidée par la sémantique. Les éléments atomiques en GND sont des traits unaires – P pour 'prédicativité' et N pour 'référentialité – qui se combinent afin de définir les catégories lexicales et leurs projections syntagmatiques. La relation entre tête et dépendant(s) est centrale à la construction des unités lexicales et syntaxiques, relation qui est formalisable sous forme d'arbres de dépendances. Les traits P et N sont conçus comme étant les pôles extrêmes d'un continuum, et non comme des traits parallèles. La pondération entre P et N permet ainsi de complexifier le continuum catégoriel :

- Catégories lexicales primaires en GND :

{P}	**{P;N}**	**{P:N}**	**{N;P}**	**{N}**
Auxiliaires	Verbes	Adjectifs	Noms communs	Noms propres
P/N = 4/0	3/1	2/2	1/3	0/4

Ces catégories primaires peuvent être associées à une catégorie secondaire, par exemple, le temps et l'aspect pour un verbe, le genre et le nombre pour un nom ; par exemple, {P;N{progressive}}.

- Catégories fonctionnelles primaires en GND :

Finitness	{P}	Verbe	Catégories lexicales prototypiques
Determinative	{N}	Nom	
Comparator	{P.N}	Adjectif	
Functor	{ }		

Les catégories fonctionnelles étiquettent la tête syntagmatique et peuvent se refléter dans la langue de diverses manières, que ce soit par l'emploi de mots distincts ou par une marque flexionelle. Notons ici que la GND élimine les catégories vides par la notion d'incorporation, à travers des opérations de sous-jonction *vs* adjonction.

Grammaire Universelle : [Linguistique générative et transformationnelle] Pensée par Noam Chomsky, la Grammaire Universelle ou GU est en fait une métagrammaire censée représenter les structures syntaxiques communes à toutes les langues du monde.

Hypostatisation : Processus mental qui consiste à matérialiser un concept – mouvement de l'abstrait au concret.

Impure textual deixis : John Lyons (1977) définit la 'deixis textuelle impure' comme la référence d'un pronom à une entité de troisième ordre précédemment évoquée au sein du discours.

Incidence : [Psychomécanique] Désigne le rapport de dépendance qu'il existe entre deux termes linguistiques. Exemple : l'adjectif est incident au substantif, en ce que l'adjectif est un 'apport de signification' au substantif, qui est son 'support de signification'.

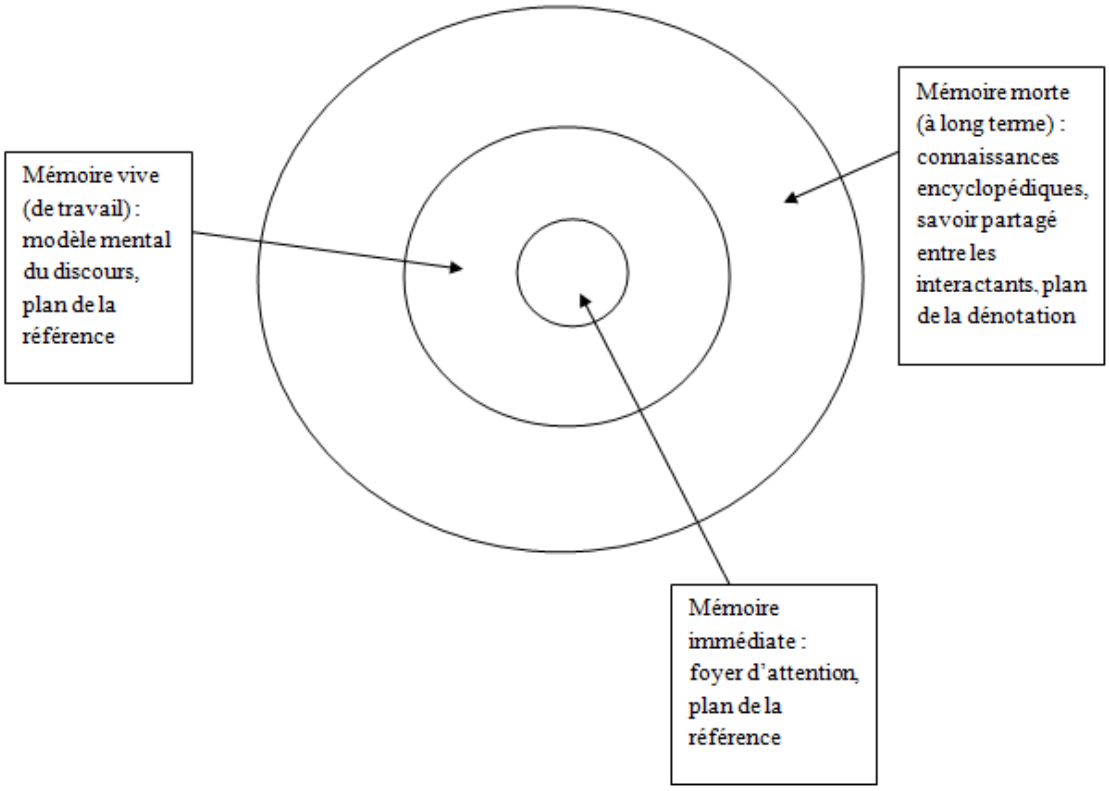
Indexicalité : Processus de pointage vers un élément du contexte situationnel, discursif, ou lié aux connaissances des interlocuteurs qu'effectue une expression langagière afin de pouvoir référer.

Lexis : [TOPE] La lexis est une opération de repérage qui agence les notions représentées en niveau 1 selon un axe à trois places, afin de rendre possible la mise en place d'une relation prédicative en niveau 2.

Macro-syntaxe : La macro-syntaxe est définie comme le « domaine des relations qu'on ne peut pas décrire à partir des rections de catégories grammaticales » (Blanche-Benveniste *et al*., 1990 : 290), la « rection verbale » étant l'« ensemble des éléments, sujet et compléments, organisés syntaxiquement par le verbe » (Blanche-Benveniste *et al.*, 1990 : 290), ou encore comme le […] domaine des relations grammaticales correspondant en partie à ce que l'on a nommé « la syntaxe détachée ». Les relations de macro-syntaxe, qui peuvent s'étendre sur des morceaux plus vastes que la 'phrase' traditionnelle, ne peuvent pas être ramenées aux notions de subordination ou de coordination. Dans *trois francs il réclame*, *trois francs* est syntaxiquement le complément de *il réclame* et en même temps, sur le plan de la macro-syntaxe c'est un élément noyau, qui peut former à lui seul un énoncé autonome, avec une mélodie de fin d'énoncé (Blanche-Benveniste, 1997 : 159).

Marqué (*vs* non marqué) : Le membre marqué d'une paire combinant élément marqué et élément non marqué est assujetti à des restrictions d'emplois, le membre non marqué étant utilisé par défaut. Exemple : en anglais, *this* est le membre marqué de la paire *this/that*, *that* pouvant en principe se substituer à toutes les occurrences de *this*, l'inverse n'étant pas avéré.

Mémoire : Nous proposons la mémoire comme facteur à la fois premier et indispensable à la dimension cognitive du langage. A la manière d'un ordinateur, nous présentons la mémoire sous forme de trois couches concentriques : la mémoire morte (ou à long terme, ou encore inactive), la mémoire vive (ou active) et la mémoire immédiate (ou focus) :



Micro-syntaxe : La micro-syntaxe – terme parfois suppléé par l'unique item 'syntaxe' – désigne un système de dépendances et d'interdépendances – aussi appelées 'rections' – construit autour du verbe, du nom et de l'adjectif. Cette combinatoire de segments est définie par des contraintes catégorielles et distributionnelles.

Notion : [TOPE] La notion est une représentation cognitive dotée de particularités physico-culturelles et antérieure à la catégorisation lexicale. La notion est représentée entre barres obliques. Exemple : /homme/.

Noyau : Dans le modèle aixois (Blanche-Benveniste *et al.*, 1990 : 114), le noyau est l'unité minimale centrale de la macro-syntaxe, qui constitue un énoncé autonome. Le noyau peut être précédé ou suivi d'unités minimales non autonomes, les préfixes et les suffixes (ou postfixes).

Occurrence : [TOPE] Elément d'ordre linguistique et référentiel qui se définit comme l'instanciation d'une notion. Exemple : « musicien », « instrument », « mélodie »,

« harmonie », « accorder » sont des occurrences linguistiques et sémantiques de la notion /musique/.

Période : Dans le modèle fribourgeois (Berrendonner, 1990 ; 1993), la période est l'unité maximale de la macro-syntaxe, composée d'une ou plusieurs clauses et se terminant par une prosodie conclusive.

Programme Minimaliste : [Linguistique générative transformationnelle] Version de la Grammaire Générative et Transformationnelle chomskyenne développée en 1995 et toujours en vigueur. L'architecture du Programme Minimaliste se veut optimale ; ainsi, le système computationnel qui prend en charge les entrées provenant du lexique se résume à deux opérations : 'Merge' ('fusion') et 'Move' ('déplacement'). Merge est une opération binaire qui, à partir de deux entités α et β, crée une troisième entité γ, dont α et β sont les constituants. Cette fusion est soumise aux contraintes définies par le schéma X-barre. Move est une opération de déplacement des constituants qui permet d'obtenir l'agencement observé en sortie. Les opérations Merge et Move sont à leur tour assujetties à un processus de vérification afin d'obtenir une bonne formation en sortie. Chaque item provenant du lexique est doté de traits formels – par exemple, le genre et le nombre ; ce sont ces traits qui déclenchent les dérivations syntaxiques, notamment la nécessité ou non d'effectuer un déplacement. Enfin, chaque élément en sortie est associé à deux niveaux de représentation : une 'Forme Phonologique' (PF) en relation avec les aspects physiques de la langue et une 'Forme Logique' (LF) concernant l'interface entre la syntaxe et le but communicatif.

Psychomécanique : Théorie du langage développée par Gustave Guillaume (*Problème de l'article et sa solution dans la langue française*, 1919). Guillaume distingue 'langue' et 'discours', le premier correspondant au système langagier – *in posse* 'en puissance' – et le second à son actualisation – *in esse* 'en effet' – par l'énonciateur lors d'un acte de langage. La Psychomécanique ou Psychosystématique décrit les opérations langagières comme un mouvement de l'esprit entre un pôle étroit et un pôle large – et *vice-versa* – sur lequel s'effectuent des arrêts momentanés ; la distance relative de cet arrêt, par rapport aux pôles large et étroit, définit la valeur d'effet que prend l'invariant de puissance propre à chaque item langagier.

Puissance : [Psychomécanique] Terme que l'on retrouve dans les appellations 'unités de puissance', '*in posse*', qui désigne les éléments systémiques de la langue. Chaque item langagier est doté d'un pôle invariant qui permet de rendre compte de tous ses emplois possibles en discours.

Référence : Selon notre définition, et suivant celle de Lyons (1977), la référence d'un terme linguistique est l'entité discursive renvoyée lors de l'activation de ce terme en contexte énonciatif.

Référent : [sémantique] Aussi nommé 'entité discursive', 'objet-de-discours', ou 'objet du discours', le référent est une représentation mentale qui résulte de l'actualisation en discours d'une expression langagière. Le référent émerge et se développe tout au long de l'activité de langage, tant que celui-ci est évoqué.

Référentielle (expression) : Selon certaines théories, une expression référentielle se distingue d'un pronom démonstratif ou clitique en ce qu'elle est référentiellement autonome, et par conséquent, désigne le plus souvent un syntagme nominal. Selon nous, les pronoms sont aussi des expressions référentielles ; la différence est qu'ils renvoient par un processus d'indexicalité.

Règle de réécriture : [Linguistiques distributionnelle et générative] Règle qui décrit formellement la structure interne d'un syntagme, chacun étant composé d'une tête rectrice, de constituants obligatoires – appelés compléments ou arguments – et de constituants facultatifs – appelés modifieurs ou adjoints.

Regroupement : Dans le modèle aixois (Blanche-Benveniste *et al*., 1990), le regroupement est l'unité maximale de la macro-syntaxe, composée au moins d'un noyau.

Relation prédicative : [TOPE] Opération de niveau 2 qui ré-agence la lexis en conférant des rôles d'actants aux notions pré-organisées par la lexis. L'opération de relation prédicative permet d'obtenir une structure syntactico-sémantique <argument sujet + verbe + argument objet (+ circonstants éventuels)>.

Repérage : [TOPE] Relation fondamentale dans la TOPE car tout terme est repéré par rapport à un terme-repère. Le repérage est une opération essentielle au niveau des

constructions énonciatives en ce qu'elle est à la base de la détermination, et, donc, de la construction d'un énoncé par rapport au point-origine du repère énonciatif.

Schwa : Le schwa – prononcé /ʃ v a/ ou /ʃ w a/ – désigne la voyelle centrale, neutre, notée /ə/ en API (Alphabet Phonétique International).

Sémantèse : [Psychomécanique] Désigne les particularités syntactico-sémantiques d'un élément du lexique.

Sens : Selon le système de Frege (1892), le sens est interne à la langue et constitue le mode de présentation de la référence d'un terme linguistique. En cela, le sens est constitutif de la langue et détermine la référence. Selon notre définition, et suivant celle de Lyons (1977), le sens d'un terme linguistique autorise la délimitation des mots, par les connaissances que nous avons des propriétés associées à ce terme : le mot *chat* est une unité lexicale qui se différencie des termes *chien*, *table* et *ciel*, par exemple, mais qui entretient des relations plus ou moins proches avec les termes *animal*, *félin* et *siamois*, en faveur de ce que l'on croit être un « chat » : « animal de compagnie, mammifère, félin, de caractère indépendant, souvent ingrat, voire sournois ».

Subduction : [Psychomécanique] Désigne l'opération de dématérialisation et de désémantisation des mots.

Submorphème : [Linguistique énonciative] Unité plus petite que le morphème dotée d'une valeur notionnelle invariante. Le submorphème se trouve à l'initiale des items lexicaux et grammaticaux, la suite de ce qui compose le morphème dans son entier s'appelant 'rime'. Exemple : TH- dans *this*, *that*, *though*, *thus* signale un préconçu cognitif, WH- dans *which*, *who*, *whose* indique un vide informationnel.

Systemic Functional Grammar : La SFG (Halliday, 1985) concerne principalement les choix de langue dont disposent les individus parlants afin de communiquer à bon escient. La SFG comporte trois niveaux d'analyse, à savoir la sémantique, la phonologie et la morphosyntaxe – 'lexicogrammar' selon la terminologie de Halliday –, régies par des bases fonctionnelles appelées 'métafonctions' : les métafonctions idéationnelle, interpersonnelle et textuelle :

- La métafonction idéationnelle est en relation avec le domaine d'application du texte – ce que Halliday nomme 'field', parfois traduit par 'champ' – et que l'on pourrait décrire comme le genre du texte considéré, le thème majeur discuté, le scénario. Ce domaine d'application détermine les choix lexicaux et l'agencement des différentes catégories morphosyntaxiques ;
- La métafonction interpersonnelle concerne les relations entre les divers participants à l'acte communicatif ainsi que les conséquences que celles-ci pourraient avoir sur les choix perlocutoires et lexicaux ;
- La métafonction textuelle prend en compte la cohérence et la cohésion internes d'un texte – relations anaphoriques, enchaînements, ellipses, etc.

Tenseur binaire radical : [Psychomécanique] Aussi nommé 'schème binaire radical', le tenseur binaire radical désigne le psychomécanisme fondamental de la pensée humaine, décrit comme un mouvement de va-et-vient entre un pôle étroit et un pôle large. Le mouvement de l'étroit au large est qualifié de 'généralisant', le mouvement du large à l'étroit est dit 'particularisant'. L'arrêt particulier et momentané de l'esprit sur l'un ou l'autre de ces axes permet l'actualisation d'une unité de puissance en unité d'effet.

Texte : Ensemble de signes verbaux et non verbaux dont l'énonciateur dispose afin d'effectuer un acte de langage. Le texte, dans cette acception, comprend non seulement les éléments linguistiques mais aussi la gestuelle.

Théorie des Opérations Prédicatives et Enonciatives : Théorie du langage développée par Antoine Culioli. Si la TOPE se classe parmi les linguistiques de l'énonciation, elle s'en démarque par l'approche choisie, radicalement différente des théories voisines. Ici, l'acte de langage n'est pas vu comme l'actualisation d'un système langagier par un énonciateur quelconque, mais comme une construction du sens. Ainsi, l'espace cognitif humain abrite un ensemble de notions – concepts dotés de particularités physico-culturelles –, antérieures à toute lexicalisation. Ces notions sont ensuite organisées selon un schéma à trois places afin de former des lexèmes (niveau 1), qui sont à leur tour dotés d'un rôle d'actant afin de former la structure syntactico-sémantique de l'énoncé à venir (niveau 2), qui ne sera effectif que lors d'un repérage énonciatif dont l'origine est l'énonciateur (niveau 3). Les opérations décrites par la

TOPE sont complexes et variées mais sont toutes centrées sur la notion de repérage : tout terme est repéré par rapport à un terme-repère.

Théorie des Rangs : Développée par Otto Jespersen, la Théorie des Rangs permet de rendre comptes des constituants syntaxiques en fonction de leurs propriétés combinatoires. Jespersen distingue trois rangs – rang 1, rang 2 et rang 3 : « Dans la combinaison *extremely hot weather* (‘temps extrêmement chaud’), le dernier mot, *weather* (‘temps’), qui représente évidemment l’idée principale, peut être désigné comme de rang 1 ; *hot* (‘chaud’), qui définit *weather*, sera alors de rang 2, et *extremely* (‘extrêmement’), qui définit *hot,* de rang 3. [...] Il est inutile de distinguer plus de trois rangs, car il n’existe aucun trait, formel ou autre, qui permette de distinguer ces mots d’un rang inférieur des mots de rang 3. » (Jespersen, 1971). Jespersen définit également deux types de construction, ‘jonction’ et ‘nexus’, qui correspondent respectivement au processus de nominalisation et à la formation d’une proposition.

Théorie du Gouvernement et du Liage : [Linguistique générative et transformationnelle] Développée principalement par Chomsky (1981), la Théorie du Gouvernement et du Liage se réfère à deux principes : le Gouvernement et le Liage. Le Gouvernement concerne l’assignation des cas à la structure syntaxique et la conservation de ces informations tout au long des transformations, conservation spécifiée par le principe de projection, au cœur de la théorie X-barre (Jackendoff, 1977). La théorie du Liage fixe les contraintes qui gèrent les relations d’anaphore – ce dernier terme ayant un sens très restreint en Grammaire Générative.

Théorie Standard Etendue : [Linguistique générative et transformationnelle] Version remaniée de la Théorie Standard, la Théorie Standard Etendue ou TSE rejette l’hypothèse des représentations logiques ou sémantiques assignées à la structure profonde de la syntaxe pour les attribuer à la structure de surface. En découle une série de modifications du traitement des transformations, notamment la contrainte de c-commande et les traces laissées en structure profonde lors du mouvement des constituants – mouvement qui permet aux constituants d’apparaître dans l’ordre observé en sortie.

Théorie Standard : [Linguistique générative et transformationnelle] Première version de la théorie de la Grammaire Générative et Transformationnelle, La Théorie Standard ou

TS (Chomsky, 1965) postule que la structure profonde détermine entièrement et *a priori* le sens d'une phrase. Cette théorie est peu à peu supplantée par la Théorie Standard Etendue qui propose que la structure de surface contribue, en partie tout du moins, à l'interprétation sémantique de la phrase.

Topique : [définition pragmatique] Partie de l'énoncé qui constitue l'information connue, souvent paraphrasé en 'ce dont il est question', le topique s'oppose au focus de l'énoncé. Le topique est aussi appelé 'thème'.

Topogénèse : [Psychomécanique] Opération de construction du sens lexical par l'actualisation du nom en discours. La topogénèse est effectuée par l'intermédiaire des déterminants.

Visée du discours : [Psychomécanique] Message que l'énonciateur a à transmettre : pour parler, il faut avoir quelque chose à dire.

Visée phrastique : [Psychomécanique] Volonté qu'a l'énonciateur de communiquer son message : pour parler, il faut avoir quelque chose à dire, mais il faut également vouloir le dire.

A

Accessibility Scale 113, 317, 339, 349
anadeixis.....112, 113, 164, 165, 166, 167
anaphore7, 8, 14, 27, 28, 29, 32, 34, 35, 36, 40, 42, 73, 76, 78, 81, 82, 112, 114, 115, 116, 118, 120, 121, 122, 124, 145, 156, 164, 165, 166, 167, 171, 182, 185, 192, 201, 202, 203, 208, 228, 231, 232, 233, 254, 261, 270, 280, 281, 296, 300, 317, 318, 319, 326, 338, 344, 349
anaphorique14, 27, 28, 36, 41, 42, 47, 48, 72, 76, 77, 112, 113, 114, 116, 118, 119, 122, 123, 143, 145, 156, 169, 171, 183, 185, 187, 191, 193, 194, 202, 208, 222, 227, 228, 231, 245, 246, 247, 254, 261, 263, 270, 280, 296, 301, 317, 318, 319, 323, 329, 330

C

cataphore34, 37, 73, 190, 200, 208, 212, 213, 247, 248, 252
contexte discursif103, 182, 189, 190, 192, 194, 199, 202, 247, 259, 322, 346, 347
contexte situationnel immédiat146, 159, 186, 189, 201, 208, 239, 246, 247, 270, 346, 347
coréférence27, 28, 29, 42, 59, 139, 156, 169, 171, 254, 293, 300, 323
co-texte.....39, 47, 101, 103, 155, 322, 323, 327

D

déclencheur d'antécédent112, 146, 157, 171, 189, 190, 191, 192, 196, 197, 198, 199, 201, 202, 220, 221, 238, 239, 241, 247, 254, 319, 323
déictique13, 34, 41, 42, 47, 48, 72, 76, 77, 110, 112, 114, 116, 119, 120, 123, 143, 145, 183, 187, 188, 189, 190, 191, 193, 194, 208, 213, 228, 231, 234, 246, 247, 261, 263, 267, 269, 270, 277, 280, 281, 296, 299, 301, 324, 325, 329
deixis7, 8, 14, 27, 34, 35, 41, 43, 48, 76, 78, 81, 82, 110, 112, 114, 115, 116, 118, 120, 121, 122, 124, 145, 146, 156, 164, 165, 166, 167, 171, 182, 185, 186, 187, 192, 201, 208, 210, 215, 228, 232, 233, 254, 270, 280, 300, 317, 324, 325, 327, 331, 344, 349
deixis *am phantasma*..... 112, 324
deixis de discours..... 325
deixis textuelle 112, 331
discours2, 7, 8, 9, 13, 15, 32, 34, 36, 37, 39, 42, 43, 51, 54, 69, 70, 73, 74, 75, 76, 77, 87, 92, 98, 101, 103, 104, 105, 107, 109, 111, 112, 114, 115, 116, 118, 119, 123, 127, 129, 137, 138, 139, 141, 148, 153, 154, 155, 156, 159, 162, 163, 170, 171, 182, 185, 186, 187, 189, 190, 191, 192, 193, 194, 196, 198, 202, 204, 206, 207, 208, 209, 213, 216, 218, 219, 220, 222, 227, 231, 232, 233, 234, 235, 238, 245, 248, 254, 256, 259, 261, 263, 266, 267, 268, 269, 270, 300, 315, 318, 319, 323, 324, 325, 326, 328, 331, 334, 335, 339, 343, 344, 345, 346, 347, 349, 351, 352

E

endophore14, 36, 40, 41, 72, 78, 81, 110, 111, 112, 115, 326, 327
exophore14, 34, 36, 41, 72, 78, 81, 110, 111, 112, 115, 326, 327

G

Givenness Hierarchy.....113, 114, 329, 349
GND129, 130, 131, 132, 133, 134, 137, 138, 140, 141, 142, 143, 146, 147, 185, 229, 230, 231, 232, 234, 235, 238, 240, 254, 269, 271, 272, 274, 275, 286, 288, 299, 301, 315, 316, 330, 331, 346, 347, 350
Grammaire Notionnelle Dépendancielle2, 6, 7, 8, 9, 97, 129, 133, 135, 137, 141, 146, 147, 182, 235, 300, 301, 330, 344, 345, 347

H

hypostatisation 96, 97, 260, 277, 302

I

indexicalité2, 82, 91, 98, 99, 101, 112, 115, 201, 213, 335, 343, 344

M

mémoire8, 41, 78, 101, 104, 115, 116, 117, 118, 120, 121, 122, 123, 124, 126, 127, 138, 156, 182, 185, 186, 187, 189, 190, 191, 192, 193, 194, 196, 197, 198, 199, 200, 202, 203, 208, 209, 213, 220, 222, 227, 228, 231, 233, 245, 246, 247, 248, 249, 250, 253, 259, 261, 263, 268, 269, 270, 280, 301, 318, 324, 332, 344, 345, 346, 347, 349, 350, 351

O

ordres d'entités.....8, 95, 97

R

référent13, 14, 42, 45, 48, 51, 52, 53, 54, 55, 75, 88, 91, 93, 97, 112, 113, 114, 116, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 138, 139, 142, 146, 156, 157, 164, 165, 166, 167, 185, 186, 187, 188, 189, 190, 191, 192, 194, 196, 197, 199, 200, 201, 202, 203, 207, 208, 209, 212, 213, 214, 217, 220, 224, 226, 227, 228, 230, 231, 233, 237, 238, 240, 241, 242, 243, 246, 247, 248, 250, 253, 255, 256, 259, 260, 261, 270, 273, 275, 280, 283, 293, 294, 301, 317, 324, 325, 329, 335

S

savoirs encyclopédiques et partagés..... 9

Table des matières

Chapitre I. Introduction générale et présentation des chapitres *7*

Chapitre II. Les démonstratifs en anglais : analyses structurale, générative et énonciatives *10*

1. L'analyse structurale des démonstratifs en anglais *10*
2. Le traitement des démonstratifs anglais dans un cadre génératif moderne *15*
   2.1. L'analyse des syntagmes dans la théorie X-barre classique *16*
   2.2. L'analyse de la phrase dans la théorie X-barre *20*
   2.3. Le déterminant comme tête du syntagme nominal *22*
   2.4. La Théorie du Liage *27*
3. Le traitement des démonstratifs anglais dans les théories énonciatives *30*
   3.1. La Grammaire Métaopérationnelle d'Adamczewski : un bref rappel théorique *30*
   3.2. La linguistique énonciative de Lapaire & Rotgé *32*
      3.2.1. Cadre théorique *32*
      3.2.2. *THIS* et *THAT* : une remise en question terminologique *33*
      3.2.3. *THIS* et *THAT* : des éléments anaphoriques *34*
         3.2.3.1. Le concept de 'submorphème' en linguistique énonciative *35*
         3.2.3.2. *THIS* et *THAT* : un microsystème *36*
   3.3. La Théorie des Opérations Prédicatives et Enonciatives de Culioli *38*
      3.3.1. Rappels théoriques *38*
      3.3.2. La forme schématique dans la TOPE *39*
      3.3.3. La forme schématique des démonstratifs *40*

Chapitre III. Les démonstratifs en français : analyses structurale, générative et énonciatives *44*

1. L'analyse structurale des démonstratifs français *46*
   1.1. Le rôle syntagmatique des démonstratifs *46*
   1.2. La structure morphologique des démonstratifs *49*
   1.3. L'opposition articles/pronoms personnels et démonstratifs *51*
   1.4. La distribution des formes pronominales *51*

1.4.1. *Celui, celle(s)* et *ceux* *52*
1.4.2. *Ce* *53*
1.4.3. *Ceci, cela* et *ça* *55*
2. L'analyse des démonstratifs français dans un cadre génératif *57*
2.1. L'analyse générative harrissienne : Gross (1977) *57*
2.2. L'analyse de Dubois & Dubois-Charlier (1970) *59*
2.3. Les analyses transformationnelles modernes *61*
2.4. L'analyse transformationnelle dans le cadre du Programme Minimaliste (Chomsky, 1995) *61*
3. Les analyses énonciatives des démonstratifs français *68*
3.1. Guillaume et la Psychosystématique du Langage (1919) *68*
3.1.1. Présentation générale *68*
3.1.2. La détermination en Psychomécanique *71*
3.1.3. Les pronoms démonstratifs en Psychomécanique *73*
3.2. Les pronoms chez Benveniste (1956) *75*
3.3. Les démonstratifs français dans la Théorie des Opérations Prédicatives et Enonciatives *76*
3.3.1. Rappels théoriques *76*
3.3.2. *Ce, cette, ces* dans la Théorie des Opérations Prédicatives et Enonciatives *77*

Bilan des chapitres II et III. Les ressemblances et les divergences entre démonstratifs anglais et français *79*

Chapitre IV. La référence et l'indexicalité *82*

1. La référence, entre langage et philosophie *83*
1.1. L'approche objectiviste *83*
1.2. L'approche constructiviste *86*
1.3. L'objectivité en termes d'intersubjectivité *87*
2. Précisions conceptuelles et terminologiques *92*
2.1. La classification ontologique des entités *92*
2.2. La 'référence', l''indexicalité', le 'texte' et le 'discours' *98*
2.3. Le 'contexte' *102*
2.4. Quelques théories discursivo-fonctionnelles *105*
2.4.1. La FDG (Functional Discourse Grammar) de Hengeveld & Mackenzie (2008) *106*

2.4.2. La SFG (Systemic Functional Grammar) de Halliday (1985) *109*
2.5. L'indexicalité : la deixis et l'anaphore..... *112*
3. Un candidat possible au chaînon manquant entre réalité et cognition : la mémoire *115*
4. Une théorie de la référence des indexicaux : un seul mode de fonctionnement pour plusieurs accès possibles *118*
4.1. Les éléments déictiques, les éléments indexicaux, la deixis et l'anaphore..... *118*
4.2. Les conséquences de la deixis comme opérant un retour sur le repère énonciatif..... *122*
4.2.1. Le fonctionnement référentiel des noms propres *122*
4.2.2. Le fonctionnement référentiel des expressions définies..... *123*
4.2.3. Le fonctionnement référentiel des expressions démonstratives *124*

Chapitre V. Représentation des expressions référentielles en Grammaire Notionnelle Dépendancielle..... *129*

1. La Grammaire Notionnelle Dépendancielle : une introduction..... *129*
2. Les catégories syntactico-sémantiques en Grammaire Notionnelle Dépendancielle *133*
2.1. Les catégories lexicales en Grammaire Notionnelle Dépendancielle..... *133*
2.2. Les catégories fonctionnelles en Grammaire Notionnelle Dépendancielle *135*
3. Les expressions référentielles en Grammaire Notionnelle Dépendancielle *137*
3.1. Les noms propres en Grammaire Notionnelle Dépendancielle *137*
3.2. Les syntagmes nominaux définis en Grammaire Notionnelle Dépendancielle *141*
3.3. Les syntagmes nominaux démonstratifs en Grammaire Notionnelle Dépendancielle *146*

Chapitre VI. Les corpus oraux en linguistique : avantages, limites et choix des données *149*

1. Quelle place accorder aux corpus en linguistique ? *149*
1.1. Définition du corpus *149*
1.2. Les enjeux épistémologiques de l'utilisation des corpus en linguistique *150*
1.3. Les avantages et les limites de l'utilisation des corpus en linguistique..... *150*
2. L'oral et l'écrit..... *151*
2.1. L'oral et l'écrit, deux facettes d'une même pièce..... *151*
2.2. Le genre et la visée communicative..... *153*
3. Structure du texte, structure du discours *156*

4. Les projets PFC et PAC *159*
4.1. La méthodologie des projets PFC et PAC *160*
4.1.1. Les points d'enquêtes et le choix des locuteurs *160*
4.1.2. Les tâches effectuées *160*
4.1.3. La méthode d'annotation *161*
4.1.3.1. Les conventions de transcription *161*
4.1.3.2. Le codage phonologique *162*
5. Les projets PFC et PAC 'au-delà de la phonologie' *163*
5.1. Comment construire un corpus voué à l'étude des démonstratifs à partir de corpus destinés aux phonologues ? *164*
5.1.1. Un essai d'annotation chiffrée 'à la PFC/PAC' *164*
5.1.2. Un essai de système de codage à attributs-valeurs *168*
5.1.3. Une plateforme open-source pour l'annotation linguistique multi-niveaux : MMAX2 ..... *168*
5.2. Les projets PFC et PAC et l'annotation syntaxique : le projet SAXO (Buscail & Eychenne) *170*
5.2.1. Quelques méthodes d'annotation existantes *170*
5.2.2. La plateforme SAXO : où en sommes-nous ? *172*
5.2.2.1. Quelques spécificités du français parlé difficiles à traiter de manière computationnelle *172*
5.2.2.2. Les étapes du développement de SAXO *173*
5.2.2.2.1. D'un sous-corpus-test à l'ensemble de la base de données PFC *174*
5.2.2.2.2. La segmentation par TreeTagger et l'adaptation de ce logiciel au corpus PFC *174*
5.2.2.2.3. La gestion de la structure syntaxique des énoncés *180*

Chapitre VII. Retour sur les démonstratifs anglais : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle *182*

1. Introduction : rappels et problématisation *182*
2. La référence indexicale par le recours aux contextes et l'activation de certains types de mémoire *185*
2.1. *This* *185*

2.1.1. La référence indexicale de *this* par l'invocation du contexte situationnel immédiat *186*
2.1.2. La référence indexicale de *this* par l'invocation du contexte discursif *189*
2.1.2.1. *This* en position de référence cataphorique *189*
2.1.2.2. *This* puise sa référence dans le discours en amont *190*
2.2. *That* *192*
2.2.1. La référence indexicale de *that* par l'invocation des contextes, d'une part, situationnel immédiat et, d'autre part, discursif *192*
2.2.2. La référence indexicale de *that* par l'invocation des savoirs partagés et encyclopédiques *195*
2.2.3. *That* et la génération d'un nouvel objet-de-discours *198*
2.3. *It* *201*
2.3.1. La référence indexicale de *it* par l'invocation du contexte situationnel immédiat *201*
2.3.2. La référence indexicale de *it* par l'invocation du contexte discursif *202*
2.3.2.1. *It* et le maintien référentiel en mémoire immédiate *202*
2.3.2.2. *It* en tant que support de prédication *204*
2.3.2.3. *It* et la référence au topique discursif *206*
2.4. La référence indexicale de *this, that* et *it* : bilan *208*
3. *This, that* et la notion de distance *210*
3.1. *This* et la proximité *210*
3.1.1. *This* et la proximité spatio-temporelle *210*
3.1.2. *This* et la proximité cognitive *211*
3.2. *That* et l'éloignement *215*
3.2.1. *That* et l'éloignement spatio-temporel *215*
3.2.2. *That* et l'éloignement cognitif *216*
3.3. La distance exprimée par *this* et *that* au niveau de l'(inter)communication *219*
4. *This, that* et *it* en tant qu'indices de co-construction du discours *222*
5. La structure interne de *this, that* et *it* et une représentation possible des pronoms démonstratifs anglais en GND *229*
5.1. La représentation de *this* et *that* en GND : rappels *229*
5.2. Une représentation possible de *this, that* et *it* en GND d'après leur structure interne *231*

Chapitre VIII. Retour sur les démonstratifs en français : référence indexicale des pronoms neutres, structure du discours et formalisation en Grammaire Notionnelle Dépendancielle *235*

1. Les pronoms démonstratifs français : rappels et précisions *235*
   1.1. *Celui-ci* et *celui-là*, *ceci* et *cela*, *ça* *235*
   1.2. *Ça* et *ce* *242*
   1.3. *Ça* : un indexical singulier au sein du système des démonstratifs français *243*
2. La référence indexicale de *ça* *245*
   2.1. Les contextes invoqués et les types de mémoire activés par *ça* *245*
      2.1.1. La référence indexicale de *ça* par le recours au contexte situationnel immédiat *246*
      2.1.2. La référence indexicale de *ça* par le recours au contexte discursif *247*
   2.2. Le rendu référentiel de *ça* : une production sémantique *251*
      2.2.1. *Ça* : un indexical relationnel *251*
      2.2.2. *Ça* : un indexical à élargissement référentiel *252*
      2.2.3. *Ça* : nominalisation et dynamisme référentiel *255*
      2.2.4. *Ça* : une référence indexicale qui résulte de deux paramètres *260*
3. Les fonctions (inter-)communicatives de *ça* *261*
   3.1. *Ça* et la fonction de résomption *262*
   3.2. *Ça* et la fonction de ratification *264*
   3.3. *Ça* et la fonction de repérage *267*
   3.4. Les fonctions (inter-)communicatives de *ça* : bilan *267*
4. La structure interne de *ça* et une représentation possible de ce pronom neutre en GND. *269*
5. Une représentation possible des pronoms démonstratifs en GND : bilan *272*

Chapitre IX. Perspectives de recherches *278*

1. Des perspectives concernant la prise en compte de la diachronie *278*
   1.1. Un bref historique du développement des démonstratifs en anglais, du Vème siècle à nos jours *278*
   1.2. Un bref historique du système des démonstratifs en français, du latin à nos jours ... *281*
      1.2.1. Du latin à l'ancien français (jusqu'au IX^{ème} siècle) *281*
      1.2.2. De l'ancien français au moyen français (IX^{ème}-XIV^{ème} siècle) *282*
      1.2.3. Du moyen français au français moderne (fin XIV^{ème}-aujourd'hui) *284*

1.3. Quelles pistes la perspective diachronique pourrait-elle nous offrir pour le traitement des pronoms démonstratifs en français ? *285*
2. Des perspectives liées à la prise en compte de la prosodie *288*
2.1. ToBI : un projet de conventions pour l'analyse des facteurs prosodiques *288*
2.2. Les accents mélodiques dans le système ToBI *289*
2.3. Des corrélations possibles entre accentuation mélodique et accès à la référence *292*
3. Et si nous nous intéressions aux autres langues du monde ? *296*

Chapitre X. Conclusion générale *300*

Table des figures

Figure 1 : « Par opération, nous entendons tout travail psychique et linguistique […] entrepris sur les signes tels que le lexique nous les propose, à des fins d'expression » (Lapaire & Rotgé, 1992 : 32) *33*

Figure 2 : Le microsystème *THIS/THAT* (Lapaire & Rotgé, 1992 : 63) *37*

Figure 3 : Le schème bi-tensif en Psychomécanique *70*

Figures 4 & 5 : Le cinétisme anti-extensif de l'article indéfini et le cinétisme extensif de l'article défini en Psychomécanique *72*

Figure 6 : 'Texte', 'contexte' et 'discours' : un système interdépendant *104*

Figure 7 : L'architecture générale de la FDG (Hengeveld & Mackenzie, 2008 : 13) *108*

Figure 8 : Les types de référence en SFG (Halliday & Hasan, 1976 : 33) *110*

Figure 9 : L'échelle de phoricité codée par certaines catégories d'expressions indexicales (Cornish, 2007 : 149) *113*

Figure 10 : L''échelle d'accessibilité des référents' (Accessibility Scale) d'Ariel (1988 : 84) *113*

Figure 11 : L''échelle des statuts cognitifs des référents, codés par certaines expressions linguistiques' (Givenness Hierarchy) de Gundel *et al*. (1993 : 275) *114*

Figure 12 : La mémoire comme transition entre réalité et cognition *116*

Figure 13 : La mémoire selon trois cercles concentriques *117*

Figure 14 : L'activation des indexicaux en discours *119*

Figure 15 : La deixis comme un mouvement rétroactif vers le repère énonciatif *120*

Figure 16 : L'anaphore comme un envoi référentiel unique vers la mémoire immédiate *121*

Figure 17 : Une construction référentielle en deux étapes pour les expressions démonstratives *125*

Figure 18 : L'accommodation d'une expression définie s'effectue dans la mémoire..... *126*

Figure 19 : La stratification schématique de la GND..... *130*

Figure 20 : La représentation schématique de la GND (Anderson, 2011 : 110)..... *130*

Figure 21 : La relation de complémentation en GND (Anderson, 2011 : 115)..... *131*

Figure 22 : La relation de modification en GND (Anderson, 2011 : 116)..... *131*

Figure 23 : La complexité grammaticale selon le genre : un continuum *154*

Figure 24 : L'implication interpersonnelle selon le genre : un continuum *155*

Figure 25 : Le fonctionnement indexical de *that* lors de l'activation des savoirs encyclopédiques ou partagés *197*

Figure 26 : Le fonctionnement indexical de *that* 'générateur de référence' *199*

Figure 27 : Une échelle descendante de référence entre *celui-ci/là* et *ça* *272*

Figure 28 : Un ton H* est lié à une montée mélodique s'il est placé au début de la syllabe proéminente, et à une descendante s'il se trouve à la fin (d'après Martins, 2008 : 3) *290*

Figure 29 : L'accent mélodique H* porté par la syllabe proéminente de *minimum* *290*

Figure 30 : Un ton L* est lié à une chute mélodique s'il est placé au début de la syllabe proéminente, et à une montée s'il se trouve à la fin (d'après Martins, 2008 : 3)..... *291*

Figure 31 : L'accent mélodique L* porté par la syllabe proéminente de *banana*..... *291*

Figure 32 : L'accent mélodique L+H* porté par la syllabe proéminente de *Amelia* *292*

Figure 33 : Des hypothèses sur le lien entre accentuation et référence des démonstratifs en anglais..... *295*

Table des tableaux

Tableau 1 : La classification des démonstratifs selon Dubois (1965 : 160).....*49*

Tableau 2 : Le système de Frege (1892) *85*

Tableau 3 : La structure du contexte (d'après Connolly, 2007)..... *105*

Tableau 4 : Les quatre principaux registres selon Biber (2002 : 13) *154*

Tableau 5 : Les catégories morphosyntaxiques du français proposées par TreeTagger *175*

Tableau 6 : Un extrait de conversation PFC étiqueté par TreeTagger *176*

Tableau 7 : Une illustration des étiquettes <TRUNC-X> et <GUESS-X>..... *177*

Tableau 8 : Les catégories morphosyntaxiques du français proposées par TreeTagger 'version PFC' *179*

Tableau 9 : Un extrait de conversation PFC étiqueté par TreeTagger 'version PFC'..... *179*

Tableau 10 : Les contextes invoqués et les types de mémoire activés par *this* et *that*..... *200*

Tableau 11 : Les contextes invoqués et les types de mémoire activés par *this*, *that* et *it* *209*

Tableau 12 : Les procédures discursives réalisées par *this*, *that* et *it*..... *209*

Tableau 13 : La sphère personnelle de *this* et la sphère interpersonnelle de *that* *222*

Tableau 14 : La sphère externe de *it* *227*

Tableau 15 : Les sphères discursives de *this*, *that* et *it*..... *227*

Tableau 16 : Les procédures indexicales réalisables par *this*, *that* et *it* en discours *232*

Tableau 17 : Les contextes invoqués et les types de mémoire activés par *this*, *that*, *it* et *ça* *250*

Tableau 18 : Les sphères discursives de *this*, *that*, *it* et *ça*..... *269*

Tableau 19 : Les procédures indexicales réalisables par *this*, *that*, *it* et *ça* en discours..... *270*

Tableau 20 : Le système des démonstratifs 'distaux' en vieil-anglais (V^{ème}-fin XI^{ème}) *278*

Tableau 21 : Le système des démonstratifs 'proximaux' en vieil-anglais (V^{ème}-fin XI^{ème}).. *278*

Tableau 22 : Le système des démonstratifs en moyen-anglais (fin XI^{ème}-XV^{ème}) *279*

Tableau 23 : L'évolution diachronique de *this*, *that*, *it* et *the* *280*

Tableau 24 : *CIST* et *CIL* jusqu'au XIII^{ème} siècle *282*

Tableau 25 : La série des déterminants neutres *ce/ces* au XIII^{ème} siècle..... *284*

Tableau 26 : Le système des démonstratifs en tongan (Churchward, 1953 : 150-152) *297*

Tableau 27 : Les marques de distance des déterminants démonstratifs dans les différentes langues du monde (Diessel, 2011) *297*

Tableau 28 : Le système des déterminants démonstratifs en hunzib, en japonais et en hausa (respectivement, Van den Berg, 1995 : 61 ; Kuno, 1973 : 27 ; Wolff, 1993 : 119-120) *298*