

La Théorie Sens-Texte

Alain Polguère

Observatoire de linguistique Sens-Texte (OLST)

Département de linguistique et de traduction

Université de Montréal

C.P. 6128, Succursale Centre-Ville

Montréal (Québec) H3C 3J7

Canada

Email : polguera@ere.umontreal.ca

1. Principes sous-jacents à l'approche Sens-Texte

En art, en science, en presque tout, il existe un concept clé de la pensée moderne : celui de « courant principal » (angl. *mainstream*). Chaque discipline semble avoir son courant principal, qui cohabite avec le reste. Quand on est actif dans une discipline, on est soit dans le *mainstream*, soit dans le marais anarchique du « reste ». Le but de cet article est de vous présenter une approche linguistique qui rôde justement dans ce marais fangeux : la théorie Sens-Texte. Je fais cela non pas parce qu'il faut bien s'intéresser aux courants marginaux, mais plutôt parce que la théorie Sens-Texte représente à mon avis la meilleure approche existant à ce jour pour mener à bien le traitement automatique des langues.

Pourquoi attacher tant d'importance à la possibilité de faire effectuer des opérations linguistiques par l'ordinateur ? On peut apporter ici au moins deux éléments de réponse complémentaires :

- 1- L'ordinateur est l'outil le plus efficace dont nous disposons pour construire et valider les modélisations formelles, quelles qu'elles soient. Il est même tellement puissant que plus aucune science « dure » ne peut maintenant se permettre de rejeter son utilisation. Dans la mesure où la linguistique formelle peut prétendre au statut de science, au même titre (ou presque) que la physique ou la chimie, il est essentiel que ses fondements théoriques et les modèles qu'elle construit puissent être manipulés par les programmes informatiques. Cela ne veut pas dire qu'il faille que la linguistique se soumette aux limitations présentes de l'informatique. Si les machines dont nous disposons actuellement, si leur mémoire, leur vitesse de calcul, les langages de programmation qu'elles peuvent interpréter, etc. ne sont pas assez puissants pour faire « tourner » les modèles linguistiques, c'est un problème que l'informatique doit régler. Mais si les modèles linguistiques sont trop flous, imprécis, etc. pour pouvoir être programmés, voilà un problème qu'il appartient à la linguistique de résoudre. Il faut noter que non seulement la linguistique doit déboucher sur la construction de modèles calculables, mais il faut aussi que les calculs que ces modèles permettent de faire présentent de l'intérêt pour la science linguistique elle-même ainsi que pour ses applications. Cette remarque permet de faire le lien avec le second élément de réponse, donné ci-dessous.
- 2- Une science ne vaut que par la compréhension qu'elle nous donne des phénomènes naturels **et** par ses applications potentielles. C'est peut-être triste à dire, mais une science sans débouchés,

une science qui n'a pas de branche appliquée, est condamnée à disparaître par le simple fait que les sociétés humaines ne seront pas prêtes à la soutenir¹. Regardons donc un peu ce qui se passe au niveau des applications de la linguistique, dont l'objet d'étude est la langue. La langue est un moyen de codification et de transmission d'informations et, à la fin du second millénaire, l'information est de plus en plus gérée par les ordinateurs. Même si les progrès du traitement automatique de la langue sont beaucoup plus lents qu'on aurait pu l'espérer, il existe un lien évident entre langue et ordinateur : l'information linguistique doit être informatisable et c'est une des tâches de la linguistique que de procurer les modèles permettant cette informatisation.

La première publication présentant les fondements de la THÉORIE SENS-TEXTE (TST) remonte à plus de trente ans : Žolkovskij et Mel'čuk (1965). La théorie linguistique proposée par I. Mel'čuk et ses collègues de Moscou peut être décrite succinctement à partir des cinq caractéristiques suivantes :

- 1) La TST rend compte de l'association que tout locuteur d'une langue L est capable de faire entre un sens donné de L et l'ensemble des énoncés paraphrastiques de L exprimant ce sens.

La théorie visée voit donc la langue **non** comme un ensemble infini d'énoncés grammaticaux ou comme la machine virtuelle permettant de produire l'ensemble infini des énoncés grammaticaux d'une langue (ce qui correspondrait à une approche de type générativiste), mais plutôt comme une machine virtuelle permettant de **traduire** des Sens en énoncés, appelés *Textes*, et vice versa. C'est cette philosophie de l'approche linguistique qui justifie le nom de *Théorie Sens-Texte*. (Dans cet article, j'écrirai *Sens* et *Texte* avec une majuscule initiale lorsqu'il s'agira de référer aux deux niveaux extrêmes de représentation des énoncés linguistiques postulés par la TST.)

- 2) La TST est universelle, c'est-à-dire qu'elle repose sur des principes généraux s'appliquant à toutes les langues.

La TST ne privilégie la description d'aucune langue en particulier. On ne peut pas considérer qu'elle porte les stigmates de sa « langue maternelle » (le russe). Il est bien évident que toutes les théories linguistiques modernes prétendent à l'universalité des principes sur lesquels elles reposent ; seule la comparaison d'études typologiques, faites selon des approches théoriques différentes, permettrait de décider qu'elles sont les approches qui, de ce point de vue, tiennent le plus leurs promesses. Ce n'est nullement le but du présent article de faire ce genre de comparaison et je m'en abstiendrai donc.

- 3) La TST est linguistique en ce sens qu'elle permet, à partir des principes généraux sur lesquels elle repose, de construire des modèles linguistiques spécifiques pour chaque langue humaine.

Même si la TST repose sur des principes généraux universaux, la détermination de ces universaux n'est pas considérée comme un fin en soi. La finalité de la théorie est, au-delà de la

¹ Cette position est peut-être un peu extrême. Il est en effet probable qu'une science qui apporterait des éléments de réponse à certaines interrogations fondamentales que nous avons sur le « monde » pourrait à la rigueur se passer de débouchés concrets. En ce sens, ses découvertes seraient ses propres débouchés et la satisfaction de notre curiosité sa seule application. Il est ainsi possible de considérer que la recherche scientifique visant à expliquer les origines de l'univers a comme débouché premier la satisfaction d'une certaine curiosité existentielle.

découverte des universaux linguistiques, la construction des modèles particuliers de chaque langue. **La TST est donc orientée vers la description** et se veut un outil pour le lexicographe et le grammairien. Une conséquence importante de cette orientation descriptive est que la TST ne prétend pas élucider la réalité psychique de la langue. Il est clair qu'il existe des principes généraux gouvernant les phénomènes tels que l'apprentissage linguistique, mais pour l'approche Sens-Texte, l'étude de tels phénomènes (i) relève de ce que Saussure (1972) a appelé la *linguistique externe* — c'est-à-dire la linguistique qui étudie la langue dans ses rapports avec son contexte externe d'utilisation et (ii) présuppose une connaissance préalable des phénomènes internes à la langue. C'est l'étude et la description de ces phénomènes internes — structures lexicales et grammaticales de la langue — qui est le but premier de l'approche Sens-Texte.

4) La TST permet de construire des modèles calculables.

Comme on le verra plus loin, un modèle linguistique Sens-Texte consiste avant tout en un lexique, une grammaire et un ensemble de procédures permettant d'activer ces deux composantes pour effectuer la connexion entre Sens et Texte, symbolisée par la formule *Sens* \Leftrightarrow *Texte*. Comme toutes les composantes d'un modèle Sens-Texte sont formalisées (voir le point 5) ci-dessous), celui-ci est donc « calculable » en ce sens qu'il peut être activé par un système logique, comme un programme informatique. On peut donc dire que les modèles Sens-Texte sont par nature computationnels. Ils peuvent être testés informatiquement et utilisés par des applications informatiques mettant en jeu les connaissances lexicales et grammaticales de la langue.

5) La TST est formelle.

C'est-à-dire qu'elle utilise des langages formels pour :

- représenter les énoncés linguistiques ;
- encoder les règles de manipulation des représentations linguistiques — de telles règles servent à modéliser la correspondance Sens \Leftrightarrow Texte.

Sous cet aspect, la TST n'est pas très différente des autres approches de linguistique formelle, qui utilisent toutes des formalismes de représentation des énoncés linguistiques (arbres syntaxiques, formalismes logiques, etc.) et d'encodage des règles linguistiques (règles de réécriture syntagmatiques, règles lexicales, etc.). Ce qui distingue ici la TST des autres approches formelles, c'est sans doute la grande richesse et la relative complexité des formalismes utilisés. Cela est dû au fait que la théorie prend en compte, comme nous le verrons plus loin, tous les niveaux de fonctionnement de la langue (de la sémantique à la phonétique, en passant par la syntaxe et la morphologie). De plus, la TST a, dès ses débuts, mis sur un pied d'égalité la description des langues « traditionnelles » et des langues « exotiques ». La nécessité d'avoir des modèles rendant compte de tous les types de structures linguistiques possibles implique, par ricochet, la nécessité d'utiliser des langages formels très riches et sophistiqués.

Une partie essentielle de la tâche de développement d'un langage formel, pour toute science, est la construction d'une terminologie. La TST met donc l'emphase sur la spécification d'une terminologie linguistique ; celle-ci ne doit pas être une plate collection de termes, mais plutôt un système logique et structuré de concepts — voir Mel'čuk (1982) et Mel'čuk (1993:9-23).

Pour résumer, la TST est donc une théorie linguistique visant la description de la correspondance Sens \Leftrightarrow Texte, au moyen de la construction de modèles formels. Ces modèles peuvent être considérés comme des machines logiques virtuelles du type suivant :

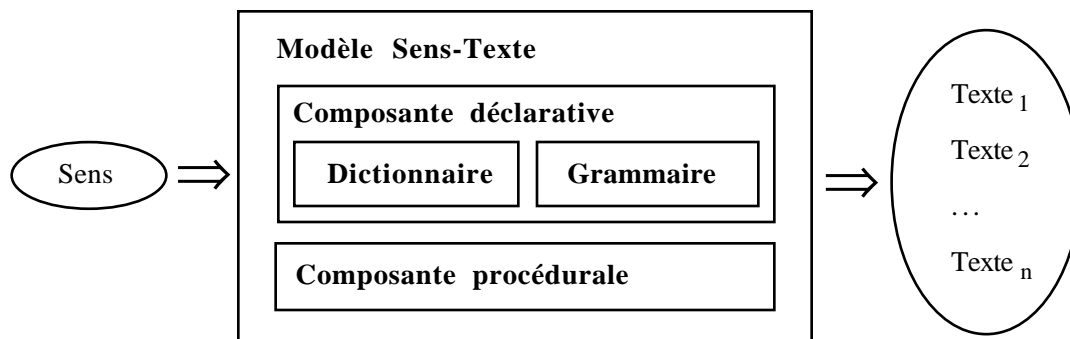


Figure 1 Structure fonctionnelle d'un modèle Sens-Texte

La Figure 1 illustre le fait qu'un modèle Sens-Texte est une machine virtuelle qui prend en entrée des (représentations de) sens d'énoncés et retourne en sortie un **ensemble** de Textes, qui contient **toutes les paraphrases permettant d'exprimer le Sens donné en entrée**. Par exemple, le modèle Sens-Texte du français permettra d'effectuer la correspondance suivante :

(1) <u>Sens</u>		<u>Textes</u>
'Norm aime sa femme Marge de façon très intense'	\Leftrightarrow	<i>Norm aime follement sa femme Marge.</i> <i>Norm aime sa femme Marge à la folie.</i> <i>Norm aime sa femme Marge comme un fou.</i> <i>Norm éprouve un amour fou pour sa femme Marge.</i> <i>Norm ressent un amour immense pour sa femme Marge.</i> etc.

Le Sens donné ci-dessus n'est pas représenté de façon formelle — si ce n'est pour l'identification précise des sens d'unités lexicales faite au moyen de numéros distinctifs (empruntés ici au *Petit Robert*). La formalisation Sens-Texte et la modélisation de la correspondance entre un Sens et un ensemble de paraphrases seront présentés plus loin.

Il convient de noter que, dans le présent article, la correspondance Sens \Leftrightarrow Texte est toujours envisagée sous l'angle de la synthèse — du Sens au Texte — plutôt que sous celui de l'analyse — du Texte au Sens. La raison en est que seule la modélisation de la synthèse linguistique permet de mettre en jeu les connaissances purement linguistiques (contenues dans le dictionnaire et la grammaire de la langue). L'analyse, elle, ne peut se faire sans que l'on soit confronté au problème de la désambiguïsation, problème qui ne peut être résolu (par le locuteur ou par une modélisation formelle) sans le recours à des heuristiques basées sur des connaissances extra-linguistiques. En résumé, la synthèse fait appel aux connaissances linguistiques du locuteur, qui doit effectuer des choix purement linguistiques entre les différentes options offertes par la langue pour l'expression d'un Sens donné. L'analyse passe par la résolution d'ambiguïtés, qui est un processus cognitif très

complexe échappant au seul domaine de la linguistique. La modélisation du processus d'analyse est pour les linguistes Sens-Texte un cadre d'**application** de la linguistique ; celle du processus de synthèse est une **méthode d'expérimentation/simulation** permettant d'identifier clairement les phénomènes linguistiques.

La primauté donnée à l'orientation Sens \Rightarrow Texte sur l'orientation Texte \Rightarrow Sens est une caractéristique fondamentale de l'approche théorique présentée ici. Pour modéliser un phénomène en linguistique Sens-Texte, il faut toujours se poser, dans l'ordre indiqué, les deux questions suivantes :

Premièrement, quel est le sens exprimé ?
Deuxièmement, quels sont les moyens utilisés pour exprimer ce sens ?

La suite du présent article est structurée de la façon suivante. La Section 2 illustrera, à partir d'un exemple simple, le processus de synthèse linguistique Sens \Rightarrow Texte — c'est-à-dire le processus permettant d'associer à une représentation d'un message linguistique l'ensemble des paraphrases exprimant, toutes, ce même message. La Section 3 fera la synthèse des informations données dans la Section 2 afin de présenter la structure générale des modèles Sens-Texte — composante déclarative, comprenant le dictionnaire et la grammaire de la langue *vs* composante procédurale, comprenant les règles permettant d'activer les descriptions de la composante déclarative. Finalement, la Conclusion offrira des pointeurs vers d'autres textes, dans lesquels le lecteur intéressé pourra trouver des informations détaillées sur la TST.

Comme je l'ai déjà signalé, ce texte ne fait pas de comparaison entre la TST et les autres approches linguistiques. De plus, il ne sera fait qu'une présentation très partielle de la théorie. En règle générale, j'éviterai de parler d'un aspect de la théorie, même important, si celui-ci ne peut être présenté de façon compacte et autosuffisante dans le présent article.

2. Le processus de synthèse linguistique Sens \Rightarrow Texte

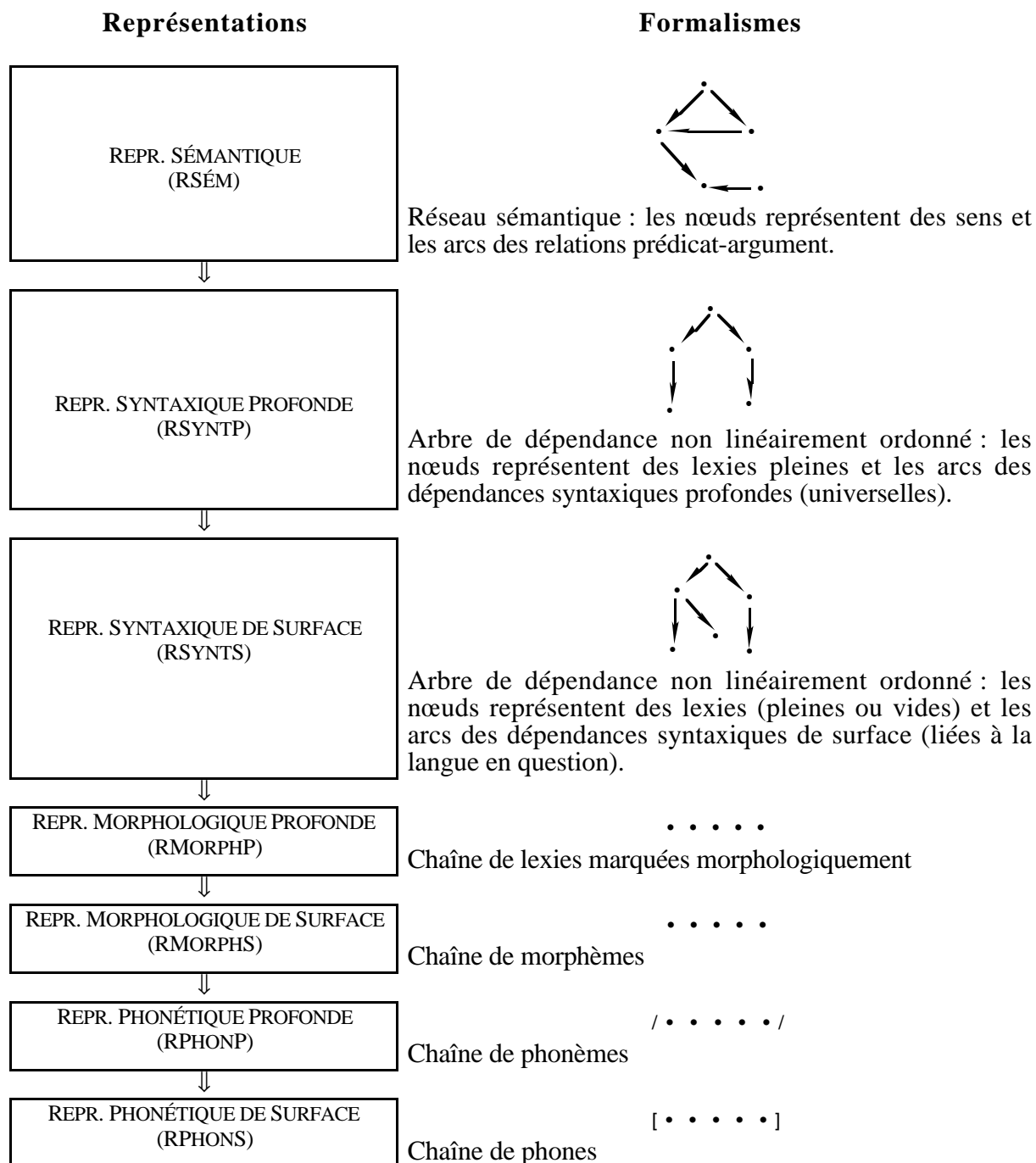
Une façon simple de présenter la structure et le fonctionnement d'un modèle Sens-Texte est de démontrer de quelle façon une représentation d'un Sens donné peut être traduite en un ensemble de paraphrases. C'est que je vais tenter de faire ici, en montrant comment, à chaque étape de ce processus de traduction, s'opèrent des choix linguistiques menant à la production d'un énoncé donné — un choix parmi toutes les paraphrases possibles permettant d'exprimer le Sens fourni en entrée au modèle.

2.1 Niveaux de représentation postulés par la théorie

La première chose à faire pour décrire le processus de synthèse Sens \Rightarrow Texte est d'examiner les différents niveaux de représentation des énoncés postulés par la théorie. Cette description sera très succincte puisque les contraintes d'espace ne me permettent pas de justifier en profondeur le recours à ces niveaux. (Il faudrait ainsi expliquer pourquoi ces niveaux sont nécessaires et suffisants pour modéliser de façon satisfaisante la correspondance Sens \Leftrightarrow Texte.)

La TST reprend à son compte la partition classique de la modélisation d'un énoncé en niveaux de représentation sémantique, syntaxique, morphologique et phonologique/phonétique. La spécificité de l'approche Sens-Texte consiste en ce que chacun des niveaux, à l'exception du niveau

sémantique, est subdivisé en niveau profond vs de surface. Le processus de synthèse Sens \Rightarrow Texte peut donc être résumé de la façon suivante :



Le tableau ci-dessus ne peut bien entendu suffire à faire comprendre au lecteur la nature des niveaux mis en jeu dans la transition Sens \Rightarrow Texte et des formalismes qui leur sont associés. C'est pourquoi nous allons maintenant examiner un cas concret de transition (basé sur l'exemple (1) donné plus haut).

2.2 Construction de la Représentation Sémantique

Le premier problème que nous devons examiner est celui de la construction de l'entrée du modèle Sens-Texte : la REPRÉSENTATION SÉMANTIQUE (RSÉM). En théorie, l'élaboration d'un message linguistique et sa traduction en une RSém est un processus qui ne relève pas du modèle linguistique proprement dit. Il est à cheval entre la modélisation du « monde » (et des processus cognitifs non linguistiques) et la modélisation linguistique elle-même. Je vais donc présupposer que la RSém est un donné de la transition Sens \Rightarrow Texte, même si dans le cadre d'applications (informatiques ou autres) de la TST, il faut bien évidemment se poser la question la production/construction de la RSém.

Reprenons le Sens donné dans l'exemple de la Section 1 : 'Norm aime_{I.2} sa femme_{II} Marge de façon très intense'. Pour pouvoir associer à ce Sens non formalisé une RSém véritable, au moins deux structures formelles doivent être construites : la STRUCTURE SÉMANTIQUE de la RSém et la STRUCTURE COMMUNICATIVE.

A) Construction de la structure sémantique de la RSém

Il s'agit ici de dégager les éléments de sens individuels, d'identifier leur nature sémantique (prédicats à un, deux, ... arguments ou objets sémantiques) et d'établir les connexions prédicat-argument les unissant. Par exemple, 'aimer_{I.2}' est un prédicat à deux arguments ('X aime_{I.2} Y') dont le premier argument est ici 'Norm' et le deuxième 'femme_{II}'. Ceci nous permet de construire le segment de réseau suivant : 'Norm $\leftarrow 1$ — aimer_{I.2} — 2 \rightarrow femme_{II}'. La structure sémantique complète de la RSém de (1) sera donnée un peu plus loin (Figure 2).

B) Construction de la structure communicative de la RSém

Il s'agit ici de subdiviser la RSém en sous-réseaux identifiant les regroupements communicatifs de sens présents dans le message. Ainsi, nous allons rendre explicite le fait que ce que nous voulons affirmer dans notre message — ce nous appellerons le RHÈME du message — est le caractère très intense du sentiment que Norm éprouve pour sa femme. La composante du message à propos de laquelle cette chose est affirmée — ce que nous appellerons le THÈME du message — est le fait même que Norm aime sa femme Marge. Nous sommes donc ici en présence d'un message qui pourrait être une réponse à une question comme :

(2) *Norm aime-t-il vraiment sa femme Marge ?*

Pour voir à quel point la spécification de la structure communicative (considérée ici sous l'angle de l'opposition Thème/Rhème) est importante, il suffit de constater que parmi les phrases ci-dessous — qui réfèrent toutes *grosso modo* au même état de choses — seule (3a) serait une réponse valide à la question (2).

(3) a. *Norm aime FOLLEMENT sa femme Marge.*

- b. *Norm aime follement sa femme **MARGE**.*
 c. *La femme de Norm, qu'il aime à la folie, s'appelle Marge.*

[Les segments en petites majuscules et en gras sont à lire avec une emphase accentuelle.]

Les RSém des phrases (3a-c) ont toutes la même structure sémantique, mais elles diffèrent par leur structure communicative.

Les deux opérations qui viennent d'être examinées en A) et B) permettent la construction de la RSém de la Figure 2 ci-dessous, pour la phrase (3a) ; cette phrase sera notre « Texte cible » dans les sous-sections suivantes² :

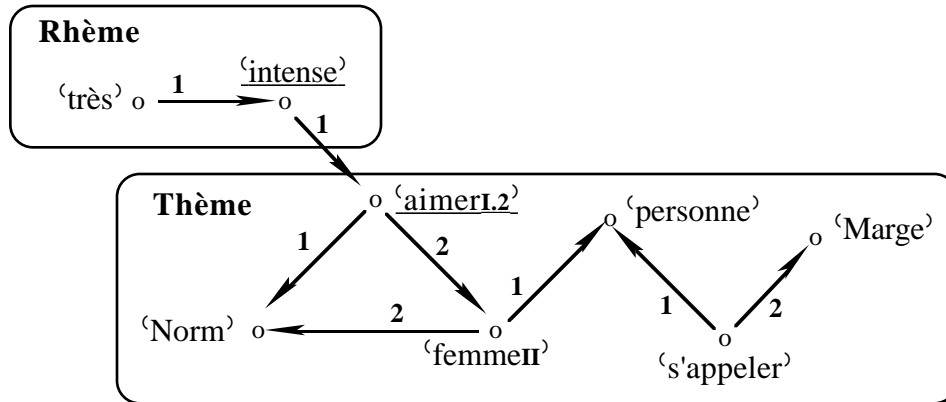


Figure 2 RSém de (3a) et de toutes ses paraphrases

Dans la figure ci-dessus, le soulignement indique que le nœud correspondant est celui qui « résume » le sens du sous-réseau dont il fait partie. Ce nœud est dit COMMUNICATIVEMENT DOMINANT. Ainsi, le Thème concerne avant tout un sentiment donné (l'amour de Norm pour sa femme) et le Rhème concerne l'intensité de ce sentiment.

Pour les lecteurs plus habitués au formalisme de la logique des prédicats qu'à celui des réseaux sémantiques, on peut indiquer que la structure sémantique de la Figure 2 est équivalente à la collection de relations prédictives suivante :

- (4) ('aimerI.2'('Norm', 'femmeII') &
 'femmeII'('personne', 'Norm') &
 's'appeler'('personne', 'Marge') &
 'intense'('aimerI.2') &
 'très'('intense')).

² Il est bien évident que, d'un point de vue théorique, c'est (3a) **et toutes ses paraphrases** qui sont la cible de la synthèse Sens ⇒ Texte. Je me concentre ici sur la synthèse d'un seul Texte pour ne pas allonger indûment mon exposé. J'essaierai cependant, dans la mesure du possible, d'indiquer au lecteur les principales options qui s'offrent à nous lors de la synthèse Sens ⇒ Texte.

Il manque cependant à la formule (4) la spécification de la structure communicative du message pour qu'elle puisse être considérée comme une RSém véritable.

2.3 Choix et construction d'une Représentation Syntaxique Profonde

Reprenant à son compte l'approche syntaxique proposée par Lucien Tesnière dans ses *Éléments de syntaxe structurale* (Tesnière 1965), la TST postule que la structure syntaxique d'une phrase est l'ensemble des liens de dépendance fonctionnelle (= les relations de fonction syntaxique) existant entre les mots de la phrase. Cette structure a pour propriété de pouvoir se représenter formellement par un arbre appelé *ARBRE DE DÉPENDANCE*. Comme il a été indiqué en 2.1 (dans le tableau décrivant les étapes du processus de synthèse Sens \Rightarrow Texte), la TST fait appel à deux niveaux de représentation syntaxique : la REPRÉSENTATION SYNTAXIQUE PROFONDE (RSyntP) et la REPRÉSENTATION SYNTAXIQUE DE SURFACE (RSyntS). Voyons tout d'abord quelles sont les caractéristiques de la RSyntP, en examinant la transition RSém \Rightarrow RSyntP à partir de notre exemple.

On pourrait dire que le but de la transition qui nous intéresse ici est d'« arboriser » la RSém, comme l'illustre la figure ci-dessous :

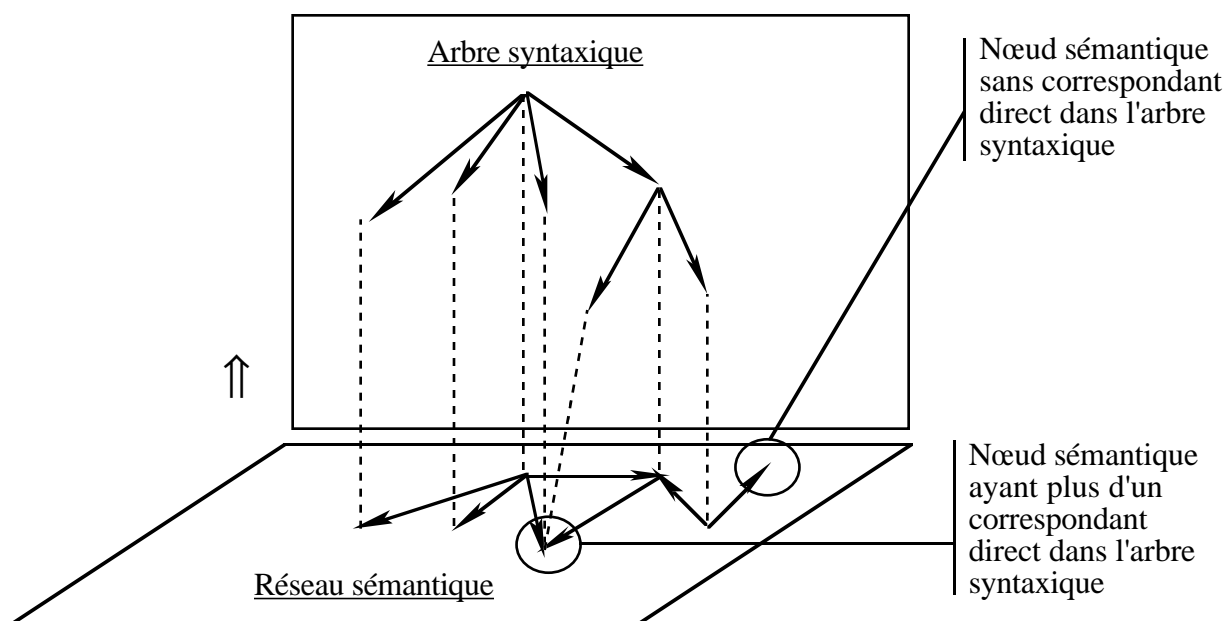


Figure 3 Exemple d'« arborisation » d'un réseau sémantique

La figure ci-dessus montre clairement que la différence formelle entre un arbre syntaxique de dépendance et un réseau sémantique réside avant tout dans les configurations de connexions entre nœuds permises par les deux types de formalismes (le réseau sémantique et l'arbre syntaxique étant tous deux des cas particuliers de graphes connectés).

Notre but étant de ne modéliser que la production de la phrase (3a) (entre toutes les phrases françaises pouvant exprimer le message représenté dans la Figure 2), nous allons identifier le sens ('aimer_{I.2}') comme la composante centrale de notre message et en dériver la racine de l'arbre syntaxique profond. Nous obtenons la RSyntP suivante :

[Prosodie déclarative neutre]

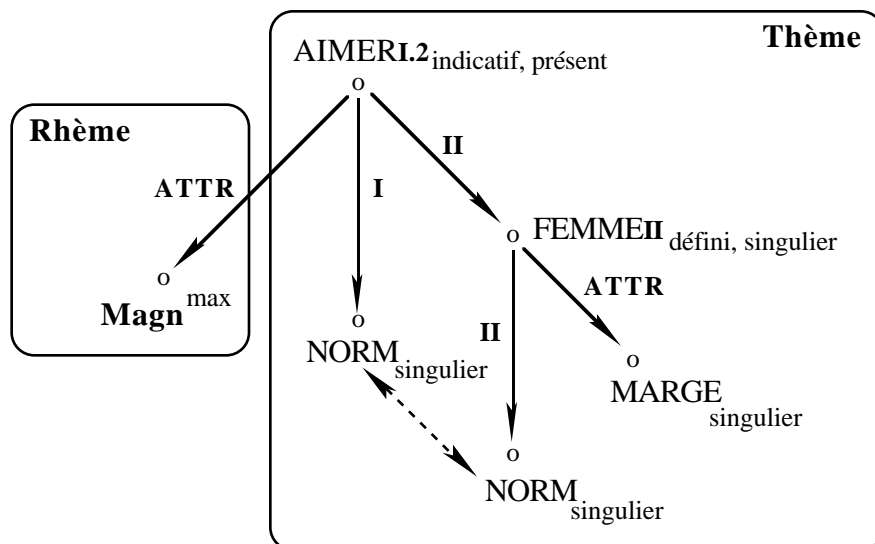


Figure 4 RSyntP de (3a)

La RSyntP est formée de quatre structures : la structure syntaxique profonde, la structure anaphorique profonde, la structure communicative profonde et la structure prosodique profonde.

A) Structure syntaxique profonde

C'est la structure principale de la RSyntP ; elle correspond à l'arborisation de la structure sémantique de la RSém et est faite des nœuds de l'arbre et des liens de dépendance (fléchés) les unissant : il s'agit de l'arbre de dépendance lui-même.

Les nœuds de l'arbre syntaxique profond sont étiquetés par deux types d'entités linguistiques :

- 1 les unités lexicales pleines qui vont figurer dans le Texte cible — ici, AIMERI.2, FEMMEII, MARGE et NORM³ ;
- 2 les FONCTIONS LEXICALES — ici, **Magn**.

³ On remarquera, dans la Figure 4, que les noms d'unités lexicales sont accompagnés de marques du type *indicatif*, *présent*, etc. correspondant aux significations grammaticales devant être exprimées auprès de ces lexies. Pour alléger la présentation, je n'ai pas introduit dans la Figure 2 la représentation de ces significations grammaticales.

Le concept de fonction lexicale est l'une des découvertes les plus intéressantes de la linguistique Sens-Texte. Il n'est malheureusement pas possible d'en faire ici une présentation satisfaisante et je vais me contenter d'une introduction très grossière.

Les fonctions lexicales sont un outil de modélisation des phénomènes collocationnels. Par exemple, dans le cas qui nous intéresse ici, on voit que la composante sémantique ('très —1→ intense') de la RSém a été traduite au niveau syntaxique profond par le nœud **Magn**^{max}. La raison en est que l'intensification est couramment exprimée en langue au moyen de constructions semi-idiomatiques, appelées *collocations*. **Magn** est le nom de la configuration collocationnelle correspondant à un modificateur intensificateur des noms, verbes, etc. Voici, à fin d'illustration, quelques exemples de **Magn** — le signe « plus petit que » (<) indique la gradation dans l'intensification :

Magn(*aimer*_{1.2}) = *beaucoup* < *de tout son cœur, intensément* < *à la folie, comme un fou*
Magn(*pluie*) = *grosse, forte* < *diluvienne, torrentielle*
Magn(*angl. rain*) = *driving, heavy* < *torrential*

L'introduction du nom de la fonction lexicale dans la RSyntP permet de modéliser le fait que ce qui compte pour le locuteur, relativement au **contenu** du Texte qu'il va produire, c'est l'expression d'un modificateur d'intensification. Le choix d'une collocation particulière (*aimer follement* vs *à la folie*) relève d'un problème relativement superficiel, plus proche des considérations de formes que de sens. C'est exactement à cela que sert la dichotomie SyntP vs SyntS : elle permet de rendre compte de la différence existant entre les choix syntaxiques liés au **contenu** à exprimer — transition RSém ⇒ RSyntP — et ceux liés à la **façon** d'exprimer ce contenu — transition RSyntP ⇒ RSyntS.

Les arcs de la RSyntP sont étiquetés par des noms de relations syntaxiques profondes. Tel qu'indiqué dans le tableau de la sous-section **2.1**, ces relations sont universelles : le jeu de relations syntaxiques profondes utilisé pour construire les RSyntP ne varie pas, quelque soit la langue considérée. La RSyntP de la Figure 4 contient les deux relations actantielles syntaxiques profondes **I** et **II**, qui sont les pendants syntaxiques des relations actantielles sémantiques **1** et **2** (premier et deuxième actant du prédicat ('aimer_{1.2}')), que l'on trouve dans la RSém. Nous verrons plus loin, lorsque nous examinerons la prochaine transition, que les relations syntaxiques profondes **I** et **II** donnent naissance, pour les dépendants syntaxiques des verbes, aux relations de surface du type sujet ou complément. Une autre relation syntaxique profonde apparaissant dans la Figure 4 est la relation attributive **ATTR**. Elle « factorise » au niveau profond tous les types de modificateurs (cf. *aimer* —**ATTR**→ *à la folie* et *sa femme* —**ATTR**→ *Marge*)⁴.

Il est très important de noter qu'un arbre syntaxique (profond ou surface) de la TST n'est pas linéairement ordonné. Ce type d'arbre ne sert qu'à décrire les liens de dépendance syntaxique existant entre les lexies de la phrase ; l'ordre des éléments dans la phrase — qui est un **moyen d'expression** des dépendances syntaxiques — ne sera calculé que lors de la transition vers le niveau morphologique. (Les représentations morphologiques sont des chaînes linéaires.) Le fait d'avoir dessiné la dépendance unissant **Magn**^{max} à **AIMER**_{1.2} à gauche de celle unissant **NORM** à

⁴ Il existe d'autres relations syntaxiques profondes (**III**, **IV**, **V**, **VI**, **APPEND** et **COORD**), que je n'introduirai pas ici afin d'alléger l'exposé.

ce même verbe n'a donc aucune signification représentationnelle. On aurait pu tout aussi bien choisir la disposition inverse.

B) Structure anaphorique profonde

La RSyntP est un arbre pour la bonne raison que les lexies de la phrase (à l'exception de celle formant la racine de l'arbre) dépendent syntaxiquement d'une et une seule autre lexie de la phrase : ce qui revient à dire qu'une lexie n'a qu'une seule fonction syntaxique dans la phrase. On peut donc trouver dans l'arbre syntaxique plusieurs nœuds ayant « germés » à partir d'un même nœud de la RSém de départ : ce sont des nœuds coréférentiels — voir les deux nœuds NORM de la Figure 4. La coréférence doit être indiquée dans la RSyntP, notamment pour pouvoir modéliser les phénomènes de pronominalisation lors du passage à la RSyntS. Cela se fait en superposant à la structure syntaxique profonde de base une structure additionnelle : la structure anaphorique profonde, représentée au moyen de flèches pointillées à double orientation reliant les nœuds coréférentiels.

C) Structure communicative profonde

Comme la RSém, la RSyntP possède une structure communicative. Si la structure communicative de la RSém est utilisée pour contrôler le processus d'arborisation, celle de la RSyntP va permettre d'influencer des choix plus superficiels comme celui des dépendances syntaxiques de surface, qui, à leur tour, auront une incidence sur la linéarisation et le calcul des structures d'intonation.

D) Structure prosodique profonde

Finalement, la RSyntP possède une structure prosodique : c'est-à-dire, une représentation formelle de la prosodie nécessaire à l'expression du type de message encodé par la RSém. Dans notre exemple, cette structure prosodique profonde est élémentaire : elle se réduit à la mention *prosodie déclarative neutre* accompagnant l'arbre SyntP.

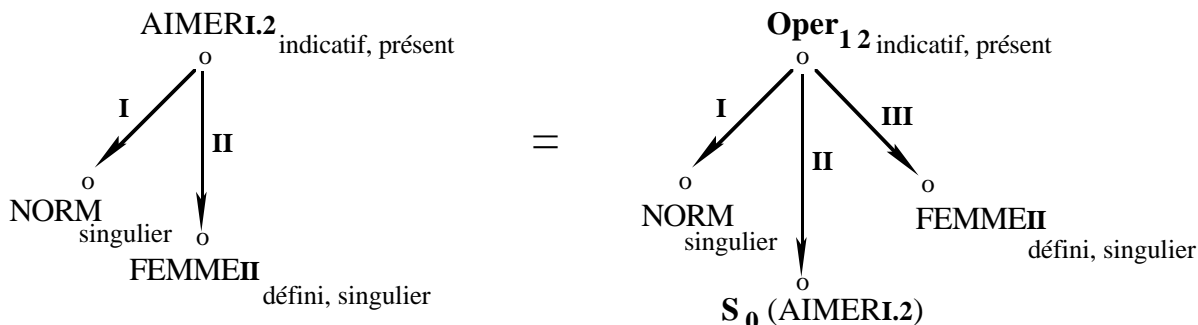
Pour conclure cette trop rapide présentation de la RSyntP, je voudrais mettre l'accent sur le fait que la Figure 4 est en réalité la RSyntP des trois premières phrases qui ont été données en (1) :

Norm aime follement sa femme Marge. [= (3a)]
Norm aime sa femme Marge à la folie.
Norm aime sa femme Marge comme un fou.

Les deux dernières phrases de (1) :

Norm éprouve un amour fou pour sa femme Marge.
Norm ressent un amour immense pour sa femme Marge.

ne peuvent être dérivées **directement** de cette RSyntP. Il faut, pour les construire, que le locuteur utilise l'égalité suivante afin de produire une autre RSyntP, à partir de la RSyntP de la Figure 4⁵ :



Le sous-arbre de droite fait appel à deux nouvelles fonctions lexicales :

- **S₀(AIMERI.2)** retourne le substantif correspondant au verbe en question, c'est-à-dire AMOURI.3 (numérotation du *Petit Robert*) ;
- **Oper₁₂** retourne la liste des verbes supports sémantiquement vides du type ÉPROUVER ou RESSENTIR (*de l'amour pour...*).

Je n'irai pas plus loin dans la description de ces fonctions lexicales. Le lecteur intéressé pourra trouver plusieurs références bibliographiques à ce sujet dans la Conclusion.

2.4 Choix et construction d'une Représentation Syntaxique de Surface

La transition vers la RSyntS va essentiellement impliquer les opérations suivantes :

- calcul des relations syntaxiques de surface, spécifiques à la langue en question, à partir des dépendances syntaxiques profondes ;
- choix parmi les valeurs possibles des fonctions lexicales présentes dans la RSyntP ;
- introduction des lexies vides (mots grammaticaux) nécessaires pour assurer la grammaticalité de la phrase (prépositions régies, etc.) ;
- pronominalisation ;
- construction des structures communicative, anaphorique et prosodique du niveau syntaxique de surface.

La RSyntS de la phrase (3a) se présente comme suit :

⁵ La nouvelle RSyntP (dérivée) est équivalente à celle de la Figure 4 en regard du Sens exprimé mais correspond à la structure syntaxique profonde des deux dernières phrases.

[Prosodie déclarative neutre]

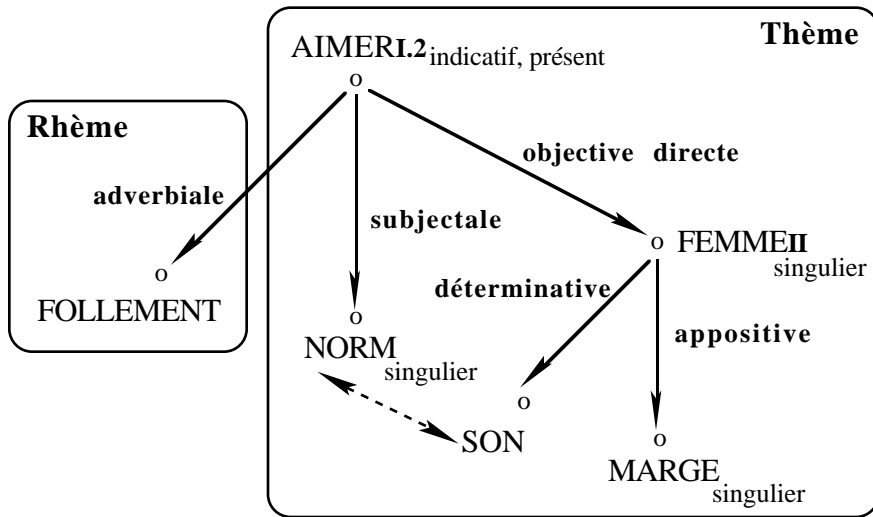


Figure 5 RSyntS de (3a)

On voit que cette représentation contient toutes les unités lexicales qui seront présentes dans le Texte final. Puisque le verbe AIMERI.2 est transitif direct, aucune préposition n'a été introduite pour « faire tenir » le sous-arbre syntaxique de surface correspondant à l'expression de son second actant sémantique. On remarquera que la pronominalisation a été effectuée, ce qui permet de répondre à une double besoin :

- 1- éviter la répétition de *Norm* dans la phrase (on aurait sinon *Norm aime follement la femme de Norm*, ce qui semble impliquer qu'il s'agit de deux Norm différents) ;
- 2- procurer un déterminant au nom commun FEMMEII.

2.5 Choix et construction d'une Représentation Morphologique Profonde

Formellement, la tâche principale à effectuer lors de la transition RSyntS ⇒ RMorphP est ce que l'on pourrait appeler la « linéarisation » de l'arbre syntaxique de surface. Celle-ci est schématisée dans la figure ci-dessous :

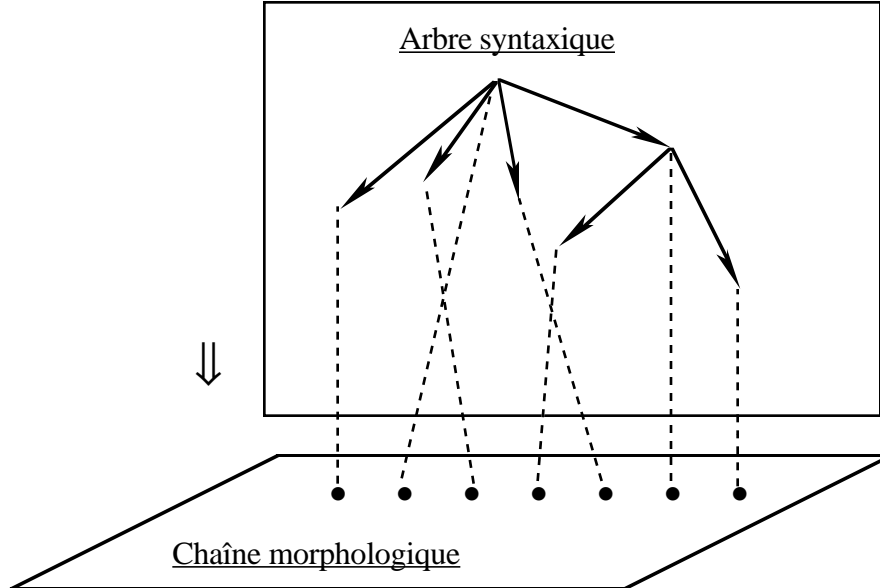


Figure 6 « Linéarisation » d'un arbre syntaxique

La Figure 6 ne doit pas tromper le lecteur : il ne s'agit pas ici de permuter l'ordre des nœuds de l'arbre syntaxique de surface afin d'obtenir l'ordre effectif des unités lexicales de la phrase. Rappelons que les nœuds des arbres syntaxiques Sens-Texte ne sont pas linéairement ordonnés, même si la représentation graphique doit nécessairement faire le choix d'une spatialisation gauche-droite donnée. La linéarisation de l'arbre ne consiste donc pas à changer un ordre préexistant, mais à **calculer** un ordre linéaire.

La linéarisation est l'opération la plus importante — la plus spectaculaire, en quelque sorte — effectuée à ce niveau de transition, mais elle n'est pas la seule. Au total, la transition RSyntS \Rightarrow RMorphP prend en charge les trois opérations suivantes :

- calcul de l'ordre linéaire des unités lexicales de la phrase ;
- calcul des différents accords morphologiques (par exemple, accord du verbe avec son sujet grammatical)⁶ ;
- calcul de la prosodie de la phrase, qui doit refléter adéquatement la structuration syntaxique, communicative et prosodique de surface.

Dans le cas de notre exemple, nous obtenons la RMorphP suivante :

⁶ C'est également à ce niveau que seront calculés les régimes casuels, pour les langues mettant en jeu des cas grammaticaux.

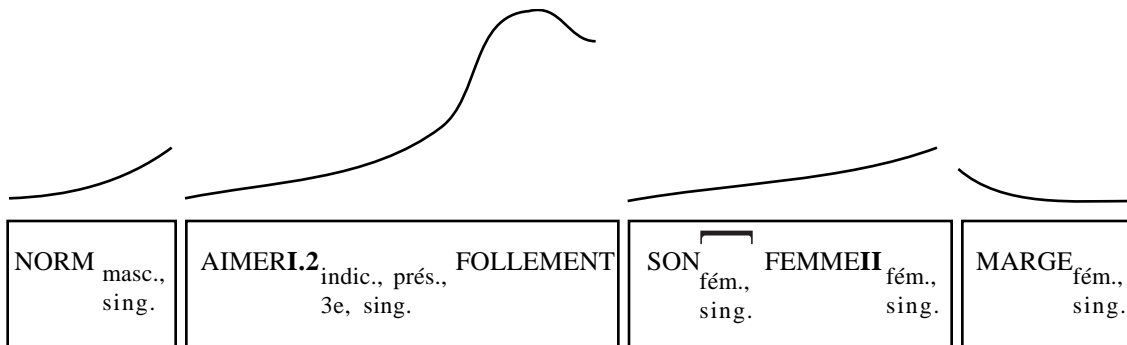


Figure 7 RMorphP de (3a)

Comme on le voit dans la figure ci-dessus, la RMorphP met en jeu deux types de structures :

- 1- La structure morphologique profonde proprement dite, qui est constituée de la chaîne des lexies de la phrase, portant les indications morphologiques pertinentes ;
- 2- la structure morphologique prosodique profonde, qui contient (i) l'indication des groupes prosodiques majeurs (ici, les quatre « boîtes »), (ii) les enchaînements prosodiques obligatoires (voir le lien entre SON et FEMMEII) et (iii) les schémas d'intonation (les courbes surmontant les blocs prosodiques)⁷.

À ce niveau de transition, des choix sont bien entendu opérés. Ainsi, à partir de la RSyntS de la Figure 5, un modèle Sens-Texte du français pourrait produire la RMorphP de la phrase (5), qui est distincte de la RMorphP de (3a) puisqu'elle présente un changement dans l'ordre linéaire des éléments de la phrase :

(5) *Norm aime sa femme Marge follement.*

Cet exemple illustre le fait que plus on s'éloigne du niveau de représentation sémantique, plus les choix disponibles sont limités et superficiels.

J'arrête ici la présentation des transitions mises en jeu dans la modélisation Sens-Texte, en espérant que la philosophie générale de l'approche est maintenant assez claire, au moins pour ce qui a trait à sa nature stratificationnelle. Dans le cadre d'un examen complet de la TST, il serait bien entendu nécessaire de présenter les autres niveaux de transition, mais je voudrais alléger ici au maximum mon exposé afin de consacrer le reste de l'article à la présentation du modèle Sens-Texte lui-même : c'est-à-dire à la présentation des règles permettant d'effectuer les transitions décrites ci-dessus. Je rappelle que la construction des modèles linguistiques (dictionnaires et grammaires) est la finalité même de l'étude linguistique Sens-Texte.

⁷ On peut noter, dans ce dernier cas, l'intonation marquée sur FOLLEMENT, qui va permettre d'exprimer sa nature rhématique — cf. la présentation de l'exemple (3a) en 2.2.

3. Structure d'un modèle Sens-Texte

Comme je l'ai mentionné précédemment, un modèle Sens-Texte d'une langue donnée peut être divisé en deux parties. La composante principale du modèle, qui comporte le dictionnaire et la grammaire, est un ensemble de règles déclaratives. Une règle est dite *déclarative* si elle ne correspond pas à une suite d'instructions, comme par exemple « Faire A puis B, puis si le résultat est C faire D ... ». En gros, on peut dire qu'une règle linguistique déclarative décrit un fait de langue. L'autre composante du modèle Sens-Texte est un ensemble de règles procédurales : ce sont des règles servant à activer les règles déclaratives de la composante centrale. Dans la présente section, je vais brièvement passer en revue quelques types de règles déclaratives (3.1) et procédurales (3.2). Le but de mon exposé sera de fournir quelques illustrations, sans plus, afin de donner un avant-goût de ce à quoi ressemble un modèle Sens-Texte. Il serait en effet impossible de présenter ici toutes les règles Sens-Texte utilisées pour effectuer les transitions entre les différentes représentations examinées dans la section précédente, ces règles se chiffrant, littéralement, par centaines.

3.1 Composantes déclaratives

3.1.1 Le dictionnaire Sens-Texte : Dictionnaire Explicatif et Combinatoire

Contrairement à la plupart des approches modernes en linguistique formelle, la TST voit le dictionnaire d'une langue comme le cœur du modèle de cette langue. Cela est en partie une conséquence de l'orientation sémantique de l'approche Sens-Texte, puisque le dictionnaire est avant tout le répertoire des significations de la langue. Il existe bien entendu des significations non lexicales — c'est-à-dire des significations grammaticales (temps grammatical, etc.) — qui doivent être décrites dans la grammaire de la langue ; mais les significations linguistiques se retrouvent pour la plupart dans l'ensemble des significations lexicales de la langue et sont décrites dans le dictionnaire. Le « linguiste Sens-Texte » est donc nécessairement un lexicographe. Je vais consacrer cette partie de l'article à une très brève description de ce qu'est un dictionnaire Sens-Texte.

Le dictionnaire postulé par la TST s'appelle *DICTIONNAIRE EXPLICATIF ET COMBINATOIRE* (DEC). Je vais le présenter en justifiant les deux qualificatifs *explicatif* et *combinatoire* qui lui sont associés.

Le DEC est **explicatif**, car il décrit le sens de l'unité lexicale au moyen d'une définition analytique qui décompose ce sens en termes de sens plus simples (c'est-à-dire, de sens qui sont plus proches des primitifs sémantiques que le sens de départ). L'explication du sens correspond donc très directement à une analyse : décomposition en termes d'éléments constitutifs. Il est impossible d'entrer ici dans le détail des principes qui régissent la construction des définitions du DEC — voir Mel'čuk *et al.* (1995), et je vais me contenter de proposer un exemple : la définition de AMOURI.3, qui est le substantif correspondant à AIMERI.2⁸.

⁸ On considérera que 'X aimeI.2 Y' est strictement équivalent à 'X a de l'amourI.3 pour Y'. En d'autres termes, il suffit de donner dans le dictionnaire la définition du nom et de renvoyer à celle-ci dans la définition du verbe. Pour ne

Ce tableau indique que le premier actant sémantique de AMOURI.3 ('X') se réalise en syntaxe profonde par le premier actant syntaxique profond de la lexie et en syntaxe de surface par un complément du nom introduit par *de* ou un déterminant du type adjectif possessif ; le deuxième actant sémantique ('Y') se réalise en syntaxe profonde par le deuxième actant syntaxique profond de la lexie et en syntaxe de surface par un complément du nom introduit par *envers* ou *pour*.

Il est important de remarquer que le tableau de régime équivaut à un ensemble de règles de transition qui permettent de construire les sous-arbres syntaxiques profond et de surface associés à la configuration sémantique correspondante. Par exemple, le tableau de régime de AIMERI.2 sera utilisé dans la construction des RSyntP/S des Figures 4 et 5.

La combinatoire lexicale est décrite au moyen des FONCTIONS LEXICALES, concept dont il a déjà été question en 2.3. Deux types de fonctions lexicales sont utilisés dans une entrée de DEC décrivant une lexie de la langue :

- 1- les fonctions lexicales paradigmatiques, qui connectent la lexie en cause aux autres lexies de la langue qui y sont directement reliées sémantiquement (les synonymes, antonymes, conversifs, etc.) ;
- 2- les fonctions lexicales syntagmatiques, qui décrivent la combinatoire lexicale proprement dite, c'est-à-dire les collocations contrôlées par la lexie en cause.

Je donne maintenant, sans explication additionnelle, une partie de la combinatoire lexicale de AMOURI.3, telle qu'elle doit être décrite dans l'entrée de DEC de cette lexie. Le lecteur frustré par le caractère approximatif et incomplet de ma présentation trouvera dans Mel'čuk *et al.* (1995:125-152) une description détaillée des fonctions lexicales utilisées ci-dessous.

FONCTIONS LEXICALES PARADIGMATIQUES :

Syn_n	: passion1; fam béguin; affection
Anti_n	: indifférence; répulsion < haine
Gener	: sentiment [d'~]
Magn+Figur	: feu [de l'~]
V₀	: aimerI.2
S₂	: amourI.5, fam béguin
etc.	

FONCTIONS LEXICALES SYNTAGMATIQUES :

Magn+A₁	: fou, malade [d'~]
en en souffrant Magn+A₁	: transi [d'~ sans modificateur]
Adv₁Manif	: avec ~
Magn	: tendre; grand antépos, fort; aveugle < enflammé, éperdu, passionné < infini postpos; fou
Magn^{temp}	: éternel
AntiMagn^{temp}	: passager postpos; 「de rencontre」
Oper₁	: vivre [ART ~ avec modificateur]

Oper₁₂	: éprouver, ressentir [de l'~ pour N=Y], Québ être [en ~ avec N=Y]
Magn+Oper₁₂	: brûler, se consumer, mourir [d'~ pour N=Y]
Ver+Oper₁₂	: filer [le parfait ~ avec N=Y]
Ver+Oper₁₊₂	: [N=X et N=Y] filer [le parfait ~]
IncepOper₁₂	: Québ tomber [en ~ avec N=Y]
Caus₂Func₁	: inspirer [ART ~ à N=X]
etc.	

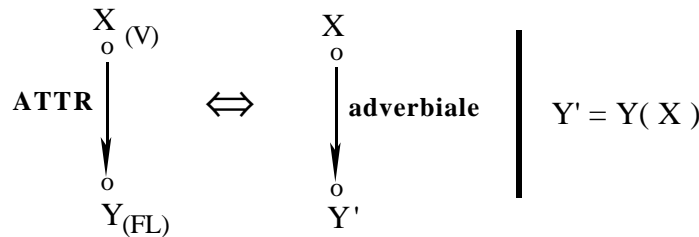
Cela termine ma présentation très succincte de la lexicologie explicative et combinatoire. Nous allons maintenant passer à un bref examen des principales caractéristiques de la composante grammaticale des modèles Sens-Texte.

3.1.2 La grammaire Sens-Texte

Le corps d'une grammaire Sens-Texte est constitué d'un ensemble de règles déclaratives de correspondance entre les niveaux de représentation adjacents postulés par la théorie. Par exemple, la composante syntaxique profonde d'une grammaire Sens-Texte est un ensemble de règles ayant le format suivant :

$$\text{Micro-structure SyntP} \Leftrightarrow \text{Micro-structure SyntS} \quad \Bigg| \quad \text{Conditions.}$$

Je ne citerai ici qu'un exemple de règle grammaticale de transition Sens-Texte, celle prenant en charge la traduction au niveau syntaxique de surface d'un modificateur du type fonction lexicale (FL) du verbe :



Ex. : *Norm aime[X] follement[Y'] sa femme Marge.*

Figure 8 Règle syntaxique profonde du français (simplifiée)

Cette règle n'est pas trop difficile à interpréter. Elle indique que le dépendant **ATTR** d'un verbe **X** qui est une fonction lexicale **Y** se réalise au niveau syntaxique de surface par une construction **adverbiale**, qui est une des valeurs possibles de la fonction lexicale **Y** pour **X**.

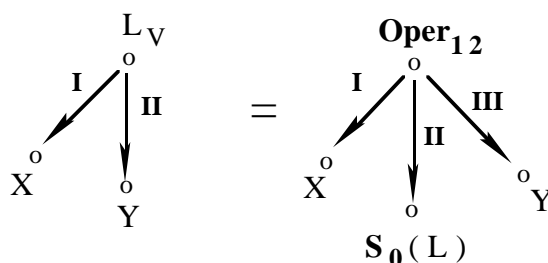
L'exemple de règle de la Figure 8 montre clairement qu'une grammaire Sens-Texte est avant tout un modèle de traduction assurant, par étapes successives, la transition du Sens au Texte.

Puisque la TST postule sept niveaux de représentation, les grammaires Sens-Texte possèdent six composantes (de transition) :

Composantes de la grammaire	Correspondances assurées
1) Sémantique	RSém ↔ RSyntP
2) Syntaxe profonde	RSyntP ↔ RSyntS
3) Syntaxe de surface	RSyntS ↔ RMorphP
4) Morphologie profonde	RMorphP ↔ RMorphS
5) Morphologie de surface	RMorphS ↔ RPhonP
6) Phonétique profonde	RPhonP ↔ RPhonS

Le lecteur intéressé trouvera dans Mel'čuk and Pertsov (1987) une syntaxe de surface (composante 3)) de l'anglais relativement complète. Il n'existe malheureusement pas à ce jour de travail équivalent publié pour le français.

En plus des règles assurant la transition entre niveaux de représentation, la grammaire Sens-Texte comprend des règles fonctionnant à un seul niveau ; ce sont en quelque sorte des règles de transformation structurale. Un bon exemple d'une telle règle est la règle de paraphrasage permettant de modéliser la relation de synonymie existant entre les constructions du type *aimer* et *éprouver de l'amour*, déjà examinées en 2.3 :



Ex. : $Norm[X] aime[L] sa femme[Y] = Norm[X] éprouve[Oper_{12}] de l'amour[S_0] pour sa femme[Y]$

Figure 9 Exemple de règle de paraphrasage

J'arrêterai là ce bref aperçu de la composante grammaticale des modèles Sens-Texte en rappelant au lecteur qu'il pourra trouver plusieurs renvois bibliographiques sur ce sujet dans la Conclusion de l'article (Section 4).

Nous passons maintenant à une brève caractérisation des règles procédurales nécessaires à l'activation des modèles Sens-Texte.

3.2 Composantes procédurales

Imaginez que l'on veuille écrire un programme informatique qui, en utilisant l'information contenue dans un DEC et une grammaire Sens-Texte du français, associe à la RSém de la Figure 2 (3a) et toutes ses paraphrases. Pour chaque niveau de transition (sémantique, syntaxique profond,

syntaxique de surface, etc.), il va falloir construire un ensemble de procédures permettant d'effectuer au moins les quatre opérations suivantes :

- 1- parcours (lecture) de la représentation de départ afin d'identifier les micro-structures à traduire ;
- 2- accès aux règles lexicales et grammaticales décrivant la traduction des micro-structures identifiées ;
- 3- construction des micro-structures de la représentation cible ;
- 4- réunion de ces micro-structures en un tout cohérent.

De telles procédures peuvent être fort complexes à déterminer et à implanter de façon formelle. On peut noter que les procédures d'activation des composantes sémantique et syntaxique de surface seront les plus sophistiquées. En effet, elles prennent en charge la transition entre des niveaux de représentation formellement très différents : des réseaux aux arbres de dépendance et des arbres de dépendance aux chaînes (linéaires).

Faute de place, je ne peux m'engager ici dans une véritable description des aspects procéduraux de la linguistique Sens-Texte. Encore une fois, je suggère au lecteur intéressé de consulter les différentes références bibliographiques citées dans la Conclusion, qui suit immédiatement.

4. Conclusion

Comme promis, je vais maintenant donner en guise de conclusion quelques références bibliographiques qui permettront au lecteur d'approfondir sa compréhension de l'approche Sens-Texte. Ces pointeurs bibliographiques sont regroupés en six catégories : autres textes généraux d'introduction, lexicologie/lexicographie Sens-Texte et fonctions lexicales, syntaxe, morphologie et, finalement, textes relatifs à l'informatisation et à l'activation procédurale des modèles Sens-Texte.

- 1) AUTRES TEXTES GÉNÉRAUX D'INTRODUCTION : Mel'čuk (1981), Mel'čuk (1993, Chapitre II), Steele (1990).
- 2) LEXICOLOGIE/LEXICOGRAPHIE SENS-TEXTE ET FONCTIONS LEXICALES : Mel'čuk *et al.* (1984, 1988 et 1992), Mel'čuk and Polguère (1987), Mel'čuk *et al.* (1995), Wanner (1996).
- 3) COMPOSANTE SÉMANTIQUE : Polguère (1992), St-Germain (1995).
- 4) COMPOSANTE SYNTAXIQUE : Mel'čuk and Pertsov (1987), Mel'čuk (1988).
- 5) COMPOSANTE MORPHOLOGIQUE : Mel'čuk (1982), Mel'čuk (1993).
- 6) INFORMATISATION DES MODÈLES SENS-TEXTE ET ASPECTS PROCÉDURAUX : Polguère (1990), Nasr (1996).

La liste donnée ci-dessus n'est en rien exhaustive. Je me suis limité aux textes écrits en français ou en anglais et qui me semblent pouvoir servir le plus adéquatement de point d'entrée dans la littérature Sens-Texte. On trouvera dans Mel'čuk *et al.* (1992:59-66) un répertoire détaillé des textes traitant de la linguistique Sens-Texte, préparé par Suzanne Mantha.

Remerciements

Je tiens à remercier mes collègues Sylvain Kahane, Marie-Claude L'Homme, Suzanne Mantha et Igor Mel'čuk pour leurs commentaires sur une première version de ce texte, dont la rédaction a été rendue possible grâce à une subvention FCAR 97-ER-2741 du Gouvernement du Québec ainsi qu'à une subvention du Conseil de Recherche en Sciences Humaines du Canada.

Références

- MEL'ČUK, I. A. 1981. « Meaning-Text Models : A Recent Trend in Soviet Linguistics », *Annual Review of Anthropology*, 10, 27-62.
- MEL'ČUK, I. A. 1982. *Towards a Language of Linguistics. A System of Formal Notions for Theoretical Morphology*, München, W. Fink.
- MEL'ČUK, I. A. 1988. *Dependency Syntax : Theory and Practice*, Albany, The SUNY Press.
- MEL'ČUK, I. A. 1993. *Cours de morphologie générale. Volume 1*, Montréal/Paris, Presses de l'Université de Montréal/CNRS Éditions.
- MEL'ČUK, I. A., A. CLAS et A. POLGUÈRE. 1995. *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot.
- MEL'ČUK, I. A. and N. PERTSOV. 1987. *Surface Syntax of English : A Formal Model in the Meaning-Text Framework*, Amsterdam/Philadelphia, Benjamins.
- MEL'ČUK, I. A. and A. POLGUÈRE. 1987. « A Formal Lexicon in the Meaning-Text Theory (or How to Do Lexica with Words) », *Computational Linguistics (Special Issue on the Lexicon)*, 13:3-4, 261-275.
- MEL'ČUK, I. A., N. ARBATCHEWSKY-JUMARIE, L. ELNITSKY, L. IORDANSKAJA et A. LESSARD. 1984. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques I*, Montréal, Presses de l'Université de Montréal.
- MEL'ČUK, I. A., N. ARBATCHEWSKY-JUMARIE, L. DAGENAIS, L. ELNITSKY, L. IORDANSKAJA, M.-N. LEFEBVRE et S. MANTHA. 1988. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques II*, Montréal, Presses de l'Université de Montréal.
- MEL'ČUK, I. A., N. ARBATCHEWSKY-JUMARIE, L. IORDANSKAJA et S. MANTHA. 1992. *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques III*, Montréal, Presses de l'Université de Montréal.
- NASR, A. 1996. « Un modèle de reformulation automatique fondé sur la Théorie Sens-Texte : Application aux langues contrôlées », thèse de doctorat, Université de Paris 7.
- POLGUÈRE, A. 1990. « Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte », thèse de doctorat, Université de Montréal.
- POLGUÈRE, A. 1992. « Remarques sur les réseaux sémantiques Sens \Leftrightarrow Texte », A. Clas (éd.): *Le mot, les mots... les bons mots*, Montréal, Presses de l'Université de Montréal, 109-148.
- SAUSSURE, F. de. 1972. *Cours de linguistique générale*, Paris, Payot.
- STEELE, J. (ed.). 1990. *Meaning-Text Theory : Linguistics, Lexicography and Implications*, Ottawa, University of Ottawa Press.
- ST-GERMAIN, J. 1995. « Incidence de la structure sémantique et communicative sur la structure syntaxique profonde des énoncés causatifs du français contemporain », thèse de doctorat, Université de Montréal.
- TESNIÈRE, L. 1965. *Éléments de syntaxe structurale*, Paris, Klincksieck.
- WANNER, L. (ed.) 1996. *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam/Philadelphia, Benjamins.
- ŽOLKOVSIIJ, A. et I. A. MEL'ČUK. 1965. « O vozmožnom metode i instrumentax semantičeskogo sinteza » [À propos d'une méthode possible pour la synthèse sémantique et des instruments correspondants], *NTI*, n° 5, 23-28.