

Collège de France
Chaire Internationale

Vers une linguistique Sens-Texte

LEÇON INAUGURALE

par

M. Igor Mel'čuk

Professeur

Monsieur l'Administrateur,
Mesdames et Messieurs les Professeurs,
Chers amis,

J'aimerais d'abord remercier tous les collègues auxquels je dois ce moment tellement précieux pour moi : vous parler de mon travail — plutôt de mon amour, de l'amour de ma vie, c'est-à-dire de la linguistique, à laquelle j'ai consacré quarante ans de mon existence. Hélas, ces collègues sont trop nombreux pour que je puisse les nommer un à un. Qu'il me soit donc permis de mentionner tout d'abord les deux idoles de ma jeunesse, les deux géants de la linguistique française : Émile Benveniste et Lucien Tesnière, dont l'œuvre linguistique m'a profondément marqué. Ensuite, j'indiquerai Yves Gentilhomme : son amitié m'a servi d'appui pendant toutes ces années. Et *last but not least*, je m'adresserai à mes deux collègues du Collège de France, MM. Claude Hagège et Jean-Marie Zemb, grâce auxquels je suis ici. Quant aux autres, mes maîtres en linguistique (A.Reformatskij et A.Xolodovič, avant tout) ainsi que mes compagnons de randonnées à travers la jungle de la Langue (A.Zholkovsky, Ju.Apresjan, L.Iordanskaja, N.Pertsov, A.Clas, A.Polguère, ...), qu'ils reçoivent ici l'expression de ma gratitude la plus chaleureuse. Et maintenant, au travail !

Je procéderai en cinq étapes.

1) Tout d'abord, la méthodologie qui se trouve au cœur de mon approche de l'étude des langues naturelles sera exposée. Cette méthodologie consiste en la construction de MODÈLES FONCTIONNELS des langues. L'idée sous-jacente en est très simple :

La langue est considérée comme un mécanisme, ou un système de règles, qui permet au locuteur de faire deux choses :

1) PARLER, c'est-à-dire, (être capable de) faire correspondre à un sens qu'il veut exprimer tous les textes de sa langue qui, d'après lui, peuvent véhiculer ce sens et choisir celui qui passe le mieux dans les circonstances concrètes d'un acte langagier donné ;

2) COMPRENDRE LA PAROLE, c'est-à-dire, (être capable de) faire correspondre à un texte qu'il perçoit tous les sens que, d'après lui, ce texte peut véhiculer et choisir celui qui passe le mieux dans les circonstances concrètes d'un acte langagier donné.

Notre tâche en tant que linguistes est donc de construire, pour la langue étudiée **L**, un système de règles — quelque chose comme un programme informatique — qui définisse les mêmes correspondances entre sens et textes que celles qu'établissent les locuteurs. De tels systèmes de règles

forment ce que nous appellerons des *modèles Sens-Texte*, ou MST. Je parlerai donc d'abord de la modélisation Sens-Texte en linguistique.

2) Dans un deuxième temps, j'esquisserai un MODÈLE SENS-TEXTE PARTICULIER, sur lequel je travaille depuis plus de 30 ans. Je donnerai des exemples de représentations linguistiques d'énoncés à tous les niveaux considérés, ainsi que des exemples de règles pour chacune des composantes du modèle.

3) Après cela, j'examinerai quelques LIENS entre PHÉNOMÈNES SÉMANTIQUES HYPOTHÉTIQUES et PHÉNOMÈNES COOCCURRENCIELS OBSERVABLES (en français), pour étayer ainsi la modélisation proposée. En particulier, je m'arrêterai sur le concept de *fonction lexicale* : un formalisme élaboré en vue de la description systématique de la combinatoire lexicale.

4) Ensuite, je caractériserai une branche de la linguistique qui découle naturellement de la modélisation fonctionnelle des langues — la LINGUISTIQUE SENS-TEXTE. J'indiquerai les grandes lignes de recherche existant dans ce cadre, ainsi que certains des résultats obtenus ; ce faisant, je développerai tout particulièrement la question du métalangage de la linguistique.

5) Pour finir, je parlerai brièvement de la place que la linguistique Sens-Texte est appelée à prendre dans la vie de la société humaine.

Mon exposé se divisera donc en cinq sections.

1. La modélisation fonctionnelle comme principe central en recherche linguistique

1.1. Les modèles dans la science

Le concept de modèle joue un rôle crucial dans toutes les sciences. Les chercheurs qui, pour une raison ou une autre, ne sont pas en mesure d'observer la structure interne de l'objet ou du phénomène **P** qu'ils étudient ont habituellement recours à des modèles : ils construisent des modèles de **P** qu'ils étudient ensuite. De même, lorsque **P** est trop compliqué ou que les données à son sujet sont trop riches ou incohérentes, on construit un modèle approximatif de **P**, qui permet d'en faire une étude approfondie. Et cela aussi bien en cosmogonie qu'en géophysique, en biologie moléculaire qu'en chimie, en neurologie ou en sociologie. En exagérant un peu, je pourrais dire que **la science, c'est la construction de modèles**. Dans toutes les situations où une observation directe des faits est impossible, on est contraint de construire un modèle de ces faits — ce qui est le cas depuis Galilée. «Il n'y a pas de savant qui ne pense continuellement par modèles — même s'il ne l'avoue ni aux autres ni à lui-même» (Auger 1965: 4).

La linguistique est dans la même situation que toute autre science de la nature. En effet, la langue, qui est un système de règles fort complexes, encodées on ne sait comment dans le cerveau des locuteurs, est inaccessible à l'observation directe des linguistes «purs»¹ : nous ne pouvons ouvrir les crânes, ni les pénétrer avec des électrodes pour observer la langue telle qu'elle est stockée dans le cerveau. Nous n'avons pour solution que de recourir à des modèles des langues. La linguistique moderne est donc une science qui doit se donner comme tâche principale la construction de modèles de langues particulières et de la Langue humaine en général ; même si c'est rarement dit explicitement, c'est ce qui se fait. N.Chomsky, avec sa théorie générative-transformationnelle, a, vers la fin des années 50, solidement implanté la modélisation en science du langage. On ne manque pas non plus de théorisation de la notion même de modèles linguistiques ; on trouvait déjà dans Nagel *et al.* 1962

quelques articles qui en discutaient de façon très précise. Citons aussi J.Molino (1985: 29) : «La morphologie — comme les autres domaines linguistiques et comme le langage en général — ne peut être décrite que par des modèles». On est donc en droit de considérer comme acquis en linguistique l'usage extensif de modèles.

La question se pose de savoir quel type de modèles la linguistique doit de préférence utiliser. J'essayerai d'ébaucher ici une réponse possible.

1.2. Les modèles fonctionnels

Débarrassons-nous tout d'abord de l'ambiguïté du mot *modèle*. Considérons les expressions suivantes : *un modèle dont on peint un portrait* — *un modèle d'avion (en papier)* — *le modèle d'atome de Bohr-Rutherford*. Elles ont en commun de faire référence à un même état de choses : deux entités, X et Y (le portrait représentant quelqu'un, un jouet en papier ressemblant à un avion, les équations décrivant l'atome), dont l'une, X, est intentionnellement créée par l'Homme de façon à posséder certaines caractéristiques de l'autre, Y². Il s'agit donc d'une relation binaire antisymétrique ('être un modèle de ...'), qui appelle trois remarques :

Premièrement, dans *un modèle dont on peint un portrait*, on appelle *modèle* l'entité donnée, alors que la représentation de cette entité créée par l'homme s'appelle *portrait* ; dans *un modèle d'avion* ou *le modèle d'atome de Bohr-Rutherford*, la relation est converse : l'entité donnée, l'avion/l'atome, n'a pas de nom spécial, alors que la représentation de cette entité s'appelle *modèle*³. Je retiendrai seulement les emplois du deuxième type, réservant donc le terme de *modèle* à une ENTITÉ CRÉÉE dans le but de représenter l'entité étudiée.

Deuxièmement, l'expression *un modèle d'avion* désigne un objet physique, alors que le modèle d'atome de Bohr-Rutherford n'est pas un objet physique, mais un système d'expressions symboliques (= d'équations). Le terme *modèle* ne sera utilisé dans la suite de mon exposé que dans ce dernier sens, pour désigner un SYSTÈME ABSTRAIT D'EXPRESSIONS SYMBOLIQUES.

Troisièmement, un modèle d'avion ressemble à un avion, même si ce modèle ne peut pas voler : c'est un modèle structural. Le modèle d'atome de Bohr-Rutherford, par contre, ne ressemble en rien à un atome ; en toute logique, il ne peut même pas y prétendre : c'est un modèle FONCTIONNEL. Il est appelé à représenter le comportement, ou le fonctionnement, de l'atome (d'où son nom de *modèle fonctionnel*).

Je peux maintenant fixer le sens dans lequel j'utiliserai dorénavant le terme *modèle* :

Modèle fonctionnel

|| *X est un modèle (fonctionnel) de Y* : X est un système d'expressions symboliques créé par le
|| chercheur dans le but de représenter le fonctionnement de l'entité donnée Y qu'il étudie.

Dans ce qui suit, il ne sera question que des modèles fonctionnels conformes à la définition ci-dessus⁴.

Bien entendu, un modèle fonctionnel n'est pertinent que si c'est le FONCTIONNEMENT de l'objet modélisé qui nous intéresse, comme c'en est le cas pour la langue. Les modèles fonctionnels sont donc particulièrement adaptés à l'étude linguistique.

Avant d'aller plus loin, signalons deux propriétés importantes des modèles fonctionnels :

1) Un modèle fonctionnel permet de modéliser une «BOÎTE NOIRE», c'est-à-dire un objet manifestant un certain comportement observable, mais qu'on ne peut «ouvrir» pour en observer et étudier la structure interne.

2) Un modèle fonctionnel ne garantit pas la vérité de la description obtenue, mais une simple APPROXIMATION de la vérité : avec un modèle fonctionnel, en n'observant que les effets, nous en déduisons les causes ; or on sait que le même effet peut avoir une infinité de causes. Une chose est sûre : plus nous étudions l'objet modélisé — c'est-à-dire plus nous recensons et prenons en charge de ses comportements —, plus nous nous approchons de la réalité. Cependant, le linguiste qui accepte de travailler avec des modèles fonctionnels doit se rappeler constamment qu'un seul contre-exemple bien choisi suffit à compromettre le modèle qu'il propose (= à démontrer son inadéquation) ; un modèle fonctionnel est donc facilement falsifiable (au sens de Popper) et, pour cette raison, il peut être considéré comme un instrument scientifique de valeur. Par contre, rien ne peut prouver la validité d'un modèle fonctionnel. Son caractère hypothétique lui est inhérent ; c'est une réalité que le linguiste doit humblement accepter.

1.3. Le modèle Sens-Texte : un modèle fonctionnel global de la langue

C'est dans le sens de la définition ci-dessus que j'ai amorcé, il y a trente ans, avec quelques collègues de Moscou (principalement, Alexandre Zholkovsky et Jurij Apresjan), l'élaboration d'un modèle fonctionnel global de la langue naturelle : le Modèle Sens-Texte [= MST] (Žolkovskij & Mel'čuk 1967, Mel'čuk 1973, 1974, 1981, 1988, 1992a). Je commencerai par ébaucher la théorie linguistique sous-jacente au MST, c'est-à-dire la Théorie Sens-Texte. Pour faciliter l'exposé, les abréviations suivantes seront utilisées :

DEC	: Dictionnaire Explicatif	-Phon	: phonique
	et Combinatoire	R-	: représentation
FL	: fonction lexicale	S-	: structure
-Morph	: morphologique	-S	: de surface
MST	: Modèle Sens-Texte	-Sém	: sémantique
-P	: profond	-Synt	: syntaxique

De plus, L réfèrera à une lexie donnée, et **L**, à une langue donnée.

La première mention d'un terme technique sera imprimée en *italique gras*.

La théorie Sens-Texte est basée sur trois postulats, de nature plutôt hétérogène. Le premier vise l'OBJET DE L'ÉTUDE de la théorie ; il exprime ma conception générale de ce qui est la langue. Le deuxième vise le RÉSULTAT ESCOMPTÉ DE L'ÉTUDE ; il exprime ma conception de ce que devrait être la recherche et la description linguistique. Quant au troisième, il vise le LIEN entre LA LANGUE et SA DESCRIPTION ; il expose certains traits fondamentaux de la langue qu'il est essentiel de refléter directement dans la description.

Postulat 1 : La langue comme correspondance «Sens-Texte»

|| La langue est un système fini de règles qui spécifie une CORRESPONDANCE multi-multivoque
|| entre l'ensemble infini dénombrable de sens et un ensemble infini dénombrable de textes.

Posons que les sens apparaissent, dans un MST, sous forme d'objets symboliques formels appelés **Représentations Sémantiques** [= RSém], et les textes sous forme d'objets symboliques formels appelés **Représentations Phoniques** [= RPhon]. Alors le Postulat 1 peut s'écrire comme suit :

$$(1) \{RSém_i\} \text{ langue}; \longleftrightarrow; \{RPhon_j\} \quad | \quad 0 < i, j < \infty$$

Le Postulat 1 présuppose le CARACTÈRE DISCRET de nos représentations et, par conséquent, de notre modèle. Je n'ai aucun argument pour justifier le choix que j'ai fait de ce type de modèles linguistiques. Comme on le sait, les scientifiques cherchent, de plus en plus, à construire des modèles linguistiques continus, ou analogiques ; je pense cependant qu'il est essentiel de continuer l'élaboration de modèles discrets, sans pour autant leur donner un caractère exclusif. Il semble que leur potentiel cognitif soit loin d'être épuisé et qu'au moins leur utilité pratique — par exemple, dans l'enseignement des langues et dans le traitement automatique des textes — justifie l'approche discrète.

Précisons que la RPhon est écrite en utilisant une transcription phonétique quelconque ; ce type de transcription est trop connu pour être discuté ici. La RSém, quant à elle, est écrite en utilisant une «transcription sémantique» ; cette transcription est spécifique à chaque langue et doit encore être élaborée ; j'en dirai quelques mots plus loin.

L'ensemble des RSém_i bien formées peut être spécifié par une grammaire formelle, de sorte que cet ensemble peut servir d'entrée à un modèle formel.

Postulat 2 : Les modèles Sens-Texte comme outil de description des langues

La correspondance (1) doit être décrite par un DISPOSITIF LOGIQUE, qui constitue un modèle fonctionnel de la langue de type Sens-Texte ; il doit être élaboré et présenté dans la direction Sens \implies Texte.

Un modèle Sens-Texte reçoit, à l'entrée, des RSém_i et produit des RPhon_j à la sortie ; il doit le faire d'une façon qui soit la plus proche possible de celle des locuteurs, c'est-à-dire que le modèle doit reproduire au mieux la correspondance entre le sens qu'un locuteur veut exprimer et le texte qui, d'après lui, véhicule ce sens.

D'un point de vue formel, le passage «Sens \implies Texte» et le passage «Texte \implies Sens» sont équivalents. Mais du point de vue de l'élaboration du modèle par des linguistes et de sa présentation, ce n'est pas du tout le cas. En conséquence, un MST sera organisé À PARTIR DU SENS VERS LE TEXTE (ce qu'indique bien son appellation), c'est-à-dire dans le sens de la *synthèse*, ou de la *production de la parole* — plutôt que dans le sens opposé, celui de l'analyse, ou de la compréhension de la parole. Autrement dit, le MST suit le parcours onomasiologique, modélisant avant tout l'activité langagière du LOCUTEUR, considérée comme plus linguistique que celle du destinataire. Construire un texte pour un sens donné présuppose essentiellement l'exercice de connaissances purement linguistiques, alors que l'extraction du sens d'un texte donné exige, dans une bien plus grande proportion, une connaissance du monde assez poussée et des capacités logiques. On peut citer beaucoup d'exemples confirmant le fait que la langue elle-même donne une priorité absolue au locuteur : par exemple, on parle de *locuteurs natifs*, pas de **destinataires natifs* ;

l'expression *parler français* est idiomatique (*français* utilisé en tant qu'adverbe — sans article; cf. **lire français*), alors que l'expression *comprendre le français* est libre ; la langue est «égocentrique» — en ce sens que beaucoup de signes linguistiques encodent le point de vue de 'moi' (= de celui qui parle), comme par exemple les déictiques ; toutes les langues ont un verbe spécial pour le sens 'produire, à partir du sens, du texte acoustique dans une langue' : *parler*, mais aucune n'a de verbe spécial pour exprimer le sens inverse 'extraire du sens à partir du texte acoustique d'une langue' : on dit *comprendre*, ce qui s'applique à n'importe quelle compréhension. Je ne vais cependant pas justifier cette vision de la langue et, en conséquence, de la linguistique : je la postule.

Postulat 3 : La phrase et le mot comme unités de base de la description linguistique

Dans la description de la correspondance (1), deux NIVEAUX INTERMÉDIAIRES de représentation des énoncés sont nécessaires pour mettre en lumière les faits linguistiques pertinents : la représentation SYNTAXIQUE [= RSynt], qui correspond aux régularités spécifiques à la PHRASE, et la représentation MORPHOLOGIQUE [= RMorph], qui correspond aux régularités spécifiques au MOT.

La phrase et le mot (plus précisément *mot-forme* au sens de Mel'čuk 1993-97, v. 1: 187 *ssq.*) sont, respectivement, l'unité maximale et l'unité minimale de la parole, unités autonomes et universelles ; les lois de la langue se limitent à la phrase, et les signes plus petits que le mot ne sont pas perçus par les locuteurs. C'est dans le cadre de la phrase qu'on doit s'occuper de l'ordre des mots, de l'accord et du régime, de la structuration communicative, de la cooccurrence lexicale restreinte, etc., alors que dans le mot, on fait face à la flexion et à la dérivation, ainsi qu'aux alternances phonémiques. C'est pour capter ces caractéristiques de la phrase et du mot que le MST postule les deux niveaux intermédiaires entre sens et textes. L'introduction de ces niveaux — le niveau syntaxique et le niveau morphologique — semble d'autant plus justifiée qu'elle correspond à une tradition respectée pratiquement par toutes les écoles de pensée linguistiques.

Il résulte des postulats 1-3 qu'un MST est un dispositif logique, ou un ensemble de règles, ayant la structure suivante :

(2) STRUCTURE GÉNÉRALE D'UN MODÈLE SENS-TEXTE

$$\{RSém_i\} \setminus a(\mathbf{sémantique}; \longleftrightarrow;) \{RSynt_k\} \setminus a(\mathbf{syntaxe}; \longleftrightarrow;) \{RMorph\}$$

$$\mathbf{morphologie +; phonologie}; \longleftrightarrow; ; \{RPhon_j\}$$

Les noms des composantes majeures, ou modules, du MST sont imprimés en gras.

Un MST possède trois particularités importantes.

1) Un MST est purement ÉQUATIF, ou TRADUCTIF ; à la différence de beaucoup de ses contemporains, ce n'est pas un modèle génératif. Il ne génère rien (comprenant le terme *générer* au sens strictement mathématique, c'est-à-dire, 'spécifier un ensemble de ...', ce que font les grammaires génératives) ; mais il fait correspondre à chaque RSém toutes les RPhon qui peuvent l'exprimer dans une langue donnée ; c'est pourquoi il est qualifié d'«équatif». Plus précisément, un MST met en rapport des représentations linguistiques des niveaux adjacents : il prend une représentation du niveau *n* — tout comme on prend une recette de cuisine pour faire un gâteau ou un plan pour construire une

maison — et il lui associe toutes les représentations correspondantes du niveau $n+1$, qui sont ainsi produites sous le contrôle de la représentation de départ, sans que celle-ci soit modifiée ; de plus, il fait la sélection de la meilleure représentation du niveau $n+1$. Un MST n'est donc pas un modèle transformationnel non plus, puisque les représentations manipulées ne subissent pas de changement. Un MST essaie de se comporter comme un locuteur, qui ne passe son temps ni à générer des ensembles des phrases grammaticalement correctes ou à distinguer entre les phrases correctes et incorrectes, ni à transformer des structures abstraites ; un locuteur parle, c'est-à-dire qu'il exprime, au moyen de textes, les sens qu'il veut communiquer. Un MST doit faire la même chose : «traduire» un sens donné en un texte qui l'exprime (voilà pourquoi ce modèle est qualifié de «traductif»).

2) Un MST est basé sur le PARAPHRASAGE, c'est-à-dire sur la synonymie des énoncés. En effet, la compétence linguistique du locuteur consiste, avant tout, en sa capacité de produire, pour un sens de départ ' σ ', tous les textes qui peuvent l'exprimer (= toutes les paraphrases possibles) et de choisir le ou les textes les mieux adaptés à une situation ou à un contexte donnés.

3) Un MST est GLOBAL ET INTÉGRAL : il tend à présenter la langue comme un tout indivisible plutôt qu'à décrire juste un fragment isolé de celle-ci (comme, par exemple, sa sémantique ou sa morphologie). Toutes les composantes du MST — son lexique et toutes les parties de sa grammaire — doivent être bien «accordées», puisqu'elles sont destinées à «collaborer» étroitement au cours du processus de synthèse des textes.

Voilà donc les grandes lignes selon lesquelles doivent se construire les modèles linguistiques de type Sens-Texte.

2. Ébauche d'un modèle Sens-Texte particulier

Pour illustrer mon propos, je vais maintenant présenter un MST particulier — celui qui fait l'objet de mes recherches depuis 30 ans. Je commencerai par introduire la distinction des sous-niveaux profonds et de surface au sein des niveaux de la représentation linguistique, pour donner ensuite des exemples des représentations pertinentes à tous les niveaux ; après quoi nous verrons des exemples des règles de chaque composante du modèle. De cette façon, j'illustrerai les étapes qui conduisent, en partant d'un sens donné, c'est-à-dire d'une entrée possible pour un MST du français, à un des textes correspondants, c'est-à-dire à une des sorties possibles du même MST. Étant donné le caractère de mon exposé, je limite mes références aux travaux réalisés dans le cadre du MST.

2.1. Sous-niveaux profond et de surface

Tous les niveaux de représentation linguistique des énoncés, sauf le niveau sémantique, sont subdivisés en un sous-niveau profond [= -P] et un sous-niveau de surface [= -S]. Le sous-niveau profond est orienté vers le sens : sa tâche est d'exprimer explicitement toutes les distinctions sémantiques pertinentes à son niveau. Le sous-niveau de surface est orienté vers le texte : sa tâche est d'exprimer explicitement toutes les distinctions formelles pertinentes à son niveau. (Je ne peux pas m'étendre davantage sur les différences entre sous-niveau profond et sous-niveau de surface, si importantes soient-elles ; je ferai quelques remarques à ce propos dans mes illustrations.)

Du fait de l'introduction de la dichotomie «profond vs de surface», nous obtenons un ensemble de sept représentations linguistiques pour un énoncé donné, de sorte que la structure détaillée du MST particulier dont il est question ici peut s'écrire de la façon suivante :

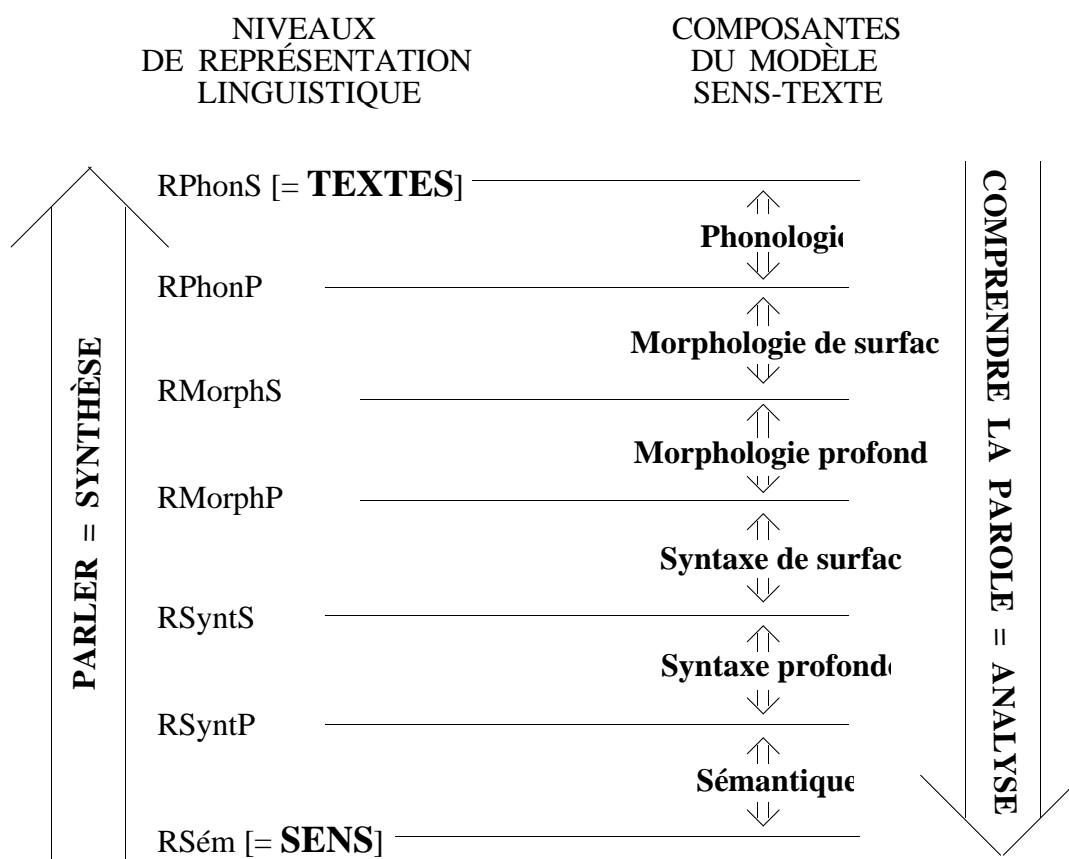
(3) STRUCTURE DÉTAILLÉE DU MODÈLE SENS-TEXTE CONSIDÉRÉ

$$\{RSém_i\} \backslash a(\text{sémantique}; \longleftrightarrow;) \{RSyntP_{k1}\} \backslash a(\text{syntaxe profonde}; \longleftrightarrow;) \{RSyntS_{k2}\} \backslash a(\text{syntaxe de surface}; \longleftrightarrow;) \{RMorphP_{l1}\}$$

$$\backslash a(\text{morphologie profonde}; \longleftrightarrow;) \{RMorphS_{l2}\} \backslash a(\text{morphologie de surface}; \longleftrightarrow;) \{RPhonP_{j1}\} \text{phonologie}; \longleftrightarrow; \{RPhonS_{j2}\}$$

[La RPhonP est une représentation PHONÉMIQUE : elle reflète les distinctions phoniques à charge sémantique ; une RPhonS est une représentation PHONÉTIQUE : elle reflète les distinctions phoniques pertinentes pour la production des sons réels.]

La figure ci-dessous présente l'agencement des toutes les composantes du modèle Sens-Texte vis-à-vis des niveaux de représentation des énoncés sous une forme légèrement différente qu'en (3) : à la verticale, c'est-à-dire, à partir du niveau le plus bas, ou le plus profond — le niveau sémantique — et en remontant vers le niveau le plus haut, ou le plus superficiel — le niveau phonétique :



Comme on le voit, le MST est constitué de six composantes, qui — mise de côté la distinction des sous-niveaux profond et de surface — correspondent aux divisions reconnues de la linguistique moderne (la sémantique, la syntaxe, la morphologie et la phonologie). Soulignons que le nom de chaque composante est déterminé par sa représentation de départ : ainsi, la SÉMANTIQUE part d'une représentation SÉMANTIQUE RSém pour construire toutes les RSyntP_{k1} qui véhiculent le même sens (c'est-à-dire le sens exprimé par cette RSém) ; la SYNTAXE PROFONDE part d'une représentation

SYNTAXIQUE PROFONDE RSyntP et fournit toutes les RSyntS_{k2} qui peuvent réaliser cette RSyntP de départ ; et ainsi de suite.

2.2. Représentations linguistiques dans le Modèle Sens-Texte

Chaque représentation utilisée par le MST est constituée de plusieurs objets formels appelés *structures*. Ainsi, une RSém comprend trois structures :

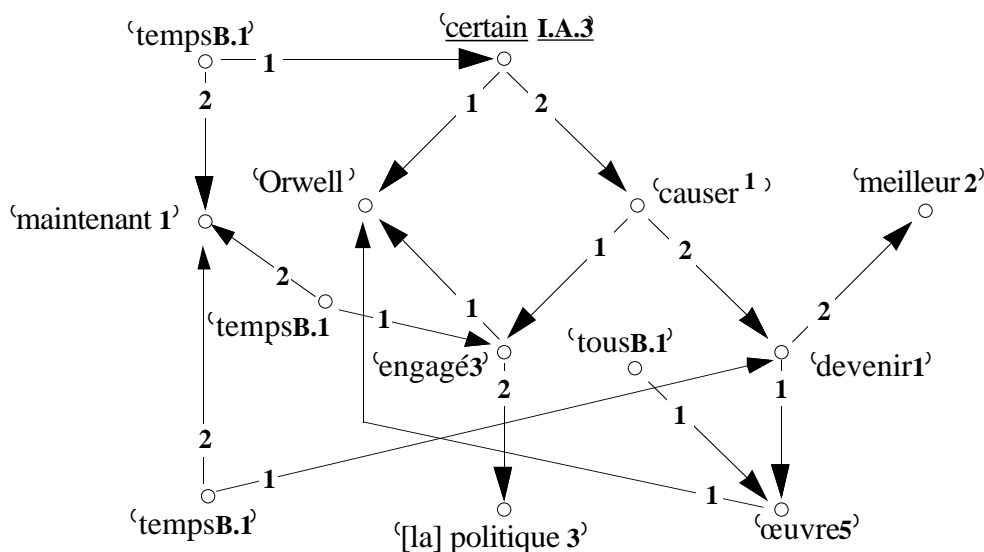
- Une structure sémantique, qui reflète le sens propositionnel, ou objectif, de l'énoncé représenté ; elle constitue le noyau de la RSém, ou sa *structure de base*, les deux autres structures lui étant superposées et la précisant.
- Une structure sémantico-communicative, qui reflète le sens communicatif, ou subjectif, de l'énoncé. C'est, de façon métaphorique, l'«itinéraire» que le locuteur suit à travers la structure sémantique lorsqu'il la «balaie» en la réalisant. La structure sémantico-communicative exprime les oppositions comme «thème vs rhème», «donné vs nouveau», «emphatisé vs neutre», etc.
- Une structure rhétorique, qui reflète les intentions «artistiques» du locuteur (l'ironie, le pathétique, les niveaux du langage différents, etc.).

De plus, une RSém est munie d'un ancrage référentiel : ses nœuds se voient associer des numéros identifiant les référents extralinguistiques des unités sémantiques. Comme on le sait, le statut référentiel/non référentiel des unités sémantiques peut avoir des répercussions sur les formes linguistiques qui les expriment (*Je cherche un collègue qui SAIT programmer* [l'existence d'un tel collègue est présupposé] vs *Je cherche un collègue qui SACHE programmer* [l'existence d'un tel collègue n'est pas présupposé]).

Les représentations des autres niveaux sont organisées de façon similaire, chacune comportant plusieurs structures, de sorte que le tableau véritable est assez compliqué. Pour simplifier, je n'indiquerai qu'une seule structure pour chaque représentation : la structure de base. Ainsi, la RSém est réduite à la structure sémantique SSém, les RSyntP et RSyntS aux structures syntaxiques profonde et de surface SSyntP et SSyntS, etc.

Commençons par la RSém. Le noyau d'une RSém — sa structure de base — est, comme je viens de le dire, la structure sémantique SSém d'une famille de phrases (quasi) synonymes, et c'est cette structure qui est représentée ci-dessous. C'est un graphe connexe orienté et étiqueté (= un réseau) dont les NŒUDS sont étiquetés par des noms d'unités sémantiques de la langue **L** (sens désambiguïsés des lexies de **L**). En l'occurrence, ce sont les acceptions des lexies françaises munies de numéros lexicographiques identificateurs (empruntés, dans mes exemples, au *Petit Robert*). Les unités sémantiques sont subdivisées en prédicats et noms (au sens du calcul des prédicats). Les ARCS du graphe sont donc étiquetés par des numéros distinctifs spécifiant les arguments du prédicat en cause.

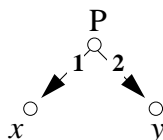
(4) Une structure sémantique



[Le soulignement de *certain* en haut du diagramme marque le nœud **communicativement dominant** ; cette indication fait partie de la structure sémantico-communicative, qui n'est pas représentée en tant que telle dans mon exposé.]

Ce réseau peut être lu littéralement ainsi : 'Orwell est certain que (le fait) qu'il est engagé dans la politique cause que ses œuvres deviennent meilleures'.

La SSém de (4) est écrite dans un langage sémantique (= la «transcription sémantique», mentionnée ci-dessus). La syntaxe de ce langage est celle des réseaux, qui servent de support formel à des expressions de type de formules du calcul des prédicats, les arcs indiquant des relations pré-

cat-argument. Le sous-réseau  représente donc une prédication du type $P(x,y)$.

Le lexique de ce langage est le lexique français désambiguïsé. De ce fait, la SSém du MST n'est pas universelle : elle représente des sens «français», tout en utilisant des moyens sémantiques «français».

La SSém (4) peut être exprimée par un nombre élevé de phrases du français. À titre d'exemple, considérons la phrase (5) :

(5) *Orwell n'a pas de doute quant à l'effet positif de son engagement politique sur la qualité de ses œuvres.*

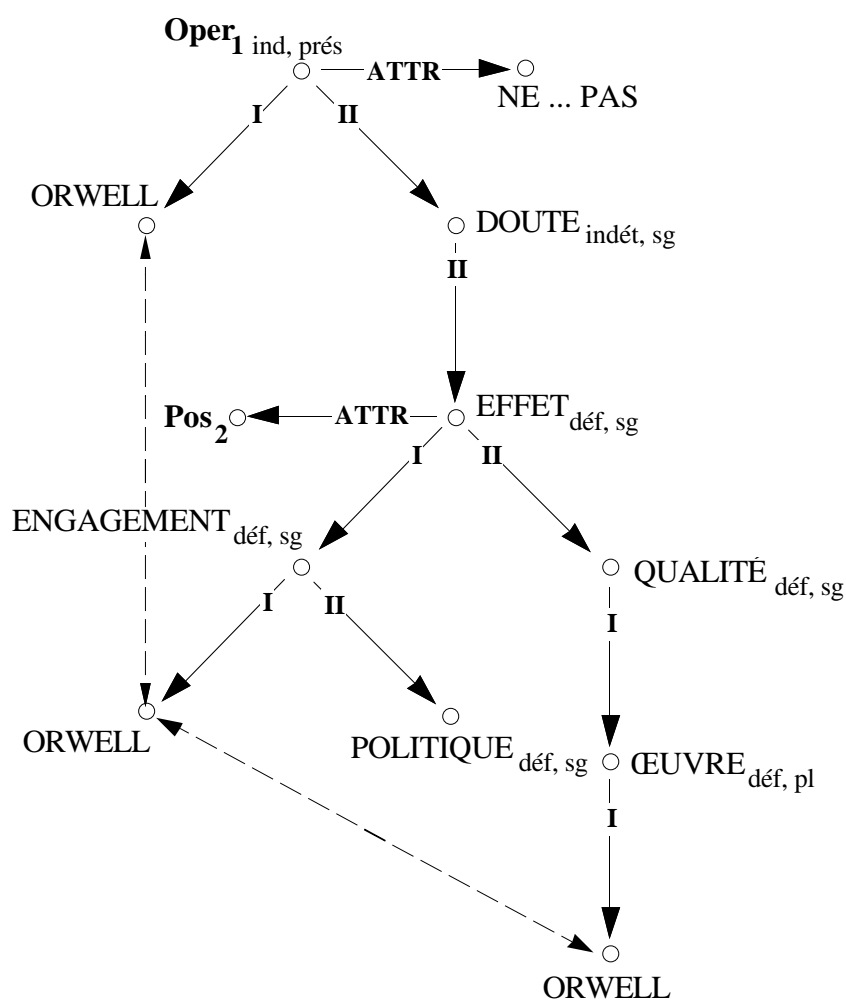
Pour obtenir cette phrase, la composante sémantique du MST construit d'abord — en se basant sur la RSém dont la SSém (4) fait partie — une RSyntP, que je présenterai en (6) par une seule de ses structures, à savoir son noyau, la SSyntP.

La structure syntaxique profonde SSyntP (d'une phrase) est un arbre de dépendance, dont les NŒUDS sont étiquetés par des lexèmes pleins : aucune servitude grammaticale, tels que les régimes, n'est représentée, et les pronominalisations ne sont pas effectuées ; sont également absents les marqueurs analytiques des valeurs flexionnelles. (Les lexèmes doivent être identifiés par leur numéros lexicographiques ; si je ne les indique pas, c'est pour ne pas surcharger la présentation.)

Les indices affectés aux lexèmes représentent les valeurs flexionnelles à charge sémantique, c'est-à-dire qui ont une source immédiate dans la RSém ; ces indices sont calculés par les règles sémantiques flexionnelles : cf. **R²** dans 2.3, p. 00.

Les ARCS de l'arbre SyntP (= branches) représentent des *relations syntaxiques universelles* — les six relations actancielles (I, II, III, ...) ; la relation **ATTR**(ibutive), qui couvre tous les cas de modification ; la relation **COORD**(inative), qui représente toute sorte de coordination ; et la relation **APPEND**(itive) pour les éléments de la phrase qui se trouvent dans un rapport vague avec son sommet (interjections, formes d'adresse, adverbes de phrase). Cet arbre n'est pas linéairement ordonné (en ce sens que l'ordonnement de ses nœuds imposé par une présentation graphique n'a aucune pertinence) : l'ordre des mots est un moyen d'expression de la structure syntaxique, il ne peut donc pas en faire partie ; il n'apparaît comme tel que dans la structure morphologique profonde de la phrase, donc à un niveau plus près du texte (voir (8)).

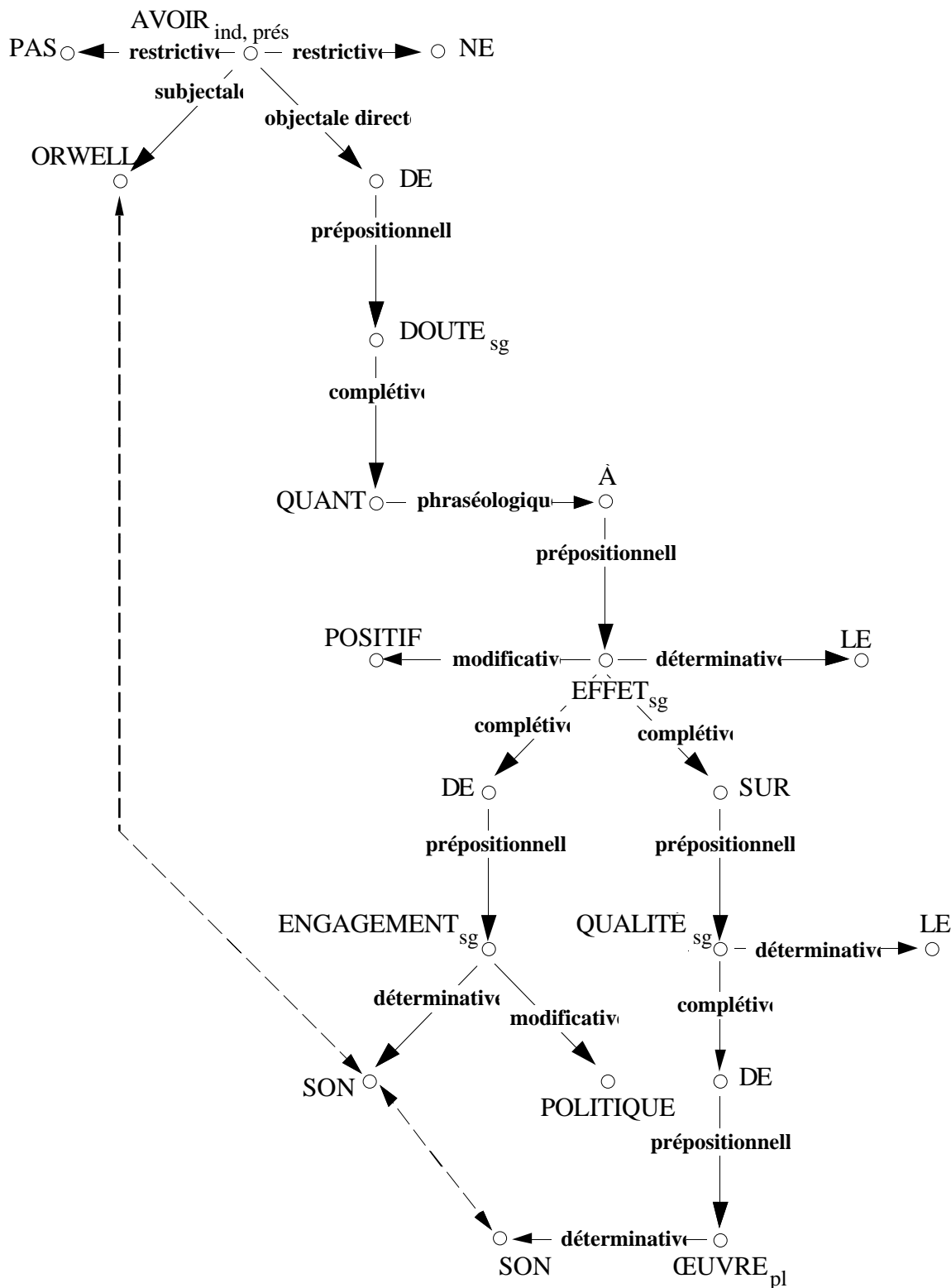
(6) Structure syntaxique profonde de la phrase (5)



[Les symboles **Oper₁** et **Pos₂** désignent des *fonctions lexicales* particulières ; ce sont des éléments introduits afin de représenter la cooccurrence lexicale restreinte (plus précisément, les collocations ; voir plus loin, 3.2). Les flèches bidirectionnelles en pointillé marquent la *coréférence* ; ces indications font partie de la structure anaphorique, qui est, comme je l'ai dit, une des composantes de la RSyntP, mais qui, pour la raison d'économie d'espace, n'est pas représentée en tant que telle dans mes exemples.]

On reconnaîtra immédiatement, dans le formalisme proposé, une élaboration poussée des «stemmas» du grand Lucien Tesnière (1959) ; c'est effectivement à lui que le MST doit ses structures relationnelles de dépendance (plutôt que les structures linéaires de constituants et leurs dérivées, plus connues dans la linguistique américaine).

(7) Structure syntaxique de surface de la phrase (5)



[Les flèches bidirectionnelles en pointillé marquent, tout comme dans la SSyntP, les relations de coréférence, mais cette fois-ci, celles entre pronoms ou adjectifs pronominaux et leurs antécédents.]

La composante syntaxique profonde du MST prend la représentation syntaxique profonde à laquelle la SSyntP (6) appartient comme point de départ, et elle en dérive toutes les représentations syntaxiques de surface RSyntS possibles, parmi lesquelles on trouve celle de la phrase (5) ; je la ferai apparaître en (7), p. 00, également sous la forme de sa structure de base : la SSyntS.

Une structure syntaxique de surface est aussi un arbre de dépendance non ordonné. Mais ici, les NŒUDS sont étiquetés de TOUS les lexèmes de la phrase, y compris tous les mots-outils et tous les éléments pronominaux ; on notera, entre autres, la présence des articles, qui, dans la structure SSyntP, sont encodés par les indices flexionnels de lexèmes pleins (les articles français étant en principe des unités sémantiques). Les ARCS de l'arbre SyntS (= branches) représentent des constructions syntaxiques particulières de la langue **L**, dans ce cas — du français. Néanmoins, les caractéristiques flexionnelles des lexèmes associés aux nœuds de la SSyntS restent incomplètes : tout comme dans la SSyntP, ne sont représentées que celles qui portent une charge sémantique ; les autres — résultants de l'accord ou du régime — n'apparaissent que plus près du texte, dans la RMorphP de la phrase, où elles sont introduites par les règles de la syntaxe de surface.

La composante syntaxique de surface du MST prend la représentation syntaxique de surface dont la SSyntS fait partie comme point de départ, et elle en dérive toutes les représentations morphologiques profondes RMorphP possibles, dont celle de la phrase (5) ; voir en (8) sa structure de base : la structure morphologique profonde SMorphP⁵. C'est une suite (= ensemble linéairement ordonné, ou chaîne) des représentations morphologiques profondes de tous les mots-formes qui constituent la phrase synthétisée. Une RMorphP d'un mot-forme est le nom du lexème auquel ce mot-forme appartient, muni de toutes les valeurs flexionnelles pertinentes, y compris ceux qui sont déterminées par les règles d'accord et de régime : par exemple, «3, sg» auprès de AVOIR ou «sg, masc» auprès de LE, POSITIF, SON et POLITIQUE.

(8) Structure morphologique profonde de la phrase (5)

ORWELL NE AVOIR_{ind, prés, 3, sg} PAS DE DOUTE_{sg} QUANT À LE_{sg, masc}
 EFFET_{sg} POSITIF_{sg, masc} DE SON_{sg, masc} ENGAGEMENT_{sg} POLITIQUE_{sg, masc}
 SUR LE_{sg, fém} QUALITÉ_{sg} DE SON_{pl, fém} ŒUVRE_{pl}

La composante morphologique profonde du MST prend la RMorphP comme entrée pour en dériver les RMorphS correspondantes, dont celle de la phrase (5) ; on en voit en (9) la structure de base, la structure morphologique de surface SMorphS. C'est une chaîne des représentations morphologiques de surface — c'est-à-dire morphémiques — des mots-formes de la phrase (5). Dans la SMorphS, chaque mot-forme est représenté comme l'ensemble des morphèmes qui le constituent ; les morphèmes, qui sont des ensembles de morphes, sont entourés d'accolades.

(9) Structure morphologique de surface de la phrase (5)

{ORWELL} {NE} {AVOIR},{IND.PRÉS},{3SG} {PAS} {DE} {DOUTE},{SG}
 {QUANT} {À} {LE},{SG.MASC} {EFFET},{SG} {POSITIF},{SG.MASC} {DE}
 {SON},{SG.MASC} {ENGAGEMENT},{SG} {POLITIQUE},{SG.MASC} {SUR}
 {LE},{SG.FÉM} {QUALITÉ},{SG} {DE} {SON},{PL.FÉM} {ŒUVRE},{PL}

La morphologie du français étant plutôt pauvre, on ne voit pas, dans mes illustrations, l'utilité de la distinction entre SMorphP et SMorphS ; on a l'impression d'avoir affaire à un simple changement de notation. Ce n'est cependant pas le cas pour les langues à morphologie plus riche, où très souvent la distribution des grammèmes — significations flexionnelles — entre morphèmes ne se fait pas de façon univoque⁶. La SMorphP vise les grammèmes, alors que la SMorphS a pour tâche de présenter explicitement les morphèmes, c'est-à-dire les ensembles de signes qui véhiculent ces grammèmes.

À partir de la représentation RMorphS dont la SMorphS (9) est la structure de base, la composante morphologique de surface du modèle Sens-Texte construit toutes les représentations phonémiques RPhonP possibles, y compris celle de la phrase (5). La composante phonologique du modèle fait correspondre à cette RPhonP une représentation phonétique RPhonS, le «plafond» du modèle : la correspondance entre la RPhonS et les véritables sons, établie par le module phonétique, se situe en dehors du MST.

Je ne citerai pas ici les représentations PhonP et PhonS, la phonologie étant peu pertinente pour mon exposé.

La même SSém (4) peut bien entendu être réalisée par d'autres phrases (plus ou moins) synonymes de (5) :

- (10) **a.** *Que s'engager politiquement apporte une contribution positive à ses œuvres ne soulève pas de doute chez Orwell.*
b. *Orwell est certain que son engagement politique influence positivement son œuvre, en en améliorant la qualité.*
c. *Être politiquement engagé permet à Orwell d'élever le niveau de ses œuvres — il en est convaincu.*
d. *Orwell n'a aucun doute : son engagement politique rend ses œuvres meilleures.*
e. *Orwell ne doute pas que son engagement politique rende ses œuvres meilleures.*
f. *L'engagement politique améliore beaucoup ses œuvres : Orwell en est tout à fait convaincu.*

On sait qu'une phrase de complexité modérée peut admettre des dizaines de millions de paraphrases (cf., par exemple, Mel'čuk 1992a: 26-27 ou Mel'čuk 1993-1997: vol. 1, 44, 76-77) ; inutile d'en énumérer ici un plus grand nombre. Les exemples (10) suffisent pour démontrer la richesse et la souplesse synonymiques de la langue, que le MST se doit de refléter. Ainsi, la SSém

(4) représente le sens d'une FAMILLE de phrases plus ou moins synonymes, autrement dit — de paraphrases. Le paraphrasage est à la fois le fondement du MST et son terrain d'essai : le modèle doit assurer toutes les paraphrases pour un sens donné et seulement les paraphrases acceptées par les locuteurs natifs.

2.3. Composantes du Modèle Sens-Texte

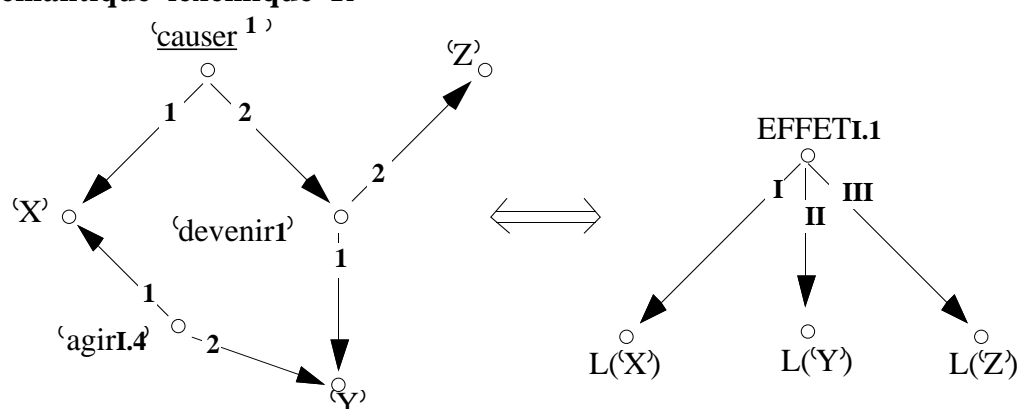
Comme je viens de le dire, un MST prend, à l'entrée, une RSém et produit, à la sortie, les phrases correspondantes en transcription phonétique, en passant par des représentations intermédiaires. Ce modèle est fonctionnel dans deux sens différents du terme *fonctionnel* : d'une part, il modélise le FONCTIONNEMENT de la langue (car les locuteurs ne font que cela : traduire des sens en textes et vice versa) ; d'autre part, il est lui-même organisé comme une FONCTION (au sens mathématique du mot) — une fonction qui fait correspondre à un sens (= son argument) l'ensemble de tous les textes synonymes qui l'expriment (= sa valeur). C'est au moyen de dispositifs de ce type que la description des langues naturelles peut (et, comme je le crois, devrait) être effectuée.

Présentons, à tour de rôle, les six composantes majeures, ou modules, du modèle Sens-Texte complet, tel que je le conçois à présent.

Une composante de ce MST est un ensemble de règles qui assurent la correspondance entre les représentations de deux niveaux adjacents : par exemple, la sémantique assure la correspondance entre RSém et RSyntP, et ainsi de suite. Je ne citerai pour chaque composante qu'une ou deux règles, choisies de façon à illustrer le passage entre les représentations de la phrase (5) indiquées ci-dessus ; le lecteur pourra suivre ainsi, au moins en partie, la construction de la phrase (5) à partir du sens (4), comme elle est effectuée par le MST.

Sémantique

Règle sémantique lexémique R^1

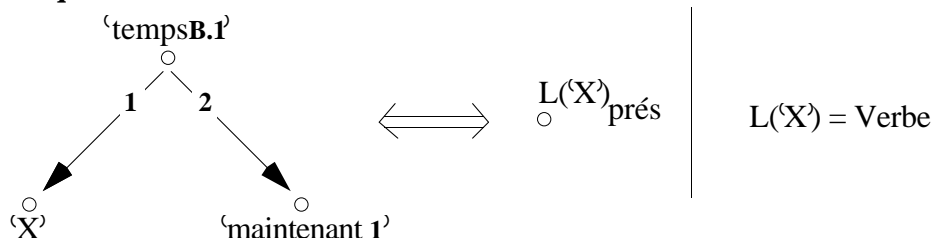


[La notation « $L(\alpha)$ » veut dire 'la lexie L qui exprime le sens ' α ''.]

La règle R^1 stipule que le sens 'X, agissant sur Y, cause que Y devient Z' peut être exprimé (entre autres, bien sûr) par le lexème EFFETI.1 : *l'effet Z de X sur Y*. Elle permet d'obtenir, à partir de la SSém (4), un fragment de l'arbre SyntP (6) : le nœud étiqueté EFFETI.1, avec les branchements correspondants. Soulignons que cette description du sens du lexème français EFFETI.1 est approximative : en réalité, la RSém de cette lexie doit la distinguer de tous ses quasi-synonymes, comme CONSÉQUENCE, RÉSULTAT, IMPACT, etc.

Une règle sémantique lexémique n'est pas autre chose que la partie centrale de l'article de dictionnaire du lexème en question. Le nombre de règles de ce type (et par conséquent d'articles de dictionnaire) dans une langue comme le français est d'à peu près un million. Leur ensemble constitue un dictionnaire formalisé d'un type nouveau : le *Dictionnaire explicatif et combinatoire* [= DEC ; voir Mel'čuk *et al.* 1984-1997), qui joue un rôle très important dans le MST.

Règle sémantique flexionnelle R^2



[La barre verticale sépare la partie conditions ; les conditions sont spécifiées par des formules booléennes qui portent a priori sur les deux parties de la règle.]

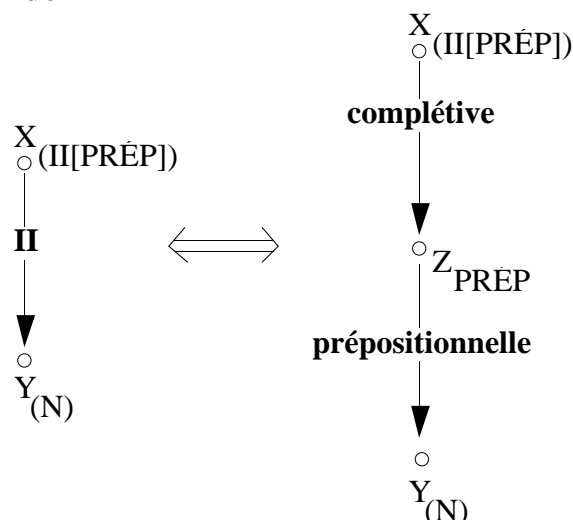
La règle R^2 stipule que le sens 'le temps de X est maintenant', c'est-à-dire que X a lieu au moment du discours, peut être exprimé par le grammème 'présent', attaché au lexème qui réalise le sens 'X', si ce lexème est un verbe. Cette règle contribue, elle aussi, à l'arbre (6) : elle ajoute l'indice «prés» au sommet de l'arbre (qui est étiqueté du symbole de la FL **Oper**₁). En morphologie, cet indice permettra la production de la forme du présent du verbe fini : voir la règle R^7 .

Les règles de ce type décrivent le sémantisme des grammèmes de la langue (= «sémantique morphologique»). Leur nombre, dans une langue donnée L , est de quelques centaines (en fonction de la richesse morphologique de L).

Il existe d'autres types de règles sémantiques dont je ne parlerai pas ici : entre autres, les règles sémantiques décrivant le sens des phrasèmes et des constructions syntaxiques⁷, ainsi que les règles qui concernent la structure communicative de la phrase. Une règle d'un tout autre type doit, par exemple, veiller à ce que la phrase ait un sommet verbal (par exemple, le nœud dominant de la SSém (4) — 'certain' — peut être réalisé, au niveau SyntP, par l'adjectif CERTAIN — comme en (10b) ; en français, un adjectif ne peut pas être le sommet syntaxique d'une phrase complète ; la règle en question introduit alors la copule ÊTRE).

Syntaxe profonde

Règle syntaxique profonde R^3

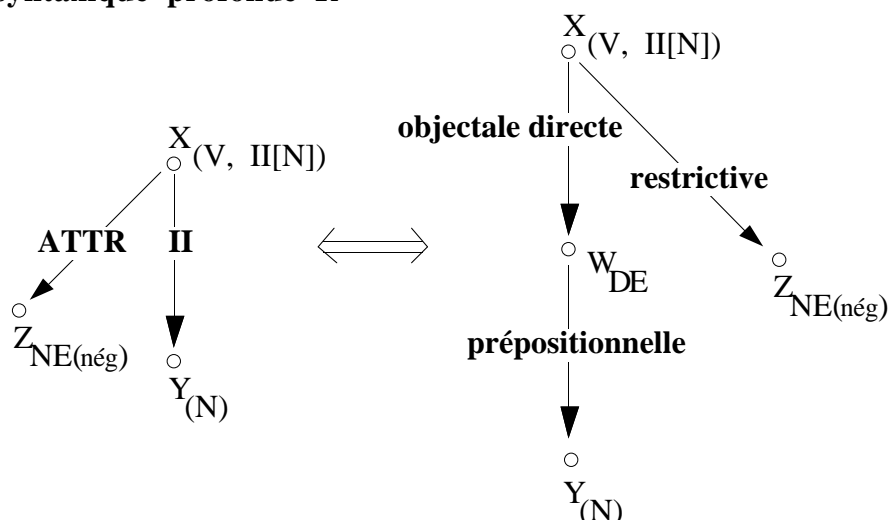


[L'indice entre parenthèses représente le syntactique de la lexie L associée au nœud en question, c'est-à-dire l'ensemble de données sur la combinatoire de L.]

C'est une règle de *régime*, qui introduit, dans la SSyntS, la préposition demandée par telle ou telle lexie. L'expression « $X_{(II[PRÉP])}$ » signifie 'l'actant SyntP II de la lexie X doit être introduit par la préposition PRÉP'; les indications de ce type font partie de ce qu'on appelle le *schéma de régime* dans un article de dictionnaire (d'un DEC). Le *schéma de régime* de la lexie L est un tableau qui spécifie la forme de l'expression — dans la phrase — des *actants* de L, c'est-à-dire des éléments essentiels qui peuvent dépendre de L syntaxiquement et qui sont prévus par sa définition (cf. les variables X, Y, Z, ..., qu'on a vues dans nos exemples de représentations sémantiques ; voir Mel'čuk *et al.* 1995: 119, 221). Ainsi, en français on *aide qqn.*, mais en allemand *à qqn.* (*jemandem helfen*) ; en français, on *attend qqn.*, en anglais *pour qqn.* (*wait for somebody*), et en allemand ou en polonais *sur qqn.* (*auf jemanden warten*; *czekać na kogoś*) ; et ainsi de suite. Ce sont toutes les informations pertinentes pour ce type de phénomène qui sont présentées dans le schéma de régime de L. Les règles de la syntaxe (profonde et de surface) s'adressent au schéma de régime de la lexie traitée pour déterminer la préposition (ou la conjonction, ou le cas, etc.) nécessaire ; c'est cela que fait la règle R^3 . Elle produit, dans l'arbre syntaxique de surface (7), les deux sous-arbres comprenant des prépositions régies qui marquent l'actant syntaxique profond II de la lexie régissante : *doute* [= X] \uparrow *quant à* [= $Z_{(PRÉP)}$] *l'effet* [= Y] et *effet* [= X] *sur* [= $Z_{(PRÉP)}$] *la qualité* [= Y].

D'autres règles de la syntaxe profonde s'appliquent à la réalisation de configurations syntaxiques profondes, qui sont universelles, par les constructions syntaxiques de surface, qui, elles, sont typiquement françaises. En voici un exemple :

Règle syntaxique profonde R^4



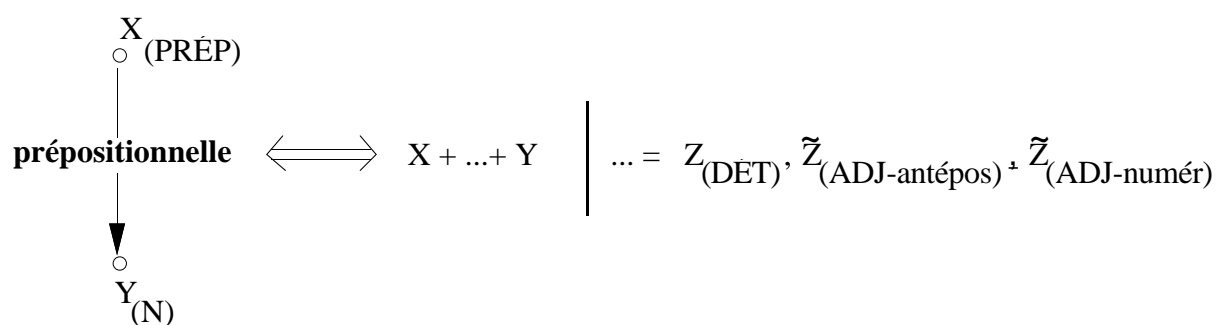
La règle SyntP R^4 introduit (dans la SSyntS) la préposition DE devant un objet direct d'un verbe qui reçoit la négation (*Je NE mange pas de viande*) ; elle fait juste cela dans l'arbre SyntS de (7).

De plus, une partie de règles SyntP «calculent» les valeurs des FL qui apparaissent dans la SSyntP, en consultant des articles de dictionnaire des lexies en question ; ainsi, la configuration $\text{Oper}_1 \text{II}; \text{---} \rightarrow ; : \text{DOUTE}$, qu'on voit dans l'arbre (6), est remplacée par la configuration $\text{AVOIRII}; \text{---} \rightarrow ; : \text{DOUTE}$ dans l'arbre (7).

Le nombre de règles SyntP est de l'ordre de quelques centaines (plus quelques dizaines de schémas de règles).

Syntaxe de surface

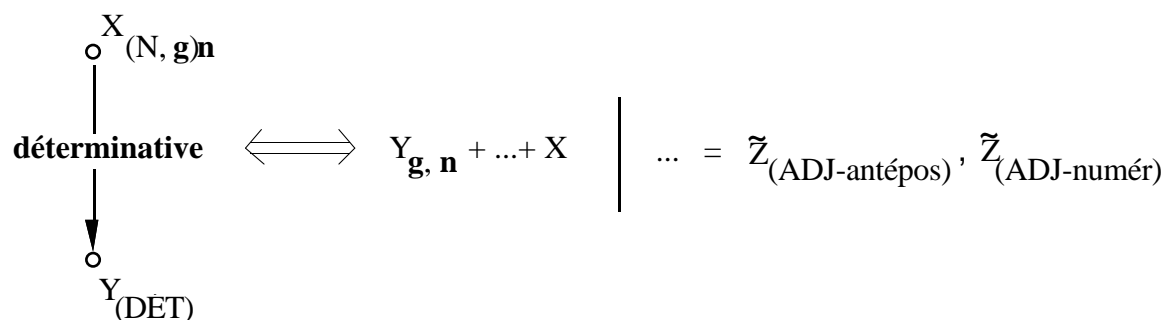
Règle syntaxique de surface R^5



[Les trois points entre les éléments dans la partie droite d'une règle SyntS indiquent la possibilité de séparation linéaire par certains mots-formes : selon R^5 , seuls des déterminants ainsi que des adjectifs antéposés et des numéraux avec leur dépendants peuvent apparaître entre la préposition X et le nom Y. Le tilde au-dessus d'un nœud Z — \tilde{Z} — signifie 'Z et ses dépendants', par exemple, $d' [= X]$ À PEU PRÈS TROIS CENTS [= \tilde{Z}] tonnes [= Y].]

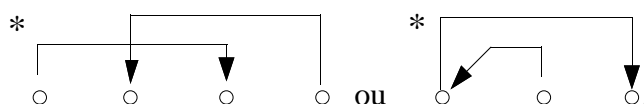
Cette règle est une règle typique de l'ordre des mots : elle positionne la préposition devant le nom qui en dépend, ne permettant entre les deux qu'un des éléments spécifiés dans la partie conditions. La règle R^5 établit la position linéaire de cinq prépositions en (8) : **DE** DOUTE, **QUANT À** [l'] EFFET, **DE** [son] ENGAGEMENT, **SUR** [la] QUALITÉ et **DE** [ses] ŒUVRES.

Règle syntaxique de surface R^6



Cette règle assure l'accord du déterminant Y en genre **g** et en nombre **n** avec le nom déterminé X (l'accord étant en général du domaine de la syntaxe de surface), ainsi que le positionnement linéaire du déterminant devant le nom. La règle R^6 produit, par exemple, les valeurs flexionnelles «pl, fém» pour le lexème pronominal SON en (8) (avant-dernier élément) et place ce lexème à gauche de ŒUVRE.

Les règles R^{5-6} n'illustrent qu'un seul type de règles syntaxiques de surface : les règles dites *locales*, dont l'opérande est juste une seule branche de l'arbre syntaxique de surface (pour une liste assez complète de règles syntaxiques de surface locales pour l'anglais, voir Mel'čuk & Pertsov 1987 ; les règles SyntS pour le sujet et les compléments d'objet de la langue aloutore sont données dans Mel'čuk & Savvina 1978). Mais la syntaxe de surface comprend également un autre type de règles — les règles *globales*, qui traitent des configurations de nœuds beaucoup plus complexes et/ou prennent en charge les données contenues dans la structure communicative de la phrase. Ce sont ces règles qui, entre autres, rendent compte de propriétés de l'ordre des mots telle que la *projectivité* (la langue évite, de façon générale, les ordonnancements non projectifs de mots-formes — du type



où deux arcs de dépendance se croisent ou où un arc en recouvre un autre qui sort d'un élément gouverneur ; cf. Iordanskaja 1963). Je ne peux cependant pas envisager ici les règles SyntS globales avec plus de détails (voir Mel'čuk 1967a et 1995a).

Le nombre total de règles syntaxiques de surface est à peu près d'un millier.

Morphologie

Morphologie profonde

Règles morphologiques profondes

R^7 ind, prés \iff {IND.PRÉS}

R^8 3, sg \iff {3SG}

R^9 pl \iff {PL}

Ce sont des règles dites *catégorielles* : grosso modo, elles regroupent les éléments de catégories flexionnelles, c'est-à-dire des grammèmes, sous des morphèmes, ensembles de signes linguistiques segmentaux élémentaires, c'est-à-dire morphes. Autrement dit, une règle catégorielle

indique qu'un tel ensemble de grammèmes est exprimé par un seul signe de la langue **L** (sans spécifier, pour autant, quel signe concret le fait).

Morphologie de surface

Règles morphologiques de surface

\mathbf{R}^{10}	{IND.PRÉS}	\iff	$-\emptyset$
\mathbf{R}^{11}	{3SG}	\iff	$-e$ verbe du groupe I, ...
		\iff	$-t$ verbe du groupe II ou III, ...
\mathbf{R}^{12}	{PL}	\iff	$-s$ nom ordinaire
		\iff	$-x$ nom du type «pl-x»
\mathbf{R}^{13}	{ÊT(-re)}, {IND.PRÉS}, {3SG}	\iff	est

Ici nous voyons la sélection — par des règles dites *morphémiques* — des morphes spécifiques appartenant aux morphèmes calculés à l'étape précédente ; cette sélection se fait selon les conditions morphologiques, que je n'ai esquissées qu'approximativement.

La règle \mathbf{R}^{13} illustre un cas de figure spécial : la réalisation d'un ensemble de morphèmes par un seul signe indivisible, appelé *mégamorphe* ; c'est un cas de supplétion.

Les règles \mathbf{R}^{7-13} permettent d'obtenir, entre autres, les mots-formes *affecte*, *est* et *œuvres*, apparaissant dans la phrase (5).

La composante morphologique de surface comprend également les règles *morphologiques* : ce sont les règles qui effectuent des changements phonologiques dans les mots-formes synthétisés. Une règle morphologique est dite profonde, si elle est contrainte par des conditions morphologiques, et de surface autrement (ses conditions d'application ne réfèrent qu'au contexte phonologique).

Règle morphologique profonde \mathbf{R}^{14}

$$/\Lambda/ + \implies /s/ + \left\{ \begin{array}{l} 1) \text{ verbe du II}^e \text{ groupe ;} \\ 2) \text{ soit } ____\{ /V/; /j/ \}, \\ \text{soit [le verbe est au subj ou à l'ind, prés, 3pl]} \end{array} \right.$$

Cette règle effectue l'insertion de /s/ à la fin du radical (indiquée par «+») dans les verbes du II^e groupe — soit devant une voyelle ou un yod, soit au subjonctif ou à la 3^e personne pluriel du présent de l'indicatif : par exemple, *grossiss+ions* [devant le yod] ou [*que je*] *grossisse* [au subjonctif]. (Le symbole Λ désigne la chaîne vide.)

Règle morphologique de surface \mathbf{R}^{15}

$$/i/ \implies /j/ \mid ___ /V/$$

La règle \mathbf{R}^{15} dévocalise tout /i/ devant une voyelle : par exemple, dans *étudie* ~ *étudions*, où l'on a /etüdi/ ~ /etüdjo/.

Le nombre total de règles morphologiques et morphonologiques des quatre types illustrés précédemment varie, selon la langue, de quelques dizaines à quelques centaines. (Pour tous les détails concernant la description de la morphologie, on peut consulter Mel'čuk 1993-1997.)

Dans cet exposé, je ne présenterai pas de règles de la composante phonologique.

En résumé, je peux dire que la caractéristique la plus générale et la plus saillante du MST en question est son caractère STRATIFICATIONNEL et MODULAIRE (ces deux propriétés étant les deux côtes d'une même médaille). Le principe directeur, qui prévaut dans tous les conflits théoriques, est de séparer le plus possible la représentation et la manipulation des phénomènes linguistiques qui, du point de vue de l'intuition du chercheur, sont de nature différente. Par exemple, chercher à construire une représentation unifiée de la phrase qui exprime SIMULTANÉMENT les liens hiérarchiques, anaphoriques et linéaires entre les éléments lexicaux, ce que certains linguistes veulent faire à tout prix, est un tabou absolu pour le MST. Son mot du jour est «distinguer et séparer» ; chaque aspect identifiable d'un fait à décrire reçoit une représentation autonome, pour établir ensuite des règles de correspondance entre ces diverses représentations.

3. La modélisation de deux phénomènes linguistiques particuliers : les choix lexicaux et la cooccurrence lexicale

Après avoir esquissé le MST dans sa globalité, j'aimerais aborder deux problèmes linguistiques particuliers tendant à montrer les avantages pratiques qu'offre la démarche «Sens-Texte» dans l'étude des langues. Le premier relève de la sémantique lexicale (le sens des mots) et le second, de la cooccurrence lexicale restreinte (la combinatoire des mots).

Comme on le sait, depuis Saussure, Hjelmslev et Jakobson, l'activité langagière se déroule selon deux axes : l'axe PARADIGMATIQUE, ou «vertical», où le locuteur fait des choix à partir des unités qui s'opposent et s'excluent («Soit X, soit Y» — lequel choisir ?) ; et l'axe SYNTAGMATIQUE, ou «horizontal», où le locuteur combine des unités déjà sélectionnées («X et Y» — peut-on les mettre ensemble ?). Autrement dit, il s'agit des CHOIX LEXICAUX et de la COOCCURRENCE LEXICALE.

3.1. Modélisation des choix lexicaux paradigmatiques

Les énoncés *Orwell est certain que P* et *Orwell ne doute pas que P* sont synonymes : 1) ils peuvent être substitués l'un à l'autre ; 2) leurs négations sont quasi synonymes (sans l'être parfaitement, voir plus loin) : *Orwell n'est pas certain que P* *Orwell doute que P* ; 3) ils sont tous les deux non factifs, c'est-à-dire qu'aucun ne présuppose la vérité de P : chacun peut être suivi de ... *mais en fait P n'a pas lieu* ; etc. Qu'est-ce que le locuteur a dans son cerveau à propos des lexies CERTAIN et DOUTER/AVOIR DES DOUTES qui lui permet de les manipuler comme il le fait ? Nous ne pouvons pas le savoir ; cependant, nous pouvons en proposer un modèle fonctionnel.

En suivant A.Zholkovsky, A.Bogusławski et A.Wierzbicka⁸, nous avançons l'hypothèse que le sens de ces lexèmes s'exprime à partir de sens plus simples, c'est-à-dire que ce sens est DÉCOMPOSABLE. Considérons la série suivante de combinaisons de propositions et d'équations sémantiques (le dièse # indiquant une continuation inacceptable) :

- (11) a. *Je crois que Pierre est venu, mais je n'en suis pas certain.*
 b. *Je suis certain que Pierre est venu, mais #je ne le crois pas.*
 c. *Je crois que Pierre est venu, mais #j'en doute.*
 d. *Je suis certain que Pierre est venu = Je ne doute pas que Pierre est venu.*
 e. *Je ne suis pas certain que Pierre est venu Je doute que Pierre soit venu.*

Pour que le locuteur (ou un programme informatique) puisse produire les expressions en (11), en établissant leur acceptabilité ainsi que leur synonymie, il suffit de représenter le sens des lexies en question comme suit :

(12) a. *X est certain que P* : ‘Ayant la croyance «P a lieu», |
X n'est pas disposé à admettre que P n'a pas lieu’.

b. *X doute que P* : ‘N'ayant pas la croyance «P a lieu», |
X est disposé à admettre que P n'a pas lieu’⁹.

[La partie d'une définition lexicographique à gauche de la barre verticale «|» est un présupposé.]

Avec ces définitions, nous obtenons, pour les phrases (11), les décompositions sémantiques suivantes (les contradictions apparaissent relevées en gras) :

(13) a. ‘Je crois que Pierre est venu, mais [tout] en ayant la croyance «Pierre est venu», | je suis disposé à admettre que Pierre n'est pas venu’.

b. ‘**Ayant la croyance «Pierre est venu»,** | je ne suis pas disposé à admettre que Pierre n'est pas venu, mais [#]**je ne crois pas que** [= je n'ai pas la croyance] **«Pierre est venu»**’.

c. ‘**Je crois que Pierre est venu,** mais [#]**n'ayant pas la croyance «Pierre est venu»,** | je suis disposé à admettre que Pierre n'est pas venu’.

d. ‘Ayant la croyance «Pierre est venu», | je ne suis pas disposé à admettre que Pierre n'est pas venu = ‘Ayant la croyance «Pierre est venu» [ce qui résulte de la négation de la composante correspondante¹⁰], | je ne suis pas disposé à admettre que Pierre n'est pas venu’.

e. ‘Ayant la croyance «Pierre est venu», | je suis disposé à admettre que Pierre n'est pas venu’.

‘N'ayant pas la croyance «Pierre est venu», | je suis disposé à admettre que Pierre n'est pas venu’.

En (11e), on a l'impression que la phrase avec *je doute* exprime un degré plus élevé d'incertitude que sa contrepartie ; cette intuition est reflétée dans les décompositions de (13e). En effet, dans les deux phrases, je suis disposé à admettre que Pierre n'est pas venu ; cependant, dans la première, j'ai quand même la croyance «Pierre est venu», qui constitue le présupposé de la phrase, alors que dans la deuxième, je ne l'ai pas.

On voit donc que, si nous adoptons les décompositions sémantiques proposées pour les lexies CERTAIN et DOUTER, nous pouvons décrire les faits observés de façon formelle, régulière et cohérente. La démarche Sens-Texte nous permet donc de postuler les règles (12a-b), qui sont des éléments constitutifs de la composante sémantique (= du module sémantique) du MST du français. Plus précisément, ce sont des règles sémantiques lexémiques (comme la règle **R**¹, p. 00) — articles ou parties d'articles de dictionnaire qui composent le *Dictionnaire explicatif et combinatoire*, ou DEC, dont il a déjà été question. Les règles (12a-b) représentent en fait les définitions lexicographiques des lexies citées. Ce sont des DÉCOMPOSITIONS SÉMANTIQUES, ou des «formules moléculaires» du sens, établies selon des règles et des critères qui sont présentés dans Mel'čuk *et al.* 1995: 78 *ssq.*

A part la définition du sémantisme de la lexie vedette L, un article du DEC comporte d'autres données lexicographiques importantes, en particulier la description des propriétés combinatoires de L, que nous allons maintenant examiner.

3.2. Modélisation des choix lexicaux syntagmatiques : les fonctions lexicales

La combinatoire lexicale restreinte constitue un élément fort problématique pour toute description dictionnaire et, par ricochet, pour la linguistique théorique. Le caractère capricieux et imprévisible de la cooccurrence lexicale est notoire. On dit *plonger dans le désespoir*, mais *mettre en rage* (<*mettre en désespoir, *plonger dans la rage). On peut être *gravement* ou *grièvement blessé*, mais seulement *gravement malade* (<*grièvement malade) (le fameux exemple de Bally) ; de plus, on a *de grands blessés*, alors qu'une blessure est *grave*, pas *grande. On est *très fatigué*, mais *bien reposé* (<*très reposé [comme dans *Je suis très reposé]), *fabuleusement riche*, mais *bien connu* (<*bien riche, *fabuleusement connu) ; et ainsi de suite. En français, on *fait* un pas, en espagnol, on le *donne* (= *dar un paso*), et en anglais, on le *prend* (= *take a step*). La pluie est *forte* en français, mais *lourde* (= *heavy rain*) en anglais. On *fait* un rêve en français, mais on l'*a* en anglais (= *have a dream*) ; vous *donnez* un cours en français, mais vous *l'enseignez* (= *teach a course*) en anglais et vous le *lisez* (= *čitat' kurs*) en russe (même si vous ne le lisez pas !). Toutes ces bizarreries d'usage sont connues depuis longtemps : ce sont des *collocations* — des expressions phraséologiques d'un certain type, qu'on ne peut pas prévoir et que la linguistique ne savait pas présenter de façon logique et facilement calculable. Or l'approche Sens-Texte propose un modèle fonctionnel de ce phénomène : les *fonctions lexicales*.

Le concept de fonction lexicale repose sur l'hypothèse que les cas de cooccurrence lexicale restreinte du type ci-dessus se rencontrent, le plus souvent, avec un nombre fort réduit de sens spécifiques — très abstraits et généraux. Le sens 'intense/intensément, très' en est un exemple typique : son expression n'est pas libre, à la différence de n'importe quel sens «normal», mais dépend de la lexie auprès de laquelle ce sens est exprimé. Cf. :

'intensément'(dormir)	= profondément, comme une souche, comme un loir, à poings fermés, ...
'intensément'(blessé _A)	= gravement, grièvement
'intense'(blessé _N)	= grand antépos
'intense'(pluie)	= forte < torrentielle, diluvienne ; battante
'intense'(froid _N)	= de canard, de chien, de loup, de tous les diables, du diable, sibérien
'intensément'(nier)	= catégoriquement

Le sens d'intensification s'exprime au voisinage d'une lexie *x* par une ou plusieurs lexie(s) *y*, et ceci de façon irrégulière, mais toujours en fonction de *x*. Par conséquent, l'expression de ce sens peut être décrite par une fonction (au sens mathématique du terme) **f** qui associe, à tout *x* pour lequel ce sens peut être exprimé, tous les *y* possibles : **f**(*x*) = *y*. La lexie *x* pour laquelle on cherche le(s) cooccurrent(s) exprimant ce sens est l'**argument** de **f**, et l'ensemble des cooccurrents est sa **valeur**. Si on donne au sens en question le nom **Magn** (du lat. *magnus* 'grand'), on peut utiliser la notation fonctionnelle bien connue :

Magn (dormir)	= profondément, comme une souche, comme un loir, à poings fermés, ...
Magn (blessé _A)	= gravement, grièvement
etc.	

Les fonctions comme **Magn** sont appelées des fonctions lexicales [= FL], puisque leurs arguments et leurs valeurs sont exclusivement des lexies.

Deux faits importants sont à noter : d'une part, les FL du type illustré ci-dessus — les FL *standard* — sont PEU NOMBREUSES (une soixantaine à peu près) et, d'autre part, elles sont UNIVERSELLES : elles existent dans toutes les langues et suffisent (presque¹¹) à décrire, de façon systématique et formelle, l'ensemble des collocations. Sans entrer dans les détails, je me limiterai à trois groupes d'exemples de FL (le lecteur intéressé pourra consulter Mel'čuk 1993c, 1996 ou Mel'čuk *et al.* 1995: 125 *ssq.*, ainsi que les volumes publiés du DEC du français).

1. **Bon** 'bon, tel que le locuteur l'approuve' [louange standard consacrée par la langue] :

Bon (conseil)	= précieux	AntiBon (victoire)	= à la Pyrrhus
Bon (temps)	= beau	AntiBon (temps)	= de chien
Bon (choix)	= heureux	AntiBon (choix)	= malheureux
Bon (se porter)	= comme un charme	AntiBon (résultat)	= piteux

2. **Oper**, **Func** et **Labor** sont trois types syntaxiquement différents de ce qu'on appelle les *verbes supports* :

Oper ₁ (suprémie)	= détenir [ART ~]	Func ₀ (réunion)	= est en cours
Oper ₁ (remarque)	= faire [ART ~]	Func ₁ (aide)	= vient [de N]
Oper ₁ (méfait)	= perpétrer [ART ~]	Func ₁ (responsabilité)	= incombe [à N]
Oper ₂ (danger)	= courir [ART ~]	Func ₂ (danger)	= menace [N]
Oper ₂ (applaudissements)		= recueillir [ART ~]	Func ₂ (liste) =
comprend [N]			
Oper ₃ (conseil)	= recevoir [ART ~]	Func ₂ (interdiction)	= frappe [N]
	Labor ₁₂ (liste)	= mettre [N sur ART ~]	
	Labor ₁₂ (note)	= prendre [N en ~]	
	Labor ₁₂ (location)	= donner [N en ~]	
	Labor ₃₂ (location)	= prendre [N en ~]	

3. **Loc** est une préposition de localisation standard — spatiale ou temporelle :

Loc _{in/ad} (gare)	= en [[~] un train, à [ART ~] des personnes]	Loc _{in} (personnel)	= 'au sein' [de ART ~]
Loc _{in/ad} (ville)	= en [~]	Loc _{in/ad} (campagne)	= à [la ~]
Loctemp;in (régime)	= sous [le ~]	Loctemp;in (période)	= pendant [ART ~]
Loctemp;in (Antiquité)	= dans [l'~]	Loctemp;ab (Antiquité)	= depuis [l'~]

Au moyen des FL, le linguiste modélise les choix lexicaux syntagmatiques faits spontanément par le locuteur. Le MST d'une langue donnée inclut un dictionnaire (mentionné en 2.3, **R**¹, p. 00) : le DEC. Un article du DEC comprend, en plus de la définition de la lexie vedette L et de son schéma de régime, dont il a déjà été question (**R**³, p. 00), une liste de FL de L avec leurs valeurs — de sorte que le DEC couvre toutes les collocations de L¹².

On ne peut pas prouver de façon rigoureuse que les FL existent telles quelles dans le cerveau des locuteurs ; mais la systématisme et la cohérence de la description qu'elles assurent constituent un argument de poids [= **Magn**(argument) !] en leur faveur. C'est d'autant plus vrai que la décomposition sémantique et les FL d'une même lexie entretiennent des liens très étroits, ce qui renforce

encore notre conviction d'avoir correctement traité aussi bien la décomposition sémantique que les fonctions lexicales. On notera une ressemblance frappante entre la corrélation «structure sémantique de la lexie ~ ses FL» et la corrélation observée entre la structure moléculaire d'une substance (= sa formule structurale) et ses propriétés chimiques et physiques, entre autres sa combinabilité avec d'autres substances. En fait, les formules structurales de certains corps chimiques ont été établies comme des modèles fonctionnels de réactions dans lesquelles ces corps étaient impliqués.

Autrement dit, à une FL donnée de la lexie L, c'est-à-dire à $f(L)$, il doit correspondre une composante sémantique particulière ' σ ' dans la définition lexicographique de L ; le sens de la valeur de cette FL, ' $f(L)$ ', porte sur ' σ '. Cette correspondance permet d'étayer les décisions prises par le linguiste concernant le sens et la cooccurrence de L.

3.3. Corrélations entre la modélisation paradigmatique et syntagmatique des comportements lexicaux

Pour illustrer les liens structuraux qui existent entre nos hypothèses sémantiques et cooccurrence, considérons les deux exemples suivants.

Exemple 1

(14) a. *Je suis *très <*fort> certain que Pierre est venu ~ Je suis absolument <tout à fait> certain que Pierre est venu.*

vs

b. *Je doute fort que Pierre soit venu ~ Je doute *absolument <*tout à fait> que Pierre soit venu.*

Pourquoi CERTAIN ['sûr'] et DOUTER ont-ils des propriétés combinatoires opposées vis-à-vis des intensificateurs (= les valeurs de la FL **Magn**) ? En utilisant les décompositions (12a-b), il est facile de le découvrir. La composante sémantique centrale de CERTAIN est '**ne pas** [être disposé à ...]', et on ne peut intensifier un état nié 'ne pas être', puisqu'un état nié n'est pas en principe graduel ; *très* ou *fort* sont donc inacceptables ici. On peut, par contre, renforcer la négation dans la mesure où le locuteur veut insister sur cette composante sémantique : *absolument* ou *tout à fait* sont par conséquent tout naturels auprès de CERTAIN. Avec DOUTER, la composante centrale est '**être disposé** [à ...]', et cet état présuppose un degré : on est nécessairement plus ou moins disposé ; d'où l'acceptabilité des intensificateurs *très* et *fort*. Dans ce cas, le choix de deux types différents de valeurs de la FL **Magn** est sémantiquement justifié. Notons de plus qu'on peut être 'PEU disposé' et qu'en conséquence on peut dire, pour exprimer ce sens, *Je doute peu* ou *J'ai peu de doutes*, ce qui prouve que 'disposé' est une meilleure composante sémantique pour 'douter/doute' que, par exemple, 'prêt' : on ne peut pas être '*peu prêt'. Notre hypothèse sur la structure sémantique des lexies CERTAIN et DOUTER se trouve ainsi confirmée.

Exemple 2

Le nom APPLAUDISSEMENT est défini dans le *Petit Robert* comme 'battement des mains en signe d'approbation, d'admiration ou d'enthousiasme'. Au premier abord, cette définition ne soulève pas d'objection. Cependant, on dit des *applaudissements nourris/clairsemés*, c'est-à-dire qu'on peut préciser le «degré» des applaudissements — alors que rien dans la définition citée n'admet la graduation. Ainsi, la cooccurrence avec NOURRI et CLAIRSEMÉ nous impose la présence, dans la définition lexicographique de APPLAUDISSEMENT, de la composante 'avec une

force et une fréquence données³ : et c'est cette composante qui «accepte» l'intensification. Aussitôt qu'on introduit cette composante, on découvre que la définition en cause a un autre défaut : elle n'indique pas que la force et la fréquence des battements sont proportionnelles au degré d'approbation, d'admiration ou d'enthousiasme ; il faut donc ajouter la composante correspondante, ce qui donne 'battement des mains en signe d'approbation, d'admiration ou d'enthousiasme, avec une force et une fréquence données, proportionnelles au degré d'approbation, d'admiration ou d'enthousiasme³. Comme on le voit, la cooccurrence lexicale restreinte d'une lexie L est liée au sémantisme de L — en ce sens que la présence de telle composante sémantique dans la définition de L 1) détermine l'existence de telle FL **f(L)** et 2) justifie, jusqu'à un certain point et dans certains cas, certains éléments de la valeur de **f(L)** — comme, par exemple, pour CERTAIN *vs* DOUTER (même si en règle générale, les éléments de la valeur d'une FL **f(L)** sont, bien entendu, imprévisibles). Puisque la cooccurrence lexicale restreinte est observable, elle nous permet de mieux découvrir le sémantisme des lexies, qui, lui, n'est pas directement observable.

3.4. Universalité des modèles Sens-Texte

Les MST — tels que présentés ci-dessus — sont universels : les techniques descriptives et les formalismes proposés s'appliquent à toutes les langues de la même façon. Je pourrais citer en exemple le cas de la cooccurrence lexicale restreinte. L'appareil de fonctions lexicales permet de décrire la cooccurrence lexicale de n'importe quelle langue avec des moyens identiques, en facilitant ainsi la comparaison des langues (et, en particulier, la traduction réciproque). Voici quelques valeurs des FL **Magn** (intensificateur) et **Oper₁** (un des verbes supports) pour plusieurs langues de structure très différente : allemand, russe, hongrois, arabe et chinois ; les *mots-clés* (= arguments des FL en question) sont sémantiquement équivalents (sauf dans le cas du sens 'excuses' [comme dans *présenter ses excuses*], pour lequel l'allemand et le hongrois n'ont pas d'expression nominale, et dans celui du sens 'résistance', dont l'expression nominale n'existe pas en chinois).

Allemand

Magn (<i>Regen</i> 'pluie')	= <i>starker</i> 'fort', <i>Platz-</i> 'éclatant'
Magn (<i>Argument</i> 'argument')	= <i>gewichtiges</i> 'de poids', <i>schlagendes</i> 'frappant', <i>unschlagbares</i> 'imbattable', <i>unwiderlegbares</i> 'irréfutable'
Magn (<i>Applaus</i> 'applaudissements')	= <i>tosender</i> 'mugissant'
Oper₁ (<i>Reise</i> 'voyage')	= [ART ~ _{acc}] <i>machen</i> 'faire'
Oper₁ (<i>Übereinkunft</i> 'accord')	= [<i>über</i> ART ~ _{acc}] <i>erzielen</i> 'obtenir'
Oper₁ (<i>Widerstand</i> 'résistance')	= [N _{dat} ART ~ _{acc}] <i>leisten</i> 'livrer'
Oper₁ (<i>Entschuldigung</i> 'pardon')	= [N _{acc} <i>um</i> ~ _{acc}] <i>bitten</i> 'prier'

Russe

Magn (<i>dožd'</i> 'pluie')	= <i>sil'nyj</i> 'fort', <i>prolivnoj</i> 'd'averse'
Magn (<i>dovod</i> 'argument')	= <i>veskij</i> 'de poids', <i>ubeditel'nyj</i> 'convaincant'
Magn (<i>aplodismenty</i> 'applaudissements')	= <i>burnye</i> 'tempétueux', <i>gromovye</i> 'de tonnerre'

Oper ₁ (<i>putešestvie</i> ‘voyage’)	= <i>sovershit</i> ‘accomplir’ [~e]
Oper ₁ (<i>soglašenie</i> ‘accord’)	= <i>pridti</i> [k ~ju] ‘venir à’
Oper ₁ (<i>soprotivlenie</i> ‘résistance’)	= <i>okazat</i> ‘[un verbe vide] manifester’ [N _{dat} ~e]
Oper ₁ (<i>izvinenija</i> ‘excuses’)	= <i>prinesti</i> ‘apporter’ [N _{dat} (svoi) ~ja]

Hongrois

Magn (<i>esó</i> ‘pluie’)	= <i>zuhogó</i> ‘torrentiel’
Magn (<i>érv</i> ‘argument’)	= <i>komoly</i> ‘sérieux’
Magn (<i>taps</i> ‘applaudissements’)	= <i>viharos</i> ‘tourbillonnant’, <i>vas-</i> ‘de fer’
Oper ₁ (<i>utazás</i> ‘voyage’)	= [~t] <i>tenni</i> ‘faire’
Oper ₁ (<i>megegyezés</i> ‘accord’)	= [~re] <i>jutni</i> ‘arriver à’
Oper ₁ (<i>ellenállás</i> ‘résistance’)	= [~t] <i>kifejteni</i> ‘développer’, <i>tanusítani</i> ‘démontrer’
Oper ₁ (<i>bocsánat</i> ‘pardon’)	= [N _{abl} ~ot] <i>kérni</i> ‘demander’

Arabe

Magn (<i>maṭar</i> ‘pluie’)	= <i>yaẓīr</i> ‘abondant’, <i>qawijj</i> ‘fort’
Magn (<i>huẓẓa</i> ‘argument’)	= <i>dāmiya</i> ‘frappant’, <i>qawijja</i> ‘fort’
Magn (<i>taṣṭīq</i> ‘applaudissements’)	= <i>ḥarr</i> ‘chaud’
Oper ₁ (<i>safar</i> ‘voyage’)	= <i>qāma</i> [bi ~] ‘se lever à, partir en’
Oper ₁ (<i>ʔittifāq</i> ‘accord’)	= <i>tawaṣṣala</i> [ʔila ~] ‘arriver à, obtenir’
Oper ₁ (<i>muqāwamat</i> ‘résistance’)	= <i>qāma</i> [bi ~] ‘se lever à, partir en’
Oper ₁ (<i>ʔiʕīdarāt</i> ‘excuses’)	= <i>qaddama</i> [ART ~] ‘avancer [trans.]’

Chinois

Magn (<i>yǔ</i> ‘pluie’)	= <i>dà</i> ‘grand’
Magn (<i>lùnjù</i> ‘argument’)	= <i>yǒulì-de</i> ‘de force’
Magn (<i>zhǎngshēng</i> ‘applaudissements’)	= <i>léidòng</i> ‘de tonnerre’ postpos
Oper ₁ (<i>lǜtú</i> ‘voyage’)	= <i>tàshang</i> [~] ‘marcher sur’
Oper ₁ (<i>xiéyí</i> ‘accord’)	= <i>dáchéng</i> [~] ‘arriver à’
Oper ₁ (<i>qiàn</i> ‘excuses’)	= <i>dào</i> [yige ‘une’ ~] ‘dire’

Ces exemples montrent à quel point les valeurs d'une même FL, pour des lexies sémantiquement équivalentes, varient d'une langue à autre : la pluie est «forte» en allemand, en russe et en arabe, mais elle est «grande» en chinois ; les applaudissements sont «de fer» en hongrois, «mugissants» en allemand et «tempétueux» en russe (*nourris* en français !) ; et ainsi de suite. Cependant les FL en garantissent une description homogène et systématique — dans le dictionnaire de chaque langue.

Bien plus, grâce à l'universalité des FL, on peut noter facilement et de façon compacte les correspondances INTERlinguistiques entre des valeurs de FL, ce qu'il est presque impossible de faire par un autre moyen. En effet, comment indiquer, dans un dictionnaire anglais-français, que, dans des contextes particuliers, [to] *develop* peut se traduire par *prendre* ou [to] *wean* par *détourner* ? Et

pourtant, [*to*] *develop a habit* se traduit par *prendre une habitude*, et [*to*] *wean from a habit* par *détourner d'une habitude*. Si on fait appel aux FL, ces correspondances s'expriment explicitement et systématiquement — en mettant les valeurs des FL pertinentes dans les articles de dictionnaire du lexème anglais HABIT et du lexème français HABITUDE. Le tableau suivant illustre ce point :

	angl. HABIT	↔	fr. HABITUDE
IncepOper₁	[<i>to</i>] <i>acquire, develop, form</i> [ART ~], <i>get</i> [<i>into</i> ART ~], <i>take</i> [<i>to</i> ART ~]		<i>contracter, prendre</i> [ART ~]
FinOper₁	[<i>to</i>] <i>drop</i> [ART ~], <i>get out, get rid</i> <i>of</i> [ART ~]		<i>abandonner, perdre</i> [ART ~]
LiquOper₁	[<i>to</i>] <i>break</i> [N <i>of</i> ART ~], <i>wean</i> [N] <i>away</i> [<i>from</i> ART ~]		<i>détacher, détourner</i> [N <i>de</i> ART ~]
Liqu₁Oper₁	[<i>to</i>] <i>break off, kick, shake off, throw</i> <i>off</i> [ART ~]		<i>se débarrasser, se défaire</i> [<i>de</i> ART ~], <i>renoncer</i> [à ART ~], <i>rompre</i> [<i>avec</i> ART ~]
CausFunc₁	[<i>to</i>] <i>instill</i> [ART ~ <i>in</i> (<i>to</i>) N]		<i>inculquer</i> [ART ~ à N]

En spécifiant de façon indépendante les valeurs des FL pour chaque lexie de chaque langue, nous obtenons automatiquement — par l'intermédiaire des FL — des équivalences intra- ou interlinguistiques. Cet aspect de l'utilisation des FL, souligné dès le moment de leur invention (voir la note 12, p. 00), ouvre une perspective nouvelle sur l'élaboration de dictionnaires bilingues, surtout dans le domaine du traitement automatique du langage (Iordanskaja & Mel'čuk 1997) ; mais ce sujet est trop pointu pour que je m'y étende.

Le caractère universaliste des modèles du type Sens-Texte se manifeste également dans d'autres domaines. Ainsi, les structures de dépendance utilisées en syntaxe permettent de décrire de façon homogène toutes les constructions trouvées dans les langues les plus «exotiques» ; les techniques proposées en morphologie (la séparation des règles catégorielles, morphémiques et morphologiques ; le concept universel de procédé morphologique ; ...) semblent couvrir tous les phénomènes que présente la littérature à ce sujet. Cependant, faute de place, je me contenterai du seul exemple de FL, analysé ci-dessus.

4. La linguistique Sens-Texte

Je suis maintenant en mesure d'ébaucher les contours du domaine spécifique de la linguistique dont je m'occupe : la linguistique Sens-Texte. Les tâches que les linguistes doivent affronter aujourd'hui sont énormes. Afin de les simplifier et d'homogénéiser les méthodes et les formalismes utilisés, je pense qu'il est bon, autant que faire se peut, de restreindre le champ d'étude. En conséquence, les quatre axes suivants ne sont pas considérés pour le moment dans le cadre Sens-Texte :

- l'axe temporel, c'est-à-dire la diachronie ;
- l'axe spatial, c'est-à-dire tout ce qui a trait au développement géographique des langues (la dialectologie et les recherches apparentées) ;
- l'axe social, c'est-à-dire la sociolinguistique ;
- l'axe psychologique/neurologique, c'est-à-dire la psycholinguistique, y compris l'apprentissage des langues et le développement du langage chez l'enfant, le lien entre la langue et l'art verbal, etc.,

ainsi que la neurologie moderne, qui fait des progrès considérables dans l'étude des bases cérébrales du langage.

Cependant, ces axes de recherche sont absolument vitaux pour la linguistique en général et pour la linguistique Sens-Texte, en particulier. Je suis convaincu que la vérification ultime de la pertinence des modèles Sens-Texte ne peut venir que des études menées dans ces directions, surtout dans celles de la psycholinguistique et de la neurolinguistique. Ce que je propose n'est donc qu'une tactique, mais une tactique d'importance : pour avancer de façon sensible dans la compréhension de la langue en tant que telle, il faut faire aujourd'hui abstraction de bon nombre de ses aspects et de ses facettes, même incontournables en principe. Pour élaborer des modèles simplifiés, mais bien formalisés, comme je veux le faire, il faut, dans un premier temps, restreindre le champ couvert (sans toutefois trop le restreindre, pour ne pas perdre d'éléments pertinents).

La linguistique Sens-Texte se réserve donc trois objectifs majeurs.

1. (a) Élaborer des descriptions formelles et précises de langues naturelles qui puissent servir de base fiable et commode à la construction des modèles Sens-Texte pour les langues correspondantes ;

(b) élaborer en parallèle la théorie de ces descriptions — la théorie Sens-Texte —, en déterminant les moyens nécessaires à la description linguistique.

C'est la linguistique Sens-Texte théorique.

2. (a) Élaborer les MST les plus complets possibles pour différentes langues ;

(b) élaborer en parallèle la théorie de ces modèles, en généralisant les observations faites sur les dispositifs particuliers qui assurent le passage «sens langagiers \implies textes langagiers».

C'est la linguistique Sens-Texte expérimentale.

3. Élaborer les «interfaces» avec deux autres disciplines voisines :

- D'une part, nous avons un domaine d'études qui doit s'occuper du passage «réalité extralinguistique \implies sens langagier» et pour lequel nous n'avons pas encore de nom. Je pense que ce passage doit se faire à travers une représentation du monde tel qu'il est perçu par le locuteur et qu'on pourrait parler ici de *Représentation conceptuelle* ; le domaine en question peut alors être appelé la CONCEPTIQUE.

- D'autre part, il s'agit de la PHONÉTIQUE, responsable du passage «textes langagiers (= transcription phonétique) \implies sons langagiers».

Je vais me concentrer, dans cette section, exclusivement sur le premier objectif (le deuxième est ébauché dans ce qui précède, et le troisième n'est pas encore suffisamment au point). Mise en garde importante : étant donné la place dont je dispose, je ne peux proposer que des caractéristiques incomplètes et approximatives ; je suis également obligé d'utiliser, sans explication, des concepts spéciaux hautement techniques, en espérant que les exemples et le contexte suffiront à les rendre intelligibles.

4.1. Les objectifs et les aboutissements de la linguistique Sens-Texte

La linguistique Sens-Texte comporte quatre divisions, également reconnues par les autres courants de la linguistique moderne, qui découlent de la structure du MST, présentée dans ce qui précède :

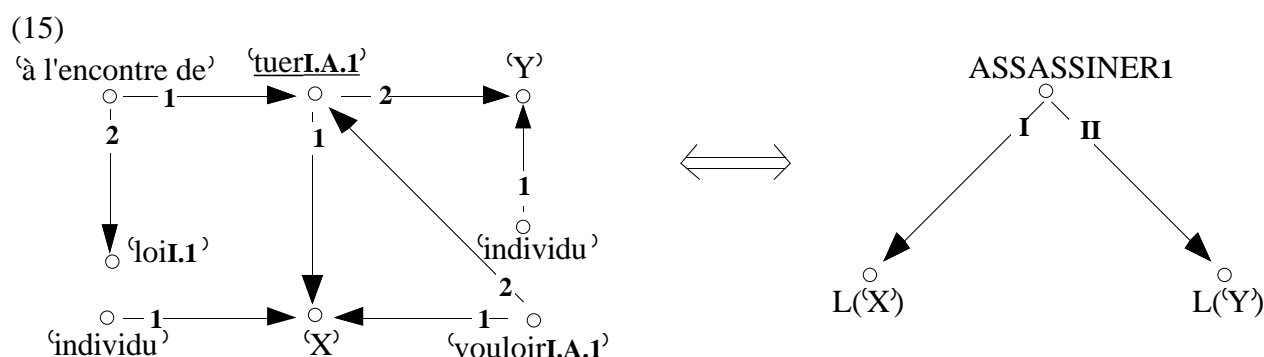
- la sémantique ;

- la syntaxe ;
- la morphologie ;
- la phonologie.

Je vais caractériser brièvement chacune de ces divisions, en indiquant certaines des lignes directrices de recherche, ainsi que certains des problèmes et des solutions que nous en proposons, dans le cadre de l'approche Sens-Texte.

SÉMANTIQUE

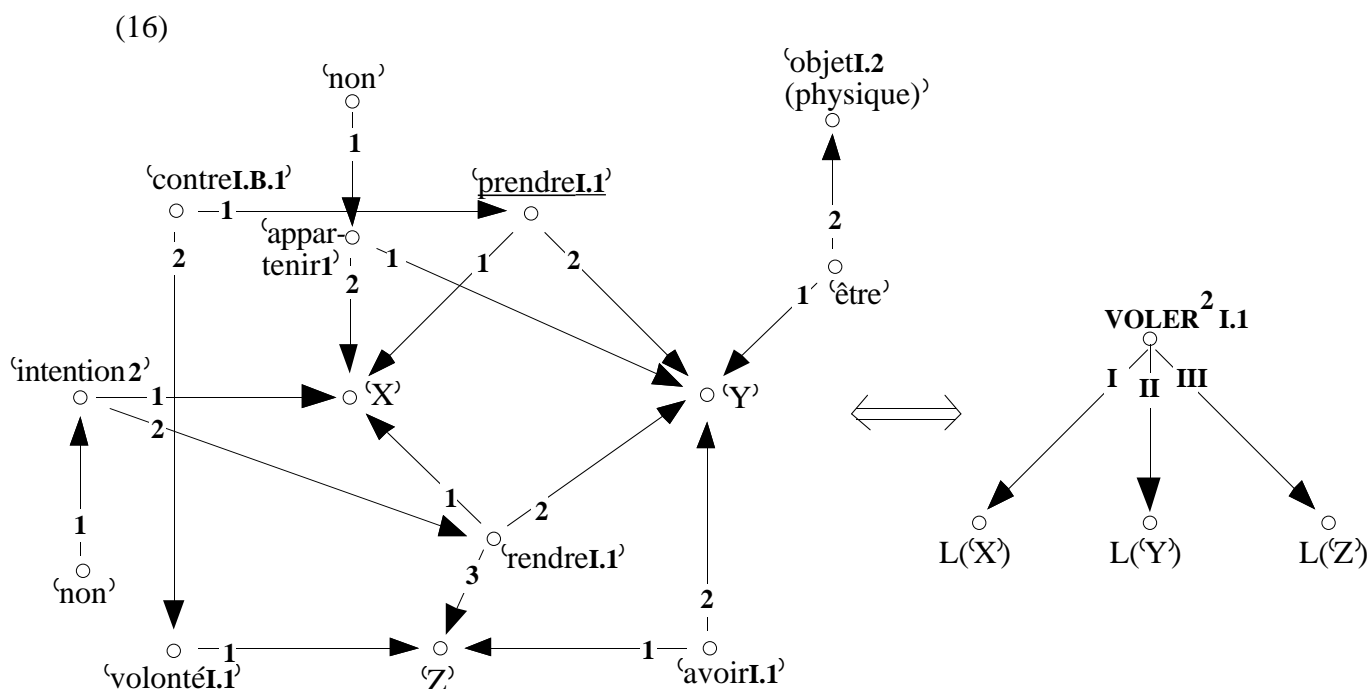
La sémantique linguistique Sens-Texte ne vise pas autre chose qu'un ensemble de règles assurant les correspondances entre les RSém et les RSyntP des phrases. Les règles les plus importantes sont les *règles sémantiques lexémiques*, dont on en a vu déjà un exemple : **R¹**, p. 00. Une règle sémantique lexémique établit la correspondance entre le sens décomposé d'un lexème et ce lexème ; par exemple :



[Le soulignement — dans la partie gauche de la règle — de l'élément 'tuer' l'identifie comme élément communicativement dominant, ou générique, du sens représenté ; 'assassiner' est donc un cas particulier de 'tuer' (sans cette indication, le réseau dans la partie gauche de (15) pourrait être interprété comme 'X veut tuer Y à l'encontre de la loi', ce qui ne correspond pas à ASSASSINER1) ; cf. la remarque après la RSém (4), p. 00. Comme je l'ai déjà dit, les numéros lexicaux identificateurs sont pris, ici et dans les autres exemples, dans le *Petit Robert*.]

Le lexème en cause ici est ASSASSINER1 ; son sens est 'l'individu X, voulant tuer l'individu Y, tue Y, ceci allant à l'encontre de la loi' (on ne peut pas **assassiner par négligence* ⟨*par hasard*⟩, ce qui justifie la composante 'voulant tuer Y' ; on ne peut ni **assassiner un animal*, ni **assassiner un ministère*, ainsi que ni un animal, ni un ministère ne peut assassiner non plus : on a donc besoin de la contrainte sémantique 'individu' sur X et Y ; tuer un soldat ennemi à la guerre ne s'appelle pas *assassiner*, de sorte que la composante 'à l'encontre de la loi' est aussi nécessaire). La partie droite de la règle précise les rôles syntaxiques profonds que doivent remplir auprès du verbe ASSASSINER1 les désignations de l'assassin X et de la victime Y ; ce sont les actants SyntP d'ASSASSINER1 : X est l'actant SyntP I, et Y l'actant SyntP II.

Les sens lexicaux peuvent être plus compliqués ; c'est le cas, par exemple, du lexème français VOLER²I.1 'prendre contre le gré de qqn.'. Citons la règle sémantique lexémique correspondante :



La RSém de $VOLER^2 I.1$ peut se lire comme 'X prend, contre la volonté de Z, l'objet physique Y qu'a Z, sans intention de rendre Y à Z' (*X vole Y à Z, comme dans *Le voisin lui a volé une poule**).

Je ne peux pas présenter ici les autres types de règles sémantiques : celles qui concernent les phrasèmes (en particulier, les fonctions lexicales), les grammèmes et les dérivatèmes, les constructions syntaxiques de toutes sortes et les prosodies) ; disons seulement que cette façon de faire — la DÉCOMPOSITION SÉMANTIQUE — s'applique à toutes les classes d'unités linguistiques, quelle que soit leur nature.

L'étude systématique de la structure «atomique» du sens langagier a été entamée à Moscou tout au début des années 60, au Laboratoire de Traduction Automatique, fondé et dirigé par V.Rosenzweig : voir Žolkovskij *et al.* 1961, Žolkovskij 1964 et MPiPL 1964. Cette étude, qui a été ensuite mise de l'avant par A.Wierzbicka dans ses nombreux ouvrages à partir du début des années 70 (voir, par exemple, son dernier livre Wierzbicka 1996) et développée dans l'œuvre importante d'Apresjan (1992, 1995), a changé la sémantique linguistique tout autant que la découverte de la structure atomique de la matière a changé la physique. Suite à cela, la sémantique Sens-Texte se consacre avant tout à la décomposition sémantique du lexique de la langue étudiée. Cette décomposition est effectuée selon des règles et des critères rigoureux, qui interdisent, entre autres, les cercles vicieux dans les définitions lexicographiques. En procédant de cette façon, nous tomberons inévitablement sur des *primitives sémantiques*, c'est-à-dire les sens langagiers indécomposables à l'intérieur de la langue. Un inventaire hypothétique universel de ces atomes sémantiques est proposé dans l'œuvre de Wierzbicka. (Dans le cadre du MST, cependant, l'existence des primitives sémantiques UNIVERSELLES n'est pas postulée. On s'attend à la découverte des primitives sémantiques dans chaque langue étudiée, après quoi on pourra se poser la question de leur universalité.)

Mais l'analyse sémantique des lexies ne peut se faire de façon isolée, sans référence à la syntaxe et à la cooccurrence lexicale (aussi bien libre que restreinte) de chaque lexie vedette. La tâche

du sémanticien Sens-Texte revient donc à élaborer des descriptions lexicales multifacettes (définition lexicographique + schéma de régime + fonctions lexicales), c'est-à-dire des articles du DEC, le dictionnaire déjà mentionné à plusieurs reprises. De cette façon, la lexicologie, ainsi que la lexicographie, sa contrepartie appliquée, se trouve au cœur des recherches sémantiques Sens-Texte. Dans cette perspective, l'importance du lexique est telle qu'on pourrait, sans trop exagérer, qualifier l'approche Sens-Texte de «lexicocentrique». Les deux DEC (partiels), l'un du russe (Mel'čuk & Zholkovsky 1984), l'autre du français (Mel'čuk *et al.* 1982-1997), sont les résultats tangibles de nos efforts dans cette voie. De plus, un récent ouvrage (Mel'čuk *et al.* 1995) traite des caractéristiques générales de l'élaboration d'un DEC.

Le travail lexicographique, si crucial soit-il, n'épuise pas, bien entendu, la sémantique Sens-Texte. Voici cinq autres tâches sémantiques prioritaires qui se posent dans le cadre Sens-Texte :

- Une élaboration plus poussée des réseaux sémantiques. Avant tout, on doit explorer le problème des primitives sémantiques (Mel'čuk 1989), ainsi que celui des éléments sémantiques aux importantes propriétés particulières qu'on appelle *étiquettes sémantiques* (par exemple, 'action', 'acte', 'état', 'événement', 'propriété', 'individu', 'substance', etc. : ces éléments influent beaucoup sur la structure des définitions lexicographiques où ils apparaissent, ainsi que sur la cooccurrence lexicale libre des lexies concernées ; une partie des étiquettes sémantiques pourraient bien être des primitives).

- La structuration interne de la représentation sémantique (Polguère 1992), qui concerne la distinction entre éléments posés et présupposés, ainsi que les contraintes sémantiques (comme, par exemple, 'individu' en (15)), les sens flexionnels, etc.

- Le rapport entre la RSém et une représentation plus profonde du monde extralinguistique : la Représentation conceptuelle (voir p. 00).

- Les structures communicative et rhétorique au sein de la RSém ; un livre sur ce sujet est en préparation (Mel'čuk 1997b).

- L'aspect procédural du passage entre RSém et RSyntP ; des résultats intéressants ce domaine se trouvent dans Iordanskaja & Polguère 1988 et Polguère 1990.

Il existe, bien sûr, nombre d'autres sujets sémantiques, envisagés par la linguistique Sens-Texte, tout aussi fascinants et tout aussi épineux.

SYNTAXE

La SYNTAXE PROFONDE présente quatre problèmes majeurs pour un syntacticien Sens-Texte :

- Le raffinement de l'inventaire des relations SyntP, c'est-à-dire, par exemple, l'exploration 1) de la distinction entre relation attributive restrictive et relation attributive descriptive, 2) de la nature de la relation entre un verbe de discours et le discours direct qu'il introduit (par exemple, [«*Mon Dieu !*»,]←--*soupira-t-elle*) et 3) de l'opposition «relations SyntP vs lexèmes marqueurs des constructions», etc.

- L'étude approfondie de la nature des actants SyntP : définition et typologie.

- L'étude de la correspondance entre actants Sém et actants SyntP : entre autres, le problème de la voix grammaticale en tant que catégorie flexionnelle (voir Mel'čuk 1993a).

- L'étude de pronominalisations de toute sorte.

La SYNTAXE DE SURFACE présente divers problèmes que nous regroupons en six points, dont trois premiers concernent la représentation SyntS et les deux derniers, le passage entre RSyntS et RMorphP :

- Le raffinement de l'inventaire des relations SyntS.
- L'étude approfondie de la nature des actants SyntS : définition et typologie.
- La correspondance entre actants SyntP et actants SyntS ; parmi les problèmes théoriques spécifiques liés à cette correspondance, signalons le problème bien connu de l'opposition «construction nominative vs la construction ergative» (Mel'čuk 1992b).
- La morphologisation de la SSyntS, c'est-à-dire le calcul des toutes les valeurs morphologiques déterminées par la syntaxe de la phrase : accord et régime (Mel'čuk 1993b).
- La linéarisation de la SSyntS : l'ordre des mots (en russe : Mel'čuk 1995 ; pour l'ordonnement des clitiques en français, voir Iordanskaja 1982).

Une description assez détaillée de la syntaxe de surface de l'anglais (en termes d'arbres de dépendance) est présentée dans Mel'čuk & Pertsov 1987.

MORPHOLOGIE

La MORPHOLOGIE PROFONDE présente tout un bouquet de problèmes linguistiques intéressants, dont les suivants :

- L'analyse et la définition du concept de catégorie flexionnelle (Mel'čuk 1991).
- La distinction entre flexion et dérivation ; à ce propos, voir une esquisse d'une théorie de la dérivation dans Mel'čuk 1990.
- La définition des catégories flexionnelles ; par exemple, une définition du concept de catégorie de cas grammatical est proposée dans Mel'čuk 1986a.
- L'inventaire des grammèmes d'une langue donnée **L**.
- La distribution des grammèmes établis par morphèmes ; voir, par exemple, le traitement d'un cas particulier — le dualisateur et le pluralisateur indépendants dans la langue aloutore — dans Mel'čuk 1986b.
- L'inventaire des dérivatèmes de **L**.

La MORPHOLOGIE DE SURFACE calcule, à partir d'une chaîne morphémique qui représente un mot-forme, la chaîne de phonèmes correspondante munie, si tel est le cas, de configurations de prosodèmes. Les problèmes correspondants peuvent être regroupés sous deux rubriques générales :

- L'inventaire et la description des alternances de **L**.
- La détermination de la position de l'accent de force dans les langues qui connaissent l'accent «ambulant» : accent qui change de position en fonction de facteurs morphologiques ; pour un exemple de modélisation du comportement de l'accent de ce type (hébreu moderne), voir Mel'čuk & Podolsky 1996.

Pour les descriptions quasi exhaustives des fragments fermés de la morphologie de trois langues typologiquement très différentes, le lecteur peut consulter Mel'čuk 1967a (la conjugaison en espagnol), 1968 (la déclinaison en hongrois) et Mel'čuk & Aroga Bessong 1983 (la conjugaison en bafia, langue semi-bantoue du Cameroun).

PHONOLOGIE

La PHONOLOGIE assure la correspondance entre les phonèmes/les prosodèmes, d'une part, et les phones/les prosodies (= symboles représentant les sons et les prosodies de la parole), d'autre part. Son intérêt réside, avant tout, dans l'établissement d'un inventaire de phonèmes/de prosodèmes de **L** tel qu'il permette une expression commode et en même temps élégante de tous les phénomènes morphologiques pertinents. Autrement dit, les problèmes phonologiques se résolvent, dans la linguistique Sens-Texte, dans le cadre de la production de mots-formes. Ainsi, en espagnol, pour assurer le bon fonctionnement d'un modèle de la conjugaison, on a besoin de trois séries de phonèmes apparentés : les voyelles /i/ et /u/, les glides /j/ et /w/, et les consonnes /j/ et /w/ (*pi*do, *pud*o ~ *pie*, *puede* ~ *hierba*, *huele*) ; voir Mel'čuk 1965. (La raison principale est l'existence des paires de verbes du type *cambiar* /kambjár/ ('changer') ~ *enviar* /embjár/ ('envoyer') : le premier a, à la 1^{re} personne du singulier (au présent de l'indicatif) la forme *cambio* /kámbjo/ <*/kambío/), et le second, *envío* /embío/ <*/émbjo/ ; pour les produire correctement, le radical du premier verbe doit être /kambj/, avec un glide, alors que celui du second a une voyelle : /embí/.)

Le survol rapide effectué ci-dessus suffit, me semble-t-il, à donner au lecteur une idée sinon très profonde, du moins assez structurée du champ d'étude dans l'approche Sens-Texte, de ses cibles et de quelques-uns de ses aboutissements. Passons maintenant au deuxième volet de la discussion : l'élaboration d'un métalangage linguistique unifié.

4.2. L'appareil conceptuel de la linguistique

«La linguistique est la seule science actuelle dont l'objet coïncide avec le discours qu'elle tient sur lui», a écrit C.Hagège (1986 : 81) ; on ne peut le dire mieux. Le problème est que les linguistes se servent des langues naturelles pour parler des langues naturelles. Il en découle inévitablement que la construction d'un appareil conceptuel assez formalisé et d'une terminologie correspondante a dans la linguistique une importance beaucoup plus grande que dans les autres sciences, où l'objet et l'instrument ne coïncident pas. C'est d'en avoir pris conscience qui m'a amené, il y a plus de 30 ans déjà, à me lancer dans l'élaboration d'un métalangage linguistique, basé sur un système conceptuel suffisamment formel pour assurer une vision générale, cohérente et homogène de la linguistique. On peut comparer ma démarche à celle du groupe mondialement connu des mathématiciens français qui, vers la fin des années trente, sous le nom d'emprunt de Nicolas Bourbaki, avaient cherché à créer une vision et un métalangage unifiés pour l'ensemble des mathématiques modernes (au sujet de l'œuvre de Bourbaki, voir Halmos 1957 et Queneau 1978). Mon travail s'est fortement inspiré des résultats bourbakistes (encore une autre *French connection* !). J'ai notamment proposé, pour la morphologie linguistique, un système déductif de concepts (Mel'čuk 1993-1997). À partir d'un nombre limité de concepts non définis (empruntés, d'une part, à la logique et aux mathématiques et, d'autre part, aux autres domaines de la linguistique — la sémantique, la syntaxe et la phonologie), je propose des définitions rigoureuses pour des concepts comme *signe (linguistique)*, *représentabilité de signes par d'autres signes*, ..., *mot-forme*, *lexème*, *racine* vs *affixe*, *catégorie flexionnelle* et *grammème*, *cas grammatical*, *voix*, ..., *morphe* vs *morphème*, *supplétion*, *réduplication*_{1/2}, *alternance*, *apophonie*, ... et des dizaines d'autres.

Je pense qu'il est utile de présenter ici un exemple concret qui démontre à mon avis de façon très claire à quel point une bonne terminologie sous-tendue par un système conceptuel bien formé peut être importante en linguistique théorique et descriptive.

Depuis des décennies on écrit que, dans les langues nilotiques du sud (le massai, le turkana, le teso, etc.), le cas nominatif et le cas accusatif ont des propriétés si exotiques que ces langues s'opposeraient à toutes les autres langues de la Terre. Ouvrons une grammaire (excellente d'ailleurs) du massai (Tucker and Mpayeei 1955) ; nous y apprenons que :

- l'accusatif est la forme de citation ;
- le nominatif se forme à partir de l'accusatif (par un changement tonal) ;
- le nominatif, mais non l'accusatif, est régi par des prépositions ;
- l'accusatif s'utilise pour marquer le thème antéposé au verbe ;
- la forme d'adresse est à l'accusatif, sauf si elle est introduite par une particule spéciale, qui la met au nominatif ;
- les adjectifs et les numéraux, en tant que compléments de la copule 'être', sont à l'accusatif ;
- le nominatif marque (mis à part le sujet syntaxique), les compléments d'objet indirects et obliques.

On ne trouve nulle part ailleurs un pareil faisceau de propriétés extravagantes pour ces deux cas, les plus importants dans les grammaires descriptives et dans la théorie linguistique. Ce traitement des cas massai, adopté pour toutes les langues de la famille nilotique, soulève de sérieux problèmes pour ce qui est de la description de la voix passive : le sujet du passif est dit être à l'accusatif, alors que le complément d'agent (qui existe dans certains parlars du massai) apparaît au nominatif. Se fondant sur les analyses disponibles à l'époque, E. Keenan (1976: 326-328) a même proposé toute une nouvelle théorie des voix. Or il s'agit tout simplement d'un quiproquo dû à une dénomination fautive : les africanistes ont appelé *nominatif* la forme du nom qui sert de sujet syntaxique à tous les verbes, et *accusatif*, la forme qui remplit le rôle d'objet direct ; c'est-à-dire qu'on a bel et bien suivi la grammaire... latine ! Pourtant, rien ne justifie d'appeler les cas grammaticaux d'après les fonctions syntaxiques qu'ils signalent : si on identifie le cas et la fonction syntaxique, il n'est plus besoin de deux concepts distincts. La dualité «cas ~ fonction syntaxique» est importante justement parce que la fonction syntaxique et les cas qui la marquent sont logiquement indépendants. Je propose que le nominatif soit TOUJOURS (= dans n'importe quelle langue) défini comme étant le cas de nomination des objets — conformément à l'étymologie du mot *nominatif*. Si nous l'admettons et, en conséquence, inversons les noms des cas dans les grammaires des langues nilotiques, tout se met en place : le «premier» cas du massai n'est pas un accusatif mais un nominatif, alors que ce qu'on appelle maintenant le nominatif est un oblique ; le nominatif est alors la forme de citation (par définition) et la forme de base à partir de laquelle l'oblique est formé — par des apophonies tonales. Le cas qui suit une préposition est bien l'oblique, comme on s'y attend, et le complément d'agent au passif est aussi à l'oblique, ce qui est aussi tout à fait naturel. Les langues nilotiques sont donc, du point de vue de leur système casuel, comme toutes les autres langues du monde. Elles possèdent quand même un trait spécifique : leur sujet syntaxique est toujours à l'oblique, et l'objet direct, au nominatif. Autrement dit, ces langues présentent — avec n'importe quel verbe, y compris la copule

— une construction ergative de la phrase (Mel'čuk 1997a). Mais la construction ergative est largement répandue et n'a rien d'exotique.

Il est inutile de multiplier des exemples : il est clair qu'en linguistique, une terminologie précise, s'appuyant sur un système de concepts rigoureux, est absolument vitale, faute de quoi de nombreuses discussions dégénèrent en dialogues de sourds et la compréhension mutuelle entre linguistes devient impossible. Je crois qu'un développement harmonieux de la théorie linguistique, ainsi que des progrès sensibles des modèles fonctionnels des langues sont impossibles sans un métalangage linguistique unifié. La théorie Sens-Texte met donc la tâche d'élaboration d'un système conceptuel et d'une terminologie correspondante au premier plan de la recherche en linguistique.

5. L'apport possible de la linguistique Sens-Texte pour la société humaine

Dans sa leçon inaugurale au Collège de France, C.Hagège (1988 : 29-31) a parlé de trois domaines où la linguistique s'avère utile à l'humanité : l'histoire des civilisations, la planification linguistique (usage normatif, orthographe, néologismes, terminologies, ...) et l'étude de l'Homme. Le premier sort tout à fait du champ Sens-Texte ; mais pour le deuxième et le troisième, la linguistique Sens-Texte a bien des choses en réserve. En effet, un modèle fonctionnel d'une langue **L** présente tout d'abord une UTILITÉ PRATIQUE au moins dans :

1) le traitement automatique de la langue naturelle, TALN (y compris la traduction automatique, la génération de textes, les systèmes d'interrogation de bases de données en langue naturelle, etc.) ;

2) l'apprentissage des langues (étrangères et maternelles) ;

et 3) la création d'ouvrages de référence — aussi bien dictionnaires de collocations, de régime, de synonymes, que grammaires de consultation, manuels, etc.

Grâce à leur caractère formalisé et surtout à leur point de vue spécifique — «Comment exprimer telle pensée dans telle langue ?» — les MST offrent des moyens commodes pour atteindre ces objectifs.

En même temps, les MST des langues me paraissent présenter un INTÉRÊT THÉORIQUE encore plus important. En effet, jusqu'aujourd'hui, l'évolution scientifique a surtout porté sur la maîtrise de l'énergie et de la matière. Nous avons fondamentalement augmenté nos forces physiques (de la machine à vapeur à la bombe H) et considérablement «renforcé» nos organes de locomotion (la navette spatiale, sans parler des T.G.V., etc.), ainsi que nos organes de perception (le microscope électronique, le spectromètre, le radiotélescope, etc.). Nous avons pénétré dans l'atome et dans les profondeurs de l'espace ; nous avons appris des choses importantes sur l'apparition de notre univers et sur la structure de nos gènes. Mais nous n'avons pas fait de progrès comparables dans le domaine du traitement de l'information par le cerveau humain. Nous connaissons trop peu de choses sur le fonctionnement informatique de notre raison, et pourtant le «renforcement» de cet organe, c'est-à-dire la création d'outils puissants capables de suppléer à certaines fonctions essentielles de la raison, est, à mon avis, la tâche la plus pressante de la science moderne. Face au problème crucial du 21^e siècle — le manque de ressources naturelles sur la Terre pour une population qui s'accroît — nous avons un besoin aigu d'un SUPERCERVEAU, c'est-à-dire de machines capables de penser à une échelle que l'Homme seul ne saurait atteindre. Il nous faut des modèles, et de bons modèles, de la pensée humaine. La communauté scientifique l'a bien compris et, selon toute probabilité, la plupart

des savants sont d'accord avec ce programme. Les nombreuses recherches en Intelligence Artificielle à travers le monde dont le but est de modéliser tous les aspects du comportement intelligent, en sont une bonne preuve. Cependant, on oublie trop souvent que, pour avoir accès à la pensée humaine, la langue est la seule clé fiable. J'insiste :

- 1) Le meilleur moyen d'accès à la compréhension du fonctionnement de la pensée est la compréhension du fonctionnement de la langue.
- 2) Le meilleur moyen d'accès à la compréhension du fonctionnement de la langue est un modèle Sens-Texte, c'est-à-dire un modèle qui permette de relier la pensée inobservable du locuteur à la seule chose qui est aujourd'hui observable pour nous — le texte que celui-ci produit.

On ne pourra jamais comprendre comment notre cerveau traite les informations sans savoir comment la langue y fonctionne. De nos jours, les modèles fonctionnels de la langue acquièrent donc une importance toute particulière : ils peuvent contribuer de façon essentielle à la compréhension que nous avons de nous-mêmes et, plus précisément, de notre pensée. J'ose penser que le modèle Sens-Texte aura là un rôle à jouer.

Remerciements

Le texte de cet exposé a été lu, critiqué et commenté par M.Alonso Ramos, L.Iordanskaja, S.Kahane, J.Lévy, J.Milićević et A.Polguère. Je les remercie de tout mon cœur pour leur aide, qui m'a permis d'améliorer grandement la formulation de mes idées.

Notes

¹ Des neurologues peuvent en principe tenter de telles observations directes, et certains progrès ont effectivement été déjà faits dans cette voie ; néanmoins, dans cet exposé, je ne veux considérer que les approches strictement linguistiques.

² Je laisse de côté les autres sens du nom MODÈLE, qu'on voit, par exemple, dans les énoncés suivants : *Harpagon est le modèle de l'avare ; des garçons modèles ; des modèles nus ; Renaud modèle 1991 ; etc.*

³ Des relations converses de ce type sont assez fréquentes dans le lexique des langues naturelles. Ainsi, on dit en français *louer* dans le sens 'prendre en location' et 'donner en location' ; de façon similaire, le terme *hôte* s'applique à celui qui reçoit quelqu'un et à celui qu'on reçoit.

⁴ Chao 1962 offre une liste d'à peu près 40 interprétations données en linguistique au terme *modèle*. Il est donc tout à fait justifié de chercher à éviter les confusions possibles !

⁵ L'autre structure faisant partie de la RMorphP d'une phrase est la structure morphologique prosodique, qui reflète la division de la phrase en syntagmes phonétiques, c'est-à-dire en vrais constituants.

⁶ Pour ne pas encombrer l'exposé, j'indiquerai, dans cette note, trois cas typiques de décalages importants entre la SMorphP et la SMorphS d'un mot-forme :

- Expression cumulative/agglutinative de grammèmes :

russe DOM 'maison'_{pl, gén} \iff {DOM-}, {PL.GÉN} (dom+ov)

vs
turc EV 'maison'_{pl, gén} \iff {EV-}, {PL}, {GÉN} (ev+ler+in)

- Présence des morphèmes vides, exigés par le type flexionnel du lexème (en l'occurrence, l'élément thématique) :

espagnol TOCAR 'toucher'_{ind, imparf, 3, pl} \iff {TOC-}, {ÉL.THÉM}, {IND.IMPF}, {3PL} (toc+a+ba+n)

- Expression double de certains grammèmes :

hébreu moderne LEJLA 'nuit'_{pl, 1sg} \iff {LEJL}, {PL}, {PL}, {1SG} (lejl+ot+a+j 'mes nuits')

L'existence de nombreux signes morphologiques non segmentaux (apophonies, reduplications, conversions) complique encore la correspondance entre SMorphP et SMorphS.

⁷ Les constructions syntaxiques ayant un sens propre ne sont pas très répandues en français, mais quand même elles y existent. Ainsi, la construction PRONOM_{interr} \leftarrow V_{inf} (*Qui épouser ?*, *À qui m'adresser ?* ou *Où aller ?*) exprime le sens 'pourrais-je ou devrais-je'. À propos des constructions syntaxiques porteuses de sens, voir Mel'čuk 1987.

⁸ Je ne peux faire ici le point sur la théorie sémantique sous-jacente à l'approche Sens-Texte ; je me limite donc à la mention des noms — des principaux chercheurs qui ont lancé la sémantique nouvelle au début des années 60.

⁹ Les deux définitions ont été simplifiées pour fins d'illustration ; pour une discussion détaillée du comportement de la négation avec les sens du type présenté en (12), ainsi que les définitions du verbe DOUTER, de l'adjectif SÛR et de quelques lexies similaires, voir Iordanskaja 1986.

¹⁰ La négation est ici phraséologisée : une négation «normale» n'affecte pas les présupposés, alors que dans *ne pas douter* (<*ne pas avoir de doute*>) le présupposé de 'douter (<avoir des doutes>)' — 'n'ayant pas la croyance «P a lieu» — se trouve dans le champ de la négation. Cette propriété de la négation (indiquée dans Iordanskaja 1986: 355-356) doit être explicitement spécifiée dans l'article de dictionnaire de DOUTER/DOUTE.

¹¹ Cette réserve est nécessaire pour deux raisons :

- D'une part, beaucoup de collocations nécessitent, pour leur description, des FL *non standard*, qui sont en nombre élevé et s'appliquent chacune à très peu de lexies : par exemple, 'sans ajout de produit laitier'(*café*) = *noir* ou 'sans ajout de produit laitier ou de fruit'(*thé*) = *nature*. Les FL non standard ne sont pas universelles non plus : leur jeu est spécifique à chaque langue et doit être établi de façon empirique.

• D'autre part, certaines collocations sont décrites par le schéma de régime plutôt que par des FL : par exemple, les collocations *assurance MALADIE* (cf. angl. *HEALTH insurance* 'assurance santé') ou *condamner À MORT/À PERPÉTUITÉ*, etc., où les expressions en petites capitales correspondent aux actants sémantiques de la lexie vedette L et, en tant que tels, doivent être spécifiées comme régime de L.

¹² J'ai découvert les fonctions lexicales au début des années 60, lors de ma participation à une expédition géologique dans la zone montagneuse semi-désertique du sud du Kazakhstan, — avec l'intention de les utiliser dans la Traduction Automatique. Quand j'ai eu l'idée de ce qui est devenu plus tard (avec la collaboration précieuse de A.Zholkovsky : Žolkovskij & Mel'čuk 1967) les FL **Magn** et **Oper**₁, j'étais en train de chercher une méthode simple permettant d'éviter les milliers de tests fastidieux alors nécessaires pour que l'ordinateur trouve les équivalents russes — déterminés par le contexte lexical — des lexèmes anglais comme HEAVY, IMPORTANT, EXTENSIVE, HIGH, etc. d'une part, et [to] DO, [to] MAKE, [to] GIVE, [to] GET, etc., de l'autre. Il suffit de prendre quelques exemples au hasard pour mesurer l'étendue de ce problème, d'ailleurs bien connu des traducteurs : HEAVY [*rain*], litt. 'pluie **lourde**', correspond en russe à SIL'NYJ [*dožd'*] 'pluie **forte**', alors que HEAVY [*losses*] 'pertes **lourdes**' est rendu par TJAŽĚLYE [*poteri*], également 'pertes **lourdes**', s'il s'agit des pertes en personnel à la guerre, et par SER'EZNYE/ZNAČITEL'NYE [*poteri*] 'pertes **sérieuses/considérables**' au cas de pertes financières ; alors que HEAVY [*prison terms*] '**lourdes** peines de prison' se traduit par DLITEL'NYE [*sroki zaključenija*], litt. 'peines **de longue durée**', et ainsi de suite; GIVE [*a talk*] '**donner** une conférence' se traduit par ČITAT' [*doklad*] '**lire** une conférence', mais GIVE [*a look*], litt. '**donner** un regard', par BROSIT' [*vzgljad*] '**jeter** un regard'. Les FL permettent d'établir les correspondances interlinguistiques entre cooccurrents contextuels de façon directe et logique. Cette visée des FL est explicitement formulée dans l'article Žolkovskij & Mel'čuk 1967, mentionné ci-dessus ; voir son application dans Danlos 1994.

Références

- Apresjan, Ju. 1992. *Lexical Semantics. User's Guide to Contemporary Russian Vocabulary*. Ann Arbor (MI), Karoma. 633 pp.
- Apresjan, Ju. 1995. *Izbrannye trudy. Tom II. Integral'noe opisanie jazyka i sistemnaja leksikografija* [Œuvres choisies. Vol. II. Description intégrale de la langue et la lexicographie systématique]. Moskva, Jazyki russkoj kul'tury. 767 pp.
- Auger, P. 1965. Les modèles dans la science. *Diogène*, n° 52, 3-15.
- Chao, Yuen Ren. 1962. Models in Linguistics and Models in General. In : Nagel *et al.* (eds) 1962: 558-566.
- Danlos, L. 1994. Coder des informations monolingues sur les noms pour éviter des règles bilingues sensibles au contexte. *Langage*, n° 116, 95-110.
- Hagège, C. 1986. *Homme des paroles*. Paris, Fayard. 406 pp.
- Hagège, C. 1988. *Leçon inaugurale*. Paris, Collège de France. 32 pp.
- Halmos, P. 1957. Nicolas Bourbaki. *Scientific American*, n° 5, 88-99.

- Iordanskaja, L. 1963. O nekotoryx svojstvax pravil'noj sintaksičeskoj struktury (na materiale rus-skogo jazyka [Quelques propriétés de la structure syntaxique bien-formée (à base du russe)]. *Voprosy jazykoznanija*, n° 4, 102-112.
- Iordanskaja, L. 1982. Le placement linéaire des clitiques pronominaux non-sujets en français contemporain. *Linguisticae Investigationes*, 6: 1, 145-188.
- Iordanskaja, L. 1986. Propriétés sémantiques des verbes promoteurs de la négation en français. *Linguisticae Investigationes*, 10: 2, 345-380.
- Iordanskaja, L., I.Mel'čuk. 1997. Le corps humain en russe et en français : Vers un Dictionnaire explicatif et combinatoire bilingue. In : *Hommage à Yves Gentilhomme* [= *Cahiers de lexicologie*, 70: 1].
- Iordanskaja, L., A.Polguère. 1988. Semantic Processing for Text Generation. In : *Proceedings of the First International Computer Science Conference '88, 19-21 December 1988, Hong Kong*, 310-318.
- Keenan, E. 1976. Towards a Universal Definition of Subject. In : Ch.Li (ed.), *Subject and Topic*, New York, etc. : Academic Press, 303-333.
- Mel'čuk, I. 1965. O fonologičeskoj traktovke "poluglasnyx" v ispanskom jazyke [Traitement phonologique des «semivoyelles» espagnoles]. *Voprosy jazykoznanija*, n° 4, 92-109. [Traduction anglaise: *Linguistics*, 1973, n° 109, 35-60.]
- Mel'čuk, I. 1968. Model' sklonenija vengerskix suščestvitel'nyx [Un modèle de déclinaison des noms en hongrois]. In: S. Šaumjan, réd., *Problemy strukturnoj lingvistiki – 1967*, Moskva, Nauka, 1968, 344-373.
- Mel'čuk, I. 1967a. Ordre des mots en synthèse automatique des textes russes. *T.A.Informations*, 8: 2, 65-84.
- Mel'čuk, I. 1967b. Model' sprjaznenija v ispanskom jazyke [Un modèle de conjugaison en espagnol]. *Masinyj perevod i prikladnaja lingvistika*, Institut des Langues Etrangères de Moscou, v. 10, 21-53. [Traduction anglaise: V.Ju. Rozencvejk (ed.), *Essays on Lexical Semantics*, I, 1974, Stockholm: SCRIPTOR, 43-94 ; une version récente améliorée : *Voz y Letra*, 1993, 4: 1, 9-85.]
- Mel'čuk, I. 1968. Model' sklonenija vengerskix suščestvitel'nyx [Un modèle de déclinaison des noms en hongrois]. In: S. Šaumjan (réd.), *Problemy strukturnoj lingvistiki – 1967*, Moskva, Nauka, 1968, 344-373.
- Mel'čuk, I. 1973. Towards a Linguistic "Meaning \iff Text" Model. In : F.Kiefer (ed.), *Trends in Soviet Theoretical Linguistics*, Dordrecht, Reidel, 33-57.
- Mel'čuk, I. 1974. *Opyt teorii lingvističeskix modelej "Smysl \iff Tekst"*. *Semantika, Sintaksis* [Esquisse d'une théorie des modèles linguistiques du type "Sens \iff Texte". La sémantique, la syntaxe]. Moskva, Nauka. 314 pp.
- Mel'čuk, I. 1981. Meaning-Text Models: A Recent Trend in Soviet Linguistics. *Annual Review of Anthropology*, v. 10, 27-62.
- Mel'čuk, I. 1986a. Toward a Definition of Case. In : R.Brecht and J.Levine (eds), *Case in Slavic*, Columbus (OH), Slavica, 35-85.

- Mel'čuk, I. 1986b. The Dualizer and Pluralizers in the Alutor Verb. In : B.Elson (ed.), *Language in Global Perspective. Papers in Honor of the 50th Anniversary of the Summer Institute of Linguistics (1935-1985)*, 1986, Dallas (TX), SIL, 421-438.
- Mel'čuk, I. 1987. Un affixe dérivationnel et un phrasème syntaxique du russe moderne. Essai de description formelle. *Revue des études slaves*, 59: 3, 631—648. [Réimprimé dans Mel'čuk, I., *The Russian Language in the Meaning-Text Perspective*, Moskva/Wien, Wiener Slawistischer Almanach/Jazyki russkoj kul'tury.]
- Mel'čuk, I. 1988. *Dependency Syntax : Theory and Practice*. Albany (NY), The SUNY Press. 428 pp.
- Mel'čuk, I. 1989. Semantic Primitives from the Viewpoint of the Meaning-Text Linguistic Theory. *Cuaderni di semantica*, 10: 1, 65-102.
- Mel'čuk, I. 1990. Slovoobrazovanie v lingvističeskix modeljax tipa "Smysl \iff Tekst" (predvaritel'nye zamečanija) [La dérivation dans les modèles linguistiques du type «Sens-Texte» (remarques préliminaires)]. In : Z.Saloni (red.), *Metody formalne w opisie języków słowiańskich*, Białystok, Wydawnictwo Filii UW w Białymstoku, 47-74.
- Mel'čuk, I. 1991. Toward a Universal Calculus of Inflectional Categories: On Roman Jakobson's Trail. In : L.Waugh and S.Rudy (eds), *New Vistas in Grammar: Invariance and Variation*, Amsterdam/Philadelphia, Benjamins, 85-109.
- Mel'čuk, I. 1992a. Paraphrase et lexique : la théorie Sens-Texte et le Dictionnaire explicatif et combinatoire. In : Mel'čuk *et al.* 1992: 9-58.
- Mel'čuk, I. 1992b. Towards a Logical Analysis of the Notion 'Ergative Construction'. *Studies in Language*, 16: 1, 91-138.
- Mel'čuk, I. 1993a. The Inflectional Category of Voice: Towards a More Rigorous Definition. In : B.Comrie and M.Polinsky (eds), *Causativity and Transitivity*, Amsterdam/Philadelphia, Benjamins, 1-46.
- Mel'čuk, I. 1993b. Agreement, Government, Congruence. *Linguisticae Investigationes*, 17: 2, 307-373.
- Mel'čuk, I. 1993c. La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère. *Études de Linguistique Appliquée*, vol. 92, 82-113.
- Mel'čuk, I. 1993-97. *Cours de morphologie générale*. Vol. 1-4. Montréal/Paris, Presses de l'Université de Montréal/Éditions du C.N.R.S. Vol. 1 (1993) : 412 pp. ; vol. 2 (1994) : 456 pp. ; vol. 3 (1996) : 326 pp. ; vol. 4 (1997).
- Mel'čuk, I. 1995. Porjadok slov v ruskoj fraze [L'ordre des mots dans la phrase russe]. In : Mel'čuk, I., *The Russian Language in the Meaning-Text Perspective*, Moskva/Wien, Wiener Slawistischer Almanach/Jazyki russkoj kul'tury, 273-321.
- Mel'čuk, I. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In : L.Wanner (ed.), *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam/Philadelphia, Benjamins, 1996, 37-102.
- Mel'čuk, I. 1997a. Grammatical Cases, Basic Verbal Construction, and Voice in Maasai: Towards a Better Analysis of the Concepts. In : W.Dressler *et al.* (eds), *Towards Progress in Morphology*, Amsterdam/ Philadelphia, Benjamins.

- Mel'čuk, I. 1997b. *Communicative Organization of Sentences in Natural Language* [manuscrit].
- Mel'čuk, I. (avec N.Arbatchewsky-Jumarie, L.Elnitsky, I.Iordanskaja, et A.Lessard ; réd. A.Clas). 1984. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques I*. Montréal, PUM. 172 pp.
- Mel'čuk, I. (avec N.Arbatchewsky-Jumarie, L.Dagenais, L.Elnitsky, L.Iordanskaja, M.-M.Lefebvre, et S.Mantha ; réd. A. Clas). 1988. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques II*. Montréal, PUM. 332 pp.
- Mel'čuk, I. (avec N.Arbatchewsky-Jumarie, L.Iordanskaja, et S.Mantha ; réd. A.Clas). 1992. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*. Montréal, PUM. 323 pp.
- Mel'čuk, I., D.Aroga Bessong. 1983. Un modèle formel de la conjugaison bafia (à l'indicatif). *Bulletin of the School of Oriental and African Studies*, 46: 3, 477-528.
- Mel'čuk, I., A.Clas, A.Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve, Duculot. 256 pp.
- Mel'čuk, I., N.Pertsov. 1987. *Surface Syntax of English. A Formal Model Within the Meaning-Text Framework*. Amsterdam/Philadelphia, Benjamins. 526 pp.
- Mel'čuk, I., B.Podolsky. 1996. Stress in Modern Hebrew Nominal Inflection. *Theoretical Linguistics*, 22 : 1-2, 155-194.
- Mel'čuk, I., E.Savvina. 1978. Toward a Formal Model of Alutor Surface Syntax: Predicative and Completive Constructions, *Linguistics*, Special Issue, 5-39. [Réimprimé dans Mel'čuk 1988.]
- Mel'čuk, I., A.Zholkovsky. 1984. *Explanatory Combinatorial Dictionary of Modern Russian*. Wiener Slawistischer Almanach, Wien. 992 pp.
- Molino, J. 1985. Où en est la morphologie? *Langages*, 78, 5-40.
- MPiPL. 1964. *Mašinnyj perevod i prikladnaja lingvistika* [Traduction automatique et linguistique appliquée]. Vol. 8. Moskva, Institut inostrannyx jazykov im. M. Toreza.
- Nagel, E., P.Suppès, A.Tarski (eds). 1962. *Logic, Methodology and Philosophy of Science*. Stanford, CA: Stanford University Press.
- Polguère, A. 1990. *Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte*. Montréal, Univ. de Montréal, Dép. de linguistique et traduction [thèse de doctorat]. 313 pp.
- Polguère, A. 1992. Remarques sur les réseaux sémantiques Sens \iff Texte. In : A.Clas (réd.), *Le mot, les mots, les bons mots*, Montréal, PUM, 109-148.
- Queneau, R. 1978. *Bords: mathématiciens, précurseurs, encyclopédistes*. Paris, Hermann.
- Tesnière, L. 1959. *Éléments de syntaxe structurale*. Paris, Klincksieck.
- Tucker, A. , J.T. ole Mpaayei. 1955. *A Maasai Grammar with Vocabulary*. London, Longmans, Green and C°.
- Wierzbicka, A. 1996. *Semantics. Primes and Universals*. Oxford—New York, Oxford University Press. 500 pp.
- Žolkovskij, A. 1964. Predislovie [Préface]. *MPiPL* 1964, 3-16. [Traduction anglaise: V.Ju. Rozencvejk (ed.), *Essays on Lexical Semantics* , I, 1974, Stockholm: SCRIPTOR, 00-00.]

Žolkovskij, A., N.Leont'eva, Ju.Martem'janov. 1961. O principial'nom ispol'zovanii smysla pri mašinnom perevode [Sur l'utilisation essentielle du sens en traduction automatique]. In : *Mašinnyj perevod*, 2, Moskva, Institut točnoj mexaniki i vyčislitel'noj texniki AN SSSR. [Traduction anglaise: V.Ju. Rozencvejg (ed.), *Essays on Lexical Semantics* , I, 1974, Stockholm: SCRIPTOR, 00-00.]

Žolkovskij, A., I.Mel'čuk. 1967. O semantičeskome sinteze [Sur la synthèse sémantique]. *Problemy kibernetiki*, v. 19, 177-238. [Traduction française: *T.A.Informations*, 1970, n° 2, 1-85.]