

PARAPHRASE ET LEXIQUE : LA THÉORIE SENS-TEXTE ET LE DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

Igor Mel'čuk

I. INTRODUCTION

En 1966, Alexandre Zholkovsky et l'auteur de ces lignes terminaient l'article «Sur la synthèse sémantique» (Žolkovskij* et Mel'čuk 1967), qui aurait plutôt dû s'intituler «Paraphrase et lexique», car tel en était le contenu. Mais notre optique étant, à l'époque, celle de la traduction automatique, nos propositions visaient la SYNTHÈSE AUTOMATIQUE DE PHRASES PAR ORDINATEUR — à partir d'une représentation sémantique d'une famille de phrases (plus ou moins) synonymes. De fait, l'article ébauchait une théorie linguistique globale, appelée la *théorie Sens-Texte* (TST). Cette théorie presuppose un modèle particulier de la langue naturelle : le *modèle Sens-Texte* (MST), qui, à son tour, comprend — en tant que composante fondamentale — un lexique d'un nouveau type, le *Dictionnaire explicatif et combinatoire* (DEC).

Plus de vingt ans se sont écoulés, des années remplies d'événements et d'expériences de toutes sortes, entre autres, de travail intensif sur la théorie Sens-Texte (TST), surtout sur le MST et le DEC (voir les références ci-dessous). Pourtant, au moment de terminer cet article, nous sommes étonnés de constater que les thèses et les principes énoncés en 1967 sont encore valables. Bien entendu, nous y avons apporté beaucoup de modifications, formulé certaines choses de façon différente, corrigé un certain nombre d'erreurs, etc. Nous avons aussi pris en considération les ouvrages traitant de paraphrase en linguistique publiés depuis (par exemple, Martin 1976; Fuchs 1980, 1982; ML 1983; Bès et Fuchs 1988). Mais pour l'essentiel, l'article Žolkovskij et Mel'čuk 1967 reste intact, si bien que parfois nous en citons des règles et des explications de façon *verbatim*. Il est vrai que, malgré le développement intensif de la linguistique générale entre 1967 et 1990, la théorie de la paraphrase est restée plus ou moins à l'écart, à l'exception des recherches mentionnées ci-dessus et des travaux effectués dans le cadre de la théorie Sens-Texte. (Pour mieux évaluer l'état de la question, le lecteur consultera Leeman 1973, la bibliographie donnée dans ML 1983 : 191-192, ainsi que les références dans Bès et Fuchs 1988.) En somme, bien que beaucoup d'outils nécessaires, absents en 1967, soient aujourd'hui disponibles, le programme énoncé dans Žolkovskij et Mel'čuk 1967, à savoir,

|| construire un système de règles formelles capables, en s'aidant d'un lexique approprié, de produire, pour une

|| phrase donnée, l'ensemble le plus vaste possible de par-phrases (quasi) synonymes,

n'est pas encore accompli.

Aussi sommes-nous amené, après plus de vingt ans de recherches, à présenter un panorama raisonnablement complet et détaillé de notre connaissance sur la paraphrase langagière dans le cadre de la théorie Sens-Texte. Un tel panorama est d'autant plus nécessaire ici que le *Dictionnaire explicatif et combinatoire* (DEC) est lié au paraphrasage de façon tout à fait cruciale. Le but central du DEC est d'assurer l'ensemble des paraphrases le plus riche possible pour chaque «pensée» donnée; le choix des informations à y consigner se fait selon les besoins de production de paraphrases. Par conséquent, le procédé logique le plus utilisé pour la vérification et la justification des articles de ce dictionnaire est la paraphrase — la raison d'être du DEC. Il est donc très important de bien comprendre le système de paraphrasage proposé pour saisir les fondements mêmes du DEC, ainsi que les considérations qui déterminent nos solutions et les tâches posées.

Or, pour des raisons purement techniques (surtout le manque de temps), nous avons déjà publié deux volumes d'un DEC du français contemporain sans avoir présenté une description de base du système de paraphrasage sous-jacent au DEC comme tel. Qui plus est, nous n'avons même pas donné, dans ces publications, une caractérisation suffisamment détaillée du modèle Sens-Texte et de la théorie Sens-Texte, alors qu'il est évident qu'un dictionnaire comme le DEC ne peut être évalué et utilisé sous tous ses aspects qu'avec une bonne compréhension de la théorie linguistique et du modèle formel dont il fait partie intégrante.

Pour corriger notre omission, nous soumettons au lecteur intéressé ces lignes sur un système de paraphrasage et les relations que ce dernier a avec le DEC en général.¹

Le texte ci-dessous se divise en cinq sections :

- II – Paraphrase langagière;
- III – Modèle Sens-Texte;
- IV – Représentation sémantique dans le modèle Sens-Texte;
- V – Système de paraphrasage dans le modèle Sens-Texte;
- VI – Remarques finales.

* Žolkovskij est une translittération directe du nom russe du chercheur, alors que Zholkovsky est l'écriture officielle américaine de ce nom.

II. PARAPHRASE LANGAGIÈRE

1. Thèse centrale

Nous aimerais commencer d'emblée par la thèse centrale de notre démarche :

Une des tâches primordiales de la linguistique théorique contemporaine est l'élaboration d'une théorie de la paraphrase langagière.

Cela admis, décrire une langue L , c'est notamment construire un système de règles formelles doué de la même capacité d'effectuer le paraphrasage en L qu'ont les locuteurs de L , ni plus, ni moins. Chercher à faire plus, par exemple, à décrire la compréhension des énoncés de L par le destinataire, serait déborder le cadre de la linguistique comme telle et faire de la psychologie. Se borner à moins, par exemple, ne décrire que des transformations syntaxiques de L , serait peu fertile : la langue est «un système où tout se tient» (F. de Saussure), et ce «tout» tient surtout au sens, cet invariant fondamental des paraphrases; si nous jetons le sens (et, du même coup, la paraphrase) par-dessus bord, nous nous privons de l'essence même de la langue. Sans paraphrase, la linguistique serait appauvrie au point que l'on pourrait se demander s'il vaut la peine d'être linguiste!

Le rôle important attribué à la paraphrase en linguistique tient profondément à la théorie Sens-Texte. Cette théorie vise la description d'une langue sous la forme d'un système de règles qui en représente le fonctionnement : c'est ce qu'on appelle le MODÈLE SENS-TEXTE, et la paraphrase en est, dans un certain sens, le pivot. Le modèle Sens-Texte (dorénavant, MST) procède d'une représentation du sens — d'une REPRÉSENTATION SÉMANTIQUE (RSém), qui n'est pas autre chose qu'une représentation formelle de l'invariant d'un ensemble de paraphrases (plus ou moins) synonymes. À partir d'une RSém, le MST produit l'ensemble des phrases porteuses du sens correspondant à cette RSém; le mécanisme clé qui assure cette production est appelé SYSTÈME DE PARAPHRASAGE, et il fait partie de la composante sémantique du modèle. Enfin, toutes les informations lexicales, c'est-à-dire celles concernant les lexies individuelles, nécessaires au fonctionnement de ce système, sont emmagasinées dans un dictionnaire spécial : le DICTIONNAIRE EXPLICATIF ET COMBINATOIRE (DEC); c'est ici que la boucle «paraphrase – lexique» se ferme.

Ainsi, pour mettre en lumière le lien entre paraphrase et lexique, il nous faut considérer au moins les quatre points suivants :

- Modèle Sens-Texte;
- Représentation sémantique;
- Système de paraphrasage;
- Dictionnaire explicatif et combinatoire.

Les trois premiers points seront traités chacun dans une section à part (les sections III-V); quant au DEC, il a déjà été décrit et présenté dans les deux volumes publiés, et nous pouvons ici tenir pour acquises les grandes lignes de son organisation.

2. Un peu d'histoire; références

La théorie Sens-Texte (TST) n'est pas une nouveauté linguistique. Comme nous l'avons déjà dit, ses fondements ont été jetés dans les années 1960, à Moscou, par A. Zholkovsky (présentement, à University of Southern California, Los Angeles) et l'auteur de ces lignes (voir Žolkovskij et Mel'čuk 1965, 1967). Peu après, Ju. Apresjan s'est joint à nous, et le noyau de ce qu'on allait appeler le CERCLE SÉMANTIQUE DE MOSCOU a été ainsi formé. Depuis, des dizaines de personnes ont apporté leur contribution au développement de la TST et du MST, tant à partir du russe qu'à partir d'un certain nombre d'autres langues (anglais, français, allemand, polonais, tatare, somalien). Des recherches ont été effectuées dans tous les domaines de la linguistique, sauf en phonologie. Un aperçu général du MST a été publié en 1974 (Mel'čuk 1974), et deux ouvrages sur le lexique et la composante sémantique, en 1974 et 1980 (Apresjan 1974, 1980). Les résultats de ces recherches sont résumés dans un DEC du russe contemporain (Mel'čuk et Zholkovsky 1984), dans une description détaillée de la syntaxe de surface de l'anglais (Mel'čuk et Pertsov 1987) et dans un livre sur la syntaxe de dépendances (Mel'čuk 1988a). En plus, quelque 150 articles ont paru dans des périodiques linguistiques; voir les bibliographies dans Mel'čuk 1981 et dans Hartenstein et Schmidt 1983. Enfin, le travail portant sur un DEC du français contemporain a été commencé en 1979 par une équipe rattachée au Département de linguistique de l'Université de Montréal; le premier volume a paru en 1984, le deuxième en 1988 (Mel'čuk *et al.* 1984, 1988) et le troisième, vous l'avez en main.

Comme la documentation sur la TST est abondante et accessible, nous nous limiterons ici à un exposé plutôt dogmatique. Nous nous concentrerons sur les points cruciaux et, à cause des contraintes d'espace, laisserons de côté plusieurs détails, justifications et explications, dont quelques-uns qui seraient pourtant pertinents. Pour ces derniers, nous renverrons le lecteur aux titres indiqués.

3. La notion de paraphrase

Avant de procéder à la description du modèle Sens-Texte, du système de paraphrasage, etc., nous apporterons quelques précisions à la notion de paraphrase. Nous ne toucherons pas aux acceptations du mot *paraphrase* liées à l'idée de 'développement explicatif/verbex et diffus [d'un texte]' (*Ce n'est pas*

une traduction, c'est une paraphrase!), ‘transformé d'un texte comme visant un effet artistique’ (*C'est une paraphrase du fameux dicton de Voltaire*), etc. Dans ce qui suit, le terme *paraphrase* sera utilisé exclusivement dans le sens suivant :

|| Nous dirons que la phrase P₁ et la phrase P₂ sont *en relation de paraphrase* si et seulement si elles sont plus ou moins synonymes.

Autrement dit, la relation de paraphrase est une relation de SYNONYMIE (et de QUASI-SYNONYMIE) ENTRE PHRASES. Par extension, nous appliquerons le terme *paraphrase* aussi aux syntagmes.²

Nous dirons également, pour deux phrases (ou syntagmes) synonymes P₁ et P₂, que l'une est *une paraphrase de* l'autre ou que l'une *paraphrase* l'autre (et *vice versa*, puisque la relation de paraphrase est symétrique). L'action de *paraphraser* sera appelée aussi *paraphrase* (ou *paraphrasage*). Bref, voici nos emplois du terme *paraphrase* :

X et Y sont en relation de paraphrase = Les phrases X et Y sont (plus ou moins) synonymes.

{*X est une paraphrase de Y*} = La phrase X est synonyme de la phrase Y.

On paraphrase X en/par Y = On construit la phrase Y, une paraphrase de X.

La paraphrase de X = L'action de paraphraser X [= ‘paraphrasage de X’].

À ce point, cinq remarques importantes s'imposent.

3.1. Synonymie

Dans notre approche, LA SYNONYMIE DES PHRASES EST PRISE COMME UNE NOTION INTUITIVE DE DÉPART. Nous posons que la capacité de juger si deux phrases données sont équivalgantes (= ont le même sens) fait partie de la compétence linguistique des locuteurs, et nous batissons notre système formel sur cette capacité. Comme on le verra (pp. 12 et 14), nous définissons la notion de ‘sens’ à partir de la notion de ‘même sens’, cette dernière étant considérée comme acquise. Par conséquent, nous n'essayons pas ici de définir la paraphrase ou la synonymie; plutôt, nous postulons la notion d’identité du sens comme indéfinissable.³

3.2. Sens langagier

Par SYNONYMIE, nous entendons l'identité (ou la quasi-identité) du SENS LANGAGIER, et non l'identité des informations contenues dans les textes en cause. Si deux textes renvoient à la même situation et comportent la même information,

mais que, pour établir ce fait, on doit mettre en valeur les connaissances extralinguistiques ou les opérations logiques, mathématiques, etc., ces textes ne sont pas synonymes pour nous : leurs sens langagiers ne sont pas identiques (même si leurs contenus informationnels le sont). Dans Mel'čuk et Žolkovskij 1970 : 17-19, on a montré explicitement que les paires de phrases suivantes ne sont pas des paraphrases :

(1) a. *Le peuple allemand luttait contre le fascisme.*

vs

Des Alpes à la Mer du Nord, du Rhin à l'Oder, le peuple continuait la lutte d'Edgar André.

b. *Plus de la moitié du territoire de l'Égypte est occupée par des déserts.*

vs

En Égypte, les déserts occupent un territoire plus vaste que la France.

Pour établir l'équivalence informationnelle de ces phrases, la maîtrise du français ne suffit pas : il faut savoir quel pays est situé entre le Rhin et l'Oder, qui était Edgar André, quelle est la surface de l'Égypte et de la France, etc.; on ne considérera donc pas les paires de phrases de ce type comme des paraphrases.⁴

3.3. Exactitude nécessaire des paraphrases

La question de l'EXACTITUDE NÉCESSAIRE DES PARAPHRASES soulève beaucoup de problèmes et engendre beaucoup de malentendus. Deux paraphrases doivent-elles avoir strictement le même sens? Est-il possible, de façon générale, que deux expressions linguistiques différentes aient strictement le même sens? Et qu'est-ce que ‘strictement le même sens’? Nos réponses à toutes les questions de ce type s'appuient sur les considérations suivantes :

TROIS TYPES DE SENS

Le sens d'une phrase est composé d'éléments de trois types :

- SENS SITUATIONNEL, ou sémantique à proprement parler, qui représente l'état de choses dont il s'agit (les objets, les événements, etc., et toutes les relations que ceux-ci entretiennent entre eux; le sens situationnel inclut également les données sur l'état interne du locuteur : ses attitudes, ses opinions, etc.);

- SENS COMMUNICATIF, qui représente l'organisation du message par le locuteur (ce qu'il veut présenter comme connu ou comme nouveau, ce qu'il veut faire ressortir ou, au contraire, reléguer à l'arrière-plan, etc.);

- SENS RHÉTORIQUE, qui représente les effets expressifs ou artistiques visés par le locuteur (niveau soutenu, poétique,

argotique, officiel, familier, ...; exposé explicatif, ironique, humoristique, ...).

(Ces trois types de sens sont reflétés dans les trois structures qui constituent la représentation sémantique adoptée dans le MST : structure sémantique, structure communicative et structure rhétorique : voir pp. 27-28.)

Or, pour les paraphrases prises dans notre optique, seul le sens situationnel est pertinent. Les paraphrases peuvent différer du point de vue communicatif et rhétorique; en effet, très souvent cette différenciation est même le but de la paraphrase. Dans ce qui suit, nos exigences quant aux paraphrases se limitent à l'identité ou à la proximité suffisante du sens situationnel; les deux autres types de sens peuvent varier.

CARACTÈRE RELATIF DE TOUTE IDENTITÉ DE MESURE

Beaucoup de linguistes n'admettent pas la «vraie» synonymie : selon eux, les expressions linguistiques X et Y quelconques peuvent être considérées synonymes seulement parce que l'analyse sémantique utilisée pour X et Y n'est pas suffisamment raffinée; en principe, disent-ils, on peut toujours trouver des différences sémantiques entre n'importe quelles expressions linguistiques, pourvu qu'on s'y applique. Nous aimerais cependant y opposer la considération suivante. On ne peut trouver deux tiges métalliques qui aient, strictement parlant, la même longueur : en utilisant des dispositifs de mesure de plus en plus précis, on peut toujours trouver des différences de longueur entre deux tiges quelconques — jusqu'au point où l'erreur inévitable de mesure dépasse la résolution du dispositif utilisé, après quoi c'est l'indétermination qui règne. La même chose est vraie à propos du poids, de la vitesse, du temps écoulé, etc., bref, à propos de toute mesure. Pourtant, cela ne gêne personne (à part, peut-être, les philosophes) : pour chaque tâche pratique, on établit un degré raisonnable de précision. En effet, on ne prend pas une balance trop précise pour peser des légumes dans une épicerie; on se contente de poids relativement identiques à 1 kg à une approximation raisonnable près, et on se met d'accord pour ne pas tenir compte de cette approximation. (Pour peser les ingrédients d'un médicament, une balance beaucoup plus précise est nécessaire; et ainsi de suite pour chaque tâche.)

Le sens d'une expression linguistique ne se distingue pas, de façon substantielle, des paramètres physiques mesurables, sauf que l'instrument de «mesure» est ici l'intuition linguistique du locuteur. Le réglage de cet instrument, c'est-à-dire le degré de précision ou de résolution exigé, doit être approprié à la tâche; cela veut dire, entre autres, que le chercheur ne doit pas être trop précis dans sa quête des nuances sémantiques (on ne pèse pas des pommes de terre avec une balance de pharma-

cie!). Beaucoup de distinctions sémantiques qui existent dans la langue sont ignorées par les locuteurs dans le discours : le contexte réel «éteint» certaines oppositions virtuelles. À notre avis, en sémantique il est très important de ne pas aller trop loin sans nécessité, en cherchant des distinctions qui ne sont pas nécessairement observées dans la réalité langagière. Tout en admettant la relativité de l'identification de sens, nous fondons notre modèle linguistique sur la notion primitive de «même sens».

PARAPHRASES ET CHANGEMENT DE SENS

En règle générale, nous n'exigeons même pas que la paraphrase ne modifie pas le sens situationnel de l'expression de départ. Une paraphrase particulière peut modifier le sens, mais de façon complètement contrôlée (le chercheur sait ce qu'elle modifie et comment) et sans que la transformation ne soit trop grande (le degré tolérable est établi intuitivement, en fonction de la tâche). Les paraphrases approximatives ne contredisent pas notre conception de la paraphrase, puisque nous considérons aussi, à côté de la synonymie exacte, la quasi-synonymie (les synonymes plus larges, plus étroits et à intersection, voir la liste des fonctions lexicales, pp. 127 *ssq.*).

3.4. Neutralisation sémantique contextuelle

Pour notre approche de la paraphrase, le fait suivant est crucial : souvent, on peut construire des paraphrases (presque) exactes, c'est-à-dire sémantiquement équivalentes, à partir d'unités linguistiques — lexèmes, phrasèmes, constructions syntaxiques, etc. — qui ne sont pas du tout équivalentes (mais qui partagent, bien entendu, une partie sémantique importante). Cela est dû à la NEUTRALISATION SÉMANTIQUE CONTEXTUELLE. Dans un contexte particulier, qu'il soit linguistique (= le choix d'unités lexicales, de constructions syntaxiques spécifiques, etc.) ou extralinguistique (= une situation concrète dont il est question), deux lexies ou deux constructions qui, comme telles, s'opposent sémantiquement, perdent leur caractère oppositif et — sans devenir équivalentes! — apportent à l'énoncé en question (à peu près) le même «morceau» de sens. (Nous avons fait allusion à cet état de choses ci-dessus, en parlant du contexte, qui peut «éteindre» certaines distinctions.) Le phénomène de neutralisation sémantique a été étudié en profondeur par Ju. Apresjan (1974 : 156-163, 239-242, 281-283), et nous nous limiterons ici à trois exemples que nous lui avons empruntés.

— Les verbes *s'arrêter* et *cesser* (dans les acceptations correspondantes) ne sont pas synonymes : cf. *Ses visites se sont arrêtées durant le mois de mai* [plutôt temporairement] vs *Ses visites ont cessé durant le mois de mai* [plutôt de façon permanente]; cependant, dans le contexte du nom *pluie*, ils ne s'opposent plus : 'La pluie s'est arrêtée' = 'La pluie a cessé'.

– Les expressions *tout homme* [indéfini, ‘n’importe quel?’] et *chaque homme* [défini] ne sont pas synonymes, mais dans un contexte générique, elles ne s’opposent pas : *Tout homme / Chaque homme a le droit de vivre en paix.*

– ‘Empirer’ ≠ ‘se gâter’, et pourtant ‘Le temps a visiblement empiré’ = ‘Le temps s’est visiblement gâté’.

Une autre illustration convaincante est la série suivante de titres de journaux qu’on peut voir au moment des grands départs en vacances :

$\left\{ \begin{array}{l} \text{Chaos} \\ \text{Confusion} \\ \text{Pagaille} \\ \text{Désordre} \end{array} \right\}$	$\left\{ \begin{array}{l} \text{fou} \\ \text{démentiel} \\ \text{incroyable} \\ \text{terrible} \end{array} \right\}$	<i>sur les autoroutes du</i>	<i>Midi.</i>
--	--	------------------------------	--------------

Chaos n’est pas du tout synonyme de *confusion*, pas plus que *démentiel* ne l’est de *terrible*. Cependant, les seize titres qu’on peut obtenir en combinant de toutes les façons possibles les noms et les adjectifs donnés le sont (en tout cas, plus ou moins) : dans ce contexte particulier, les différences sémantiques indéniables sont «éteintes» (exemple emprunté à l’article de M. Dickson dans *Le français dans le monde*, 1990, n° 235, p. 66 ssq.).

C’est la neutralisation sémantique dans un contexte donné, un phénomène des plus typiques dans les langues naturelles et dans le comportement langagier des locuteurs, qui assure la richesse paraphrastique dont nous avons besoin comme fondement de la théorie Sens-Texte (et, en particulier, du *Dictionnaire explicatif et combinatoire*).

À ce propos, l’éclaircissement suivant nous paraît indispensable. Beaucoup de linguistes s’opposent de façon plus catégorique à l’idée même de synonymie des textes. Au niveau morphologique, ils reconnaissent l’identité sémantique de marqueurs différents, par exemple, du nombre, tels que *house – house + s* vs *ox – ox + en* vs *tooth – teeth*, ou du temps, tels que *answer – answered* vs *sing – sang*, en anglais. Mais ce niveau mis à part, ils pensent que toute différence de forme linguistique entraîne nécessairement une différence de sens, de sorte que l’on peut toujours trouver des distinctions sémantiques entre deux expressions, pourvu que l’analyse soit suffisamment raffinée et sophistiquée. Un des tenants les plus connus de ce point de vue est probablement D. Bolinger (1977), qui postule et défend la thèse que «in syntax there is no such thing as two different surface structures with the same deep structure (that is, with the same meaning)» (p. 4). Il propose des analyses très subtiles et convaincantes, en démontrant qu’on peut découvrir des nuances entre des constructions anglaises qui, en règle générale, sont jugées synonymes : les actifs vs les passifs, les phrases avec ou sans *that* relatif (*The*

forecast says (that) it is going to rain), *any* vs *some*, les constructions avec *it* et *there* anticipatifs, etc.

Cependant, même si nous acceptons la justesse des analyses et des remarques de Bolinger et des autres linguistes de la même école de pensée, nous croyons pouvoir maintenir notre position pour la raison suivante :

Les distinctions sémantiques dont il est question PEUVENT être trouvées — si l’on fait très attention à ce qu’on dit; mais de façon générale, les locuteurs ne font pas très attention à ce qu’ils disent : ils ne tiennent tout simplement pas compte de beaucoup de distinctions latentes. Dans la pratique quotidienne de la parole, on ne se soucie pas de différences potentielles dont on n’a que faire dans un contexte donné.⁵

En d’autres termes, le locuteur traite comme identiques — dans la plupart des cas, de façon tout à fait inconsciente, bien sûr — les expressions linguistiques qu’il aurait pu en principe opposer (et qu’il fera contraster, probablement, dans un contexte différent). Les locuteurs ne sont pas aussi rigoureux que les linguistes qui étudient les possibilités sémantiques virtuelles des expressions. Nous voudrions refléter, dans notre modèle linguistique, ce comportement sémantique un peu nonchalant qui exploite la proximité des expressions en ignorant leurs distinctions, trop subtiles pour la tâche donnée. En nous appuyant sur les neutralisations contextuelles, nous nous permettons de travailler avec des paraphrases approximatives en les traitant comme (presque) identiques — de la même façon que le font les locuteurs.

3.5. Paraphrases linguistiques réellement utilisées

Les paraphrases qui nous intéressent sont exclusivement des PARAPHRASES LINGUISTIQUES, c'est-à-dire, des paraphrases réellement utilisées par les locuteurs dans leur comportement langagier ordinaire : pour s’exprimer de façon plus appropriée, pour changer de style, pour rendre plus claire leur pensée, etc. De nombreux exemples de ce type de paraphrases sont donnés aux pp. 37-43, où nous illustrons les règles lexicales de paraphrase. Nous laissons de côté les paraphrases métalinguistiques, qui servent à expliquer le sens d’une lexie inconnue — par exemple, *acculer (qqn) au pied de mur* = *l’obliger à parler, à agir, à prendre part, à se prononcer*. (Bien entendu, ces dernières font, elles aussi, partie du comportement langagier, mais d’un comportement assez spécial — tourné vers le langage lui-même. Cf. la note 2.)

En résumé, nous pouvons dire que dans notre cadre, les paraphrases sont orientées vers la préservation du sens langagier, et plus précisément — situationnel, bien que même pour ce sens, des modifications soient permises dans des conditions

contrôlées; l'identité ou la proximité des sens n'est qu'une question de jugement intuitif des locuteurs; le contexte y joue un rôle primordial, en entraînant la neutralisation, à l'intérieur

des paraphrases, de certaines oppositions sémantiques; seules les paraphrases purement linguistiques sont considérées, à l'exclusion des paraphrases métalinguistiques.

III. MODÈLE SENS-TEXTE

1. Les postulats de base

La théorie Sens-Texte (= la TST), ainsi que le modèle correspondant (= le MST), s'appuient sur les trois postulats suivants :

1.1. Postulat 1

|| La langue naturelle est (considérée comme) une correspondance multi-multivoque entre un ensemble dénombrable de sens et un ensemble dénombrable de textes.

Symboliquement, le postulat 1 peut s'écrire comme suit :

$$(2) \{SENS_i\} \xrightarrow{\text{langue}} \{TEXTE_j\} \mid 0 < i, j < \infty$$

Un sens relève de la face interne de la parole; c'est un phénomène psychique — une information transmise (ou à transmettre) dans un événement langagier. Un texte relève de la face externe de la parole; c'est un phénomène physique — un ensemble de vibrations acoustiques servant de moyen de transmission dans un événement langagier.

Le postulat 1 exige trois commentaires importants.

Primo, le terme *sens* doit être interprété ici de la façon la plus restrictive possible. Il ne s'agit aucunement du sens dégagé à la suite d'une bonne compréhension d'un énoncé, c'est-à-dire extrait grâce à la logique, à nos connaissances extralinguistiques, etc. En d'autres termes, il ne s'agit pas du «vrai» sens, qui est, en fin de compte, la seule raison d'être de la communication. Dans (2), nous ne visons que le SENS PUREMENT LANGAGIER : le plus superficiel, le plus littéral, celui qui est accessible uniquement grâce à la mise en œuvre de la langue considérée. C'est seulement ce sens-là qui nous intéresse et que nous définissons comme suit :

|| Le sens est l'invariant des paraphrases langagières, ou, de façon plus développée,

|| Le sens est la seule propriété commune à tous les énoncés ayant le même sens (voir pp. 26-27).

De ce fait, les problèmes de la vérité, de l'absurdité, de la contradiction logique, etc. sont intentionnellement laissés en dehors de notre étude.

Secundo, le terme *texte* doit aussi être interprété d'une façon particulière. Il ne s'agit aucunement du texte cohérent, c'est-à-dire d'un discours organisé, donc du texte au sens de la

grammaire du texte ou des théories de la narration. Dans (2), nous appliquons le terme *texte* à la face externe, physique de toute manifestation langagière. Ainsi, nous considérons comme des textes les signifiants des morphes, des mots-formes, des phrases, des alinéas, etc., de même que des nouvelles et des romans entiers. *Texte* est, dans le cadre du MST, un terme technique qu'il faut utiliser strictement dans les limites indiquées.

Tertio, bien que les sens et les textes aient leur existence réelle, le MST, modèle rigoureusement linguistique, ne doit et ne peut pas les traiter dans leur réalité psychique (= neurologique) et physique. Le MST ne s'occupe que de leurs REPRÉSENTATIONS LINGUISTIQUES, c'est-à-dire de leurs descriptions au moyen de langages formels élaborés à cette fin. La représentation du sens est appelée *représentation sémantique* (RSém), la représentation du texte, *représentation phonologique* (RPhon). Divers langages formels, ou transcriptions, pour la RPhon sont connus en linguistique; mais pour la RSém, la linguistique ne dispose pas encore d'un langage formel universellement admis. La construction d'un tel langage, c'est-à-dire d'un langage formel sémantique, est, dès lors, une des tâches les plus pressantes de la linguistique moderne. Nous en reparlerons dans la section IV, pp. 27 ssq.

Étant donné le rôle primordial des représentations dans le MST, nous pouvons réécrire (2) comme (3) :

$$(3) \{RSém_i\} \xrightarrow{\text{langue}} \{RPhon_j\} \mid 0 < i, j < \infty$$

C'est sous cette forme que nous utilisons le postulat 1.

N.B. : En conformité avec la tradition de la linguistique, nous considérons surtout la langue parlée. Rien n'empêche, pourtant, que la langue écrite soit, elle aussi, prise en charge. Dans ce cas-là, on utilise la forme graphique conventionnelle du texte, c'est-à-dire le texte au sens courant du terme. Si l'on ne s'intéresse ni à la morphologie de surface ni à la phonologie, la forme graphique du texte est suffisante à toutes fins pratiques. Dans ce qui suit, nous nous limitons aux textes graphiques.

1.2. Postulat 2

|| La correspondance (3) doit être représentée par un dispositif logique, qui constitue un modèle fonctionnel de la langue.

Les sens et les textes d'une langue donnée sont directement perçus par les locuteurs, mais ce n'est pas le cas des règles qui les relient. Par conséquent, nous nous trouvons dans une situation classique, devant une «boîte noire» (*black box*) : en tant que linguistes, nous ne contrôlons que les entrées (= sens) et les sorties (= textes), la correspondance entre les deux étant inobservable. La seule solution qui se présente est un système de règles formelles simulant cette correspondance de la meilleure façon possible : c'est un MODÈLE FONCTIONNEL de la langue naturelle, ou un modèle Sens-Texte (MST).

Par sa nature, un MST complet doit être dynamique : il assure le passage entre un sens donné et tous les textes qui l'expriment (ou bien entre un texte donné et tous les sens qu'il exprime). Cependant, il est logique de distinguer, dans un MST, deux sous-modèles :

- 1) un système de règles purement linguistiques qui spécifient la correspondance $\{RSém_i\} \Leftrightarrow \{RPhon_j\}$ comme telle de façon STATIQUE;
- 2) un système de règles procédurales DYNAMIQUES qui spécifient le processus de passage entre les sens et les textes, c'est-à-dire qui effectuent la correspondance $\{RSém_i\} \Leftrightarrow \{RPhon_j\}$.

Il est important de souligner que les règles procédurales manipulent les données présentées par les règles linguistiques ; en d'autres termes, le second système ne fonctionne qu'à partir du premier. Or, nous croyons que ce second système de règles n'est pas spécifique à la linguistique : le même type général de règles procédurales est nécessaire partout où des règles emmagasinées (représentant des connaissances statiques) doivent être appliquées à la solution d'une tâche. Pour cette raison, nous faisons abstraction de ce second système (= des règles procédurales). Par conséquent, nous présentons le MST comme un système entièrement statique de correspondances entre sens élémentaires et textes élémentaires — si bien que, par exemple, le problème de l'ordonnancement des règles de correspondance ne se pose même pas.

Il est essentiel aussi d'ajouter que le MST n'est pas un système génératif ou transformationnel. Sa vocation n'est pas d'engendrer (= d'énumérer) l'ensemble de toutes les phrases correctes d'une langue et seulement de telles phrases, ni de transformer certaines entités linguistiques en d'autres entités. Un MST est un système purement équatif, ou transductif : il doit associer des RPhon à des RSém (et *vice versa*) conformément à l'intuition des locuteurs.

Nous avons déjà indiqué que la correspondance entre les RSém_i et les RPhon_j est multi-multivoque.

En effet, d'une part, pour un sens suffisamment complexe, on peut souvent construire des (centaines de) milliers de

paraphrases plus ou moins synonymes. Ainsi, la phrase anglaise (4) :

- (4) *The Food and Drug Administration has seriously cautioned expectant mothers to avoid one of life's simple pleasures: a cup of coffee* 'La Régie des denrées et des médicaments a adressé une mise en garde sérieuse aux futures mamans, pour qu'elles évitent un des simples plaisirs de la vie : le café'.

possède plus de 200 000 paraphrases (Mel'čuk 1981 : 31). La phrase française (8), pour laquelle nous citons des représentations illustratives, a (environ) un million de paraphrases (on le verra en détail dans la section IV). La synonymie est en effet extrêmement riche dans les langues naturelles.

D'autre part, un même texte peut correspondre à plusieurs sens ; ainsi, la phrase (5), citée dans Gentilhomme 1980 : 85-93 :

- (5) *Voici un buste superbe en marbre blanc de Carrare du duc d'Aumale quatrième fils de Louis-Philippe exécuté par Paul Dubois.*

admet, si l'on n'a pas recours à des moyens prosodiques et à des connaissances encyclopédiques, 9 072 interprétations (*buste* → *blanc* ou *marbre* → *blanc*; *marbre* → *de Carrare* ou *blanc* → *de Carrare*; *buste* → *exécuté*, *fils* → *exécuté* ou *L.-Ph.* → *exécuté*; etc.). L'ambiguïté (= homonymie et polysémie) est elle aussi une caractéristique des langues naturelles.

La complexité extraordinaire de la correspondance $\{RSém_i\} \Leftrightarrow \{RPhon_j\}$ nous pousse à postuler des niveaux de représentation intermédiaires, qui permettent de la diviser en composantes plus simples.

1.3. Postulat 3

Pour décrire la correspondance $\{RSém_i\} \Leftrightarrow \{RPhon_j\}$, deux niveaux intermédiaires de représentation linguistique sont nécessaires : la R(représentation) Synt(axique) et la R(représentation) Morph(ologique).

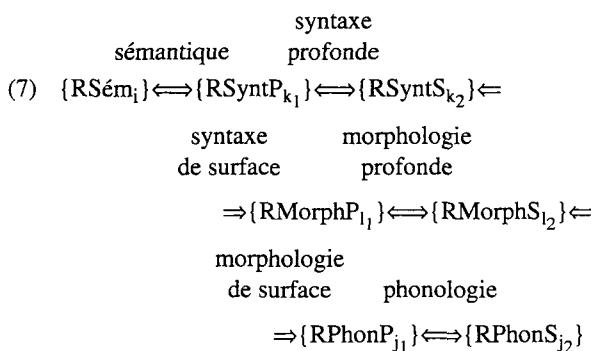
Par conséquent, nous pouvons réécrire (3) sous une forme plus développée, soit :

$$\begin{array}{ccccccc} & & & & & & \text{morphologie +} \\ & & & & & & \text{sémantique} & \text{syntaxe} & \text{phonologie} \\ (6) \quad \{RSém_i\} & \Leftrightarrow & \{RSynt_k\} & \Leftrightarrow & \{RMorph_l\} & \Leftrightarrow & \{RPhon_j\} \end{array}$$

Nous ne pouvons pas justifier ici les deux niveaux additionnels ; remarquons seulement qu'ils apparaissent dans toutes les traditions grammaticales, parce que toutes les langues possèdent des lois d'organisation dans deux domaines plus ou moins indépendants : la phrase et le mot, couverts respective-

ment par la syntaxe et la morphologie. (Pour quelques explications théoriques, voir Mel'čuk 1988a : 17 ssq., 48-49.)

De plus, tous les niveaux, sauf le niveau sémantique, doivent être «scindés» en deux sous-niveaux chacun : profond (-P), orienté vers le sens et relevant de celui-ci, et de surface (-S), orienté vers la forme et relevant de celle-ci. Par conséquent, le système entier des représentations linguistiques aura la forme suivante :



Dans (7), la R(eprésentation) Phon(ologique) de Surface est identique à la RPhon(ologique) de (3), c'est-à-dire la représentation phonétique, et la RPhonP n'est pas autre chose que la représentation phonologique au sens courant en linguistique. Les flèches doubles bidirectionnelles entre les ensembles de représentations linguistiques représentent, à leur tour, les composantes du MST, c'est-à-dire les ensembles de règles. Ainsi, la sémantique est, d'après (7), l'ensemble de règles qui associent à une RSém donnée toutes les RSyntP qui expriment le même sens (et *vice versa* : elles associent à une RSyntP toutes les RSém qui peuvent être exprimées par cette dernière). Dans les deux sous-sections suivantes, nous examinerons les représentations linguistiques des énoncés et les composantes du modèle. Mais avant d'aborder cette partie, nous devons formuler une thèse fort importante pour notre exposé.

De façon formelle, la correspondance $\{\text{SENS}_i\} \leftrightarrow \{\text{TEXTE}_j\}$ est symétrique, si bien que du point de vue purement logique, on pourrait décrire le MST dans n'importe quelle direction : soit à partir du sens vers le texte (= SYNTHÈSE), soit à partir du texte vers le sens (= ANALYSE). (Dans la plupart des théories linguistiques qui nous sont connues, la description se fait à partir du texte vers le sens, donc dans la direction de l'analyse.) Cependant, du point de vue linguistique, la direction Sens \Rightarrow Texte, c'est-à-dire la synthèse, nous semble beaucoup plus avantageuse. Il serait trop long d'étayer cette affirmation ici, aussi nous limiterons-nous à quelques remarques :

- La langue même accentue le rôle du locuteur par rapport à celui du destinataire : toutes les langues possèdent un verbe signifiant 'parler', mais aucune ne possède, à notre connaissance, de verbe spécial signifiant 'écouter / comprendre la parole'.

- Le locuteur est absolument indispensable pour l'acte langagier, mais non le destinataire. Ainsi, l'émission de parole peut avoir lieu même s'il n'y a pas de réception, tandis que la réception de parole n'est possible que si une émission a lieu. Il en découle qu'un acte langagier de communication n'est qu'un cas particulier d'un acte langagier tout court.

- Beaucoup de signes linguistiques font référence au locuteur (ce qu'on appelle des *shifters*, par exemple) ou aux attitudes du locuteur, etc.; les signes faisant référence au destinataire indépendamment du locuteur ou aux attitudes du destinataire sont inconnus.

- Dans un cas idéal, le locuteur a une information complète sur ce qu'il dit; il n'a que des obstacles purement linguistiques à surmonter, c'est-à-dire qu'il doit trouver les meilleurs moyens d'expression pour un contenu donné. (Bien entendu, ce tableau simplifie beaucoup la réalité.) Le destinataire, lui, ne connaît pas au préalable le contenu exact du message qui lui arrive; il doit l'extraire en mettant en œuvre non seulement ses connaissances linguistiques, mais aussi (et souvent, surtout) ses capacités logiques et ses connaissances extralinguistiques. Les activités du locuteur sont plus linguistiques que celles du destinataire.

- Un grand nombre de phénomènes linguistiques ne sont observables que dans la direction Sens \Rightarrow Texte. Ainsi, le syntagme *très fatigué* ne présente aucun problème (et aucun intérêt) du point de vue de l'analyse, alors que pour la synthèse, ce syntagme est extrêmement intéressant : on peut dire *très fatigué*, mais non pas, par exemple, **très armé*, bien que sémantiquement, cette expression agrammaticale soit correcte; comparez de même *très malade* vs **très blessé*, etc. Ces exemples illustrent le phénomène des *fonctions lexicales* (dont il sera question plus loin, pp. 31 ssq.), qu'on ne peut découvrir que si l'on adopte le point de vue de la synthèse.

Nous sommes donc amené à caractériser le MST exclusivement dans la direction de la synthèse et de façon strictement statique.

2. Représentations linguistiques (= représentations des énoncés)

Étant donné les contraintes d'espace, nous introduirons les niveaux de représentation linguistique adoptés dans le MST sans définitions rigoureuses et sans descriptions abstraites — à partir d'exemples accompagnés de quelques commentaires. Nous travaillerons à partir de la phrase suivante, tirée d'un magazine français :

- (8) *Le style des persécutions policières des gens de lettres en Union Soviétique a évidemment connu, depuis un demi-siècle, des changements sérieux.*

Pour cette phrase, nous présentons d'abord la RSém et ensuite la RSyntP et la RSyntS, ainsi que la RMorphP et la RMorphS. En même temps, nous citerons les RSyntP et les RSyntS de quelques phrases synonymes de (8), afin de montrer les possibilités de production des paraphrases dont dispose le MST. Les représentations sont accompagnées d'explications succinctes; nous prions le lecteur de garder à l'esprit le fait que ces représentations sont simplifiées et que beaucoup de choses importantes ne sont pas explicitées.

Puisque même un survol du MST et des représentations linguistiques qu'il utilise implique l'emploi systématique d'une vingtaine de concepts et de termes correspondants, nous sommes obligé d'avoir recours à des notations abrégées. Nous en donnons tout de suite une liste alphabétique.

-Anaph	: anaphorique	Rel-	: relation
-Comm	: communicatif	ṛ	: rhème
FL	: fonction lexicale	S-	: structure
-Morph-	: morphologique	-S	: de surface
MST	: modèle Sens- Texte	-Sém-	: sémantique
-P	: profond	SG	: sujet grammatical
-Phonét-	: phonétique	-Synt-	: syntaxique
PND	: prosodie neutre	⊥	: thème
R-	: représentation	$X + Y + \dots + Z$: expression phraséo- logique composée de lexèmes X, Y, ..., Z

2.1. Représentation sémantique

La nature et l'organisation de la RSém dans le cadre du MST seront spécialement discutées dans la section IV; ici, nous nous limiterons à l'exemple qui suit (fig. 1, p. 18). Cependant, quelques clarifications semblent nécessaires.

1. Les lettres A et B (en haut du diagramme) ainsi que les chiffres 1 et 2 (à gauche) ne font pas partie de la RSém, mais servent à identifier les zones du diagramme pour fins de référence.

2. La ligne hérissée (|||||) marque les frontières du thème (= ⊥) et du rhème (= ḙ), voir plus loin, pp. 27-28.

3. Le soulignement d'un élément sémantique (par exemple, 'devenir' dans B2) indique son rôle sémantiquement dominant (par rapport aux nœuds adjacents): cet élément est ICI un élément générique et/ou posé (= affirmé), alors que les autres éléments sémantiques sont soit des éléments spécifiques, soit des présupposés.

4. En lisant la RSém (9) [= fig. 1], il faut tenir compte du fait que chaque nœud d'où partent des flèches représente un élément prédicatif (au sens logico-sémantique); si un tel nœud est étiqueté avec un nom ou un adjectif, il faut y ajouter mentalement le verbe 'être' pour faciliter la compréhension. Ainsi, 'distance 3' [en A1] doit se lire comme '... est la distance temporelle entre ... et ...', 'façon II' [B1] — comme '... est la façon dont ...', etc.

5. Les chiffres romains et arabes qui suivent certains mots français dans le diagramme sont des numéros distinctifs; ils identifient les acceptations correspondantes des mots polysémiques, comme il est fait dans le DEC. (Ces numéros ont été empruntés au *Petit Robert*.)

La RSém (9) peut se lire en français de la façon suivante :

(9') 'La façon α dont la police de l'URSS, ayant le pouvoir institutionnalisé sur les gens de lettres qui habitent ce pays [A2], les tourmente usuellement sans relâche, par un traitement injuste, est devenue [= 'devenir complété'] très différente [B2] de la façon β [B1] dont la même police tourmente les mêmes gens; α est avant β [A1-2], la distance [temporelle] entre α et β étant de 50 ans [A1-B2]; ce fait est maintenant évident à tout le monde [B2].'

2.2. Représentation syntaxique profonde

Au niveau de la syntaxe profonde, la phrase (8) apparaît comme dans la figure 2, p. 19.

Pour donner une idée plus approfondie de la RSyntP, nous présenterons encore une RSyntP, cette fois pour la phrase (10), qui est synonyme (= une paraphrase) de la phrase (8) :

(10) *Chose certaine, la manière dont la police soviétique que persécute les gens de lettres est maintenant très différente de celle d'il y a cinquante ans.*

Les phrases (8) et (10) ont la même RSém : celle de la figure 1; par conséquent, les RSyntP des figures 2 et 3 (p. 19) sont toutes les deux synthétisées à partir de cette RSém.

Une RSyntP est composée, de façon générale, de quatre structures :

- Une structure syntaxique profonde (SSyntP), qui est un arbre de dépendances (et qui indique les dépendances syntaxiques entre les éléments terminaux de la phrase — entre les occurrences des lexies; remarquons que cette structure ne comporte pas de syntagmes, c'est-à-dire de non-terminaux).

– Une structure syntaxique communicative profonde (SSynt-CommP), qui consiste, *grossso modo*, en une spécification du thème et du rhème (tout comme dans la RSém).

– Une structure syntaxique anaphorique profonde (SSynt-AnaphP) : dans la figure 3, la flèche bidirectionnelle en pointillée spécifie la coréférence des deux unités lexicales MANIÈRE dans la partie gauche de l'arbre.

– Une structure syntaxique prosodique profonde (SSynt-Prosp) : elle est marquée, de façon extrêmement simpliste, par le sigle PND (= prosodie neutre déclarative).

Nous caractériserons plus en détail la SSyntP. Pour ce faire, nous indiquerons ses trois propriétés majeures :

1. L'arbre de la SSyntP n'est pas ordonné linéairement ; donc, la distribution spatiale de ses nœuds dans un diagramme n'a aucune pertinence. L'ordre linéaire des mots-formes de la phrase réelle est déterminé par les règles de la composante syntaxique de surface (voir plus loin, pp. 25-26, *linéarisation*).

2. Les branches de la SSyntP sont étiquetées de noms de *relations syntaxiques profondes* (RelSyntP), qui sont des

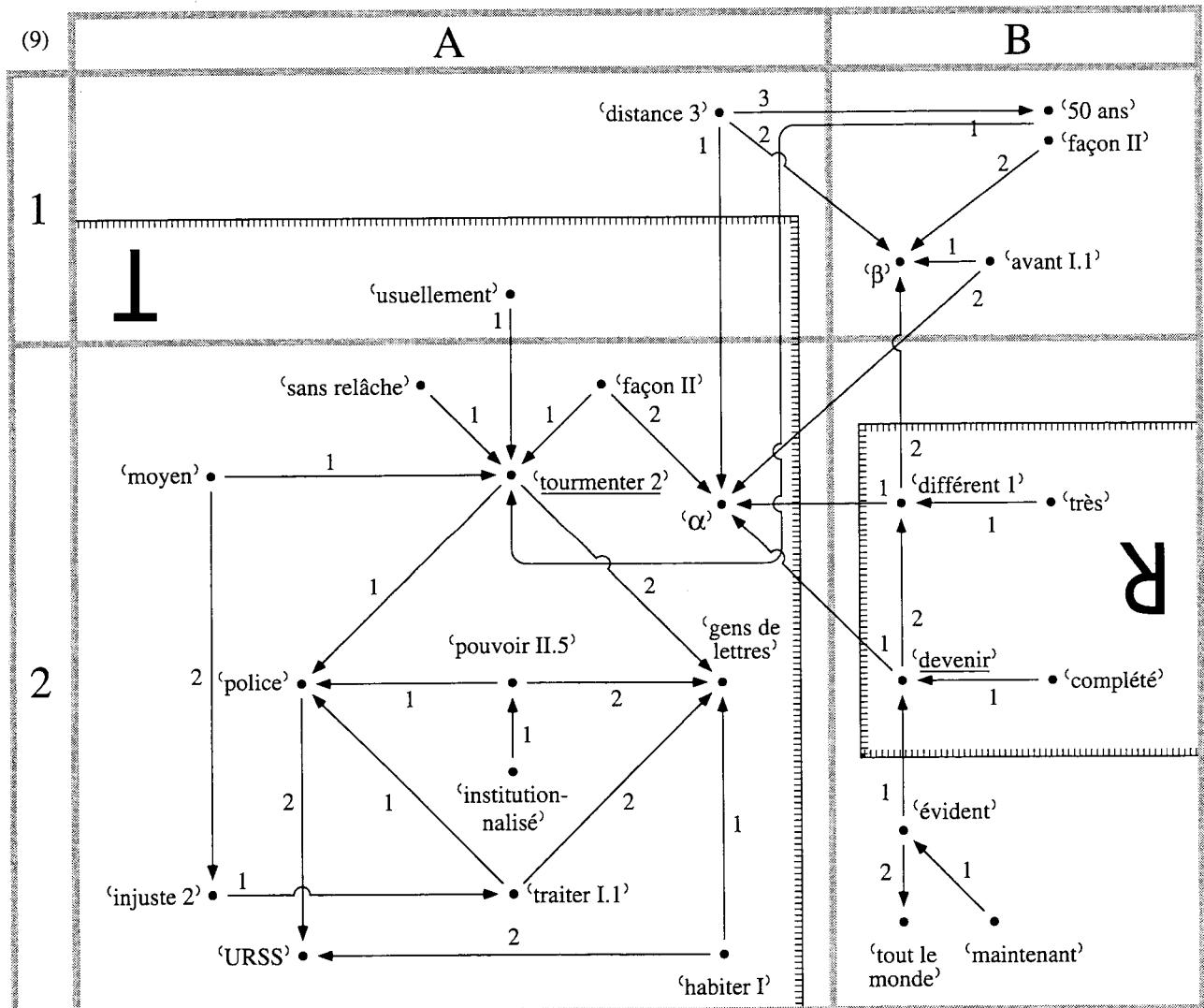


Figure 1 – Représentation sémantique de la phrase (8) [et de toutes les phrases qui lui sont synonymes]

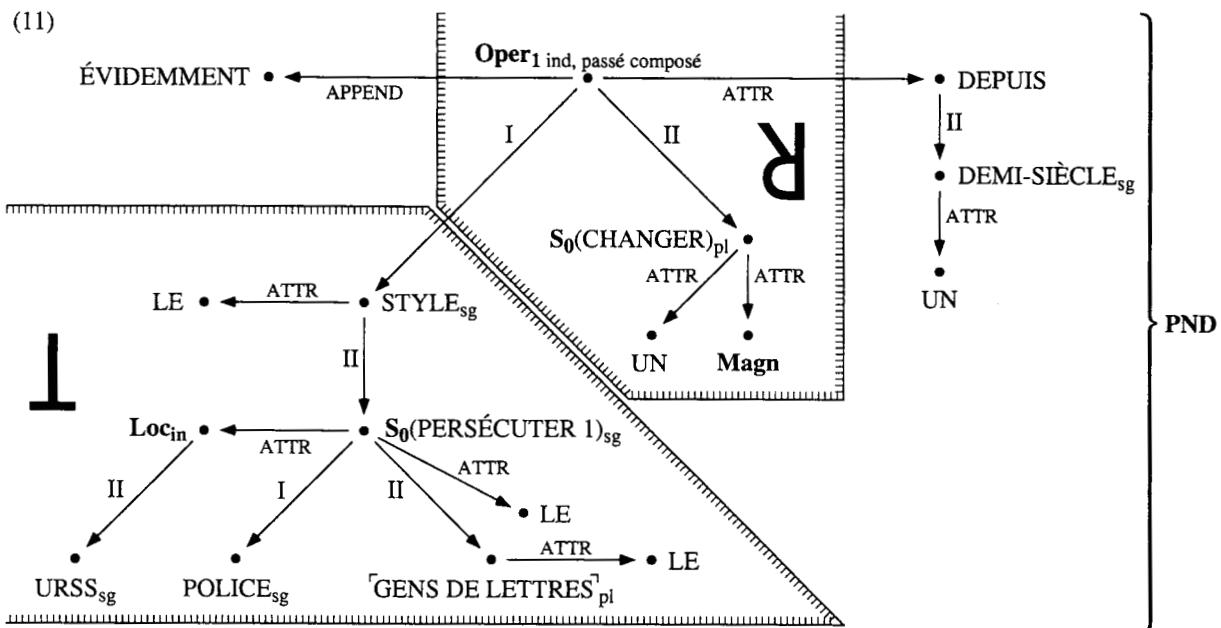


Figure 2 – Représentation syntaxique profonde de la phrase (8)

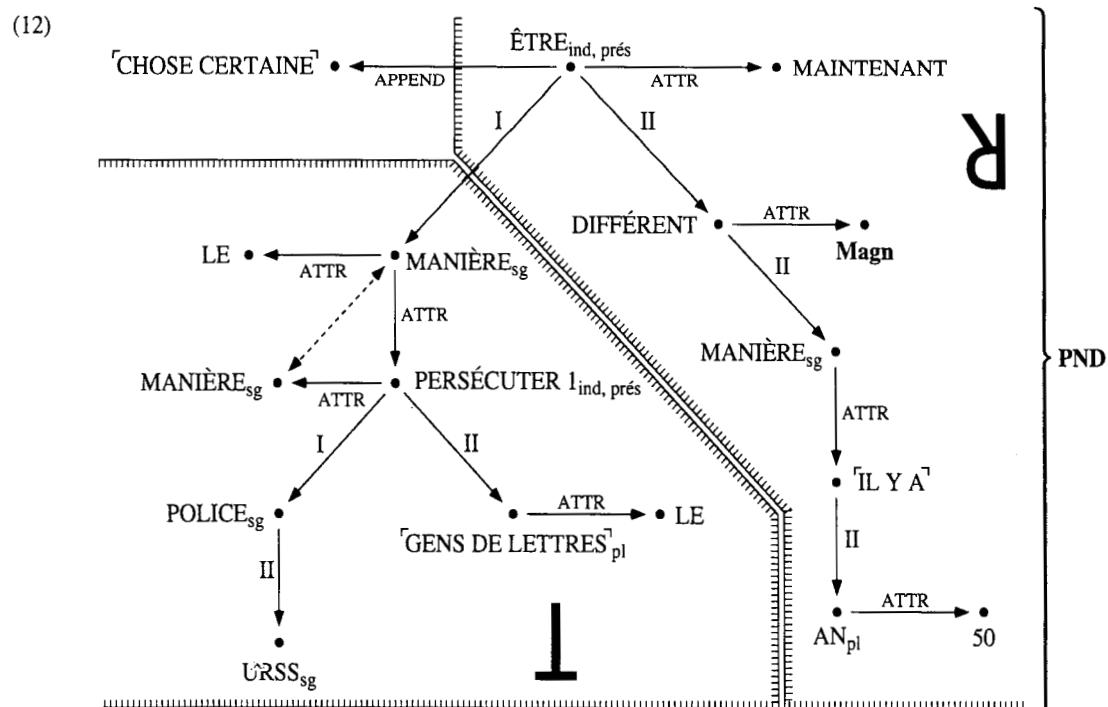


Figure 3 – Représentation syntaxique profonde de la phrase (10)

relations syntaxiques universelles, au nombre de neuf : six RelSyntP actancielles (I, II, ..., VI), une RelSyntP modificative (ATTR), une RelSyntP coordinative (COORD) et une RelSyntP «appenditive» (APPEND). Chaque RelSyntP représente, de façon générale et abstraite, toute une famille de diverses constructions syntaxiques des langues naturelles. Ainsi, la RelSyntP I correspond à la construction prédicative (= V(erbe)_{fini} —→ S(ujet) G(rammatical)) et à tous ses «transformés» (par exemple, *persécution* —→ *policière*, puisque ‘La police persécuté’, etc.); la RelSyntP II correspond aux constructions avec le complément d’objet direct (CO^{dir}) et à ses «transformés», la RelSyntP III — aux constructions avec le CO^{indir}, les RelSyntP IV-VI, aux constructions avec le CO^{oblique}; ATTR représente toutes les constructions modificatives (épithètes, compléments de nom, etc.); COORD, toutes sortes de coordinations; et APPEND, toutes les constructions «extrastructurales» (parenthétiques, d’adresse, etc.).

3. Les nœuds de la SSyntP sont étiquetés de lexèmes de la langue naturelle en question, avec la restriction suivante : ces lexèmes doivent être ce que nous appelons des *lexèmes profonds*. Par conséquent, on trouve quatre différences principales entre le lexique de la phrase représentée et le lexique de sa SSyntP (= lexique profond) :

(i) Les lexèmes «vides» (verbes auxiliaires; prépositions et conjonctions régies, qui sont censées être des moyens d’expression syntaxique) n’apparaissent jamais dans la SSyntP. Ainsi, les deux prépositions DE dans la phrase (8), celle reliant STYLE à PERSÉCUTIONS et celle reliant PERSÉCUTIONS à ‘GENS DE LETTRES’, sont absentes de la figure 2. En sont aussi absents les auxiliaires : une forme verbale composée n’est représentée dans la SSyntP que par un seul nœud (cf. *a connu*, qui apparaît dans la figure 2 comme Oper₁ indicatif, passé composé).

(ii) Un idiotisme, ou phrasème (= expression multilexémique ayant des propriétés qu’on ne peut déduire des propriétés de ses constituants), est représenté par un seul nœud : ‘gens de lettres’, figures 2 et 3, ‘chose certaine’ et ‘il y a’, figure 3.

(iii) La SSyntP n’admet pas de pronoms ni de mots pronominaux : ici, seuls leurs antécédents apparaissent, telles les deux occurrences du lexème MANIÈRE, figure 3, qui doivent être remplacées dans la structure syntaxique de surface par DONT et CELLE.

(iv) La SSyntP n’admet pas non plus certains lexèmes qui sont choisis en fonction d’autres lexèmes avec lesquels les premiers sont syntaxiquement liés; les lexèmes ainsi exclus sont représentés dans la SSyntP par des symboles spéciaux. Par exemple, on dit CONNAÎTRE des changements (= ‘changer’), MENER une lutte (= ‘lutter’), PORTER plainte (= ‘se

plaindre’), FAIRE un voyage (= ‘voyager’), DONNER un ordre (= ‘ordonner’), etc. : le verbe en majuscules est choisi en fonction du nom correspondant, qui est son complément d’objet. Le tout est une locution phraséologique d’un type particulier, le verbe en question représentant la valeur de ce que nous appelons la fonction lexicale Oper₁, dont l’argument est le nom objet. Comparez :

fonction(argument)	= valeur
Oper ₁ (changement)	= connaître
Oper ₁ (lutte)	= mener
Oper ₁ (plainte)	= porter

et ainsi de suite. (Pour Oper₁, voir plus loin, p. 32.) Les fonctions lexicales (FL), phénomène de la syntaxe profonde, sont universelles (en ce sens qu’on les trouve dans toutes les langues : pp. 33-34); leur nombre est d’une cinquantaine. Elles forment un système très cohérent et sont sous-jacentes au système de paraphrasage, qui sera décrit dans la section V; c’est à ce moment-là que nous reprendrons la question des fonctions lexicales. Ici, il suffit de dire que dans la SSyntP, la valeur d’une FL est systématiquement représentée par le symbole de la FL elle-même. Ainsi, au lieu de connaître —→ change-ments, nous écrivons, dans la SSyntP, Oper₁ —→ S₀(chan-ger)_{pl} [S₀ étant un nom d’action]; et au lieu de change-ments ATTR. sérieux, on trouve S₀(changer)_{pl} ATTR. Magn [Magn étant un intensificateur consacré par l’usage]; voir figure 2.

Avant de passer à la discussion de la représentation syntaxique de surface, signalons encore que les nœuds de la RSyntP sont munis de caractéristiques (= variables) morphologiques, ou flexionnelles, sémantiquement pleines : le mode et le temps pour les verbes, le nombre pour les noms. Les autres caractéristiques flexionnelles (le nombre et le genre des déterminatifs et des adjectifs; la personne et le nombre des verbes) sont des moyens d’expression syntaxiques (donc sémantiquement vides), et elles doivent être calculées — à partir de la RSyntS — par les règles de la syntaxe de surface (voir pp. 25-26, la *morphologisation*).

2.3. Représentation syntaxique de surface

Nous présenterons ici les RSyntS pour les phrases (8) et (10), c’est-à-dire les RSyntS qui correspondent aux RSyntP (11) et (12). Ces RSyntS ont la même organisation générale que les représentations syntaxiques profondes : elles comportent, au niveau de surface, les mêmes quatre structures (la SSyntS, la SSynt-CommS, la SSynt-AnaphS et la SSynt-ProspS), qui remplissent les mêmes rôles que leurs partenaires profonds.

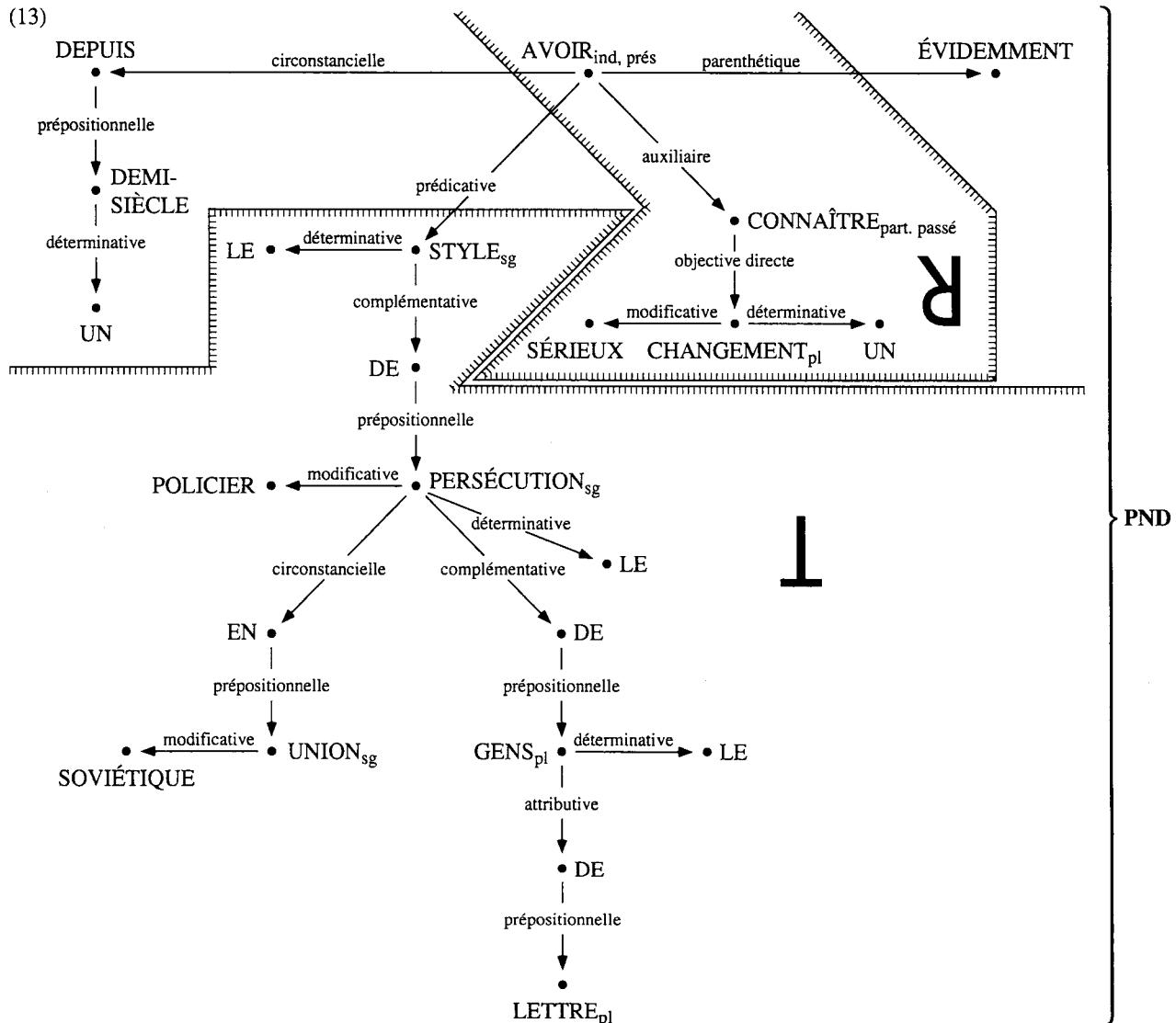


Figure 4 – Représentation syntaxique de surface de la phrase (8)

La structure syntaxique de surface (SSyntS) possède, de façon générale, les mêmes propriétés formelles que la SSyntP : c'est aussi un arbre de dépendances linéairement non ordonné et dont les nœuds ne comportent que les caractéristiques morphologiques sémantiquement pleines. Par conséquent, ce que nous avons dit au sujet de la linéarisation et de la morphologisation à propos de la SSyntP vaut également pour la SSyntS. Cependant, la composition et l'étiquetage de la SSyntS sont différents de ce que nous trouvons dans la SSyntP.

Quant aux branches de la SSyntS, elles représentent des constructions syntaxiques particulières et spécifiques de la

langue en cause (en l'occurrence, le français), dont elles portent les noms. (Dans le cadre de cet exposé, il nous est impossible de jeter plus de lumière sur le concept de construction syntaxique de surface ou de justifier notre choix de constructions et/ou de noms pour notre illustration.)

Quant aux nœuds de la SSyntS, chacun correspond à un mot-forme de la phrase et est étiqueté par le lexème correspondant muni de toutes les caractéristiques flexionnelles. Cela veut dire, entre autres, qu'à ce niveau-là, les mots outils (les servitudes grammaticales) sont introduits, les phrasèmes sont représentés par des sous-arbres spécifiant leur structure interne

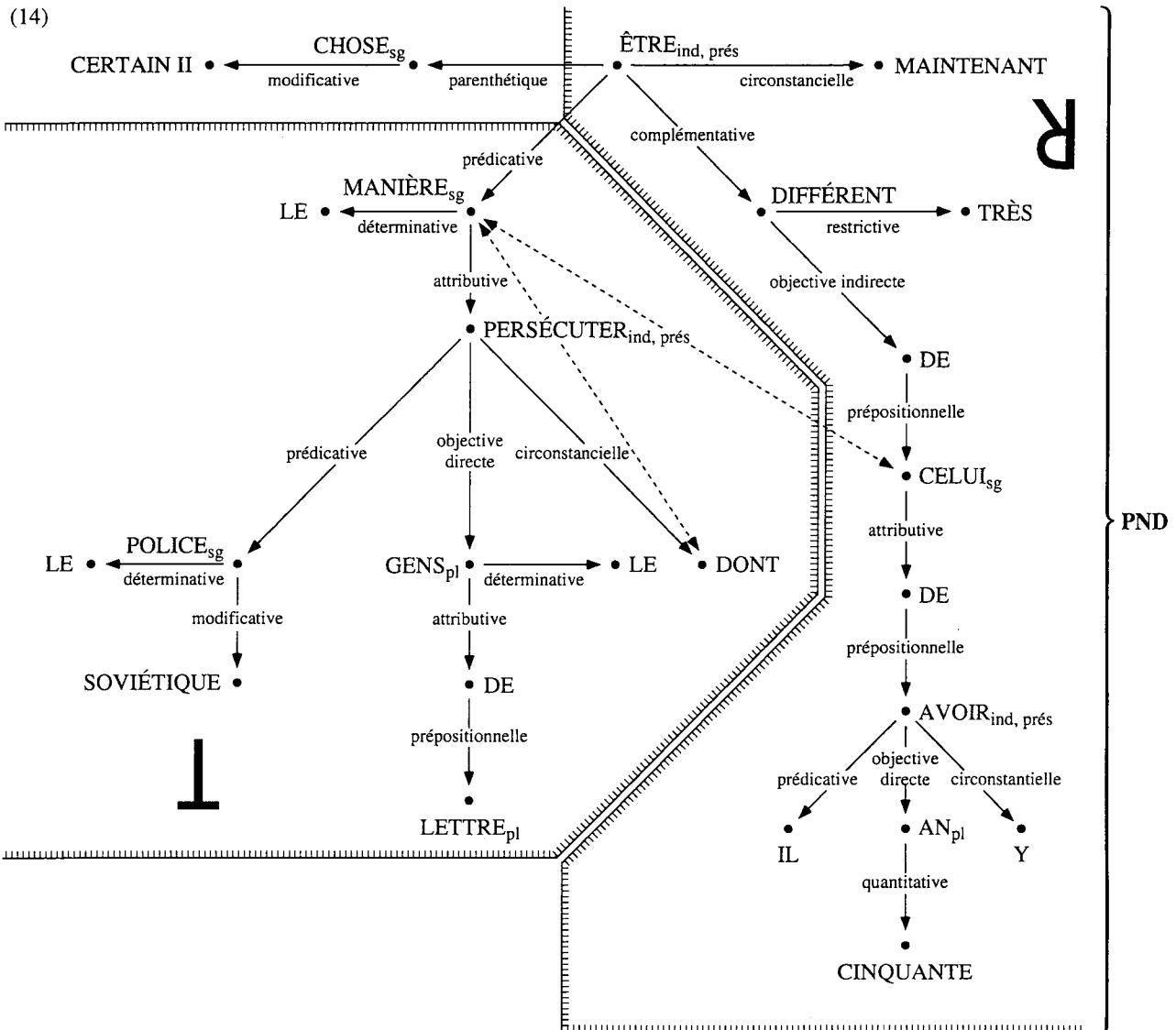


Figure 5 – Représentation syntaxique de surface de la phrase (10)

en termes de lexèmes constituants, toutes les pronominalisations sont effectuées et les FL sont remplacées par leur valeur calculée à partir de l'argument correspondant.

2.4. Représentation morphologique profonde

Nous présentons maintenant les RMorphP des phrases (8) et (10), c'est-à-dire les RMorphP qui correspondent aux RSyntS (13) et (14); voir fig. 6 et 7 à la page suivante.

La RMorphP d'une phrase est constituée de deux structures : une structure SMorphP de phrase et une structure morphologique prosodique SMorph-Prosp de phrase.

La première est la chaîne de RMorphP des mots-formes qui composent la phrase. La deuxième comporte les indications des pauses (I, II et III, signalant les pauses de plus en plus importantes), des intonations (—, —, etc.), et des groupes rythmiques (X Y signifie que X et Y forment un groupe rythmique et ne peuvent être séparés par une pause). C'est à ce niveau que les constituants (ou les syntagmes) sont indiqués; pour nous, ils relèvent de la représentation morphologique, et non syntaxique, de la phrase.

La RMorphP d'un mot-forme est le nom du lexème correspondant muni des valeurs de toutes les variables

(15) LE_{masc,sg} STYLE_{sg} DE LE_{fém,sg} PERSÉCUTION_{sg}
 POLICIER_{fém,sg} | DE LE_{masc,pl} GENS_{pl} DE
 LETTRE_{pl} EN UNION_{sg} SOVIÉTIQUE_{fém,sg} ||
 AVOIR_{ind,prés,3sg} ÉVIDEMMENT
 CONNAÎTRE_{part.passé} DE_{masc,pl} CHANGEMENT_{pl}
 SÉRIEUX_{masc,pl} | DÉPUIS
 UN_{masc,sg} DEMI-SIECLE_{sg} |||

Figure 6 – Représentation morphologique profonde de la phrase (8)

(16) CHOSE_{sg} CERTAIN_{fém,sg} || LE_{fém,sg} MANIÈRE_{sg} |
 DONT LE_{fém,sg} POLICE_{sg} SOVIÉTIQUE_{fém,sg}
 PERSÉCUTER_{ind,prés,3sg} LE_{masc,pl} GENS_{pl} DE
 LETTRE_{pl} || ÊTRE_{ind,prés,3sg} MAINTENANT TRÈS
 DIFFÉRENT_{fém,sg} | DE CELUI_{fém,sg} | DE IL_{sujet}
 Y AVOIR_{ind,prés,3sg} CINQUANTÉ AN_{pl} |||

Figure 7 – Représentation morphologique profonde de la phrase (10)

morphologiques (= flexionnelles) nécessaires pour spécifier ce mot-forme de façon univoque; les caractéristiques flexionnelles qui servent comme marqueurs de relations syntaxiques, c'est-à-dire les caractéristiques déterminées par l'accord et le régime, sont ajoutées par les règles de syntaxe de surface.

2.5. Représentation morphologique de surface

La RMorphS d'une phrase est différente de la RMorphP de la même phrase seulement en ce que les RMorphP des mots-formes sont remplacées par les RMorphS des mêmes mots-formes. La RMorphS d'un mot-forme représente celui-ci par des morphèmes et des opérations morphologiques significatives, reflétant ainsi sa structure interne (ce que la RMorphP d'un mot-forme ne fait pas). Puisque, d'une part, la RMorphS d'une phrase est en général peu pertinente pour la paraphrase et que, d'autre part, la morphologie du français est plutôt pauvre, la RMorphS d'un mot-forme est peu intéressante pour nous; nous ne donnons ci-dessous la RMorphS que d'une partie de la phrase (8):

Nous ne citerons pas les représentations phonologiques, étant donné leur non-pertinence pour la paraphrase.

(17) {LE} + {MASC.SG} {STYLE} + {SG} {DES} {PER-
 SÉCUTION} + {PL} {POLICIER} + {FÉM} + {PL}
 {DES} {GENS} {DE} {LETTRE} + {PL} {EN}
 {UNION} + {SG} {SOVIÉTIQUE} + {FÉM} +
 {SG} || {AVOIR} + {IND.PRÉS} + {3SG} ...

Figure 8 – Représentation morphologique de surface (partielle) de la phrase (8)

3. Les composantes du modèle Sens-Texte

La correspondance entre deux représentations de niveaux adjacents est établie par une composante spéciale du MST. Le MST peut être compris, sous cet angle, comme une chaîne de production, dont chaque machine-outil (= une composante) usine la pièce en cause (= la phrase à produire) d'une façon spécifique. Le MST possède six composantes principales :

- la sémantique,
- la syntaxe profonde,
- la syntaxe de surface,
- la morphologie profonde,
- la morphologie de surface,
- la phonologie,⁶

dont nous ne caractériserons — très brièvement — que les trois premières (le reste étant marginal du point de vue de la paraphrase).

3.1. Sémantique

Cette composante établit la correspondance {RSém_j} ↔ {RSyntP_k} au moyen de huit opérations, que nous ne décrirons pas ici (par exemple, le sectionnement de la RSém de départ en « morceaux » dont la taille correspond à la taille sémantique d'une phrase; réunion des unités sémantiques en faisceaux correspondant aux sens des lexèmes; etc.). Ces opérations sont effectuées par des règles sémantiques, qui peuvent être divisées en deux classes majeures :

- les règles sémantiques au sens propre, soit les règles qui font correspondre à une RSém quelques RSyntP de base qui portent le même sens (nous ne préciserons pas quelles RSyntP s'obtiennent à ce niveau) et *vice versa*;
- les règles sémantiques de paraphrase qui font correspondre à une RSyntP de base toutes les RSyntP qui lui sont synonymes — et *vice versa*.

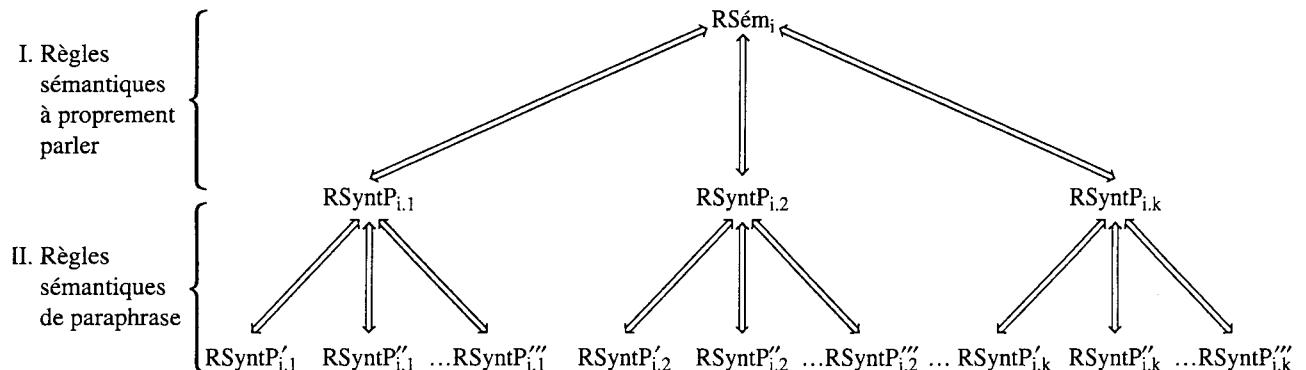
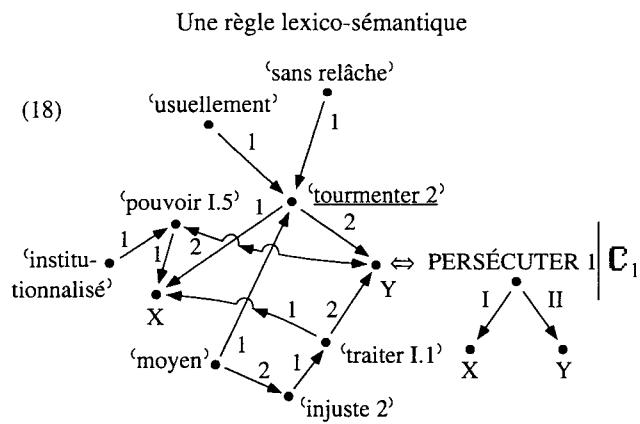


Figure 9 – Correspondances entre une RSém donnée et toutes les RSyntP qui portent le même sens

Ainsi nous établissons la correspondance entre les deux ensembles : {RSém_i} et {RSyntP_k} — en deux temps, ce qu'on peut représenter par le diagramme ci-dessus, fig. 9 :

Les règles sémantiques de paraphrase — ou le système de paraphrasage — sont décrites dans la section V. Dans cette section, nous ne traiterons que des règles sémantiques au sens propre et n'en illustrerons que trois types (il y en a beaucoup plus!).

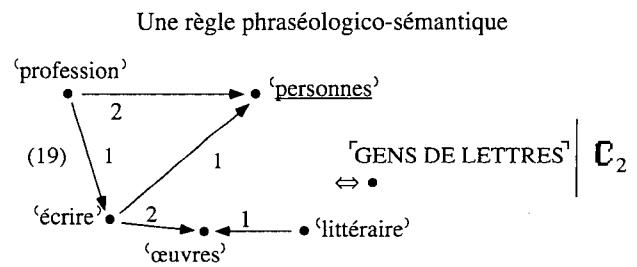
Les règles citées ici, de même que celles citées pour les autres composantes du MST, font partie des règles nécessaires à la construction de la phrase (10); de cette façon, le lecteur sera mieux à même de vérifier l'application des règles du MST.



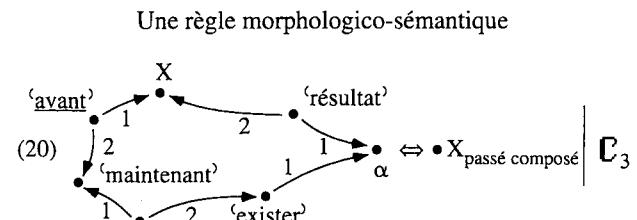
Une telle règle n'est pas, *grossso modo*, autre chose que l'article de dictionnaire pour le lexème en question (= PERSÉCUTER 1 ayant le sens 'persécuter 1'), la partie de gauche étant la définition (c'est-à-dire une décomposition du sens 'persécuter 1'), et l'arbre dans la partie de droite représentant (très approximativement) le régime. (Le soulignement d'un élément dans la partie gauche de la règle (18) joue le même

rôle que dans la RSém, c'est-à-dire qu'il indique l'élément sémantiquement dominant.) De plus, un article de dictionnaire est censé contenir les données sur la combinatoire syntaxique et lexicale du lexème vedette, c'est-à-dire ses traits syntaxiques, ses fonctions lexicales, etc. Sous forme d'une règle lexico-sémantique, ces données devraient constituer la partie conditions (de la règle), que nous avons notée par C pour raisons de simplicité.

Le nombre de règles lexico-sémantiques est égal au nombre de significations lexicales (\approx sémantèmes) existant dans la langue, soit de l'ordre de $10^5 \sim 10^6$ (environ un million).⁷



Une règle phraséologico-sémantique est aussi un article de dictionnaire, cette fois-ci, pour le phrasème en question. Le nombre de telles règles est de l'ordre de $10^3 \sim 10^4$ (quelques milliers environ).



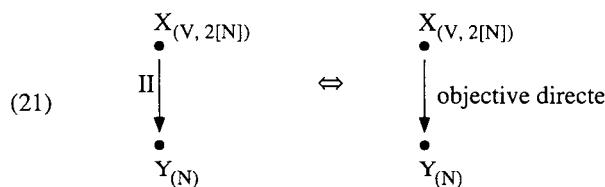
Cette règle représente — encore une fois, de façon très approximative — une des significations du passé composé français (= ‘une action passée dont le résultat existe à présent’). Le nombre de règles similaires est, en moyenne, de l’ordre de $10^1 \sim 10^2$ (quelques dizaines).

L’application des règles sémantiques à une RSém de départ en vue de l’obtention des RSyntP synonymes est un processus assez complexe, dont nous ne traiterons pas ici. La même chose est vraie pour les règles des autres composantes du modèle : elles sont toutes présentées sous leur aspect statique, en ignorant leur aspect dynamique ou procédural. (Au sujet de l’implémentation d’une procédure d’application des règles sémantiques, voir Iordaneskaja et Polguère 1988.)

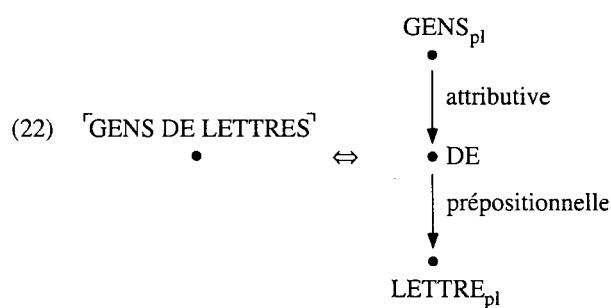
3.2. Syntaxe profonde

La composante syntaxique profonde (SyntP) établit la correspondance $\{\text{RSyntPk}_1\} \Leftrightarrow \{\text{RSyntSk}_2\}$ au moyen de cinq opérations (effectuées par les règles syntaxiques profondes), dont trois seront citées en tant qu’illustration.

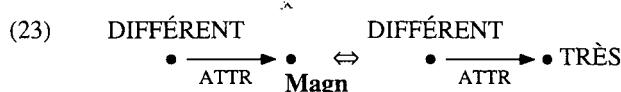
Une règle syntaxique profonde de structuration



Une règle syntaxique profonde phraséologique (impliquant un phrasème)

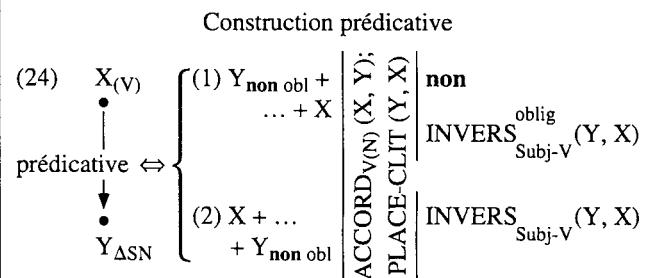


Une règle syntaxique profonde lexico-fonctionnelle (impliquant une fonction lexicale)

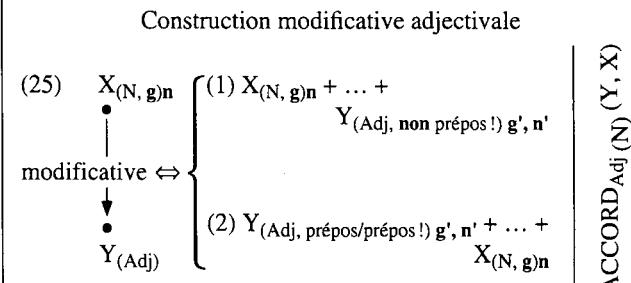


3.3. Syntaxe de surface

La composante SyntS établit la correspondance $\{\text{RSyntSk}\} \Leftrightarrow \{\text{RMorphP}_1\}$, ce qui revient, *grossièrement*, à la **morphologisation** et à la **linéarisation** de la SSyntS. (Nous ne mentionnerons pas ici les autres opérations effectuées par cette composante, telles l’ellipse, la formation des constituants et la prosodisation; pour le texte écrit, la prosodisation est représentée par la ponctuation.) Les nœuds de la SSynt reçoivent les marqueurs morphologiques et sont linéairement ordonnés à l’aide de règles SyntS, qui sont l’outil de base de cette composante. Nous en citerons deux, nécessaires pour le passage (14) \Leftrightarrow (16), c’est-à-dire de la SSyntS de la phrase (10) à sa SMorphP.



Cette règle stipule que, pour réaliser une construction prédictive avec un syntagme nominal SN ($= \Delta \text{SN}$) comme sujet, on doit mettre le sujet soit avant le verbe (la sous-règle (1) : seulement si la fonction standard «inversion obligatoire du sujet et du verbe» ne s’applique pas), soit après le verbe (la sous-règle (2) : seulement si la fonction standard «inversion du sujet et du verbe» s’applique); le sujet ne doit pas être au cas oblique (l’accusatif ou le datif; cette exigence n’est pertinente que pour le pronom : *me* vs *je*, *le / lui* vs *il*, *que* vs *qui*); le verbe doit s’accorder avec le sujet (la fonction standard ACCORD_{V(N)}); et le sujet manifesté par un pronom clitique ne doit pas être séparé du verbe si ce n’est par d’autres clitiques (la fonction standard PLACE-CLIT). Le ΔSN et les fonctions standard sont définis ailleurs dans la syntaxe de surface; il n’en sera pas question ici.



La règle (25) réalise la construction modificative «nom + adjectif» : l’adjectif suit le nom s’il n’est pas marqué «seulement (= !) prépos(itif)»; autrement, il précède le nom; s’il est

marqué «prépos(itif)» (mais sans «!»), alors il peut soit suivre, soit précéder le nom, et le choix se fait selon d'autres règles. De plus, l'adjectif s'accorde avec le nom selon le genre (= g) et le nombre (= n), ce qui est assuré par la fonction standard ACCORD_{Adj(N)}.

Nous ne toucherons pas aux autres types de règles SyntS (par exemple, règles d'ordonnancement de mots dans les syn-

tagmes, règles globales de l'ordre des mots et des syntagmes dans les phrases, etc.).

Il semble que notre description soit suffisante pour donner une idée des représentations manipulées par le modèle Sens-Texte et du modèle lui-même.⁸ Nous pouvons maintenant étudier la représentation sémantique, sujet qui a une pertinence spéciale du point de vue de la paraphrase.

IV. REPRÉSENTATION SÉMANTIQUE DANS LE MODÈLE SENS-TEXTE

Nous avons déjà dit que le sens linguistique (= langagier) est l'invariant des paraphrases (synonymes); la représentation sémantique est un moyen formel de décrire cet invariant — dans le but d'en dériver l'ensemble des paraphrases possibles. Voyons de plus près ce moyen, qui est en fait un langage artificiel que nous créons en vue de décrire les sens langagiers — une (MÉTA)LANGUE SÉMANTIQUE.

La phrase (8), qui nous a servi de point de départ, et sa paraphrase (10) admettent beaucoup d'autres paraphrases; en voici deux :

(26) *Bien entendu, le style des persécutions des gens de lettres soviétiques par la police a sérieusement changé au cours des cinquante dernières années.*

(27) *Il est clair que la manière dont la police en U.R.S.S. persécute les gens de lettres est aujourd'hui très différente de ce qu'elle était il y a cinq décennies.*

Pour évaluer, ne serait-ce que très approximativement, le nombre total de paraphrases dans cette famille, nous diviserons les phrases (8), (10), (26) et (27) en «tranches» sémantiques et nous indiquerons des variantes d'expression synonymes ou quasi synonymes pour chaque tranche; voir figure 10 ci-dessous.

Le chiffre en gras en bas d'une colonne de la figure 10 indique le nombre de variantes dans cette colonne. Supposant que chacune des variantes dans une colonne peut se combiner

A	B	C	D
bien {sûr entendu}			
évidemment			
certainement			
certes			
Il va de soi que			
Il est {évident certain clair indéniable} que	le style de {la façon la manière} dont	subir les persécutions {policières par la police} la police {persécute poursuit harcèle}	gens de lettres auteurs écrivains
chose {évidente certaine}			
Il n'est pas niable que			
On ne peut nier que			
On est certain que			
15	3	5	3

Figure 10 – Variantes d'expression pour les paraphrases des phrases (8), (10), (26) et (27) [voir la suite à la p. 19]

avec n'importe quelle variante d'une autre colonne, nous pouvons, tout simplement, multiplier ces chiffres, ce qui donne $15 \times 3 \times 5 \times 3 \times 3 \times 15 \times 9 \times 4 \times 4 = 4\,374\,000$ paraphrases!

Soulignons que ce ne sont pas des «paraphrases» du type dont s'occupait M. Jourdain : *Belle marquise, vos beaux yeux me font mourir d'amour ~ D'amour mourir me font, belle marquise, vos beaux yeux ~ ...*, où la seule différence est la permutation des mots et des syntagmes. Les paraphrases dont il est question ici se distinguent entre elles par au moins une lexie (et dans la plupart des cas, par la construction syntaxique); les différences d'ordonnancement ne sont pas prises en compte.

Beaucoup de combinaisons obtenues mécaniquement seront éliminées en raison de contraintes de cooccurrence; cependant, on pourrait ajouter d'autres variantes au tableau de la figure 10 (par exemple, *ont connu des modifications radicales ~ différent beaucoup, aujourd'hui ~ maintenant, il y a 50 ans ~ 50 ans plus tard ~ au cours de 50 années*, etc.). On obtient ainsi deux ou trois millions de paraphrases pour notre phrase. Cela montre clairement à quel point la synonymie des langues naturelles est riche.⁹

La théorie Sens-Texte pose comme tâche primordiale une description formelle et exhaustive de la synonymie de phrases,

c'est-à-dire de la richesse paraphrasique. Cette description présuppose, à son tour, la description de l'invariant sémantique des paraphrases observables, cette dernière n'étant pas autre chose que la représentation sémantique (RSém) dont on traite ici.

Du point de vue formel, la RSém proposée est l'ensemble de trois objets appelés *structures* :

- La structure sémantique (SSém) représente la partie «objective» du sens en question : elle correspond au sens situationnel (p. 11) et spécifie les objets qui composent la situation décrite par les paraphrases ainsi que toutes les relations entre eux; voir plus loin.

- La structure sémantico-communicative (SSém-Comm) représente l'itinéraire que le locuteur emprunte à travers la situation évoquée : ce qu'il veut prendre comme point de départ vs ce qu'il prendra comme point d'arrivée; ce qu'il veut poser d'abord vs ce qu'il veut discuter ensuite; ce qu'il affirme vs ce qu'il présuppose; ce qu'il veut mettre en relief vs ce qu'il veut renvoyer à l'arrière-plan, etc. La division en thème et rhème (\mathbb{L} vs \mathbb{R}), marquée dans la figure 1, appartient à la SSém-Comm, aussi bien que le soulignement des deux nœuds [B2], qui signale la dominance sémantique (le caractère générique

E	F	G	H	I
en {U.R.S.S. Union Soviétique} soviétique	a changé a {subi connu} des changements des changements sont survenus s'est transformé a {subi connu} des transformations des transformations sont survenues est devenu {différent autre} n'est plus le même {a été s'est} modifié a {subi connu} des modifications	bien beaucoup profondément sérieusement de façon {profonde sérieuse} tout à fait très [différent] tout [autre]	au cours de depuis il y a pendant	50 {années ans} cinq décennies demi-siècle
3	15	9	4	4

Figure 10 (suite) – Variantes d'expression pour les paraphrases des phrases (8), (10), (26) et (27)

et/ou l'affirmation comme opposée à la présupposition : 'Il est évident que X est devenu ...').

- La structure rhétorique (SRhét) représente les buts ou les effets artistiques visés par le locuteur : veut-il que son texte soit pathétique, ironique ou neutre, qu'il fasse rire ou pleurer, qu'il soit de registre neutre, familier ou élevé, etc.

La SSém-Comm et la SRhét ne seront pas décrites ici malgré leur importance pour la paraphrase; nous ne nous occuperons que de la SSém.

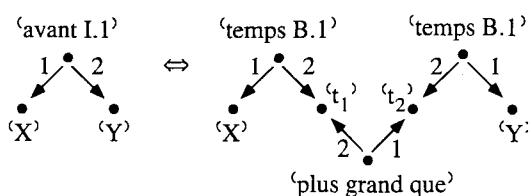
La STRUCTURE SÉMANTIQUE spécifie le sens langagier de l'énoncé (et de tous les énoncés qui lui sont synonymes, donc paraphrasables) indépendamment de sa forme langagière. Ainsi, la distribution du sens en mots, syntagmes, propositions et phrases, son expression par des moyens prosodiques ou par des constructions syntaxiques particulières, etc. ne sont pas prises en considération dans la SSém. Notamment, les différences entre affirmation et question, entre communication neutre, moquerie, menace ou pitié, entre ordre, requête ou imploration — qui très souvent sont exprimées par des phénomènes prosodiques — doivent être reflétées dans la SSém de la même façon technique que les différences entre 'aider' et 'assister' ou entre 'tasse', 'gobelet' et 'verre'.

Du point de vue formel, une SSém est un graphe connexe orienté étiqueté, c'est-à-dire un RÉSEAU : un ensemble de points (= sommets, ou nœuds) reliés par des flèches (= arcs); les nœuds et les arcs sont porteurs d'étiquettes, voir *infra*.

Un nœud d'une SSém est étiqueté par une unité sémantique, ou un sémantème. Un SÉMANTÈME correspond (*grossost modo*) au sens d'une lexie de la langue en question, c'est-à-dire à une acceptation bien déterminée d'un mot ou d'un phrasème. Ce sens doit être identifié par un numéro distinctif qui accompagne l'étiquette lexicale et qui provient d'un dictionnaire, en principe, notre *Dictionnaire explicatif et combinatoire*; avant que le DEC ne soit suffisamment avancé, ces numéros peuvent être pris dans un bon dictionnaire de la langue, par exemple, pour le français, dans le *Petit Robert*. (Ainsi, 'façon II' = 'forme d'être ou d'agir particulière', opposé à 'façon I' = 'action de donner une forme particulière à ...' et à 'façon III' = 'manière d'être extérieure d'une personne'; 'habiter I' = 'avoir sa demeure', opposé à 'habiter II' = 'occuper un logis de façon durable', etc.)

Un SÉMANTÈME peut être élémentaire (= un SÈME) ou complexe, c'est-à-dire constitué de sèmes ou de sémantèmes plus simples. Dans ce cas, le sémantème doit être représenté (dans le dictionnaire) par un réseau sémantique qui en spécifie la DÉCOMPOSITION. Ainsi, le sémantème 'avant I.1' [fig. 1, B1] se décompose de la façon suivante :

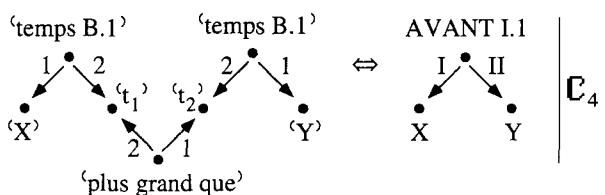
(28) a.



Remarquons que (28a) est une règle lexico-sémantique (bien qu'incomplète), étant donné que le sens 'avant I.1' est bel et bien le signifié d'un lexème français : AVANT I.1. Ainsi, la décomposition sémantique d'un sémantème est donnée par le dictionnaire.

La forme canonique de cette règle (cf. la règle (18) à la p. 24) est comme suit :

(28) b.



La décomposition sémantique peut se poursuivre jusqu'au niveau nécessaire ou voulu — mais pas plus loin que les sèmes, qui, eux, sont indécomposables (= non représentables en termes d'autres mots de la même langue). Pour le moment, nous n'avons pas de liste définitive de sèmes — ÉLÉMENTS SÉMANTIQUES PRIMITIFS; ils doivent apparaître comme résultat de notre recherche lexicographique, qui vise à la décomposition successive de tous les lexèmes du français. Nous pouvons, cependant, supposer que les éléments suivants sont de bons candidats au statut de sèmes : 'quelque chose', 'plus grand que', 'ne ... pas' (= 'non'), 'et', 'ou', 'ensemble' (au sens mathématique), 'espace', 'temps', 'dire', 'cet acte de parole'.¹⁰

Puisqu'une décomposition complète en sèmes rendrait illisible la SSém, nous cherchons à y utiliser des sémantèmes complexes. Autrement dit, de façon générale, la RSém d'un énoncé donné doit être la moins profonde possible. *Grosso modo*, la profondeur minimale nécessaire est fonction du nombre de paraphrases dans la famille dont nous voulons représenter le sens : plus nombreuses sont les paraphrases dans la famille de départ, plus profonde devra être sa RSém. Bien entendu, chaque RSém peut toujours subir des décompositions ultérieures (= les nœuds de la SSém sont remplacés par leurs réseaux correspondants), jusqu'au niveau des sèmes.¹¹ Une conséquence importante de ce principe est que la RSém n'est pas canonique : deux phrases synonymes d'une même langue

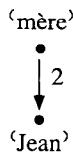
peuvent obtenir deux RSém non identiques, bien qu'équivalentes (= elles peuvent être réduites à une même RSém si nous poussons la décomposition suffisamment loin). Il n'est pas *a priori* exclu que la RSém ne soit pas non plus universelle dans le sens suivant : nous admettons que deux phrases synonymes de deux langues différentes aient deux RSém non équivalentes (même si elles sont très proches). Cela peut se produire parce que nos unités sémantiques — y compris les sèmes — sont déterminées par le stock lexical (ou, de façon plus générale, par le stock de signes) disponible dans \mathbf{L} ; les langues peuvent différer à cet égard. Mais nous ignorons si c'est vraiment le cas.

Nous distinguons deux grandes classes de sémantèmes :

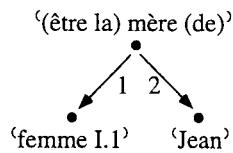
1. Les FONCTEURS, qui se subdivisent en trois sous-classes inégales :
 - PRÉDICATS (= événements, actions, états, propriétés, relations, etc.), dont le nombre est de quelques centaines de milliers (c'est une grande partie de l'ensemble des significations lexicales, voir p. 24);
 - QUANTIFICATEURS ('pour tout X', 'il existe X tel que', tous les numéraux);
 - CONNECTEURS LOGIQUES ('et', 'ou', 'si', 'non', etc.).
2. Les NOMS (DE CLASSES) D'OBJETS, y compris les noms propres; ceux-ci sont aussi très nombreux.

Dans la SSém, des variables sont également utilisées : voir, par exemple, α et β dans la figure 1 [A2 et B1].

Bien que la distinction logique entre foncteurs et noms soit en principe très nette, nous nous permettons les abréviations du type suivant, qui l'estompent quelque peu : pour représenter *la mère de Jean*, nous écrivons



au lieu de l'écriture plus rigoureuse :



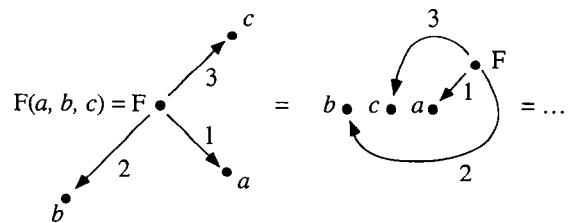
(= 'la femme qui est la mère de Jean'), en faisant «l'incorporation» du premier actant dans le prédictat. De cette façon, nous avons *la mère de X* = 'la femme qui est la mère de X'. Par conséquent, les sémantèmes tels que 'mère', 'gouvernement', 'ministre', 'police', etc. sont traités, dans nos réseaux sémantiques, comme des foncteurs n'ayant que le deuxième actant séman-

tique (avec le premier actant «incorporé»); on peut les appeler, par abus de langage, *semi-foncteurs*.

Le formalisme de la SSém (dans la théorie Sens-Texte) est fondé sur la notation du calcul de prédictats. Pour représenter, au niveau sémantique, la phrase *La police soviétique persécuté les gens de lettres*, on pose deux noms : 'U.R.S.S.' (= *soviétique*) et 'gens de lettres', ainsi que quatre foncteurs : 'police' (de quel pays?), 'persécuter 1' (qui persécuté qui?), 'identique', et 'être le temps B.1 de'; on y ajoute, comme point de référence nécessaire, 'cet acte de parole'. L'écriture selon la notation du calcul de prédictats est alors (le symbole \wedge désignant la conjonction logique) :

'police' ('U.R.S.S.') \wedge
 'persécuter 1' ('police'; 'gens de lettres') \wedge
 'temps B.1' ('persécution 1'; 't₁') \wedge
 'temps B.1' ('cet acte de parole'; 't₂') \wedge
 'identique' ('t₁', 't₂').

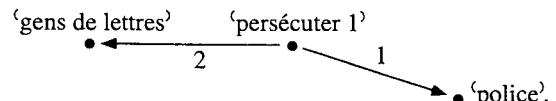
C'est pour éviter la linéarité de cette notation et tous les problèmes théoriques et techniques qu'elle entraîne (par exemple, l'indication de la coréférence), que nous avons recours au moyen graphique du réseau : les nœuds ne sont pas linéairement ordonnés, et les liens entre un foncteur F et ses arguments a, b, c, \dots sont signalés par des arcs numérotés dont la flèche est orientée à partir de F vers son argument; cf. :



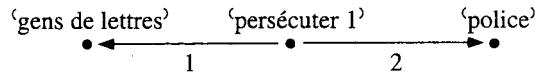
La disposition spatiale des nœuds et des arcs n'a donc aucune pertinence.

On notera que les arcs peuvent aboutir aux noms aussi bien qu'aux foncteurs : tout sémantème, indépendamment de son type, peut recevoir une flèche. Cependant, un arc ne peut partir que d'un foncteur : seulement un sémantème foncteur peut être à l'origine d'une flèche.

Un ARC d'une SSém est étiqueté par un numéro qui n'a aucune signification en soi mais qui sert à identifier l'argument du foncteur en cause. Quand on écrit



1 signifie seulement que ‘police’ est le premier argument de ‘persécuter 1’; et 2 signifie que ‘gens de lettres’ en est le deuxième argument. L’écriture



signifierait ‘Les gens de lettres persécutent la police’.

Le rôle de chaque argument par rapport à son foncteur doit être fixé dans l’article de dictionnaire de ce dernier; il est spécifié par la décomposition sémantique (= définition lexico-graphique) du foncteur. Par exemple, (18) à la p. 24 précise que le premier argument de ‘persécuter 1’ [= X], qui est le signifié du lexème PERSÉCUTER 1, tourmente usuellement le deuxième [= Y] au moyen de l’injustice du traitement que X inflige à Y. En décomposant ‘tourmenter’, on obtient que X ‘cause que Y subisse des souffrances physiques ou morales’; en décomposant ‘souffrance’, on montre que Y ‘éprouve’ des états spéciaux de son organisme; et ainsi de suite, jusqu’à ce qu’on aboutisse aux sèmes (= éléments sémantiques primaires), pour lesquels le rôle de leurs arguments est établi par définition.

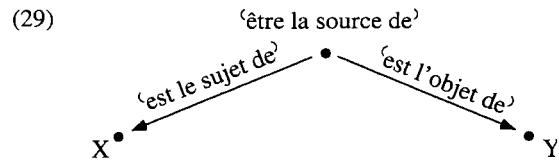
Les arguments d’un foncteur s’appellent ses ACTANTS SÉMANTIQUES : X et Y dans (18), p. 24, sont les actants sémantiques de ‘persécuter 1’. (Les expressions françaises, en l’occurrence les syntagmes nominaux, qui expriment les actants sémantiques dans une phrase auprès du lexème signifiant le foncteur en cause, sont ses ACTANTS SYNTAXIQUES : *La police* et *les gens de lettres* par rapport au verbe *persécuter* dans *La police persécuté les gens de lettres* ou par rapport au nom *persécution* dans *les persécutions policières* (= *par la police*) des *gens de lettres*.)

Le caractère purement distinctif (= asémantique) des arcs constitue une particularité importante de la SSém du MST — par rapport aux réseaux sémantiques employés dans de nombreuses approches courantes en linguistique et surtout en intelligence artificielle. Là, on étiquette les arcs en y mettant les noms de relations reliant le foncteur à ses arguments : ‘être l’agent de’, ‘être la source de’, ‘être le moyen de’, etc. Nous avons rejeté une telle démarche pour deux raisons principales.

D’une part, nous exigeons un caractère strictement homogène de présentation d’informations sémantiques : un même nom de relation ne doit pas apparaître tantôt sur des nœuds, tantôt sur des arcs.

D’autre part, nous voulons éviter la régression infinie introduite fatalement par le fait d’étiqueter TOUS les arcs. Soit, par exemple, X • ‘est la source de’ • Y; il faut alors étiqueter les

relations entre le foncteur ‘est la source de’ et ses arguments X et Y, et on peut le faire, par exemple, comme cela :



Immédiatement, on se voit obligé d’étiqueter aussi les relations entre ‘X’ et ‘être le sujet de’, entre ‘être la source de’ et ‘être le sujet de’, et ainsi de suite, indéfiniment. La seule façon de s’arrêter de tourner en rond, c’est de déclarer arbitrairement que certaines relations (données par une liste) ne seront pas étiquetées. Nous préférons éviter l’arbitraire, en nous appuyant surtout sur les décompositions sémantiques et en ne permettant pas d’étiquetage «sémantique» des arcs dans la SSém.

Pour compléter cet aperçu sur la RSém dans la théorie Sens-Texte, il convient de souligner que les mêmes principes, techniques et formalismes qui sont utilisés pour représenter le sens de phrases (ou de textes plus longs), le sont aussi pour représenter le sens de lexies. Autrement dit, toutes les définitions dans le lexique du MST, c’est-à-dire dans le DEC, sont des RSém de la même nature et du même type que celle présentée dans la section IV. Cela garantit l’applicabilité des règles lexico-sémantiques (= des articles de dictionnaire du DEC) à des RSém de départ, en vue de la production d’ensembles paraphrastiques.¹²

Une RSém — d’une phrase ou d’une lexie — est écrite, suivant notre approche, dans un langage formel : la (méta)langue sémantique dont il était question plus haut, pp. 14 et 27 ssq. La particularité centrale de ce langage est la suivante : sa syntaxe est universelle (celle des réseaux étiquetés), contrairement à son lexique, qui est national; c’est quasiment le lexique de la langue décrite, avec, cependant, une distinction capitale : ce dernier est un lexique désambiguisé, dans lequel les unités sont des sémantèmes (= des sens bien déterminés) et non des mots naturels. Ce lexique doit être spécifié par un dictionnaire, en l’occurrence, un DEC. Notre (méta)langue sémantique pour une langue naturelle *L* est *L* elle-même débarrassée de tous les «caprices» de sa syntaxe et de sa morphologie et complètement désambiguisée. De cette façon, chaque langue naturelle a sa propre (méta)langue sémantique qui partage sa syntaxe avec les autres (méta)langues sémantiques, mais qui possède un lexique particulier. Du reste, la question de la relation entre une langue naturelle et sa (méta)langue sémantique, ainsi que celle de l’universalité de la (des) langue(s) sémantique(s), est trop complexe et, en même temps, trop spécialisée pour être traitée ici. Notons seulement que, pour le moment, il ne s’agit nullement d’une métalangue sémantique complète. Nous n’avons, à ce jour, que les lignes directrices

pour sa construction et seulement un fragment plutôt modeste : les sémantèmes élaborés dans les deux volumes publiés auparavant ainsi que dans le présent volume du DEC du français contemporain. Au bout du compte, compléter notre mé-

talangue sémantique pour le français revient à compléter le DEC du français — plus, évidemment, élaborer les descriptions sémantiques pour les grammèmes du français, ainsi que pour les structures syntaxiques pertinentes.

V. SYSTÈME DE PARAPHRASAGE DANS LE MODÈLE SENS-TEXTE

Les paraphrases des langues naturelles sont de deux types majeurs :

– Paraphrases SÉMANTIQUES produites par la distribution différente du sens de départ entre lexies; deux paraphrases sémantiques se distinguent obligatoirement par au moins une lexie pleine. Exemple typique :

(30) *Il s'est emparé d'un territoire immense.*

vs

Un territoire immense est tombé sous son pouvoir.

– Paraphrases SYNTAXIQUES produites par le choix de constructions syntaxiques différentes; deux paraphrases syntaxiques «pures» affichent le même jeu de lexies pleines, qui sont pourtant arrangées différemment du point de vue syntaxique. Exemple typique :

(31) ... [la somme] récupérée par le ministre

vs

...[la somme] que le ministre a récupérée.

Les paraphrases syntaxiques sont (en partie) décris par les transformations syntaxiques, qui ont été tellement en vogue dans les années 1960-1970. Malgré leur intérêt incontestable pour la théorie de la paraphrase, nous ne nous occuperons pas des paraphrases syntaxiques et nous nous concentrerons entièrement sur les paraphrases sémantiques.

Ces dernières ne sont pas, cependant, homogènes; elles se laissent classer, comme nous l'avons déjà indiqué dans la section III, pp. 23-24, en deux sous-classes : celles qui sont effectuées par des règles sémantiques au sens propre et celles qui le sont par des règles sémantiques appelées «règles de paraphrase». Les paraphrases de la première sous-classe sont engendrées par des CHOIX LEXICAUX différents : en principe, la RSém de départ peut être articulée en sous-réseaux correspondant aux lexèmes de plusieurs façons (l'exemple (30) ci-dessus en est un spécimen). De telles paraphrases sont plus profondes que les autres; elles sont extrêmement intéressantes pour la recherche sémantique mais pour le moment, elles ont été mal étudiées et semblent (peut-être justement à cause de cela) peu systématiques. Par contre, les paraphrases de la seconde sous-classe sont engendrées par différentes SUBSTITUTIONS LEXICALES : elles manifestent les liens sémantiques existant entre les lexies dans la langue. Les paraphrases de ce type, qu'on pourrait appeler paraphrases lexicales, se sont avé-

rées à la fois très riches et très systématiques; (32) en est une bonne illustration.

(32) *La police persécute brutalement les gens de lettres.*
vs
Les gens de lettres subissent des persécutions brutales de la part de la police.

Ce sont ces dernières paraphrases qui sont traitées par le système de paraphrasage faisant partie de la composante sémantique du modèle Sens-Texte. Dans ce qui suit nous ne nous occuperons que de celles-ci.

La description des paraphrases lexicales et l'élaboration du système de paraphrasage ont nécessité la création d'un nouveau concept linguistique — la FONCTION LEXICALE — et le développement de l'appareil formel correspondant, que nous allons maintenant caractériser.

1. Fonctions lexicales

Une fonction lexicale (FL) est une dépendance, ou correspondance, f qui associe à une unité lexicale, ou lexie (DEC 2 : 28), L , appelée l'ARGUMENT de f , un ensemble d'unités lexicales, ou lexies, $f(L)$ — la VALEUR de f . Chaque FL correspond à un sens très général (qui peut, à la limite, être vide) et à un rôle syntaxique profond. Plus précisément, pour qu'une telle dépendance lexicale soit une fonction lexicale, les deux conditions suivantes doivent être vérifiées simultanément :

Quelles que soient les lexies L_1 et L_2 d'une langue L , si $f(L_1)$ et $f(L_2)$ existent toutes les deux, alors :

1. $f(L_1)$ et $f(L_2)$ entretiennent respectivement une même relation avec L_1 et L_2 en ce qui concerne le sens et le rôle syntaxique profond, ce que l'on note :

$$\frac{f(L_1)}{L_1} = \frac{f(L_2)}{L_2}$$

[Cette condition ne dépend pas de la langue L .]

2. Au moins, pour certaines lexies arguments, $f(L_1) \neq f(L_2)$.

[Cette condition dépend uniquement de la langue L : elle signifie qu'en L , la valeur de f est phraséologiquement liée par son argument.]

L'importance de la condition 2 peut être illustrée par l'exemple suivant. Le sens 'cher II.1' [= 'qui est d'un prix élevé'] ne correspond pas à une fonction lexicale en français, puisque avec n'importe quelle lexie, il peut toujours être exprimé par le lexème CHER II.1 : son expression ne dépend pas du lexème modifié. Par contre, le sens 'très' est une FL en français : 'très'(malade) = très, gravement, mais 'très'(blessé) = grièvement; 'très'(grippe) = carabinée, mais 'très'(prix) = haut, élevé, ...; 'très'(lutter) = sans relâche, à corps perdu, mais 'très'(battre) = à plate couture, etc. Comme on le voit, l'expression de 'très' [= de l'idée d'intensité] dépend du lexème modifié. Ce sens correspond, en effet, à la FL **Magn**, voir ci-dessous.

L'argument d'une fonction lexicale (*malade*, *blessé*, *grippe*, *prix*, etc. par rapport à 'très') est aussi appelé MOT CLÉ ou LEXÈME CLÉ. On évitera ainsi l'homonymie fâcheuse du terme *argument* (d'une FL vs d'un prédicat sémantique) dans certains contextes.

Parmi les FL, il convient de distinguer une sous-classe importante : les FL STANDARD SIMPLES. Ces FL vérifient deux conditions supplémentaires :

3. **f** est définie pour un grand nombre d'arguments, c'est-à-dire que son domaine est grand. (Autrement dit, **f** a une vaste cooccurrence sémantique : le sens '**f**' est suffisamment abstrait et général pour être compatible avec beaucoup d'autres sens.)
[Cette condition ne dépend pas de la langue.]
4. **f** possède un grand nombre de valeurs différentes. (Autrement dit, l'ensemble de toutes les valeurs **f(L_i)**, pour un grand nombre de **L_i**, est suffisamment riche.)
[Cette condition dépend uniquement de la langue.]

Prenons un exemple. Le sens 'additionné de ...' (en particulier, de rien) est une FL en français, puisqu'on dit *café crème* mais *fraises à la crème* (**café à la crème*, **fraises crème*), *café au lait* (**café lait*; cf., cependant, it. *caffè e latte*, litt. 'café et lait', ou *caffelatte*), *café arrosé* (it. *caffè coretto*), *café noir*, *thé nature*, etc. Cependant, ce n'est pas une FL standard simple : les conditions 3 et 4 ne sont pas respectées, étant donné que le sens en cause est très concret (par conséquent, il ne possède qu'un petit nombre d'arguments : noms de certains plats et boissons) et qu'il n'a que très peu d'expressions (*crème* en apposition, *à la crème*, *avec de la crème*, *noir*, *arrosé*, ...). C'est une FL non standard, dont il ne sera pas question ici.

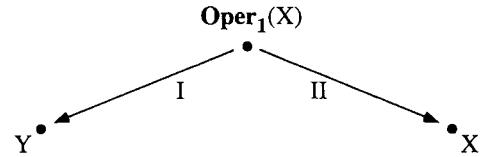
Voici deux exemples de fonctions lexicales standard simples.

- La FL **Magn** (les noms de FL viennent toujours du latin; dans ce cas, de *magnus* 'grand'), qui est un intensificateur :

Magn(riposte)	= <i>foudroyante</i>	Magn(sourd) = <i>comme un pot</i>
Magn(cri)	= <i>d'orfraie</i>	Magn(fort) = <i>comme un Turc</i>
Magn(applau- dissements)	= <i>nourris,</i> <i>frénétiques</i>	Magn(soûl) = <i>comme un Polo-</i> <i>nais, comme une grive (une bouri-</i> <i>rière)</i>

Magn(apprécier)	= <i>grandement</i>
Magn(ivre)	= <i>mort</i>
Magn(recourir [à qqch.])	= <i>largement</i>
Magn(dormir)	= <i>profondément, comme un loir,</i> <i>comme une bûche, à poings fermés</i>
Magn(surveiller)	= <i>étroitement, de près</i>

– La FL **Oper₁** (lat. *operari* 'travailler'), qui est un verbe sémantiquement vide (ou vidé dans le contexte de son mot clé) et qui prend le mot clé comme son complément d'objet direct (CO^{dir}) ou comme son complément principal (si le verbe n'est pas transitif) et le 1^{er} actant SyntP du mot clé, comme son sujet grammatical (SG) :



où Y est le 1^{er} actant SyntP de X. Comparez les exemples suivants :

Oper₁(plainte)	= <i>porter [~]</i>	Oper₁(pas)	= <i>faire [un ~]</i>
Oper₁(cri)	= <i>pousser [un ~]</i>	Oper₁(ordre)	= <i>donner [un ~]</i>
Oper₁(grippe)	= <i>avoir [une ~]</i>	Oper₁(supré- matie)	= <i>détenir, exer-</i> <i>cer [une ~]</i>
Oper₁(déses- poir)	= <i>être [au ~]</i>	Oper₁(regard)	= <i>jeter [un ~]</i>
Oper₁(geste)	= <i>faire [un ~]</i>	Oper₁(précau- tion)	= <i>prendre [des ~s]</i>
Oper₁(santé)	= <i>jouir [d'une</i> <i>(bonne) ~]</i>	Oper₁(attaque)	= <i>mener [une ~]</i>

Oper₁ représente une famille de verbes qui ont reçu le nom de VERBES OPÉRATEURS ou SUPPORTS dans les travaux de M. Gross et de son équipe (Giry-Schneider 1978 et M. Gross 1981, où l'on trouve d'autres références).¹³

Une FL n'est aucunement une unité sémantique bien précise (et encore moins, un élément sémantique primitif). D'une part, il y a des FL (comme **Oper₁**) qui sont sémantiquement vides ou quasi vides; d'autre part, les lexies comprises dans la valeur d'une FL donnée pour un mot clé donné ne sont pas nécessairement des synonymes exacts : ainsi, 'comme une bûche' ≠ 'à poings fermés' auprès de *dormir*, et pourtant les deux expressions font partie de la valeur de **Magn(dormir)**. Il suffit que toutes les lexies qui constituent la valeur de la FL f

pour un mot clé donné partagent une composante sémantique non triviale suffisamment importante, d'habitude de caractère assez abstrait; elles peuvent sémantiquement différer dans beaucoup de nuances.

Les FL sont introduites comme éléments du lexique profond de **L**, et elles n'apparaissent que dans la SSyntP (voir la section III, p. 20). Leur vocation est double :

1) Elles servent à décrire la cooccurrence lexicale restreinte : phénomène voulant qu'un certain sens (ou un rôle sémantico-syntaxique) s'exprime de façon phraséologisée selon la lexie au sens de laquelle le sens en question se rattache.

2) Elles servent à décrire la synonymie entre phrases, reposant sur des relations sémantiques entre lexies, c'est-à-dire les paraphrases lexicales. C'est cette capacité des FL qui nous intéresse surtout ici.

Nous n'entrerons pas dans les détails concernant les FL (voir Mel'čuk 1982; Mel'čuk *et al.* 1984 : 6-13 et 1988 : 91-94), ni ne présenterons une liste des FL (puisque une telle liste est donnée aux pp. 127 *ssq.* de ce volume). Nous nous limiterons à signaler deux propriétés importantes des fonctions lexicales qui n'ont pas été suffisamment discutées auparavant.

Première propriété des FL. Tout d'abord, il convient de signaler le caractère universel des FL : les mêmes FL sont valables pour toutes les langues. Qu'il nous soit permis d'illustrer les deux FL mentionnées (et quelques autres) à partir de langues très variées (les gloses françaises sont partout littérales) :

Anglais

Magn(*rain* 'pluie') = *heavy* 'lourd'

Magn(*argument* 'argument') = *strong* 'fort', *weighty* 'pesant'

Magn(*applause* 'applaudissements') = *thunderous* 'de tonnerre', *deafening* 'assourdisant', *boisterous* 'bruyants', *whirl-wind* 'tourbillonnants', ...

Oper₁(*trip* 'voyage') = *take* [*a ~*] 'prendre'

Oper₁(*deal* 'accord, transaction') = *strike* [*a ~*] 'frapper'

Oper₁(*apologies* 'excuses') = *offer* [*~*] 'offrir'

Russe

Magn(*dožd'* 'pluie') = *sil'nyj* 'fort', *prolivnoj* 'd'averse'

Magn(*dovod* 'argument') = *veskij* 'pesant', *ubeditel'nyj* 'convaincant'

Magn(*applodismeny* 'applaudissements') = *burnye* 'tempétueux'

Oper₁(*putešestvie* 'voyage') = *sovershit'* [*~e*] 'accomplir'

Oper₁(*soglašenie* 'accord') = *zaključit'* [*~e*] 'contracter', *priđti* [*k ~ju*] 'arriver à'

Oper₁(*izvinenija* 'excuses') = *prinesti* [*~ja*] 'apporter'

Allemand (Reuther 1978, 1980)

Magn(*Regen* 'pluie') = *starker* 'fort', *Platz-* 'à éclat'

Magn(*Frost* 'froid') = *klirrender* 'cliquetant'

Magn(*Schweigen* 'silence') = *tiefes* 'profond'

Oper₁(*Scham* 'honte') = [*~*] *empfinden* 'percevoir'

Oper₁(*Macht* 'pouvoir') = [*an der ~*] *sein* 'être à'

Oper₁(*Verzweiflung* 'désespoir') = [*in ~*] *sein* 'être au'

Hongrois

Magn(*eső* 'pluie') = *zuhogó* 'torrentiel'

Magn(*érv* 'argument') = *komoly* 'sérieux', ...

Magn(*taps* 'applaudissements') = *viharos* 'tourbillonnant', *vas-* 'de fer'

Oper₁(*utazás* 'voyage') = [*~t*] *tenni* 'faire'

Oper₁(*lépés* 'pas') = [*~t*] *tenni* 'faire'

Oper₁(*hatalom* 'pouvoir') = [*~t*] *birtokolni* 'posséder'

Polonais

Janus (1971)

Magn(*klęska* 'débâcle') = *ciężka* 'lourd', *wielka* 'grand', ..., *druzgocąca* 'écrasant'

Magn(*zwycięstwo* 'victoire') = *duże* 'gros', *ogromne* 'énorme', *walne* 'décisif'

Oper₁(*klęska*) = *ponieść* [*~ę*] 'porter'

Oper₁(*zwycięstwo*) = *odnosić* 'emmener', *osiągnąć* [*~o*] 's'emparer de'

Somalien

(Žolkovskij 1970, 1971 : 225-261)

Oper₁(*birmad* 'attaque') = *kiñi* 'lever', *dhufan* 'frapper'

Oper₁(*rajoda* 'espoir') = *qabi* 'tenir'

Oper₁(*dagaal* 'lutte') = *jiri* 'se trouver dans'

Oper₁(*jawabta* 'réponse') = *felin* 'tourner'

Oper₁(*firo* 'attention') = *lahaan* 'avoir', *yeelan* 'faire'

Oper₁(*moqif* 'position') = *taagan* 'être debout dans'

Albanais

(Hendriks 1980)

Oper₁(*besim* 'confiance') = *ka* 'avoir'

Oper₁(*be* 'serment') = *bën* 'faire'

IncepOper₁(*bela* 'problèmes') = *bie* [*në ~*] 'tomber dans'

Oper₂(*qotek* 'rossée') = *ha* 'manger'

Real₁(*borxhin* 'dette') = *bën* 'faire' [= 'payer une dette' et non *'faire une dette']

Persan

(Sheintuch 1976)

Oper₁(*kotak* 'rossée') = *zadan* 'frapper'

Oper₁(*gālabe* 'victoire') = *kardan* 'faire'

Oper₁(*soāl* 'question') = *kardan* 'faire'

- Oper₁**(*qose* ‘ennui, tracas’) = *xordan* ‘manger’
Oper₁(*fahm* ‘compréhension’) = *dāštan* ‘avoir’
Oper₁(*kalame* ‘parole(s)’) = *harf zadan* ‘parler’
Real₁(*Jaru* ‘balai’) = *kardan* ‘faire’ [= ‘balayer’]
Real₁(*češm* ‘œil’) = *duxtan* ‘coudre’ [= ‘observer’]

En persan, la plupart des sens verbaux (à peu près 90 %) sont exprimés non pas par des lexèmes individuels, mais par des expressions bilinguiques du type illustré, comprenant des fonctions lexicales comme **Oper** et **Real**. On voit que dans cette langue, les FL occupent une place vraiment importante.

La liste des fonctions lexicales standard établies jusqu'à présent semble être exhaustive en ce sens que si un fait de la cooccurrence lexicale restreinte d'une langue peut être décrit par une FL qui mérite le statut de fonction standard, f se trouve déjà parmi les FL énumérées. Bien sûr, la question de savoir si nos FL standard sont suffisantes est une question purement empirique; cependant, il semble peu probable qu'on en découvre un grand nombre de nouvelles.

Deuxième propriété des FL. Ensuite, nous aimerions souligner qu'en principe, chaque lexie apparaissant comme valeur d'une FL d'un mot clé donné doit être munie d'un régime, de restrictions portant sur ses actants (le cas échéant) et d'autres informations normalement attachées aux entrées lexicales. Tout cela est nécessaire parce que le comportement d'une lexie en tant que valeur d'une FL peut être très différent de celui qu'elle a dans ses emplois non lexico-fonctionnels. Citons quelques exemples (empruntés à Giry-Schneider 1978, livre qui contient beaucoup de données précieuses sur l'emploi du verbe support FAIRE).

Dans un premier temps, on notera la propriété des expressions lexico-fonctionnelles suivante : elles sont (plus ou moins) phraséologisées, et la phraséologisation se traduit en français assez souvent par l'emploi particulier du déterminant avec le nom faisant partie du phraséologisme. On dit, notamment, *faire partie* (**faire* {une} partie); *porter plainte* (**porter* {une} plainte), mais *faire une plainte* (**faire* plainte); *pousser* {un} cri (**pousser cri*), etc. Par conséquent, dans le régime assigné aux valeurs d'une fonction lexicale, il faut spécifier les déterminants appropriés, ce qui n'est pas normalement le cas avec les lexies pleines (où les déterminants apparaissent selon le sens et suivant les règles générales de la grammaire). Parfois même la distribution des déterminants dans les expressions lexico-fonctionnelles peut être liée au sens. En voici un exemple : **Oper₁**(*offense*) = *faire*; mais selon

la présence ou l'absence du déterminant, la construction aura un sens volontaire ou non. Comparez :

- (33) $\left\{ \begin{array}{l} \text{Jean} \\ \text{Cette histoire} \\ \text{Cela} \end{array} \right\}$ *a fait offense à Paul.*
 vs
 $\left\{ \begin{array}{l} \text{Jean} \\ \text{*Cette histoire} \\ \text{*Cela} \end{array} \right\}$ *a fait une offense à Paul.*
 ou
 $\left\{ \begin{array}{l} \text{Jean} \\ \text{*Cette histoire} \\ \text{*Cela} \end{array} \right\}$ *a fait à Paul l'offense de le caricaturer.*

Par conséquent, nous écrirons sous **Oper₁** (dans l'article de dictionnaire de OFFENSE) :

faire [~]; *faire* [ART ~] | M₁ désigne une personne

Dans l'article de NEZ I.2 (DEC 2 : p. 239, au sens de ‘organe de l'odorat’), la description de la FL **Oper₁** = *avoir* doit contenir l'indication suivante : le lexème *nez* prend l'article indéfini ou reste sans article s'il est précédé de *bon*, mais il prendra l'article défini, s'il est suivi de *fin*, cf. : *Jean a (un) bon nez* (**Jean a le bon nez*), mais *Jean a le nez fin* (**Jean a (un) nez fin*).

Assez souvent, on doit aussi indiquer si tel ou tel verbe lexico-fonctionnel admet/n'admet pas le passif, la relativisation, etc. (cela est également une conséquence du caractère phraséologique des expressions en cause) :

- (34) a. *analyser* = *faire l'analyse*, et on peut dire *L'analyse des données a été faite par Jean*, *l'analyse des données que Jean a faite*, ...
 vs
 b. *ruiner* = *faire la ruine*, mais on ne peut pas dire **La ruine de la famille a été faite par Jean*, **la ruine de la famille que Jean a faite*, ...

Par conséquent, il est nécessaire d'indiquer, pour **Oper₁**(*ruine*) = *faire*, qu'il n'accepte pas le passif et la relativisation.

Notons encore la particularité suivante : certaines constructions à **Oper₁** admettent la forme passive du verbe mais n'acceptent le complément d'agent qu'avec une condition supplémentaire (par exemple, la présence d'un déterminant auprès du nom). Comparez :

- (35) *Jean a fait allusion à cette histoire* \Leftrightarrow *Allusion a été faite (*par Jean) à cette histoire* \Rightarrow *Une allusion a été faite (par Jean) à cette histoire*.

Enfin, certains **Oper₁** qui ont, dans leur régime, le syntagme «à N» admettent que ce syntagme soit pronominalisé par Y, alors que d'autres ne l'admettent pas :

- (36) *Jean prête attention à ce problème* ⇒ *Jean y prête attention*,
 vs
Jean fait attention à ce problème ≠ **Jean y fait attention*.

Il faut alors écrire (dans l'article de ATTENTION) :

Oper₁ : faire [~ à N/*y]; prêter [~ à N/y].

Ces trois exemples semblent suffire pour montrer la complexité éventuelle de description de chaque lexie faisant partie de la valeur d'une FL. En exagérant légèrement, on pourrait dire qu'une telle description constitue un mini-article de dictionnaire au sein d'un article de dictionnaire.

Comme nous l'avons signalé, les FL sous-tendent le système de paraphrasage, propos principal de cet article. Après *avoir passé* [= **Labor₁₂(revue)**] en revue les FL, nous pouvons *concentrer* [= **Caus₁Func₂(attention)**] notre attention sur le système de paraphrasage à proprement parler.

2. Système de paraphrasage¹⁴

Le système de paraphrasage se compose de règles de deux types : les règles lexicales (ci-dessous) et les règles syntaxiques (p. 45).

Les RÈGLES LEXICALES DE PARAPHRASAGE représentent soit des (quasi-)équivalences sémantiques, soit des implications sémantiques formulées en termes de FL. Nous connaissons une cinquantaine de règles lexicales de paraphrasage, valables pour toutes les langues.

La liste proposée de ces règles est fondée sur l'ensemble original des règles lexicales de paraphrasage présenté dans Žolkovskij et Mel'čuk 1967. Nous l'avons augmenté et modifié, en nous aidant de Mel'čuk 1974 : 149-161, d'Apresjan 1974 : 324-335 et de Žolkovskij 1977. La présente liste est organisée de la façon suivante :

- D'abord, nous présentons les équivalences sémantiques (les règles 1-47) : section A; ensuite, nous donnons les implications sémantiques (les règles 48-53) : section B.

- Dans la section A, nous présentons d'abord les règles du type $C_0 \Leftrightarrow \text{FL}(C_0)$, c'est-à-dire les règles qui remplacent une lexie «indépendante» par sa (ou ses) FL et *vice versa* : sous-section A.I; ensuite, nous mettons les règles du type $\text{FL}_1(C_0) \Leftrightarrow \text{FL}_2(C_0)$, c'est-à-dire les règles qui établissent les correspondances entre les fonctions lexicales d'une même lexie : sous-section A.II.¹⁵

- Dans les sous-sections A.I et A.II, nous plaçons d'abord les règles du type $X \Leftrightarrow Y$, soit des «substitutions», ensuite, les règles du type $X \Leftrightarrow Y + Z$, soit des «fissions».

- Quant à la section B, elle n'a pas de subdivisions.

La forme d'une règle lexicale de paraphrasage est évidente et n'a pas besoin d'explications générales. Une règle lexicale établit que tels nœuds lexicaux de la SSyntP donnée (un ou deux nœuds) peuvent être remplacés *salva significatione* par tels autres nœuds (là aussi un ou deux), les derniers étant les FL des premiers ou bien les premiers et les derniers étant les FL du même mot clé. Pour chaque règle lexicale, nous spécifions la suite des règles syntaxiques de paraphrasage qui la «desservent» : elles effectuent les restructurations syntaxiques imposées par les changements lexicaux. Cette suite est identifiée par la chaîne des numéros placée à la droite de la règle lexicale; le lien entre celle-ci et «ses» règles syntaxiques s'effectue à l'aide de variables (X, Y, ...) qui sont associées aux composantes de la règle lexicale et qui figurent dans les règles syntaxiques correspondantes. L'application des règles syntaxiques de paraphrasage sera expliquée au moment opportun, un peu plus loin.

Quelques commentaires concernant les détails techniques seront ajoutés à la fin de la liste des règles lexicales (pp. 43 ssq.); les références à ces commentaires sont placées aux endroits pertinents — après chaque règle particulière que le commentaire en question intéresse : ce sont les numéros encerclés (notés ①, ②, ...). Ici, nous discuterons de trois particularités générales des règles lexicales de paraphrasage sous leur forme actuelle :

- l'absence d'exactitude sémantique absolue;
- l'absence de conditions d'application;
- l'incomplétude et la redondance de l'ensemble des règles.

1. Aucune règle lexicale ne garantit, à elle seule, l'exactitude sémantique des paraphrases qu'elle décrit. Elle exprime une équivalence au niveau abstrait, idéal, où il est vrai, par exemple, qu'un verbe signifie la même chose que son nom déverbal plus le verbe support approprié (tel que *voyager* = *faire un (des) voyage(s)*). Mais au niveau concret, sur le plan de la réalisation langagière où des lexèmes français spécifiques apparaissent à la place des FL, cette équivalence peut subir des entorses, puisque les valeurs réelles des FL s'écartent souvent de leur signification «idéale». Ainsi, *en larmes* ≈ **A₁(pleurer)**, en sorte que *Elle pleurait* ≈ *Elle était en larmes*; or cette égalité n'est qu'approximative, si bien que, dans de nombreux contextes, elle ne peut pas se réaliser, ainsi :

- (37) *Elle pleurait bruyamment.*
 vs
 **Elle était en larmes bruyamment.*

La raison en est évidente : *en larmes* veut dire ‘ayant le visage baigné de larmes comme résultat d’avoir pleuré’; *en larmes* = $A_{1C}(pleurer)$, où le symbole «_C» [= sens plus étroit, c'est-à-dire plus spécifique] indique explicitement un écart sémantique.

De façon similaire, *en flammes* = $A_{1C}(brûler)$ ‘se consumer par le feu’); cf. :

- (38) a. *La maison brûlait* = *La maison était en flammes*.
 b. *La cigarette brûlait*.
 vs
**La cigarette était en flammes*.
 c. *Tout le charbon qui leur restait brûlait* (**était en flammes*) *dans la cheminée*.
 vs
Tout le charbon qui leur restait [= leur stock de charbon] *était en flammes*.

En fait, *en flammes* = ‘brûlant intensément en tant qu’acte de destruction’.

Par conséquent, le paraphrasage doit s’effectuer sous LE CONTRÔLE DE LA SÉMANTIQUE, c'est-à-dire en prenant en compte les descriptions sémantiques de toutes les lexies qui participent au processus. Quand nous paraphrasons *Elle pleurait* par *Elle était en larmes* à l'aide d'une règle lexicale de paraphrase (n° 24, une fission à verbe support, voir la page 40 : $V \Leftrightarrow A_1(V) + Oper_1(A_1(V))$), nous devons consulter les définitions de PLEURER et de EN LARMES (dans l’article de dictionnaire LARME) et voir les différences sémantiques qu'une telle paraphrase entraîne inévitablement. En nous appuyant sur la connaissance des écarts sémantiques surgiants entre les deux paraphrases, nous sommes en mesure de les évaluer et de décider si ces paraphrases sont admissibles dans le contexte donné.

2. L'applicabilité des règles lexicales de paraphrasage est contrainte par plusieurs conditions syntaxiques (au sens très large de ce terme), — qu'on pourrait appeler des FILTRES —, puisqu'elles trient les paraphrases logiquement possibles pour rejeter celles qui sont linguistiquement interdites. Ainsi, une règle de conversion impliquant le 1^{er} actant ne peut pas être appliquée à un verbe si ce verbe est modifié par un adverbe syncatégorématique (= qui porte sémantiquement sur son sujet grammatical) :

- (39) a. *Jean, très généreusement, a fait cette proposition à Marie.* $\not\Rightarrow$ *Marie, très généreusement, a reçu* [= $Conv_{21}(faire)$] *cette proposition de Jean*.

[Plus précisément, faire = $Oper_1(proposition)$ et recevoir = $Oper_2(proposition)$.]

En fait, le problème réside ici dans la nécessité de contrôler toutes les paraphrases sous l’angle de la S(tructure) Sém(an-

tico)-Comm(unicative) de la phrase, ainsi que des liens logiques entre les phrases et les parties de phrase. Cf., par exemple, (39b) :

- b. *Tout le monde loue la générosité de Jean, qui a fait cette proposition à Marie.* $\not\Rightarrow$ *Tout le monde loue la générosité de Marie, qui a reçu cette proposition de Jean*.

Cependant, dans ce texte, nous caractériserons le système de paraphrasage comme tel, sans toucher aux données qui devraient contraindre les paraphrases possibles.

Une autre illustration : pour qu'une règle lexicale de fission (par exemple, $V \Leftrightarrow S_0(V) + Oper_1(S_0(V))$) puisse être appliquée à un verbe, il faut spécifier le nombre grammatical du nom S_0 qui apparaît dans la partie droite de la règle (*punir* \Leftrightarrow *infliger une punition / des punitions*, ...); ce nombre dépend de la signification lexicale du verbe (surtout de son côté aspectuel) ainsi que de sa forme grammaticale et probablement du contexte. Cependant, les filtres précis des types indiqués ne sont pas encore formulés, de sorte que nos règles peuvent mener à des phrases fautives.

Bien plus, on n'a pas encore exploré suffisamment les problèmes théoriques liés à de tels filtres. Considérons les paraphrases suivantes :

- (40) a. *Ali-Pacha établit son pouvoir sur la région de Yanina.* \Leftrightarrow *Ali-Pacha prit le pouvoir dans la région de Yanina*.

Cependant, si le nom *pouvoir* a une certaine épithète du type **Magn**, par exemple, *dictatorial*, la deuxième paraphrase devient impossible :

- b. *Ali-Pacha établit un pouvoir dictatorial sur la région de Yanina*.
 vs
**Ali-Pacha prit le pouvoir dictatorial dans la région de Yanina*.

La raison en est évidente : le syntagme semi-figé *prendre le pouvoir* [où *prendre* = $Caus_1Func_1(pouvoir)$] n'admet pas de modification adjetivale du nom. C'est un cas plutôt typique : il arrive assez souvent que les valeurs de deux fonctions lexicales pour le même mot clé, lexème L, ne se combinent pas entre elles auprès de L — de telle sorte qu'on a L— f(L) et L— g(L), mais pas *f(L)— L— g(L). Pour le moment, nous ne possédons pas de moyen systématique pour exclure les combinaisons interdites. (Une voie possible : élaborer une échelle pour caractériser le degré auquel une expression donnée inscrite dans un article de dictionnaire en tant que valeur d'une FL est figée; munir, dans le DEC, toutes les expressions figées d'indication de ce degré; formuler des règles bloquant

certaines combinaisons en fonction du degré pour lequel l'expression en question est figée.)

Apresjan (1974 : 335-344) présente un survol des filtres possibles dans un système de paraphrasage. Bien qu'élaborées à partir de données du russe, la plupart des suggestions qu'Apresjan formule valent aussi pour une théorie générale, indépendamment de la langue objet.

3. Notre liste de règles est loin d'être complète. Il reste à entreprendre des recherches approfondies visant à découvrir d'autres équivalences sémantiques exprimables en termes de

fonctions lexicales. De plus, cette liste n'est pas exempte de redondance : on pourrait introduire une écriture plus générale pour certains groupes de règles ou exclure des règles établissant des équivalences qui peuvent être obtenues à l'aide de suites d'autres règles. Cependant, nous ne l'avons pas fait parce que nous croyons qu'il est peut-être encore trop tôt pour faire le ménage à l'intérieur du système de paraphrasage : il faut d'abord bien le roder et s'assurer de sa complétude relative. Par ailleurs, nous croyons qu'il est utile et intéressant de pouvoir obtenir certaines paraphrases par des voies différentes.

3. Règles lexicales de paraphrasage

Dans ce qui suit, les conventions d'écriture suivantes sont adoptées :

1. C_0 (de «[mot] clé») dénote la lexie de départ, c'est-à-dire la lexie à laquelle la règle en question est appliquée et qui est le mot clé de la fonction lexicale correspondante.
2. Les actants syntaxiques profonds de C_0 sont notés, comme avant, I, II, III, ...
3. Les variables X, Y, Z, W identifient les nœuds correspondants dans les règles syntaxiques de paraphrasage desservant la règle lexicale donnée (voir les règles syntaxiques).
4. La flèche au-dessus du numéro d'une règle syntaxique signale son application dans le sens inverse (de droite à gauche); les parenthèses autour du numéro signalent le caractère facultatif de la règle. Voir par exemple, $\overline{13}$ dans la règle lexicale n° 21 et $\overline{10}$ dans la règle lexicale n° 22.

A. ÉQUIVALENCES SÉMANTIQUES

A.I. RÈGLES DU TYPE $C_0 \Leftrightarrow FL(C_0)$

Règle lexicale	Règles syntaxiques desservant la règle lexicale	Exemples
X Y	SUBSTITUTIONS <u>Substitutions synonymiques</u>	
1. $C_0 \Leftrightarrow \text{Syn}(C_0)$	—	<i>Ce phénomène s'explique facilement</i> [C_0] \Leftrightarrow <i>Ce phénomène s'explique sans difficulté</i> [Syn(C_0)]
	<u>Substitutions conversives</u>	
2. $C_0(v) \Leftrightarrow \text{Conv}_{21}(C_0)$	Conversifs à deux arguments	<i>Je [I] crains [C_0] les conséquences</i> [II] \Leftrightarrow <i>Les conséquences [I] m'[II] effraient</i> [Conv ₂₁ (C_0)]
	Conversifs à trois arguments	
3. $C_0(v) \Leftrightarrow \text{Conv}_{132}(C_0)$	22	<i>russe On [I] obuchaet [C_0] detej [II] jazyku [III], litt. 'Il instruit les enfants dans la langue'</i> \Leftrightarrow <i>On [I] prepodaet [Conv₁₃₂(C_0)] jazyk [II] detjam [III]</i> 'Il enseigne la langue aux enfants'

Règle lexicale	Règles syntaxiques desservant la règle lexicale	Exemples
X Y		
4. $C_0(v) \Leftrightarrow \text{Conv}_{321}(C_0)$	21	<i>Il [I] enseigne [C_0] la langue [III] aux enfants [III] \Leftrightarrow Les enfants [I] apprennent [Conv_{321}(C_0)] la langue [III] avec lui [III]</i>
5. $C_0(v) \Leftrightarrow \text{Conv}_{312}(C_0)$	21, 22	<i>russe On [I] zaplatil [C_0] za ètu knigu [III] 30 rublej [II] 'Il a payé ce livre 30 roubles' \Leftrightarrow Èta kniga [I] obošlas' [Conv_{312}(C_0)] emu [III] v 30 rublej [III], litt. 'Ce livre lui a coûté dans 30 roubles' = 'Ce livre lui a coûté 30 roubles'</i>
	Conversifs à quatre arguments	
6. $C_0(v) \Leftrightarrow \text{Conv}_{1432}(C_0)$	23, 22, 23	<i>Il [I] a acheté [C_0] ce livre [II] à Jean [III] 30 francs [IV] \Leftrightarrow Il [I] a payé [Conv_{1432}(C_0)] à Jean [III] 30 francs [II] pour ce livre [IV]</i>
7. $C_0(v) \Leftrightarrow \text{Conv}_{3214}(C_0)$	21	<i>Jean [I] lui [III] a vendu [Conv_{3214}(C_0)] ce livre [II] pour 30 francs [IV]</i>
N.B. : Dans cette règle, ainsi que dans les deux règles suivantes, $C_0 = \text{acheter}$, et la phrase de départ est toujours <i>Il a acheté ce livre à Jean 30 francs</i> .		
8. $C_0(v) \Leftrightarrow \text{Conv}_{142}(C_0)$	23, 22	<i>Il [I] a dépensé [Conv_{142}(C_0)] 30 francs [II] pour ce livre [III]^①</i>
9. $C_0(v) \Leftrightarrow \text{Conv}_{241}(C_0)$	23, 21	<i>Ce livre [I] lui [III] a coûté [Conv_{241}(C_0)] 30 francs [II]</i>
10. $C_0(v) \Leftrightarrow \text{Conv}_{412}(C_0)$	23, 22, 20	<i>russe U nego [II] ušlo [Conv_{412}(C_0)] na knigu [III] 30 rublej [I], litt. '30 roubles sont partis de chez lui pour ce livre'^②</i>
	Substitutions dérivatives (= adverbiales)	
11. $C_0(v) \Leftrightarrow \text{Adv}_{1A}(C_0)$	25	<i>Il semble [C_0] [que cette approche est valable] [I] \Leftrightarrow À ce qu'il semble [Adv_{1A}(C_0)], cette approche est valable^③</i>
12. $C_0(v) \Leftrightarrow \text{Adv}_{1B}(C_0)$	27	<i>Il [I] se hâta [C_0] de sortir [II] \Leftrightarrow Il [I] sortit en hâte [Adv_{1B}(C_0)]</i>
13. $C_0(v) \Leftrightarrow \text{Adv}_{1C}(C_0)$	28	<i>Ils [I] ont terminé [C_0] le voyage [II] par une visite [III] à Samarkande \Leftrightarrow À la fin [Adv_{1C}(C_0)] du voyage, ils [I] ont visité Samarkande</i>

X	Y	Z	Règle lexicale	Règles syntaxiques desservant la règle lexicale	Exemples
				<u>Substitution métonymique</u>	
14.	$C_{0(N)} \Leftrightarrow \left\{ \begin{array}{l} S_1 \text{Real}_1(C_0) \\ S_1(C_0) \end{array} \right\}$	$(\xrightarrow{\text{II}} C_0)$		—	<i>La voiture [C₀] ralentit \Leftrightarrow Le conducteur [S₁Real₁(C₀)] (de la voiture) ralentit; L'agression [C₀] dépasse toutes les bornes \Leftrightarrow L'agresseur [S₁(C₀)] dépasse toutes les bornes</i>
15.	$C_{0(V)} \Leftrightarrow \text{Conv}_{12\text{-est-}3}(C_0)$	$\xrightarrow{\text{II}} \text{ÊTRE}$	6	<u>FISSIONS SIMPLES</u> <u>Fissions à copule</u>	<i>Posons [C₀] le diamètre [II] égal [III] à 10 \Leftrightarrow Supposons [Conv₁₂-est-₃(C₀)] [que le diamètre est égal à 10] ④</i> <i>Cette approche [I] me [II] semble [C₀] valable [III] \Leftrightarrow Il me [II] semble [Conv₁-est-₂₃(C₀)] [que cette approche est valable] [I]</i>
16.	$C_{0(V)} \Leftrightarrow \text{Conv}_{1\text{-est-}23}(C_0)$	$\xrightarrow{\text{II}} \text{ÊTRE}$	5		
17.	$\left\{ \begin{array}{l} C_{0(V)} \\ \text{ÊTRE} \longrightarrow C_{0(A)} \end{array} \right\}$	$\Leftrightarrow \text{Anti}(C_0) \xrightarrow{\text{ATTR}} \text{NE...PAS}$	—	<u>Fission à négation</u>	<i>Jean est resté [C₀] \Leftrightarrow Jean n'est pas parti [Anti(C₀)]; Lucie est courageuse [C₀] \Leftrightarrow Lucie n'est pas peureuse [Anti(C₀)] ⑤</i>
18.	$C_{0(V)} \Leftrightarrow S_0(C_0)$	$\xleftarrow{\text{II}} \text{Oper}_1(S_0(C_0))$	1, (12)	<u>Fissions à verbe support</u> ¹⁶	<i>Jean nous a bien accueillis [C₀] \Leftrightarrow Jean nous a réservé [Oper₁(S₀(C₀))] un bon accueil [S₀(C₀)]</i>
19.	$C_{0(V)} \Leftrightarrow S_0(C_0)$	$\xleftarrow{\text{II}} \text{Oper}_2(S_0(C_0))$	1, 17, (11)		<i>Jean nous a bien accueillis [C₀] \Leftrightarrow Nous avons trouvé [Oper₂(S₀(C₀))] chez Jean un bon accueil [S₀(C₀)]</i>
20.	$C_{0(V)} \Leftrightarrow S_0(C_0)$	$\xleftarrow{\text{II}} \text{Func}_0(S_0(C_0))$	3		<i>En Afghanistan, le peuple lutte [C₀] contre l'envahisseur soviétique \Leftrightarrow En Afghanistan, se déroule [Func₀(S₀(C₀))] la lutte [S₀(C₀)] du peuple contre l'envahisseur soviétique</i>
21.	$C_{0(V)} \Leftrightarrow S_0(C_0)$	$\xleftarrow{\text{II}} \text{Func}_1(S_0(C_0))$	4, (13)		<i>... sans qu'il nous aide [C₀] \Leftrightarrow ... sans qu'une (aucune) aide [S₀(C₀)] nous vienne [Func₁(S₀(C₀))] de lui</i>
22.	$C_{0(V)} \Leftrightarrow S_0(C_0)$	$\xleftarrow{\text{II}} \text{Func}_2(S_0(C_0))$	3, (10)		<i>Ils se préoccupent [C₀] surtout des luttes d'influence \Leftrightarrow Leurs préoccupations [S₀(C₀)] vont [Func₂(S₀(C₀))] surtout aux luttes d'influence</i>

Règle lexicale	Règles syntaxiques desservant la règle lexicale	Exemples
X	Y	Z
23. $C_0(V) \Leftrightarrow S_0(C_0)$ \xrightarrow{III} $Labor_{12}(S_0(C_0))$	7	<i>L'aviation soviétique a farouchement bombardé [C₀] le village</i> \Leftrightarrow <i>L'aviation soviétique a soumis [Labor₁₂(S₀(C₀))] le village à un bombardement [S₀(C₀)] farouche</i>
24. $C_0(V) \Leftrightarrow S_1/A_1(C_0)$ \xrightarrow{II} $Oper_1(S_1/A_1(C_0))$	1	<i>Belfort, qui a écrit [C₀] cet article, ...</i> \Leftrightarrow <i>Belfort, qui est [Oper₁(S₁(C₀))] l'auteur [S₁(C₀)] de cet article, ...</i> ^⑥
25. $C_0(V) \Leftrightarrow S_2/A_2(C_0)$ \xrightarrow{II} $Oper_1(S_2/A_2(C_0))$	20, 1	<i>Je le hais [C₀]</i> \Leftrightarrow <i>Il m'est [Oper₁(A₂(C₀))] détestable [A₂(C₀)]</i>
26. $C_0(V) \Leftrightarrow S_1(C_0)$ \xrightarrow{III} $Labor_{21}(S_1(C_0))$	20, 7	<i>Je suis content que vous la soigniez [C₀]</i> \Leftrightarrow <i>Je suis content qu'elle vous ait [Labor₂₁(S₁(C₀))] comme médecin [S₁(C₀)]</i>
27. $C_0(V) \Leftrightarrow S_2(C_0)$ \xrightarrow{III} $Labor_{21}(S_2(C_0))$	7	<i>Il est plaisant de le servir [C₀]</i> \Leftrightarrow <i>Il est plaisant de l'avoir [Labor₂₁(S₂(C₀))] comme client [S₂(C₀)]</i>
28. $C_0(V) \Leftrightarrow S_{instr}(C_0)$ \xrightarrow{II} $Real_1(S_{instr}(C_0))$	1, (12)	<i>Pierre conduisait [C₀]</i> \Leftrightarrow <i>Pierre tenait le (était au) [Real₁(S_{instr}(C₀))] volant [S_{instr}(C₀)]</i>
29. $C_0(V) \Leftrightarrow S_{res}(C_0)$ \xrightarrow{II} $CausFunc_{0/2}(S_{res}(C_0))$	1, (12)	<i>russe On skopiroval [C₀] ètot dokument</i> ‘Il a copié ce document’ \Leftrightarrow <i>On snjal [CausFunc₂(S_{res}(C₀))] s ètogo dokumenta kopiju [S_{res}(C₀)], litt. ‘Il a ôté de ce document une copie’</i>
30. $C_0(V) \Leftrightarrow S_{med}(C_0)$ \xrightarrow{II} $CausOper1(Smed(C0))$	2	<i>Ils ont asphalté [C₀] ce chemin</i> \Leftrightarrow <i>Ils ont couvert [CausOper₁(S_{med}(C₀))] ce chemin d'asphalte [S_{med}(C₀)]</i>
<u>Fission à Gener</u>		
31. $C_0 \Leftrightarrow Gener(C_0)$ \xrightarrow{ATTR} $A_0/Adv_1(C_0)$	—	<i>un liquide [C₀]</i> \Leftrightarrow <i>un corps [Gener(C₀)] liquide [A₀(C₀)]</i> ; russe <i>prošepтал</i> [C ₀] ‘murmura’ \Leftrightarrow <i>skazal [Gener(C₀)] šepotom</i> [Adv ₁ (C ₀)], litt. ‘dit en murmure’
<u>Fissions métonymiques et métaphoriques</u>		
32. $C_0 \Leftrightarrow Figur(C_0)$ \xrightarrow{ATTR} C_0	—	<i>Le brouillard [C₀] nous cache le sommet</i> \Leftrightarrow <i>Le rideau [Figur(C₀)] de brouillard nous cache le sommet</i>

Règle lexicale	Règles syntaxiques desservant la règle lexicale	Exemples		
X	Y	Z		
33. $C_0(N) \Leftrightarrow \text{Sing}_{\text{pl}}(C_0)$	$\xrightarrow{\text{II}} C_0$	_____	_____	<i>La pluie [C₀] tambourinait sur le toit \Leftrightarrow Les gouttes [Sing(C₀)] de pluie tambourinaient sur le toit</i> ^⑦
34. $C_{0(N)\text{pl}} \Leftrightarrow \text{Mult}(C_0)$	$\xrightarrow{\text{II}} C_0$	_____	_____	<i>les professeurs [C₀] de l'Université de Montréal \Leftrightarrow le corps [Mult(C₀)] professoral de l'Université de Montréal</i>
FISSIONS COMPLEXES				
X	W	Y	Z	
35. $C_{0(V)} \xrightarrow{\text{ATTR}} \text{Loc}_{\text{in}} \Leftrightarrow S_{\text{loc}}(C_0)$		$\xrightarrow{\text{II}} \text{Oper}_1(S_{\text{loc}}(C_0))$	8	<i>Les milices musulmanes combattaient [C₀] dans [Loc_{in}] les rues de Saïda \Leftrightarrow Les rues de Saïda sont [Oper₁(S_{loc}(C₀))] l'arène [S_{loc}(C₀)] du combat [S₀(C₀)] des milices musulmanes</i>
36. $C_{0(V)} \xrightarrow{\text{ATTR}} \text{Instr} \Leftrightarrow S_{\text{instr}}(C_0)$		$\xrightarrow{\text{II}} \text{Oper}_1(S_{\text{instr}}(C_0))$	8	<i>Il a été assassiné [C₀] avec [Instr] ce stylet \Leftrightarrow Ce stylet a été [Oper₁(S_{instr}(C₀))] l'arme [S_{instr}(C₀)] de son assassinat [S₀(C₀)]</i>
37. $C_{0(V)} \xrightarrow{\text{ATTR}} A \Leftrightarrow \left\{ \begin{array}{l} S_{\text{mod}}(C_0) \\ S_0(C_0) \end{array} \right\}$		$\xleftarrow{\text{I}} \text{Pred} \xrightarrow{\text{II}} A$	26	<i>Lucie écrit [C₀] clairement [A] \Leftrightarrow Le style [S_{mod}(C₀)] de Lucie est [Pred] clair [A]</i>
A.II. RÈGLES DU TYPE $\text{FL}_1(C_0) \Leftrightarrow \text{FL}_2(C_0)$				
SUBSTITUTIONS				
X	Y			
38. $\text{Oper}_1(C_0) \Leftrightarrow \text{Oper}_2(C_0)$		1) 21, 11		<i>Les partisans leur ont opposé [Oper₁(C₀)] une résistance [C₀] acharnée \Leftrightarrow Ils se sont butés [Oper₂(C₀)] sur une résistance [C₀] acharnée des partisans</i>
		2) 12, 21		<i>Il a manifesté [Oper₁(C₀)] une bonne compréhension [C₀] de ce point de vue \Leftrightarrow Ce point de vue a rencontré [Oper₂(C₀)] chez lui une bonne compréhension [C₀]</i>
39. $\text{Func}_1(C_0) \Leftrightarrow \text{Func}_2(C_0)$		16		<i>L'idée [C₀] d'appliquer l'opérateur λ appartient [Func₁(C₀)] au Pr. Cusson \Leftrightarrow L'idée [C₀] du Pr. Cusson consiste [Func₂(C₀)] à appliquer l'opérateur λ</i>

Règle lexicale X	Règles syntaxiques desservant la règle lexicale Y	Exemples
40. Labor ₁₂ (C ₀) ⇔ Labor ₃₂ (C ₀)	21	russe <i>Ivan vzjal</i> [Labor ₁₂ (C ₀)] <i>v dolg</i> [C ₀] <i>u Petra 100 rublej</i> , litt. 'Jean a-pris en prêt à-Pierre 100 roubles' ⇔ <i>Pétr dal</i> [Labor ₃₂ (C ₀)] <i>Ivanu v dolg</i> [C ₀] <i>100 rublej</i> , litt. 'Pierre a-donné à-Jean en prêt 100 roubles'
41. Oper ₁ (C ₀) ⇔ Func ₁ (C ₀)	20	<i>Le général Clas a donné</i> [Oper ₁ (C ₀)] <i>cet ordre</i> [C ₀] ⇔ <i>Cet ordre</i> [C ₀] <i>est venu</i> [Func ₁ (C ₀)] <i>du général Clas</i>
42. Oper ₁ (C ₀) ⇔ Labor ₁₂ (C ₀)	1) (12) 2) 22	<i>Il a effectué</i> [Oper ₁ (C ₀)] <i>une analyse</i> [C ₀] <i>profonde du phénomène</i> ⇔ <i>Il a soumis</i> [Labor ₁₂ (C ₀)] <i>le phénomène à une analyse</i> [C ₀] <i>profonde</i> <i>L'algorithme effectue</i> [Oper ₁ (C ₀)], <i>sur la phrase donnée, la transformation</i> [C ₀] <i>suivante</i> ⇔ <i>L'algorithme soumet</i> [Labor ₁₂ (C ₀)] <i>la phrase donnée à la transformation</i> [C ₀] <i>suivante</i>
43. Real _i (C ₀) ⇔ Adv _{1B(B)} Real _i (C ₀)	29	<i>Il a suivi</i> [Real ₃ (C ₀)] <i>mon conseil</i> [C ₀] <i>de rester chez lui</i> ⇔ <i>Il est resté chez lui suivant</i> [Adv _{1B(B)} Real ₃ (C ₀)] <i>mon conseil</i> [C ₀] ⑧
44. Magn ₂ (C ₀) ⇔ MagnS ₂ (C ₀)	24	<i>Ça coûte</i> [C ₀] <i>très-très cher</i> [Magn ₂ (C ₀)] ⇔ <i>Ça coûte</i> [C ₀] <i>la peau des fesses</i> [MagnS ₂ (C ₀)] ¹⁷
FISSIONS (À NÉGATION)		
45. Liqu(C ₀) $\xrightarrow{\text{II}}$ C ₀ ⇔ Caus(Anti(C ₀)) $\xrightarrow{\text{II}}$ Anti(C ₀)	—	<i>D'un geste menaçant, il a mis fin</i> [LiquFunc ₀ (C ₀)] <i>au brouhaha</i> [C ₀] <i>dans la salle</i> ⇔ <i>D'un geste menaçant, il a imposé</i> [CausFunc ₀ (Anti(C ₀))] <i>le silence</i> [Anti(C ₀)] <i>dans la salle</i> ⑨
46. Perm(C ₀) ⇔ Liqu(C ₀) $\xrightarrow{\text{ATTR}}$ NE...PAS	—	<i>L'administration ferme les yeux</i> [PermFunc ₀ (C ₀)] <i>sur l'alcoolisme</i> [C ₀] <i>des travailleurs</i> ⇔ <i>L'administration ne combat</i> [LiquFunc ₀ (C ₀)] <i>pas l'alcoolisme des travailleurs</i>
47. Fin(C ₀) $\xrightarrow{\text{II}}$ C ₀ ⇔ Incep(Anti(C ₀)) $\xrightarrow{\text{II}}$ Anti(C ₀)	—	<i>Jean a cessé</i> [Fin(C ₀)] <i>de soigner</i> [C ₀] <i>sa goutte</i> ⇔ <i>Jean a commencé</i> [Incep(Anti(C ₀))] <i>à négliger</i> [Anti(C ₀)] <i>sa goutte</i> ⑩
48. Cont(C ₀) ⇔ Fin(C ₀) $\xrightarrow{\text{ATTR}}$ NE...PAS	—	<i>garder</i> [ContOper ₁ (C ₀)] <i>son calme</i> [C ₀] ⇔ <i>ne pas perdre</i> [FinOper ₁ (C ₀)] <i>son calme</i>

B. IMPLICATIONS SÉMANTIQUES

X	Y		
49. CausX($C_0(N)$) \Rightarrow IncepX($C_0(N)$)		20	<i>Jean a mis en marche [CausFact₀(C_0)] le moteur [C_0] \Rightarrow Le moteur [C_0] a démarré [IncepFact₀(C_0)]</i>
50. CausX($C_0(V)$) \Rightarrow X($C_0(V)$)		20, 22	<i>Jean a introduit [Caus(entrer)] le boulon dans l'orifice \Rightarrow Le boulon est entré [C_0] dans l'orifice</i>
51. PermX($C_0(N)$) \Rightarrow X($C_0(N)$)		20, 22	<i>russe Ivana dopustili [PermOper₂(C_0)] k èkzamenu [C_0], litt. ‘[On] a-admis Jean à-l'examen’ \Rightarrow Ivan deržal [Oper₂(C_0)] èkzamen, litt. ‘Jean a-tenu l'examen’ [=‘Jean a fait son examen, mais on ne sait pas s'il l'a réussi ou pas’]</i>
52. LiquX($C_0(N)$) \Rightarrow FinX($C_0(N)$)		20	<i>Jean a freiné [LiquFact₀(C_0)] (le camion [C_0]) \Rightarrow Le camion [C_0] s'est arrêté [FinFact₀(C_0)]</i>
53. IncepX \Rightarrow X		—	<i>Le moteur [C_0] a démarré [IncepFact₀(C_0)] \Rightarrow Le moteur [C_0] tourne [Fact₀(C_0)]</i>
54. ContX \Rightarrow X		—	<i>Jean gardait [ContOper₁(C_0)] son poste [C_0] \Rightarrow Jean était [Oper₁(C_0)] à son poste [C_0]</i>

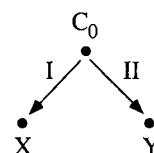
COMMENTAIRES SUR LES RÈGLES LEXICALES DE PARAPHRASE

①, p. 38, n° 8. La règle n° 8, ainsi que les règles n°s 9 et 10, décrivent la conversion avec la perte d'un actant syntaxique : en l'occurrence, ACHETER possède quatre actants, mais DÉPENSER et COÛTER n'en ont que trois (ils n'admettent pas la désignation du vendeur). De telles règles ne sont applicables qu'aux SSyntP où l'actant amputé est de toute façon absent : dans *Jean a acheté ce livre à Lucie 30 francs*, le verbe *acheter* ne saurait pas être remplacé par *dépenser* ou *coûter*, mais il peut l'être dans *Jean a acheté ce livre 30 francs*.

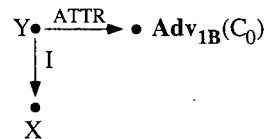
②, p. 38, n° 10. Nous n'avons pas cité ici toutes les règles de conversion logiquement possibles; il est d'ailleurs très simple de les fournir.

③, p. 38, n° 11. **Adv₁** dénote — comme il a été indiqué dans la liste de FL aux pp. 127-131 — l'adverbe caractérisant, par rapport à la situation ' C_0 ', le premier actant de C_0 ($Adv_1 \approx$

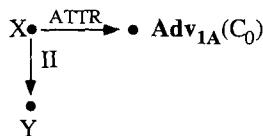
'en faisant C_0 '). La lettre majuscule en indice signale l'actant du mot clé (**A** = I, **B** = II et **C** = III) qui devient le sommet syntaxique de la proposition, si le mot clé est remplacé par son **Adv₁**. Par exemple, si l'on a



et que l'on remplace C_0 par $Adv_{1B}(C_0)$, c'est Y qui sera le sommet :



mais si l'on devait remplacer C_0 par $\text{Adv}_{1A}(C_0)$, le résultat en serait différent : le sommet doit être X; comparez :

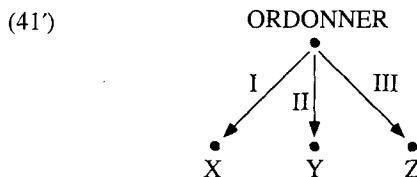


[dans une langue comme le français ou l'anglais, le verbe X devenu le sommet syntaxique doit être muni d'un SG factice : *il* ou *it*].

④, p. 39, n° 15. **Conv_{12-est-3}** et **Conv_{1-est-23}** dénotent les conversions similaires : dans le premier cas, les actants II et III forment, à l'aide du verbe ÊTRE, un seul actant — une proposition complétive; dans le deuxième cas, la même chose arrive aux actants I et II. Remarquons que les phrases similaires du type (41a) et (41b) :

- (41) a. *Il ordonna à Lucie de partir.*
- b. *Il ordonna que Lucie parte.*

ne sont pas converses l'une par rapport à l'autre, puisqu'elles dérivent de la même structure SyntP :



(*X ordonne à Z que [Y soit fait*). Alors que dans (41b), le troisième actant de *ordonner* n'est pas exprimé (cf. *Il m'ordonna que Lucie parte*), dans (41a), le troisième actant d'*ordonner* et le premier actant de son deuxième actant (= de *partir*) sont identiques (c'est *Lucie*), ce qui entraîne, suivant la règle générale de la syntaxe française, l'utilisation de l'infinitif au lieu d'une subordonnée complétive. Entre (41a) et (41b), on observe une nuance évidente : dans (41b), il n'est pas dit à qui il a adressé son ordre, et ce n'est pas nécessairement à *Lucie*.

⑤, p. 39, n° 17. Comme m'a fait remarquer A. Tutin, le remplacement par antonyme s'applique beaucoup mieux pour les antonymes «binaires», tels *absent ~ présent* ou *savoir ~ ignorer*, puisque 'absent' = 'ne pas présent', etc. Par contre, pour les antonymes «gradués», tels *vieux ~ jeune* ou *aimer ~ détester*, la paraphrase n'est pas précise : 'vieux' ≠ 'ne pas jeune' et 'aimer' ≠ 'ne pas détester'. Dans ce dernier cas, on a besoin d'un intensificateur de négation (*Il n'est pas du tout jeune!*) et/ou d'une intonation spéciale.

⑥, p. 40, n° 24. La valeur de la FL **Oper₁** d'un **S₁** est invariablement ÊTRE, en sorte qu'on pourrait mettre ce verbe particulier dans les règles n°s 24 et 25 au lieu de **Oper₁**. Cependant, nous avons préféré écrire ces règles sous la forme la

plus générale possible et les garder dans la section des règles de fissions à verbe support. La même chose est vraie à propos des règles n°s 26 et 27, puisque la valeur de **Labor₂₁** d'un **S₁** ou d'un **S₂** est aussi invariablement AVOIR.

⑦, p. 41, n° 33. Les substitutions décrites par les règles n°s 32 et 33 ne sont possibles que dans le contexte de quelques FL qui acceptent **Figur**(C_0) ou **Sing_{pl}**(C_0) au lieu de C_0 . Ainsi, *La pluie tambourinait* = *Les gouttes de pluie tambourinaient*, mais *La pluie a duré trois heures* n'est pas équivalent, bien sûr, à **Les gouttes de pluie ont duré trois heures*. Les paraphrases incorrectes de ce dernier type doivent être bloquées, d'une part, par des conditions de combinatoire que nous ne formulons pas dans notre liste de règles, et d'autre part, par les incompatibilités sémantiques dont il a été question plus haut (p. 35, point 1).

⑧, p. 42, n° 43. Le problème des paraphrases utilisant les adverbes du type **AdvReal** est discuté dans Žolkovskij 1977; dans ce qui suit, nous avons recours aux résultats de cet article.

— Le symbole **Adv_{B(B)}(X)** désigne un dérivé adverbial de X qui peut être caractérisé comme suit : quand X est remplacé par **Adv_{B(B)}(X)**, c'est le II^e actant SyntP du II^e actant SyntP de X qui devient le sommet syntaxique de la structure transformée. L'indice "1" exprime, comme c'est l'habitude dans le formalisme des fonctions lexicales, l'orientation syntaxique de l'adverbe en cause par rapport à son mot clé X (**Adv₁(X)** ≈ 'X-ant' plutôt que 'étant X-é'). Dans notre exemple, *rester* est le II^e actant SyntP de *conseil*, qui, lui, est, à son tour, le II^e actant SyntP de *suivre* [= X]; **Adv_{1B(B)}Real₃(conseil)** = *suivant* [ART ~]. Voici deux autres exemples — pour **Real₁** et **Real₂** :

Adv_{1B(B)}Real₁(règle) = *d'après* [ART ~] :

Jean observe [= **Real₁(règle)**] *cette règle d'accord des adjectifs* ~ *Jean accorde les adjectifs d'après cette règle*.

Adv_{1B(B)}Real₂(invitation 1a) = *à, sur* [ART ~] :

Jean a répondu [= **Real₂(invitation 1a)**] *à l'invitation du rectorat de venir* ~ *Jean est venu sur l'invitation du rectorat*.

⑨, p. 42. Les règles n°s 45 et 47 peuvent probablement être réécrites sous une forme plus générale (suggérée dans Apresjan 1974 : 331) :

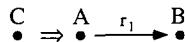
(*) $X \rightarrow Y \Leftrightarrow \text{Anti}(X) \rightarrow \text{Anti}(Y)$.

La règle (*) peut être illustrée par des exemples *Il s'en souviendra* [X] *toujours* [Y] \Leftrightarrow *Il ne l'oubliera* [$\text{Anti}(X)$] *jamais* [$\text{Anti}(Y)$], *On doit* [X] [*le prendre en considération*] [Y] \Leftrightarrow *On ne doit pas* [$\text{Anti}(X)$] *le négliger* [$\text{Anti}(Y)$], etc. En effet, *Liqu* = $\text{Anti}(\text{Caus})$ et *Fin* = $\text{Anti}(\text{Incep})$. La règle (*) serait d'un nouveau type : substitution antonymique complexe — et devrait être comprise dans le groupe A.I, SUBSTITUTIONS.

Les RÈGLES SYNTAXIQUES DE PARAPHRASAGE décrivent les changements de la SSyntP qui doivent accompagner les substitutions lexicales effectuées par des règles lexicales. Les règles syntaxiques de paraphrasage correspondent, dans un certain sens, aux transformations de la grammaire transformationnelle : les transformations des arbres étiquetés de dépendance qui représentent les SSyntP des phrases.

Remarquons que, d'une part, il n'y a que trois opérations syntaxiques fondamentales (= élémentaires) au niveau profond :

- la fission d'un nœud en deux, ou l'ajout d'une branche

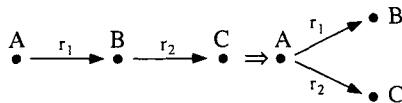


vs la fusion de deux nœuds, ou la contraction d'une branche

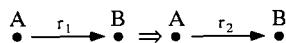


c'est-à-dire l'inverse de la fission;

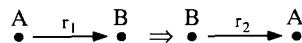
- le transfert d'une branche avec son nœud terminal



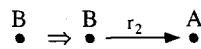
- le réétiquetage d'une branche



Les autres opérations syntaxiques concernant la structure syntaxique profonde, telles que l'inversion de la subordination, du type



peuvent s'exprimer comme des compositions de ces opérations fondamentales. Par exemple, l'inversion de subordination que nous venons de mentionner peut être remplacée par la suite d'une fission :



et d'une fusion :



D'autre part, les branches de l'arbre d'une SSyntP ne portent que neuf étiquettes différentes (voir les relations syntaxiques profondes : pp. 18 ssq.) et sont soumises à des contraintes de cooccurrence rigoureuses; ainsi, deux branches étiquetées

par un même numéro actanciel (I, II, ..., VI) ne peuvent partir d'un même noeud, etc.

Par conséquent, on peut énumérer toutes les transformations élémentaires définies sur les SSyntP du modèle Sens-Texte. Ces transformations sont introduites et analysées dans ce que nous appelons la GRAMMAIRE Δ (Δ vient du grec $\Delta\text{EN}\Delta\text{PON}$ 'arbre', il s'agit donc d'une «dendro-grammaire», ou «grammaire d'arborescences»; voir Gladkij et Mel'čuk 1975 (1983)). Toutes les règles syntaxiques de paraphrasage peuvent être formulées à partir de ces transformations élémentaires. Cependant, dans ce qui suit, nous n'avons pas recours à des transformations vraiment élémentaires dans tous les cas, puisque cela risque de surcharger l'exposé. Nous utilisons un répertoire de règles syntaxiques de base, qui sont les plus simples possibles (sans être nécessairement élémentaires) tout en gardant leur caractère linguistique; chacune de nos règles syntaxiques de base correspond à un phénomène syntaxique connu des langues naturelles. Dans nombre de cas, une seule règle syntaxique de base ne suffit pas pour desservir une substitution lexicale (évoquée par une règle lexicale de paraphrasage). Dans ce cas, nous avons recours à des suites de règles syntaxiques dont chacune doit être appliquée au résultat de l'application de la règle précédente (cf., par exemple, les règles lexicales n°s 5, 6, 8-10, etc.).

Nous présentons les règles syntaxiques de paraphrasage en quatre classes, dont les trois premières correspondent aux trois types d'opération décrit ci-dessus et la quatrième fait intervenir tous les trois.

- A – les opérations de fission / de fusion;
- B – les opérations de transfert;
- C – les opérations de réétiquetage;
- D – les opérations d'inversion de subordination.

Auprès de chaque règle syntaxique, nous indiquons, à des fins de référence, les numéros des règles lexicales qu'elle dessert.

L'application des règles syntaxiques de paraphrasage est soumise à certaines conditions générales, qui sont en fait des MÉTA-RÈGLES. Nous citerons ici, en guise d'illustration, les deux méta-règles concernant le traitement des nœuds autres que ceux faisant partie de la règle syntaxique.

Méta-règle 1. Lors de l'application d'une règle syntaxique de paraphrasage **R** à une structure SyntP, tout nœud de cette structure qui n'est pas mentionné dans **R** reste attaché à son gouverneur par la même RelSyntP, sauf mention expresse du contraire dans une autre autre méta-règle.

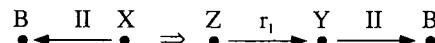
Méta-règle 2. Si, lors de l'application d'une règle syntaxique de paraphrasage **R** à une structure SyntP, un nœud **X**

est remplacé par deux nœuds $Z \xrightarrow{r_1} Y$, alors les dépendants de X qui ne sont pas mentionnés dans R sont distribués entre Z et Y de la façon suivante :

– L'actant I de X reste auprès du sommet du nouveau sous-arbre, c'est-à-dire auprès de Z :



– Les actants II-VI de X passent à Y :



– Les modificateurs circonstanciels (temps, lieu, ...) restent auprès de Z.

– Les modificateurs qualitatifs (y compris, par exemple, Magn) passent à Y, si Z est un verbe (semi-)vide, c'est-à-dire Oper_i, Func_i, Labor_{ij}; ou bien ils peuvent soit rester auprès de Z, soit — si Z est un verbe sémantiquement plein — passer à Y.

4. Règles syntaxiques de paraphrasage

Les conventions d'écriture sont évidentes. Signalons, néanmoins, que les variables X, Y, Z, W identifient les nœuds affectés par les substitutions lexicales et renvoient aux éléments correspondants des règles lexicales. Ainsi, dans la règle syntaxique n° 1, lorsqu'elle dessert la règle lexicale n° 19, X doit être un C₀(v), Z — un Oper₂(S₀(C₀)), et Y — un S₀(C₀).

Certaines règles syntaxiques de paraphrasage ne desservent aucune règle lexicale particulière (les règles syntaxiques n°s 9, 14, 15, 18 et 19). Leur vocation est de décrire des transformations plus générales de structures syntaxiques profondes.

A. FISSIONS

	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
1.	$X \bullet \Leftrightarrow \begin{array}{c} Z \\ \\ \bullet \\ \downarrow \\ II \\ \\ \bullet \\ Y \end{array}$	18, 19, 24, 25, 28, 29	<i>Ils luttent [X] ⇔ Ils mènent [Z] une lutte [Y]</i>
2.	$X \bullet \Leftrightarrow \begin{array}{c} Z \\ \\ \bullet \\ \downarrow \\ III \\ \\ \bullet \\ Y \end{array}$	30	<i>Il beurre [X] son pain ⇔ Il tartine [Z] son pain avec du beurre [Y]</i>
3.	$\begin{array}{c} X \\ \\ \bullet \\ \downarrow \\ I \\ \\ \bullet \\ Y \\ \\ \bullet \\ A \end{array} \Leftrightarrow \begin{array}{c} Z \\ \\ \bullet \\ \downarrow \\ I \\ \\ \bullet \\ Y \\ \\ \bullet \\ A \end{array}$	20, 22	<i>Partout en Afrique, les peuples [A] luttent [X] contre les dictatures communistes ⇔ Partout en Afrique, se déroule [Z] la lutte [Y] des peuples [A] contre les dictatures communistes</i>
4.	$\begin{array}{c} X \\ \\ \bullet \\ \downarrow \\ A \end{array} \Leftrightarrow \begin{array}{c} Z \\ \\ \bullet \\ \downarrow \\ I \\ \\ \bullet \\ Y \\ \\ \bullet \\ II \\ \\ \bullet \\ A \end{array}$	21	<i>Jean [A] hait [X] Lucie profondément ⇔ La haine [Y] envers Lucie remplit [Z] le cœur de Jean [A]</i>

	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
5.	<pre> graph TD X((X)) -- I --> A((A)) X -- III --> C1((C)) C1 -- II --> C2((C)) C1 -- I --> Y((Y)) C2 -- I --> A </pre>	16	<i>Cette proposition [A] me [C] semble [X] raisonnable \Leftrightarrow [Il] me [C] semble [Z] que cette proposition [A] est [Y] raisonnable</i>
6.	<pre> graph TD X((X)) -- II --> B1((B)) X -- III --> C1((C)) B1 -- I --> B2((B)) B1 -- II --> C2((C)) C2 -- II --> C </pre>	15	<i>Je la [B] crois [X] malade [C] \Leftrightarrow Je crois [Z] qu'elle [B] est [Y] malade [C]</i>
7.	<pre> graph TD X((X)) -- II --> B1((B)) X -- III --> C1((C)) B1 -- II --> B2((B)) B1 -- III --> C2((C)) C2 -- III --> C </pre>	23, 26, 27	<i>Il analyse [X] le texte [B] \Leftrightarrow Il soumet [Z] le texte [B] à l'analyse [Y]</i>
8.	<pre> graph TD X((X)) -- I --> A((A)) X -- II --> B1((B)) B1 -- II --> B2((B)) B1 -- I --> Z((Z)) Z -- II --> Y((Y)) Y -- "I/II" --> A </pre>	35, 36	<i>Ils [A] ont fait l'amour [X] sur [W] ce vieux sofa [B] \Leftrightarrow Ce vieux sofa [B] a été [Z] le théâtre [Y'] de leurs [A] ébats amoureux [Y"]</i>

B. TRANSFERTS

Les opérations de transfert presupposent un changement de niveau de subordination des actants, c'est-à-dire le passage d'un actant d'un étage de la SSyntP à un autre.

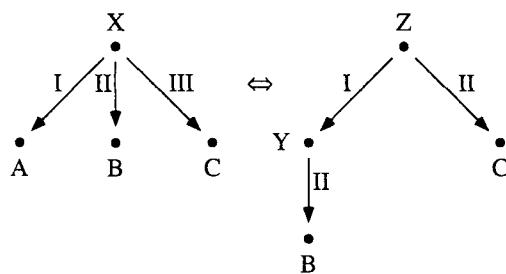
	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
9.			<i>Un ordre [A] nous est parvenu [X] du général Morin [B] ⇔ Un ordre [A] du général Morin [B] nous est parvenu [Y]</i>
10.			22 <i>Le film de Busacker [A], qui est à l'affiche partout, présente [X] la France médiévale [B] ⇔ Le film de Busacker [A] sur la France médiévale [B] est à l'affiche [Y] partout</i>
11.			19, 38 <i>Jean bénéficie [X] auprès de Lise [C] d'une entière confiance [B] ⇔ Jean bénéficie [Y] de l'entièle confiance [B] de Lise [C]</i>
12.			18, 28, 29, 38, 42 <i>Jean a consacré [X] un roman [B] aux linguistes [C] ⇔ Jean a écrit [Y] un roman [B] sur les linguistes [C]</i>

Règle syntaxique

Règles lexicales qu'elle dessert

Exemples

13.



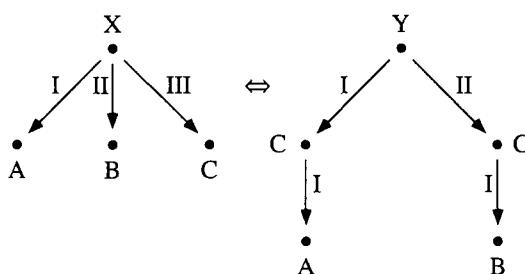
21

*L'aide importante [A] nous [B] vient [X] de lui [C] \Leftrightarrow *L'aide importante [Y] à nous [B] vient [Z] de lui [C]*

N.B. :

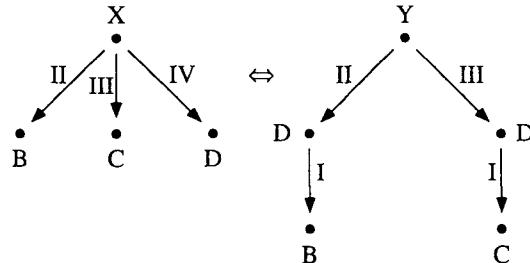
Cette règle est nécessaire pour corriger les résultats intermédiaires agrammaticaux d'autres règles.

14.



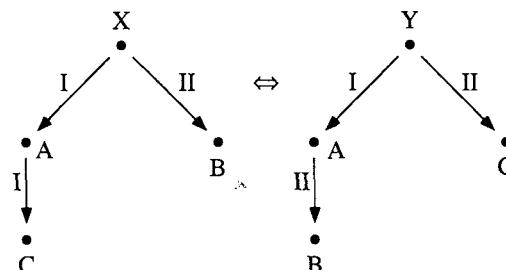
Jean [A] dépasse [X] Lise [B] par son intelligence [C] \Leftrightarrow L'intelligence [C] de Jean [A] dépasse [Y] celle [C] de Lise [B]

15.



Jean compare [X] Marie [B] avec Lise [C] d'après leur intelligence [D] \Leftrightarrow Jean compare [Y] l'intelligence [D] de Marie [B] à celle [D] de Lise [C]

16.



39

Le plan [A] de Jean [C] prévoit [X] un voyage dans les Alpes [B] \Leftrightarrow Le plan [A] de voyage dans les Alpes [B] vient [Y] de Jean [C]

	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
17.		19	<i>Les partis d'opposition [A] se prononcent [X] pour le soutien [B] de cette loi [C] ⇔ Cette loi [C] a [Y] le soutien [B] des partis d'opposition [A]</i>
18.		—	<i>Une envie [A] le saisit [X] irrésistiblement [P] ⇔ Une envie [A] irrésistible [P] le saisit [X]</i>
19.		—	<i>*... réserve [X] bien [P] un accueil [B] ⇔ ... réserve [X] un bon [P] accueil [B]</i>
20.		2, 10, 25, 26, 41, 49, 50, 51, 52	<i>Cet inventaire [A] comprend [X] la règle de déplacement [B] ⇔ La règle de déplacement [B] appartient [Y] à cet inventaire [A]</i>
21.		4, 5, 7, 9, 38, 40	<i>Jean [A] a vendu [X] la maison à Pham [C] ⇔ Pham [C] a acheté [Y] la maison à Jean [A]</i>
22.		3, 5, 6, 8, 10, 42, 50, 51	<i>Le président a envoyé [X] au Congrès [C] une lettre de remerciement [B] ⇔ Le président s'est adressé [Y] au Congrès [C] par une lettre de remerciement [B]</i>

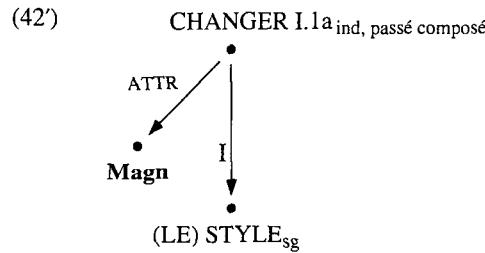
	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
23.	<pre> graph TD X((X)) -- III --> C((C)) X -- IV --> D((D)) Y((Y)) -- IV --> - Y -- III --> D </pre>	6, 8, 9, 10	<i>Jean a payé [X] 100 mille dollars (à Pham [C]) pour la maison [D] ⇔ Jean a dépensé [Y] 100 mille dollars pour la maison [D]</i>
N.B. : L'exemple illustrant la règle n° 23 présente le cas d'une conversion avec perte d'un actant (= C); cf. le commentaire ①, à la p. 43.			
24.	<pre> graph TD X((X)) -- II --> B((B)) X -- ATTR --> B </pre>	44	<i>payer [X] une fortune [B] ⇔ payer [X] cher [B]</i>
D. INVERSION DE SUBORDINATION			
25.	<pre> graph TD X((X)) -- I --> P((P)) X -- ATTR --> Y((Y)) </pre>	11	<i>[L'augmentation des prix] [P] accompagne [X] la crise ⇔ Les prix augmentent [P] avec [Y] la crise</i>
26.	<pre> graph TD Z((Z)) -- I --> Y((Y)) Z -- II --> W((W)) X((X)) -- ATTR --> W </pre>	37	<i>Son sommeil [Y] est [Z] profond [W] ⇔ Il dort [X] profondément [W]</i>
27.	<pre> graph TD X((X)) -- I --> A((A)) X -- II --> P((P)) X -- III --> B((B)) P -- ATTR --> Y((Y)) A -- I --> A B -- II --> B </pre>	12	<i>Jean [A] persistait [X] [à nier sa faute] [P] ⇔ Jean [A] [niait sa faute] [P] avec persistance [Y]</i>
28.	<pre> graph TD X((X)) -- I --> A((A)) X -- III --> P((P)) P -- ATTR --> Y((Y)) A -- I --> A </pre>	13	<i>Il [A] a commencé [X] son discours [par une brève présentation] [P] du budget ⇔ Au début [Y] de son dis- cours, il [A] a brièvement présen- té [P] le budget</i>

	Règle syntaxique	Règles lexicales qu'elle dessert	Exemples
29.		43	<i>Lise [A] tint [X] sa promesse [B] de venir [P] ⇔ Lise [A] vint [P], comme [Y] [elle l'avait promis] [B]</i>

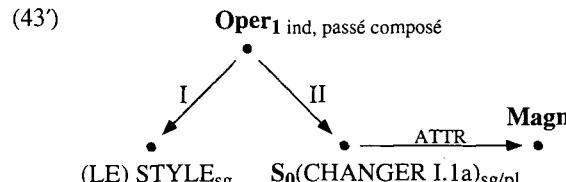
5. Un exemple de fonctionnement du système de paraphrasage

Nous montrerons ici comment le système ébauché permet de dériver des SSyntP synonymes à partir d'une SSyntP donnée. Pour ce faire, prenons un fragment d'une phrase déjà vue, à savoir la phrase (26) :

(42) [cf. (26), p. 26] *Le style ... a sérieusement changé*, qui a la SSyntP (42') :



La structure SyntP (42') sera soumise au système de paraphrasage. Tout d'abord, on peut appliquer à (42') une règle lexicale quelconque, par exemple, n° 18. Posant CHANGER I.1a = C₀ = X, on le remplace par S₀(CHANGER I.1a) = Y et Oper₁(S₀(CHANGER I.1a)) = Z; en utilisant la règle syntaxique correspondante n° 1, nous obtenons la SSyntP (43') :



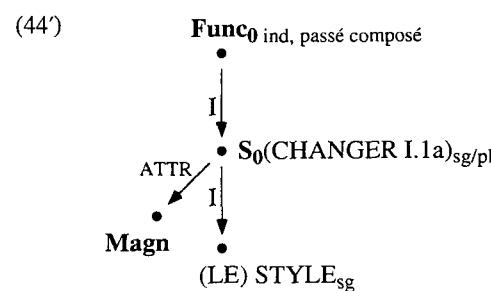
ce qui peut se lire comme la phrase (43) :

(43) *Le style ... a connu (subi) un changement / des changements sérieux.*

(Le nœud étiqueté **Magn** a été transféré du sommet Z à son dépendant Y suivant la métarègle 2 des pp. 45-46.)

N.B. : Remarquons que les paraphrases du type présenté en (43') et en (44') soulèvent le problème de l'article, puisque dans les deux cas, il faut introduire, dans la SSyntP synthétisée, l'article indéfini déterminant le nom S₀(CHANGER I.1a), c'est-à-dire CHANGEMENT I.1a. Lorsqu'il s'agit des transformations impliquant les fonctions lexicales, le choix de l'article peut être déterminé par la description lexicographique de la fonction lexicale en cause. Ainsi, Oper₁(change-ment I.1a) = connaître, subir [un <des> ~<s>], etc. Cependant, dans beaucoup d'autres cas, il faudrait des règles sémantiques spéciales qui pourraient introduire des articles dans la SSyntP à partir des données trouvées dans la RSém de départ. Ce sujet dépasse, quand même, le cadre de cet ouvrage.

On pourrait également appliquer à la SSyntP de (42') une autre règle lexicale de paraphrase : n° 20, avec la règle syntaxique n° 3; le résultat serait (44') :



c'est-à-dire

- (44) *Un changement / Des changements sérieux du style ... est / sont survenu / survenus.*

Une autre voie serait de remplacer d'abord CHANGER I.1a par son synonyme (plus étroit) SE TRANSFORMER (la règle lexicale de paraphrase n° 1) et appliquer ensuite au résultat les règles mentionnées. Cela nous donnerait des phrases comme (45) :

- (45) a. *Le style ... s'est beaucoup transformé.*
 b. *Le style ... s'est transformé de façon sérieuse (radicale, profonde).*
 c. *Le style ... a subi (connu) des transformations sérieuses (radicales, profondes).*
 d. *Des transformations sérieuses ... du style ... ont eu lieu (sont survenues).*

VI. REMARQUES FINALES

Après l'exposé de la section V, il doit être clair à quel point un lexique est important pour notre approche de la paraphrase et dans quelle mesure notre lexique (= DEC) est fondé sur la paraphrase. La paraphrase s'appuie sur les fonctions lexicales de façon tout à fait cruciale : les règles de paraphrasage sont formulées en termes de FL; or, les valeurs des FL pour chaque lexie sont spécifiées dans le lexique. Le paraphrasage se fait sous le contrôle des représentations séman-

tiques de lexies et en prenant en compte leurs propriétés syntaxiques (le nombre et le type d'actants, la combinatoire particulière); or, c'est encore une fois dans le lexique que toutes ces informations sont consignées. La paraphrase et, de façon plus générale, un modèle linguistique formel, ne sont même pas concevables sans un lexique formalisé fort complexe et très volumineux. Au sein du MST, le lexique en question est réalisé comme *Dictionnaire explicatif et combinatoire*.

REMERCIEMENTS

Gilles Houde et Henri Béjoint ont apporté des corrections au français de la première version de mon article. Lidija Iordaneskaja, Jean-Luc Descamps, Jurij Apresjan, Nicolas Pertsov, Alain Polguère et Bernard Levrat ont lu l'article en entier

et m'ont indiqué nombre d'incohérences et d'omissions importantes. André Clas et Victor Rosenzweig ont assuré la révision du texte final. C'est avec grand plaisir que je leur exprime ici ma gratitude la plus cordiale.

NOTES

¹ (Page 9.) Une version préliminaire du présent article a été publiée comme Mel'čuk 1988b et une version abrégée, comme Mel'čuk 1988c.

² (Page 11.) On pourrait se passer du terme *paraphrase* et ne se servir que du terme *synonymie / synonyme*. Cependant, nous croyons utile de le garder pour la raison suivante : dans l'usage courant, *synonymie* s'applique surtout aux lexèmes, si bien qu'il faut préciser explicitement chaque fois qu'il s'agit de synonymie de syntagmes ou de phrases. Par contre, le terme *paraphrase* implique la synonymie de syntagmes/phrases, et de ce fait, est très commode pour nos buts. Il faudrait noter que déjà le livre de Gleitman et Gleitman (1970 : 25 *ssq.*) insistait sur l'importance de la paraphrase pour la linguistique : «*a display of paraphrasing abilities would establish a crucial connection between grammars and human linguistic organization ... It is surprising that the psycholinguists have failed to study this question very extensively*». Cependant, ce même livre se limite à l'étude des paraphrases très spéciales : celles des composés nominaux en anglais (du type *glass bird-house* = *a bird-house made of glass*, *glass-bird house* = *a*

house for glass-birds, *house-glass bird* = *a bird made from house-glass*, ...), en vue d'éclaircir les procédures d'interprétation grammaticale des composés. De telles paraphrases ne sont pas (ou ne sont que très rarement) employées par les locuteurs dans les circonstances normales du comportement langagier; ce sont des paraphrases métalinguistiques, qui ne nous intéressent pas ici (voir *infra*, pp. 13-14). À notre connaissance, la psycholinguistique est restée, jusqu'à nos jours, indifférente à l'étude de la paraphrase : un examen rapide de six revues majeures de psychologie (*Cognition*, 1983-1989; *Cognition and Brain Theory*, 1981-1989; *Cognitive Psychology*, 1982-1989; *Journal of Verbal Learning and Verbal Behavior*, 1982-1989; *Psychological Bulletin*, 1983-1989; and *Psychological Review*, 1983-1989) a révélé l'absence totale d'articles consacrés à la paraphrase, bien que ces revues abondent de matériaux et de sujets linguistiques (traitant surtout de reconnaissance et de compréhension de la parole).

³ (Page 11.) Pour préciser, sinon définir, la notion de synonymie, on a très souvent recours à la substituabilité mutuelle *salva significatione et correctione* [«préservant le sens et la

grammaticalité du texte³]. On dit que X est synonyme de Y si et seulement si X et Y sont mutuellement substituables dans tous les contextes sans affecter le sens et la qualité du texte en quoi que ce soit. Mais dans la vaste majorité des cas, les expressions linguistiques les plus synonymes ne sont pas substituables l'une par l'autre sans affecter la qualité du texte! Il suffit de se souvenir du fameux cas *gravement* [*malade*] vs *grièvement* [*blessé*], où *gravement* et *grièvement* ne peuvent pas être échangés, bien qu'ils soient parfaitement synonymes. (Des exemples encore plus révélateurs abondent en morphologie. Signalons, en particulier, les suffixes du pluriel en anglais — ou dans n'importe quelle autre langue — qui, tout en étant parfaitement synonymes, ne peuvent jamais se substituer l'un à l'autre : *box* ('boîte') ~ *box + es* (**box + en*), mais *ox* ('bœuf') ~ *ox + en* (**ox + es*).) On ne doit parler de substituabilité des expressions synonymes que *salva significatione* ('préservant le même sens'), avec la violation possible des contraintes syntaxiques, des restrictions sur la cooccurrence lexicale, etc. Mais alors une telle formulation ne peut servir de définition de la synonymie, étant donné qu'elle renvoie à la notion du 'même sens'.

Beaucoup plus prometteuse nous semble l'approche fondée sur les inférences logiques : les phrases P₁ et P₂ sont jugées synonymes si et seulement si l'ensemble de toutes les inférences que P₁ entraîne et celui de toutes les inférences que P₂ entraîne sont identiques. Malheureusement, cette approche n'est pas, elle non plus, sans problèmes. Le sujet est, cependant, trop spécial et trop éloigné de la ligne centrale de notre présentation pour être discuté ici plus en détail (cf., par exemple, Smith et Wilson 1980 : 150-171).

⁴ (Page 11.) Ne sont pas non plus considérées comme paraphrases les paires de phrases du type (i) ou (ii) :

- (i) *Le résultat est la somme de 7 et 3.*
vs
Le résultat est le produit de 2 et 5.
- (ii) *Le Général s'est retiré définitivement en 1969 à Colombey-les-deux-Églises.*
vs
De Gaulle a quitté définitivement le pouvoir après le référendum de 1969 [exemple tiré de Martin 1976 : 82].

Bref, la paraphrase dite pragmatique ou situationnelle est laissée hors de notre propos.

⁵ (Page 13.) Cela n'est vrai que si l'on ignore les types de texte très particuliers, par exemple, les textes juridiques ou mathématiques, qui visent, de façon très consciente, une précision extrême. Il en est de même des jeux de mots, des calembours, des dictions artistiques, des expressions métaling-

guistiques, etc., où il est parfois essentiel d'observer des distinctions sémantiques très subtiles ou même où l'on fait jouer des distinctions de forme. Nos raisonnements ne valent que pour des textes du type «neutre», textes qui ne relèvent que de l'emploi le plus banal de la langue.

⁶ (Page 23.) On remarquera qu'à la différence des deux composantes intermédiaires, les deux composantes «extrêmes» du MST — la sémantique et la phonologie — ne sont pas divisées en des sous-composantes profonde et de surface. Ce qu'on pourrait appeler la «phonologie de surface» (la phonétique, dans l'usage courant) est un mécanisme qui établit la correspondance entre la RPhoS (c'est-à-dire une transcription phonétique) et les sons réels de la parole. Nous pensons que ce mécanisme se trouve en dehors du MST.

⁷ (Page 24.) Cette estimation se fonde sur la considération suivante : le nombre de mots vedettes dans un gros dictionnaire de langue est, en règle générale, de l'ordre de 10⁵ (de 20 à 100 mille mots); le nombre moyen d'acceptions (= de significations) par mot vedette est 5; le produit se situe donc entre 100 mille et un million.

⁸ (Page 26.) Le présent état du MST et, en particulier, du système de paraphrasage mentionné, est tel qu'il peut être informatisé. Boyer et Lapalme (1985a, b) proposent un programme d'ordinateur qui réalise le modèle partiellement (la composante sémantique et la composante syntaxique; la composante morphologique a été remplacée par des procédures *ad hoc*). En s'aidant d'un petit dictionnaire du type DEC, ce programme produit des paraphrases françaises pour une RSém donnée. À leur suite, A. Polguère (Iordanskaja et Polguère 1988) a exploré les procédures de transition entre les RSém et les RSyntP, permettant de mettre en jeu de véritables règles sémantiques du MST. Le trait caractéristique de cette approche est l'exploitation simultanée de la SSém et de la SSém-Comm.

⁹ (Page 27.) Les raisons pour lesquelles les langues naturelles affichent cette richesse synonymique sont multiples. Premièrement, étant donné de nombreuses contraintes et restrictions sur la combinatoire des signes linguistiques, la langue doit offrir au locuteur beaucoup de souplesse pour qu'il soit en état de continuer son énoncé, en contournant tous les obstacles dont elle parsème son chemin. Deuxièmement, la variété synonymique sert à des fins expressives, c'est-à-dire artistiques. Troisièmement, la paraphrase aide souvent à mieux saisir le contenu : la même idée, mais exprimée en d'autres mots, semble plus claire, etc. Remarquons qu'en mathématiques, la synonymie joue un rôle primordial : des expressions comme $3 \times 4 = 12$, $a(b + c) = ab + ac$, ou «Le carré de l'hypoténuse est égal à la somme des carrés des deux autres côtés» (Pytha-

gore), ainsi que beaucoup d'autres règles, lois et théorèmes, ne sont pas autre chose qu'une manifestation de la synonymie!

¹⁰ (Page 28.) Le lecteur familier avec l'œuvre sémantique de A.Wierzbicka (1972, 1980, 1985, 1987, 1991) notera la différence entre son approche du problème d'éléments sémantiques primitifs et la nôtre. Wierzbicka commence par la recherche de ces éléments, elle en établit quinze : 'want', 'not want', 'think of', 'imagine', 'say', 'become', 'be part of', 'place', 'time', 'something', 'someone', 'I', 'you', 'world', 'this' (Wierzbicka 1987 : 31), et elle les applique à l'analyse sémantique de centaines de lexèmes, de morphèmes et de constructions syntaxiques de plusieurs langues naturelles. (Plus tard, Wierzbicka a élargi son inventaire de primitifs sémantiques de façon considérable : leur nombre a atteint une trentaine; voir Wierzbicka 1991 : 8.) Nous, par contre, commençons par l'analyse sémantique des lexèmes et d'autres unités linguistiques en termes d'unités toujours plus simples, en espérant ainsi aboutir aux éléments sémantiques primitifs (= nos sèmes) en tant que résultat final de notre recherche. Pour plus de détails, voir Mel'čuk 1989.

¹¹ (Page 28.) De telles décompositions peuvent être déclenchées selon les besoins de la manipulation du texte en question : par exemple, pour la recherche des équivalents dans une autre langue ou pour l'extraction de certaines informations encyclopédiques, l'accès direct à des composantes sémantiques de plus en plus profondes d'une signification lexicale peut être nécessaire.

¹² (Page 30.) Bien que les articles du DEC soient en principe applicables à une RSém, une telle application ne peut être réalisée immédiatement, sans efforts considérables. En effet, il s'agit de développer des règles formelles de type procédural, en vue de manipuler les articles de dictionnaire et de les utiliser lors du passage d'une RSém aux RSyntP correspondantes. De telles règles sont étudiées par A. Polguère (voir l'article mentionné Iordanskaja et Polguère 1988).

¹³ (Page 32.) Quand nous disons que la valeur de la FL Oper₁ (et des FL qui lui sont apparentées, comme Oper₂, Func_{0,1,2} ou Labor_{12,21}, voir p. 129) est un verbe sémantiquement vide, trois commentaires s'imposent.

Primo, ce verbe n'est pas nécessairement vide comme tel en français : dans d'autres contextes, il peut bien avoir un sens (ou des sens). Ainsi, JETER ou EMBRASSER ne sont pas, bien entendu, intrinsèquement vides; ils ne le deviennent que dans les contextes du type *jeter un regard* ou *embrasser un point de vue*, c'est-à-dire là où ils représentent les valeurs des FL considérées.

Secundo, ce verbe peut (totalement ou partiellement) garder son sens même en tant que valeur d'une FL du type Oper_i.

Ainsi, le verbe ÉPROUVER 'ressentir' a ce même sens dans éprouver un besoin (*de la tendresse*), où il est Oper₁; de façon similaire, FAIRE dans faire une action (*un mouvement*) veut dire 'faire'. (Comme on le voit, le sens d'un Oper₁ peut être très abstrait et général, comme 'faire', ou bien assez concret et spécifique, comme 'éprouver'.) Cependant, même dans les exemples cités, le sens du verbe remplissant le rôle de Oper_i, etc., répète, pour ainsi dire, une partie du sens du mot clé (= de l'argument de la FL); par conséquent, ce verbe ne contribue en rien à la RSém de l'expression entière, et c'est pour cette raison qu'on le qualifie de vide.

Tertio, la valeur de Oper_i pour un mot clé donné peut comprendre des verbes qui s'opposent sémantiquement, donc qui ne sont pas complètement vides. Par exemple, Oper₁(colère) est être [en] et éprouver (ressentir); mais être en colère diffère sémantiquement d'éprouver (ressentir) de la colère! Pour le voir, il suffit de comparer la phrase normale (i) :

- (i) [- Laisse-la tranquille!] Tu vois bien qu'elle est en colère!

à la phrase bizarre (ii) :

- (ii) [- Laisse-la tranquille!] ?Tu vois bien qu'elle éprouve (ressent) de la colère!

Le caractère déviant de la phrase (ii) est dû au fait qu'éprouver actualise l'émotion, et qu'il n'est pas facile de voir une émotion; être [en] implique plutôt l'état général (de la personne) causé par une émotion en plus de l'émotion elle-même; et seul l'état général, incluant certaines manifestations physiques, est bien visible. Dans de tels cas, les expressions de la valeur de Oper₁ qui portent une charge sémantique doivent être munies de distinceurs sémantiques :

être [en colère] [état général et/ou émotion];
éprouver, ressentir [de la colère] [émotion].

Il est évident qu'ici les valeurs des FL du type Oper₁ ne sont pas complètement vides.

N.B. : Il est possible que la solution suivante du problème des Oper_i non vides soit préférable. Un nom du type COLÈRE sera défini comme '(α) Émotion de X ... ou (β) état général de X causé par cette émotion'. Étant donné une telle définition disjonctive, on pourrait postuler qu'être [en] vaut pour le sens global de COLÈRE, tandis qu'éprouver et ressentir doivent être munis de la restriction suivante : «seulement pour la première partie de la définition», c'est-à-dire «pour COLÈRE(α)». Le problème exige une discussion plus poussée.

¹⁴ (Page 35.) La description du système de paraphrasage qui suit reprend, en grande partie, celle de Žolkovskij et

Mel'čuk 1967 (bien entendu, avec beaucoup de modifications).

¹⁵ (Page 35.) Comme l'a signalé A. Žolkovskij (1977), la division des règles lexicales de paraphrasage en ces deux sous-groupes — A.I (= règles qui s'appliquent aux lexies «indépendantes») et A.II (= celles qui s'appliquent aux fonctions lexicales) — n'est pas nécessaire du point de vue logique. Au contraire, il serait même plus logique de n'avoir qu'un seul ensemble de règles où la manipulation des fonctions lexicales serait effectuée par les mêmes règles que la manipulation des lexies indépendantes. En effet, au niveau de la SSyntP, les FL et les lexies indépendantes sont égales en tant qu'éléments du lexique profond; leurs différences se font sentir seulement au niveau de la SSyntS — au moment de calculer les valeurs des FL; on n'a donc pas de raisons substantielles pour les traiter différemment lors du paraphrasage. Par conséquent, on pourrait représenter le remplacement $\text{Oper}_1(C_0) \Leftrightarrow \text{Oper}_2(C_0)$ comme un cas particulier du remplacement $C_0 \Leftrightarrow \text{Conv}_{321}(C_0)$, puisque $\text{Oper}_2(C_0) = \text{Conv}_{321}(\text{Oper}_1(C_0))$, cf. *Jean donne un coup de pied à Pierre ~ Pierre reçoit un coup*

de pied de Jean. De même, le remplacement $\text{Oper}_1(C_0) \Leftrightarrow \text{Func}_1(C_0)$ apparaît comme un cas particulier du remplacement $C_0 \Leftrightarrow \text{Conv}_{21}(C_0)$, puisque $\text{Func}_1(C_0) = \text{Conv}_{21}(\text{Oper}_1(C_0))$, cf. *Jean donne à Pierre l'ordre de partir ~ L'ordre, pour Pierre, de partir vient de Jean*. Et ainsi de suite. Comme on peut le voir, une telle démarche réduit considérablement le nombre de règles dans le système de paraphrasage et, en même temps, explicite certaines généralisations importantes. Si nous ne l'avons pas suivie, c'est strictement pour les raisons de facilité de perception par le lecteur. La description des règles de paraphrasage est déjà tellement complexe que nous craignons qu'une extrême généralité empêche davantage la compréhension, de sorte que nous avons retenu la division des règles lexicales de paraphrasage en sous-groupes A.I et A.II.

¹⁶ (Page 39.) Une étude détaillée des fissions à verbe support en français contemporain est présentée dans G. Gross 1989. On y trouve de riches données et une bibliographie pratiquement exhaustive.

¹⁷ (Page 42.) Pour la FL composée MagnS_2 , voir Elnitsky et Mel'čuk 1988 : 77 ssq.

RÉFÉRENCES

- Apresjan, Jurij (1974). *Leksičeskaja semantika. Sinonimičeskie sredstva jazyka* [Sémantique lexicale. Moyens synonymiques de la langue]. Moscou : Nauka.
- Apresjan, Jurij (1980). *Tipy informacii dlja poverxnostno-semantičeskogo komponenta modeli "Smysl ↔ Tekst"* [Les types d'information pour la composante sémantique de surface du modèle «Sens-Texte»]. Vienne : Wiener Slawistischer Almanach.
- Bès, Gabriel et Catherine Fuchs, réd. (1988). *Lexique et paraphrase [= Lexique 6]*. Lille : Presses Universitaires de Lille.
- Bolinger, Dwight (1977). *Meaning and Form*. London — New York : Longman.
- Boyer, Michel et Guy Lapalme (1985a). Generating Sentences from Semantic Networks. In : V. Dahl, P. Saint-Dizier, eds., *Natural Logic Understanding and Logic Programming*, Amsterdam : North-Holland, 181-189.
- Boyer, Michel et Guy Lapalme (1985b). Generating Paraphrases from Meaning-Text Semantic Networks. *Computational Intelligence*, vol. 1, 103-117.
- Elnitsky, Léo et Igor Mel'čuk (1988). Le champ sémantique 'GRANDEUR' (\cong 'PARAMÈTRES') : description lexicographique de quelques cas problématiques de la cooccurrence lexicale (actants à valeur numérique et modificateurs de degré). In : Mel'čuk et al. 1988 : 73-80.
- Fuchs, Catherine (1980). *Paraphrases et théories du langage*. Paris : Université de Paris VII. (Thèse de doctorat d'État.)
- Fuchs, Catherine (1982). *La paraphrase*. Paris : PUF.
- Gentilhomme, Yves (1980). Un microsystème didactique. *Bulletin de linguistique appliquée et générale* (Université de Besançon), n° 7, 81-94.
- Giry-Schneider, Jacqueline (1978). *Les nominalisations en français. L'opérateur «faire» dans le lexique*. Genève — Paris : Librairie Droz.
- Gladkij, Aleksej et Igor Mel'čuk (1975). Tree Grammars : A Formalism for Syntactic Transformations in Natural Languages. *Linguistics*, vol. 50, 47-82. [Réimprimé dans : A. Gladkij et I. Mel'čuk, *Elements of Mathematical Linguistics*, 1983, Berlin etc. : Mouton, 151-187.]
- Gleitman, Lila R. et Henry Gleitman (1970). *Phrase and Paraphrase. Some Innovative Uses of Language*. New York : W.W. Norton.
- Gross, Gaston (1989). *Les constructions converses du français*. Genève — Paris : Librairie Droz.
- Gross, Maurice (1981). Les bases empiriques de la notion de prédicat sémantique. *Langages*, vol. 63, 7-52.

- Hartenstein, Klaus et Peter Schmidt (1983). Kommentierte Bibliographie zum «Smysl ↔ Tekst»-Modell. *Wiener Slawistischer Almanach*, vol. 11, 355-409.
- Hendriks, P. (1980). Lexical Functions in Albanian and Macedonian. In : A.A. Barentsen, B.M. Groen and R. Sprenger, eds., *Studies in Slavic and General Linguistics*, vol. 1, Utrecht : Rodopi, 161-170.
- Iordanskaja, Lidija et Alain Polguère (1988). Semantic Processing for Text Generation. In : *Proceedings of the International Computer Science Conference '88* (Hong-Kong), 310-318.
- Janus, Elżbieta (1971). *Pjat' pol'skix slovarnyx statej* [Cinq articles de dictionnaire polonais] : KLĘSKA, PASAZER, WALKA, ZWYCIEŻSTWO, ZWYCIEŻAĆ. *Naučno-texničeskaja informacija*, Serija 2, n° 3, 21-24.
- Leeman, Danielle, réd. (1973). *La paraphrase [= Languages*, 29]. Paris : Didier et Larousse.
- Martin, Robert (1976). *Inférence, antonymie et paraphrase*. Paris : Klincksieck.
- Mel'čuk, Igor' (1974). *Opty teorii lingvističeskix modelej "Smysl ↔Tekst"* [Essai d'une théorie de modèles linguistiques du type Sens-Texte]. Moscou : Nauka.
- Mel'čuk, Igor (1981). Meaning-Text Models : A Recent Trend in Soviet Linguistics. *Annual Review of Anthropology*, vol. 10, 27-62.
- Mel'čuk, Igor (1982). Lexical Functions in Lexicographic Description. *Proceedings of the 8th Annual Meeting of the Berkeley Linguistics Society*, 427-444.
- Mel'čuk, Igor (1988a). *Dependency Syntax : Theory and Practice*. Albany, N.Y. : SUNY Press.
- Mel'čuk, Igor (1988b). Paraphrase et lexique dans la théorie linguistique Sens-Texte. *Cahiers de lexicologie*, 52 : 1, 5-50, 53 : 2, 5-53.
- Mel'čuk, Igor (1988c). Paraphrase et lexique dans la théorie linguistique Sens-Texte. In : Bès et Fuchs 1988, 13-54.
- Mel'čuk, Igor (1989). Semantic Primitives from the Viewpoint of the Meaning-Text Linguistic Theory. *Quaderni di semantica*, 10 : 1, 65-102.
- Mel'čuk, Igor et al. (1984). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques I*. Montréal : Les Presses de l'Université de Montréal.
- Mel'čuk, Igor et al. (1988). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques II*. Montréal : Les Presses de l'Université de Montréal.
- Mel'čuk, Igor et Nikolaj Pertsov (1987). *Surface Syntax of English : A Formal Model Within the Meaning-Text Framework*. Amsterdam — Philadelphia : Benjamins.
- Mel'čuk, Igor et Alexander Zholkovsky (1984). *Explanatory Combinatorial Dictionary of Modern Russian* [en russe]. Vienne : Wiener Slawistischer Almanach.
- Mel'čuk, Igor et Aleksandr Žolkovskij (1970). Towards a Functioning 'Meaning-Text' Model of Language. *Linguistics*, vol. 57, 10-47.
- ML (1983) *Modèles linguistiques*, 5 : 3.
- Reuther, Tilmann (1978). Plädoyer für das Wörterbuch. *Linguistische Berichte*, vol. 57, 25-48.
- Reuther, Tilmann (1980). Nemeckie frazeologičeskie slovosčetaniya tipa *in völliger Verzweiflung sein* i ix russkie ěkvivalenty [Locutions allemandes du type *in völliger Verzweiflung sein* et leurs équivalents russes]. *Wiener Slawistischer Almanach*, vol. 5, 207-219.
- Sheintuch, Gloria (1976). Periphrastic Verb Formation in Persian — a Dynamic Process. In : Christie, ed., *Current Progress in Historical Linguistics*, Amsterdam : North-Holland, 139-146.
- Smith, Neil et Deirdre Wilson (1980). *Modern Linguistics. The Results of Chomsky's Revolution*. Bloomington, IN : Indiana University Press.
- Wierzbicka, Anna (1972). *Semantic Primitives*. Frankfurt : Athenäum.
- Wierzbicka, Anna (1980). *Lingua Mentalis*. Sydney etc. : Academic Press.
- Wierzbicka, Anna (1985). *Lexicography and Conceptual Analysis*. Ann Arbor, MI : Karoma.
- Wierzbicka, Anna (1987). *English Speech Verbs : A Semantic Dictionary*. Sydney etc. : Academic Press.
- Wierzbicka, Anna (1991). *Cross-Cultural Pragmatics. The Semantics of Human Interaction*. Berlin — New York : Mouton de Gruyter.
- Žolkovskij, Aleksandr (1970). Materialy k russko-somalijskomu slovarju [Matériaux pour un dictionnaire russe-somalien]. *Mašinnyj perevod i prikladnaja lingvistika*, vol. 13, 35-63.
- Žolkovskij, Aleksandr (1971). *Sintaksis somali* [La syntaxe du somalien]. Moscou : Nauka.
- Žolkovskij, Aleksandr (1977). O nekotoryx popravkax k sisteme perifrazirovaniya (zamečanija o pravilax s **AdvReal**) [Sur quelques corrections dans le système de paraphrasage (remarques sur les règles avec **AdvReal**)]. *Naučno-texničeskaja informacija*, Serija 2, n° 5, 28-30.

Žolkovskij, Aleksandr et Igor' Mel'čuk (1965). O vozmožnom metode i instrumentax semantičeskogo sinteza [Sur une méthode possible de synthèse sémantique et les outils correspondants]. *Naučno-texničeskaja informacija*, n° 5, 23-28.

Žolkovskij, Aleksandr et Igor' Mel'čuk (1967). O semantičeskom sinteze [Sur la synthèse sémantique]. *Problemy kibernetiki*, 19, 177-238. [Une traduction française : *T.A. Informations*, 1970, n° 2.]