

1988

Paraphrase et lexique dans la théorie linguistique Sens-Texte

Igor Mel'čuk

1. INTRODUCTION

Nous aimerions commencer par énoncer la thèse suivante :

La linguistique théorique présuppose nécessairement une théorie de la paraphrase langagière.

Décrire une langue \mathcal{L} signifie alors construire un système de règles formelles qui manifeste la même capacité de produire des textes de \mathcal{L} pour un sens donné, entre autres d'effectuer des paraphrases en \mathcal{L} , qu'ont les locuteurs de \mathcal{L} .

La conception du rôle exceptionnel de la paraphrase est profondément ancrée dans une théorie linguistique particulière : la théorie Sens-Texte. Cette théorie vise la description d'une langue sous 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 le pivot. Le modèle Sens-Texte (dorénavant, MST) procède d'une représentation du sens — d'une REPRÉSENTATION SÉMANTIQUE (RSém), qui est une représentation formelle de l'invariant d'un ensemble de paraphrases (plus ou moins) synonymes. A 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 PARAPHRASES, et il fait partie de la composante sémantique du modèle. Enfin, toutes les informations lexicales (c'est-à-dire concernant les lexèmes individuels) nécessaires au fonctionnement de ce système sont emmagasinées dans un dictionnaire spécial : 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 faudra considérer au moins ces quatre points : 1) modèle Sens-Texte ; 2) représentation sémantique ; 3) système de paraphrases ; 4) dictionnaire explicatif et combinatoire. Chaque point sera traité dans une section à part (les sections 2-5).

La théorie Sens-Texte n'est pas une nouveauté linguistique. Ses fondements ont été jetés dans les années 60, à Moscou, par A. Zholkovsky (présentement, à l'University of Southern California, Los Angeles) et

l'auteur de ces lignes (Žolkovskij & Mel'čuk (1965, 1967), Mel'čuk & Žolkovskij (1970)). 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 & Žolkovskij (1984)), dans une description détaillée de la syntaxe de surface de l'anglais (Mel'čuk & Pertsov (1986)) et dans deux livres exposant la syntaxe de dépendance (Mel'čuk (1979) et (1988)). En plus, quelque 150 articles ont paru dans des périodiques linguistiques ; voir les bibliographies dans Mel'čuk (1981) et dans Hartenstein & Schmidt (1982). Enfin, le travail portant sur un DEC du français contemporain a été commencé en 1979 ; le premier volume a paru en 1984 (Mel'čuk & *al.* (1984)), et le deuxième est sous presse.

Comme la documentation sur la théorie Sens-Texte est accessible, nous nous limiterons ici à un exposé plutôt dogmatique, laissant de côté plusieurs détails, justifications et explications. En particulier, nous ne toucherons pas à la notion même de paraphrase. Pour la discussion théorique, nous renvoyons le lecteur à Leeman (1973), Martin (1976), Fuchs (1980, 1982), Fuchs & Le Goffic (1983, 1985), *ML* (1983) (pour mieux évaluer l'état de la question, on peut avoir recours à la bibliographie donnée dans *ML* (1983 : 191-192)).

2. MODÈLE SENS-TEXTE

2.1. *Les postulats de base*

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

Postulat 1

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

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

$$(1) \quad \{ \text{SENS}_i \} \xleftrightarrow{\text{langue}} \{ \text{TEXTE}_j \} \quad | \quad M \ 0 < i, j, < \infty$$

Un sens relève du côté intérieur de la parole ; c'est un phénomène psychique — une information transmise (ou à transmettre) dans un événement langagier. Un texte relève du côté extérieur 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 deux clarifications importantes.

Primo, le terme *sens* doit être interprété ici de la façon la plus étroite possible. Il ne s'agit aucunement du sens que nous obtenons comme résultat d'une bonne compréhension d'un énoncé quelconque, que nous

en extrayons grâce à la logique, à nos connaissances extralinguistiques, etc. Dans (1), nous ne voyons que le sens purement langagier : le plus superficiel, le plus littéral, celui qui est accessible uniquement grâce à la maîtrise de la langue en cause. C'est seulement ce sens-là qui nous intéresse et que nous définissons comme suit :

Le sens est l'invariant des paraphrases langagières, c'est-à-dire la seule propriété commune à tous les énoncés ayant le même sens.

Secundo, bien que les sens et les textes aient leur existence réelle, le MST, modèle linguistique par excellence, ne doit et ne peut pas les traiter dans leur réalité psychique (= neurologique) et physique. Ce 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 phonétique* (RPhonét).

Etant donné le rôle primordial des représentations dans le MST, nous pouvons réécrire (1) comme (2) :

$$(2) \quad \{RSém_i\} \xleftrightarrow{\text{langue}} \{RPhonét_j\} \quad 0 < i, j < \infty$$

Postulat 2

L'application (1)/(2) doit être représentée par un dispositif logique ou, plus précisément, par un modèle cybernétique, ou 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 un *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).

Nous avons déjà indiqué que la correspondance entre les RSém_i et les RPhonét_j est multi-multivoque. 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 (3), tirée d'un magazine français :

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

pour laquelle nous citons plus loin des représentations illustratives, a entre deux et trois millions de paraphrases : voir la section 3, pp. 24-25. La synonymie est donc extrêmement riche dans les langues naturelles. D'autre part, comme on le sait bien, un même texte peut correspondre à plusieurs sens ; ainsi, la phrase (4), citée dans Gentilhomme (1980 : 85-93) :

- (4) 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, 9072 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 donc également une caractéristique des langues naturelles.

La complexité extraordinaire de la correspondance $\{RSém_i\} \Leftrightarrow \{RPhonét_j\}$ nous pousse à postuler des niveaux de représentation INTERMÉDIAIRES, qui permettraient de la diviser en composantes plus simples :

Postulat 3

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

Par conséquent, nous pouvons réécrire (2) sous une forme plus développée, soit (5) :

$$(5) \quad \{RSém_i\} \xleftrightarrow{\text{sémantique}} \{RSynt_k\} \xleftrightarrow{\text{syntaxe}} \{RMorph_j\} \xleftrightarrow{\text{morphologie} + \text{phonologie}} \{RPhonét_j\}$$

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 :

$$(6) \quad \begin{array}{ccccccc} \text{sémantique} & & \text{syntaxe} & & \text{syntaxe} & & \text{morphologie} \\ & & \text{profonde} & & \text{de surface} & & \text{profonde} \\ \{RSém_i\} & \Leftrightarrow & \{RSyntP_{k_1}\} & \Leftrightarrow & \{RSyntS_{k_2}\} & \Leftrightarrow & \{RMorphP_{j_1}\} & \Leftrightarrow & \\ & & \text{morphologie} & & \text{phonologie} & & & & \\ & & \text{de surface} & & & & & & \\ & & \{RMorphS_{j_2}\} & \Leftrightarrow & \{RPhonétP_{j_1}\} & \Leftrightarrow & \{RPhonétS_{j_2}\} & & \end{array}$$

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 (6), 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.

2.2. Représentations linguistiques des énoncés

Etant donné les contraintes d'espace, nous caractériserons les niveaux de représentations linguistiques adoptés dans le MST sans définitions rigoureuses — à partir d'exemples accompagnés de quelques commentaires. Nous travaillerons à partir de la phrase (3), pour laquelle nous présentons d'abord la RSém et ensuite la RSyntP. En même temps, nous citerons la RSyntP d'une phrase synonyme de (3), afin de montrer les possibilités de production des paraphrases dont le MST dispose. 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.

Représentation sémantique

La nature et l'organisation de la RSém dans le cadre du MST seront spécialement discutées dans la Section 3 ; ici, nous nous limiterons à l'exemple qui suit : Figure 1. 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 (سلسله) marque les frontières de thème (= **L**) et de rhème (= **R**), voir plus bas, p. 25.
3. Le soulignement d'un élément sémantique (par exemple, 'devenir' dans B2) indique son rôle assertif : cet élément est posé (= affirmé), et non présupposé, comme les autres.
4. En lisant la RSém (7) [= Fig. 1], il faut tenir compte du fait que chaque nœud dont partent des flèches représente un élément prédicatif (au sens logico-sémantique) ; si un tel nœud est étiqueté d'un nom ou d'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 acceptions voulues des mots polysémiques. (Ces numéros ont été empruntés au *Petit Robert*.)

(7)

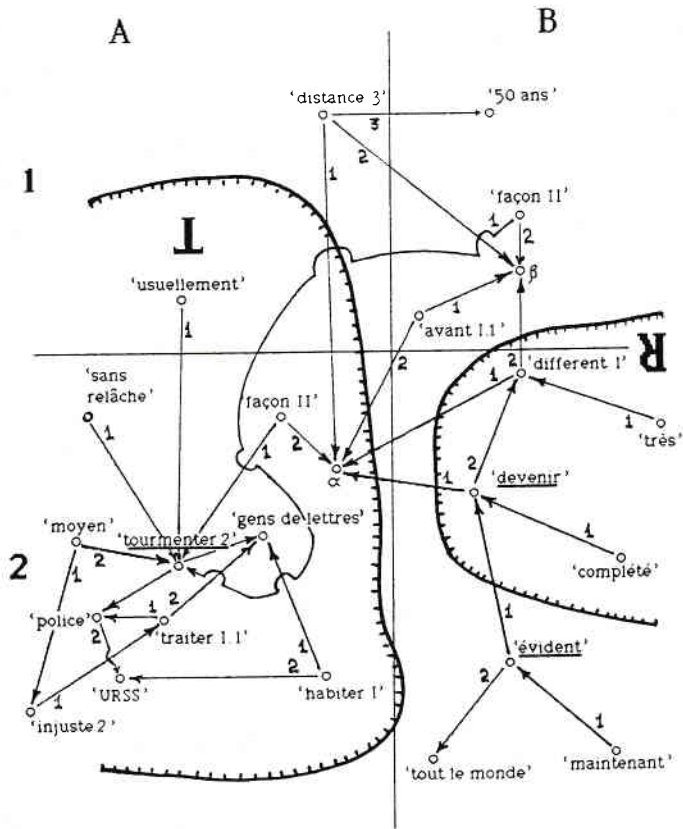


Figure 1

La représentation sémantique de la phrase (3)
[et de toutes les phrases qui lui sont synonymes]

La RSém (7) peut se lire en français de la façon suivante :

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

Représentation syntaxique profonde

Au niveau de la syntaxe profonde, la phrase (3) apparaît comme dans la Figure 2.

Pour donner une meilleure idée de la RSyntP, nous présenterons encore une RSyntP, cette fois pour la phrase (8), qui est synonyme (= une paraphrase) de la phrase (3) :

- (8) Chose certaine, la manière dont la police soviétique persécute les gens de lettres est maintenant très différente de celle d'il y a 50 ans.

Les phrases (3) et (8) ont la même RSém : celle de la Figure 1 ; par conséquent, les RSyntP des Figures 2 et 3 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 unités lexicales).

— Une structure syntaxique communicative profonde (SSynt-CommP), qui consiste, *grosso 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 pointillé spécifie la corréférence des unités lexicales (en l'occurrence, MANIÈRE).

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

Nous ne caractériserons plus en détail que 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.

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

(= $V(\text{erbe}) \xrightarrow[\text{fini}]{I} S(\text{ujet}) G(\text{rammatical})$) et à tous ses « transformés » (par exemple, *persécution* \xrightarrow{I} *policrière*, puisque 'La police persécute', etc.) ; la RelSyntP II correspond aux constructions avec

le complément d'objet direct CO^{dir}, la RelSyntP III, aux constructions avec le CO^{indir}, les RelSyntP IV-VI, aux constructions avec le CO^{oblique} ; ATTR représente toutes les constructions modificatrices (épithètes, compléments de nom, ...); COORD, toutes sortes de coordinations ; et APPEND, toutes les constructions « extrastructurales » (parenthétiques, d'adresse, ...).

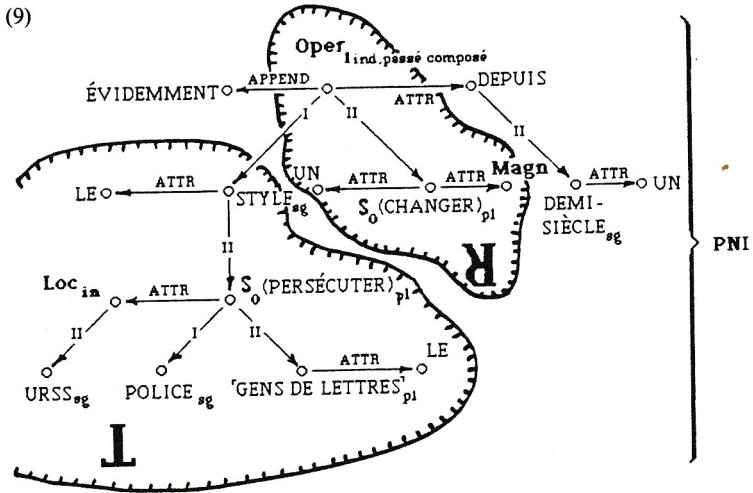


Figure 2

La représentation syntaxique profonde de la phrase (3)

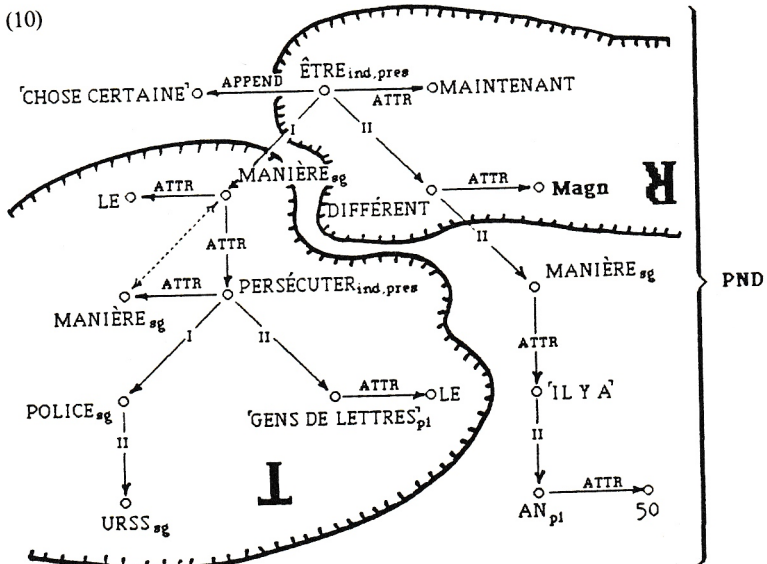


Figure 3

La représentation syntaxique profonde de la phrase (8)

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 la SSyntP :

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

(ii) Un idiotisme, ou phrasème (= expression multilexémique ayant des propriétés imprévisibles à partir des propriétés de ses constituantes) 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 à la surface syntaxique 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. : tous les verbes en majuscules sont idiomatiquement liés par le nom correspondant et représentent la valeur de ce que nous appelons la fonction lexicale Oper₁. Cf. :

fonction	argument	valeur
Oper ₁	(<i>changement</i>)	= <i>connaître</i>
Oper ₁	(<i>lutte</i>)	= <i>mener</i>
Oper ₁	(<i>plainte</i>)	= <i>porter</i>

(Oper₁ n'est pas autre chose qu'un des *verbes supports* bien connus de la linguistique française : voir, par exemple, Gross (1981)). Les fonctions lexicales (FL), phénomène de la syntaxe profonde, sont universelles (en ce sens qu'on les trouve dans toutes les langues) ; leur nombre est d'une cinquantaine. Les FL forment un système très cohérent et sont sous-jacentes au système de la paraphrase (section 4) ; c'est dans la section 4 que nous présenterons une liste de FL avec quelques explications supplémentaires. 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* $\xrightarrow{\text{II}}$ *changements*, nous écrivons Oper₁ $\xrightarrow{\text{II}}$ S₀ (*changer*) [S₀ étant un nom d'action], et au lieu de *changements* $\xrightarrow{\text{ATTR}}$ *sérieux*, S₀ (*changer*) $\xrightarrow{\text{ATTR}}$ Magn [Magn étant un intensificateur consacré par l'usage] ; voir Figure 2.

Nous ne toucherons pas aux niveaux de représentation plus proches de la surface, puisqu'ils sont peu pertinents pour la paraphrase, telle que conçue dans le cadre du MST.

2.3. *Les composantes du modèle Sens-Texte*

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

- 1) la sémantique ;
- 2) la syntaxe profonde ;
- 3) la syntaxe de surface ;
- 4) la morphologie profonde ;
- 5) la morphologie de surface ;
- 6) la phonologie profonde ;
- 7) la phonologie de surface ;

dont nous ne caractériserons que les deux premières (le reste étant marginal du point de vue de la paraphrase).¹

La sémantique

Cette composante établit la correspondance $\{RSém_i\} \iff \{RSyntP_k\}$ au moyen de huit opérations, que nous ne décrirons pas ici (par exemple, sectionnement de la RSém de départ en des « morceaux » dont la taille correspond à la taille sémantique d'une phrase ; réunion des unités sémantiques en faisceaux correspondant aux 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* ;

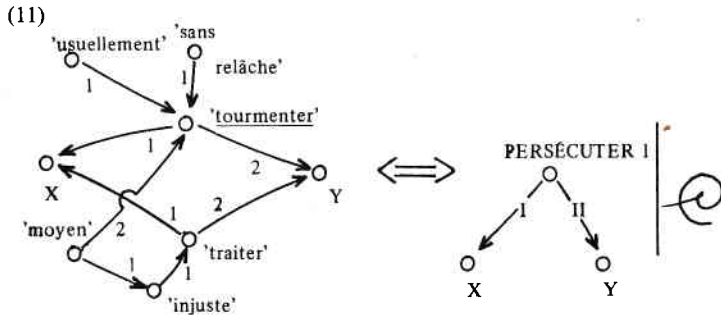
— les *règles sémantiques de paraphrase* qui font correspondre à une RSyntP de base toutes les RSyntP qui en sont synonymes, et *vice versa*.

Ainsi nous établissons la correspondance entre les deux ensembles : $\{RSém_i\}$ et $\{RSyntP_k\}$ — en deux temps.

Les règles sémantiques de paraphrase, ou le système de paraphrases, sont décrites dans la section 4. Dans la présente section, nous ne traiterons que des règles sémantiques au sens propre et n'en illustrerons que trois types (il y en a beaucoup plus !).

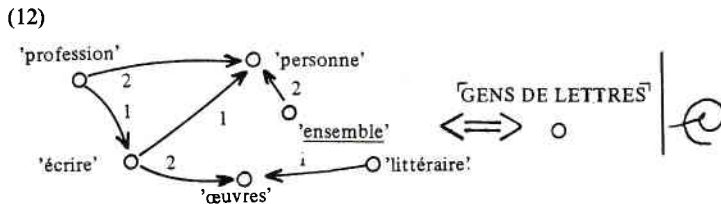
Les règles du MST citées ici appartiennent à celles nécessaires à la construction de la phrase (8) ; ainsi le lecteur sera plus en mesure de vérifier leur application.

- Une règle lexico-sémantique



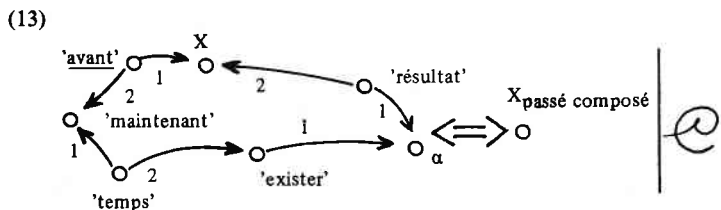
Une telle règle n'est pas autre chose que l'article de dictionnaire pour le lexème en question (= PERSÉCUTER 1), la partie de gauche étant la définition, et l'arbre dans la partie de droite représentant (très approximativement) le régime. Un article de dictionnaire contient encore les données sur la combinatoire lexicale du lexème vedette, c'est-à-dire ses fonctions lexicales. Sous forme d'une règle lexico-sémantique, ces données devraient constituer la partie conditions (de la règle), que nous avons notée par \textcircled{e} pour raisons de simplicité. (Le lecteur trouvera quelques articles de dictionnaire — du type adopté dans le MST, c'est-à-dire du DEC — dans la section 5.)

- Une règle phraséologico-sémantique



Une règle phraséologico-sémantique est aussi un article de dictionnaire, cette fois-ci pour le phrasème en question.

- Une règle morphologico-sémantique

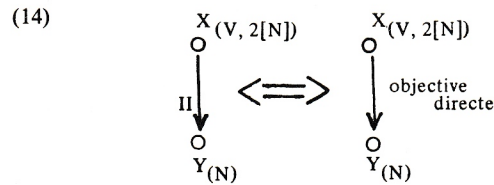


Cette règle représente une des significations du passé composé français (= 'une action passée dont le résultat existe à présent').

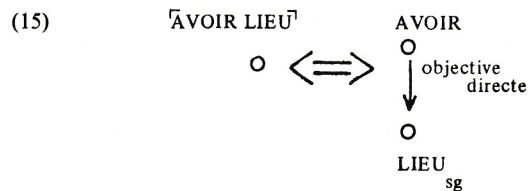
La syntaxe profonde

La composante syntaxique profonde établit la correspondance $\{\text{RSyntP}_{k_1}\} \leftrightarrow \{\text{RSyntS}_{k_2}\}$ au moyen de cinq opérations effectuées par les règles syntaxiques profondes, dont trois seront citées en tant qu'illustration.

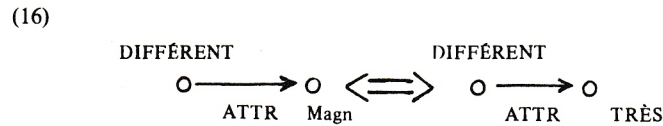
- Une règle syntaxique profonde de structuration



- Une règle syntaxique profonde phraséologique (impliquant un phrasème)



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



3. LA REPRÉSENTATION SÉMANTIQUE

Nous avons déjà dit que le sens linguistique est l'invariant des paraphrases (synonymes) ; la représentation sémantique est alors un moyen formel de décrire cet invariant — dans le but d'en dériver l'ensemble des paraphrases possibles. Voyons de plus près ce moyen.

La phrase (3), qui nous a servi de base, et sa paraphrase (8) admettent beaucoup d'autres paraphrases ; en voici encore deux exemples :

- (17) a Bien entendu, le style de persécution des gens de lettres soviétiques par la police a sérieusement changé au cours des 50 dernières années.
- b 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 grossièrement, le nombre total de paraphrases dans cette famille, nous diviserons les phrases (3), (8) et (17) en « tranches » sémantiques et nous indiquerons des variantes d'expression synonymes ou quasi synonymes pour chaque tranche ; voir la Figure 4.

Le chiffre en gras en bas d'une colonne de variantes indique leur nombre. Supposant que chacune des variantes dans une colonne peut se combiner avec n'importe quelle variante d'une autre colonne, nous pouvons, tout simplement, multiplier ces chiffres :

$$15 \times 3 \times 5 \times 3 \times 3 \times 15 \times 9 \times 4 \times 4 = 4\,374\,000 \text{ paraphrases}$$

Beaucoup de combinaisons seront éliminées en raison de contraintes de cooccurrence ; cependant, on pourrait ajouter d'autres variantes au tableau de la Figure 4 (par exemple, *ont connu des modifications radicales* ~ *diffèrent beaucoup*, *aujourd'hui* ~ *maintenant*, *il y a 50 ans* ~ *50 ans plus tard* ~ *au cours des 50 années*, etc.). On a donc deux ou trois millions de paraphrases pour notre phrase de départ. Cela démontre clairement à quel point la synonymie des langues naturelles est riche.

La théorie Sens-Texte se donne, en tant que tâche tout à fait primordiale, une description formelle et exhaustive de la synonymie de phrases, c'est-à-dire de la richesse paraphrastique. Cette description présuppose, à son tour, la description de l'invariant sémantique des paraphrases observables c'est-à-dire la représentation sémantique (RSém).

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 spécifie les objets qui composent la situation décrite par nos paraphrases ainsi que toutes les relations entre eux ; voir plus loin.
- La structure sémantico-communicative (SSém-Comm) représente l'itinéraire que le locuteur choisit à travers la situation décrite : ce qu'il veut prendre comme point de départ, ce qu'il veut discuter d'abord et ce qu'il veut discuter ensuite, ce qu'il présuppose *vs* ce qu'il affirme, ce qu'il veut mettre en relief, etc. La division en thème et rhème (**L** *vs* **R**) marquée dans la Figure 1 appartient à la SSém-Comm, ainsi que le soulignement des deux nœuds [B2], qui signale l'affirmation ('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, etc.

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

E	F	G	H	I
a changé a { subi } des changements des changements sont survenus s'est transformé a { subi } des transformations des transformations sont survenues est devenu { différent } { autre } n'est plus le même { a été } modifié { s'est } a { connu } des modifications { subi }	bien beaucoup profondément sérieusement de façon { profonde } { sérieuse } tout à fait très (<i>différent</i>) tout (<i>autre</i>)	au cours de depuis il y a pendant	demi-siècle 50 { années } { ans } cinq décennies	
3	15	9	4	4

Figure 4
 Variantes d'expression pour les paraphrases
 des phrases (3), (8) et (17)

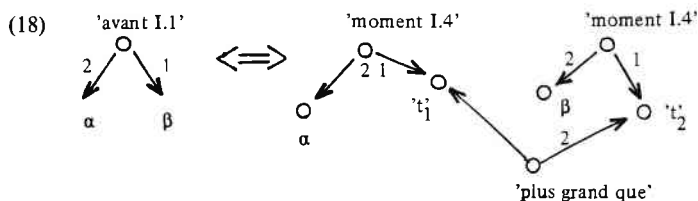
La SSém-Comm et la SRhét ne seront pas décrites ici malgré leur importance pour la paraphrase; nous ne nous occuperons que de la SSém.

La structure sémantique spécifie le sens de l'énoncé (= de tous les énoncés synonymes, donc paraphrasables) indépendamment de sa forme linguistique. Ainsi la distribution du sens en mots, propositions et phrases, son expression par des prosodies ou par des constructions syntaxiques particulières, etc., ne sont pas prises en considération dans la SSém. Entre autres, 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'.

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 au sens d'une unité lexicale, c'est-à-dire à une acception bien déterminée d'un mot ou d'un phrase-sème. Ce sens doit être identifié par un numéro distinctif qui accompagne l'étiquette lexicale et qui provient d'un dictionnaire. En principe, cela doit être le Dictionnaire explicatif et combinatoire — DEC, dont nous traiterons dans la section 5; mais avant que le DEC soit élaboré, ces numéros peuvent être pris dans un bon dictionnaire de la langue, par exemple, pour le français, dans *Le Petit Robert*. (Ainsi, 'façon II' est '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' est '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:



(Remarquons que (18) est une règle lexico-sémantique, étant donné que AVANT I.1 est bel et bien un lexème français. De cette façon, la décomposition sémantique d'un sémantème est donnée par le dictionnaire).

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 la décomposition sémantique 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'.²

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. 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 peuvent être remplacés par leurs réseaux correspondants) : jusqu'au niveau des sèmes.

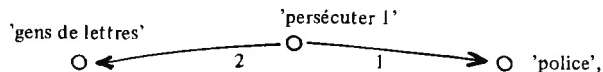
Nous distinguons deux grandes classes de sémantèmes :

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

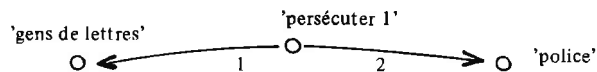
Dans la SSém, des variables sont également utilisées : voir, par exemple, α et β dans la Figure 1, A2 et B1.

Les arcs du réseau peuvent aboutir autant aux noms qu'aux foncteurs (= un sémantème quelconque, 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

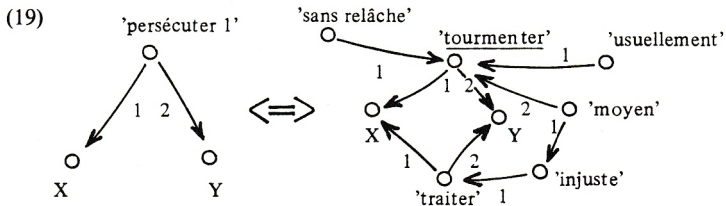


1 signifie seulement que 'police' est le premier argument de 'persécuter', et 2, que 'gens de lettres' en est le deuxième argument. L'écriture



signifierait 'Les gens de lettres persécutent la police'.

Le rôle exact 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 lexicographique) du foncteur : par exemple, (19) précise que le premier argument [= X] de 'persécuter 1' tourmente le deuxième [= Y] au moyen d'un traitement injuste usuel que X réserve à Y. En décomposant 'tourmenter', on verra que X 'cause' que Y subisse des souffrances physiques ou morales ; en décomposant 'souffrance', on montrera que Y 'éprouve' des états spéciaux de son organisme ; et ainsi de suite, jusqu'à ce qu'on aboutisse aux sèmes (= éléments sémantiques primitifs), 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 (19) sont les actants sémantiques de 'persécuter 1'. (Les expressions françaises, en l'occurrence les syntagmes nominaux, qui expriment les actants sémantiques dans la 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 persécute les gens de lettres* ou par rapport au nom *persécution* dans *les persécutions policières* < = *par la police* > *des gens de lettres*.)

4. LE SYSTÈME DE PARAPHRASES

4.1. Généralités

Les paraphrases des langues naturelles sont de deux types majeurs :
 — Les paraphrases SÉMANTIQUES sont produites par la distribution différente du sens de départ entre des unités lexicales ; donc deux paraphrases sémantiques se distinguent obligatoirement par au moins une unité lexicale ;

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

vs

Un territoire immense **est tombé sous son pouvoir**.

— Les paraphrases SYNTAXIQUES sont produites par le choix des constructions syntaxiques différentes ; deux paraphrases syntaxiques « pures » affichent le même jeu d'unités lexicales pleines mais qui sont arrangées différemment du point de vue syntaxique ;

- (21) ... [la somme] récupérée par le ministre
 vs
 ... [la somme] que le ministre a récupérée

Malgré l'intérêt incontestable des paraphrases syntaxiques pour la théorie de la paraphrase, nous laisserons les paraphrases syntaxiques de côté — pour mieux nous concentrer sur les paraphrases sémantiques.

Ces dernières ne sont pas, cependant, homogènes ; elles se laissent classer, comme nous l'avons déjà indiqué dans la section 2.3, p. 22, en deux sous-classes : celles qui sont effectuées par des règles sémantiques au sens propre et celles qui le sont par des règles de paraphrase. Les premières sont engendrées par différents choix lexicaux : en principe, la RSém de départ peut être articulée en sous-réseaux correspondant aux lexèmes de plusieurs façons. De telles paraphrases sont plus profondes que les autres ; elles sont extrêmement intéressantes pour la recherche sémantique mais pour le moment, elles sont très peu étudiées et semblent (peut-être justement à cause de cela) peu systématiques. Par contre, les secondes sont engendrées par des substitutions lexicales : elles exploitent les liens sémantiques existant entre les unités lexicales dans le lexique. Les paraphrases de ce type se sont avérées très riches et très systématiques en même temps ; nous parlons des paraphrases du type (22) :

- (22) 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 paraphrases qui sont traitées par le système de paraphrases faisant partie de la composante sémantique du modèle Sens-Texte ; dans ce qui suit nous ne nous occuperons que d'elles.

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

4.2. Fonctions lexicales

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

Quelles que soient les unités lexicales L_1 et L_2 , si $f(L_1)$ et $f(L_2)$ existent tous les deux, alors :

1. $f(L_1)$ et $f(L_2)$ affichent une même relation avec L_1 et L_2 , respectivement, en ce qui concerne le sens et le rôle syntaxique profond :

$$'f(L_1)' / 'L_1' = 'f(L_2)' / 'L_2'.$$

Cette condition ne dépend pas de la langue.

2. Au moins, pour certaines unités lexicales, $f(L_1) \neq f(L_2)$.

Cette condition dépend complètement de la langue : elle signifie qu'en \mathcal{L} , la valeur de f est phraséologiquement liée par l'argument.

L'importance de cette condition 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 puisque avec n'importe quelle unité lexicale il sera toujours exprimé par le lexème CHER II.1 : son expression ne dépend pas du lexème modifié. Pourtant le sens 'très' est une FL : '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. En fait, c'est la FL Magn, voir ci-dessous.

Parmi les FL, il convient de distinguer une sous-classe importante : ce que nous appelons les FL STANDARD SIMPLES. Ces FL satisfont deux conditions supplémentaires :

3. f est définie pour un grand nombre d'arguments. (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 tous les $f(L_i)$, pour beaucoup de L_i , est suffisamment riche.) Cette condition dépend complètement de la langue.

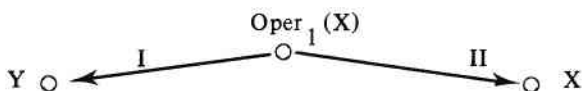
Par exemple, le sens 'additionné de crème' est une FL en français, puisqu'on dit *café crème* mais *fraises à la crème* < **café à la crème*, **fraises crème* >. Cependant, ce n'est pas une FL standard simple : les conditions 3 et 4 sont violées, étant donné que le sens en cause est très concret et qu'il n'a que deux expressions. 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') :

Magn(<i>riposte</i>)	= <i>foudroyante</i>	Magn(<i>sourd</i>)	= <i>comme un pot</i>
Magn(<i>cri</i>)	= <i>d'orfraie</i>	Magn(<i>fort</i>)	= <i>comme un Turc</i>
	Magn(<i>apprécier</i>)		= <i>grandement</i>
	Magn(<i>recourir [à qqch]</i>)		= <i>largement</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}) et le premier actant SyntP de X, comme son sujet grammatical (SG) :



Oper ₁ (<i>plainte</i>)	= <i>porter</i>	Oper ₁ (<i>pas</i>)	= <i>faire</i>
Oper ₁ (<i>cri</i>)	= <i>pousser</i>	Oper ₁ (<i>ordre</i>)	= <i>donner</i>
Oper ₁ (<i>grippe</i>)	= <i>avoir, faire</i>	Oper ₁ (<i>suprématie</i>)	= <i>détenir</i>
Oper ₁ (<i>désespoir</i>)	= <i>être au</i>	Oper ₁ (<i>regard</i>)	= <i>jeter</i>

Sans entrer dans les détails concernant les FL (voir Mel'čuk (1982) et Mel'čuk & al. (1984 : 6-13)), nous allons présenter ici une liste abrégée de FL (leur ordre dans cette liste correspond à l'ordre de leur apparition dans un article de dictionnaire du DEC).

Liste de fonctions lexicales

1. Syn, Syn_C, Syn_⊃, Syn_∩ [synonymes et quasi-synonymes : C signifie « plus étroit », ⊃ signifie « plus large », tandis que ∩ dénote une intersection des sens] : Syn(*espoir*) = *espérance* ; Syn_C(*respect*) = *vénération* ; Syn_⊃(*vénération*) = *respect* ; Syn_∩(*seconder*) = *assister II*.

2. Conv_{ijk} [co^oversif, ou un lexème qui dénote une relation converse de la relation exprimée par le mot clé de cette FL ; les indices montrent l'ordre des actants syntaxiques associés au conversif relativement à l'ordre des actants associés au mot clé du conversif, qui est toujours « 123 »] : Conv₂₁(*être derrière*) = *être devant* [*Je me trouvais derrière une vieille dame* = *Devant moi se trouvait une vieille dame*] ; Conv₂₁(*plus*) = *moins* ; Conv_{3214∩}(*vendre*) = *acheter*.

.....

8. S₀, V₀, A₀, Adv₀ [dérivés syntaxiques — nom, verbe, adjectif, adverbe] : S₀(*acheter*) = *achat* ; A₀(*école*) = *scolaire* ; V₀(*promesse*) = *promettre* ; Adv₀(*honnête*) = *honnêtement*.

9. S₁, S₂, S₃, ... [nom typique pour le premier, deuxième, troisième, ... actant sémantique du mot clé] : S₁(*crime*) = *criminel* ; S₂(*crime*) = *victime* [*du crime*] ; S₁(*acheter*) = *acheteur* ; S₂(*acheter*) = *marchandise* ; S₃(*acheter*) = (*vendeur*) ; S₄(*acheter*) = *prix*.

.....

32. Oper₁, Oper₂, ... [verbe sémantiquement vide qui prend le nom du premier, deuxième, ... actant de la situation 'C₀' comme son sujet grammatical (SG) et le mot clé C₀ comme son complément d'objet (CO) principal] : Oper₁(*attention*) = *faire* ; Oper₂(*attention*) = *attirer* ; Oper₁(*conseil*) = *donner* ; Oper₂(*conseil*) = *recevoir*.

33. Func₀, Func₁, Func₂, ... [verbe sémantiquement vide qui prend C₀ comme SG et, au cas où la situation C₀ a des actants, le nom du premier, deuxième, ... actant de C₀ comme son CO principal] : Func₀(*silence*) = *régner*, Func₁(*aide*) = *venir, provenir* [*de qqn*] ; Func₂(*liste*) = *contenir, comprendre* [*qch*].

34. Labor_{ij} [verbe sémantiquement vide qui prend le nom de l'actant *i* comme son SG, le nom de l'actant *j* comme son CO principal, et C₀ comme son deuxième CO] : Labor₁₂(*analyse*) = *soumettre* [*qch à l'analyse*] ; Labor₁₂(*soin*) = *entourer* [*qqn de soins*].

.....

39-41. Caus, Liqu, Perm[« causer » (= 'faire en sorte que quelque chose commence à avoir lieu'), « liquider » et « permettre », respectivement ; ces FL s'emploient le plus souvent en combinaison avec d'autres FL] ; CausOper₁(*désespoir*) = *pousser, réduire [qqn au désespoir], jeter [qqn dans le désespoir], frapper [qqn de désespoir]* ; CausFunc₀(*difficulté*) = *créer, poser* ; LiquFunc₀(*assemblée*) = *dissoudre* ; LiquFunc₂(*attention*) = *détourner [l'attention de qqn de qch]* ; Perm₁Fact₀(*colère*) = *s'abandonner [à la colère]* ; Caus₂Func₂(*attention*) = *accaparer* ; Liqu₁Func₂(*peur*) = *surmonter, vaincre [sa (propre) peur]* ; Liqu₁Func₀(*colère*) = *étouffer [sa (propre) colère]*.

42. Real₁, Real₂, ... [« réaliser » : verbe qui prend C₀ comme son CO principal et le nom du premier, deuxième, ... actant de C₀ comme son SG] : Real₁(*problème*) = *résoudre* ; Real₁(*piège*) = *tendre* ; Real₂(*piège*) = *tomber dans [le piège]* ; Real₂(*ordre*) = *exécuter*.

43. Fact₀, Fact₁, Fact₂, ... [« se réaliser » ; verbe qui prend C₀ comme son SG et le nom du premier, deuxième ... actant de C₀ (si C₀ a un actant) comme son CO principal] : Fact₀(*soupçon*) = *se confirmer* ; Fact₂(*mémoire*) = *garder, retenir [qch]* ; AntiFact₁(*mémoire*) = *trahir [qqn]*.

4.3. Le système de paraphrases

Le système de paraphrases se compose de règles de deux ordres : les règles dites lexicales et les règles dites syntaxiques.

- *Les règles lexicales de paraphrase* représentent soit des (quasi-)équivalences sémantiques soit des implications sémantiques formulées en termes de FL. Nous connaissons une cinquantaine de règles lexicales de paraphrase, valables pour toutes les langues. Nous présenterons ici une liste abrégée (empruntée à Žolkovskij & Mel'čuk (1967) et Mel'čuk (1974)).

Une règle lexicale de paraphrase constate que tels nœuds lexicaux de la SSyntP (un ou deux) peuvent être remplacés *salva significatione* par tels autres nœuds (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 séquence de règles syntaxiques de paraphrase qui la « desservent » : elles effectuent les restructurations nécessitées par les changements lexicaux. Cette séquence 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. Avant de procéder, nous discuterons de trois particularités générales des règles lexicales de paraphrases sous leur forme présente :

- absence de précision sémantique ;
- absence de conditions d'application ;
- incomplétude et redondance de l'ensemble.

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 déverbatif plus le verbe support approprié (tel que *voyager* = *faire un < des > voyage > s >*). Mais au niveau concret, sur le plan linguistique, où les lexèmes français apparaissent à la place des FL, cette équivalence peut subir des entorses, puisque les valeurs réelles des FL s'écartent souvent de la signification « idéale ». Ainsi, *en larmes* = A_1 (*pleurer*), de façon que *Elle pleurait* = *Elle était en larmes*; pourtant cette égalité n'est qu'approximative, si bien qu'elle ne peut pas se réaliser dans beaucoup de contextes, par exemple :

(23) 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 indique explicitement la présence d'un décalage sémantique.

Par conséquent, les paraphrases doivent s'effectuer sous LE CONTRÔLE DE LA SÉMANTIQUE, c'est-à-dire en prenant en compte les descriptions sémantiques de toutes les unités lexicales qui participent au processus. Quand nous paraphrasons *Elle pleurait* par *Elle était en larmes* à l'aide d'une règle lexicale de paraphrase (N° 25, une fission à verbe support : $V \Leftrightarrow A_1(V) + \text{Oper}_1(A_1(V))$), 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.

2. L'applicabilité des règles lexicales de paraphrase est contrainte par plusieurs conditions syntaxiques (au sens très large de ce terme). Ainsi, la règle de conversion ne peut pas être appliquée si le verbe de départ est modifié par un adverbe syncatégorématique (= qui porte sémantiquement sur le sujet grammatical) :

(24) Jean, **très généreusement**, a fait cette proposition à Marie. ~~↔~~
Marie, **très généreusement**, a reçu [= Conv₂₁(faire)] cette proposition de Jean.

Une autre illustration : pour qu'une règle lexicale de fission (par exemple, $V \Leftrightarrow S_0(V) + \text{Oper}_1(S_0(V))$) puisse être appliquée, il faut spécifier le nombre grammatical du nom qui apparaît (*punir* ↔ *infliger une punition/des punitions*, ...); ce nombre dépend de la signification lexicale du verbe (surtout du côté aspectuel) ainsi que de sa forme grammaticale. Cependant, les règles précises à ce propos ne sont pas encore formulées.

3. D'une part, notre liste de règles n'est nullement complète. Il reste à entreprendre des recherches approfondies visant à découvrir d'autres équivalences sémantiques exprimables en termes de fonctions lexicales.

les. D'autre part, 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, croyant qu'il est peut-être encore trop tôt pour faire le ménage à l'intérieur du système de paraphrases. Il faut d'abord bien le roder et s'assurer de sa complétude relative. Aussi croyons-nous qu'il est utile et intéressant de pouvoir obtenir certaines paraphrases par des voies différentes.

Règles lexicales de paraphrase

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

1. C_0 dénote le lexème de départ, c'est-à-dire le lexème auquel 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 paraphrase desservant la règle lexicale donnée (voir les règles syntaxiques).
4. La flèche sur le numéro d'une règle syntaxique signale son application dans le sens inverse (de la droite à la gauche) ; les parenthèses autour du numéro signalent le caractère facultatif de la règle (par exemple, $\bar{11}$ dans la règle lexicale N° 37 et (11) dans la règle lexicale N° 18). Voir pp. 40 et 18.

• *Les règles syntaxiques de paraphrase* décrivent les changements de la SSyntP qui doivent accompagner les substitutions lexicales effectuées par des règles lexicales. Ces règles correspondent, dans un certain sens, aux transformations de la grammaire transformationnelle : elles effectuent des transformations des arbres étiquetés de dépendance qui représentent les SSyntP des phrases (pour les transformations élémentaires définies sur les SSyntP du modèle Sens-Texte voir Gladkij & Mel'čuk (1975 (1983 : 151-187))).

Règles syntaxiques de paraphrase

Nous présentons les règles syntaxiques de paraphrase réunies en quatre groupes : A – les opérations de fission/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, pour fins de référence, les numéros des règles lexicales qu'elle dessert.

A. Fissions

Règles lexicales desservies

- | | | | |
|----|---|---------------------------------------|--|
| 1. | $\begin{array}{c} z \\ \circ \\ \downarrow \\ \circ \\ Y \end{array}$ | $X \circ \Leftrightarrow \Rightarrow$ | 18, 19, <i>Ils luttent</i> [X] \Leftrightarrow <i>Ils mènent</i>
24, 25, [Z] <i>une lutte</i> [Y]
28, 29 |
|----|---|---------------------------------------|--|

RÈGLES LEXICALES DE PARAPHRASE

SUBSTITUTIONS	X	Y	Exemples
	1. C ₀ (V) ↔ Syn(C ₀)		— Ce phénomène s'explique facilement [C ₀] ↔ Ce phénomène s'explique sans difficulté [Syn(C ₀)]
	2. C ₀ (V) ↔ Conv ₂₁ (C ₀)		19 Je [I] crains [C ₀] les conséquences [III] ↔ Les conséquences [I] m' [III] effraient [Conv ₂₁ (C ₀)]
	4. C ₀ (V) ↔ Conv ₃₂₁ (C ₀)		20 Il [I] enseigne [C ₀] la langue [II] aux enfants [III] ↔ Les enfants [I] apprennent [Conv ₃₂₁ (C ₀)] la langue [II] avec lui [III]

A. Équivalences sémantiques

I. Règles du type C₀ ↔ FL(C₀)

SUBSTITUTIONS

Substitutions synonymiques

Règles syntaxiques de paraphrase

Substitutions conversives

Conversifs à deux arguments

Conversifs à trois arguments

				Conversifs à quatre arguments
6.	$C_0(V) \Leftrightarrow \text{Conv}_{1432}(C_0)$			22, 21, 22 <i>Il [I] a acheté [C₀] ce livre [III] à Jean [III] 30 francs [IV] \Leftrightarrow Il [I] a payé [Conv₁₄₃₂(C₀)] à Jean [III] 30 francs [II] pour ce livre [IV]</i>
...				
11.	$(C_0(V) \Leftrightarrow \text{Adv}_{1A}(C_0))$			Substitutions dérivatives (= adverbiales)
23				<i>Il semble [C₀] que cette approche est valable [I] \Leftrightarrow À ce qu'il semble [Adv_{1A}(C₀)], cette approche est valable</i>
12.	$C_0(V) \Leftrightarrow \text{Adv}_{1B}(C_0)$			25 <i>Il [II] se hâta [C₀] de sortir [III] \Leftrightarrow Il [I] sortit en hâte [Adv_{1B}(C₀)]</i>
...				
				FISSIONS SIMPLES
	X	Y	Z	Fissions à copule
15.	$C_0(V) \Leftrightarrow \text{Conv}_{12\text{-est-}3}(C_0) \xrightarrow{\text{II}} \text{ÊTRE}$			5 <i>Posons [C₀] le diamètre [III] égal [III] à 10 \Leftrightarrow Supposons [Conv_{12-est-3}(C₀)] que le diamètre est égal à 10 [II]</i>
...				Fission à négation
17.	$C_0(V) \left\{ \begin{array}{l} \Leftrightarrow \text{Anti}(C_0) \text{ ATTR} \text{ NE...PAS} \\ \text{ÊTRE} \rightarrow C_0(A) \end{array} \right.$			— <i>Jean est resté [C₀] \Leftrightarrow Jean n'est pas parti [Anti(C₀)] ; Lucie est courageuse [C₀] \Leftrightarrow Lucie n'est pas poltronne [Anti(C₀)]</i>
18.	$C_0(V) \Leftrightarrow S_0(C_0) \xrightarrow{\text{II}} \text{Oper}_1(S_0(C_0))$			Fissions à verbe support
				1, (11) <i>Jean nous a bien accueillis [C₀] \Leftrightarrow Jean nous a réservé [Oper₁(S₀(C₀))] un bon accueil [S₀(C₀)]</i>

19.	$C_0(V) \Leftrightarrow S_0(C_0) \xrightarrow{II} \text{Oper}_2(S_0(C_0))$	1, 15, (10)	<i>Jean nous a bien accueillis [C₀] ⇔ Nous avons trouvé [Oper₂(S₀(C₀))] chez Jean un bon accueil [S₀(C₀)]</i>
20.	$C_0(V) \Leftrightarrow S_0(C_0) \xrightarrow{II} \text{Func}_0(S_0(C_0))$	2	<i>En Afghanistan, le peuple lutte [C₀] contre l'envahisseur soviétique ⇔ En Afghanistan, se déroule [Func₀(S₀(C₀))] la lutte [S₀(C₀)] du peuple contre l'envahisseur soviétique</i>
...			
23.	$C_0(V) \Leftrightarrow S_0(C_0) \xrightarrow{III} \text{Labor}_{12}(S_0(C_0))$	6	<i>L'aviation soviétique a farouchement bombardé [C₀] le village ⇔ L'aviation soviétique a soumis [Labor₁₂(S₀(C₀))] le village à un bombardement [S₀(C₀)] farouche</i>
24.	$C_0(V) \Leftrightarrow S_1/A_1(C_0) \xrightarrow{II} \text{Oper}_1(S_1/A_1(C_1))$	1	<i>Belfort, qui a écrit [C₀] cet article, ... ⇔ B., qui est [Oper₁(S₁(C₀))] l'auteur [S₁(C₀)] de cet article, ...</i>
25.	$C_0(V) \Leftrightarrow S_2/A_2(C_0) \xrightarrow{II} \text{Oper}_1(S_2/A_2(C_0))$	19, 1	<i>Je le hais [C₀] ⇔ Il m'est [Oper₁(A₂(C₀))] odieux [A₂(C₀)]</i>
...			
FISSIONS COMPLEXES			
X	W	Y	Z
34.	$C_0(V) \xrightarrow{\text{ATTR}} \text{Loc}_{in} \Leftrightarrow S_{loc}(C_0) \xrightarrow{II} \text{Oper}_1(S_{loc}(C_0))$	7	<i>Les milices musulmanes combattent [C₀] dans [Loc_{in}] les rues de Saïda ⇔ Les rues de Saïda sont [Oper₁(S_{loc}(C₀))] l'arène [S_{loc}(C₀)] du combat [C₀] des milices musulmanes</i>
35.	$C_0(V) \xrightarrow{\text{ATTR}} \text{Instr} \Leftrightarrow S_{instr}(C_0) \xrightarrow{II} \text{Oper}_1(S_{instr}(C_0))$	7	<i>Jean a été assassiné [C₀] avec [Instr] ce stylet ⇔ Ce stylet a été [Oper₁(S_{instr}(C₀))] l'arme [S_{instr}(C₀)] de son meurtrier [S₀(C₀)]</i>

36. $C_0(V) \xrightarrow{ATTR} A \Leftrightarrow \left\{ \begin{array}{l} S_{mod}(C_0) \\ S_0(C_0) \end{array} \right\} \xrightarrow{I} \text{Pred}(A)$ 25 *Lucie écrit [C₀] clairement [A] ⇔ Le style [S_{mod}(C₀)] de Lucie est clair [Pred(A)]*

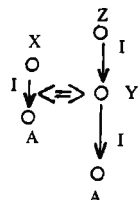
II. Règles du type FL₁(C₀) ⇔ FL₂(C₀)

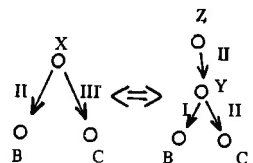
SUBSTITUTIONS

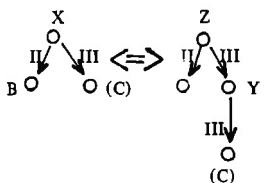
- | | | |
|---|------------------------|---|
| X | Y | |
| 37. Oper ₁ (C ₀) ⇔ Oper ₂ (C ₀) | 1) 20, 10
2) 11, 20 | <i>Les partisans leur ont opposé [Oper₁(C₀)] une résistance [C₀] acharnée ⇔ Ils ont buté [Oper₂(C₀)] sur une résistance [C₀] acharnée des partisans</i> |
| 38. Func ₁ (C ₀) ⇔ Func ₂ (C ₀) | 15 | <i>Il a manifesté [Oper₁(C₀)] une bonne compréhension [C₀] de ce point de vue ⇔ Ce point de vue a rencontré [Oper₂(C₀)] chez lui une bonne compréhension [C₀]</i>
<i>L'idée [C₀] d'appliquer l'opérateur λ appartient [Func₁(C₀)] au Prof. Cusson ⇔ L'idée [C₀] du Prof. Cusson consiste [Func₂(C₀)] en l'application de l'opérateur λ</i> |
| 39. Labor ₁₂ (C ₀) ⇔ Labor ₃₂ (C ₀) | 20 | <i>russe Ivan vzial [Labor₁₂(C₀)] v dolg [C₀] u Petra 100 rublej, lit. 'Jean a pris en prêt à Pierre 100 roubles' ⇔ Pêtr dal [Labor₃₂(C₀)] Ivanu v dolg [C₀] 100 rublej, lit. 'Pierre a donné à Jean en prêt 100 roubles'</i> |

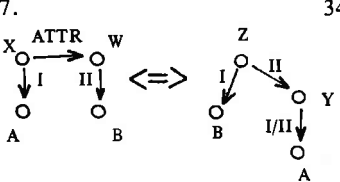
X	Y		
40. Oper ₁ (C ₀) ⇔ Func ₁ (C ₀)		19	<i>Le général Clas a donné [Oper₁(C₀)] cet ordre[C₀] ⇔ Cet ordre vient [Func₁(C₀)] du général Clas</i>
41. Oper ₁ (C ₀) ⇔ Labor ₁₂ (C ₀)		1) 11	<i>Il a effectué [Oper₁(C₀)] une analyse [C₀] profonde du phénomène ⇔ Il a soumis [Labor₁₂(C₀)] le phénomène à une analyse [C₀] profonde</i>
...		2) 21	<i>L'algorithme effectuée [Oper₁(C₀)], sur la phrase donnée, la transformation [C₀] suivante ⇔ L'algorithme soumet [Labor₁₂(C₀)] la phrase donnée à la transformation [C₀] suivante</i>
FISSIONS (à négation)			
45. Perm(C ₀) ⇔ Lique(C ₀)	ATTR NE...PAS	—	<i>L'administration ferme les yeux [Perm(C₀)] sur l'alcoolisme [C₀] des travailleurs ⇔ L'administration ne combat pas [Liqu(C₀)] l'alcoolisme des travailleurs</i>
46. Fin(C ₀) ⇔ Incep(C ₀)	II Anti(C ₀)	—	<i>Jean a cessé [Fin(C₀)] de soigner [C₀] sa goutte ⇔ Jean a commencé [Incep(C₀)] à négliger [Anti(C₀)] sa goutte</i>
B. Implications sémantiques			
48. CausX(C ₀ (N)) ⇒ IncepX(C ₀ (N))	X Y	19	<i>Jean a allumé [CausFact₀(C₀)] le moteur [C₀] ⇒ Le moteur [C₀] a démarré [IncepFact₀(C₀)]</i>

49. CausX(C ₀ (V)) ⇒ X(C ₀ (V))	19	<i>Jean a introduit [Caus(entrer)] le boulon dans l'orifice ⇒ Le boulon est entré [C₀] dans l'orifice</i>
50. PermX(C ₀ (N)) ⇒ X(C ₀ (N))	19, 21	<i>russe Ivana dopustili [PermOper₂(C₀)] k èkzameni [C₀], lit. 'On a admis Jean à l'examen' ⇒ Ivan deržal [Oper₂(C₀)] èkzamen, lit. 'Jean a tenu l'examen'</i>
...		

2.  20, 22 *Partout en Afrique, les peuples [A] luttent [X] contre les dictatures communistes* \Leftrightarrow *Partout en Afrique, se déroule [Z] la lutte [Y] des peuples [A] contre les dictatures communistes*

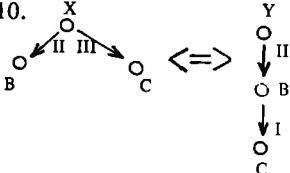
5.  15 *Je la [B] crois [X] malade [C]* \Leftrightarrow *Je crois [Z] qu'elle [B] est [Y] malade [C]*

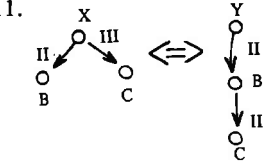
6.  23, 26, 27 *Il analyse [X] le texte [B]* \Leftrightarrow *Il soumet [Z] le texte [B] à l'analyse [Y]*

7.  34, 35 *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'']*

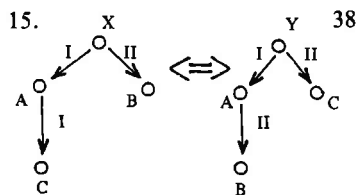
B. Transferts

Les opérations de transfert présupposent 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.

10.  19, 37 *Jean bénéficie [X] auprès de Lise [C] d'une entière confiance [B]* \Leftrightarrow *Jean bénéficie [Y] de l'entière confiance [B] de Lise [C]*

11.  18, 28, 29, 37, 41 *Jean a consacré [X] un roman [B] aux linguistes [C]* \Leftrightarrow *Jean a écrit [Y] un roman [B] sur les linguistes [C]*

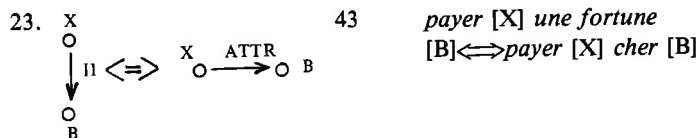
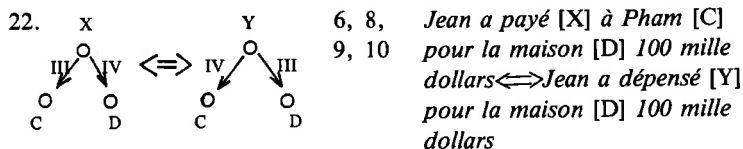
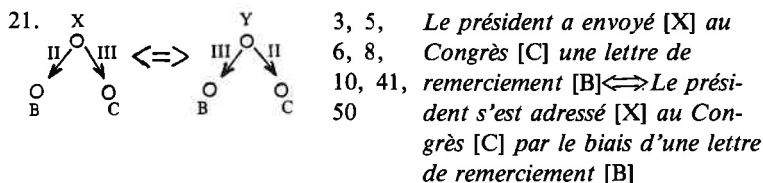
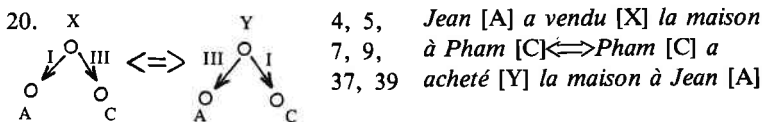
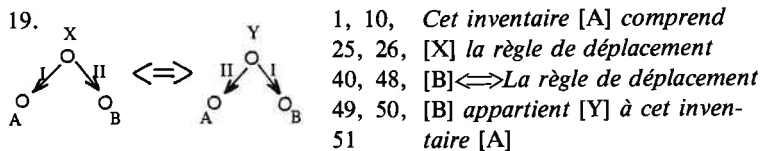
...



Le plan [A] de Jean [C] prévoit [X] un voyage dans les Alpes [B] ↔ Le plan [A] de voyage dans les Alpes [B] provient [Y] de Jean [C]

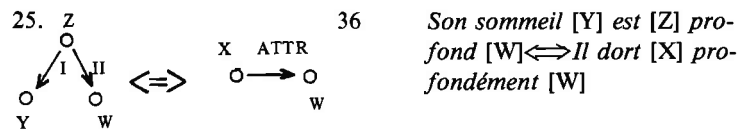
...

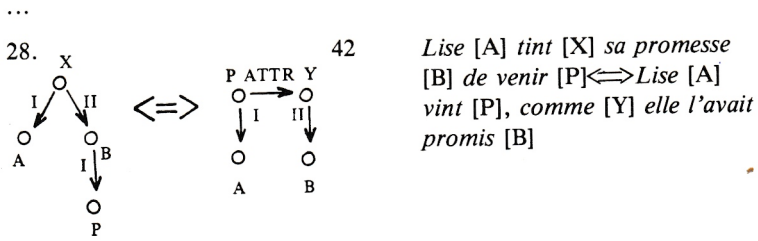
C. Réétiquetages des branches actancielles (= opérations de conversions)



...

D. Inversion de subordination



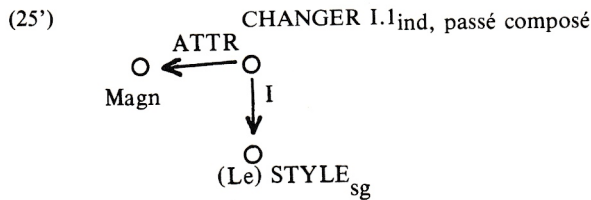


Un exemple de fonctionnement du système de paraphrases

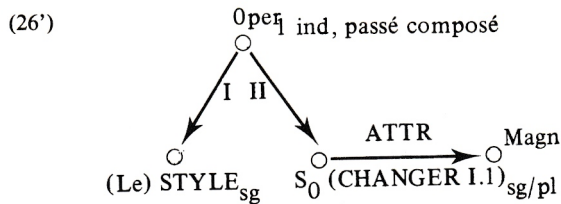
Voyons maintenant comment le système ébauché peut être appliqué à une SSyntP pour en dériver des SSyntP synonymes. Pour ce faire, prenons un fragment d'une phrase déjà vue, à savoir la phrase (17a) :

(25)[= (17a)] Le style... a sérieusement changé,

qui a la SSyntP (25') :



Tout d'abord, on peut y appliquer une règle lexicale quelconque, par exemple, N° 18 ; posant $CHANGER\ I.1 = C_0 = X$, on le remplace par $S_0(CHANGER\ I.1) = Y$ et $Oper_1(S_0(CHANGER\ I.1)) = Z$; en utilisant la règle syntaxique correspondante N° 1, nous obtenons la SSyntP (26') :

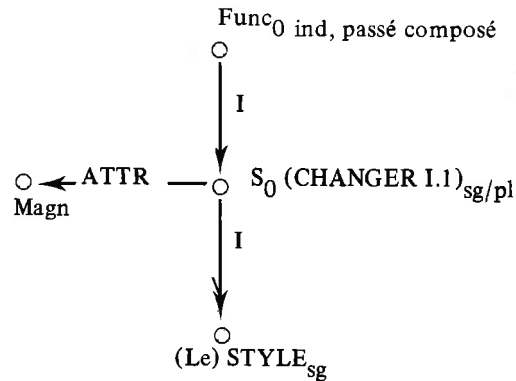


ce qui peut se lire comme

(26) *Le style... a connu <subi> un changement/des changements sérieux.*

On pourrait également appliquer à la SSyntP de (25') une autre règle lexicale de paraphrase : N° 20, avec la règle syntaxique N° 2 ; le résultat en serait (27') :

(27')



c'est-à-dire

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

Une autre voie serait de remplacer d'abord CHANGER I.1 par son synonyme SE TRANSFORMER (la règle lexicale de paraphrase N° 1) et d'appliquer ensuite au résultat les règles mentionnées ; cela nous aurait donné des phrases comme (28) :

- (28) 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 *Les transformations sérieuses ... du style ... ont eu lieu < sont survenues > .*

5. DICTIONNAIRE EXPLICATIF ET COMBINATOIRE

Après la présentation de la section précédente, il doit être clair à quel point un lexique est important pour notre approche de la paraphrase. La paraphrase s'appuie sur les fonctions lexicales de façon tout à fait cruciale : les règles de paraphrase sont formulées en terme de FL ; or les valeurs des FL pour chaque unité lexicale sont spécifiées dans le lexique. La paraphrase se fait sous le contrôle des représentations sémantiques d'unités lexicales 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 fort complexe et très volumineux. Au sein du MST, le lexique en question est réalisé comme Dictionnaire explicatif et combinatoire (DEC).

Etant donné que les deux volumes du DEC du français contemporain ont déjà été produits, nous pouvons nous limiter ici à une caractérisation très brève.

Un article du DEC est consacré à un LEXÈME ou à un PHRASÈME, c'est-à-dire à un mot (ou à une locution idiomatique) pris dans une seule acception bien déterminée. Les lexèmes qui ont le même signifiant et dont les signifiés partagent une composante sémantique non triviale sont réunis dans un VOCABLE.

Un article du DEC a une structure standard ; il comporte trois zones majeures (pertinentes pour notre discussion) :

- zone sémantique, ou *grosso modo* la définition ;
 - zone de combinatoire syntaxique, ou *grosso modo* le schéma de régime ;
 - zone de combinatoire lexicale, ou les fonctions lexicales.
- (Nous ne mentionnons pas ici la zone morphologique, la zone de marques d'usage, la zone de phraséologie et la zone illustrative.)

5.1. Zone sémantique

La particularité des définitions du DEC est que, dans le cas d'un lexème prédicat (au sens logique), on ne définit pas le lexème, mais une expression à variables qui inclut le lexème en cause ; cette expression est appelée FORME PROPOSITIONNELLE. Les variables de la forme propositionnelle d'un lexème L expriment les actants sémantiques de L, c'est-à-dire les participants de la situation décrite par L. Ainsi, on ne définit pas le verbe INTERDIRE, mais la forme propositionnelle *X interdit à Y de Z-er* ; on ne définit pas le nom PERSECUTION, mais la forme propositionnelle *la persécution de Y par X* ; etc.

Le définissant est écrit en français, mais en français restreint, amputé de nombre de ses particularités idiomatiques ; c'est en fait le (méta)-langage sémantique que nous avons mentionné dans la section 2. Il importe de souligner que tout définissant peut être mécaniquement converti en une RSém de forme standard (= en réseau sémantique).

Les définitions du DEC respectent les six principes suivants :

- Principe d'UNIVOCITÉ : Le langage définitoire exclut la synonymie et l'ambiguïté ; autrement dit, chaque élément dans le définissant n'a qu'un seul sens et chaque sens n'est exprimé que par un seul élément. Pour aboutir à cet état de choses, d'une part, les mots que nous utilisons dans le définissant sont munis de numéros distinctifs, tel qu'indiqué dans la section 2 (p. 17). D'autre part, une fois choisie l'expression pour un sens donné, nous écrivons toujours, par exemple, 'destiné à [être utilisé]' et évitons les variantes comme 'pour qu'on [l'utilise]', ou 'qu'on doit [utiliser]', etc.
- Principe d'ADÉQUATION : Toutes les composantes d'un définissant sont nécessaires et le définissant comme tel est suffisant. Il s'ensuit que le définissant et le défini doivent être mutuellement substituables dans n'importe quel contexte *salva significatione* — sans que le sens en soit

affecté (la combinatoire et le style peuvent être violés). La substitution mutuelle du défini et du définissant est l'outil principal de vérification des définitions du DEC.

- Principe de DÉCOMPOSITION : Le définissant ne doit contenir que des lexèmes qui sont sémantiquement plus simples que le défini. Autrement dit, chaque définition représente une décomposition sémantique du lexème défini ; cette façon de procéder évite les cercles vicieux dans le système de définitions et garantit, après un nombre suffisant de pas, la réduction du stock lexical à un jeu d'éléments lexicaux primitifs, qui sont donc sémantiquement indécomposables (à l'intérieur d'une langue naturelle).
- Principe de BLOC MAXIMAL : Le définissant ne doit pas contenir une expression $L_1 + L_2$ si elle signifie 'L' ; cette expression doit être remplacée par L. Ce principe assure la décomposition graduelle de façon à expliciter les liens sémantiques entre tous les lexèmes et, en même temps, permet de garder une taille maniable des définitions. (Sans ce principe qui exige la décomposition la moins profonde possible, on serait obligé d'accepter une profondeur de décomposition soit arbitraire, soit la plus grande possible ; dans ce dernier cas, les définitions deviennent tout à fait non maniables.)
- Principe de PONT SÉMANTIQUE : Le lien sémantique entre deux lexèmes appartenant au même vocable doit être exprimé explicitement — par la présence d'une composante commune (qui est appelée *pont sémantique*).
- Principe d'UNIFORMITÉ : Les vocables appartenant au même champ lexical doivent être présentés selon un même schéma.

Les définitions d'un DEC obéissent également à un certain nombre de critères qui prennent en considération la cooccurrence du lexème, mais nous ne pouvons pas développer ici ce sujet.

5.2. Zone de combinatoire syntaxique

Le pivot de cette zone est le schéma de régime : un tableau rectangulaire dont chaque colonne correspond à un actant sémantique de l'unité lexicale vedette et chaque ligne spécifie le moyen de surface utilisé pour exprimer cet actant. Ainsi, INTERDIRE I.1 [= '... X ... ne donne pas à Z le droit de Y-er'] a le schéma de régime suivant :

1 = X	2 = Y	3 = Z
1. N	1. N 2. <i>de</i> V_{inf} 3. <i>que</i> PROP _{subj} obligatoire	1. à N

Cela signifie que X est exprimé comme le premier actant SyntP de INTERDIRE, qui, à son tour, est réalisé à la surface par un syntagme nominal sujet ; Y correspond au deuxième actant SyntP et peut être exprimé soit par un SN objet (*interdire les sports violents*), soit par un infinitif introduit par DE (*interdire de skier*), soit par une proposition subordonnée au subjonctif (*Le directeur a interdit qu'on le dérange*) ; Z est exprimé par le troisième actant réalisé par un SPrép « à N » ou par un clitique datif (*interdire à Jean/lui interdire [d'y aller]*).

On indique encore des restrictions imposées à la compatibilité de divers actants syntaxiques du même lexème, mais nous n'entrerons pas dans les détails.

5.3. Zone de combinatoire lexicale

Les fonctions lexicales ont été décrites dans la section 4, et nous n'y reviendrons plus.

En guise de conclusion, nous citons quelques articles de dictionnaire (du type DEC), en illustrant surtout des lexèmes qui apparaissent dans nos exemples de paraphrases, pp. 45-46. Ces articles ne sont présentés qu'en ébauche, et pour les mots choisis, qui sont tous polysémiques, un seul sens (parfois deux) est élaboré.

CHANGEMENT, nom, masc.

I.1. *Changement de X* = (S₀)X change I.1] [= 'le fait que X change I.1']

Régime	
X	= I
1.	de N
2.	A _{poss}

le changement de son travail ; son changement

Fonctions lexicales

Syn _C	: modification, transformation I.1 ; variation ; évolution ; amélioration ; augmentation ; diminution
V ₀	: changer I.1
Magn	: profond, sérieux, radical < complet, total
AntiMagn	: léger < imperceptible
Oper ₁	: subir, connaître [ART ~]
Func ₀	: 'avoir lieu', survenir

CHANGER, verbe

I.1. *X change* = X devient 1 différent 1 de ce que X a été.

Régime

X = I	
1.	N

Le travail < Marie > change.

Fonctions lexicales

Syn _C	: se modifier, se transformer ; varier ; évoluer ; améliorer ; empirer ; augmenter ; diminuer
S ₀	: changement I.1
Magn	: beaucoup ; profondément, sérieusement, radicalement < complètement, totalement
AntiMagn	: peu ; légèrement < imperceptiblement
Able ₁	: changeant ; variable ; instable
Qual ₁	: léger, mobile ; versatile ; volage
AntiAble ₁	: stable ; invariable ; immuable ; inaltérable ; durable, constant ; fixe
A ₁ Perf	: différent 1 ; autre
AntiA ₁ Perf	: le même

PERSÉCUTER, verbe, trans.

1. *X persécute Y = X* tourmente 2 Y usuellement (et) sans relâche en traitant 1 Y de façon II injuste 2.

N.B. : La composante 'usuellement' est nécessaire puisque PERSÉCUTER 1 ne peut pas décrire un événement actuel [*J'entrais dans la pièce au moment où les policiers *persécutaient le capitaine Morin*].

Régime

X = I	Y = II
1. N	1. N

Le gouvernement soviétique persécute les intellectuels.

Fonctions lexicales

Syn _D	: tourmenter
Syn _N	: réprimer ; opprimer ; martyriser ; poursuivre I.5 ; harceler
non	: 'laisser en paix < en repos >'
Anti	: protéger
S ₀	: persécution1
S ₁	: persécuteur
S ₂	: persécuté
Magn	: brutalement, cruellement

1) C1.2

: A = A₀(N) [*policière, ...*]
les persécutions des gens de lettres par la police

Fonctions lexicales

- Syn_∩ : répression ; oppression ; harcèlement
 Anti : protection
 S₂ : victime [des ~s]
 Magn : brutales, cruelles ; sanglantes
 Oper₁ : mener [ART ~s contre N] | au passif [*les persécutions menées contre les chrétiens*]
 Oper₂ : subir [ART ~s] ; être en butte [aux ~s]
 Labor₁₂ : soumettre [N aux ~s]

RADICAL, adj.

1. Qui vise à agir sur la cause profonde du fait en cause.

méthode < mesure, réforme, cure > radicale

- Syn_∩ : drastique ; profond ; complet
 Anti_∩ : superficiel ; limité

2. Magn[≅ 'intense', 'très']

changement radical

TRANSFORMATION, nom, fém.

I.1. *Transformation de X en Y = S₀(X se transforme en Y) [= 'le fait que X se transforme en Y']*

Régime

X = I	Y = II
1. de N	1. en N
2. A _{poss}	

la transformation de l'enfant < sa transformation > en femme adulte

Fonctions lexicales

Syn _⊃	: changement I.1
Syn _∩	: modification ; variation ; évolution, développement ; amélioration ; augmentation ; diminution
V ₀	: se transformer
Magn	: profonde, sérieuse, radicale < complète, totale
AntiMagn	: légère < imperceptible
Oper ₁	: subir, connaître [ART ~]
Func ₀	: « avoir lieu », survenir

[SE] **TRANSFORMER**, verbe pron.

X se transforme (en Y) = X change I.1 (en devenant Y).

Régime

X = I	Y = II
1. N	1. en N

L'enfant se transforme en femme adulte.

Fonctions lexicales

Syn _⊃	: changer I.1
Syn _∩	: tourner [à N]
S ₀	: transformation I.1
Magn	: profondément, sérieusement, radicalement < complètement, totalement
Able ₁	} : ↑ CHANGER I.1
AntiAble ₁	
Qual ₁	

Remerciements

MM. Gilles Houde et Henri Béjoint ont apporté des corrections au français de mon manuscrit. Mme Lidija Iordanskaja a lu l'article en entier et m'a indiqué nombre d'incohérences et d'omissions importantes. MM. André Clas et Alain Polguère ont assuré la révision du texte final. C'est avec grand plaisir que je leur exprime ici ma gratitude la plus cordiale.

IGOR MEL'ČUK
Département de linguistique
Université de Montréal

NOTES

1. Le présent état du MST et, en particulier, du système de paraphrases, mentionné *supra*, est tel qu'il peut être informatisé. Boyer & Lapalme (1985a,b) proposent un programme 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 de DEC, le programme produit des paraphrases françaises pour une RSém donnée.

2. Le lecteur qui connaît l'œuvre sémantique de A. Wierzbicka (1972, 1980, 1985) notera la différence entre son approche du problème d'éléments sémantiques primitifs et la nôtre. Elle commence par la recherche de ces éléments, elle en établit treize ('want', 'don't want', 'think of', 'imagine', 'say', 'become', 'be part of', 'something', 'someone', 'I', 'you', 'world', 'this') — et elle les applique lors de l'analyse sémantique de centaines de lexèmes, de morphèmes et de constructions syntaxiques de plusieurs langues naturelles. 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.

BIBLIOGRAPHIE

- APRESJAN J. (1974), *Leksičeskaja semantika. Sinonimičeskie sredstva jazyka* [Sémantique lexicale. Moyens synonymiques de la langue], Moscou, Nauka.
(1980), *Tipy informacii dlja poverxnostno-semantičeskogo komponenta modeli « Smysl ↔ Tekst »* [Les types d'information nécessaires à la composante sémantique de surface du modèle « Sens-Texte »], Vienne, Wiener Slawistischer Almanach.
- BOYER M. & LAPALME G. (1985a), « Generating Sentences from Semantic Networks », in Dahl V. & Saint-Dizier P. eds., *Natural Logic Understanding and Logic Programming*, Amsterdam, North-Holland, pp. 181-189.
(1985b), « Generating Paraphrase from Meaning-Text Semantic Networks », *Computational Intelligence* 1, pp. 103-117.
- FUCHS C. (1980), *Paraphrase et théories du langage*, Thèse de Doctorat d'État, Paris, Université de Paris VII.
(1982), *La paraphrase*, Paris, Presses Universitaires de France.
- FUCHS C. & LE GOFFIC P. (1983) et (1985), « Ambiguïté, paraphrase et interprétation » (1ère et 2^e parties), *Modèles Linguistiques* V.2, pp. 109-136 et VII.2, pp. 27-51.
- GENTILHOMME Y. (1980), « Un micro-système didactique », *Bulletin de linguistique appliquée et générale* 7, Université de Besançon, pp. 81-94.
- GLADKIJ A. & MEL'ČUK I. (1975), « Tree Grammars: A Formalism for Syntactic Transformations in Natural Languages », *Linguistics* 150, pp. 47-82 ; réimprimé dans Gladkij A. & Mel'čuk I., *Elements of Mathematical Linguistics*, Berlin etc., Mouton, 1983, pp. 151-187.
- GROSS M. (1981), « Les bases empiriques de la notion de prédicat sémantique », *Langages* 63, pp. 7-52.
- HARTENSTEIN K. & SCHMIDT P. (1982), « Kommentierte Bibliographie zum 'Smysl-Tekst'-Modell », *Wiener Slawistischer Almanach* 11, pp. 355-409.
- LEEMAN D. éd. (1973), « La paraphrase », *Langages* 29.
- MARTIN R. (1976), *Inférence, antonymie et paraphrase*, Paris, Klincksieck.

- MEL'ČUK I. (1974), *Opyt teorii lingvističeskix modelej « Smysl ↔ Tekst »* [Essai d'une théorie de modèles linguistiques du type « Sens-Texte »], Moscou, Nauka.
- (1979), *Studies in Dependency Syntax*, Ann Arbor, Karoma.
- (1981), « Meaning-Text Models : A Recent Trend in Soviet Linguistics », *Annual Review of Anthropology* 10, pp. 27-62.
- (1982), « Lexical Functions in Lexicographic Description », *Proceedings of the 8th Annual Meeting of the Berkeley Linguistics Society*, pp. 427-444.
- (1988), *Dependency Syntax : Theory and Practice*, New York, New York University Press.
- MEL'ČUK I. & al. (1984), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I*, Montréal, Les Presses de l'Université de Montréal.
- MEL'ČUK I. & PERTSOV N. (1987), *Surface Syntax of English : A Formal Model within the Meaning-Text Framework*, Amsterdam-Philadelphia, Benjamins.
- MEL'ČUK I. & ŽOLKOVSKIJ A. (1970), « Towards a Functioning 'Meaning-Text' Model of Language », *Linguistics* 57, pp. 10-47.
- (1984), *Explanatory Combinatorial Dictionary of Modern Russian* [en russe], Vienne, Wiener Slawistischer Almanach.
- ML = Modèles linguistiques* V.2, 1983.
- WIERZBICKA A. (1972), *Semantic Primitives*, Frankfurt, Athenäum.
- (1980), *Lingua Mentalis*, Sydney etc., Academic Press.
- (1985), *Lexicography and Conceptual Analysis*, Ann Arbor, Karoma.
- ŽOLKOVSKIJ A. & MEL'ČUK I. (1965), « O vozmožnom metode i instrumentax semantičeskogo sinteza » [Sur une méthode possible de synthèse sémantique et les outils correspondants], *Naučno-texničeskaja informacija* 5, pp. 23-28.
- (1967), « O semantičeskom sinteze » [Sur la synthèse sémantique]. *Problemy kibernetiki* 19, pp. 177-238. (Une traduction française : *T.A. Informations* 1970, n° 2.)