

Grammaires graphiques

Nicolas Mazziotta, Sylvain Kahane

DANS TRAVAUX DE LINGUISTIQUE 2023/2 (N° 87), PAGES 111 À 127

ÉDITIONS DE BOECK SUPÉRIEUR

ISSN 0082-6049

ISBN 9782807380318

DOI 10.3917/tl.087.0111

Article disponible en ligne à l'adresse

<https://www.cairn.info/revue-travaux-de-linguistique-2023-2-page-111.htm>



CAIRN.INFO
MATIÈRES À RÉFLEXION

Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...

Flashez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour De Boeck Supérieur.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

GRAMMAIRES GRAPHIQUES

Nicolas MAZZIOTTA et Sylvain KAHANE*

Cette contribution porte sur l'expression formalisée de règles de grammaire d'une langue naturelle au moyen de diagrammes. Nous défendrons les idées fondamentales suivantes :

1. Les diagrammes syntaxiques sont des *objets formels* aptes à modéliser des règles de grammaire.
2. Un formalisme graphique déployé dans les deux dimensions est un *bon formalisme*. Son caractère spatial n'enlève rien à sa rigueur.
3. Les diagrammes permettent de réaliser des *opérations* au sens formel du terme. Cette notion d'*opération* est centrale.

Par ailleurs, les diagrammes des grammaires graphiques présentent certains avantages sur les formalismes s'appuyant sur des formules linéaires :

4. Ils ne sont pas astreints à une présentation séquentielle de l'information, selon l'unique dimension d'un texte ordinaire. En se déployant dans les deux dimensions de la page, ils permettent *différents parcours de lecture* de l'information (même si les conventions tendent généralement à privilégier un parcours de gauche à droite et de haut en bas).
5. Ils permettent une identification plus immédiate des différents objets en jeu, un *typage explicite* de ces objets par la forme et par les variables d'ajustement graphiques (teinte, saturation, valeur) et une mise en évidence des relations entre les objets par leurs positionnements respectifs.

Ce travail se situe dans la continuité de travaux précédents sur les relations entre constituance et dépendance (Kahane et Mazziotta, 2015), sur la manière dont la connaissance est représentée en linguistique (Mazziotta, 2016), sur l'histoire des diagrammes syntaxiques (Mazziotta et Kahane,

* Nicolas Mazziotta, Université de Liège, Traverses, nicolas.mazziotta@uliege.be ; Sylvain Kahane, Université Paris-Nanterre et Institut Universitaire de France, MoDyCo, U.M.R. 7114, skahane@parisnanterre.fr. Les deux auteurs ont contribué de manière équivalente à cette publication.

2017 ; Mazziotta, 2019, 2020 ; Kahane, 2020) et sur la modélisation graphique de règles de grammaire (Kahane, 2001, 2006, 2015). Nous reprenons ici la terminologie et les concepts théoriques que nous avons introduits dans ces études.

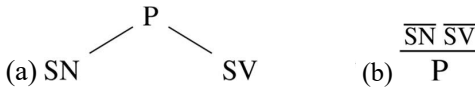
L'exposé qui suit est structuré en trois sections. Dans la première, nous présentons la notion d'*inscription* et nous expliquons en quoi un diagramme syntaxique peut être qualifié d'*inscription* d'une connaissance linguistique (section 1). La section suivante est consacrée aux premières grammaires formelles. Nous montrons déjà comment les diagrammes permettent de réaliser des opérations mentales et de calculer / générer des structures (section 2). La section 3 est quant à elle consacrée aux formalismes de *Tree adjoining grammar* (TAG ; Joshi, Levy et Takahashi, 1975) et de *Grammaire d'unification polarisée* (GUP ; Kahane et Lareau, 2005 ; Kahane, 2006, 2015), qui font intervenir les deux dimensions du plan pour rendre compte d'opérations linguistiques. La conclusion synthétise notre exposé (section 4).

1. Représenter la syntaxe : les diagrammes syntaxiques comme des *inscriptions*

Selon la théorie du support (Bachimont, 2007, 2010), la connaissance est abstraite et inaccessible si elle n'est pas convertie en une *inscription* au travers de laquelle elle est matérialisée. Ainsi, par exemple, écrire un livre sur un sujet permet de transmettre une connaissance qui, autrement, resterait inaccessible. Toutefois, en lui donnant une substance, l'inscription d'une connaissance hérite des propriétés et des limites du *medium*. Ainsi, un livre n'aura jamais d'hyperliens, alors qu'un site web contenant le même texte peut en avoir, mais ne pourra être feuilleté. Il est vrai que les inscriptions sont des phénomènes perceptibles par nos sens, non nécessairement visuels, toutefois nous nous intéresserons ici aux inscriptions visuelles de la connaissance (ou, plus précisément, aux inscriptions spatialisées dans un espace en deux dimensions). Dans cette section, nous allons envisager les diagrammes syntaxiques comme des inscriptions d'une certaine connaissance syntaxique, qui utilisent différents procédés pour inscrire cette dernière sur un plan.

Une même connaissance syntaxique peut ainsi être représentée dans différentes inscriptions. L'énoncé en langue naturelle de la règle *La phrase est la combinaison d'un syntagme nominal et d'un syntagme verbal* se veut une inscription discursive en langue naturelle de la même connaissance que celle inscrite par la règle de réécriture [1] (Chomsky, 1969 [1957]) et celle inscrite par les représentations bidimensionnelles arborescentes de la Figure 1.

[1] $P \rightarrow SN SV$

Figure 1 : Inscriptions bidimensionnelles d'une règle de dérivation

Bien entendu, les linguistes sont plus familiers avec l'une ou l'autre des notations selon leur formation et les habitudes associées aux écoles de pensée auxquelles ils adhèrent. Néanmoins, la différence entre les inscriptions n'est pas qu'une question de cosmétique ou de préférence : les propriétés du support impliquent que l'information véhiculée n'est pas exactement la même dans les trois cas, comme on le verra ci-dessous. En outre, les manipulations mentales que les inscriptions permettent ne sont pas non plus identiques ; p. ex. la Figure 1b renvoie à l'idée de *proportion* et favorise un raisonnement analogue à un raisonnement mathématique, alors que la Figure 1a met en évidence le caractère supérieur du concept de *phrase* (P) dans la description. Ce point sera également approfondi.

En accord avec la pensée du logicien C.S. Peirce, nous considérons qu'il faut qualifier de *diagrammes* l'inscription sous forme de règle de réécriture et celles de la Figure 1. En effet, Peirce décrit essentiellement les diagrammes comme des icônes de relations permettant le raisonnement créatif (voir l'introduction de ce volume). Les diagrammes inscrivent des entités conceptuelles ainsi que des relations entre ces entités. C'est la structure de l'inscription des relations qui est iconique de la structure des relations représentées. La règle de réécriture monodimensionnelle [1] inscrit principalement le concept de phrase sous la forme d'un symbole « P », ainsi que la relation entre cette phrase et ses constituants (inscrits par les symboles « SN » et « SV ») sous la forme d'une flèche positionnée entre le symbole « P » et les autres. On peut dès lors réfléchir à ce qui pourrait remplacer SN et construire une dérivation comme le fait Chomsky (1969 [1957] : 30). La Figure 1 montre que les mêmes relations peuvent être représentées de différentes manières sur un plan.

Comme les inscriptions en général, les diagrammes ne sont pas nécessairement des objets visuels, mais nous allons restreindre l'acception de *diagramme* à celle de *diagramme visuel* dans la suite de l'exposé – les diagrammes présentés seront pour la plupart planaires (bidimensionnels) et statiques (sans dimension temporelle). En effet, nous nous intéresserons uniquement à des signes qui sont à la fois des diagrammes et des signes visuels. Cela a pour conséquence principale, d'une part, que ces diagrammes sont déployés dans le plan et, d'autre part, qu'il y a deux moyens d'y inscrire la connaissance (Mazziotta, 2016) :

- par *réification*, sous la forme d'une entité graphique sémiotisée (Groupe μ , 1992), c'est-à-dire d'un signe discret, qui exprime en l'occurrence un concept, une unité, une opération ; p. ex. le symbole P réifie le groupe de mots qu'on appelle « phrase » ; les traits de la Figure 1a réifient les relations « partie-tout » (au détriment d'autres relations comme l'expression de l'extension) ;
- par *configuration*, c'est-à-dire par l'agencement relatif d'entités ; ainsi, dans les diagrammes de la Figure 1, comme dans [1], SN précède SV sur l'axe horizontal, ce qui inscrit iconiquement l'ordre linéaire des constituants.

Ces moyens sont complémentaires et interdépendants : on ne peut réifier un concept sans le positionner et toute configuration concerne nécessairement des concepts réifiés. Ils sont également *concurrents*, au sens où certains concepts peuvent être représentés soit par réification, soit par configuration. Ainsi la notation par parenthésage [2] réifie le groupe formé par SN et SV au moyen de parenthèses.

[2] P(SN SV)

Similairement, la ligne horizontale étiquetée P dans la Figure 1b s'étend de manière à couvrir la longueur occupée par les deux lignes étiquetées SN et SV auxquelles elle est parallèle. La Figure 1b réifie iconiquement (par des traits) l'information qui concerne la longueur relative des segments de textes identifiés par SN, SV et P : 1/ le début et la fin de P correspondent respectivement au début de SN et à la fin de SV ; 2/ la fin de SN correspond au début de SV. Cette identité de recouvrement n'est pas inscrite en tant que telle dans la règle de réécriture [1] – SN et SV sont à droite de la flèche – et dans la Figure 1 – SN et SV sont connectés à P par un trait et situés plus bas sur la page. Ces inscriptions alternatives ne sont pas complètement équivalentes. L'accès à cette information concernant la longueur des segments se fait de manière moins efficace et moins directe dans [1] et dans la Figure 1a que dans Figure 1b. Le choix de l'une ou de l'autre représentation dépendra des besoins de l'utilisateur des diagrammes (Mazziotta, 2021).

Tout l'enjeu de notre exposé est de montrer comment ces inscriptions diagrammatiques peuvent être utilisées pour représenter des opérations sur les termes linguistiques.

2. Des grammaires formelles aux grammaires graphiques

On peut définir une *grammaire formelle* comme un système formel linguistique, en entendant par *système formel* « ensemble défini d'unités et de

règles de combinaison de ces unités ». Pour une grammaire, les unités sont les mots, les syntagmes, etc., c'est-à-dire les signes linguistiques, alors que les règles de combinaison permettent de grouper ces signes linguistiques pour former des énoncés. Les grammaires formelles ont des propriétés algébriques, au sens où elles modélisent de manière abstraite des opérations sur leurs termes. Elles permettent de calculer, et donc de prouver, qu'une suite de mots est une phrase, ou de générer des structures. Ce calcul s'opère non sur une idée abstraite, mais directement grâce à des diagrammes au moyen desquels il est inscrit. C'est ce que nous allons illustrer dans cette section.

Nous prendrons pour exemple le fonctionnement des grammaires catégorielles (Ajdkiewicz, [1935] 2007; Bar-Hillel, 1953; Lambek, 1958). Ces dernières ont effectivement été pensées comme des systèmes formels permettant le calcul sur les langues. Selon cette approche, une langue est constituée d'un vocabulaire (les « mots ») et de catégories syntaxiques élémentaires (comme « P » ou « N »). Les deux opérateurs « / » et « \ » permettent de construire de nouvelles catégories à partir de ces catégories élémentaires : « P/N » est la catégorie qui correspond à une phrase à laquelle il manque un nom à droite, « N\P » est la catégorie qui correspond à une phrase à laquelle il manque un nom à gauche. Chaque mot du vocabulaire et chaque groupe de mots bien formé est associé à une catégorie : *Alfred* est un N, *Alfred dort* est une P, *dort* est une N\P, etc.

L'idée fondamentale est que le calcul sur la combinaison de ces catégories fonctionne similairement à un calcul algébrique sur des fractions : le diagramme [3] suit une logique similaire au calcul inscrit par [4].

$$[3] \quad (P/N) \cdot N = P$$

$$[4] \quad (7/3) \times 3 = 7$$

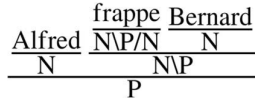
On peut ainsi prouver qu'*Alfred frappe Bernard* est une phrase, car vérifier qu'une suite de mots est une phrase devient un calcul algébrique, qui peut être noté en superposant les diagrammes impliqués dans le raisonnement, soit [5], comme dans une démonstration mathématique¹.

$$[5] \quad \begin{array}{l} \textit{Alfred} : N \\ \textit{Bernard} : N \\ \textit{frappe} : N\backslash P/N \\ \textit{frappe Bernard} : (N\backslash P/N) \cdot N = N\backslash P \\ \textit{Alfred frappe Bernard} : N \cdot (N\backslash P) = P \end{array}$$

Les diagrammes monodimensionnels (horizontaux) sont combinés selon une configuration spécifique à l'exposé de la démonstration qui inscrit de manière très implicite les relations entre eux : c'est au lecteur d'identifier la relation entre chaque ligne et la suivante.

Une inscription alternative de cette preuve est le diagramme de la Figure 2, qui combine en un seul les diagrammes plus petits comme $\frac{\text{Alfred}}{N}$ et $\frac{\text{frappe}}{N \backslash P / N}$.

Figure 2 : Inscription bidimensionnelle d'un calcul selon la grammaire catégorielle



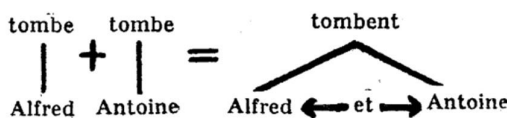
La Figure 2 inscrit les relations entre les catégories de la phrase de manière bidimensionnelle. Par rapport à la configuration précédente, pareil diagramme explicite les relations entre les diagrammes plus petits qui le constituent. Il n'impose pas de sens de lecture et se débarrasse d'une partie de la contingence associée à l'ordination des diagrammes : on peut ainsi lire la Figure 2 de haut en bas ou de bas en haut, ou se focaliser sur un de ses composants. Indépendamment du fait qu'il montre la logique algébrique du raisonnement, ce diagramme est similaire à un arbre de dérivation d'une grammaire de réécriture, lequel peut être présenté avec le symbole P en haut, ce qui privilégie alors une lecture du type « P peut être décomposé en N et N \backslash P » (Figure 1a) ou le symbole P en bas privilégiant une lecture du type « N et N \backslash P se combinent pour donner P » (Figure 1b).

La Figure 2 montre bien comment les différentes parties du calcul s'agrègent, mais d'autres types de diagrammes vont plus loin dans l'exploitation des deux dimensions pour inscrire les opérations entre diagrammes. Cette question fait l'objet de la section suivante.

3. Opérations graphiques

Les diagrammes peuvent donc être combinés pour effectuer un calcul sur la langue au travers d'une formalisation visuelle. Ce genre de calcul procède de l'utilisation d'opérations qui peuvent être réifiées, comme le suggérait déjà Tesnière (Tesnière, 1959 : chap. 50.10, diagramme unique en son genre dans l'ouvrage) :

Figure 3 : Réification de l'opération constituant la jonction chez Tesnière



Stemma 93

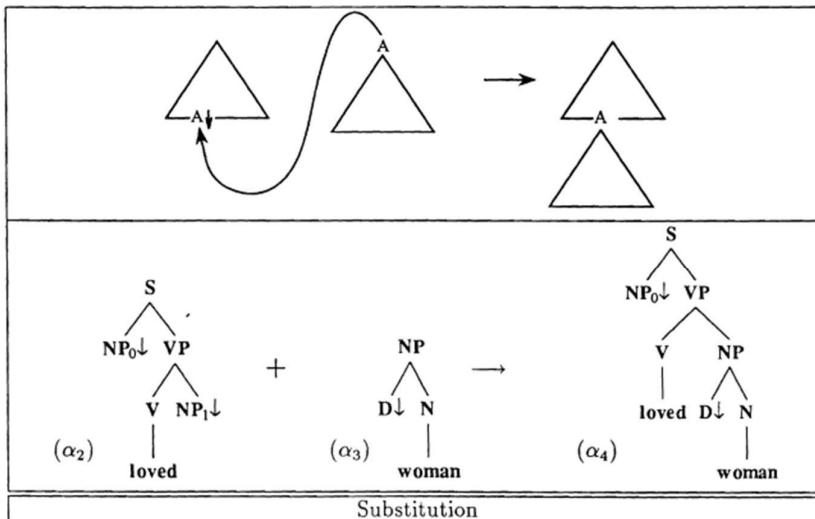
La Figure 3 peut être qualifiée de *métadiagramme* (Mazziotta, 2022), car elle inscrit des relations entre plusieurs diagrammes, ici au travers de la configuration relative des diagrammes concernés autour des symboles « + » et « = », qui réifient les opérations de combinaison et de mise en équivalence, dans une perspective assez métaphorique du fonctionnement de l'arithmétique fondamentale.

Dans cette section, nous présenterons tout d'abord le fonctionnement général de la *Tree adjoining grammar* (TAG), qui exploite explicitement la notion d'*opération* sur les concepts réifiées (3.1). Nous verrons ensuite comment nous proposons d'utiliser la *Grammaire d'unification polarisée* (GUP) pour effectuer des opérations directement sur la configuration des entités (3.2).

3.1. Opérations dans la *Tree adjoining grammar*

Si le diagramme de Tesnière est un hapax, TAG (Joshi, Levy, et Takahashi, 1975; Joshi et Schabes, 1991; pour une introduction, voir Abeillé, 1993) a exploité l'idée d'opération en définissant formellement deux d'entre elles : la substitution et l'adjonction. Cela permet de rendre compte de la valence et de l'équivalence distributionnelle des constructions syntaxiques (le terme *construction* étant entendu ici dans son sens général). Le principe général de TAG est que des arbres de constituants dits « élémentaires » peuvent se combiner pour former des arbres dérivés. Joshi et Schabes (1991 : 5) représentent la substitution par la Figure 4, dont la portion supérieure inscrit le fonctionnement de l'opération, alors que la partie inférieure l'illustre.

Figure 4 : Inscription bidimensionnelle de l'opération de substitution en TAG



TAG définit des symboles classiques comme « S », « NP » et « VP », qui sont combinables avec des entités qui réifient la possibilité de substitution : « $\alpha_0 \downarrow$ », « $\alpha_1 \downarrow$ », etc. Nous pourrions parler ici de *points d'ancrage visuels*. En lisant le métadiagramme de la portion inférieure de la Figure 4, on comprend que l'arbre élémentaire étiqueté « α_3 », dominé par le nœud « NP », se substitue au nœud « $\text{NP}_1 \downarrow$ » dans l'arbre étiqueté « α_2 ». Cette opération a pour résultat de former l'arbre étiqueté « α_4 ». Par la partie supérieure de la Figure 4, l'auteur nous invite à visualiser, au moyen d'une flèche qui le réifie, un mouvement dans l'espace. On comprend ainsi comment interpréter le symbole « + » : au départ positionné en dehors de α_2 , l'arbre élémentaire α_3 y est *déplacé* pour former l'arbre dérivé α_4 . L'opération porte donc sur la configuration des arbres.

3.2. Opérations dans la *Grammaire d'unification polarisée* : le cas des verbes à montée

Les idées d'arbres élémentaires et de réification des points d'ancrage sont également exploitées dans le formalisme de GUP (Kahane et Lareau, 2005; Kahane, 2006, 2015). La manière dont cette dernière modélise la transposition d'une structure sémantique en une structure syntaxique dépendancielle envisage toutefois la combinaison des diagrammes en prenant davantage en considération la configuration des entités sur le plan, selon un raisonnement spatial qu'on peut qualifier de « topologique » en référence aux mathématiques.

Le formalisme GUP est basé sur l'approche dépendancielle de la Théorie Sens-Texte (TST, Mel'čuk, 1988), qui distingue plusieurs modules de description linguistique. Parmi ces modules, ceux qui nous intéressent sont la structure sémantique (SSém) et la structure syntaxique dite « de surface ». Il faut noter trois aspects théoriques qui définissent la spécificité de l'approche que nous choisissons :

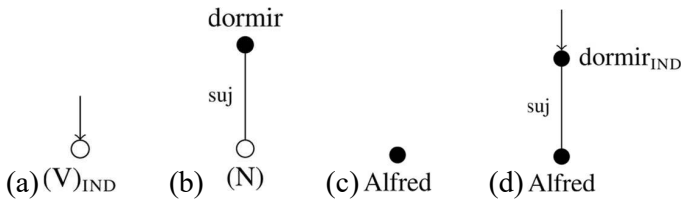
1. En TST, la SSém (Mel'čuk, 1988 : 53-58) est un graphe orienté équivalant à une hiérarchie de prédicats (origines des arcs dans le graphe) en relation avec leurs arguments (destination des arcs dans le graphe). Comme nous le verrons plus loin, les arêtes de la structure sont typées à l'aide d'un indice numérique qui correspond à l'ordre des arguments. La SSém ne tient pas compte des rôles sémantiques (« cas profonds », rôles-thêta, etc.). Comme on le verra plus bas avec notre exemple, la SSém se veut être un invariant de paraphrase, c'est-à-dire une structure commune à un ensemble de phrases de même sens².
2. La TST distingue normalement la structure syntaxique profonde (SSyntP), qui convient censément à la description de toutes les

langues, de la structure syntaxique de surface (SSyntS), spécifique à chaque langue. Notre exposé s'intéresse à cette dernière, que nous nommerons *structure syntaxique* pour plus de simplicité (SSynt) et qui prend la forme d'un arbre.

3. Nous commentons ici la perspective de *synthèse* (ou *génération*), orientée du sens au texte. Dans cette perspective, la SSém est définie préalablement à la SSynt qui la réalise. Le processus inverse pourrait évidemment être envisagé, mais il est moins utile à notre propos.

GUP fonctionne de manière similaire à TAG, au sens où les structures syntaxiques sont construites par assemblages de structures fondamentales. On peut l'illustrer par les diagrammes syntaxiques de la Figure 5.

Figure 5 : Structures syntaxiques GUP



En simplifiant, dans les diagrammes syntaxiques, les nœuds blancs sont des sortes de points d'ancrage similaires à ceux marqués par une flèche en TAG. Ils sont pourvus d'une étiquette identifiant une classe lexicale éventuellement associée à des informations grammaticales (dites « grammèmes » en TST). Dans chaque point d'ancrage peut venir s'insérer le nœud noir racine d'une sous-structure, qui est étiquetée à l'aide d'un mot appartenant à cette classe lexicale. Ainsi, la Figure 5a indique qu'un verbe à l'indicatif peut être la racine d'un arbre, la Figure 5b modélise la valence du verbe *dormir* car elle comporte un nœud blanc qui peut accueillir un nom, qui jouera le rôle de sujet (p. ex. : *Alfred*, Figure 5c). La combinaison de ces trois sous-structures permet de construire l'arbre de la Figure 5d.

Voyons comment GUP fonctionne et comment nous pouvons illustrer notre propos à l'aide du cas des verbes dits « à montée ». On qualifie ainsi les verbes dans [8], comme *semble* qui ne prennent pas leur sujet comme argument sémantique, comme le montrent les paraphrases [6-7]³.

- [6] Il semble qu'Alfred dorme.
- [7] Alfred dort, semble-t-il.
- [8] Alfred semble dormir.

On voit qu'il existe une correspondance entre l'emploi impersonnel du verbe, son emploi parenthétique et son emploi comme semi-auxiliaire. Dans

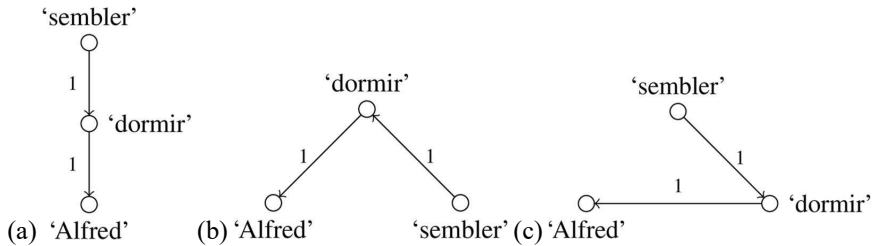
la phrase [8], on dit métaphoriquement que le mot *Alfred* est « monté » de la position de sujet de la subordonnée qu’il occupe dans [6] à la position de sujet de la principale. Cette terminologie, issue de la grammaire transformationnelle, implique que la phrase [6] est plus primitive que la phrase [8]. Nous allons donner ici une autre interprétation à la « montée » dans la relation entre les structures sémantique et syntaxique de [8].

Nous pouvons dire que, malgré leur SSynt différente, les trois phrases sont équivalentes du point de vue de leur SSém, au sens où elles sont des paraphrases les unes des autres. Le sens (commun) des trois phrases peut être modélisé comme une structure de prédicats [9]⁴.

- [9] e : ‘sembler’(e[^])
 e’ : ‘dormir’(x)
 x : ‘Alfred’

En TST, on considère généralement que ce genre de SSém peut être représentée indifféremment par n’importe lequel des graphes de la Figure 6, qui ont chacun une structure complètement identique à celle des deux autres graphes – même nombre de nœuds, mêmes arcs entre ces nœuds et mêmes étiquettes pour les nœuds et les arêtes. Si bien que la différence de configuration des entités de ces graphes est le plus souvent vue comme une caractéristique contingente.

Figure 6 : Inscriptions bidimensionnelles alternatives d’une même SSém en TST

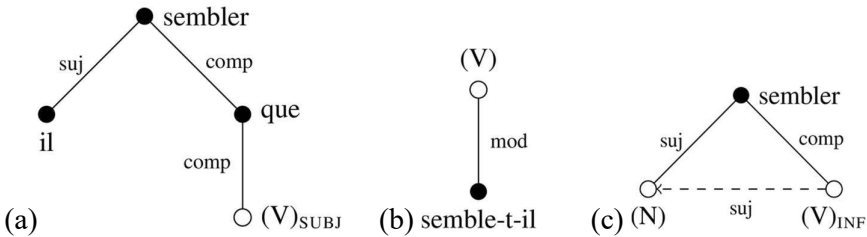


Nous pouvons utiliser GUP pour introduire des *distinctions*, au sens saussurien du terme, entre ces diagrammes habituellement considérés comme équivalents. Cela apparaît quand nous mobilisons la configuration de la SSém dans une perspective de synthèse pour expliquer le passage de la SSém aux trois SSynt possibles. On peut voir la synthèse, c’est-à-dire le passage du sens à un texte, comme la nécessité de hiérarchiser, puis de linéariser l’information que l’on souhaite communiquer. On peut projeter cette hiérarchisation de l’information sur les graphes de la Figure 6 en donnant maintenant à la dimension verticale une valeur configurationnelle.

Les diagrammes de la Figure 6 ne se lisent plus alors comme de purs graphes (au sens mathématique du terme, où seuls les contacts entre nœuds et arêtes sont signifiants, indépendamment de leur position dans le plan), mais comme des structures plus riches comportant à la fois une structure de graphe et une structure hiérarchique.

Si les trois phrases [6-8] ont en commun le même sens abstrait donné en [9] et inscrit de trois manières différentes par les diagrammes de la Figure 6, elles correspondent à des hiérarchisations différentes de l'information que l'on peut symboliser par les différentes configurations verticales du graphe de la Figure 6. Cela devient évident si l'on met en évidence les SSynt élémentaires associées aux différentes réalisations du sens 'sembler' proposées dans la Figure 7. Soulignons que de tels diagrammes constituent, dans une grammaire GUP, l'unique inscription du fonctionnement des lexèmes, illustrée ici par trois constructions du verbe *sembler*. D'autres inscriptions, sous forme de formules ou de code informatique sont bien sûr possibles, mais l'inscription sous forme de diagramme suffit à constituer une formalisation.

Figure 7 : Sous-structures GUP correspondant à la différence entre les SSynt de [6], [7] et [8]



Dans la structure (a), le sémantème 'sembler' est réalisé comme la racine d'une proposition impersonnelle et son unique argument sémantique est réalisé par un verbe subordonné au subjonctif ; dans la structure (b), ce même sémantème correspond au modifieur *semble-t-il* de *dormir*⁵ ; la structure (c), qui illustre la « montée », correspond au déplacement de la fonction sujet⁶ d'*Alfred*, de *dormir* à *sembler*.

La comparaison visuelle entre les inscriptions alternatives de la SSém (Figure 6) et les diagrammes de la Figure 7 fait ressortir un parallélisme évident. Dans la Figure 6a, 'sembler' est l'élément supérieur et 'Alfred' n'est pas situé directement en-dessous, comme les nœuds correspondant à ces sémantèmes dans la Figure 6a. Dans la Figure 6b, 'sembler' se trouve en-dessous de 'dormir', ce qui correspond à la position hiérarchique inférieure de *sembler* par rapport à *dormir* dans la Figure 7b. Dans la Figure 6c, 'Alfred' est placé directement sous 'sembler', ce qui

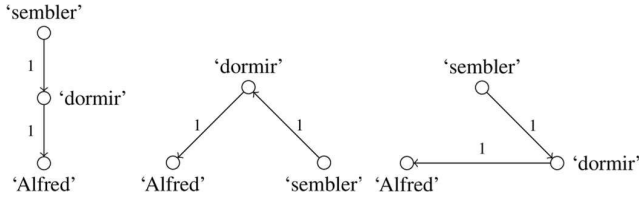
correspond à la « montée », représentée dans la Figure 7c par une flèche discontinue. Toutefois, alors que la dimension verticale est signifiante dans les diagrammes syntaxiques de la Figure 7, qui sont hiérarchiques par nature, la disposition des entités sur cette dimension est *contingente* lorsqu'on représente par un graphe le sens sous-jacent commun aux trois phrases [6-8]. De la même façon, la mobilisation de la dimension horizontale n'est pas *a priori signifiante* dans les Figure 6 et 7 (elle ne sert qu'à empêcher les superpositions, cf. Mazziotta, 2019 : sect. 2.2), alors qu'elle l'était dans les diagrammes de la Figure 1. Il est vrai que même si la dimension horizontale n'est pas signifiante *a priori*, il est très difficile pour le scripteur comme pour le lecteur de s'en abstraire totalement. Pour des raisons de lisibilité du diagramme, l'ordre dans lequel sont placés les nœuds qui dépendent d'un même nœud n'est jamais complètement arbitraire. En effet, il suit habituellement l'ordre des éléments dans l'énoncé modélisé. Ce genre d'habitude de représentation, convoquée à des fins le plus souvent ergonomiques, ne participe cependant pas à la structure signifiante du diagramme.

La description du processus de synthèse qui permet de générer chacune des SSynt à partir de la SSém qui leur est commune peut profiter de l'inscription bidimensionnelle de cette dernière pour définir la topologie des nœuds du graphe de manière pertinente. La première étape de la synthèse consiste à sélectionner la configuration qui correspond à la SSynt ciblée. Il y a donc une opération topologique qui est effectuée sur l'inscription de la SSém, et dont le résultat est une configuration spécifique, sémiotiquement (et grammaticalement) distincte des configurations alternatives. Une fois cette configuration déterminée, la sous-structure correspondante peut être sélectionnée et combinée aux sous-structures communes. Les étapes de la synthèse de la SSynt de [6] à partir de la SSém peuvent être représentées, comme inscrit dans la Figure 8, où chaque opération est explicitée.

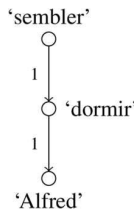
Comme toute inscription diagrammatique suppose une disposition des entités sur le plan, il est souvent nécessaire de choisir entre plusieurs alternatives. En TST, l'inscription d'une SSém n'impose rien d'autre que la distinction entre les nœuds et les arcs et laisse une grande latitude pour le placement de ces derniers. La pertinence de la configuration spatiale n'est généralement pas renseignée dans les diagrammes. Ainsi, en tant que systèmes diagrammatiques, la TST et GUP ne rendent compte iconiquement que d'un raisonnement au sein duquel la Figure 6a est équivalente à la Figure 6b. Dans ce genre de raisonnement, la disposition des entités sur le plan où est inscrite la SSém n'est pas corrélée à la SSynt. GUP laisse cependant à l'utilisateur la possibilité de sémiotiser lui-même les différentes dispositions s'il le désire : il le fait en enrichissant le système. Il serait sans doute utile de formaliser davantage l'iconicité topologique

pour montrer qu'il y a différentes manières d'organiser la SSém de façon à rendre compte de la SSynt vers laquelle la synthèse est orientée.

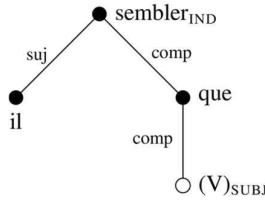
Figure 8 : Étapes de la synthèse SSém → SSynt pour [6]



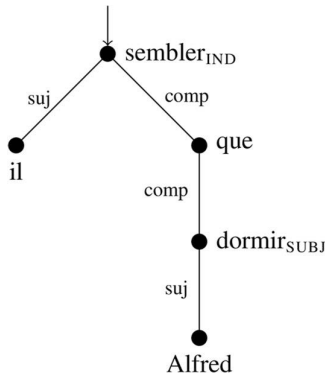
↓ Sélection de la topologie de la SSém



↓ Choix de la sous-structure SSynt élémentaire appropriée



↓ Saturation des nœuds blancs



4. Conclusion

Nous avons montré que les grammaires formelles peuvent être inscrites à l'aide de diagrammes. Ces derniers permettent de représenter les opérations impliquées dans des calculs comme la vérification de la conformité de structure d'une phrase (2) ou la génération d'une structure syntaxique à partir d'une structure sémantique dans une perspective de synthèse (3.2). Ces opérations peuvent consister en la substitution ou l'adjonction d'entités (*cf.* TAG ; 3.1), mais également en la modification / sélection de configurations spécifiques (3.2).

Les opérations figurent le raisonnement en l'inscrivant dans un métadiagramme (3), c'est-à-dire un diagramme qui inscrit des relations entre plusieurs diagrammes. Le métadiagramme réifie les opérations de manière symbolique et rend possible l'inscription de déplacements topologiques, soit en les réifiant (Figure 4), soit en montrant, par comparaison, qu'il est possible de choisir entre différentes alternatives (Figure 6).

La recherche de la connaissance inscrite par les diagrammes est une recherche des distinctions pertinentes entre les procédés de réification et de configuration, qui font système. Le problème de la distinction des unités discrètes, connu depuis les débuts de la linguistique structurale, s'applique aisément à nos objets. Ainsi, les procédés de réification, qui impliquent la forme ou la texture des objets, permettent de distinguer des signes par leur forme d'expression. Les exemples sont nombreux : remplissage blanc *vs* remplissage noir des disques dans le formalisme GUP, trait continu *vs* trait discontinu des arêtes de la Figure 7c, etc. La réflexion sur les opérations qui portent sur la topologie des diagrammes (3.2) nous a en outre permis d'illustrer la question de la pertinence / contingence des dimensions horizontale et verticale, c'est-à-dire au niveau des procédés de configuration. En conséquence, pour qu'une grammaire graphique soit un véritable système formel, il est essentiel que soient précisés sans ambiguïté les éléments signifiants du diagramme, qu'ils relèvent de procédés de réification ou de configuration.

En dernier lieu, nous voudrions envisager une piste de prolongement. La capacité qu'ont les formalismes graphiques déployés dans les deux dimensions du plan de rendre compte des opérations en manipulant les entités et les configurations les rend efficaces pour mener des raisonnements créatifs du même type que celui que nous avons mené dans la section 3.2. Cette propriété invite en outre à penser la capacité de manipuler des diagrammes en mobilisant une troisième dimension, non spatiale, mais temporelle, au moyen de séquences vidéos. Celles-ci permettraient de réifier, par l'inscription d'un déplacement, comment plusieurs structures peuvent fusionner en une seule ou comment la topologie d'une structure peut être modifiée. Nous laissons cette question à de futures études.

NOTES

1. Lambek (1958) a remarqué que les catégories pouvaient être interprétées comme des implications (NVP signifiant « si tu me donnes un N à gauche, alors nous formerons ensemble un P ») et le calcul comme une règle de déduction de type *modus ponens* (« Si p, alors q. Or p. Donc q »). Le calcul devient alors réellement une preuve. À noter que cette preuve appartient à une logique sensible aux ressources (si on donne deux fois un N, NVP ne pourra en consommer qu'un) et à l'ordre (c'est seulement si on donne un N avant que NVP pourra le consommer).

2. Mel'čuk (1988 : 52) rappelle que le sens est défini à partir du concept « avoir le même sens », qui constitue ainsi une primitive.

3. L'idée (généraliste) derrière cette notion est que les verbes à montée ont un sujet explétif et qu'ils ne peuvent apparaître dans une construction avec leur vrai sujet que si ce dernier est déplacé (« monté ») depuis la conjonctive que le verbe régit. Abeillé (1998) a fourni une description de ce problème de syntaxe française selon le formalisme TAG, où elle met en évidence le fait que la description purement syntaxique ne permet pas d'expliquer complètement les contraintes qui portent sur le déplacement du sujet.

4. Selon cette inscription, 'Alfred' est l'unique argument du prédicat 'dormir' et ce dernier est lui-même argument du prédicat 'sembler'.

5. L'approche étant dépendancielle, il n'est pas possible de faire dépendre *semble-t-il* du reste de la phrase dans son ensemble.

6. Sans rentrer dans les détails techniques, la flèche au trait discontinu indique que la relation sujet de *dormir*, qu'on voit dans la Figure 5a, sera supprimée par *sembler*.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ABEILLÉ A., 1993, *Les nouvelles syntaxes: grammaires d'unification et analyse du français*, Paris, Armand Colin.
- ABEILLÉ A., 1998, « Verbes « à montée » et auxiliaires dans une grammaire d'arbres adjoints », *Linx*, 39, p. 119-158. [<https://doi.org/10.4000/linx.882>]
- AJDUKIEWICZ K., 2007 [1935], « La connexion syntaxique », trad. de GAN-KRZYWOSZYŃSKA K., *Philosophia Scientiae*, 11, 2, p. 97-120. [<https://doi.org/10.4000/philosophiascientiae.344>]
- BACHIMONT B., 2007, *Ingénierie des connaissances et des contenus: le numérique entre ontologies et documents*, Paris, Lavoisier.
- BACHIMONT B., 2010, *Le sens de la technique: le numérique et le calcul*, 1, Paris, Encre marine / Les Belles Lettres.
- BAR-HILLEL Y., 1953, « Some linguistic problems connected with machine translation », *Philosophy of Science*, 20/3, p. 217-225. [<https://doi.org/10.1086/287266>]
- CHOMSKY N., 1969 [1957], *Structures syntaxiques*, trad. de BRAUDEAU M., Paris, Le Seuil, [édition anglaise : 1957].
- GROUPE μ , 1992, *Traité du signe visuel: pour une rhétorique de l'image*, Paris, Le Seuil.

- JOSHI A. K., LEVY L. S. et TAKAHASHI M., 1975, « Tree Adjunct Grammars », *Journal of Computer and System Sciences*, 10/1, p. 136-163. [[https://doi.org/10.1016/S0022-0000\(75\)80019-5](https://doi.org/10.1016/S0022-0000(75)80019-5)]
- JOSHI A. K. et SCHABES Y., 1991, « Tree-Adjoining Grammars and Lexicalized Grammars », *MS-CIS-91-22*, Philadelphia, Department of Computer and Information Science, School of Engineering and Applied Science, University of Pennsylvania. [https://repository.upenn.edu/cis_reports/445/]
- KAHANE S., 2001, « Grammaires de dépendance formelles et théorie Sens-Texte », *Actes de TALN*, 2, p. 17-76.
- KAHANE S., 2006, « Polarized Unification Grammars », in *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the ACL (ACL-COLING)*, Association for Computational Linguistics, p. 137-144. [<https://doi.org/10.3115/1220175.1220193>]
- KAHANE S., 2015, « Les trois dimensions d'une modélisation formelle de la langue : syntagmatique, paradigmatique et sémiotique », *TAL*, 56/1, p. 13-37.
- KAHANE S., 2020, « How dependency syntax appeared in the French *Encyclopedia*: from Buffier (1709) to Beauzée (1765) », in IMRÉNYI A. et MAZZIOTTA N., *Chapters of Dependency Grammar: A Historical Survey from Antiquity to Tesnière*, Amsterdam, John Benjamins, p. 85-131. [<https://doi.org/10.1075/slcs.212.05maz>]
- KAHANE S. et LAREAU F., 2005, « Grammaire d'Unification Sens-Texte: modularité et polarisation », in *Actes de TALN*, p. 23-32.
- KAHANE S. et MAZZIOTTA N., 2015, « Syntactic polygraphs. A formalism extending both constituency and dependency », in *Proceedings of the 14th Meeting on the Mathematics of Language (MoL 2015)*, Association for Computational Linguistics, p. 152-164. [<https://doi.org/10.3115/v1/W15-2313>]
- LAMBEK J., 1958, « The mathematics of sentence structure », *The American Mathematical Monthly*, 65/3, p. 154-170. [<https://doi.org/10.1080/00029890.1958.11989160>]
- MAZZIOTTA N., 2016, *Représenter la connaissance en linguistique. Observations sur l'édition de matériaux et sur l'analyse syntaxique*, Mémoire de synthèse d'Habilitation à diriger des recherches, Université Paris Nanterre. [<https://hdl.handle.net/2268/204408>]
- MAZZIOTTA N., 2019, « The evolution of spatial rationales in Tesnière's stemmas », in *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest)*, Association for Computational Linguistics, p. 69-80. [<https://doi.org/10.18653/v1/W19-7709>]
- MAZZIOTTA N., 2020, « Dependency in early sentence diagrams: Stephen W. Clark », in IMRÉNYI A. et MAZZIOTTA N., *Chapters of Dependency Grammar: A historical survey from Antiquity to Tesnière*, Amsterdam, John Benjamins, p. 134-62. [<https://doi.org/10.1075/slcs.212.05maz>]
- MAZZIOTTA N., 2021, « Drawing the syntactic space: choices in diagrammatic reasoning », in *Proceedings of the Sixth International Conference on Dependency Linguistics (Depling, SyntaxFest)*, Sofia (Bulgaria), Association for Computational Linguistics, p. 81-90. [<https://aclanthology.org/2021.depling-1.8>]

- MAZZIOTTA N., 2022, « Employer les diagrammes pour raisonner : usage dynamique des stemmas chez Lucien Tesnière », in NEVEU F. et ROIG A., *L'œuvre de Lucien Tesnière*, De Gruyter, p. 45-66. [<https://doi.org/10.1515/9783110715118-003>]
- MAZZIOTTA N. et KAHANE S., 2017, « To what extent is Immediate Constituency Analysis dependency-based? A survey of foundational texts », in *Proceedings of the Fourth International Conference on Dependency Linguistics (Depling 2017)*, Pisa, Linköping University Electronic Press, p. 116-126. [<https://aclanthology.org/W17-6515>]
- MEL'ČUK I. A., 1988, *Dependency syntax: theory and practice*, Albany, State University of New York Press.
- TESNIÈRE L., 1959, *Éléments de syntaxe structurale*, Paris, Klincksieck.