HAL open science

# Utilisation des informations sémantiques en traduction automatique

Dominique Augereau

## To cite this version:

Dominique Augereau. Utilisation des informations sémantiques en traduction automatique. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1962. Français. NNT : . tel-00278554

**HAL Id: tel-00278554**

**https://theses.hal.science/tel-00278554v1**

Submitted on 13 May 2008

**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.

N° d'ordre :

# THÈSE

présentée à

LA FACULTÉ DES SCIENCES DE L'UNIVERSITÉ DE GRENOBLE

pour obtenir

LE TITRE DE DOCTEUR DE TROISIÈME CYCLE

Mathématiques Appliquées

---

par

Dominique AUGEREAU

Ingénieur I. E. G., I. M. A. G.

Licencié ès Sciences

---

# UTILISATION DES INFORMATIONS SEMANTIQUES EN TRADUCTION AUTOMATIQUE

*Thèse soutenue le 6 Octobre 1962, devant la Commission d'examen :*

Monsieur J. KUNTZMANN, Président

Messieurs B. VAUQUOIS, Examinateurs

N. GASTINEL

T H E S E

présentée à

LA FACULTE DES SCIENCES DE L'UNIVERSITE DE GRENOBLE

pour obtenir

le titre de Docteur de troisième cycle

Mathématiques Appliquées

par

Dominique AUGEREAU

Ingénieur I.E.G.-I.M.A.G.
Licencié ès sciences

# UTILISATION DES INFORMATIONS SEMANTIQUES EN TRADUCTION AUTOMATIQUE

Thèse soutenue le 6 octobre 1962 devant la Commission d'examen :

MM. J. KUNTZMANN Président
B. VAUQUOIS
N. GASTINEL Examinateurs

FACULTE DES SCIENCES

DE L'UNIVERSITE DE GRENOBLE

-:-:-

Doyens honoraires :

M.FORTRAT P.
M.MORET L. , Membre de l'Institut

Doyen :

M.WEIL L.

Professeurs :

MM.NEEL L, Membre de l'Institut, Physique expérimentale
MOREL L, Géologie et minéralogie
WOLFERS F., Physique
DORIER A,Zoologie
HEILMANN R, Chimie organique
KRAVTCHENKO J, Mécanique rationnelle
PARDE M, Potamologie
BENOIT J, Radioélectricité
CHENE M, Chimie papetière
NOBECOURT P, Micrographie papetière
BESSON J, Chimie
WEIL L, Thermodynamique
FELICI N, Electrostatique
KUNTZMANN J, Mathématiques Appliquées
BARBIER R, Géologie Appliquée
SANTON L, Mécanique des fluides
CHABAUTY C, Calcul différentiel et intégral
OZENDA P, Botanique
FALLOT M, Physique industrielle
GOLVANI O, Mathématiques
MOUSSA A, Chimie nucléaire et radioactivité
TRAYNARD P, Chimie générale
CRAYA A, Hydrodynamique
SOUTIF M, Physique générale
REEB G, Statistiques mathématiques
REULOS R, Théorie des champs
AYANT Y, Physique approfondie
GALLISSOT F, Mathématiques pures
Mlle LUTZ E, Mathématiques générales
MM.BLAMBERT M, Mathématiques
BOUCHEZ R, Physique nucléaire
LLIBOUTRY R, Géophysique
MICHEL R, Géologie et minéralogie
BONNIER E, Electrochimie

Professeurs sans chaire :

MM. SILBER R, Mécanique des fluides
DUSSAUX G, Physiologie animale
MOUSSIGNT J, Electronique
PILLET J, Electrotechnique
BARBIER J.C, Physique
BUYLE-BODIN M, Electronique
PAUTHENET R, Electrotechnique
Mme KOFLER L, Botanique

Maîtres de Conférences :

MM. VAILLANT F, Zoologie et hydrobiologie
DREYFUS B, Thermodynamique
Mlle NAIM L, Mathématiques
MM. PERRET R, Servomécanismes
ARNAUD P, Chimie
Mme BARBIER M.J, Electrochimie
MM. BRISSONNEAU P, Physique
COHEN J, Physique
DEBELMAS J, Géologie et minéralogie
Mme SOUTIF J, Physique
MM. VAUQUOIS B, Mathématiques Appliquées
DEPASSEL R, Mécanique des fluides
GERBER R, Mathématiques
ROBERT A, Chimie papetière
ANGLES D'AURIAC, Mécanique des fluides
BIARNEZ, Mécanique physique
COUMES A, Electronique
DODU J, Mécanique des fluides
DUCROS P, Minéralogie et cristallographie
GIDON P, Géologie et minéralogie
GLENAT R, Chimie
HACQUES G, Calcul numérique
LANCIA R, Physique automatique
PEBAY-PEROULA, Physique
GASTINEL, Chargé d'enseignement, Mathémtiques appliquées
LACAZE A, Chargé d'enseignement, Thermodynamique

-:-:-

Je tiens à exprimer ma profonde reconnaissance :

à Monsieur le Professeur VAUQUOIS , sous la direction duquel ce travail a été effectué, pour les conseils judicieux qu'il m'a donnés au cours de cette année et pour tout le temps qu'il a consacré à mon étude ;

à Monsieur le Professeur KUNTZMANN , Directeur du Laboratoire de Calcul de l'Université de Grenoble, pour l'honneur qu'il m'a fait en acceptant de présider le Jury de thèse ;

à Monsieur GASTINEL , Maître de Conférences, qui a bien voulu en faire partie.

Je remercie également les membres du Centre d'Etudes pour la Traduction Automatique de Grenoble, et notamment Mademoiselle D. ANDRON dont la collaboration me fut précieuse.

J'exprime toute ma gratitude à Mademoiselle M. BOUVIER qui a assuré la réalisation matérielle de cette thèse.

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

En toute langue, chaque phrase est construite à partir de la pensée qu'elle représente. Quand nous traduisons nous-mêmes nous comprenons d'abord le texte. De même en traduction mécanique le passage idéal d'une langue à une autre devrait se fonder sur la nature sémantique du texte à traiter. Dans ses travaux, le Professeur CECCATO [1] a montré que dès la lecture d'une phrase on pense, c'est-à-dire que l'on fait de la sémantique. Ceci nous explique pourquoi un enfant comprend immédiatement des phrases ambigues du type suivant :

"Il vit une directrice de société très enrhumée".

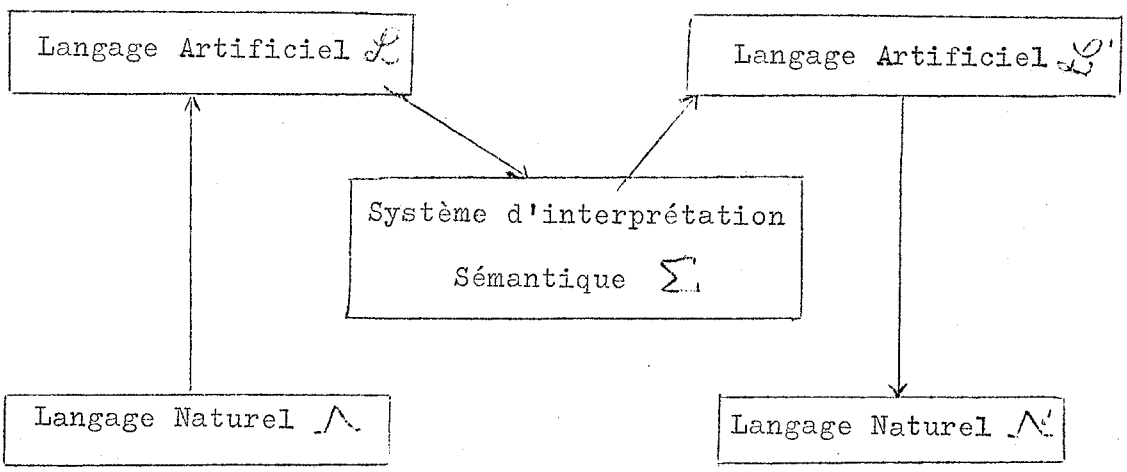
La grammaire est née après la langue et ceci nous convainc de l'importance primordiale de la sémantique.

Avec le Professeur VAUQUOIS [2] , nous pouvons envisager la Traduction Automatique comme le passage entre deux langages artificiels $\mathcal{L}$ et $\mathcal{L}'$. $\mathcal{L}$ est un modèle de la langue naturelle source $\Lambda$ . $\mathcal{L}'$ représentant la langue cible $\Lambda'$. $\mathcal{L}$ et $\mathcal{L}'$ sont des systèmes formels. Il nous faudra résoudre trois problèmes distincts :

-1°) Construction du langage Artificiel $\mathcal{L}$.
-2°) Construction du langage Artificiel $\mathcal{L}'$
-3°) Transfert $\mathcal{L} \longrightarrow \mathcal{L}'$

Mais le transfert $\mathcal{L} \rightarrow \mathcal{L}'$ se fera en passant par un Système d'interprétation Sémantique $\Sigma$ qui sélectionnera les élément valables.

Nous avons donc le schémas suivant :




Ce besoin de la sémantique apparait sous de nombreuses formes. En particulier, les schémas représentant les structures syntaxiques peuvent présenter des ambiguïtés si l'on n'a pas de support sémantique et de même, pour recomposer la phrase à partir d'éléments donnés, les critères syntaxiques seront souvent insuffisants.

L'algorithme permettant de passer de $\mathcal{L}$ à $\mathcal{L}'$ se fractionne en plusieurs étapes :

Analyse syntaxique - Analyse sémantique - Transfert de structures

mais préalablement, on devra définir les langages artificiels $\mathcal{L}$ et $\mathcal{L}'$

Dans les grammaires structurales classiques on discerne avec assez de facilité la limite entre la morphologie et la syntaxe-sémantique. Par contre, pour celle existant entre la syntaxe et la sémantique c'est tout à fait différent, car elle n'apparait pas clairement. Voici donc une difficulté supplémentaire et pourtant pour $\mathcal{L}$ et $\mathcal{L}'$ les deux niveaux seront marqués par des règles de construction de types différents.

Dans cette thèse, nous nous proposons :

1°) D'examiner les différents types de difficultés rencontrées au cours de l'Analyse syntaxique

2°) D'étudier sur le plan sémantique la constitution des langages Artificiels formalisés

3°) D'ébaucher l'Analyse sémantique du Transfert $\mathcal{L} \longrightarrow \mathcal{L}'$ et de montrer son aptitude à résoudre les difficultés envisagées précédemment.

---

- I -

# RESULTAT DE L'ETUDE SYNTAXIQUE-SES LIMITES

## I-. DEFINITION RENCONTREES EN MORPHOLOGIE ET EN SYNTAXE

Au cours de l'étude morphologique et syntaxique, on a été amené à définir des valeurs qui sont utilisées dans la formulation des codes morphologique et syntaxique et dans les tables représentant les diverses fonctions syntaxiques.

Ces valeurs et leurs symboles sont les suivants :

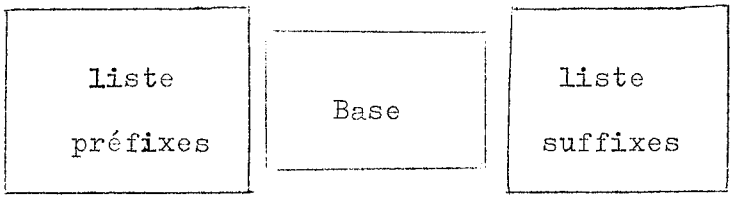
Nous ajoutons pour chacune de ces valeurs un exemple

| | | |
|---|---|---|
| - Unité lexicale | : $U_l$ | Chien couleur |
| - catégorie syntaxique | : $K_{u,v}$ | $K_{1,1}$ Substantif commun<br>$K_{4,8}$ Verbe à l'infinitif |
| -variable grammaticale | : $V_g$ | $V_g$ permanente : genre (masculin féminin) |
| - relation syntaxique | : $\varphi_p$ | liaison (sujet)-(verbe...) |
| - relation entre $K_u$ | : $\psi_q$ | liaison subst. commun-subst. commun |

Nous allons préciser ces diverses valeurs :

- Pour les $V_g$ et les $K_{u,v}$, nous conservons les définitions données par B. VAUQUOIS, J. VEYRUNES [3] et G. VEILLON [4.]

-Nous appelons Unité lexicale (cf. 2 page 38) l'ensemble de toutes les formes qui appartiennent à une certaine classe morphologique et que l'on obtient en prenant une base et en y ajoutant une liste de préfixes et d'affixes. Une Ul peut se représenter par le schéma suivant :




En langue russe, les affixes sont de trois sortes :

suffixe, désinence et suffixe de réflexivité.

Les relations $\Psi_q$ et $\varphi_p$ ont été introduites par A.AUROUX [5] $\varphi_p$ en particulier est un opérateur qui s'applique à un couple de syntagmes élémentaires $S_i$ , $S_j$ et lui fait correspondre un élément $S_k$ de même nature que $S_i$ et $S_j$ et que l'on appelle "Syntagme".

Nous pouvons rappeler quelques définitions d'éléments du langage artificiel : celles du terme, du terme ordonné et du syntagme élémentaire <s.e.> (cf [2] , page 23)

Nous avons :

<terme> ::= <forme> <code syntaxique> <code sémantique>
<terme ordonné> ::= <$\nu$> <N° Ul> <code syntaxique> <code sémantique>

Le syntagme élémentaire : c'est le terme ordonné privé de son code sémantique d'où :

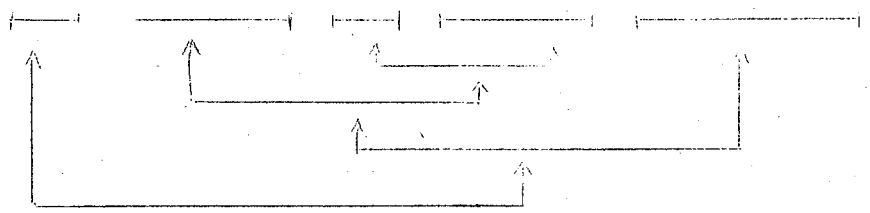
<s.e.> ::= <$\nu$> <N° Ul> <$K_{u,v}$> <CΔ> <$V_{gp}$> <$V_{gc}$> <$C_g$> <$C_d$>

où les symboles ont la signification suivante :

$\nu$ : numéro séquentiel
N° Ul : Numéro d'unité lexicale
$K_{u,v}$ : catégorie syntaxique
CΔ : code de dérivation
$V_{gp}$ : variable grammaticale permanente
$V_{gc}$ : variable grammaticale contingente
$C_g$ : code de gouvernement
$C_d$ : code de dépendance.

## II-. BUT DE L'ETUDE SYNTAXIQUE

Dans l'étude syntaxique, on cherche la structure d'une phrase, par combinaison binaire de mots ou de groupes de mots d'après des schémas syntaxiques semblables au suivant :



Le schéma syntaxique est une simple représentation graphique des liaisons syntaxiques existant entre les différents mots de la phrase. On doit bien le différencier de la structure syntaxique de la phrase qui est formée par le schéma précédent et par l'ensemble des relations $\varphi_p$ qui y sont utilisées à des places bien précises. Voir les exemples du chapitre V (V).

Pour une phrase, l'analyse syntaxique va donner naissance à toutes les structures valables du point de vue des règles de construction formelle et ceci sans faire appel à la signification des éléments qui entrent en jeu.

L'étude syntaxique donne donc une ou plusieurs structures syntaxiques.

On veut pouvoir traduire cette phrase de la langue source en une phrase de la langue cible ; c'est-à-dire : on veut obtenir une nouvelle structure syntaxique appliquée à des mots de la langue cible.

Pour atteindre ce but, on utilise ;

- la "structure syntaxique" précédente
- une "table de transfert syntaxique" permettant le passage des relations syntaxiques du langage source aux relations syntaxiques du langage cible.
- un "dictionnaire" donnant l'équivalence en langue cible des mots de la phrase exprimés en langue source.

Si nous étudions pour langue source, une langue riche en flexion, le dictionnaire sera formé en admettant que l'analyse morphologique se fasse suivant la méthode algorithmique exposée par G. VEILLON [4] . C'est le cas de la langue russe.

Alors le dictionnaire est considérablement réduit et il participe à la création dynamique des syntagmes élémentaires.

Au cours de la recherche de la nouvelle structure syntaxique, nous rencontrons bien des difficultés, les principales sont les suivantes:

1) Le schéma syntaxique vu précédemment est rarement unique. L'étude syntaxique aura généralement donné de la phrase en langue source, plusieurs schémas valables.

Et même au cas où il n'existerait qu'un seul schéma en langue source, la "table de transfert syntaxique" donnera souvent à partir de celui-ci, plusieurs schémas syntaxiques valables en langue cible.

Nous classons ces difficultés sous la rubrique : "Pluralité des Constructions syntaxiques".

2°) Le dictionnaire ne donnera pas toujours d'un mot de la langue source, un et un seul équivalent.

Tout d'abord, ce mot n'a peut-être pas d'équivalent direct dans la langue cible.

C'est le cas des mots composés.

Ce type de difficultés sera vu dans la rubrique : "mots composés".

D'autre part, un mot de la langue source, peut avoir plusieurs représentations dans la langue cible.

Nous verrons ce type de difficultés dans la rubrique : "pluralité des mots cible ".

## III-. ETUDE DES DIFFICULTES RENCONTREES

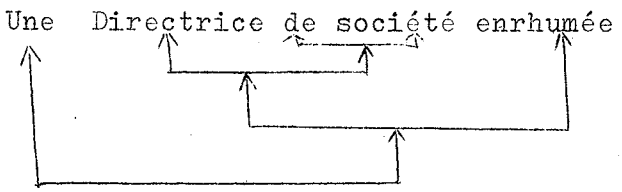
Nous allons les représenter sous forme d'un tableau puis les commenter. (Voir page 8 bis)

### 1°) Pluralité des constructions syntaxiques

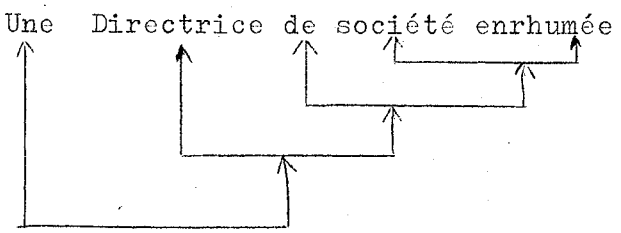
a) Il existe plusieurs schémas syntaxiques possibles en langue source et la syntaxe seule, ne permet pas de faire le choix déterminant le schéma syntaxique sémantiquement valable.

Exemple N° 1

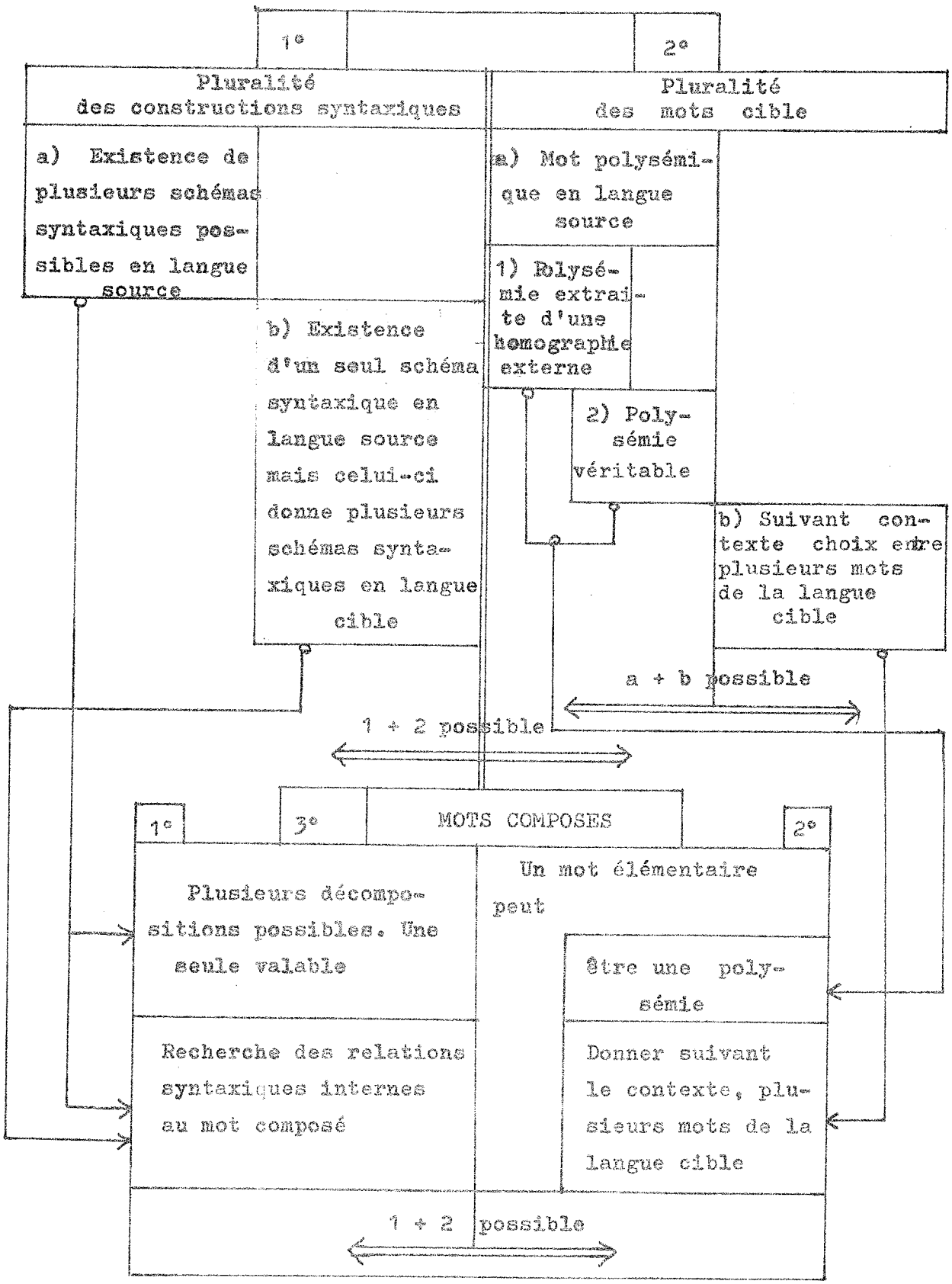
- Schéma 1
- Vrai




- Schéma 2
- Faux

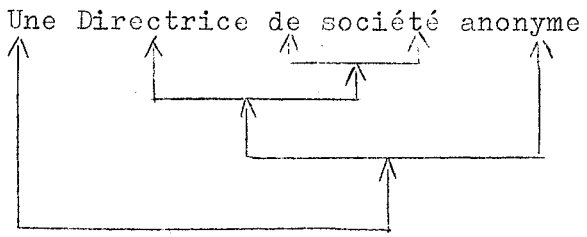


DIFFICULTES

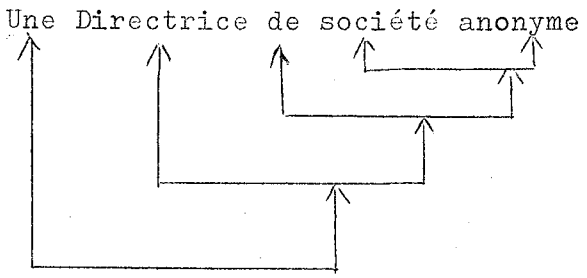


Exemple N° 2

- Schéma 1
- Faux




- Schéma 2
- Vrai




D'après les exemples N° 1 et N° 2 on voit que :

1- Au moins deux schémas syntaxiques sont valables
2- La syntaxe est bien incapable de définir celui qui est sémantiquement valable

Le schéma 1 est sémantiquement valable pour l'exemple N°1 mais ne l'est pas pour l'exemple N° 2

Et inversement, le schéma 2 est sémantiquement valable pour l'exemple N°2 mais ne l'est pas pour l'exemple N° 1.

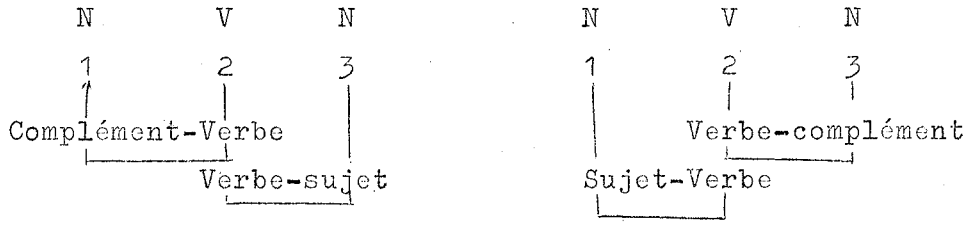
Un cas particulier qui sera très difficile à résoudre, peut être illustré par l'exemple qui suit :

On a la suite des trois mots suivants :

$Nom_1$ Verbe $Nom_3$

et il est impossible d'une façon syntaxique de trouver le nom sujet

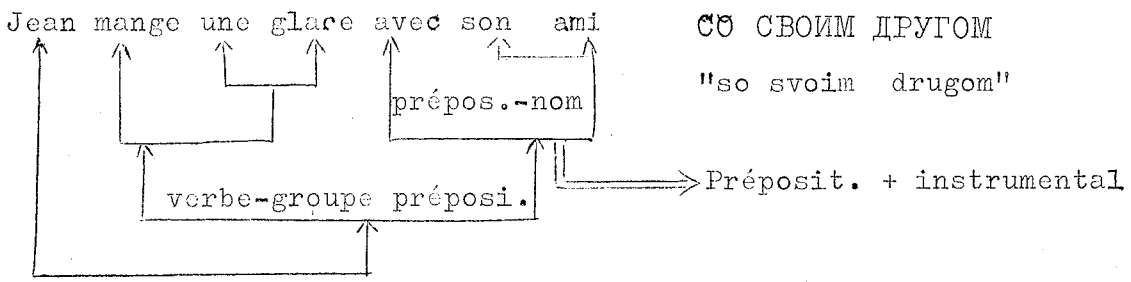
et le nom complément. On obtient alors, deux schémas possibles :

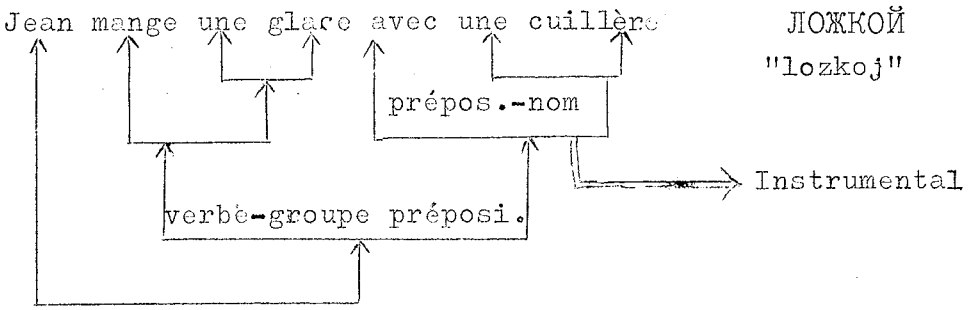



Si on admet que l'inversion du sujet est syntaxiquement possible, on est en présence de deux schémas syntaxiques valables.

b) Il n'existe qu'un seul schéma syntaxique en langue source mais celui-ci donne naissance à plusieurs schémas syntaxiques en langue cible.

Exemple 1




Ainsi la construction syntaxique Préposition-Nom peut donner en russe les deux constructions syntaxiques suivantes :

{ Préposition + Instrumental
{ Instrumental

Exemple 2

a) Jean travaille avec son ami

(Inst.) СО СВОИМ ДРУГОМ "so svoim dr u gom" - en russe
with his friend - en anglais

b) Jean travaille avec ardeur

(Inst.) С УСЕРДИЕМ "s userd'em" - en russe
eagerly - en anglais.

Ici en français et en russe, les deux phrases a et b présentent le même schéma syntaxique. Et celui-ci donne naissance à deux schémas différents en anglais pour a et pour b.

Les constructions syntaxiques Français : préposition-nom
russe : préposition-instrumental

peuvent donner les deux constructions syntaxiques anglaises :

- préposition - nom
- adverbe

Il y a deux raisons principales pour qu'à un mot "source" corresponde plus d'un mot "cible".

1- Les mots sont polysémiques en langue source : c'est-à-dire que les occurrences peuvent désigner plus d'un chose. Nous sommes en présence soit d'une polysémie soit d'une homographie externe.

Souvent dans certaines langues, comme le russe, ce n'est le cas que du mot écrit ; tandis que les différentes désignations sont séparées en langage parlé au moyen d'une accentuation différente. Nous avons ici, des homographes externes.

Ceux-ci appartiennent à deux unités lexicales distinctes et différent aussi par une variable grammaticale permanente.

Exemple : La 3è personne du singulier СТОИТ "stoit" peut provenir de deux infinitifs : СТОИТЬ "stoit" (coûter) et СТОЯТЬ sto'at' (se tenir debout). Quand on le prononce, une accentuation différente marque les deux significations qui ne peuvent pas être discriminées quand le mot est écrit.

Nous pouvons donc considérer deux types de polysémies:

1°) Une polysémie extraite d'une homographie externe

En langue française, un exemple nous est fourni par le mot "manche". Ce n'est pas une polysémie véritable, car le genre (vgp) varie ; mais ce genre peut-être indéterminé : "les manches" est une polysémie. Et si l'on écrit : "les manches du veston ", comment repérer s'il s'agit de l'Ul à genre masculin ou de l'Ul à genre féminin ?

Ce type de polysémie est rencontré, toutes les fois que l'étude syntaxique est incapable de déterminer quel est le syntagme élémentaire valable.

2°) La polysémie véritable

Ici nous sommes en présence d'un phénomène différent. L'analyse syntaxique a bien déterminé le syntagme élémentaire valable puisque nous n'avons qu'une unité lexicale. Mais celle-ci représente plusieurs concepts sémantiques et en vue de la traduction future en langue cible, il nous faut déterminer quel concept est ici représenté.

Exemple : A l'Ul "pile" correspond cinq concepts sémantiques

pile →
- côté d'une pièce de monnaie
- amas de choses
- générateur d'électricité
- générateur d'énergie nucléaire
- massif de maçonnerie

De toutes façons, nous remarquons que dans les deux cas de polysémie, l'étude que l'on doit faire pour résoudre le problème se déroulera dans la seule langue source et indépendamment de la langue cible choisie.

La seconde raison fait intervenir la précision avec laquelle chaque mot de la langue est défini.

Ainsi les mots russes, et en particulier les verbes, expriment une situation avec quelques détails. Le verbe de la langue cible qui lui correspond contient généralement plus de détails que le verbe russe. Celui-ci a donc besoin d'être enrichi.

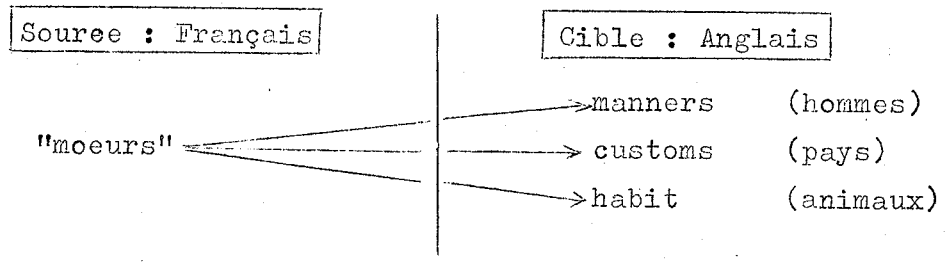
Exemple : ИДТИ "idti" se traduira en français soit par "aller"
soit par "venir"

Dans ИДТИ il y a une idée de mouvement mais non de direction ; tandis que ces deux idées sont rencontrées dans les verbes français.

Ainsi un mot de la langue source, non polysémique donne naissance suivant le contexte à plusieurs mots de la langue cible.

Ici on a une unité lexicale et un concept sémantique

Exemple :




Nous voyons ici apparaître une hiérarchie dans les concepts sémantiques Nous préciserons celle-ci au chapitre IV (II).

Le concept sémantique de "moeurs" est : "habitudes naturelles ou acquises". Nous remarquons qu'ici nous faisons intervenir la langue cible et ceci contrairement au cas précédent.

De plus, une autre différence vient du fait qu'en langue source, à une unité lexicale correspond dans le premièr cas : plusieurs concepts sémantiques et dans le second cas : un seul concept sémantique.

Bien sûr, les deux raisons qui font qu'à un mot "source" correspond plus d'un mot "cible" peuvent se trouver conjuguées.

On fait alors apparaître distinctement ces deux raisons.

Exemple :

A l'U1 "glace" correspond trois concepts sémantiques

| | source | cible |
|---|---|---|
| glace 1 | eau congelée | ice |
| glace 2 | crême aromatisée | ice-cream |
| glace 3 | lame de verre | mirror / window |

Si nous prenons l'anglais comme langue cible, le concept sémantique N° 3 peut être représenté par deux mots "cible" distincts.

Ici nous avons bien une polysémie compliquée par une imprécision de la langue source par rapport à la langue cible.

3°) Mots composés

Ils se classent en deux groupes :

1- Le premier groupe comprend les mots composés qui usités depuis longtemps ont vu leur sens originel modifié et se trouvent donc dans le dictionnaire comme formant une nouvelle unité lexicale

Ce premier groupe est formé d'un nombre limité de mots : c'est un groupe fermé.

2 - Le second groupe est un ensemble ouvert de mots. Leur étude fait appel aux deux notions vues précédemment : pluralités des mots "cible" - pluralité des constructions syntaxiques.

a) En effet, il existe en général, plusieurs décompositions possibles mais une seule valable.

Exemple : kulturinfiltrierung : mot composé allemand

kult/urin/filtrierung , kult/ur/infiltrierung , kultur/infiltrierung
culte filtration culte infiltration culture infiltration
urine primitive

Il faudra donc trouver la décomposition sémantiquement valable. Mais même en supposant que celle-ci soit trouvée, (le cas le plus simple étant celui où elle est unique) il faudra découvrir aussi les relations syntaxiques existant entre les mots élémentaires qui forment ce mot composé.

En effet celui-ci pourra être traduit en langue cible

- soit par un ou deux mots simples

Exemple rund/sicht/wind/schutz/scheibe pare-brise panoramique

- soit par une véritable phrase qui nécessitera la connaissance d'une structure syntaxique très précise.

b) On peut rencontrer un mot élémentaire qui ait une ou les deux propriétés suivantes :

- être une polysémie (Ex. :pile)
- donner naissance suivant le contexte à plusieurs mots de la langue cible.

## CONCLUSION

Nous venons de mettre en évidence des problèmes auxquels la seule syntaxe est incapable de donner une solution.

Nous allons donc introduire un autre facteur : la sémantique qui va apporter une contribution à l'étude de ces différents problèmes.

- II -

## SOUS QUELLE FORME INTRODUIRE LA SEMANTIQUE ?

En nous inspirant des valeurs syntaxiques vues au chapitre I, nous avons été amenés à considérer l'existence des valeurs sémantiques suivantes :

Nous ajoutons pour chacune de ces valeurs, son symbole :

I - unité sémantique : $U_s$
II- champ sémantique : Ch
III- variable sémantique : $V_s$
IV- relation sémantique : $\varphi'_p$
V - relation entre champs : $\psi'_q$

Nous allons préciser ces différentes valeurs et en ajouter une autre que les études sémantiques nous ont conduits à introduire, celle de :

VI - élément notionel : $E_n$

### I-. UNITE SEMANTIQUE

Prenons le mot français "racine" et voyons par quels mots il serait traduit en anglais et en russe. On obtient le tmbleau suivant :

| Cible | source | cible |
|---|---|---|
| Anglais ← | Français | → russe |
| root | - racine (d'une plante ) | KOPEHb |
| root | - racine ('un cheveu) | --- |
| root | - racine (d'une équation) | --- |
| root | - racine (d'un nombre) | --- |
| stem | - racine (d'un mot) | --- |

Le mot "racine" est une unité lexicale. Comment définir les unités sémantiques qui lui sont attachées ?

Considérons d'abord en français les différents concepts attachés au mot "racine".

Nous en trouvons quatre :

1- racine : Partie par laquelle un organe est implanté
2- racine : élément irréductible des mots
3- racine : expression algébrique qui élevée à une puissance reproduit l'expression proposée.
4- racine : valeur de la quantité satisfaisant une équation.

Reprenons la partie centrale du tableau précédent : nous pouvons en face de chaque mot "racine" mettre le concept qui lui correspond.

| | |
|---|---|
| 1 | - racine (d'une plante) |
| 1 | - racine (d'un cheveu) |
| 3 | - racine (d'une équation) |
| 4 | - racine (d'un nombre) |
| 2 | - racine (d'un mot) |

Nous voyons que plusieurs concepts peuvent être représentés par le même mot en langue cible.

Exemple :
"root" représente les concepts 1, 3 et 4 de "racine"

de même un même concept peut être représenté suivant le contexte, par plusieurs mots différents en langue cible.

Exemple :
le mot "moeurs" possède un seul concept : "habitudes naturelles ou acquises" ; et si on le traduit en anglais, on a le choix entre trois mots "cible" :

| | | |
|---|---|---|
| | manners | (gens) |
| moeurs | customs | (pays) |
| | habits | (animaux) |

Si l'on envisageait la traduction d'un mot de la langue source (Ex. : le français) en toutes les autres langues cible (Ex. : Anglais, russe ...) on pourrait prendre pour ensemble d'unités sémantiques la réunion des concepts sémantiques attachés aux mots de traduction qui existent dans chaque langue cible. On pourrait faire correspondre une unité sémantique à chacun de ces concepts.

De cette façon, il serait possible de définir une "langue intermédiaire" qui permettrait de passer de la langue source à n'importe quelle langue cible. Mais à ce moment là un concept pourrait être représenté par plusieurs $U_s$.

Or, ceci présenterait un gros inconvénient : il empêcherait l'utilisation d'un langage imagé, celui-ci étant lié à un concept déterminé du mot source.

Pour la même raison, on ne peut déterminer les $U_s$ en considérant les différents champs où le mot est utilisé. En effet, ce mot peut être utilisé dans des champs tout à fait différents, mais avec le même concept. De plus dans le même champ, le mot peut être employé avec deux concepts différents :

Exemple : "racine" en mathématiques $\Longrightarrow$ concepts 3 et 4 de "racine"

Nous allons donc définir une unité sémantique de la façon suivante :

Soit une unité lexicale, considérons les différents concepts qui lui sont attachés ; à chaque concept nous faisons correspondre une unité sémantique.

Exemple :

1- 1 $U_l$ ⟸ "racine" ⟹ 4 $U_s$ (définies précédemment)

2- Soit l'unité lexicale "glace"

Il lui correspond trois unités sémantiques définies par les concepts suivants :

1) eau congelée
2) Rafraichissement composé de crème aromatisée et congelée
3) lame de verre.

## II-. CHAMP SEMANTIQUE

Le but recherché était de répartir l'ensemble des mots en un certain nombre de classes. A cette fin, nous avons utilisé un classement déjà employé dans les bibliothèques, à savoir, la classification décimale universelle.

Nous avons adopté celle-ci pour n'en conserver que mille subdivisions principales.

Nous avons ainsi obtenu mille champs sémantiques numérotés de 0 à 999.

Une unité sémantique appartiendra à un ou à plusieurs champs. Ainsi l'$U_s$ de "racine" : "valeur de la quantité satisfaisant à une équation" appartiendra en particulier aux champs suivants : Algèbre, trigonométrie, analyse, astronomie théorique, mécanique rationelle, analyse qualitative chimique ....

Chaque champ sera défini par trois chiffres que nous appellerons CDU (centaine, dizaine, unité) chacune des trois lettres C, D, U peut prendre les dix valeurs possibles : 0, 1, ... 8, 9. En annexe N° 1 nous donnons les champs dont le code commence par le chiffre 5.

Comment fabriquer l'ensemble des mots appartenant à un champ ?

Nous prendrons des textes traitant d'un sujet précis et faisant partie d'un champ bien déterminé, nous y enlèverons les exemples ou considérations n'ayant pas de rapport normal avec le champ en question et nous inclurons les mots restants, à l'intérieur du champ que nous cherchons à former.

A l'intérieur de chaque champ, nous classerons ces mots par ordre alphabétique et par l'étude des textes nouveaux nous enrichirons peu à peu l'ensemble des mots appartenant au champ.

## III-. VARIABLE SEMANTIQUE

A l'intérieur de chaque champ, nous considérons des notions qui interviennent fréquemment dans le sens des mots appartenant à ce champ. Pour cela nous utiliserons l'ensemble des mots qui appartiennent à chaque champ dont nous venons de voir la formation. Chacune de ces notions est une variable sémantique. Pour l'instant, nous pensons qu'au plus 32 variables sémantiques seront nécessaires pour représenter d'une façon précise les notions essentielles d'un champ. Ces $V_s$ seront numérotées de 0 à 31.

Chacune de ces variables sémantiques peut prendre des valeurs particulières qui précisent la notion qu'elle représente. Une $V_s$ comportera au maximum huit valeurs distinctes ; celles-ci seront numérotées de 0 à 7.

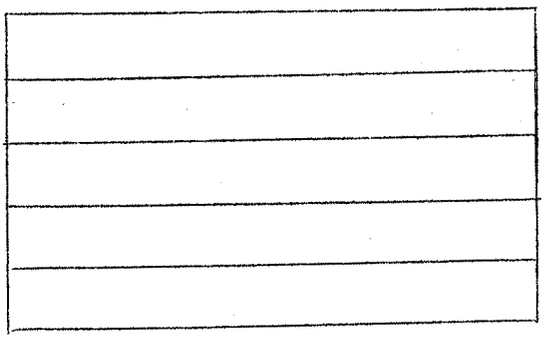
Un exemple de variables sémantiques et des valeurs de celles-ci nous est donné par l'étude du champ 512 représentant "l'Algèbre" : cf. page suivante. Pour l'instant nous y avons dénombré 14 $V_s$ .

# CHAMP 512 : ALGEBRE

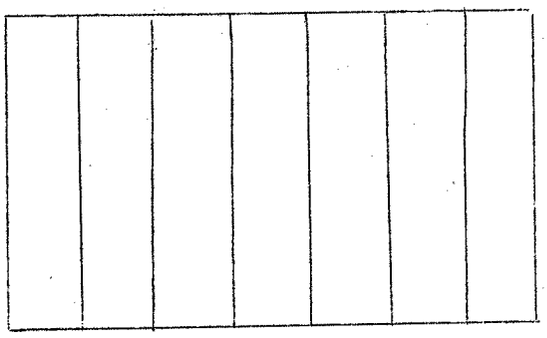
| Variables sémantiques | | Quelques valeurs de ces variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 0 | Opération | addi-tion | Diffé-rence | Réduc tion | éga-lité | | | | |
| 1 | Nature mathématique | groupe | sys-tème | déter-minant | Vec-teur | Sym-bole | Frac-tion | | |
| 2 | Rencontré dans | théo-rème | Rela-tion | Défi-nition | Propri-été | For-mule | Equa-tion | | |
| 3 | Utilisation en | théo-rie | Appli-cation | Calcul | | | | | |
| 4 | Interprétation | Géomé trie | | | | | | | |
| 5 | Etre mathématique | Fonc-tion | Vari-able | | | | | | |
| 6 | Forme | Elémen taire | 2 | 3 | ... | ... | | multi. | |
| 7 | Espace | 1 | 2 | 3 | 4 | ... | .... | n | $\infty$ |
| 8 | Degré | 1 | 2 | 3 | 4 | ... | ... | | n |
| 9 | Symétrie | /0 | /droite | / plan | | | | | |
| 10 | Exactitude | Pré-cis | Impré-cis | 1/100 | ... | | | | |
| 11 | Réalité | Réel | Imagi-naire | Com-plexe | | | | | |
| 12 | Rationalité | Oui | Non | | | | | | |
| 13 | Borne | In-finie | In-finie | Supé-rieure | Infé-rieure | | | | |
| ⋮ | | | | | | | | | |

Nous pouvons représenter l'ensemble des mots appartenant à un champ par un carré.

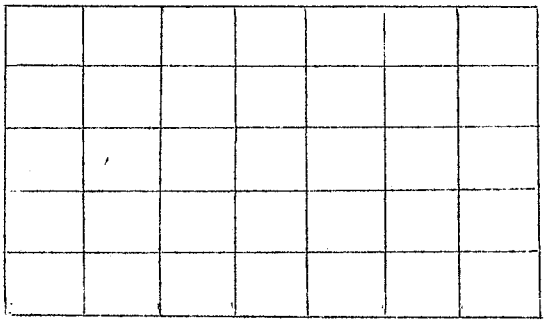
Considérons la première variable sémantique $V_s^1$ celle-ci peut prendre $n_1$ valeurs distinctes aves $1 \leq n_1 \leq 8$. Cette $V_s^1$ va donc ranger l'ensemble des mots du champ, en $n_1$ classes que l'on représente de la façon suivante :



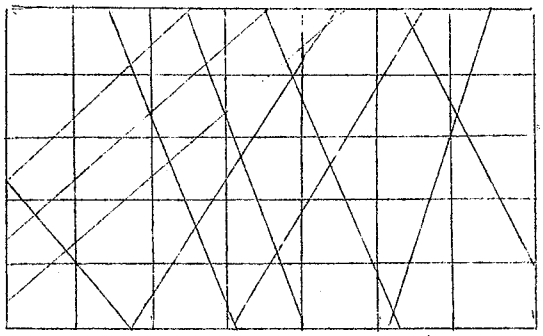
Considérons la seconde variable sémantique $V_s^2$ , celle-ci peut prendre $n_2$ valeurs distinctes et $V_s^2$ va engendrer une autre partition représentée de la façon suivante :



Considérons maintenant la partition résultant de l'action conjuguée des deux variables sémantiques $V_s^1$ et $V_s^2$. L'ensemble des mots du champ est cette fois partagé en : $n_1 \times n_2$ classes distinctes et l'on obtient la représentation de celles-ci en superposant les deux schémas précédants:



D'une façon plus générale si dans le champ existent m variables sémantiques $V_s^i$ prenant respectivement, $n_1$, $n_2$;.. $n_m$ valeurs distinctes, ces $V_s$ auront partagé l'ensemble des mots appartenant à ce champ en $n_1 \times n_2 \ldots \times n_m$ classes distinctes. Il existe un schéma (certes complexe) représentant cette partition, qui est obtenu par superposition des m schémas élémentaires représentant la répartition plane due à chaque $V_s^i$.



## IV-. RELATION SEMANTIQUE

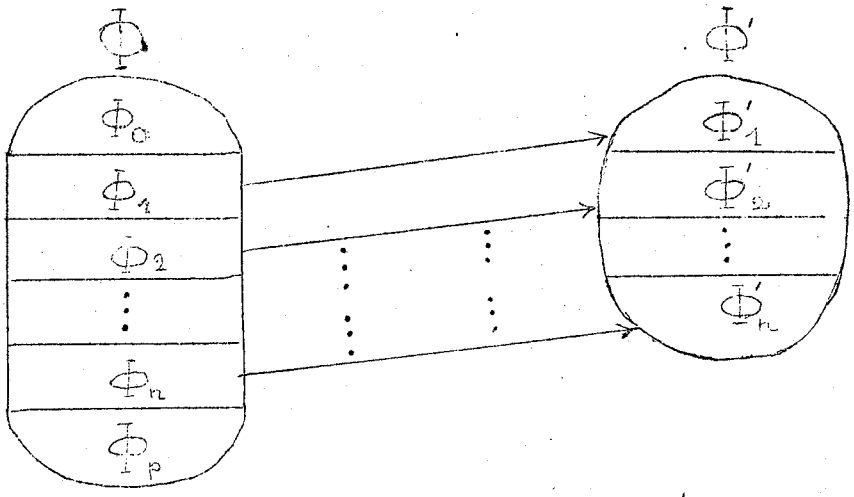
Entre deux mots liés par une relation syntaxique, il existe en général une relation sémantique.

Exemple : Si je considère la relation syntaxique "substantif commun - groupe prépositionnel gouverné par "de"" je trouve un ensemble de relations sémantiques comme :

- l'appartenance — Ex. : le tronc de l'arbre
- la fonction — Ex. : le chef de gare
- l'objet — Ex. : la culture du thé

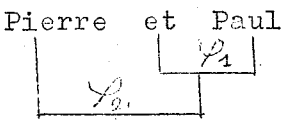
Nous verrons au chapitre III (III) la représentation des relations syntaxiques et sémantiques.

Considérons l'ensemble $\Phi$ des $\varphi$ et l'ensemble $\Phi'$ des $\varphi'$ Ceux-ci peuvent être subdivisés en plusieurs sous-ensembles ayant des liaisons entre eux. Nous avons un schéma du type suivant :




A $\Phi_0$ ne correspond aucun sous-ensemble de $\Phi'$: les $\varphi$ qui en font partie ne sont reliés à aucun $\varphi'$. Ces relations syntaxiques ne sont utilisées qu'aux premiers étages du schéma syntaxique, et à ce stade-là, les deux termes ainsi reliés ne possèdent parfois pas suffisamment de renseignements sémantiques pour être reliés par un $\varphi'$.

Exemple : Pierre et Paul




A $\varphi_1$ ne correspond aucun $\varphi'$

A $\Phi_1$ correspond $\Phi'_1$ ; chacun de ces $\varphi$ est relié à un $\varphi'$
A $\Phi_2$ correspond $\Phi_2$ ; chacun de ces $\varphi$ est relié à deux $\varphi'$

⋮

A $\Phi_n$ correspond $\Phi'_n$ les $\varphi$ qui y sont inclus sont reliés à n $\varphi'$
A $\Phi_p$ ne correspond aucun sous ensemble de $\Phi'$.
dans $\Phi_p$ seront contenues les relations syntaxiques les plus complexes qui n' interviennent qu'au moment de l'obtention de la phrase.

## V-. RELATION ENTRE CHAMPS ET DEBUT DU CODE SEMANTIQUE

Nous voyons que des mots de deux champs différents peuvent être liés par une relation sémantique. Si nous prenons tous les champs deux à deux, nous verrons que certains couples de champs présentent entre leurs mots des possibilités de relation sémantique ; tandis qu' un autre couple de champs ne présentera aucune possibilité. Nous pourrons ainsi former une matrice (cf. page suivante) donnant la possibilité ou l'impossibilité pour deux mots appartenant à deux champs différents d'être reliés par une relation sémantique.

Les relations entre champs nous donneront de suite ce renseignement.

Cette matrice symétrique se présentera de la façon suivante: le nombre "1" marquera l'existence d'une relation entre deux champs ; et le nombre "0" la non-existence d'une relation. Si la fabrication des champs a été bien faite, le nombre de "0" devra être bien supérieur à celui des "1".

La diagonale principale n'est remplie que de nombre "1" en effet, à l'intérieur d'un champ deux mots peuvent souvent être reliés.

Nous appellerons cette matrice "matrice de relation entre champs".

| $C_H A_M P$ \ $C_H A_M P$ | 000 | 001 | 002 | | ---- | 511 | 512 | 513 | ------ | 997 | 998 | 999 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 000 | 1 | | | | | 1 | 0 | 0 | | 0 | 1 | 0 |
| 001 | | 1 | | | | 0 | 0 | 1 | | 1 | 1 | 0 |
| 002 | | | 1 | | | 0 | 1 | 0 | | 1 | 0 | 1 |
| | | | | 1 | | | | | | 0 | 1 | 0 |
| ⋮ | | | | | ⋱ | | | | | | | |
| 511 | 1 | 0 | 0 | | | 1 | 1 | 0 | | 0 | 1 | 0 |
| 512 | 0 | 0 | 1 | | | 1 | 1 | 0 | | 1 | 0 | 1 |
| 513 | 0 | 1 | 0 | | | 0 | 0 | 1 | | 1 | 0 | 0 |
| ⋮ | | | | | | | | | ⋱ | | | |
| 997 | 0 | 1 | 1 | 0 | | 0 | 1 | 1 | | 1 | | |
| 998 | 1 | 1 | 0 | 1 | | 1 | 0 | 0 | | | 1 | |
| 999 | 0 | 0 | 1 | 0 | | 0 | 1 | 0 | | | | 1 |

Nous allons utiliser ces définitionssur la sémantique et indiquer de quelle façon une unité sémantique va être précisée.

On a vu qu'une $U_s$ appartient à un ou plusieurs champs et que dans chaque champ elle est indicée par une ou plusieurs $V_s$. Nous allons représenter ces renseignements de la façon suivante :

$$U_s \left\| chi \left| V_s^1 \ \ldots \ V_s^m \right| , chj \left| \ldots \right| , \ldots , chk \left| V_s^n \ldots \ V_s^p \right| \right\|$$

Bien entendu, chaque variable sémantique est précisée par la valeur qu'elle prend.

Nous venons de définir ici le début du <u>code sémantique</u>

## VI-. ELEMENTS NOTIONELS

Considérons une Unité sémantique particulière d'une unité lexicale. Dans l'étude des variables sémantiques appartenant aux différents champs où l'on rencontre cette $U_s$, on remarque que certaines apparaissent dans tous les champs (et, de plus, avec la même valeur pour $V_s$) Ces quelques variables correspondent au concept général de l'Unité sémantique et il y a intérêt à les considérer d'une façon particulière D'autre part les études effectuées pour trouver les fonctions syntaxiques et sémantiques nous ont montré que dans la différentiation de celles-ci intervenaient aussi des notions particulières. Ces notions se rapprochent de celle de "sphère notionnelle" présentée par le Professeur CECCATO [6].

Ces notions particulières et les variables sémantiques correspondant au concept général s'appellent <u>Eléments notionels: En</u>

Il y aura d'ailleurs intérêt à classer ces éléments notionels suivant l'ordre de rencontre le plus probable.

Ainsi nous allons modifier légérement la configuration du code sémantique $C_{\Gamma}$ et le représenter de la façon suivante :

$$C_{\Gamma}(Us) = Us \mid E_n^p \ldots E_n^q \mid \underset{i}{<} chi \mid V_s^1 \ldots V_s^m \mid >$$

ou en résumé

$$E_{\Gamma}(Us) = \Gamma(Us, En, chi, V_s)$$

Le symbole : $A \underset{i}{<} \alpha_i >$ représente l'énumération de tous les $\alpha_i$ concernant A . Ainsi $U_s \underset{i}{<} chi \mid V_s^1 .. V_s^m \mid >$ représente l'énumération :

$$Us \| chi \mid V_s^1 \ldots V_s^m \mid , chj \mid \ldots \mid , \ldots , ch_K \mid V_s^n \ldots V_s^p \|$$

- $C_{\Gamma}(Us)$ est caractéristique d'une unité sémantique Us et indépendant du texte ou du champ où a pu être rencontrée cette Us.

Les éléments notionels seront donnés progressivement quand on avancera dans l'étude des variables sémantiques des unités sémantiques et surtout dans celle des fonctions syntaxiques et sémantiques.

Nous pouvons déjà donner un exemple d'éléments notionels pour les verbes ; les travaux effectués par le centre de cybernétique de l'Université de Milan [7] ont été pour nous un guide précieux.

## Eléments notionels du verbe

Les verbes peuvent êtres rangés en deux grandes classes : ceux d'action et ceux d'état.

Ceux d'action peuvent être séparés d'une façon un peu plus précise en verbe de :

- 1 - Action en soi
- 2 - Action évaluatrice d'un "objet"
- 3 - Action constructive d'un "objet"
- 4 - Action modificatrice d'un "objet"

et ceux d'état en :

- 5 - Etat moral
- 6 - Etat physique.

Les autres éléments notionels rencontrés sont les suivants :

- catégorie mentale (Psychique/Impression/Sentiments...)
- royaume des activités musculaires(Participation/...)
- temps
  - durée
    - continue
    - avec intervalles
  - Instantanée
- changement
  - dans l'espace
  - dans le temps
  - d'état
    - physique
    - moral
- Déplacement/mouvement
  - Direction définie
  - en tous sens
  - spontané
  - préparé
- Comparaison/ Evaluation/Différentiation
- Modification/Transformation/Développement
- Conservation/Non-développement
- Perception/Sensation
- Distance X Contact
- Présence X Absence

Ainsi, nous avons dénombré 26 En distincts.

Nous voyons que les éléments notionels de par leur généralité auront un rôle bien plus actif que les variables sémantiques ordinaires. Ils seront utilisés dans l'étude de passage des $\varphi$ aux $\varphi'$. Ils participeront ainsi à la construction de la phrase en langue source : Sélection du $\varphi$ de sortie .... tandis que les variables sémantiques auront leur rôle principal, dans l'étude des Unités lexicales et des unités sémantiques. Elles rechercheront l'unité lexicale et l'unité sémantique ici représentées ; cette utilisation demandera des critères plus fins, mais elle sera plus passive.

## Le code sémantique d'une unité lexicale

Si une unité lexicale correspond à une seule unité sémantique on fera suivre cette Ul du code sémantique de l'Us. Mais si une unité lexicale correspond à plusieurs unités sémantiques, on fera suivre cette Ul des codes sémantiques des différentes Us.

Nous aurons donc :

$$C_{\sigma}(Ul) = U_s^1 \left[ C_{\sigma}(U_s^1) \right] U_s^2 \left[ C_{\sigma}(U_s^2) \right] \dots U_s^n \left[ C_{\sigma}(U_s^n) \right]$$

avec

$$C_{\sigma}(U_s^i) = U_s^i \mid E_n^p \dots E_n^q \mid \langle chj \mid V_s^1 \dots V_s^m \mid \rangle$$

C'est $C_{\sigma}(Ul)$ qui intervient dans la constitution du terme ordonné.

Exemple : L'Ul "glace" aura pour code sémantique, la succession des codes sémantiques des trois Us :

" eau gelée" , "sorbet", "lame de verre".

## VII-. LE CODE SEMANTIQUE EN MACHINE

Essayons de voir la place prise sur machine, en mémoire rapide par le code sémantique.

Un champ est représenté par le symbole CDU, chaque lettre correspond à 4 bits. Donc un champ correspondra à 3 X 4 = 12 bits.

Pour chaque champ, on a prévu 32 variables sémantiques ; donc une $V_s$ sera représentée par 5 bits.

Pour chaque variable sémantique, on a prévu 8 valeurs possibles Donc la valeur de la $V_s$ sera représentee par 3 bits.

Ainsi un mot appartenant à un champ et précisé par n variables sémantiques, occupera en mémoire rapide : $\left[12 + n\ (5+3)\right]$ bits.

Pour se fixer les idées, prenons n = 5

Le mot occupera pour chaque champ auquel il appartient :

12 + 5X8 = 12+40 = 52 bits

Supposons encore que le mot appartienne à cinq champs différents, la place occupée en machine par

$\langle_j chj \left| V_s^1 \ldots V_s^m \right| \rangle$ sera de 5 X 52 = 260 bits

Si nous prévoyons 32 éléments notionels possibles (soit 5 bits pour une unité sémantique et si chaque $U_s$ peut en avoir quatre $\left| E_n^1 \ldots E_n^4 \right|$ occupera en machine : 20 bits.

Donc le code sémantique $C_{\sqrt{}}(U_s)$ sera de :

260 + 20 = 280 bits soit un peu moins de huit mémoires (288 bits).

Ce calcul est certes très grossier, car nous avons pris des chiffres d'une façon parfois très arbitraire et nous avons omis de considérer que des variables sémantiques pouvaient être communes aux différents champs.

Mais de toutes façons, nous voyons au moins que la longueur du code sémantique est énorme par rapport à celle des codes morphologiques et syntaxiques.

Ainsi se pose le problème d'essayer de réduire la longueur de ce code $C_{\sigma}$

Est-il possible d'adopter une écriture permettant de rendre plus faible le nombre de bits utilisé ?

Nous remarquons que le nombre de $C_{\sigma}(U_s)$ différents est égal au produit du nombre de :

| | | |
|---|---|---|
| - champs | : | $1000 \sim 2^{10}$ |
| - variables sémantiques (par champ) | : | $32 = 2^5$ |
| - valeurs de chaque $V_s$ | : | $8 = 2^3$ |
| - Eléments notionels (par $U_s$) | : | $4 = 2^2$ |
| - valeur de chaque En | : | $32 = 2^5$ |
| | | $2^{25}$ |

Nous avons donc $2^{25}$ $C_{\sigma}(U_s)$ distincts. Ce qui peut donc être représenté par 25 bits. Or une mémoire contient 36 bits. Ainsi le code sémantique pourrait être contenu en une seule mémoire. Mais il faudrait alors faire appel à une table intermédiaire et à un programme qui permettrait de passer de cette écriture condensée à l'écriture symbolique vue précédemment. Or cette table doit contenir les codes sémantiques de toutes les $U_s$. Son occupation en mémoire sera pour chaque code, la même que précédemment et le procédé envisagé sera sans doute plus long que de considérer successivement l'ensemble des huit ou dix mémoires représentant le code sémantique.

On prévoit que 20 000 mémoires rapides peuvent être disponibles dans la machine et que le nombre d'Unités Sémantiques donc de codes $C_{\sigma}(U_s)$ est d'environ 20 000 (10 000 U1 X 2 nombre moyen d'$U_s$ par U1). Il faut donc envisager d'avoir 8 passages, à effectuer en machine à chaque recherche de $C_{\sigma}(U_s)$. Le procédé de LAMB [8] permettrait peut-être de diviser par deux ou trois le nombre de passages, mais il faudrait alors se limiter dans la longueur des textes étudiés.

Donc de toutes façons, nous voyons que l'étude sémantique demandera du temps.

## III-. POSSIBILITES DE CONCENTRER LES CHAMPS

On a vu qu'un mot pouvait appartenir à plusieurs champs. Si ce mot apparait dans tous les champs (C D U) où U prend toutes les valeurs de 0 à 9 et où C et D restent fixes on dira que ce mot appartient aux champs (CD.)

D'une façon semblable, en considérant les champs (CD.) et en faisant prendre à D, toutes les valeurs de 0 à9 on définira les champs (C..) et de même les champs (...) où C prend toutes les valeurs de 0 à 9.

Les mots pourront suivant la généralité de leur sens, appartenir à des champs représentés de la façon suivante :

| | C'est un mot : |
|---|---|
| (CDU) | - "très précis" |
| (CD.) | - "précis" |
| (C..) | - "général" |
| (...) | - "très général". |

Supposons qu'un mot appartienne à plusieurs champs pour lesquels seul "U" varie.

On a donc un ensemble de champs

CDα
CDβ
CDγ
CDδ

A partir de combien de "U" peut-on remplacer l'ensemble de ces champs par la seule représentation (CD.) ?

De suite, se présente l'objection suivante :
Les variables sémantiques peuvent être différentes suivant les (CDi). Comment les représenter ? . Il y a trois possibilités :

-1- Supprimer toutes les $V_s$ non communes

-2- Garder toutes les $V_s$

-3- Adopter une représentation permettant de garder le maximum d'informations, à savoir : les $V_s$ communes et les $V_s$ particulières à (CDi) . On obtient : (CD.) ($V_s$ communes) + $\alpha$ ($V_s^{\alpha}$) +... + $\delta$ ($V_s^{\delta}$)

Le problème se repose pour déterminer le nombre de "D" nécessaire pour remplacer l'ensemble des champs :

$C\alpha.$
$C\beta.$
$C\gamma.$
$\vdots$
$C\nu.$

par la représentation (C..)
et de même à partir de combien de "C" faut-il remplacer par la représentation (...) l'ensemble des champs :

$\alpha..$
$\beta..$
$\vdots$
$\nu..$

il faudra encore choisir une politique pour le choix des $V_s$.

# - III -

## LIAISONS ENTRE LES ELEMENTS SYNTAXIQUES ET LES ELEMENTS SEMANTIQUES

### I-. FORMATION DES UNITES LEXICALES ET SEMANTIQUES

D'après les définitions de l'unité lexicale et de l'unité sémantique vues aux chapitres précédents, nous allons voir comment former ces unités en utilisant un dictionnaire du type classique, comme par exemple le Larousse.

Nous pouvons rencontrer trois cas bien distincts :

a) dans le dictionnaire classique, il n'existe qu'un seul mot écrit de cette façon.

Dans ce cas, à ce mot de $\Lambda$ correspond dans $\mathcal{L}$ une unité lexicale. En étudiant ce mot, nous recherchons les concepts sémantiques et à chacun de ceux-ci nous faisons correspondre dans $\mathcal{L}$ une $U_s$.

Exemple :

$\Lambda \longrightarrow \mathcal{L}$

glace $\Longrightarrow$ 1 Ul $\longrightarrow U_s^1$ Eau congelée

$\longrightarrow U_s^2$ crême aromatisée et congelée

$\longrightarrow U_s^3$ lame de verre

(mot issu du latin : glacies)

b) Dans le dictionnaire classique, il existe plusieurs mots écrits de cette façon, ils possèdent chacun un concept sémantique.

Si à chacun d'eux on faisait correspondre dans $\mathcal{L}$ une unité lexicale, ces différentes Ul auraient le même Ku et les mêmes Vgp. La raison pour laquelle on rencontre plusieurs fois ces mots

dans le dictionnaire est que généralement leur étymologie est différentes. A ces mots, nous ne faisons correspondre qu'une seule Ul. Nous considérons le concept sémantique de ces mots pris un à un et à chacun d'eux nous lui associons dans $\mathcal{L}$ une $U_s$. Ainsi à ces m différents mots, correspondent dans $\mathcal{L}$ une seule Ul mais m $U_s$

Exemple : Il existe trois mots "son" substantif-commun masculin :

| | Concept sémantique |
|---|---|
| son (issu du latin sonus) | Effet des vibrations excitant l'organe de l'ouie |
| son (issu de l'ancien français : saonner) | Partie périphérique des graines de céréales |
| son | Echeveau de laine cardée. |

Nous avons dans $\mathcal{L}$ une Ul "son" et trois "$U_s$".

c) C'est la solution mixte à partir de a) et b) : Nous avons plusieurs mots écrits de cette façon ; si à chacun d'eux on faisait correspondre dans $\mathcal{L}$ une unité lexicale, ces différentes Ul auraient le même Ku et les mêmes Vgp ; de plus certains de ces mots ont plusieurs concepts sémantiques. A ces mots nous ne faisons correspondre dans $\mathcal{L}$ qu'une seule unité lexicale mais à chacun des concepts des différents mots, nous associons une $U_s$.

Exemple : Il existe deux mots "rayon" : substantif commun masculin

| | concept sémantique |
|---|---|
| rayon (emprunté au francique) | 1 gateau de cire<br>2 tablette d'une armoire. |
| rayon (latin : radius) | 3 trajectoire suivie par la lumière<br>4 distance d'un cercle/sphère à chaque point de ce cercle/sphère.<br>5 choses partant et divergeant d'un centre commun<br>6 sillon tracé en labourant.<br>7 radiations particulières<br>8 zone d'activité. |

Nous avons dans $\mathcal{L}$ une U1 "rayon" mais huit "$U_s$".

## II-. INTRODUCTION DES RELATIONS SYNTAXIQUES

L'étude syntaxique [5] a montré que dans le langage artificiel $\mathcal{L}$, à une relation $\psi_q$ entre catégories syntaxiques, correspond un ensemble de relations $\varphi_p$. Nous allons considérer dans la syntaxe traditionnelle du langage naturel $\Lambda$, ce qui est représenté par un $\varphi$ dans le langage $\mathcal{L}$. Les liaisons syntaxiques entre deux mots de $\Lambda$ sont introduites par l'intermédiaire soit d'une variable grammaticale soit d'un "mot de transition" soit de ces deux notions.

J'appelle "mot de transition", un mot du type des prépositions qui sert de lien entre deux mots. Une liaison syntaxique par une variable grammaticale se fait soit par l'intermédiaire d'un "cas" soit par simple juxtaposition avec concordance des v.g.

Exemple :

En russe КНИГА СЕСТРЫ (génitif)(le livre de sa soeur)
"kniga sestry "

En français : la petite maison (juxtaposition)

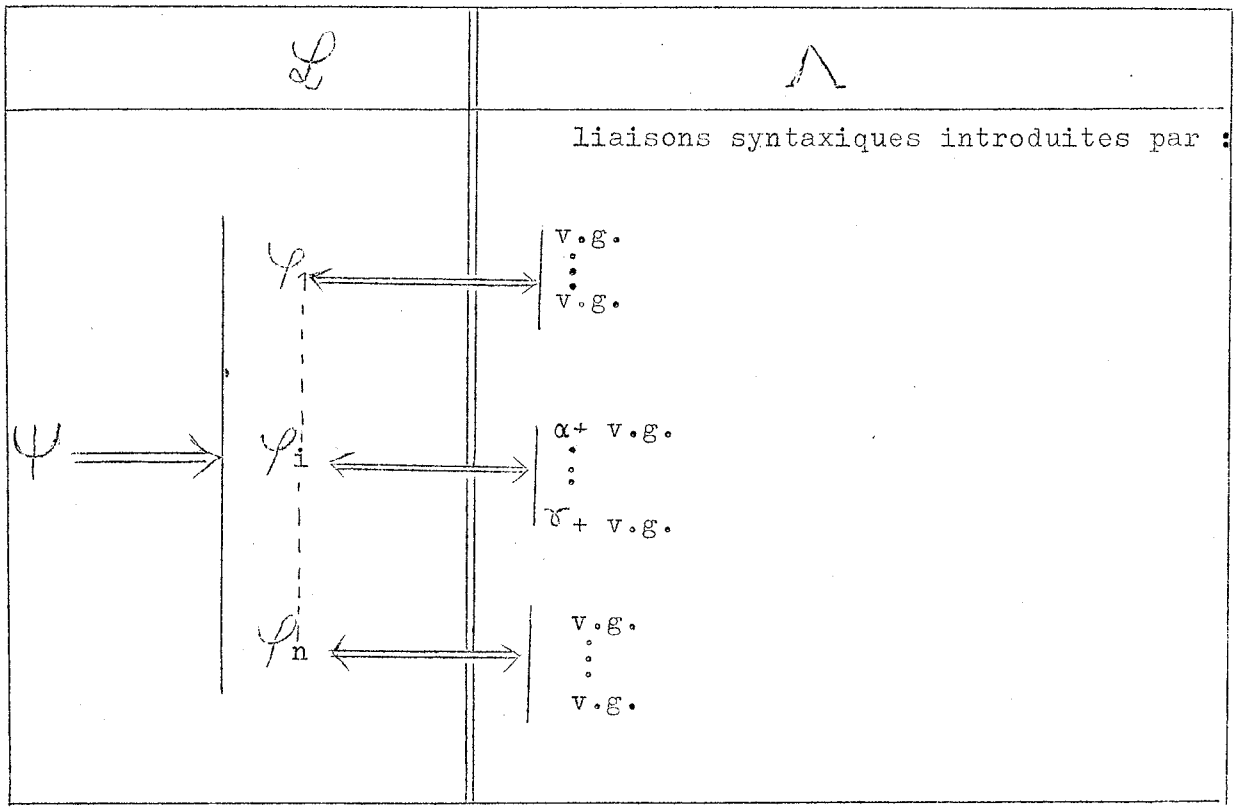
En anglais : little house (place immuable de l'Adjectif)

Exemple: de liaison syntaxique par un "mot de transition" :

En français : travailler avec ardeur

Exemple : de liaison syntaxique par un "mot de transition" + une v.g.
en russe : УЕХАТЬ В СИБИРЬ (Acc) (Partir en Sibérie)
"u'echat' v Sibir' "

Ces liaisons syntaxiques sont représentées par des $\varphi$ dans $\mathcal{L}$
Nous avons donc entre $\mathcal{L}$ et $\Lambda$ le schéma de correspondance syntaxique suivant :




Dans $\mathcal{L}$ un $\Psi_q$ correspond à n $\varphi_p$ ; et un $\varphi_p$ est un ensemble de règles formelles.

Chaque $\varphi_p$ est la représentation dans $\mathcal{L}$ d'une liaison syntaxique dans $\Lambda$, liaison qui est introduite soit par une ou plusieurs v.g. soit par un ou plusieurs "mot de transition + v.g.". Les v.g. et les mots de transition dépendent de la liaison syntaxique ainsi que des deux mots reliés par celle-ci.

Tous ces éléments vont apparaitre dans la fabrication des règles formelles des $\varphi_p$.

Nous allons donner des exemples illustrant le schéma précédent :

Exemple :

| | $\varphi$ | Introduction des liaisons syntaxiques par | |
|---|---|---|---|
| | | En russe | En français |
| $\Psi$ Subst./Subst. Commun/commun ⟹ | nom/nom en apposition ⟷ | même cas | |
| | ⋮ | | |
| | nom/complément de nom ⟷ | génitif | |
| Verbe.../Substantif commun ⟹ | Verbe.../Comp$^t$ indirect ⟷ | datif | |
| | ⋮ | | |
| | Verbe.../Comp$^t$ direct ⟷ | accusatif<br>génitif | |
| Verbe../Groupe prépositionnel ⟹ | Verbe.../Préposition de lieu + substantif ⟷ | B + Subst. "v" {Acc. loc.}<br>HA +Subst. "na" {Acc. Loc.} | à + substantif<br>{vers + substantif<br>dans + substantif} |
| | ⋮ | | |
| | Verbe.../Préposition de but + substantif ⟷ | Для + subst. gén. "dl'a" | pour + Substantif |

## III-. INTRODUCTION DES RELATIONS SEMANTIQUES ET "LIAISONS EN TRIANGLE"

D'autre part, mettons en évidence les relations sémantiques.

Entre deux mots liés par une relation syntaxique il existe en général une relation sémantique.

Donc si je considère une $\varphi_p$, celle ci représentera généralement une $\varphi'_p$ prise parmi un ensemble de $\varphi'_p$.

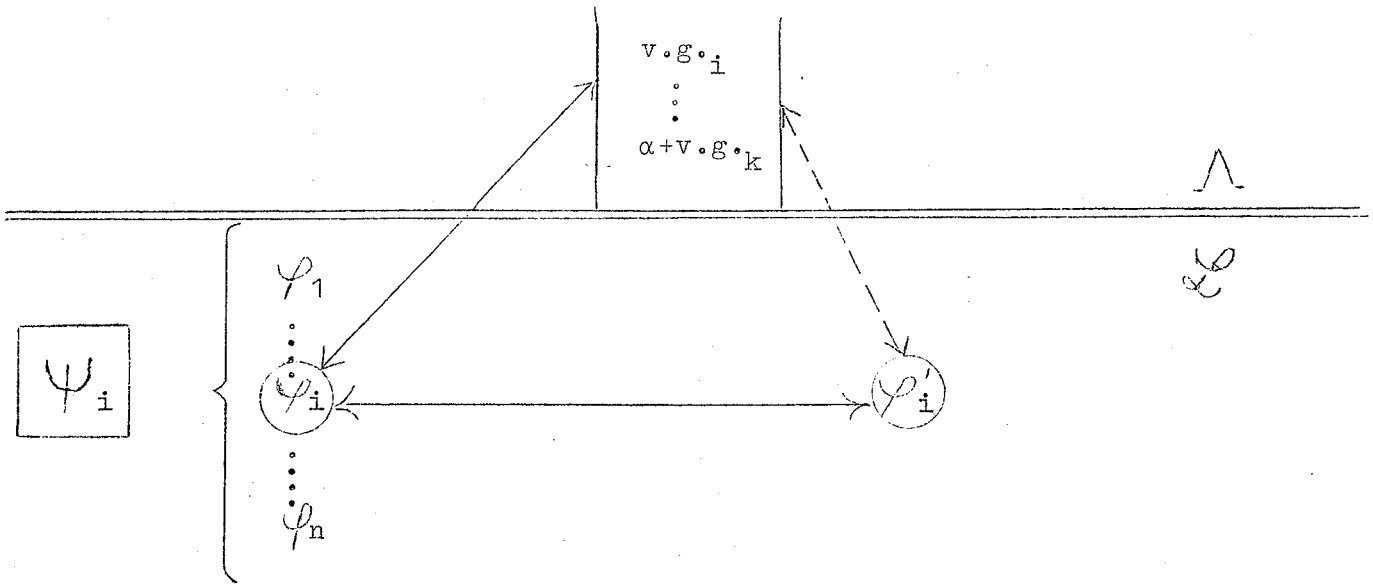
Ce choix dépendra des deux termes reliés par cette relation. Ainsi en général à un $\varphi_p$ on fait correspondre un ensemble de relations sémantiques : $\varphi_{\underset{i}{<}\varphi'_i>}$.

Nous rappelons que le symbole $A_{\underset{i}{<}\alpha_i>}$ représente l'énumération de tous les $\alpha_i$ concernant A. Donc $\varphi_{\underset{i}{<}\varphi'_i>}$ représente l'énumération de tous les $\varphi'_i$ concernant $\varphi$ soit $\varphi \| \varphi'_1 | \varphi'_2 | \cdots | \varphi'_n \|$.

Exemple :

| | $\varphi_{\underset{i}{<}\varphi'_i>}$ | Exemples |
|---|---|---|
| $\varphi_p$ Nom/ Complément de nom | Appartenance | le tronc de l'arbre |
| | direction | le train de Paris |
| | fonction | le chef de gare |
| | partitif | la tasse de thé |
| | objet | la culture du thé |

A l'intérieur de chaque $\psi$, nous avons donc un ensemble de liaisons qui peuvent être représentées par le schéma suivant :



Il serait intéressant de faire un recensement des $\varphi$ et des $\varphi'$, et de trouver à quelles fonctions $\varphi_p$ et à quelles fonctions $\varphi'_p$ correspond chaque v.g. ou chaque ensemble "α + v.g." et après de relier les $\varphi_p$ et les $\varphi'_p$.

Pour l'instant, nous nous sommes limités dans ce recensement et dans l'étude des $\varphi_p$, aux relations syntaxiques qui ont pour premier terme : un verbe. Ce sont les plus importantes. La v.g. "cas" étudiée est alors celle du second terme.

De plus ces relations $\varphi_p$ où le premier terme est un verbe, sont plus simples à étudier, car seul le second terme est qualifié par un "cas", tandis que dans d'autres $\varphi_p$, les deux termes sont qualifiés par un "cas".

Pour trouver ces "liaisons en triangle" v.g. (2ème terme) - $\varphi_i$ - $\varphi'_i$ nous opérons en étudiant chaque v.g. (ou chaque ensemble α + v.g.) du 2ème terme.

Nous allons débuter par l'ensemble des $\varphi_p$ appartenant aux $\psi_q$ suivants :

- Verbe... + substantif commun
- Verbe... + groupe prépositionnel

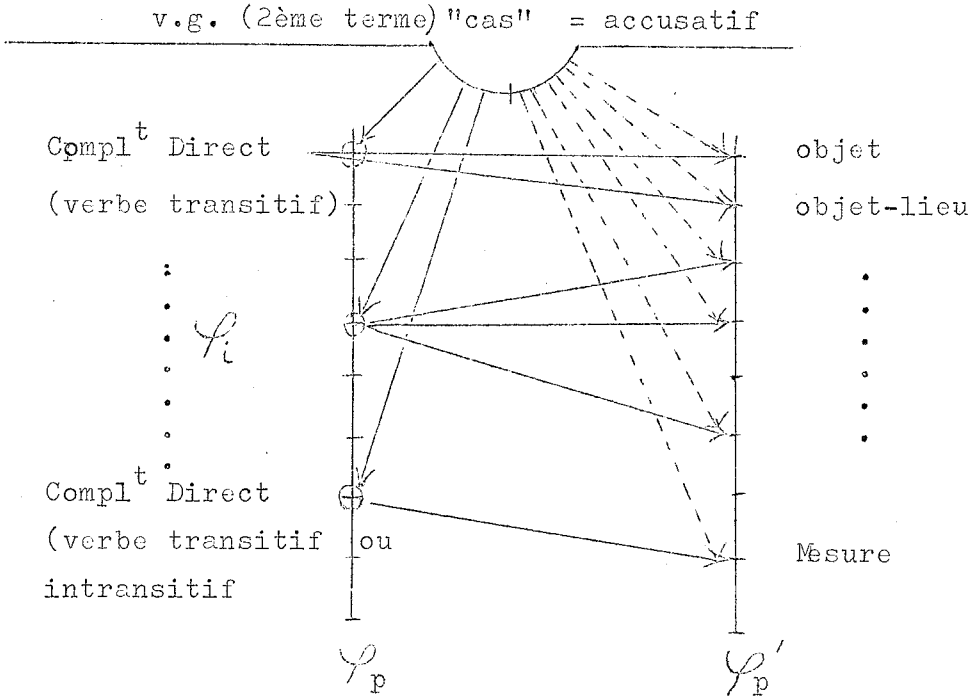
et essayer de trouver les liaisons ⟶
- v.g. "cas" (2ème terme) → $\psi_i \langle \varphi_i \rangle$
- v.g. "cas" (2ème terme) → $\varphi_k \langle \varphi'_k \rangle$
- $\varphi_i \longrightarrow \varphi_k \langle \varphi'_k \rangle$

Ces liaisons peuvent être localisées par le schéma de la page suivante :

les liaisons ⟶ représentent des liaisons directes

v.g. "cas" (2ème terme) $\longrightarrow \varphi_i \longrightarrow \varphi'_p$

Exemple : $\Psi$ = verbe .../Substantif commun




Dans ces diverses liaisons, nous introduirons des valeurs sémantiques : les éléments notionels définis au chapitre II. Celles-ci qualifieront les deux termes des relations $\varphi_p$ et $\varphi'_p$.

Nous espérons ainsi pouvoir distinguer entre elles ces différentes relations.

Nous voyons de nouveau intervenir le code sémantique, mais il nous faut pour cela, préciser la représentation des relations syntaxiques et sémantiques.

## IV-. REPRESENTATION DES $\varphi$ ET DES $\varphi'$

### 1- Cas particulier des prépositions

Parmi l'ensemble des catégories syntaxiques, il en existe une qui joue un rôle particulier, c'est celle des "prépositions".

Une "préposition" est un mot invariable exprimant les rapports entre les mots. Mais quand on traduit, cette préposition peut disparaitre en langue cible. Cette disparition peut se produire de deux façons différentes :

a) La préposition est remplacée en langue cible par un "cas" qualifiant le mot qui la suivait en langue source.

| Exemple : Langue source : français | Langue cible : russe |
|---|---|
| Il travaille avec un marteau | ОН РАБОТАЕТ МОЛОТОМ (instru.) "on rabota'et molotom" |
| le livre de la soeur | КНИГА СЕСТРЫ (génitif) "kniga sestry" |

b) La préposition se combine avec le mot qui la suivait en langue source, pour donner en langue cible un mot dont le $K_{u,v}$ est différent du $K_{u,v}$ d'origine.

| Exemple : Langue source | Langue cible |
|---|---|
| Il travaille avec ardeur | He works eagerly (Anglais) |
| La maison du père | ОТЦОВ ДОМ (Russe) "otcov dom" |

Dans le premier exemple, nous obtenons un adverbe
Dans le second exemple, nous avons un adjectif d'appartenance.

Cette possibilité de disparition nous entraine à considérer d'une façon particulière la catégorie syntaxique des prépositions.

Dans l'étude syntaxique de la langue source, on considère des relations syntaxiques $\varphi_p(s_i, s_j) = s_k$ . Dans le cas où la $K_{u,v}$ de $s_i$ est celle des prépositions, nous noterons "$s_i$" et "$s_k$" de la façon suivante: $s_k^{\varphi} = \varphi_p^{\varphi}(s_i^{\varphi}, s_j)$

et $\varphi_p^{\varphi}$ sera représentée de la façon suivante : la préposition elle-même suivie des renseignements syntaxiques concernant le syntagme $s_j$.

$\varphi_p^{\varphi}$ : | $s_i^{\varphi}$ | Renseignements syntaxiques / $s_j$ |

Exemple : Il travaille avec ardeur

$\varphi_p^{\wp}$

La relation syntaxique appliquée à "avec ardeur" est représentée par :

| avec | Substantif commun... |
|---|---|

Renseignement syntaxique/"ardeur"

## 2- Représentation des $\varphi'$

Considérons d'abord ([2] page 42) la représentation du terme ordonné élémentaire que nous représentons par le symbole $T_i$.

$\langle$Terme ordonné élémentaire$\rangle$ ::= $\langle \nu \rangle \langle$N° U1$\rangle \langle K \rangle \langle C\Delta \rangle \langle vgp \rangle \langle vgc \rangle \langle cg \rangle \langle cd \rangle \langle c\sigma \rangle$

Au syntagme élémentaire $S_i$ correspond le terme ordonné élémentaire $T_i$ et au syntagme $S_k$ correspond le terme ordonné $T_k$.

L'évolution des renseignements syntaxiques dans le passage du syntagme élémentaire $S_i$ au syntagme $S_k$ a été précisée par AUROUX [5] Il a défini :

$\langle$syntagme$\rangle$ ::= $\langle \nu \rangle \langle K_{u,v} \rangle \langle vgp \rangle \langle vgc \rangle \langle C_s \rangle \langle N \rangle \langle F(\varphi_p, S_i, S_j) \rangle$

Nous pouvons définir le terme ordonnée $T_k$ comme suit :

$\langle$terme ordonné$\rangle$ ::= $\langle$syntagme$\rangle \langle$code sémantique$\rangle$

Nous avons défini les syntagme $S_i^{\wp}$ et $S_k^{\wp}$ ; à ceux-ci correspondent les termes ordonnés $T_i^{\wp}$ et $T_k^{\wp}$.

Une $\varphi_p$ représente une relation syntaxique entre deux syntagmes $S_i$ et $S_j$. Ces syntagmes peuvent être soit élémentaires soit complexes.

Notre but est de rechercher en langue source les $\varphi'$ correspondant à ce $\varphi$ ; soit $\varphi \underset{i}{<} \varphi'_i >$ .

Un $\varphi'$ est appliqué à deux termes ordonnés $T_i$ et $T_j$ reliés par une relation syntaxique $\varphi$; nous notons ceci sous la forme : $\varphi'(T_i, T_j)$ mais aucun de ces termes ne doit être un $T_i^{\varphi}$.

Nous considérons donc les $\varphi'$ des types suivants $\begin{cases} \varphi'(T_i , T_j) \\ \varphi'(T_i, T_k^{\varphi}) \end{cases}$

mais non ceux du type $\varphi'(T_i^{\varphi}, T_j)$

Les deux premiers types de $\varphi'$ sont appelés "$\varphi'$ véritables" c'est-à-dire réellement utilisables dans la transition source-cible.

Chaque $\varphi'$ peut se représenter de la façon suivante :

| $\varphi'$ | 1er terme | 2ème terme |
|---|---|---|

Chaque terme est caractérisé par ses "Eléments Notionels caractéristiques"

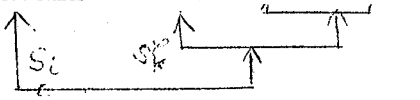
Nous avons donc :

| $\varphi'$ | En caract. 1er terme | En caract. 2ème terme |
|---|---|---|

Pour que $\varphi'_i$ soit différent de $\varphi'_j$ il faut et il suffit qu'un des éléments notionels caractéristiques du premier ou du second terme soit différent pour $\varphi'_i$ et $\varphi'_j$.

Exemple :

travailler avec un marteau




| $\varphi'$ N° | Moyen | Activité musculaire | Etat moral | Action | ... | ... | Instrument | Maniabilité facile | ... |
|---|---|---|---|---|---|---|---|---|---|

$\varphi'$ 1er terme 2ème terme

Nous voyons que pour définir un $\varphi'$ nous avons besoin de deux termes "véritables" c'est-à-dire ayant chacun un code sémantique complet. C'est le raison pour laquelle nous considérons d'une façon particulière les prépositions. En effet, nous avons vu que celles-ci n'étaient qu'un mot exprimant le rapport entre deux autres mots. Or ce"mot de liaison" peut disparaitre en langue cible.

Donc, si nous considérions des $\varphi'(T_i^{\varphi}, T_j)$ ce $\varphi'$ pourrait être représenté en langue cible par un seul terme. C'est ce que nous avons voulu éviter.

Ainsi défini, chaque $\varphi'$ est appliquée en langue source et en langue cible à deux termes ordonnés. Et dans ces conditions, nous pourrons bien parler de la conservation des $\varphi'$ au cours de la transition "source—cible".

Exemple :

Il travaille avec ardeur

$S_i$ $S_k^{\varphi} = \varphi^{\varphi}(\ )$

a) $\varphi'(T_i, T_k^{\varphi})$

$\varphi'$ 1er terme : travaille 2è terme : avec ardeur

| N° | Complément manière | Activité | Etat | ... | Manière | ... |
|---|---|---|---|---|---|---|

En anglais comme langue cible, on obtient : "He works eagerly"

$\varphi'$ 1er terme : works 2è terme : eagerly

| N° | Complement of manner | Activité | State | ... | Manner ... |
|---|---|---|---|---|---|

b) $\varphi'(T_i, T_j)$ où $T_i$ et $T_j$ sont tous deux élémentaires

Il travaille courageusement

$S_i$ $S_j$

$\varphi'$ 1er terme : travaille 2è terme : courageusement

| N° | complément manière | Activité | Etat | ... | Manière | ... |
|---|---|---|---|---|---|---|

En anglais comme langue cible, on obtient : "he works courageousely"

$\varphi'$ 1er terme : works 2è terme : courageousely

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| N° | complement of manner | Activity | state | ... | Manner | ... | |

c) $\varphi'(T_i, T_k)$ où un des deux termes est complexe

Il travaille le dimanche

$S_K = \varphi_P(S_L, S_J)$

$\varphi'$ 1er terme : travaille 2è terme : le dimanche

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N° | complément de temps | Activité | Etat | ... | Date | jour | ... | |

En anglais comme langue cible, on obtient : " He works during sunday

$S_L$ $S_K$

$\varphi'$ 1er terme : works 2è terme : during sunday

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| N° | complément of time | Activity | State | ... | Date | day | ... | |

d) $\varphi'(T_k, T_l)$ où les deux termes sont complextes

Il travaille peu le dimanche

$S_K$ $S_\ell$

$\varphi'$ 1er terme : travaille peu 2ème terme : le dimanche.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| N° | Complément temps | Activité | Etat | ... | Date | jour | ... |

Dans ces exemples avec les termes ordonnés, on voit qu'il sera nécessaire de préciser l'évolution du code sémantique et en particulier des éléments notionels d'un terme quand celui-ci passe de l'état "terme ordonné élémentaire" à l'état "terme ordonné".

Dans ce but, on utilisera les renseignements sur l'importance comparée des deux termes, renseignements qui auront été fournis dans le passage syntaxique du syntagme élémentaire en syntagme

## 3- Représentation des $\varphi$ donnant un $\varphi'$ véritable

Ainsi nous éliminons de suite les $\varphi$ de la forme suivante :

$\varphi(S_i^{\varphi}, S_j)$ où la $K_{u,v}$ de $S_i^{\varphi}$ est celle des prépositions.

a) Considérons $\varphi(S_i, S_j)$ où $S_i$ et $S_j$ sont tous deux élémentaires
La représentation sera la suivante

| | $\varphi$ | 1er terme | 2è terme |
|---|---|---|---|
| N° $\varphi$ | | Renseignements syntaxiques du 1er terme | Renseignements syntaxiques du 2è terme |

Exemple :

Il travaille courageusement

$S_i$ $S_j$

| | $\varphi$ | 1er terme | | | 2è terme | |
|---|---|---|---|---|---|---|
| N° | verbe Adverbe | Verbe | Transitif | ... | Adverbe | ... |

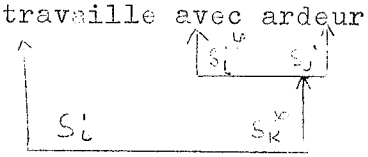
b) Soit $\varphi(S_i, S_k^{\varphi})$ où $S_k^{\varphi} = \varphi(S_i^{\varphi}, S_j)$

la représentation sera la suivante :

| | | 1er terme | | 2éterme |
|---|---|---|---|---|
| N° $\varphi$ | / | Renseignements syntaxiques du 1er terme | $S_i^{\varphi}$ | Renseignements syntaxiques . / $S_j$ |

Exemple :

Il travaille avec ardeur



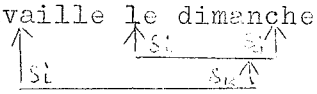

| | $\varphi$ | 1er terme | | 2ème terme | |
|---|---|---|---|---|---|
| N° | Verbe --- | Verbe | Transitif... | avec | Sujet commun... |

c) Considérons $\varphi(S_i, S_k)$ où l'un des deux est complexe ou le cas $\varphi(S_k, S_l)$ où les deux sont complexes. Nous allons encore avoir la représentation syntaxique classique :

| | $\varphi$ | 1er terme | 2ème terme. |
|---|---|---|---|
| N° $\varphi$ | | Renseignements syntaxiques du 1er terme | Renseignements syntaxiques du 2ème terme |

Exemple :

Il travaille le dimanche




| | $\varphi$ | 1er terme | | | 2ème terme | |
|---|---|---|---|---|---|---|
| N° | verbe Compt de temps | verbe | Transitif | ... | Substantif | commun... |

d) D'une façon générale, nous adopterons la représentation suivante où $S_i^{\varphi}$ existera réellement ou non.

| | | 1er terme | 2ème terme |
|---|---|---|---|
| N° $\varphi$ | | Renseignements syntaxiques du 1er terme | $S_i^{\varphi}$ Renseignements syntaxiques du 2ème terme |

Pour que $\varphi_i$ soit différent de $\varphi_j$, il faut et il suffit que , soit $S_i^{\varphi}$ soit un des renseignements syntaxiques du premier ou du second terme soit différent pour $S_i$ et $S_j$.

## V-. LIAISONS ENTRE LES $\varphi$ ET LES $\varphi'$ DANS UNE LANGUE

Une langue peut être considérée de deux points de vue soit comme langue source soit comme langue cible.

a) Prenons la comme langue source alors on a le passage de $\varphi$ à un $\varphi'$ issu de $\varphi_i \langle \varphi'_i \rangle$ et seule une étude sémantique permettra de choisir la relation ici employée.

Il se peut que l'on ait la correspondance simple : $\varphi \rightarrow \varphi'$

Exemple :

| | | | | | | |
|---|---|---|---|---|---|---|
| $\varphi$ N° | Verbe | transitif ∪ intransitif | ... | Substantif | commun | accu.. |
| $\varphi'$ N° | | | | Mesure espace | Mesure poids grandeur | Mesure temps ... |

mais cette correspondance peut être aussi plus complexe :

Exemple : Ici $\varphi_i\langle\varphi'_i\rangle$ devient $\varphi \| \varphi'_1 | \varphi'_2 \|$

| | | | | | | |
|---|---|---|---|---|---|---|
| $\varphi$ N° | Verbe | transitif | ... | Subst. | Commun | Acc... |

| | | | | | |
|---|---|---|---|---|---|
| $\varphi'_1$ N° | | | Objet | ... | |
| $\varphi'_2$ N° | Mouvement | ... | Objet | Lieu | ... |

b) Considérons la langue comme langue cible Alors on a le passage de $\varphi'$ à un $\varphi$ issu de $\varphi'_j\langle\varphi'_j\rangle$ et de même seule une étude sémantique permettra de choisir la relation $\varphi_p$ qui sera ici à employer. Cette sélection sera faite par un "critère sémantique". Celui-ci précisera quels éléments notionels devront être trouvés dans ceux du premier et du second terme pour que telle $\varphi_p$ soit choisie.

Il se peut qu'il n'y ait besoin d'aucun critère sémantique, c'est le cas où à un $\varphi'_p$ correspond un seul $\varphi_p$ . Nous le verrons dans l'exemple I. Souvent le critère sémantique porte seulement sur le second terme : Nous le verrons dans l'exemple II. Il est plus rare d'avoir besoin des éléments notionels du premier et du second terme. De toutes façons nous admettons qu'il y a conservation des $\varphi'$ au cours du transfert langue source $\longrightarrow$ langue cible : c'est-à-dire qu'à tout $\varphi'$ source correspond un et un seul $\varphi'$ cible avec les mêmes éléments notionels. Il n'est donc pas nécessaire de préciser si l'on est en langue source ou en langue cible et dans les deux cas, la relation sera désignée par le seul symbôle $\varphi'$.

Exemple N° 1 :

Souvent nous avons le passage direct $1\varphi' \Longrightarrow 1\varphi$ cible.
Ainsi considérons le $\varphi$ source suivant :

| | | | | | |
|---|---|---|---|---|---|
| N° $\varphi$ source | Verbe... | πoΔ "pod" | Subst. | Accusatif | ... |

Il lui correspond trois $\varphi'$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| N° $\varphi'_1$ | Mouvement | Etat | | Surface | endroit | nom Abstrait | .... |
| N° $\varphi'_2$ | Décision | Assignation | .. | | | | |
| N° $\varphi'_3$ | | | | Nom personne | ... | | |

Cherchons par quelle $\varphi_p$ chaque $\varphi'$ est représentée en langue cible, nous avons :

| | | | | |
|---|---|---|---|---|
| $\varphi'_1 \rightarrow$ | $\varphi$ N° | Verbe ... | sous | Substantif... |
| $\varphi'_2 \rightarrow$ | $\varphi$ N° | Verbe ... | pour | Substantif... |
| $\varphi'_3 \rightarrow$ | $\varphi$ N° | Verbe ... | à la manière de | Substantif propre... |

Exemple N° 2 Ici le passage $\varphi' \Longrightarrow \varphi$ cible sera plus complexe et nous aurons un choix à faire entre les différents $\varphi$ cible.

Considérons le $\varphi'$ suivant

| | 1er terme | | | | 2ème terme | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| N° $\varphi'$ | Compl$^t$ aboutis-sement | mouvement | action dynami-que | ... | Pays | Ville | Surface | Espace fermé | milieu | Disposition morale |

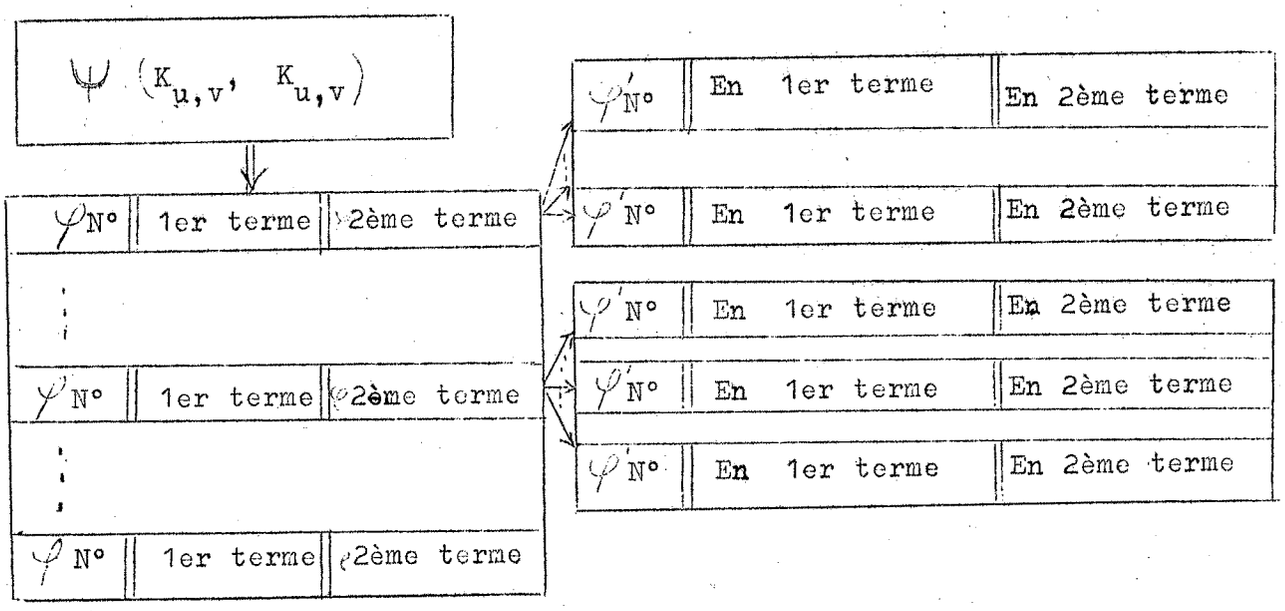
A ce $\varphi'$ correspond quatre $\varphi$ cible distincts :

| | critère sémantique | | 1er terme $\varphi$ cible | | 2ème terme | |
|---|---|---|---|---|---|---|
| N° $\varphi'$ aboutissement | 2ème terme : | ville | N° α | Verbe... | à | substantif |
| | " | pays | N° β | verbe... | en | substantif |
| | " | surface | N° γ | verbe... | sur | substantif |
| | " | espace fermé | N° δ | verbe... | dans | substantif |
| | " | milieu | N° δ | verbe... | dans | substantif |
| | " | disposition morale | N° δ | verbe... | dans | substantif |

## VI-. TABLEAU GENERAL D'ETUDES DES $\varphi$ ET DES $\varphi'$

Dans le cadre des travaux du C.E.T.A. de GRENOBLE, nous avons cherché à déterminer un grand nombre de relations syntaxiques $\varphi_p$ et de relations sémantiques $\varphi'_p$. Cette étude s'effectuait sous forme de réalisations de grands tableaux formés de la façon suivante : (Cf. page suivante).

a) Présentation du tableau On se place d'abord à l'intérieur d'un $\Psi(K_{u,v}, K_{u,v})$. A ce $\Psi$ correspond un ensemble de $\varphi$. Nous avons donc $\Psi_i \langle \varphi_L \rangle$ De même, nous savons qu'à chaque $\varphi$ correspond un ensemble de $\varphi'$ soit $\varphi_L \langle \varphi'_L \rangle$. Tout ceci peut être schématisé par le tableau suivant :




Nous allons donner deux exemples d'études.
Le 1er où les prépositions interviennent dans les $\varphi$
le 2è où elles n'interviennent pas.

Ces exemples sont tirés de l'étude de la traduction du russe en français ; la langue source est donc le russe et la langue cible est le français.

b) Exemple I : $\Psi$ (verbe ..., groupe prépositionnel)

A ce $\Psi$ correspond un ensemble de $\varphi(S_i, S_k^{\varphi})$

| | | |
|---|---|---|
| $\varphi$ | $S_i$ | $S_k^{\varphi} = (S_i^{\varphi}, S_j)$ |
| | Verbe.. | prépo + Substantif |
| | $K_4$ | $S_i^{\varphi}$ $S_j$ |

Parmi ceux-ci, on étudie ici deux $\varphi$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\varphi_0$ N° | verbe trans. ou intrans. | ... | B | substantif | ... | Acc. | .... |
| $\varphi_1$ N° | Verbe trans. | ... | B | substantif | ... | Acc. | .... |

et à ces correspondent des $\varphi'$ $\varphi_{0i}(\varphi'_{0,i})$ et $\varphi_{1i}(\varphi'_{1,i})$

| $\varphi'_{0,i}$ | En 1er terme | | | En 2ème terme | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\varphi'_{0,1}$ Compl$^t$ aboutissement | mouvement | Action dynamique | ... | pays | ville | surface | Espace fermé | milieu | Disposition morale |
| $\varphi'_{0,2}$ Compl$^t$ résultat | Transformation | Action transformation | ... | Substance | Etat physique | .... | | | |
| $\varphi'_{0,3}$ Compl$^t$ jour | | | | nom jour | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| $\varphi'_{1,1}$ Compl$^t$ temps (délai) | Action qui dûre | ... | | Subdivision de temps dans sa durée | ... |

Pour chaque $\varphi'$ on recherche les $\varphi$ cible qui lui correspondent et les critères sémantiques qui permettent de les différencier. On a en effet l'ensemble $\varphi' \underset{j}{\langle} \varphi_j \text{ cible} \rangle$

| $\varphi'$ | critère sémantique | | $\varphi$ cible | | | |
|---|---|---|---|---|---|---|
| $\varphi'_{0,1}$ Compl$^t$ aboutissement | 2ème terme | pays | $\varphi$ N° | Verbe | en | substantif |
| | | ville | $\varphi$ N° | Verbe | à | substantif |
| | | surface | $\varphi$ N° | verbe | sur | substantif |
| | | espace fermé | $\varphi$ N° | verbe | dans | substantif |
| | | milieu | | | | |
| | | Disposition morale | | | | |
| $\varphi'_{0,2}$ Compl$^t$ résultat | | | $\varphi$ N° | verbe | en | substantif |
| $\varphi'_{0,3}$ Compl$^t$ jour | | | $\varphi$ N° | verbe | substantif | |
| $\varphi'_{1,1}$ Compl$^t$ temps (délai) | | | $\varphi$ N° | verbe | en | un + substantif |

c) Exemple II :

$\psi$ (Verbe ... , Substantif ...)

A ce $\psi$ correspond un ensemble de $\varphi(S_i, S_j)$

| $\varphi$ | $S_i$ verbe ... $K_4$ | $S_j$ substantif... $K_1$ |
|---|---|---|

Parmi ceux-ci, on étudie ici les deux $\varphi$ suivants :

| | | | | |
|---|---|---|---|---|
| $\varphi_2$ N° | Verbe transitif | substantif | Acc. | ... |
| $\varphi_3$ N° | Verbe transitif $\cup$ intransitif | substantif | Acc. | .... |

A ces $\varphi$ correspondent des $\varphi'$ $\varphi_2 \underset{i}{<} \varphi'_{2,i} >$ et $\varphi_3 \underset{i}{<} \varphi'_{3,i} >$

| $\varphi'_{2,i}$ | En 1er terme | | En 2ème terme | | |
|---|---|---|---|---|---|
| $\varphi'_{2,1}$ Compl$^t$ objet | | | objet | ... | |
| $\varphi'_{2,2}$ compl$^t$ objet lieu | mouvement | | objet | lieu | ... |

| | | | | | |
|---|---|---|---|---|---|
| $\varphi'_{1,3,1}$ Compl$^t$ mesure | | Mesure espace | Mesure poids grandeur | Mesure temps | ... |

Pour chaque $\varphi'$ on recherche comme précédemment les $\varphi$ cible et les critères sémantiques de sélection.

| $\varphi'$ | critère sémantique | $\varphi$ cible | | |
|---|---|---|---|---|
| $\varphi'_{2,1}$ Compl$^t$ objet | | $\varphi$ N° | Verbe | substantif |
| $\varphi'_{2,2}$ Compl$^t$ objet-lieu | 1er terme dictionnaire | $\varphi$ N° | Verbe.. | substantif |
| | | $\varphi$ N° | Verbe.. | Préposition Substantif |
| $\varphi'_{3,1}$ Compl$^t$ mesure | | $\varphi$ N° | Verbe.. | Substantif... |

Signification du critère sémantique : 1er terme dictionnaire

C'est le dictionnaire qui nous indique le $\varphi$ de sortie d'après le 1er terme. Si le verbe français correspond rigoureusement au verbe russe, lui même obtenu par préverbation alors on utilise le 1er $\varphi$.

S'il n'existe pas de verbe français correspondant au verbe russe, on utilise le 2ème $\varphi$.

Exemple :

ПЕРЕ/НОСИТЬ
p'er'e nosit' ——→ "transporter" et le 1er $\varphi$

ВЫ/БЕЖАТЬ
vyb'ezat' ——→ "courir hors de" et le 2ème $\varphi$

---

- IV -

ETUDE DU PASSAGE MOT "SOURCE" ⟶ MOT "CIBLE"

## I-. PASSAGE D'UNE Ul de LA LANGUE SOURCE A UNE Ul DE LA LANGUE CIBLE

Si l'on cherche à traduire l'Ul "glace" en anglais et en russe, on obtient le tableau suivant :

| cible | source | cible |
|---|---|---|
| Anglais ⟵ | Français | ⟶ russe |
| ice | - $U_s^1$ glace (eau gelée) | ЛЕД "l'od" |
| ice-cream | - $U_s^2$ glace (sorbet) | МОРОЖЕНОЕ "mororenoé" |
| mirror | - $U_s^3$ glace (miroir) | ЗЕРКАЛО "zerkalo" |
| window | (vitre) | ОКНО "okno" |

On remarque que l'$U_s^3$ "glace" se traduit en anglais soit par "mirror" soit par "window" et en russe, soit par ЗЕРКАЛО soit par ОКНО.

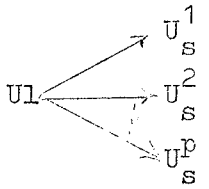
Ceci est général, une unité sémantique peut souvent être traduite en langue cible par plusieurs mots différents, le choix entre ceux-ci ayant lieu à partir du contexte.

Exemple :

| Langue source | langue cible | |
|---|---|---|
| $U_s$ moeurs | manners | (gens) |
| | customs | (pays) |
| | habits | (animaux) |

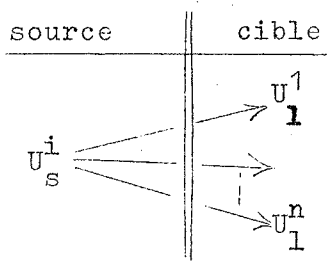
Nous voyons que la traduction en langue cible d'une Unité lexicale de la langue source, se décompose en deux parties bien distinctes :

en langue source on a



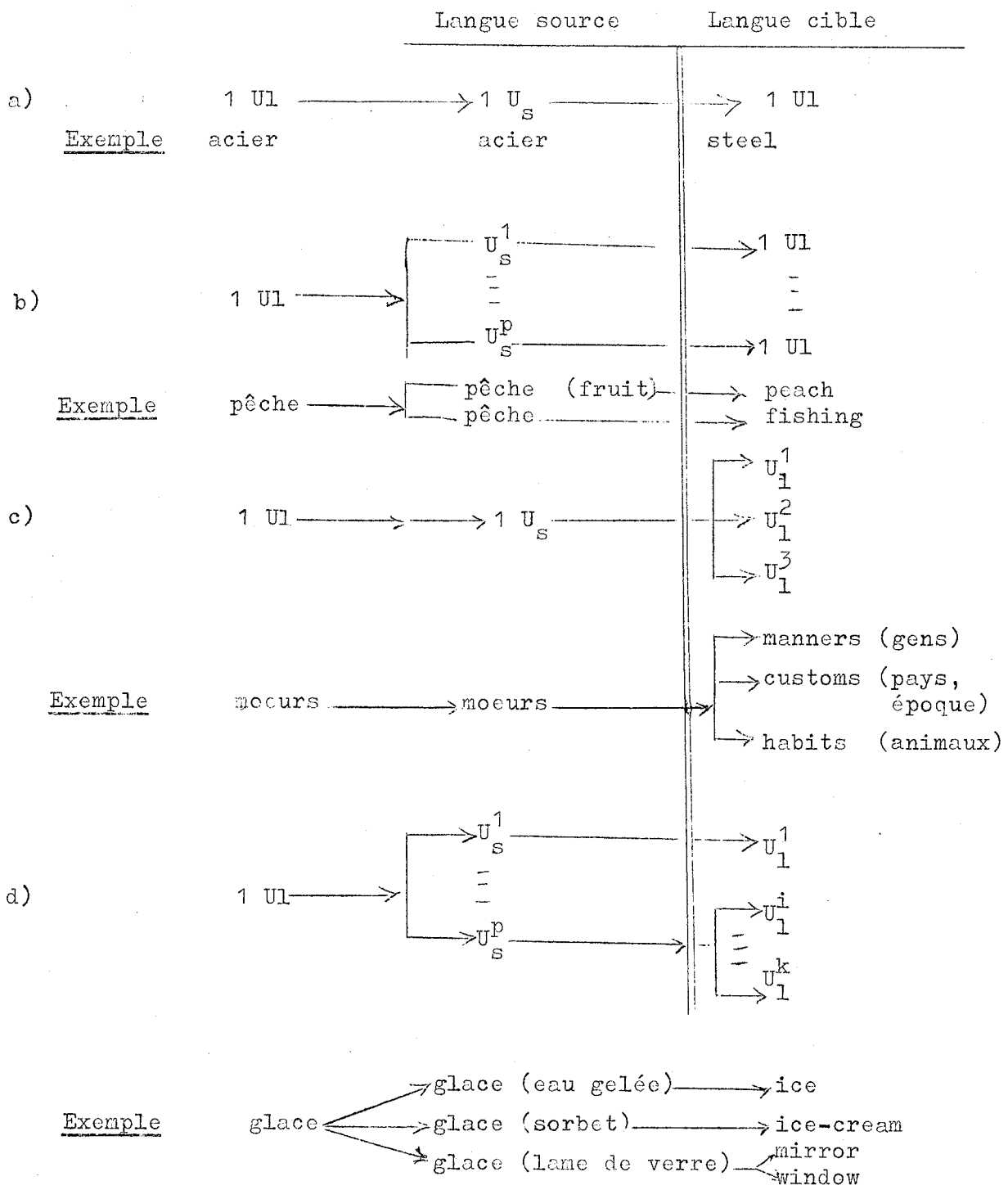

Il faudra en langue source, repérer l'unité sémantique $U_s^i$ ici employée

puis on a le passage

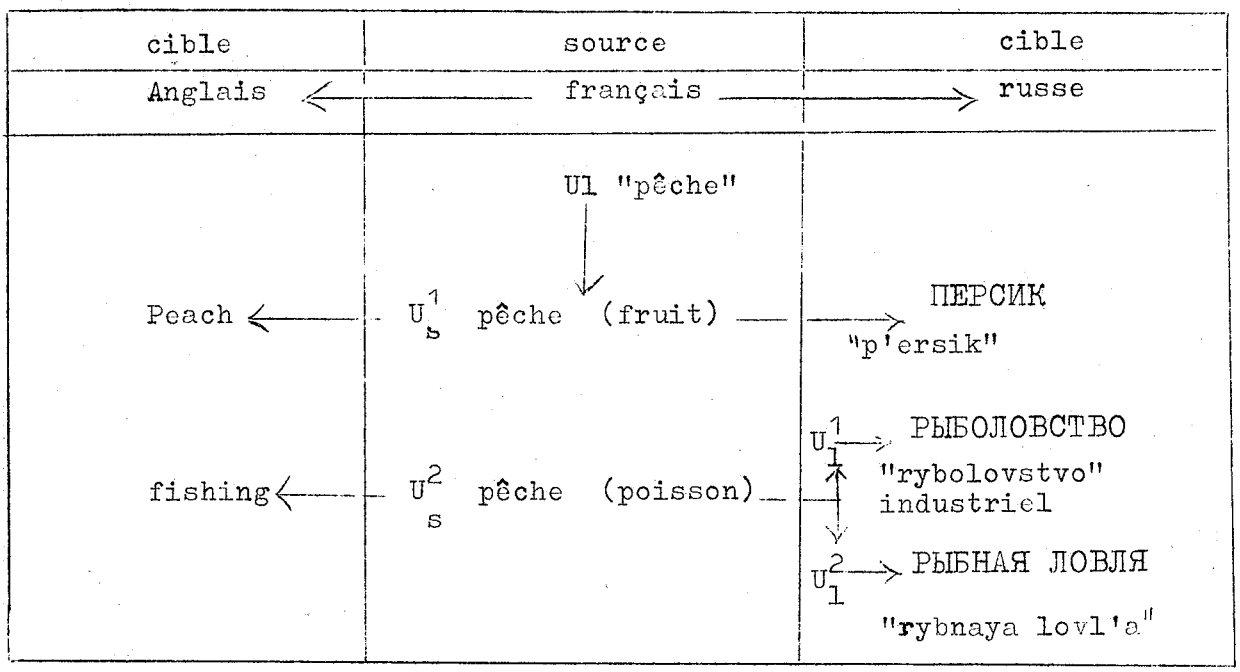



Il faudra au cours du transfert source → cible repérer d'après le contexte, l'unité lexicale cible $U_l^i$ à employer.

Toutes les combinaisons de ces deux types sont possibles. Dans les représentations schématiques, on part toujours de la langue source pour aller vers la langue cible. On peut avoir :


Langue source
Langue cible
a)
1 Ul
1 $U_s$
1 Ul
Exemple
acier
acier
steel
b)
1 Ul
$U_s^1$
$U_s^p$
1 Ul
1 Ul
Exemple
pêche
pêche (fruit)
pêche
peach
fishing
c)
1 Ul
1 $U_s$
$U_l^1$
$U_l^2$
$U_l^3$
Exemple
moeurs
moeurs
manners (gens)
customs (pays, époque)
habits (animaux)
d)
1 Ul
$U_s^1$
$U_s^p$
$U_l^1$
$U_l^i$
$U_l^k$
Exemple
glace
glace (eau gelée)
ice
glace (sorbet)
ice-cream
glace (lame de verre)
mirror
window

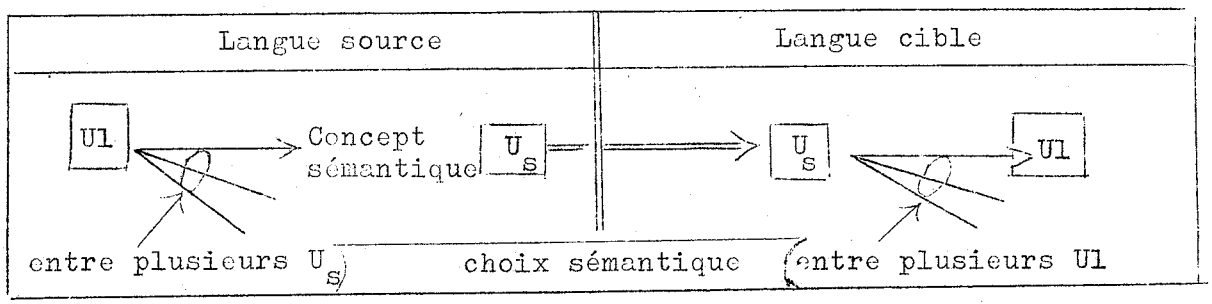
Cette décomposition en deux étapes de la traduction d'une unité lexicale présente le gros intérêt suivant : la première étape est indépendante de la langue de sortie choisie, donc son étude sera la même quelle que soit cette langue cible. La seconde étape peut être différente suivant la langue de sortie. Prenons un exemple :




Dans ce cas, la première étape consistera à reconnaitre l'unité sémantique "pêche" ici utilisée.

La deuxième étape : celle du transfert source cible sera très simple en anglais, mais nécessitera encore un choix sémantique en russe.

Le passage d'une Ul de la langue source à une Ul de la langue cible peut donc se représenter par le schéma suivant :



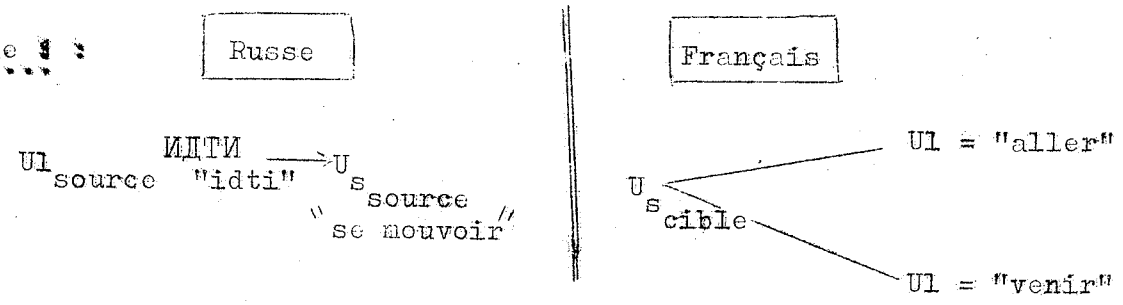
Mais ceci suppose la conservation d'une $U_s$ dans le passage langue source-langue cible.

Qu'entendons-nous par cette expression ?

## II-. CONSERVATION D'UNE $U_s$ DANS LE PASSAGE SOURCE-CIBLE

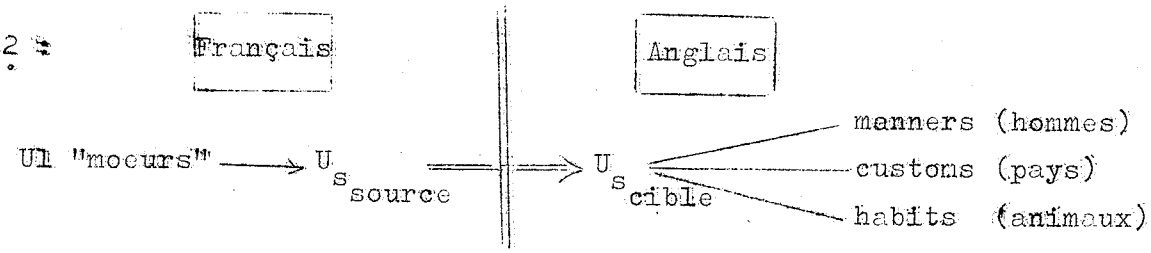
Ceci ne signifie pas que le concept sémantique attaché à l'$U_s$ source est lui-même considéré comme un concept sémantique en langue cible. Mais il est soit contenu dans un concept sémantique cible
- soit contenant un concept sémantique **cible**

Exemple 1 :

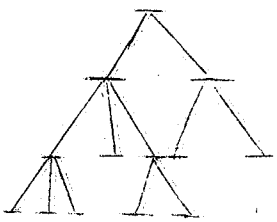



Le concept source "se mouvoir" est contenu dans chacun des concepts cible "aller" et "venir"

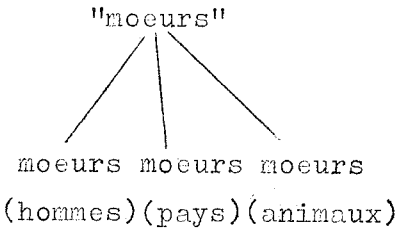
Exemple 2 :



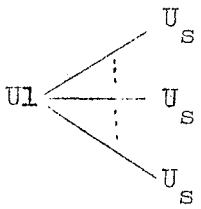

Le concept source "moeurs" contient les trois concepts cible anglais. On peut représenter suivant une pyramide les concepts sémantiques de toutes les langues qui sont attachés à une "idée".



Exemple :

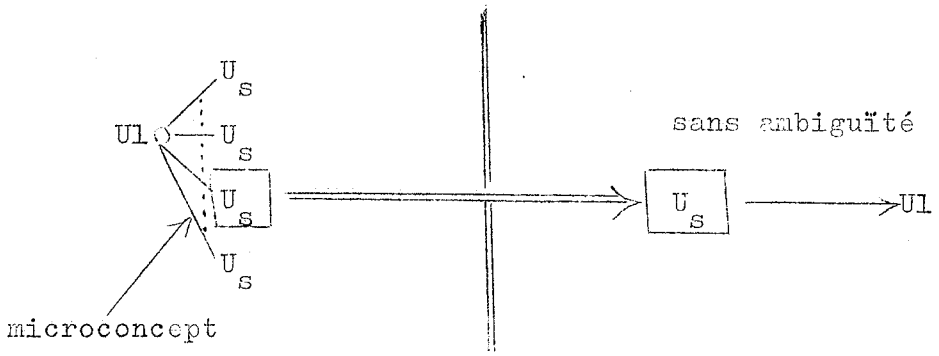



La solution idéale serait de connaître pour toutes les langues les "microconcepts" attachés à chaque "idée". Alors à chaque "micro concept" on ferait correspondre une nouvelle unité sémantique. En langue source on aurait donc :




On sélectionnerait l'$U_s$ ici employée. Celle-ci se trouverait bien conservée au cours du transfert source-cible, et cette $U_s$ serait représentée par une seule Ul en langue cible. Ainsi on n'aurait qu'un seul choix à faire pour passer d'un $U_l$ source à un $U_l$ cible. Ce choix serait certes très délicat mais unique. L'ensemble de ces "microconcepts" qui donnent naissance à des $U_s$, formerait le vocabulaire de la langue de passage Internationale.

Ce passage pourrait se représenter par le schéma suivant :



Exemple : Si l'on veut traduire "les moeurs des castors". Une étude sémantique du mot "castor" montre que pour "moeurs" on a le microconcept "moeurs d'animaux " auquel correspond une $U_s$. Celle-ci est conservée dans le passage français-anglais, et à l' $U_s$ "moeurs d'animaux" correspond en anglais l'U1 "habits". Malheureusement, cette solution est utopique, car il est presque impensable de trouver un cerveau qui puisse connaître suffisamment de langues pour faire ce travail. Aussi est-on dans l'obligation d'adopter l'autre méthode déjà exposée.

Nous allons préciser cette méthode en examinant l'étude faite sur les verbes Russes au C.E.T.A. de GRENOBLE en collaboration avec le centre de MILAN.

## II.- METHODE EMPLOYEE POUR TROUVER LE VERBE CIBLE EN VUE DE LA TRADUCTION RUSSE-FRANCAIS

Nous utilisons ici des exemples donnés par Sergei PERSCHKE [9]

... Le verbe russe est analysé pour son contenu opérationel. On voit ainsi s'il a une ou plusieurs unités sémantiques. S'il n'en a qu'une, on regarde s'il existe un verbe français désignant le même concept et possédant les mêmes compléments. Si oui : tout va très bien; si non :

a) En associant les compléments on essaye de recréer par le verbe français le plus approché, la situation qui correspond à ce que le verbe russe indique.

Exemple : Soit le verbe ДАТЬ "dat' " il possède un concept sémantique auquel correspond une unité sémantique. Son concept est : "passage de quelque chose d'une personne à une autre". A ce concept correspondent en Français les verbes suivants : donner, laisser

Les compléments du verbe russe sont :

1- 1 Accusatif : chose qui passe
2- 1 Datif : celui qui reçoit
3- 1 Infinitif : Activité exercée par celui qui reçoit.

Si on rencontre les constructions syntaxiques russes suivantes, on en déduit le verbe français à utiliser.

Si : verbe + acc. + datif ⟹ donner

Si : verbe + Inf. + datif ⟹ laisser

b) Nous n'avons jusqu'ici fait appel qu'à des critères syntaxiques, mais un examen de l'association des compléments est parfois insuffisant. Il faut considérer leur contenu sémantique.

Exemple : Soit le verbe ПОВЕСИТЬ "povesit'" et la construction syntaxique "Verbe + acc."; suivant le contenu sémantique du mot à l'accusatif, on déduira le verbe français de traduction.

Mot à l'accusatif → être humain ⟹ pendre
Mot à l'accusatif → autre chose ⟹ suspendre

Ainsi l'analyse syntaxique seule ne permet pas toujours de lever les ambiguïtés rencontrées en langue cible et de sélectionner le verbe de sortie adéquate. Il faut alors faire appel à l'étude sémantique.

c) Un petit nombre seulement de verbes russes ont plusieurs unités sémantiques.

Exemple : Le verbe ТОПИТЬ "topit'" possède trois concepts sémantiques, c'est-à-dire trois $U_s$. Il peut indiquer :

$U_s^1$ 1) Le chauffage des maisons, pièces ...

$U_s^2$ 2) La fonte des substances à basse température

$U_s^3$ 3) L'immersion dans les liquides.

Les verbes français choisis comme verbes "cible" sont les Ul suivantes :

A) Chauffer ; B) fondre ; C) couler ; D) noyer.

$U_s^1$ correspond à A ; $U_s^2$ correspond à B ; $U_s^3$ correspond à C et D.

Pour déterminer les $U_s$ et les verbes cible, les compléments directs sont divisés en groupes de variables sémantiques :

1) Batiments, parties de batiments, poêles...
2) Choses fondant à basse température
3) Bateaux,...
4) Etres vivants, douleur, sentiments,...

les $U_s$ et les verbes cible sont déterminés de la façon suivante :

| | | | | |
|---|---|---|---|---|
| $U_s^1$ | et A | chauffer | : sont caractérisés par l'ensemble | : verbe + acc.1 |
| $U_s^2$ | et B | Fondre | : ⟶ | : Verbe + acc.2 |
| $U_s^3$ | et C | couler | : ⟶ | : verbe + acc.3 |
| | et D | Noyer | : ⟶ | : verbe + acc.4 |

d) Les autres verbes russes n'ont qu'une unité sémantique

1- Souvent l'étude sémantique des compléments est suffisante pour déterminer le verbe cible :

Exemple : БРОСИТЬ "brosit'"

Son concept est : Interruption d'une relation ou d'un contact
Ses compléments sonts : 1° accusatif
2° instrumental
3° infinitif.

Le complément direct à l'accusatif appartient à un des groupes suivants de variables sémantiques :

1°) Petits objets mobiles
2°) Etres humains ou animaux domestiques
3°) Endroits
4°) Activités.

Les verbes français choisis comme verbes "cible" sont les Ul suivantes :

A) Jeter
B) Abandonner
C) Cesser

Ces Ul sont déterminées de la façon qui suit :

A) Jeter

V + Acc. 1
V + Inst.
V + Acc.1 + Inst.
V + Acc.2 + direction

B) Abandonner

V + Acc.2
V + Acc.3
V + Acc.4

C) Cesser

V + inf.

2- Parfois à l'aide des seuls renseignements sémantiques des compléments, aucun verbe français ne peut exprimer même approximativement le verbe russe.

On doit alors faire aussi une étude sémantique du sujet, en plus de celle des compléments.

Exemple : ВЕЗТИ "vezti" Il possède deux concepts d'où 2 $U_s$

$U_s^1$ : déplacement d'une chose par un moyen de transport

$U_s^2$ : déplacement d'un moyen de transport.

Pour le sujet nous avons trois groupes de $V_s$ et pour le complément direct :quatre.

Ces groupes sont les suivants :

Sujet : 1 êtres humains
2 animaux
3 moyens de transport

Complément
direct - 1 moyen de transport mu en tirant
- 2 moyen de transport mu en poussant
- 3 chose transportable
- 4 êtres humains.

Nous avons 4 verbes "cible" possibles et nous obtenons le tableau de passage suivant :

$U_s^2$ ( $U_1$ A) Tirer ⟶ sujet 1 + V + Acc.1 / Sujet 2 + V + Acc.1
$U_1$ B) Pousser ⟶ sujet 1 + V + Acc.2 )

$U_s^1$ ( $U_1$ C) transporter ⟶ sujet 1 + V + $Acc_c$.3 / sujet 2 + V + Acc.3 / sujet 2 + V + Acc.4
$U_1$ D) conduire ⟶ sujet 1 + V + Acc. 4 )

Nous remarquons que si nous considérons la relations syntaxique :

| | 1er terme | | 2ème terme | | |
|---|---|---|---|---|---|
| φ N° | Verbe | ... | Subst. | Acc. | ... |

et les critères sémantiques du second terme, nous pouvons déjà sélectionner l'$U_s$ du verbe russe.

Nous avons en effet :

critères sémantiques du complément direct

| | | |
|---|---|---|
| 1 | Moyen de transport mu en tirant | $U_s^2$ |
| 2 | Moyen de transport mu en poussant | |
| 3 | Choses transportables | $U_s^1$ |
| 4 | Etres humains | |

Après la sélection de l'$U_s$ source à laquelle on fait correspondre une $U_s$ cible, cherchons le moyen de sélectionner l'U1 cible.

Soit le φ

| φ N° | verbe | ... | Substantif | Acc. | ... |
|---|---|---|---|---|---|

il permet le passage suivant le 2ème terme :

A $U_s^2$ cible correspond

| | |
|---|---|
| Subst. Acc 1 | → U1 A) |
| Subst. Acc 2 | → U1 B) |

De même considérons la construction syntaxique et le passage suivant :

A $U_s^1$ cible correspond

| | | |
|---|---|---|
| Sujet + Verbe + subst. Acc. | | |
| Sujet 2 | | → U1 C) |
| Sujet 1 | Acc 3 | |
| Sujet 1 | Acc 4 | → U1 D) |

Ainsi toute ambiguïté a pu être levée et l'U1 cible a été sélectionnée.

D'après cette étude, nous pouvons maintenant indiquer ce qu'il est nécessaire de trouver dans le dictionnaire pour faire le choix de l'Ul cible.

## V-. CONTENU DU DICTIONNAIRE POUR L'OBTENTION DE L'UI CIBLE

Considérons une Ul source et nous voulons d'abord sélectionner.

a) L'$U_s$ source Ceci se fait dans l'étude de la langue source et ce choix est tout à fait indépendant de la langue dible ; nous pourrons donc mettre dans le dictionnaire les indications suivantes qui seront immuables :

1°) Le nombre des $U_s$ : $U_s^1$ , $U_s^2$ , ... $U_s^n$ .

2°) Le $\varphi$ à considérer pour choisir l'unité sémantique. Ceci n'est pas toujours possible, mais l'est pour les verbes.

3°) Le terme de la relation $\varphi$ (1er ou 2ème terme) qui est occupé par l'Ul dont on cherche l'$U_s$.

4) Les variables sémantiques caractéristiques du terme auquel est lié par $\varphi$ l'unité sémantique recherchée.

Ceci peut se schématiser de la façon suivante :

| "Ul" | nombre de $U_s$ | $\varphi$ N° | 1er terme X | 2ème terme | |
|---|---|---|---|---|---|
| puis | $V_s$ caractéristiques | $U_{s1}$ | ... | $V_s$ caractéristiques | $U_{sn}$ |

Exemple :

| ТОПИТЬ "topit' " | 3 $U_s$ | | N° $\varphi$ | 1er terme Verbe X | 2ème terme Subs. Acc. |
|---|---|---|---|---|---|

puis

| batiment, parties de batiments... | | $U_s$ 1 | choses fondant à basse tempé. | ... | $U_s$ 2 |
|---|---|---|---|---|---|
| bateaux | ... Etre vivant | sentiments | | $U_s$ 3 | |

Remarque : Dans le cas, où l'on ne peut pas indiquer de $\varphi$ à considérer, les $V_s$ caractéristiques de chaque $U_s$ seront bien moins précises.

b) Nous admettons la conservation $U_{s\ source} \longrightarrow U_{s\ cible}$ et nous voulons sélectionner l'Ul cible.

Si l'analyse syntaxique permet à elle seule de faire le choix de l'Ul cible, alors il faut indiquer.

1°) Le nombre d'Ul cible possibles
2°) Les $\varphi$ à considérer
3°) le terme de la relation $\varphi$ (1er ou 2eme terme) qui est occupé par l'$U_s$ que l'on étudie
4°) les relations syntaxiques caractéristiques de chaque Ul cible.

On le schématise comme suit :

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| " $U_{s_i}$ " | Nombre d'Ul cible | | $\varphi$ N° | 1er terme X | 2ème terme | $U_{l_1}$ | ... | $\varphi$ | 1er terme X | 2ème terme | $U_{l_n}$ |

Ceci est le cas de certains verbes comme ДАТЬ "dat' "

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ДАТЬ "dat' " | 2 Ul cible | | $\varphi$ N° | 1er terme verbeX | 2ème terme Subs. Acc... | donner | $\varphi$ N° | 1er terme verbe X | 2ème terme Verbe Inf. | laisser. |

Mais dans la plupart des cas, l'analyse syntaxique seule sera insuffisante, alors il faut indiquer :

1°) Le nombre d'Ul cible possibles
2°) La relation syntaxique $\varphi$ à considérer (si elle existe)
3°) Le terme de la relation $\varphi$ qui est occupé par l'$U_s$ que l'on étudie
4°) Les variables sémantiques caractéristiques des termes auxquels l'Ul cible cherchée est reliée par l'intermédiaire de $\varphi$

On le schématise de la façon suivante :

| | | | | |
|---|---|---|---|---|
| "$U_{s_i}$" | nombre d'Ul cible | $\varphi$ N° | 1er terme X | 2ème terme |

| | | | | |
|---|---|---|---|---|
| $V_s$ caractéristiques | $Ul_1$ | .... | $V_s$ caractéristiques | $Ul_n$ |

Ce cas est celui de bien des verbes :

Exemple :

| | | | | |
|---|---|---|---|---|
| ПОВЕСИТЬ "povesit" | 2 Ul cible | $\varphi$ N° | 1er terme verbe X | 2ème terme subs. Acc. |

| | | | |
|---|---|---|---|
| être humain | pendre | différent d'être humain | suspendre. |

Il se peut que le $\varphi$ à considérer soit bien plus complexe, et possède trois termes dont deux auront des $V_s$ caractéristiques. Ce sera le cas de verbes comme ВЕЗТИ "vezti" . Par contre, parfois il ne sera pas possible de préciser le $\varphi$ , alors les $V_s$ caractéristiques seront bien moins précises. De toute façon, on utilisera d'abord le plus possible de $V_s$ indépendantes du champ et après les $V_s$ dépendantes du champ.

Tous les renseignements que nous venons de voir en b) dépendent de la langue cible choisie. Ils seront contenus dans une bande de transfert $source_i \rightarrow cible_j$ et si l'on change de langue cible, il faudra changer de bande de transfert.

## V-. SUPPLEMENT SYNTAXIQUE DANS LE DICTIONNAIRE POUR LE CAS OU L'U1 EST UN VERBE

Jusqu'ici on a utilisé la sémantique pour déterminer :

1- soit la bonne unité sémantique $U_s$

2- soit le choix de l'U1 cible.

Dans la détermination de l'U1 cible, on a considéré une construction syntaxique ; celle-ci existe dans la bande de transfert puisque nous considérons un verbe ; il doit être utile d'obtenir aussi la construction syntaxique du verbe cible avec ses compléments.

Donc, dans la bande de transfert, en plus de ce que nous avons vu, il doit y avoir un tableau permettant d'après la sélection du verbe cible le passage direct du $\varphi$ source au $\varphi$ cible.

Exemple :

Pour ВЕЗТИ "vezti" nous avons la construction syntaxique suivante : Sujet + verbe + substantif Acc.

on a quatre verbes cible possibles : A B C D donc quatre constructions syntaxiques qui dans ce cas particulier sont identiques :

sujet + verbe A + substantif
sujet + verbe B + substantif
sujet + verbe C + substantif
sujet + verbe D + substantif.

# CHAPITRE V

## UTILISATION DES ELEMENTS SEMANTIQUES POUR RESOUDRE DES DIFFICULTES RENCONTREES AUX CHAPITRES PRECEDENTS.

### I- Pluralité des Constructions Syntaxiques

a- Il existe plusieurs schémas syntaxiques possibles en Langue Source.

Rappelons que chaque unité lexicale est indiçée par son code sémantique :

$$C_{\sigma} = U_1 \; {}_j\langle U_{S_j} \mid E_n^{\,p} \ldots\ldots E_n^{\,q} \mid {}_i\langle \text{chi} \mid V_s^{\,1} \ldots V_s^{\,m} \mid \rangle \rangle$$

et qu'il existe en langue source un système de correspondance $\varphi_p \; {}_i\langle \varphi'_{p_i} \rangle$

1) Voyons à quelles conditions une relation sémantique $\varphi'_p$ peut être impossible entre deux termes.

1- Il peut y avoir incompatibilité entre les champs de ces deux termes. Nous faisons alors intervenir la "Matrice de Relations entre champs".

2- Une $\varphi'_p$ est caractérisée par certains Eléments Notionels, c'est-à dire :

pour qu'une relation sémantique $\varphi'_p$ existe entre deux termes, il est nécessaire que parmi les Eléments Notionels de chacun des termes, il en existe au moins un qui appartienne au groupe des $E_n$ caractéristiques de la $\varphi'_p$.

Pour chaque $\varphi'_p$ on a déterminé les $E_n$ caractéristiques du premier terme et les $E_n$ caractéristiques du second terme.

On a ainsi un tableau du type suivant :

| $\varphi'_p$ | $E_n$ caractéristiques 1° terme | | | $E_n$ caractéristiques 2° terme | | |
|---|---|---|---|---|---|---|
| —— | X | X | ..... | X | X | ..... |
| —— | X | X | ..... | X | X | ..... |
| —— | | | ------ | | | ------ |

Si pour l'un des deux termes, il n'existe pas un des $E_n$ pouvant coïncider avec un des $E_n$ caractéristiques, alors la relation $\varphi'_p$ est impossible.

## 2) Structure syntaxique de la Phrase

Une phrase (ou une portion de phrase) P est construite suivant une structure syntaxique utilisant un ensemble de relations $\varphi_p$.

Ici nous avons plusieurs constructions syntaxiques possibles. Ainsi la phrase P peut être représentée par plusieurs structures :

$$P_1 \langle \varphi^1_i \rangle_i \quad , \quad P_2 \langle \varphi^2_j \rangle_j \quad \ldots\ldots\ldots , \quad P_p \langle \varphi^M_k \rangle_k$$

Les relations $\varphi^1_i$ , $\varphi^2_j$ ....., $\varphi^M_k$ peuvent être parfois identiques, mais il existe au moins une $\varphi^1_i$ différente d'une $\varphi^2_j$ , différente d'une $\varphi^M_k$.

Une $\varphi^M_k$ est caractérisée par la relation syntaxique $\varphi_{k'}$ mais aussi par les deux termes qu'elle relie. Aussi pour que $\varphi^M_k$ soit différent de $\varphi^V_k$ , il suffira que l'un des deux termes soit différent. En réalité il se peut qu'ils soient légèrement différents, car nous voyons réapparaître les difficultés de passage du terme élémentaire au terme complexe ; et ce qui nous intéresse ici, c'est le terme "principal" qui a formé le terme complexe.

Ce sont ces termes "principaux" qui doivent être différents pour que $\varphi^M_k$ soit différent de $\varphi^V_k$

Reprenons l'exemple vu au chapitre I :

Schéma 1 : Une Directrice de Société enrhumée

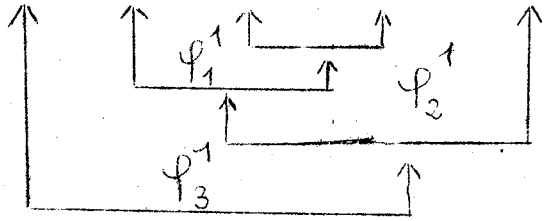
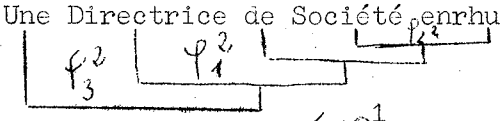



Schéma 2 : Une Directrice de Société enrhumée




Pour le schéma 1, on a $P_1 \langle_i \varphi_i^1 \rangle = P_1 \| \varphi_1^1 | \varphi_2^1 | \varphi_3^1 \|$

Pour le schéma 2, on a $P_2 \langle_j \varphi_j^2 \rangle = P_2 \| \varphi_1^2 | \varphi_2^2 | \ {}_3^2 \|$

Précisons les $\varphi_p$ :

| | | | | | |
|---|---|---|---|---|---|
| N° $\varphi_1$ | Subst | .... | de | Subst | .... |
| N° $\varphi_2$ | Adjectif | | Subst | ..... | |
| N° $\varphi_3$ | Article | ... | Subst | ..... | |

Nous voyons qu'ici $\begin{cases} \varphi_1^1 = \varphi_1^2 \\ \varphi_2^1 \neq \varphi_2^2 \\ \varphi_3^1 = \varphi_3^2 \end{cases}$

3) Continuons le Raisonnement Général, en nous limitant par mesure de simplicité au cas de deux constructions syntaxiques différentes.

Nous avons donc $P_1 \langle_i \varphi_i^1 \rangle$ et $P_2 \langle_j \varphi_j^2 \rangle$

Introduisons maintenant les relations sémantiques $\varphi'_p$.

Nous avons $\varphi_i^1 \langle_{i'} \varphi'^1_{i'} \rangle$ et $\varphi_j^2 \langle_{j'} \varphi'^2_{j'} \rangle$

Recherchons d'abord les $\varphi^1_i$ différents des $\varphi^2_{j'}$

Notons les $\overset{\varphi}{\varphi}{}^1_i$ et $\overset{\varphi}{\varphi}{}^2_j$

Pour ces $\varphi_p$ seulement, nous allons étudier les relations sémantiques :

$$\overset{\varphi}{\varphi}{}^1_i \underset{i'}{<} \overset{'}{\varphi}{}^1_{i'} > \quad \text{et} \quad \overset{\varphi}{\varphi}{}^2_j \underset{j'}{<} \overset{'}{\varphi}{}^2_{j'} >$$

Mais nous avons vu que des relations sémantiques $\varphi'_p$ étaient parfois impossibles.

Supposons que l'on montre par ces considérations sémantiques, que toutes les $\varphi'^1_{i'}$ sont impossibles. Alors $\overset{\varphi}{\varphi}{}^1_i$ est impossible, donc $P_1$ aussi.

Nous venons ainsi d'éliminer une construction syntaxique. Dans le cas présent où il n'y en avait que deux, celle qui reste est la seule valable.

Mais si on avait n schémas $P_\mu$, on pourrait de la même façon opérer en les prenant deux par deux : $P_\mu$ et $P_\nu$ ; ce qui permettrait de déterminer les $\varphi^\mu_i$ différents des $\varphi^\nu_j$ et de continuer le raisonnement précédent.

On pourrait à chaque fois éliminer : 0, 1 ou 2 constructions syntaxiques.

<u>4) Continuons l'étude de l'exemple précédent.</u>

Nous avons trouvé $\varphi^1_2 \neq \varphi^2_2$

Nous passons alors à l'étude sémantique des deux relations.

<u>Schéma 1</u> Directrice enrhumée $\uparrow \varphi^1_2 \uparrow$ <u>Schéma 2</u> Société enrhumée $\uparrow \varphi^2_2 \uparrow$

Nous concevons fort bien une impossibilité de compatibilité entre les Champs auxquels "Société" appartient et ceux auxquels "enrhumée" appartient.

La relation $\varphi^2_2$ devient donc impossible et de Schéma 2 aussi. Il ne nous reste plus que la première construction syntaxique.

b- Il n'existe qu'un seul schéma syntaxique en langue source.

Mais celui-ci donne naissance à plusieurs schémas syntaxiques en langue cible.

Nous avons une construction syntaxique en Langue Source, celle-ci est représentée par une $\varphi_p$.

Mais cette $\varphi_p$ correspond à un ensemble de relations sémantiques : $\varphi_p \underset{i}{\langle} \varphi'_i \rangle$

Il nous faut rechercher quelle $\varphi'_p$ est représentée dans la phrase par la relation syntaxique $\varphi_p$.

Pour cela, nous allons regarder quelles $\varphi'_p$ sont possibles, entre les deux termes considérés.

Dnas ce but, nous allons utiliser, comme nous l'avons déjà vu, l'incompatibilité des champs puis les Eléments Notionels des $\varphi'_p$.

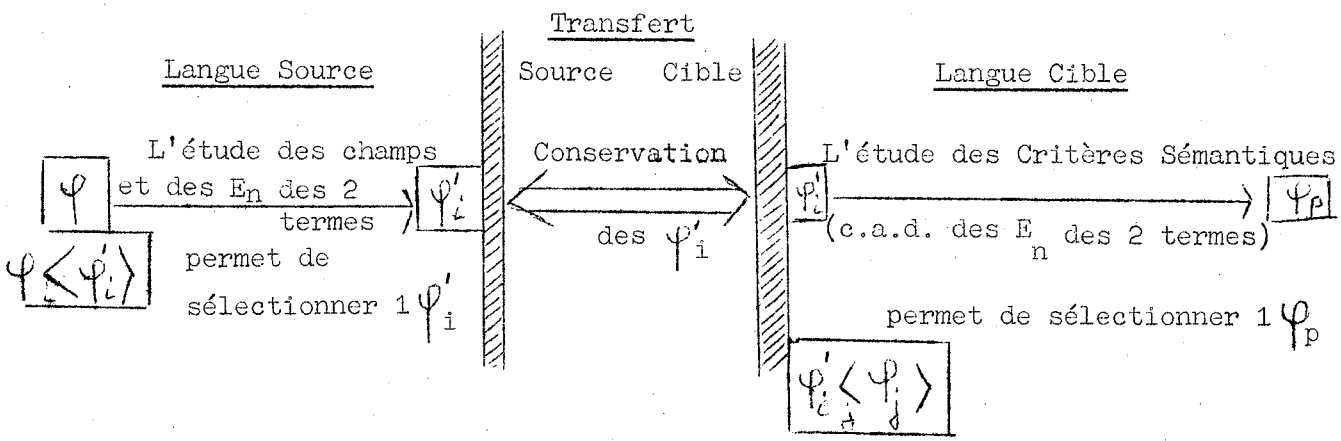
Ceci nous permet de sélectionner la $\varphi'_p$ représentée dans la phrase par la $\varphi_p$.

Nous avons admis qu'au cours du passage de la Langue Source à la Langue Cible, il y a conservation des relations Sémantiques $\varphi'$.

La $\varphi'_p$ précédente, se retrouve donc en Langue Cible. Mais dans celle-ci, on a $\varphi'_p \underset{j}{\langle} \varphi_j \rangle$ en utilisant les critères sémantiques définis au chapître III

nous pouvons ainsi sélectionner la relation syntaxique $\varphi_p$ dont on tirera la construction syntaxique valable de la Langue Cible

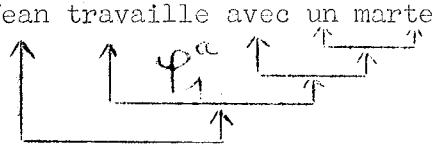
Tout ceci peut être représenté par le Schéma suivant :



Reprenons l'exemple 2 du chapître I :

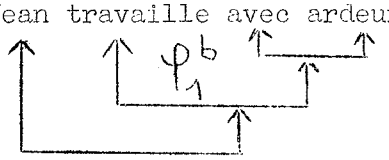
Anglais

a) Jean travaille avec un marteau



with a hammer

b) Jean travaille avec ardeur



eagerly

Dans les deux cas, nous avons la même relation syntaxique

| | | | | | |
|---|---|---|---|---|---|
| N° $\varphi_1$ | Verbe | ..... | Avec | Subst Commun | ..... |

Mais nous avons la correspondance : $\varphi_1 \rightarrow \varphi_1 \quad \underset{i}{<} \quad \varphi'_i >$

| | | | | | | |
|---|---|---|---|---|---|---|
| N° $\varphi'_1$ | compl^t Accompagnement | Activité | Etat | ..... | être animé | .... |
| N° $\varphi'_2$ | compl^t Moyen | Activité | Etat | ..... | Instrument | .... |
| N° $\varphi'_3$ | compl^t Manière | Activité | Etat | ..... | Manière | .... |

Pour a) l'étude sémantique (champs + $E_n$ caractéristiques) fait correspondre à $\varphi_1^a \Longrightarrow \varphi'_2$

Pour b) l'étude sémantique fait correspondre à $\varphi_1^b \Longrightarrow \varphi'_3$.

Mais en anglais, $\varphi'_2$ est représenté par la relation syntaxique $\varphi_i$

| | | | |
|---|---|---|---|
| N° $\varphi_i$ | Verbe ..... | with | Subst Commun ..... |

et $\varphi'_3$ par $\varphi_j$ :

| | | |
|---|---|---|
| N° $\varphi_j$ | Verbe ..... | Adv..... |

Donc $\varphi_1^a$ se traduira en anglais par

| | | |
|---|---|---|
| Verbe | With | Subst. Commun ..... |

à savoir "Verbe" with a Hammer.

De même $\varphi_1^b$ se traduira en Anglais par

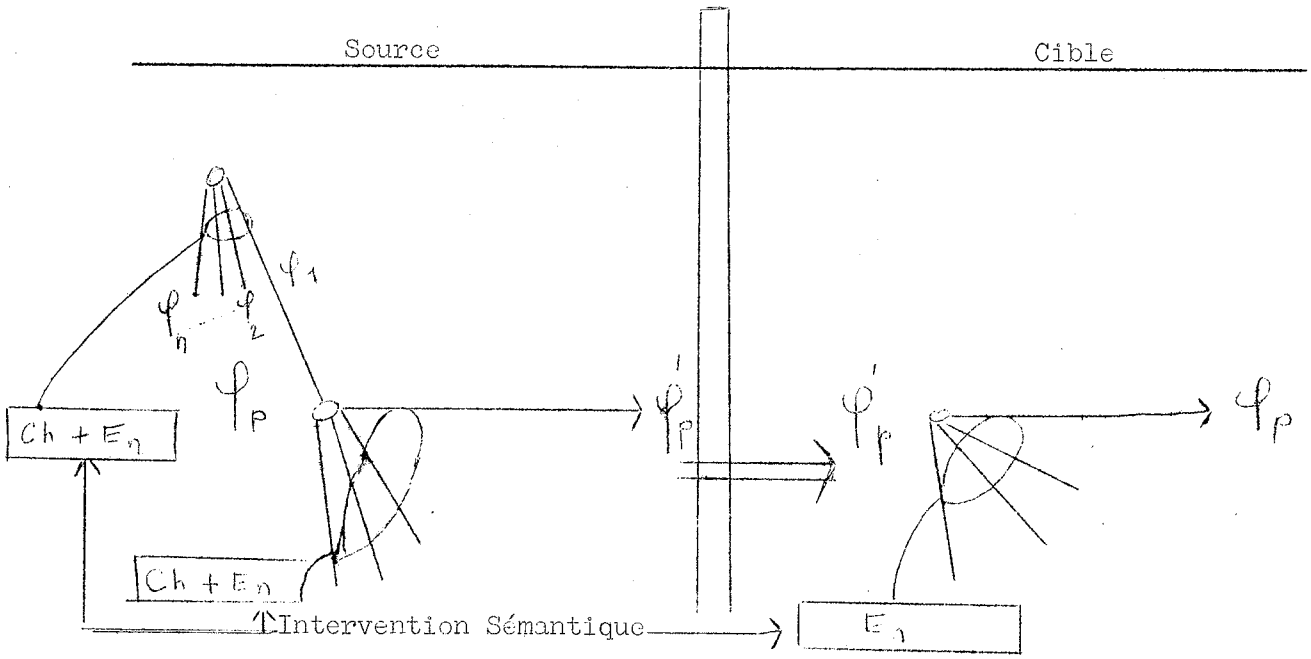
| Verbe ..... | Adv..... |
|---|---|

à savoir : "Verbe" eagerly.

Ainsi partant de la même construction syntaxique en Français, nous avons bien trouvé et choisi deux constructions syntaxiques différentes en Anglais.

c- Schéma général

Par un schéma général nous pouvons représenter le passage d'un $\varphi$ source à un $\varphi$ cible et préciser les éléments sémantiques qui interviennent à chaque fois et permettent ainsi de faire un choix. Nous avons :




II- Pluralité des mots "Cible"

1- Un mot de la langue Source est une polysémie à résoudre en langue source

Nous avons vu qu'il existait deux types de polysémie, nous allons d'abord considérer le cas le plus fréquent :

a) La Polysémie Véritable

Nous avons donc une unité lexicale $U_L$ qui conduit en langue source à plusieurs unités sémantiques : $U_L^1 \underset{i}{<} U_{Si}^1 >$. Regardons dans le Dictionnaire s'il existe une $\varphi$ qui permette en principe de choisir l'$U_S$ (cf. chap. IV- IV a).

- Si elle n'existe pas de toute façon cette unité lexicale $U_L^1$ est reliée à une autre unité lexicale $U_L^2$ par une relation syntaxique $\varphi_p$.

Cette $U_L^2$ correspond à une seule unité sémantique $U_S^2$, sinon on décompose le Problème en autant de problèmes partiels que $U_L^2$ a d'unités sémantiques.

Nous étudions alors la compatibilité d'une liaison entre $U_{S_1}^1$ et $U_S^2$ :

D'abord y-a-t'il compatibilité entre les champs de ces deux $U_S$? Ceci nous est indiqué par la "Matrice de Relation entre Champs".
Puis nous regardons si dans les $V_S$ de $U_{S_2}$ il y en a une qui appartient aux $V_S$ Caractéristiques indiquées dans le Dictionnaire et définies au chap. IV-IV,a.

S'il y a compatibilité entre les champs et s'il existe une $V_S$ caractéristique, alors $U_{S_1}^1$ est possible.

Sinon on étudie $U_{S_2}^1$ avec $U_S^2$ etc....
On détermine ainsi l'"unité sémantique $U_{S_i}$ qui d'après le contexte a la plus grande chance d'être ici employée. L'ambigrité est ainsi levée.

<u>Exemple</u> On a l'$U_L$ "glace" $\Longrightarrow$ $\exists U_s$ $\begin{cases} \text{eau gelée} & U_{S1} \\ \text{sorbet} & U_{S_2} \\ \text{lame de verre} & U_{S3} \end{cases}$

Le Dictionnaire n'indique pas de $\varphi$.

Si on a le prédicat suivant :

"le tain de la glace"

$U_L^1 \Longrightarrow U_S^1$ $\uparrow \varphi_p \uparrow$ $\qquad\qquad$ $U_L^2 \Longrightarrow U_L^2 \underset{i}{\langle} U_{Si}^2 \rangle$

On étudie la compatibilité d'une liaison entre $U_S^1$ avec $U_{Si}^2$ .
On conçoit bien qu'il y ait incompatibilité entre les champs de "tain" et ceux "d'eau gelée"

entre ceux de "tain" et ceux de "sorbet"
tandis qu'il y a compatibilité entre les champs de "tain"et ceux de "lame de verre"
Donc l'unité sémantique $U_{S\,3}^{\,2}$ "lame de verre" est valable.

- Supposons que dans le Dictionnaire, il existe une $\varphi$ qui permette de choisir l'$U_S$.
Regardons si dans la phrase étudiée, cette relation syntaxique existe réellement.
Sinon, on est ramené au cas précédent.
Si oui, cette $\varphi$ nous sélectionne une $U_L^2$ .

Nous supposons de nouveau que $U_L^2$ ne possède qu'une $U_S$ et nous étudions la compatibilité entre $U_{S1}^{\,1}$ et $U_S^2$ de la même façon qu'au cas précédent.

L'avantage d'avoir trouvé une $\varphi$ dans le dictionnaire tient à ce que les $U_S$ caractéristiques permettent alors un choix bien plus sûr.

b) Polysémie extraite d'une homographie externe

Ici nous avons le choix entre deux unités lexicales = $U_L^1$ et $U_L^2$ mais chacune de ces $U_L$ possède plusieurs unités sémantiques. Ainsi :

"Manche" (masc) $U_L^1 \Longrightarrow$

| | |
|---|---|
| $U_{S1}^{\,1}$ | Partie par laquelle on tient un instrument |
| $U_{S2}^{\,1}$ | OS apparent des côtelettes et gigots |

"Manche" (fém.) $U_L^2 \Longrightarrow$

| | |
|---|---|
| $U_{s1}^{\,2}$ | Partie de vêtement qui couvre le bras |
| $U_{s\,2}^{\,2}$ | Au jeu, une des parties liées |
| $U_{s\,3}^{\,2}$ | Conduit où passe l'air |
| $U_{s\,4}^{\,2}$ | Bras de mer resserré entre deux terres. |

Ce mot dont on veut déterminer l'$U_L$ est relié à un autre mot représenté dans $\mathcal{L}$ par l'unité lexicale $U_L^3$ suivant une liaison syntaxique représentée par la relation $\varphi$ p.

Comme auparavant, nous supposerons que $U_L^3$ ne possède qu'une $U_S^3$.

Nous allons conjuguer le choix de l'$U_L$ et de l'$U_S$ en étudiant la compatibilité d'une liaison entre les $U_S$ appartenant aux deux $U_L$ polysémiques et $U_S^3$.

Ainsi, ici nous étudions successivement :

L'étude de la compatibilité se fait comme on a vu au ch. V(II.1.a)

| | |
|---|---|
| $U_{S\,1}^{1}$ | avec $U_S^3$ |
| $U_{S2}^{1}$ | avec $U_S^3$ |
| $U_{S\,1}^{2}$ | avec $U_S^3$ |
| ⋮ | ⋮ |
| $U_{S4}^{2}$ | avec $U_S^3$ |

Cette étude nous choisit donc $U_L$ et $U_S$ en même temps.

Exemple

Soit le prédicat :

"Les Manches du Veston".

Dans le dictionnaire nous avons les renseignements suivants :

$U_L^1$

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| "Manche" (masc) | 2Us | Instrument | ... | $U_{s1}^1$ | Viande | ..... | $Us_2^1$ |

$U_L^2$

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| "Manche" (fém) | 4Us | Vêtement | ... | $Us_1^2$ | Jeu | ... | $Us_2^2$ | Air | Aération | .... | | |
| | | | | | | | $Us_3^2$ | Mer | ... | $Us_4^2$ | | |

L'$U_L$ "veston" possède la $V_S$ : "vêtement".

On conçoit que l'étude des champs et des variables sémantiques permet de choisir $U_L^1$ et $U_{s1}^1$.

c) Différentes possibilités pour l'étude des Polysèmies

Cette étude est basée sur la recherche de la compatibilité d'une liaison entre $U_{S_i}^1$ et $U_S^2$

1- On peut étudier les $U_{Si}$ l'une après l'autre et s'arrêter dès qu'il y a compatibilité alors on prend le risque d'avoir oublié une autre possibilité.

2- On peut étudier toutes les $U_S$, il y aura peut-être plusieurs solutions mais on sera sûr de ne pas en avoir oubliées.
Dans l'hypothèse du 1), il faudrait dans le Dictionnaire, classer les $U_S$ de chaque $U_L$ par ordre de sortie le plus probable. Ainsi les risques d'oubli diminueraient.

D'autre part, il se peut que les $V_S$ caractéristiques d'une $U_S$ soient bien plus précises que celles d'une autre $U_S$.
Alors dans l'hypothèse 1) l'$U_S$ bien précisée doit précéder l'autre.
Ainsi dans cette hypothèse on aurait l'exemple suivant :

| son | 3 Us | Céréales | .... | Us (péricarde) | Laines | ... | Us (Echeveau) | Instrument Musique |
|---|---|---|---|---|---|---|---|---|
| | | fréquent et précis | | | peu fréquent et précis | | | fréquent |

| Choc | Intensité | ..... | Us (bruit) |
|---|---|---|---|
| et imprécis | | | |

2- Un mot (non polysémique) de la langue Source, donne naissance suivant le contexte à plusieurs mots de la langue Cible.

Soit $U_L^1$ l'unité lexicale non polysémique il lui correspond $U_S^1$. En langue cible, nous avons donc n $U_L$ : $U_S^1 \Longrightarrow$ n $U_{Li}^1$.
Regardons dans le dictionnaire s'il existe une $\varphi$ qui permet en principe de choisir l'$U_L$ (cf. chap. IV-IV b)

- Si elle n'existe pas, de toute façon cette unité lexicale $U_L^1$ est reliée à une autre unité lexicale $U_L^2$ par une relation syntaxique $\varphi p$.

Cette $U_L^2$ correspond à une seule unité sémantique $U_S^2$, sinon on décompose le problème en autant de problèmes partiels que $U_L^2$ a d'unités sémantiques et l'Etude perd alors sa précision.

On a vu en ch. IV- IV-b que chaque $U_{Li}^1$ est précisée par les $V_S$ caractéristiques.

Nous prenons donc les $V_S$ caractéristiques de $U_{L1}^1$ une à une et cherchons si on en retrouve une dans les $V_S$ de $U_S^2$.

Si oui, $U_{L1}^1$ est possible.

Si non, on considère les $V_S$ caractéristiques de $U_{L2}^1$ etc.....
on détermine ainsi l'unité lexicale $U_{Li}^1$ qui d'après le contexte a la plus grande chance d'être ici employée. L'ambiguïté est ainsi levée.

<u>Exemple</u> l'$U_L$ "moeurs" n'est pas polysémique mais possède en langue anglaise trois $U_L$ cible possibles :

$U_S^1$ — $U_{L1}^1$ manners
$U_S^1$ — $U_{L2}^1$ customs
$U_S^1$ — $U_{L3}^1$ habits

Le Dictionnaire n'indique pas de $\varphi$.
mais il y a :

| "moeurs" | 3 $U_L$ cible | hommes | ... | manners | pays | ... | customs | Animaux | .. | habits |
|---|---|---|---|---|---|---|---|---|---|---|

Si on a le prédicat suivant :

"Les moeurs des Castors"

$U_L^1 \Longrightarrow U_S^1 \xleftarrow{\varphi p} U_L^2 \Longrightarrow U_S^2$

On conçoit bien que les $V_S$ caractéristiques : "hommes" "pays" ne seront pas trouvées dans les $V_S$ de $U_S^2$ "castors" tandis que la $V_S$ caractéristique "Animaux" y sera.
Donc l'unité lexicale $U_{L3}^1$ "habits" est la seule valable.

- Supposons que dans le dictionnaire il existe une $\varphi$ qui permet de choisir l'$U_L$ cible.
On regarde d'abord si dans la phrase étudiée, cette relation syntaxique existe réellement, sinon on est ramené au cas précédent.
Si oui, cette $\varphi$ nous sélectionne une $U_L^2$ que nous supposerons encore non polysémique et nous faisons comme auparavant l'étude de compatibilité des $V_S$.

Le fait d'avoir trouvé une $\varphi$ dans le dictionnaire aura augmenté la sûreté de la Méthode.

<u>Remarque</u>

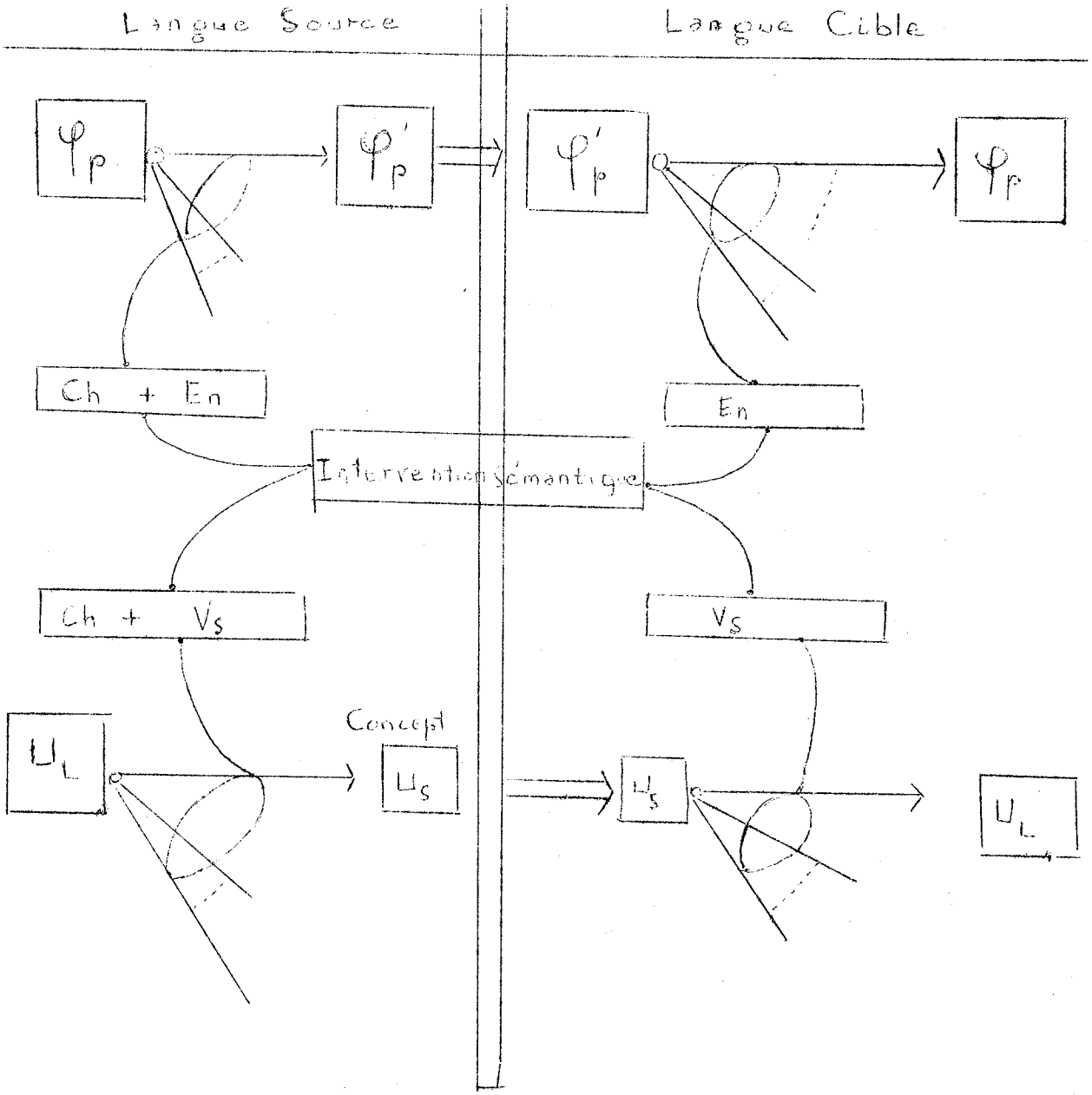
On a vu au chapître I que des $U_L$ comme "glace" conjuguaient les difficultés vues en (1) et (2).
Pour lever ces ambiguïtés on emploiera successivement les deux Méthodes : la première déterminera en langue source $U_{Si}$.

La seconde choisira l'$U_L$ cible.

## III- Comparaison des méthodes d'étude.

Nous avons vu que la méthode utilisée pour résoudre les problèmes posés par la pluralité des constructions syntaxiques était basée sur le passage d'une $\varphi$ source à une $\varphi$ cible et de même pour résoudre la pluralité des mots "cible" on utilisait une méthode basée sur le passage d'une $U_L$ source à une $U_L$ cible.

Il est intéressant de noter la similitude de ces deux passages où la sémantique intervient en des points analogues.
Nous mettons ceci en valeur dans le schéma suivant



## IV- Cas particuliers d'emploi de la sémantique.

On peut envisager d'employer aussi la sémantique dans certains cas particuliers d'Etude.
Le but alors recherché est un gain de temps et une plus grande sûreté quant au résultat dans l'état sémantique.

- Ainsi certains verbes se construisent souvent avec telle préposition et avec un second terme possédant telle variable sémantique. Nous avons donc en langue source le système syntaxo-sémantique suivant :

Verbe + préposition + 2e terme avec telle $V_S$.
On connait alors de suite en langue cible,

1°- La traduction du verbe c.a.d. l'$U_L$ cible,

2°- La construction syntaxique.

Si on trouve très fréquemment le système syntaxo-sémantique que nous venons de préciser, on aura intérêt à l'indiquer dans le dictionnaire ainsi que les renseignements 1° et 2° pour la langue cible. Ces indications seront données dans les codes se rapportant au verbe. On gagnera ainsi beaucoup de temps en effet celui-ci aurait été occupé à faire une étude syntaxique et sémantique bien plus longue et bien moins sûre en langues source et cible.

Exemples

a- Soit le système syntaxo-sémantique en langue Russe :

ИГРАТЬ "igrat'" + В + 2e terme avec $V_S$ : "noms de jeux". Le dictionnaire indiquera aussi :

1°) L'$U_L$ cible : "jouer"

2°) La construction syntaxique

| | | | | | |
|---|---|---|---|---|---|
| $\varphi_{N°}$ | Verbe | ..... | à<br>à la<br>au, aux | Subst | Commun ..... |
| | jouer | | | | nom de jeu |

Exemple : "igrat' v t'enis" jouer au tennis
ИГРАТЬ В ТЕННИС

b- Soit le système syntaxo-sémantique en langue Russe :

ПЕРЕВЕСТИ + НА + 2e terme avec $V_S$ : "langue"
" p'e r' e v' e s t i + na"

Le dictionnaire indiquera aussi :

1°- L'$U_L$ cible : "traduire"

2°- La construction syntaxique

| | | | | | |
|---|---|---|---|---|---|
| $\varphi_{N°}$ | Verbe | .... | en | Subst | ..... |
| | "traduire" | | | | Langue |

## V- INTERACTION SYNTAXE-SEMANTIQUE EN MACHINE

### a- Analyse syntaxique et analyse sémantique

On peut faire l'analyse morphologique d'un texte tout entier puis en commencer son analyse syntaxique ; par contre il n'en sera pas de meme pour les analyses syntaxique et sémantique.

En effet, nous avons vu que toute étude sémantique se faisait avec une base syntaxique.
Considérons la liaison entre les deux analyses.

Pour l'analyse syntaxique nous avons adopté la stratégie dite de "toutes les possibilités" : On recherche toutes les structures acceptables. On obtient ainsi toutes les structures syntaxiquement valables mais il peut en résulter une arborescence rapide et une exploration de branches inutiles.
Pour cette analyse syntaxique nous pouvons envisager deux cas distincts :

1°- Cas n° 1 : L'étude syntaxique est bien adaptée à la langue source et elle conduit par phrase à une moyenne de deux constructions syntaxiques possibles. On pense que les langues Russe, Allemande....appartiendront à ce groupe de langue source.

2°- Cas n° 2

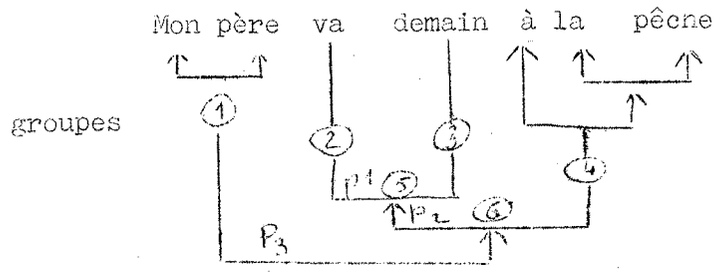
L'étude syntaxique est très difficile à faire et dans une même phrase très nombreuses sont les ambiguïtés que l'on ne peut pas résoudre par la syntaxe. On pense que les langues Italienne anglaise....appartiendront à ce groupe de langues source.

### b Hypothèses sur la représentation de la structure syntaxique d'une phrase

On rassemble d'abord les syntagmes élémentaires en "groupes". Nous prenons pour le mot "groupe" le sens qui lui est donné par Grévisse dans un modèle d'analyse grammaticale ( 10, page 1093 )
La représentation de la structure syntaxique d'une phrase peut se faire de façons bien différentes. Nous allons en voir deux :

1) On y indique les groupes de la phrase, suivis de leurs relations syntaxiques $\psi p$. Chaque $\psi p$ est représenté par son numéro : $p_i$

Exemple




(1) groupe du sujet : mon père

(2) groupe verbal : va

(3) groupe du complément de temps : demain

(4) groupe du complément de lieu : à la pêche.

D'où la représentation

$$\| \underbrace{1,\ 2,\ 3,\ 4}_{\text{groupes}} \quad \underbrace{(p_1,\ 2,\ 3) \Rightarrow 5,\ \ (p_2,\ 5,\ 4) \Rightarrow 6,\ (p_3,\ 1,\ 6)}_{\text{relations syntaxiques}} \|$$

2°- Toujours en considérant les groupes, on peut envisager une autre représentation, plus concise :

$$\| 1,\ 2,\ 3\ p_1,\ 4\ p_2,\ p_3 \|$$

Ici la fonction syntaxique est appliquée aux deux groupes (ou ensemble de groupes) qui la précède dans la représentation.
Quelle que soit la représentation de la structure syntaxique qui sera admise, on voit que les groupes et les fonctions $\varphi$p seront bien mis en valeur et facilement accessibles.

C Voyons à quel moment introduire la sémantique.

1°- D'abord pour sélectionner la seule construction syntaxique sémantiquement valable :
toute construction syntaxique valable de la phrase sera représentée. Il y aura donc autant de représentations que de constructions syntaxiques différentes.
Pour que deux représentations soient différentes, il faut et il suffit qu'un groupe ou une des $\varphi$p soit différent.
Dans le cas n° 1, l'analyse syntaxique de la phrase sera faite complètement, puis

si plusieurs constructions sont trouvées, nous ferons un choix entre les différentes représentations de la structure syntaxique.
Dans le cas n° 2, l'analyse syntaxique de la phrase ne peut être faite complètement, aussi nous opérerons en deux étapes.

1- Nous aurons d'abord à sélectionner la liste des groupes de la phrase, sémantiquement valables.

2- Puis nous ferons un choix entre les différentes représentations de la structure syntaxique.

Le fait d'avoir sélectionné la liste des groupes sémantiquement valables nous aura bien diminué le nombre des différentes représentations de la structure syntaxique. Le choix entre celles-ci sera donc bien plus facile et plus rapide qu'en employant la méthode n° 1 : nous aurons raccourci le temps d'exécution de l'analyse syntaxique.
Dans la méthode n° 2, nous aurons une imbrication de la sémantique dans la syntaxe :

1°- formation syntaxique de tous les groupes possibles.
2°- Sélection sémantique de la liste des groupes de la phrase.
3°- Formation des constructions syntaxiques de la phrase à partir des groupes.
4°- Sélection sémantique de la construction syntaxique sémantiquement valable.

2°- Ce n'est qu'après avoir obtenu pour la phrase la construction syntaxique sémantiquement valable, que l'on fera l'étude des polysémies pour déterminer l'unité sémantique représentée ici par chaque unité lexicale.

## IV- Nécessité d'introduire d'autres éléments dans l'étude sémantique

Nous avons vu que l'étude sémantique était longue et parfois imprécise. Nous allons montrer ce que l'on pourrait envisager de faire pour améliorer ces imperfections.
Considérons par exemple le choix entre plusieurs $U_S$ en langue source ou celui entre plusieurs $U_L$ en langue cible : cf chap. IV-IV.

Ces choix se font suivant des critères sémantiques et on obtient des tableaux du type de ceux dont parle le professeur Hays (11) :

critère sémantique

| A | $C_1$ | $C_2$ | --- | $C_n$ |
|---|---|---|---|---|
| $U_{S1}$ | $\alpha$ | a | ... | b |
| $U_{S2}$ | c | B | ... | d |
| ⋮ | ⋮ | ⋮ | ⋱ | ⋮ |
| $U_{Sn}$ | x | y | z | V |

critère sémantique

| B | $C_1$ | $C_2$ | $C_3$ | $C_4$ | $C_5$ | $C_6$ |
|---|---|---|---|---|---|---|
| $U_{L1}$ | $\underline{\alpha}$ | a | $\underline{B}$ | b | c | d |
| $U_{L2}$ | e | f | g | h | $\underline{\gamma}$ | i |
| $U_{L3}$ | $\underline{\alpha}$ | $\underline{\vartheta}$ | k | $\underline{\varepsilon}$ | $\underline{\ell}$ | $\underline{V}$ |

La solution idéale pour le choix des $U_S$ et des $U_L$ serait que le choix des critères sémantiques soit fait de façon à obtenir :

pour A

$$\alpha = \beta = \ldots\ldots = V = 1$$

$$a = b = \ldots = z = 0$$

pour B

$$\alpha + \beta = 1$$

$$\gamma = 1$$

$$\vartheta + \varepsilon + V = 1$$

$$a = b = \ldots\ldots = l = 0$$

Mais de part la nature d'une langue ceci se révèle impossible à faire. Malgré tout, en choisissant les variables sémantiques caractéristiques nous supposons implicitement que ces égalités sont vérifiées.

- Il en est de même pour le choix des relations syntaxiques et sémantiques suivant certains critères.

Nous sommes donc conduits a introduire dans l'étude sémantique ce que le professeur Lecerf (12) appelle :
" Le calcul d'erreurs en linguistique automatique ".

Exemple : Une $U_L$ peut posséder plusieurs $U_S$.
Si elle est dans ce cas-là, on pourra faire une étude des fréquences de rencontre de chacune des $U_S$. Si on se fixe un"seuil d'erreur " et si pour une $U_S$ la fréquence de rencontre est inférieure à ce seuil on pourra l'éliminer.

Nous pouvons donc envisager l'introduction volontaire et calculée d'une erreur dans le dictionnaire par suppression d'une $U_S$ ou d'une $U_L$ cible pour un mot.

De plus, si l'étude sémantique était mise en défaut on pourrait choisir l'$U_S$ (ou l'$U_L$) qui a la fréquence de rencontre la plus élevée.

La sémantique a été introduite dans l'algorithmique de Passage $\mathcal{L} \rightarrow \mathcal{L}'$ dans le but de lever les ambiguïtés.

Elle essaye de le faire, mais le vide d'erreurs est forcément impossible, aussi pour les réduire au minimum le "calcul d'erreurs" est-il indispensable.

Ainsi, grâce à l'aide combinée de ces nouveaux éléments et de bien d'autres encore inconnus, onréalisera le système d'interprétation sémantique $\Sigma$ .

# BIBLIOGRAPHIE


[1] S. CECCATO — "How language is understood by man and how it has been been decided to make a machine understand it" (Nato Advanced Study Institute, July 1962)

[2] B. VAUQUOIS — "Langages artificiels, systèmes formels et traduction automatique" (Nato Advanced Study Institute, July 1962)

[3] B. VAUQUOIS et J. VEYRUNES — "Présentation de l'Analyse morphologique du Russe" (C.E.T.A.-G document G-100-C)

[4] G. VEILLON — "Consultation d'un dictionnaire et analyse Morphologique en traduction automatique" (Thèse de 3ème Cycle- Université de Grenoble - Juin 1962)

[5] A.AUROUX — "Contribution à la reconnaissance des structures syntaxiques en traduction mécanique" (Thèse de 3ème cycle - Université de Grenoble - Juin 1962)

[6] S. CECCATO et Bruna ZONTA — "Human translation and translation by machine" ( First International conference on machine translation of languages - Teddington 1961)

[7] R. BELTRAME — "List of classifications for machanical translation from russian into english" Extrait de : "Linguistic Analysis and programming for mechanical translation" (Centre de Cybernétique de l'Université de Milan .)

[8] S. LAMB and W. JACOBSON — " A high-speed-capacity Dictionnary System " (Mechanical Translation - Vol. 6 - Nov. 1961)

[9] S.PERSCHKE "Some notes concerning Ambigous Russian verbs"
Extrait de "Linguistic Analysis and Programming for Mechanical Translation" (Centre de Cybernétique de l'Université de Milan)

[10] M. GREVISSE "Le bon usage" - Grammaire française - 7ème édition 1961

[11] D. HAYS "Research procedures in machine translation"
(Nato Advanced study Institute July 1962)

[12] Y. LECERF "Frequence d'erreurs. Concepts de haute ou de moyenne qualité en traduction automatique"
(Nato Advanced Study Institute July 1962)

[13] I. SAKAI "Syntax in Universal Translation"
(First International Conference on Machine Translation of languages - Teddington 1961)

[14] K. SPARCK JONES "Mechanised Semantic Classification"
(first international conference on Machine translation of languages - Teddington 1961)

[15] M. MASTERMAN "Semantic Message detection for Machine Translation, Using an interlingua"
(First International Conference on machine translation of languages - Teddington 1961)

---

# ANNEXE 1

## CHAMPS SEMANTIQUES TIRES DE LA CLASSIFICATION DECIMALE UNIVERSELLE

**50 Généralités sur les sciences pures**

501 Sur les sciences exactes
502 sur les sciences naturelles

**51 Mathématiques**

511 Arithmétique
512 Algèbre
513 Géométrie
514 Trigonométrie
515 Géométrie descriptive
516 Géométrie analytique
517 Analyse
518 Calcul graphique
519 Probabilités

**52 Astronomie - Géodésie**

521 Astronomie théorique
522 Astronomie pratique
523 Astronomie descriptive
524
525 La terre
526 Géodésie
527 Navigation
528 Ephémérides
529 Chronologie

**53 Mécanique rationelle - Physique**

531 Mécanique rationnelle
532 Hydrostatique et hydraulique
533 Gaz-Pneumatique-Aérodynamique
534 Vibrations-Ondulations-Acoustique
535 Rayonnement-Lumière-optique
536 Chaleur-Thermodynamique
537 Electricité
538 Magnétisme et électromagnétisme
539 Physique moléculaire et atomique

**54 Chimie**

541 Théorie chimique
542 Chimie opératoire et manipulations
543 Chimie analytique en général
544 Analyse qualitative
545 Analyse quantitative
546 Chimie inorganique
547 Chimie organique
548 Cristallographie
549 Minéralogie.

**55 Sciences géologiques et géophysiques**

551 Géologie physique et dynamique
552 Lithologie-Etude des roches,météorites
553 Géologie économique

56 Paléontologie

561 Paléobotanique

562 Paléozoologie- Invertébrés

563 Protozoaires

564 Mollusques

565 Vers

566 Vertébrés

567 Poissons- Amphibiés

568 Sauropsides

569 Mammifères

57 Sciences biologiques

571 Archéologie préhistorique

572 Antropologie. Ethnologie

573

574 Homologies- Hétérologies

575 Evolution - Hérédité

576 Origine de la vie - Cytologie

577 Matière vivante - Biochimie Enzymologie

578 Technique et microscopie biologique

579 Collections d'histoire naturelle

58 Botanique

581 Botanique analytique ou biologique

582 Botanique systèmatique

583 ...

⋮

$58_i$

⋮

589 Botanique systématique

59 Zoologie- Animaux

591 Zoologie analytique ou biologique

592 Invertébrés

593 Radiés

594 Malacologie- Mollusques-Coquillages

595 Articulés

596 Vertébrés

597 Poissons-Ichtyologie

598 Reptiles-Oiseaux

599 Mammifères.

---

# PLAN GENERAL

INTRODUCTION Page

CHAPITRE I :

RESULTATS DE L'ETUDE SYNTAXIQUE - SES LIMITES 4

-I- Définitions rencontrées en morphologie et en syntaxe
-II- But de l'étude syntaxique
-III- Etude des difficultés rencontrées

CHAPITRE II :

SOUS QUELLES FORMES INTRODUIRE LA SEMANTIQUE ? 17

-I- Unité sémantique
-II- Champ sémantique
-III- Variable sémantique
-Iv- Relation sémantique
-V- Relation entre champs et début du code sémantique
-VI- Eléments notionels et représentation du code sémantique
-VII- Le code sémantique en machine
-VIII-Possibilité de concentrer les champs.

CHAPITRE III :

LIAISONS ENTRE LES ELEMENTS SYNTAXIQUES ET LES ELEMENTS SEMANTIQUES 35

-I- Formation des unités lexicales et sémantiques
-II- Introduction des relations syntaxiques et sémantiques
-III- Introduction des relations sémantiques et "liaisons en triangle"
-IV- Représentation des $\varphi$ et des $\varphi'$
-V- Liaisons entre les $\varphi$ et les $\varphi'$ dans une langue
-VI- Tableau général d'Etude des $\varphi$ et des $\varphi'$

CHAPITRE IV : Page

ETUDE DU PASSAGE MOT "SOURCE" ⟶ MOT "CIBLE" 58

-I- Passage d'une Ul de la langue source à une Ul de la langue cible

-II- Conservation d'une $U_s$ dans le passage source-cible

-III- Méthode employée pour trouver le verbe cible en vue de la traduction russe-français

-IV- Contenu du dictionnaire pour l'obtention de l'Ul cible

-V- Supplément syntaxique dans le dictionnaire pour le cas où l'Ul est un verbe

CHAPITRE V :

UTILISATION DES ELEMENTS SEMANTIQUES POUR RESOUDRE LES DIFFICULTES RENCONTREES AUX CHAPITRES PRECEDENTS 74

-I- Pluralité des constructions syntaxiques

-II- Pluralité des mots "Cible"

-III- Comparaison des méthodes d'étude

-IV- Cas particuliers d'emploi de la sémantique

-V- Interaction syntaxe-sémantique en machine

-VI- Nécessité d'introduire d'autres éléments dans l'étude sémantique

BIBLIOGRAPHIE 93

ANNEXE 1 CHAMPS SEMANTIQUES TIRES DE LA CLASSIFICATION DECIMALE UNIVERSELLE 95