

Des outils graphiques pour étudier le langage et les langues. Les diagrammes en linguistique

Nicolas Mazziotta, Jacques François, Sylvain Kahane

DANS **TRAVAUX DE LINGUISTIQUE** 2023/2 (N° 87), PAGES 7 À 28
ÉDITIONS **DE BOECK SUPÉRIEUR**

ISSN 0082-6049

ISBN 9782807380318

DOI 10.3917/tl.087.0007

Article disponible en ligne à l'adresse

<https://www.cairn.info/revue-travaux-de-linguistique-2023-2-page-7.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.

I. TRAVAUX

DES OUTILS GRAPHIQUES POUR ÉTUDIER LE LANGAGE ET LES LANGUES. LES DIAGRAMMES EN LINGUISTIQUE

Nicolas MAZZIOTTA, Jacques FRANÇOIS, Sylvain KAHANE *

Si les diagrammes sont des outils d'analyse, ils peuvent faire eux-mêmes l'objet d'analyses visant à en expliquer le fonctionnement. De manière générale, outre les écrits de Peirce et de ses commentateurs, plusieurs travaux généraux, comme *La sémiologie graphique* de Jacques Bertin (2005 [1967]) ou les ouvrages d'Edward R. Tufte (2012) ont contribué à mettre au jour les procédés à l'œuvre dans les représentations diagrammatiques, ainsi que les stratégies développées pour faire de l'image un outil d'heuristique, de démonstration et de pédagogie (ou de communication ciblée en général). Les sémioticiens, en particulier les spécialistes de Peirce comme Frederik Stjernfelt (*Diagrammatology*, 2007), ont montré comment le raisonnement et le processus de découverte scientifique étaient associés à l'utilisation de diagrammes. Les spécialistes des inférences diagrammatiques ont proposé des modèles qui permettent de rendre compte de leur puissance informationnelle (Larkin et Simon, 1987 ; Shimojima, 1996, 2015 ; Stapleton *et al.*, 2017). Les diagrammes permettent de donner du sens aux données complexes organisées, qui sont devenues un enjeu

* Nicolas Mazziotta, Université de Liège, Traverses, nicolas.mazziotta@uliege.be ; Jacques François, Université de Caen-Normandie, CRISCO, E.A. 4255, jfrancois@interlingua.fr ; Sylvain Kahane, Université Paris-Nanterre et Institut Universitaire de France, MoDyCo, U.M.R. 7114, skahane@parisnanterre.fr.

crucial de la recherche dès la seconde moitié du 20^e s., quand la science prend la pleine mesure des analyses multivariées (Lima, 2011 : 45). À cet égard, à l'heure de la mode de « big data », les outils visuels revêtent une importance capitale (Dondero, 2019, 2020).

Les démarches réflexives sur la forme et l'utilisation des diagrammes sont restées peu nombreuses en linguistique : la thèse de Brittain (1973) sur les premiers diagrammes syntaxiques, le livre de Stewart (1976) sur les représentations graphiques en linguistique sont des hapax dans l'histoire de la discipline. Les colloques consacrés à la question sont assez rares (*Les linguistes et leurs graphiques, Cahiers Parisiens* n° 5, 2009), mais la linguistique apparaît régulièrement dans les contributions du colloque international *Diagrammatic representation and inference* (depuis 2000). Le récent ouvrage de Noah Bubenhofer, *Visuelle Linguistik* (2020) est encourageant, mais constitue une approche du domaine parmi d'autres possibles. Similairement, l'article programmatique de Smessaert *et al.* (2018), les travaux de Kahane et Mazziotta sur les diagrammes syntaxiques (Kahane et Mazziotta, 2015 ; Mazziotta et Kahane, 2017) correspondent à une analyse qui leur est spécifique et qui ont mené à des études des diagrammes de Tesnière (Mazziotta, 2019, 2022) et de grammairiens américains du 19^e siècle (Mazziotta, 2020a, 2020b). Des synthèses générales voient le jour sur certains types de diagrammes, comme en témoigne l'article de Georgakopoulos et Polis (2018) et il faut souhaiter que de multiples études particulières et générales viendront enrichir ces premières études encourageantes. Le domaine est loin d'être épuisé et les enjeux pratiques sont cruciaux. Par exemple, les développements de bases de données arborées ont rendu nécessaire et fréquente la mise en place d'algorithmes de génération de diagrammes syntaxiques¹. Les problèmes informatiques associés au traitement et à la mise en page des graphes sont largement discutés dans la littérature spécialisée, mais la démarche de construction visuelle spécifique à nos besoins de linguistes n'est généralement pas décrite.

Le présent volume se donne pour but, au travers d'une sélection d'articles dont cette introduction montrera la complémentarité : 1/ de donner un aperçu des enjeux de l'utilisation de diagrammes en linguistique ; 2/ de mettre à profit cet aperçu pour proposer des pistes de constitution d'un inventaire analytique des représentations graphiques en linguistique, prolongeant ainsi les démarches similaires que nous venons de présenter. La suite de notre exposé commence (section 1) par l'examen d'un diagramme classique élaboré selon les principes de la théorie X-barre. Nous montrerons que les diagrammes possèdent à la fois des propriétés visibles et des propriétés cachées, liées à la manière dont est théorisée la description. Nous illustrons une sélection de représentations graphiques utilisées en linguistique dans la section 2, en montrant notamment que certains de ces schémas sont utilisés dans d'autres domaines du savoir. Pédagogiques,

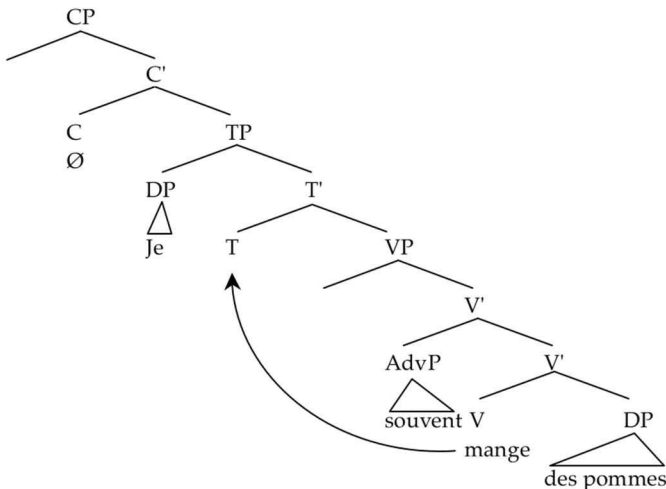
manipulables et formalisants, ces outils visuels sont hautement structurants. Ils sont même, si l'on en croit Charles S. Peirce, la condition nécessaire au raisonnement. La section 3 est consacrée à la notion de *raisonnement diagrammatique* telle que conçue par ce père de la sémiotique. Enfin, la construction et la lecture des diagrammes ne vont pas de soi : elles reposent sur des *conventions*, des règles qu'il faut apprendre pour en saisir le fonctionnement et discerner l'essentiel de l'accessoire (section 4). La section 5 propose un résumé des articles rassemblés.

1. Diagrammes connus... aux propriétés cachées

La plupart des étudiants et des chercheurs en langues et lettres ont été confrontés au moins une fois dans leur formation ou dans leur carrière à un diagramme comme celui de la Figure 1, qui représente l'analyse de la phrase [1] selon l'approche X-barre, qui nous servira de point de départ². Nous commençons par commenter ce type de diagramme car ils sont emblématiques d'un basculement méthodologique en linguistique qui s'est opéré dans la 2^e moitié du 20^e siècle, même si, comme le montrera le présent volume, les raisonnements diagrammatiques en linguistique sont antérieurs à cette période.

[1] Je mange souvent des pommes.

Figure 1 : Diagramme X-barre illustrant la notion de mouvement (Carnie, 2007 [2002] : 247)



X-barre repose sur une analyse dite « en constituants », selon laquelle les mots, dits « constituants terminaux », sont groupés en constructions³ dont ils sont les constituants immédiats. Ces constructions peuvent être elles-mêmes groupées à d'autres mots ou à d'autres constructions. Ce principe récursif est représenté par des traits / flèches et des étiquettes nommant les mots (*Je, souvent, mange, des, pommes*) et les constructions (*CP, C', TP, T'*, etc.) agencés les uns par rapport aux autres selon certaines règles. Ainsi, dans la Figure 1 on peut dire que la construction *V'* a *V* et *DP* comme constituants immédiats. Ces principes de représentation sont décrits dans de nombreux manuels de syntaxe générative dans le but de les mettre en pratique. Avoir à sa disposition un outil conceptuel comme ce diagramme permet au linguiste de donner une *forme* à l'analyse selon une logique qui correspond à la théorie.

Toutefois, la Figure 1 ne se borne pas à appliquer les principes que nous venons de décrire : elle fait usage d'autres conventions, comme la superposition des étiquettes de nature aux étiquettes identifiant les mots (superposition de *V* et de *mange* sans trait pour les relier), ou le triangle (sous *DP* et *AdvP*), qui indique que la construction identifiée au sommet pourrait être davantage décomposée, ou encore la flèche, qui lie ici le verbe *mange* à la position identifiée par le constituant *T*. Cette dernière convention, qui indique le « mouvement » de *mange* d'une position à une autre, est complexe, car elle repose sur une autre convention, nécessitant d'effectuer un certain nombre d'inférences, ce qui augmente la complexité de traitement du diagramme (Shimojima, 1996, 2015 ; Stapleton *et al.*, 2017). Certes, l'ordre des mots dans la phrase est normalement représenté par l'ordre d'apparition des étiquettes des constituants terminaux, de gauche à droite. Toutefois il est nécessaire de réaligner les mots sur l'axe vertical pour restituer la chaîne de l'ordre des mots. En outre, la flèche indique que l'ordre des constituants terminaux ne correspond pas à l'ordre des mots et, simultanément, elle montre comment il faudrait déplacer les constituants pour qu'il y ait correspondance. Grâce à cette convention, on peut distinguer un ordre des mots « de surface » d'un ordre « profond ».

Qu'on adhère ou non à la modélisation X-barre et à la notion de mouvement⁴, cet exemple de diagramme classique permet, pour la plupart des linguistes, de mettre en évidence le fait que ces objets ont un fonctionnement relativement complexe, même quand ils paraissent familiers. Nous proposerons ici des concepts particuliers pour l'analyse des diagrammes, qui nous permettront de nous focaliser sur cet aspect. Un diagramme représente une connaissance qui, sans cela, demeurerait abstraite. Nous dirons que le diagramme est une *inscription* de la connaissance (*cf.* la théorie du support ; Bachimont, 2007, 2010). Il y a deux moyens complémentaires d'inscrire la connaissance dans un signe visuel⁵ :

- On peut inscrire un concept par *réification*, sous la forme d'une entité graphique (Groupe μ , 1992) sémiotisée, c'est-à-dire d'un signe discret exprimant un concept, une unité, une opération. Par exemple, le symbole *CP* de la Figure 1 réifie le groupe de mots qu'on appelle « phrase ».
- On peut inscrire un concept par *configuration*, c'est-à-dire par l'agencement relatif d'entités, envisagée de manière globale ou locale. Ainsi, dans la Figure 1, de manière globale, l'ordre linéaire des constituants est inscrit sur l'axe horizontal (*DP* précède *T'*). De manière locale, la flèche de mouvement est tracée *entre* deux symboles représentant des constituants.

Idéalement, une inscription proprement diagrammatique est **formelle**, au sens où les entités réifiées font partie d'un inventaire clos et où les règles configurationnelles qui les régissent peuvent également être inventoriées (on peut parler de « formalisme diagrammatique »).

2. Variété des représentations graphiques des données et des analyses linguistiques

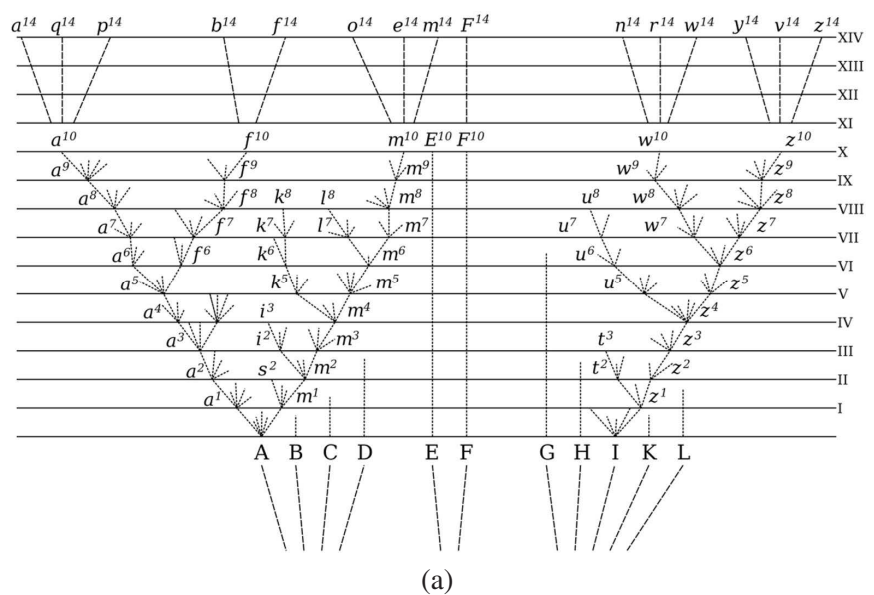
Les diagrammes syntaxiques correspondent à une structure *grosso modo* arborescente. Ils peuvent convenir à la description de groupements syntagmatiques de mots dans une phrase, mais ils sont insuffisants pour représenter des relations sémantiques de coréférence par exemple : les linguistes emploient donc une grande variété de diagrammes en fonction de leurs besoins. Dans cette section, nous donnons un aperçu des diagrammes utilisés en linguistique qui sont étudiés dans ce volume et nous les comparons brièvement à des diagrammes similaires employés dans d'autres domaines du savoir – nous nous focalisons ici sur des types de diagrammes récurrents plutôt que sur des représentations sporadiques.

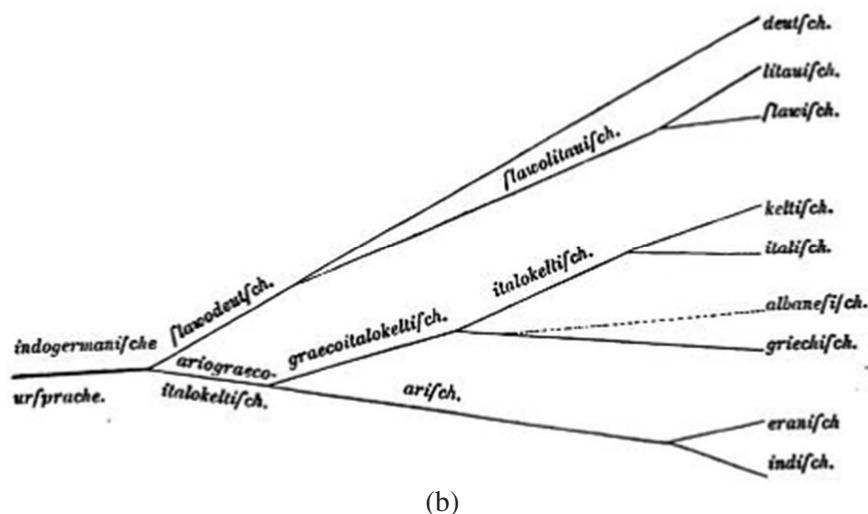
Les visualisations **arborescentes** sont des moyens d'inscription de la connaissance très répandus (Lima, 2014) car elles permettent de représenter de nombreux types de relations : généalogies, taxinomies, organigrammes, groupements imbriqués, etc. En pratique, les diagrammes arborescents sont efficaces dans la plupart des cas où il s'agit d'inscrire des hiérarchies (Bosveld-de Smet et Verheggen, 2016). En linguistique, les arborescences permettent de représenter de multiples sortes de hiérarchies (voir Bubenhof, 2020 : Sec. 5.5 ; Stewart, 1976 : 6-38). On peut partir de la Figure 1 pour se faire une idée de la variété des relations qui peuvent être représentées de manière arborescente. Des arbres comme celui de la Figure 1 sont couramment employés par certains syntacticiens, alors que d'autres leur préfèrent des arborescences dépendanciennes, qui représentent des relations hiérarchisées directes entre les mots (Mel'čuk, 1988).

Les deux approches se basent sur des diagrammes qui correspondent à des structures graphiquement similaires, mais les structures ne sont pas naturellement sémiotisées : selon leur utilisation, des entités graphiques et configurations apparemment similaires peuvent revêtir des sens différents (Mazziotta, 2020b : 69-70).

Les représentations graphiques de la généalogie des langues sont arborescentes depuis Schleicher. Les indoeuropéanistes, et de manière générale tous les linguistes qui travaillent sur les rapports phylogénétiques entre les langues, ont pris l’habitude de représenter cette généalogie sous la forme d’une arborescence de clades similaires à celle utilisée en biologie pour classer les taxons (Figure 2 ; voir François, ce volume) ou les manuscrits d’une tradition manuscrite en philologie. Il y aurait beaucoup à dire sur la manière dont les groupes de langues sont réifiés dans la Figure 2b. Nous nous contenterons ici de remarquer que le diagramme de Schleicher (1871) a une forme ressemblant à celle du diagramme de Darwin (1859), que reproduit la Figure 2a.

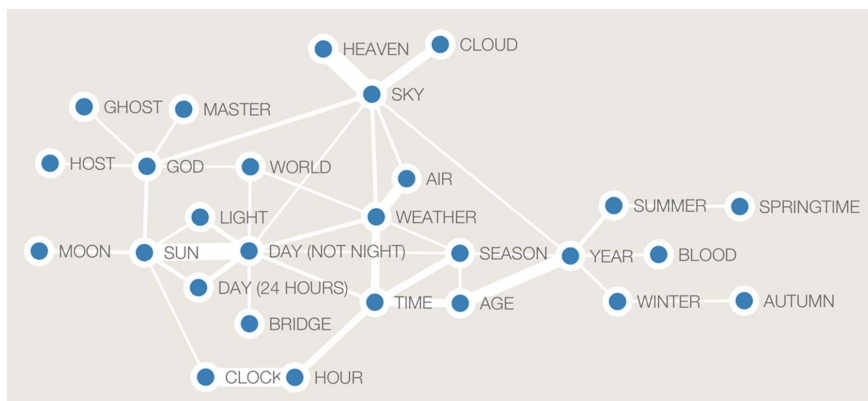
**Figure 2 : (a) Reproduction du premier diagramme publié de Darwin (1859, cf. Lima, 2014 : 39) ;
(b) Généalogie des langues indo-européennes selon Schleicher (1871 [1861] : 9).**





Le caractère formel des diagrammes rend leur emploi fréquent en mathématiques. On peut constater de nombreux échanges entre les conventions graphiques utilisées dans les différents domaines des mathématiques et en linguistique. Le concept d'*arbre* correspond à une structure qui se laisse définir mathématiquement et qui est plus contrainte que les structures nommées **graphes**⁶, dont les représentations (« tracés ») ont la forme de réseaux de nœuds connectés par des arêtes qui peuvent être orientées ou non. Les graphes sont très utilisés en sémantique, que ce soit pour encoder la structure sémantique de groupes de mots et d'énoncés (voir notamment Mel'čuk et Milićević, 2014a), ou pour décrire les relations de proximité entre des mots (*cf.* les graphes de synonymes du CRISCO⁷) ou des concepts. C'est ce dernier cas qui est illustré par la carte sémantique Figure 3 (v. Badir et Polis, ce volume).

Figure 3 : Carte sémantique de concepts temporels (Georgakopoulos et Polis, 2021, Figure 8); voir Badir et Polis, ce volume



La mise en image de la connaissance et du raisonnement se manifeste également au travers de notations mathématisantes. Ainsi, celles qui représentent les dérivations de Chomsky (Figure 4) ressemblent à s’y méprendre à un empilement d’opérations algébriques (Kahane et Mazziotta, ce volume).

Figure 4 : Dérivation proposée par Chomsky
(1969 [1957] : 30)

<i>Sentence</i>	
<i>NP + VP</i>	(i)
<i>T + N + VP</i>	(ii)
<i>T + N + Verb + NP</i>	(iii)
<i>the + N + Verb + NP</i>	(iv)
<i>the + man + Verb + NP</i>	(v)
<i>the + man + hit + NP</i>	(vi)
<i>the + man + hit + T + N</i>	(ii)
<i>the + man + hit + the + N</i>	(iv)
<i>the + man + hit + the + ball</i>	(v)

La notion d’opération, rendue explicite dans la Figure 4, nous montre ici que les diagrammes sont utilisés pour conduire un raisonnement visant à prouver des hypothèses.

Nous terminerons notre aperçu des diagrammes linguistiques similaires à ceux employés en mathématiques en comparant les diagrammes de la Figure 5. Le diagramme Figure 5a, qui représente une bijection, (ou « correspondance biunivoque », v. Halmos, 1997 [1965] : 41-43) entre les éléments des **ensembles** X et Y, présente des similarités partielles avec le diagramme Figure 5b, qui inscrit la construction progressive du sens textuel comme une progression de correspondances entre différents espaces de croyance (Col, ce volume).

À côté des diagrammes qui rappellent les concepts et les structures mathématiques, certains domaines de la linguistique privilégient des outils graphiques clairement issus d’autres disciplines. Ainsi, la géolinguistique est un domaine de la linguistique où les représentations graphiques à caractère iconique sont absolument nécessaires. La géolinguistique est une méthode de dialectologie qui consiste à reporter sur des **cartes** la répartition géographique de phénomènes linguistiques (Bubenhofer, 2020 : Sec. 5.2). Elle rend ainsi compte de la variation diatopique de la langue (Figure 6a) de la même manière que la géographie cartographie la répartition de données physiques, sociales ou historiques (Figure 6b).

Figure 5 : (a) Diagramme de bijection⁸ ; (b) Progression narrative selon Fauconnier (1997) ; v. Col (ce volume)

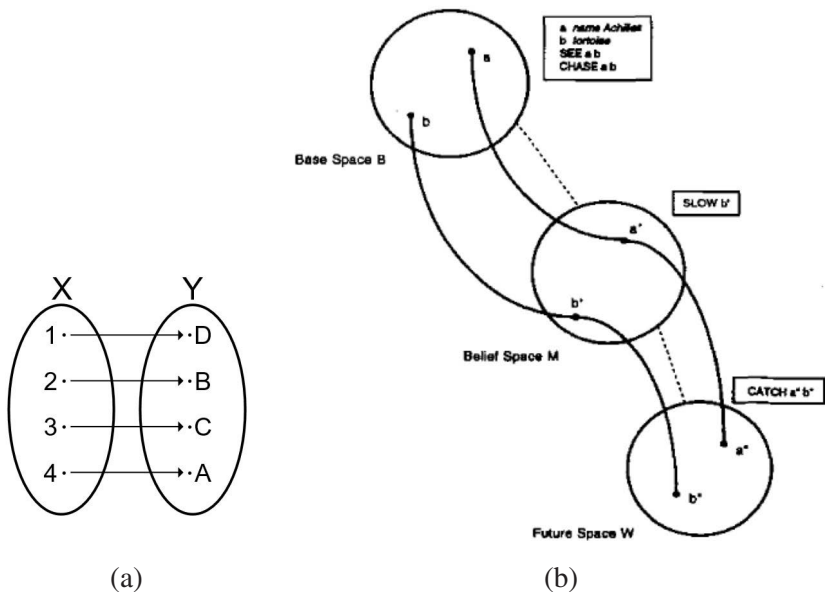
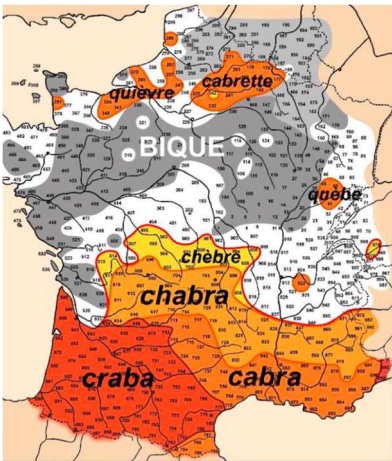
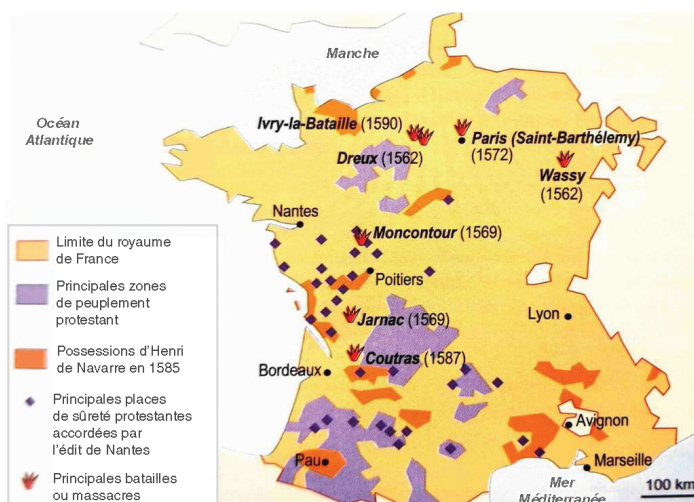


Figure 6 : (a) « ALF 272 Chèvre : aires lexicales et aires phonétiques » (Le Dû et al., 2005 : 46) ; v. Baiwir (ce volume) ; (b) Carte synthétisant la répartition des terres protestantes en France au 16^e s. (Bourel et al., 2019 : 165)



(a)



(b)

En important ces outils, la dialectologie pose des questions spécifiques, notamment liées à la relation entre les tableaux de classement de ces données, qui accompagnent les cartes dans le cas d'atlas, comme l'*Atlas linguistique de la Wallonie* (Baiwir, ce volume).

La variété des diagrammes en linguistique est évidemment beaucoup plus riche que ne le laisse penser la petite revue que nous venons d'en faire (qui n'a d'autre prétention que d'introduire les articles de ce volume). Les diagrammes sont très présents chez certains linguistes (comme Guillaume, Culioli ou Pottier en France, et Hockett ou Langacker aux États-Unis), à qui nous ne pourrions malheureusement rendre hommage que de manière indirecte.

3. Raisonner à l'aide de diagrammes

Jusqu'ici, nous avons employé le terme *diagramme* de manière lâche. Cette section est consacrée à la définition peircienne du terme (3.1), à son intégration à la conception peircienne du raisonnement (3.2), ainsi qu'au processus de *raisonnement diagrammatique* (3.3).

3.1. Les diagrammes sont des icônes de relations

Plusieurs exégètes de la pensée logique et épistémologique de Charles S. Peirce ont étudié la place éminente des diagrammes dans cette théorie (v. en particulier Stjernfelt, 2007). Peirce définit les diagrammes comme des « icônes de relations » :

« A Diagram is mainly an Icon, and an Icon of intelligible relations. [...] Now since a diagram, though it will ordinarily have Symbolide [= relatives aux symboles (?)] Features, as well as features approaching the nature of Indices, is nevertheless in the main an Icon of the forms of relations in the constitution of its Object, the appropriateness of it for the representation of necessary inference is easily seen » (*Collected papers* [1906] : § 4.531).

Les relations inscrites dans les diagrammes forment une structure similaire à celle que forment les relations représentées. Cette iconicité a comme conséquence qu'un raisonnement qui s'appuie sur un diagramme permet de découvrir des propriétés jusqu'alors inconnues des phénomènes et observations modélisés par le diagramme, ce qui en fait un **outil de raisonnement** puissant. Selon Peirce, le diagramme permet le raisonnement « théorématique », qui ne se contente pas de l'observation de prémisses, mais qui ajoute du contenu pour mener à une nouvelle interprétation créative (Chauviré, 2008 : 36-42).

La conception et la manipulation des diagrammes reposent sur la combinaison de plusieurs types de raisonnement : par induction et par déduction, mais aussi par *abduction* (notion mise en évidence par Peirce), seul type de raisonnement qui nous fasse effectivement découvrir de nouvelles connaissances et que nous allons brièvement aborder au travers des diagrammes. Stjernfelt rappelle ceci :

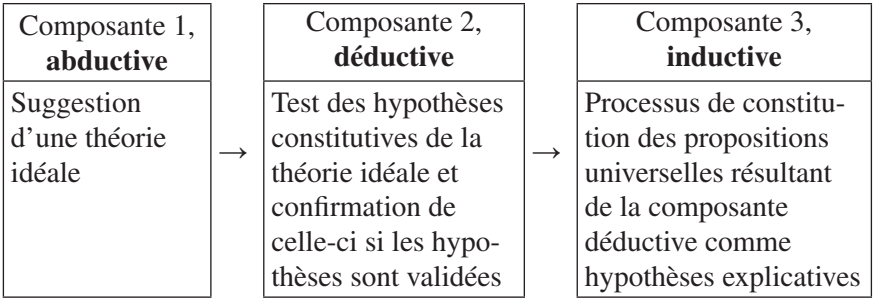
« Premièrement, il [Peirce] a proposé que toutes les mathématiques soient développées à l'aide d'un raisonnement diagrammatique. Cette proposition vient de l'idée que les mathématiques portent sur des formes de relations idéalisées, hypothétiques – et que cela procède par expérimentation avec des diagrammes représentant de telles hypothèses idéalisées [...] Deuxièmement, il a proposé que tout raisonnement déductif soit entrepris à l'aide d'un tel raisonnement diagrammatique » (2014 : 3, nous traduisons).

Bien que Peirce ne réduise pas la notion de diagramme à celle de « formalisme graphique », force est de constater que les pratiques diagrammatiques sont, à tout le moins en linguistique et dans l'exemple que nous présenterons ci-après (3.2), quasi exclusivement graphiques. Nous raisonnons sur des signes visuels ; c'est cette propriété qui nous permet d'ailleurs de les manipuler. Si nous revenons aux notions de *réification* et de *configuration* (1), c'est la seconde qui paraît la plus spécifique aux diagrammes visuels : c'est la sémiotisation des deux axes du plan graphique qui permet la définition des configurations sur lesquelles le raisonnement peut être élaboré (Badir, 2007 ; Klinkenberg, 2010).

3.2. Composantes du raisonnement peircien

La présentation de Michael Hoffmann (2003 : 128-140), dont nous synthétisons ici l'exposé, permet de saisir les relations entre les trois composantes du raisonnement. Peirce introduit la notion d'**abduction** comme première composante du raisonnement. En observant de manière sélective la masse des faits, nous acquérons de nouvelles idées ouvrant la porte à des hypothèses explicatives. La composante abductive du raisonnement selon Peirce a pour fonction de suggérer une « théorie idéale », elle a donc un caractère « sauvage » que la deuxième composante, la *déduction*, est chargée de réguler⁹. Peirce définit la fonction de la déduction comme l'établissement d'une variété de conséquences. La vérification de l'adéquation de l'hypothèse à une sélection de faits valide son applicabilité à l'ensemble des faits (*Collected papers* : § 8.209). Hoffmann (2003 : 128) en conclut que la troisième composante du raisonnement, l'*induction*¹⁰, « se limite à confirmer ou infirmer graduellement des propositions universelles données sur la base de données particulières. Le raisonnement, qui peut être qualifié de « triphasé », se décompose donc ainsi :

Figure 7 : Composition du mode de raisonnement « triphasé » de Peirce



Autrement dit, l'abduction est concernée par le problème de la découverte, tandis que l'induction vise le problème de la justification d'assertions générales, lesquelles proviennent progressivement de la démarche déductive.

3.3. Le raisonnement diagrammatique : exemple peircien

Hoffmann montre que le raisonnement triphasé de Peirce se trouve déjà en germe dans le dialogue maïeutique présenté par Platon dans le *Ménon* entre Socrate et un jeune garçon. Socrate demande à ce dernier comment obtenir géométriquement, à partir d'un carré ABCD de deux pieds de côté, un second carré de surface deux fois plus grande.

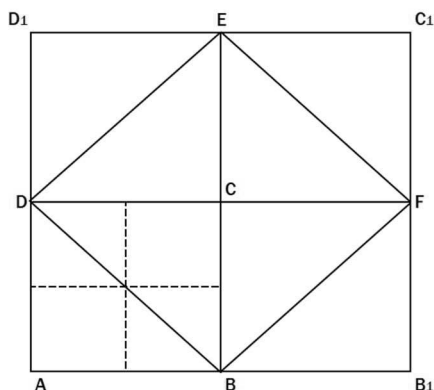
En proposant de doubler la longueur de chaque côté, le garçon montre qu'il ne saisit pas la différence entre le doublement d'un périmètre et celui de la surface enclose. Face à cette proposition, qui représente la première phase du raisonnement abductif, Socrate dessine un diagramme qui représente conjointement le carré initial ABCD et le carré $AB_1C_1D_1$ imaginé par le garçon (cf. Figure 8). Cette représentation correspond à **une icône des idées** du garçon. Au premier regard porté sur ce diagramme, le garçon comprend que le doublement des côtés produit un carré qui est quatre fois plus grand (Hoffmann, 2003 : 139). Il s'agit là de la phase intermédiaire d'invalidation de sa « théorie idéale ». Le diagramme rend la compréhension immédiate car les propriétés structurales de l'inscription manifestent automatiquement la multiplication de la surface dès que les côtés du carré sont allongés. La conclusion est en quelque sorte cognitivement « gratuite » (Shimojima, 1996, 2015).

L'observation de la différence de taille entre les carrés permet de déduire que la longueur du côté doit être comprise entre celle de la figure initiale (deux pieds) et celle de sa première réponse (quatre pieds). Le garçon suggère alors qu'elle mesure trois pieds, à égale distance entre deux et quatre pieds. Dans cette seconde phase de déduction, il a donc ajusté son système de représentation (nouvelle abduction). Toutefois, un tel carré ne se laisse pas aisément dessiner et rien ne lui garantit que trois pieds est exactement la bonne longueur. Pour trouver celle-ci, il va falloir d'abord **introduire un nouveau concept** : celui de *diagonale*, et imaginer un carré orienté différemment du carré initial ABCD et du carré $AB_1C_1D_1$ qu'il a lui-même imaginé en premier (cf. Figure 8). Le carré sur pointe, inscrit dans le carré de quatre pas de côté, a une surface inférieure de moitié à celle de ce dernier, dont chaque quart a été partagé en deux. Et Hoffmann conclut :

« Tout le processus de diagrammatisation permet au garçon d'apprendre non seulement certaines règles de la géométrie, mais aussi le nouveau concept de la « diagonale ». La solution est obtenue par le biais de ce que Peirce appelait une « abstraction hypostatique » car une partie du diagramme devient une entité propre, un nouvel objet, qui peut être désormais sujet à de nouvelles opérations » (2003 : 140).

Une partie de la puissance des diagrammes par rapport aux inscriptions discursives de la même connaissance est qu'ils rendent explicites des relations sans que le lecteur ne soit obligé d'effectuer des inférences (Larkin et Simon, 1987 ; Norman, 2003 ; Shimojima, 1996, 2015 ; Stapleton *et al.*, 2017). La présentation générale du potentiel des diagrammes s'applique de toute évidence aux diagrammes linguistiques, dont l'utilité est de rendre perceptible et manipulable la connaissance linguistique que l'on désire

Figure 8 : Diagramme résumant l'aboutissement du dialogue maïeutique de Socrate sur la diagonale dans le Ménon (Hoffmann, 2003 : 140)



manipuler en vue de preuve ou à la recherche de nouvelles propriétés de la matière inscrite. Cette utilisation dynamique des diagrammes est clairement mise en évidence dans les contributions du présent volume.

4. Analyser les diagrammes

Les diagrammes dont nous nous occupons sont des signes visuels. Pour fonctionner, ils doivent donc être « lus », c'est-à-dire perçus et compris. La perspective externe qui est la nôtre consiste à *analyser* les diagrammes et non simplement à les utiliser. Nous visons une compréhension fine de systèmes diagrammatiques, ce qui passe nécessairement par l'établissement du caractère distinctif des réifications et des configurations à l'œuvre qui les composent, leur *pertinence* (4.1). Nous visons ensuite à introduire certaines des difficultés de lecture et d'interprétation des entités qui composent les diagrammes (4.2).

4.1. Pertinence des composantes diagrammatiques

Pour comprendre un diagramme, il faut être capable de cerner la valeur des réifications et des configurations qui servent à le construire. Certains diagrammes courants comme la Figure 1 sont employés par une collectivité qui en partage les conventions de lecture¹¹. D'autres sont plus rares, anciens, voire abscons. En cela, ils posent une question philologique cruciale : comment transcrire ou utiliser des outils dont nous n'avons parfois qu'une compréhension partielle ?

Par exemple, lorsqu'ils inscrivent ce que Tesnière appelle une « translation », c'est-à-dire un changement de classe d'un mot opéré par un morphème grammatical dit « translatif », les diagrammes de l'auteur

indiquent l'emplacement du translatif par l'orientation du « crochet inférieur » de la hampe de l'entité en forme de « T » stylisé qui sert à inscrire une translation (Figure 9).

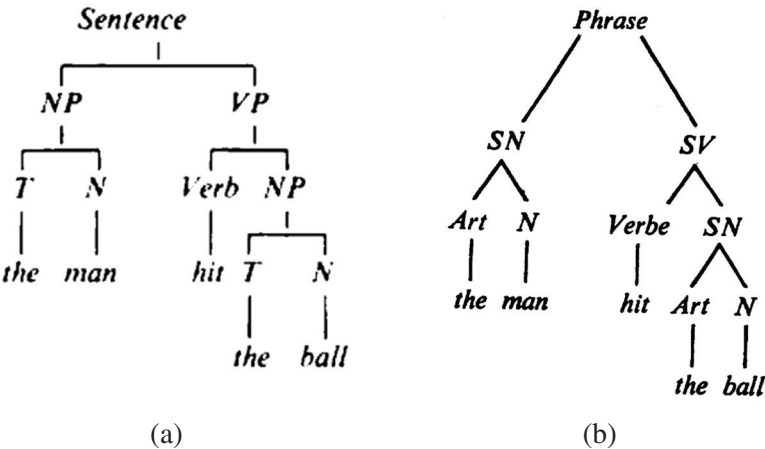
Figure 9 : Explications de Lucien Tesnière concernant une des conventions de diagrammatisation dont il fait usage (1959 : Chapitre 155, 11-12)

11. — Mais de façon à éviter toute confusion entre le transférende et le translatif, nous aurons soin de toujours tourner le **crochet inférieur de la hampe** du T vers le translatif.
12. — Le sigle de la translation sera donc T ou T selon que le translatif ou marquant de la translation sera préposé ou postposé au transférende fr. de de T Pierre ang. Peter T's

Comprendre correctement l'inscription de la translation implique de saisir le caractère distinctif de l'orientation du crochet inférieur, sa **pertinence sémiotique**. Un système où tous ces crochets seraient orientés vers la gauche est possible¹², mais différent de celui de Tesnière : les distinctions qui le fondent ne sont pas identiques.

Autre exemple : dans leur analyse des premiers diagrammes d'analyse en constituants immédiats, Mazziotta et Kahane (2017) présentent un contraste supposé entre les arbres de la Figure 10.

Figure 10 : (a) Nouvelle composition du premier diagramme arborescent publié de Chomsky (2002 [1957]) ; (b) Premier diagramme arborescent publié de Chomsky, reproduit dans la traduction française (1969 [1957])



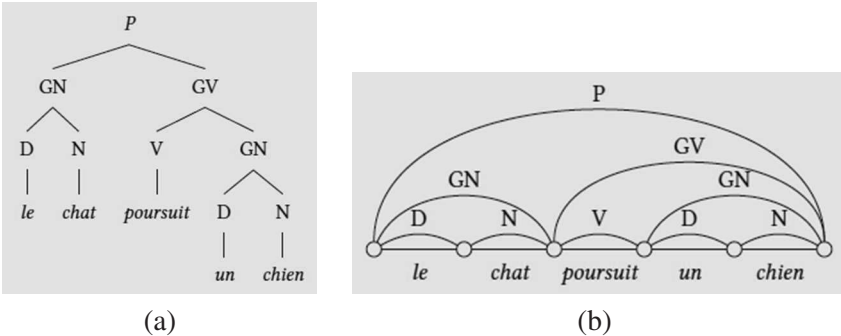
Selon eux, les traits de Figure 10a doivent être analysés comme suit : les traits verticaux inscrivent la relation de réécriture (ils équivalent à la flèche dans $NP \rightarrow T + N$), alors que les crochets qui relient les constituants immédiats d'une même construction inscrivent précisément que ces constituants sont groupés (ils équivalent au signe « + » dans la règle précédente). Ils ajoutent que la Figure 10b ne fonctionne pas de la même manière : la relation de réécriture et la relation de groupement ne sont pas discrétisées séparément. Toutefois, les auteurs négligent que la Figure 10a n'est pas l'arbre original publié, qui avait exactement la même forme que la Figure 10b (nonobstant la langue des étiquettes). En réalité, la Figure 10a est un nouveau dessin de la même structure. Il introduit des entités nouvelles. À moins que cet arbre ne reconstitue celui présent dans le manuscrit de l'auteur, Kahane et Mazziotta ont sur-interprété en mentionnant des distinctions qu'ils jugeaient pertinentes selon leur interprétation¹³. La présence de caractéristiques accidentelles (« accidental features » Shimojima, 1996), qui pousse certains lecteurs à trouver du sens où il n'y en a pas, est un défaut inhérent aux diagrammes.

La question de la pertinence des entités observées pose constamment des problèmes pratiques dans le processus d'édition et de publication des textes contenant des diagrammes. Nous sommes nombreux à avoir fait l'expérience d'une malheureuse composition typographique altérant notre propos. Jusqu'à quel point faut-il reproduire les variations les plus fines de tracé des entités ? Elle se pose de façon encore plus complexe quand il s'agit de décrire les conventions d'anciens systèmes diagrammatiques – voir par exemple Col (ce volume) sur l'analyse des diagrammes de Vandeloise, ainsi que les travaux sur les grammaires américaines du 19^e siècle (Brittain, 1973 ; Mazziotta, 2016, 2020b).

4.2. Sémiotisation des entités

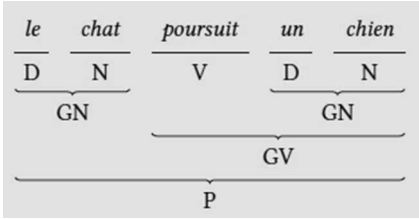
Il est important de noter que des diagrammes apparemment différents peuvent être des inscriptions d'une même structure et que, inversement, des entités apparemment similaires peuvent inscrire des concepts différents. Ainsi les deux diagrammes de la Figure 11 inscrivent une même hiérarchie arborescente de constituants. Dans la Figure 11a, selon l'usage, les rapports partie-tout qui constituent cette hiérarchie sont réifiés, tandis que l'ordre linéaire est représenté de manière configurationnelle par le placement horizontal des nœuds. La Figure 11b est un arbre « en cellules ». Dans ce dernier, ce sont les rapports de précedence correspondant à l'ordre linéaire qui sont réifiés : les nœuds du graphe correspondent aux frontières des mots et les constituants sont représentés par des arêtes, tandis que les rapports méreologiques sont inscrits de manière configurationnelle par l'alignement vertical des arêtes.

**Figure 11 : (a) Arbre de constituants classique ;
(b) Arbre de constituants « en cellules »
(arbres reproduits de Kahane et Gerdes, 2022 : 515)**



L’inscription « en cellules » peut elle-même connaître des variantes où les frontières ne sont plus explicitement réifiées, comme on l’observe dans la Figure 12. On revient là sur une représentation plus traditionnelle, généralement utilisée à des fins didactiques (voir Kahane et Mazziotta, ce volume).

**Figure 12 : Arbre de constituants « en cellules » simplifié
(Kahane et Gerdes, 2022 : 516)**



Inversement, il saute aux yeux que les mêmes entités n’inscrivent pas nécessairement les mêmes contenus. Les arêtes de la Figure 11a ont une forme d’expression proche des arêtes des cartes sémantiques et des arbres phylogénétiques (→2), mais il est évident que, n’appartenant pas au même système, elles ont une valeur différente. Le contraste entre les Figure 11a et 11b nous incite à être attentifs : malgré la proximité conceptuelle des deux diagrammes et la ressemblance des entités qui les composent, les traits y réifient des contenus différents.

5. Articles rassemblés

Nous avons pu évoquer ci-dessus les articles du volume, notamment en mentionnant les types de diagrammes sur lesquels ils se penchent. Nous

terminons cette introduction par un résumé succinct de chacune des contributions et de la conclusion.

Dans le premier article de ce volume, « Les styles collectifs de pensée visuelle en sciences du langage : canons et variations », Jacques François évoque le développement de schémas et de diagrammes linguistiques, d'abord intuitif et à finalité pédagogique notamment chez A. Schleicher, H. Schuchardt dans la seconde moitié du 19^e siècle et chez Saussure et Hockett au 20^e siècle. Ils sont ensuite destinés à visualiser le résultat du traitement de données impliquant le croisement de plusieurs facteurs et finalement celui de l'application d'un algorithme en synchronie. On observe notamment leur utilisation dans les études lexicométriques ou géolinguistiques, ou en diachronie, dans celles du changement linguistique. Pour mener son étude, l'auteur développe une méthode de calcul des écarts entre modes de visualisation réunis par paires.

Les deux articles suivants, « Les cartes sémantiques en typologie des langues. La médiation iconique entre qualification et quantification dans des représentations visuelles du discours linguistique », par Sémir Badir et Stéphane Polis, et « La place et les limites des diagrammes en sémantique cognitive », par Gilles Col, questionnent le statut et l'utilisation de diagrammes en sémantique. Dans le premier, les auteurs mènent une analyse sémiotique structurée de pratiques concrètes d'utilisation de cartes sémantiques. Dans le second, l'auteur montre quelles alternatives de représentations ont jalonné l'évolution de la sémantique cognitive. L'article se termine par des propositions concrètes de nouveaux diagrammes.

Dans le quatrième article, « Ce que contiennent (et ce que ne contiennent pas) des cartes géolinguistiques », Esther Baiwir part de son expérience de dialectologue pour questionner le rapport entre les cartes représentant la structure diatopique des matériaux, les données elles-mêmes et le texte d'analyse que ces cartes peuvent accompagner.

Le cinquième et le sixième article portent sur la modélisation en syntaxe. Dans « Grammaires graphiques », Sylvain Kahane et Nicolas Mazziotta montrent que les règles de la grammaire d'une langue peuvent être formalisées par des diagrammes, sur lesquels il est possible d'effectuer de véritables opérations. Dans « Délimitation des diagrammes syntaxiques dans la grammaire scolaire », Nicolas Gregov se questionne sur la définition et la typologie de diagrammes à visée pédagogique.

Dans un bref article conclusif, nous proposons une réflexion programmatique pour l'élaboration d'un inventaire raisonné et outillé des représentations visuelles en sciences du langage.

NOTES

1. À l'aide de logiciels comme *Gephi* <<https://gephi.org/>> ou *Cytoscape* : <<https://cytoscape.org/>>, ou de bibliothèques spécialisées comme le classique *Graphviz* <<https://graphviz.org/>>, qui peut être utilisé dans la plupart des langages de programmation, *ggraph* pour le logiciel *R* <<https://cran.r-project.org/web/packages/ggraph/>>, *NetworkX* pour Python <<https://networkx.org/>>.
2. Les notions fondamentales du cadre théorique X-barre sont ici supposées connues (voir Carnie, 2002/2007 pour une introduction).
3. Le terme *construction* est ici à comprendre dans un sens classique de l'analyse en constituants immédiats telle que présentée par Gleason (1969 [1955] : paragraphe 10.9).
4. Pour les approches en constituants sans mouvement, voir notamment les propositions de la *Radical construction grammar* (Croft, 2007 [2001], ou la *Role and Reference Grammar* (Van Valin, s.d.).
5. Nous reprenons ici les définitions de Kahane et Mazziotta (ce volume), qui mobilisent ces notions, de même Gregov (ce volume).
6. Pour une définition accessible au linguiste, voir l'exposé de Mel'čuk et Milićević (2014b : 296). En toute rigueur, la Figure 1 n'est pas un arbre, mais un graphe (Mel'čuk et Milićević, 2014b : 294 et sv.).
7. <<https://criso2.unicaen.fr/des/>>
8. <https://en.wikipedia.org/wiki/Injective_function#/media/File:Bijection.svg>
9. Cf. *deduction* in *Enc. Brit.* : « a chain of statements, each of which is either a premise or a consequence of a statement occurring earlier in the proof [...] Some experts regard all valid inference as deductive in form and, for this and other reasons, reject the supposed contrast between deduction and induction ».
10. Cf. *inductive logic* in *Enc. Brit.* : « Inductive reasoning means reasoning from known particular instances to other instances and to generalizations ».
11. Parfois au terme d'une élaboration progressive. V. la question de l'élaboration des arbres phylogénétiques traitée par François (ce volume).
12. Comme dans l'exposé de Soutet (1995/2007 : 280-286).
13. Contacté à ce propos, Noam Chomsky a confirmé qu'il ne « voyait aucune différence substantielle » (courriel du 9/11/2022).

RÉFÉRENCES CHOISIES

- 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*, Paris, Les Belles Lettres.
- BADIR S., 2007, « Rhétorique des graphiques dans les médias contemporains », *Actes du VIII^e congrès de l'ASV-IASV, Cultures du visible*, p. 279-299.
- BERTIN J., 2005 [1967], *Sémiologie graphique : Les diagrammes, les réseaux, les cartes*, Paris, EHESS.

- BOSVELD-DE SMET L. et VERHEGGEN R.-J., 2016, « Comparison of diagrams in producing and understanding hierarchies in three different application domains », in JAMNIK M., USAKA Y. et ELZER SCHWARTZ S., *Diagrams 2016. Diagrammatic representation and inference*, Vol. 9781, Berlin, Springer, https://doi.org/10.1007/978-3-319-42333-3_23, p. 272-279
- BOUREL G., CHEVALLIER M., GUILLAUSSEAU A. et JOUBERT G., 2019, *L'histoire de France : Des origines à nos jours*, Paris, Hatier.
- BRITTAIN R. C., 1973, *A critical history of systems of sentence diagramming in English*, University of Texas at Austin, <https://books.google.se/books?id=6Ve8pwAACAAJ>.
- BUBENHOFER N., 2020, *Visuelle Linguistik: Zur Genese, Funktion und Kategorisierung von Diagrammen in der Sprachwissenschaft*, Berlin, De Gruyter.
- CARNIE A., 2007 [2002¹], *Syntax: A generative introduction*, Oxford, Blackwell.
- CHAUVIRÉ C., 2008, *L'œil mathématique : Essai sur la philosophie mathématique de Peirce*, Paris, Kimé.
- CHOMSKY N., 1969 [1957¹], *Structures syntaxiques*, trad. de BRAUDEAU M., Paris, Le Seuil [édition anglaise 1957].
- CHOMSKY N., 2002 [1957¹], *Syntactic structures*, Berlin, Mouton De Gruyter.
- CROFT W., 2007 [2001], *Radical construction grammar: Syntactic theory in typological perspective*, reprinted, Oxford University Press.
- DONDERO M. G., 2019, « De l'énonciation et du métavisuel dans le cadre des Big Data », in ESTAY STANGE V., HACHETTE P. et HORREIN R., *Sens à l'horizon! Hommage à Denis Bertrand*, Limoges, France, Lambert Lucas, p. 285-298.
- DONDERO M. G., 2020, *Les langages de l'image. De la peinture aux Big Visual Data*, Paris, Hermann.
- FAUCONNIER G., 1997, *Mappings in thought and language*, Cambridge, Cambridge University Press.
- GEORGAKOPOULOS T. et POLIS S., 2018, « The semantic map model: State of the art and future avenues for linguistic research », *Language and Linguistics Compass*, 12, 2, e12270. <https://doi.org/10.1111/lnc3.12270>.
- GEORGAKOPOULOS T. et POLIS S., 2021, « Lexical diachronic semantic maps: Mapping the evolution of time-related lexemes », *Journal of Historical Linguistics*, 11, 3, <https://doi.org/10.1075/jhl.19018.geo>, p. 367-420.
- GLEASON H. A., 1969, *Introduction à la linguistique*, trad. de DUBOIS-CHARLIER F., Paris, Larousse [édition anglaise 1955].
- GROUPE μ , 1992, *Traité du signe visuel : pour une rhétorique de l'image*, Paris, Le Seuil.
- HALMOS P. R., 1997 [1965], *Introduction à la théorie des ensembles*, trad. de GARDELLE J., reprod. en fac-sim., Paris, Jacques Gabay [édition anglaise 1965].
- HOFFMANN M. H. G., 2003, « Peirce's 'Diagrammatic Reasoning' as a solution of the learning paradox », in DEBROCK G., *Process Pragmatism: Essays on a quiet philosophical revolution*, Value Inquiry Book Series 137, Brill, p. 121-143.
- KAHANE S. et GERDES K., 2022, *Syntaxe théorique et formelle. Vol. 1. Modélisation, unités, structures*, Language Science Press, <https://langsci-press.org/catalog/book/241>.

- KAHANE S. et MAZZIOTTA N., 2015, « Syntactic Polygraphs. A Formalism Extending Both Constituency and Dependency », *Proceedings of the 14th Meeting on the Mathematics of Language (MoL)*, <https://doi.org/10.3115/v1/W15-2313>, p. 152-164.
- KLINKENBERG J.-M., 2010, « À quoi servent les schémas ? Tabularité et dynamisme linéaire », *Protée*, 37, 3, <https://doi.org/10.7202/038806ar>, p. 65-73.
- LARKIN J. H. et SIMON H. A., 1987, « Why a diagram is, sometimes, worth ten thousand words », *Cognitive Science*, 11, 1, [https://doi.org/10.1016/S0364-0213\(87\)80026-5](https://doi.org/10.1016/S0364-0213(87)80026-5), p. 65-100.
- LE DÛ J., LE BERRE Y., BRUN-TRIGAUD G., LE ROUX P. et FALC'HUN F., 2005, *Lectures de l'Atlas linguistique de la France de Gilliéron et Edmont : Du temps dans l'espace ; essai d'interprétation des cartes de l'Atlas linguistique de la France de Jules Gilliéron et Edmond Edmont ; à la mémoire du professeur François Falc'hun*, Aubervilliers, Éditions du Comité des Travaux Historiques et Scientifiques.
- LIMA M., 2011, *Visual complexity: Mapping patterns of information*, Princeton, Princeton Architectural Press.
- LIMA M., 2014, *The book of trees: Visualizing branches of knowledge*, Princeton, Princeton Architectural Press.
- MAZZIOTTA N., 2016, « Drawing Syntax before Syntactic Trees: Stephen Watkins Clark's Sentence Diagrams, 1847 », *Historiographia Linguistica*, 43, 3, <https://doi.org/10.1075/hl.43.3.03maz>, p. 301-342.
- MAZZIOTTA N., 2019, « The evolution of spatial rationales in Tesnière's stemmas », *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest)*, <https://doi.org/10.18653/v1/W19-7709>, p. 69-80.
- MAZZIOTTA N., 2020a, « 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 etc., Benjamins, <https://doi.org/10.1075/slcs.212.05maz>, p. 134-162.
- MAZZIOTTA N., 2020b, « Grammar and graphical semiotics in early syntactic diagrams: Clark, 1847, and Reed-Kellogg, 1876 », in FORTIS J.-M. et AUSSANT É., *Historical journey in a linguistic archipelago: Descriptive concepts and case studies*, Language Science Press, p. 67-81, <https://langsci-press.org/catalog/book/286>.
- 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*, Berlin, De Gruyter. <https://doi.org/10.1515/9783110715118-003>, p. 45-66.
- MAZZIOTTA N. et KAHANE S., 2017, « To what extent is immediate constituency analysis dependency-based ? A survey of foundational texts », *Proceedings of the Fourth International Conference on Dependency Linguistics (Depling)*, <https://aclanthology.org/W17-6515>, p. 116-126.
- MEL'ČUK I. A., 1988, *Dependency syntax: theory and practice*, New York, State University of New York Press.
- MEL'ČUK I. A. et MILIĆEVIĆ J., 2014a, *Introduction à la linguistique*, vol 1, Paris, Hermann.

- MEL'ČUK I. A., et MILIĆEVIĆ J., 2014b, *Introduction à la linguistique*, vol 3, Paris, Hermann.
- PEIRCE C. S., 1931 [1906¹], *Collected papers of Charles Sanders Peirce*, Cambridge, Harvard University Press.
- NORMAN J., 2003, « Differentiating diagrams : a new approach », in ANDERSON M., CHENG P. et HAARSLEV V., *Theory and application of diagrams. Diagrams 2000*, vol. 1889, Berlin, Springer, https://doi.org/10.1007/3-540-44590-0_13, p. 105-116.
- SCHLEICHER A., 1871 [1861¹], *Compendium der vergleichenden Grammatik der indogermanischen Sprache*, Vienne, Cologne, Böhlau.
- SHIMOJIMA A., 1996, *On the efficacy of representation*, Thèse de doctorat, Indiana University, https://www1.doshisha.ac.jp/~ashimoji/Personal_Page/Thesis.pdf.
- SHIMOJIMA A., 2015, *Semantic properties of diagrams and their cognitive potentials*, Stanford, CSLI Publications.
- SMESSAERT H. et DEMEY L., 2018, « Towards a Typology of Diagrams in Linguistics », in CHAPMAN P., STAPLETON G., MOKTEFI A., PEREZ-KRIZ S. et BELLUCCI F., *Diagrams 2018. Diagrammatic Representation and Inference*, p. 236-244.
- SOUTET O., 2007 [1995¹], *Linguistique*, Paris, PUF.
- STAPLETON G., JAMNIK M., et SHIMOJIMA A., 2017, « What makes an effective representation of information: a formal account of observational advantages », *Journal of Logic, Language and Information*, 26, 2, <https://doi.org/10.1007/s10849-017-9250-6>, p. 143-177.
- STEWART A. H., 1976, *Graphic representation of models in linguistic theory*, Bloomington, Indiana University Press.
- STJERNFELT F., 2007, *Diagrammatology: An investigation on the borderlines of phenomenology, ontology, and semiotics*, Berlin, Springer.
- STJERNFELT F., 2014, *Natural propositions: The actuality of Peirce's doctrine of dicisigns*, Boston, Docent Press.
- TESNIÈRE L., 1959, *Éléments de syntaxe structurale*, Paris, Klincksieck.
- TUFTE E. R., 2012 [1997¹], *Visual explanations : Images and quantities, evidence and narrative*, Cheshire (Connecticut), Graphics Press.
- VAN VALIN, R. Jr., s.d. [*t. a quo* 2008], *An Overview of Role and Reference Grammar*, https://rrg.caset.buffalo.edu/rrg/RRG_overview.pdf.