

UNITÉS LINGUISTIQUES ET FORMES TEXTUELLES

FRANÇOIS RASTIER
(INaLCO-Ertim)

*Tout, dans l'énonciation, est séquence, réseau,
co-occurrences, équilibre ou fuites entre formes,
scansions d'emplois affrontées.*

Maurice Tournier

RESUME: La tradition scolaire a élaboré une conception de l'unité conforme avec l'hypothèse d'un parallélisme logico-grammatical. Elle s'avère discutable, tant au palier du mot qu'à celui de la phrase et plus encore au palier du texte.

En revanche, des grandeurs comme les thèmes ou des fonctions narratives peuvent être décrites comme des formes sémantiques et expressives, susceptibles d'être représentées par des groupements structurés de traits. Au fil du texte, ces formes sont le support de métamorphoses, transformations récurrentes qui déterminent son orientation.

Les formes textuelles se détachent sur des fonds expressifs et sémantiques (les isotopies). Au cours du texte, les fonds eux-mêmes peuvent évoluer par des transpositions. La dynamique du parcours textuel comprend en outre des dissolutions de formes par diffusion dans les fonds, et des sommations synthétiques qui construisent des formes à partir d'éléments des fonds.

Il faut donc reconsidérer la question des unités et celle des règles linguistiques à la lumière du texte, en reconnaissant la détermination du global par le local et en tenant compte également de l'interprétation des sorties logicielles en sémantique de corpus.

MOTS CLÉ: unité, règle, grammaire, texte, forme sémantique, fond sémantique.

1. Problématiques

De la grammaire, qui l'a précédée de deux millénaires, la linguistique a hérité une notion d'unité bien particulière, mais devenue si évidente qu'il nous faut la détailler.

(i) L'unité est positive, en ceci qu'elle a un contenu substantiel persistant, malgré d'éventuelles variations dues au contexte ou à la situation.

(ii) L'unité est discrète, bornée à droite et à gauche, dans le temps comme dans l'espace. Le mot est l'unité typique.

(iii) Elle est immédiatement identifiable en fonction de critères défini-
toires définis par une typologie, et apparaît comme donnée, dès lors que l'on dispose d'un répertoire de catégories. Les parties du discours en sont des exemples parfaits. Sur cette base, des règles prennent en charge ces unités discrètes et typées pour les organiser en phrases (assimilées à des propositions; cf. Pinker, *Words and Rules, The Ingredients of Language*, 1999).

Par ses présupposés ontologiques, cette image de l'unité est parfaitement compatible avec la conception frégréenne du calcul logique. Elle est à tout le moins discutable. Par exemple, avec sa théorie du signifiant discontinu, Martinet entendait rendre compte des chaînes d'accords et de concordances et posait le problème du point de vue sémasiologique. Avec la notion d'isosémie, Pottier visait les mêmes phénomènes de répétition, en adoptant une démarche inverse, onomasiologique. Sans entrer ici dans ce débat, retenons qu'il y reste en syntaxe des difficultés à inventorier les unités. En outre, des grandeurs non discrètes comme les intonations prosodiques organisent crucialement le contenu phrastique et suffisent par exemple à transformer une affirmation en interrogation. Elles ne correspondent pas à la notion positive sinon positiviste de l'unité.

Outrepassant quelque peu la frontière reconnue par Benveniste qui fait de la proposition l'horizon infranchissable de la linguistique, et transposant ces caractères des mots aux propositions, la linguistique textuelle a tenté d'étendre de diverses manières la problématique logico-grammaticale: par exemple, selon Van Dijk, un texte pouvait être résumé à une ou plusieurs macropropositions. Pour sa part, la théorie des Représentations Discursives de Kamp et Asher recode le texte en propositions logiques.

Cette conception grammaticale de l'unité (lexicale, phrastique puis textuelle) est tributaire de l'alliance millénaire de la grammaire avec la logique. Dès les grammairiens alexandrins, Denys le Thrace en premier lieu, les parties du discours sont distinguées selon leurs spécificités morphologiques (en grec) – mais ces distinctions sont homologuées à des capacités logiques: ainsi de la distinction entre catégorématiques et syncatégorématiques chez Priscien, devenue de nos jours distinction entre les mots “référentiels” et les “connecteurs”.

L'alliance de la grammaire et de la logique se fait sur le fond d'une ontologie commune dérivée de l'aristotélisme: les mots dits justement substantifs désignent des choses, les verbes des actions, les adjectifs des qualités (ce que reprend Charaudeau, 1992, p. 22 – et ces distinctions demeurent à l'œuvre dans les “ontologies” comme Wordnet).

La philosophie du langage, illustrée par des auteurs comme Austin et Searle, a maintenu jusqu'à nos jours, en pragmatique, en sémantique et en

syntaxe les principes ontologiques du parallélisme logico-grammatical et la notion traditionnelle d'unité. La grammaire générative a fait de même, et Chomsky n'est jamais revenu sur ce propos: "l'ensemble des "phrases" d'un système formalisé de mathématiques peut être considéré comme une langue" (1969, p. 15).

Le principe logique de compositionnalité demeure généralement accepté: puisque le sens d'une expression est une combinaison du sens de ses sous-expressions, le sens d'une phrase est composé du sens de ses mots; la linguistique textuelle, en projetant les catégories logico-grammaticales au-delà de la phrase, postule que le sens d'un texte est composé du sens de ses phrases. Il reste que ce principe, valide en logique frégréenne, ne suffit pas en logique (Frege lui adjoint d'ailleurs un principe de contextualité) ni *a fortiori* dans l'étude des langues: par exemple, une *pommade* n'est pas une préparation à base de pomme, une *archère* n'est pas la femme d'un archer, etc. Bref, le sens des unités minimales n'est aucunement donné et celui des unités qu'elles composent l'est encore moins.

En outre, dès sa formation voici deux siècles, la linguistique a dépassé la conception logico-grammaticale de l'unité, en créant les concepts de morphème et de phonème. Les mots ne sont plus apparus comme des unités linguistiques opératoires, mais simplement des conventions graphiques. La comparaison des langues a ruiné l'inventaire des parties du discours, notamment parce qu'un bon nombre et non des moindres ne comporte pas d'oppositions verbo-nominales. Des catégories comme celle de l'adverbe se sont avérées des fourre-tout incohérents. Enfin et surtout, en réfléchissant sur les fondements de la linguistique, Saussure a établi le caractère purement différentiel des unités¹, les privant de toute substance, ce qui a été confirmé ensuite par les études en perception. La linguistique s'est trouvée émancipée du fatras ontologique hérité de la scolastique par le biais de la tradition grammaticale et de la philosophie du langage. Ainsi, loin d'être données, les unités résultent d'une interprétation par mise en contraste en paradigme et en contexte proche ou lointain.

Peut-on alors concevoir les unités textuelles à l'image des unités lexicales ou grammaticales? Cette opinion a pu prévaloir en sémiotique: Hjelmslev envisageait ainsi la commutation d'unités textuelles. À présent, la linguistique textuelle et l'analyse du discours supposent ordinairement qu'une unité textuelle, paragraphe, séquence, chapitre, serait identifiable soit comme une chaîne de caractères (dans la tradition de Harris), soit comme la trace d'une intention énonciative, ou un acte (dans la tradition de la pragmatique discursive).

¹ "Un élément n'existera, pour commencer qu'à l'instant où on peut lui donner une signification différentielle (impliquant quelque différence). C'est l'abc." (Saussure, 2002, p. 323, cf. aussi p. 351).

Pour éclairer cette question, un recul épistémologique s'impose. Deux problématiques, logico-grammaticale et rhétorique/herméneutique, se partagent l'histoire des idées linguistiques occidentales. En bref, nous appellerons la première problématique du *signe* et la seconde problématique du *texte*.

1.1. *Unités ou formes?* – Les divergences entre problématiques apparaissent clairement à propos des unités textuelles. La conception logico-grammaticale tend à faire de l'unité un élément de "vocabulaire" textuel. À l'image d'une phrase considérée comme un enchaînement de mots, un texte résulterait d'un enchaînement d'unités (propositions, séquences, fonctions narratives, etc.), la linguistique textuelle a ainsi conçu le texte comme une suite structurée de propositions, la narratologie greimassienne a représenté le discours par une concaténation de fonctions narratives. Ces unités sont considérées comme discrètes et localisables, ce qu'atteste, par exemple, le nom de séquence².

L'ontologie logico-grammaticale attribue aux unités textuelles la discrétion et la présence, l'identité à soi et l'isonomie, à l'image naïve des objets physiques. La conception rhétorique / herméneutique que partage la sémantique interprétative admet en revanche que les objectivités qu'elle construit soient continues, parfois implicites, varient dans le temps et selon leurs occurrences et leurs contextes, connaissent entre elles des inégalités qualitatives et ne relèvent pas uniformément des mêmes règles. Elle ne rapporte pas exclusivement les formes sémantiques à des localisations spatio-temporelles, car ces formes ne sont pas des objets au sens chosiste du terme; du moins ne peut-on préjuger de leur forme d'objectivation en les soumettant aux procédures de l'analyse grammaticale: localisation, commutation, hiérarchisation avec unicité des rattachements, typage univoque des relations, caractérisation formelle de l'identité catégorielle. Les manifestations des unités textuelles peuvent en effet rester diffuses (isotopies, faisceaux isotopiques) ou rhapsodiques (thèmes). Une même forme peut au demeurant connaître des manifestations diffuses ou compactes. Bref, une forme n'est pas une unité discrète, stable, identique à elle-même: loin de s'opposer à d'évasives substances, les formes sont des *figures* qui contrastent sur des fonds.

Comme les textes présentent des contours de formes que l'interprétation a pour objectif de reconnaître et de parcourir, leur identification et leur parcours restant d'ailleurs indissociables, on doit compléter et sans doute dépasser la conception distributionnelle du texte par une conception morphosémantique qui tienne compte des inégalités qualitatives entre formes.

² Il est utilisé tant en linguistique textuelle qu'en narratologie. Définir l'unité par la localisation spatio-temporelle et l'identité à soi reste un geste caractéristique de l'ontologie classique, telle qu'elle a été perpétuée par la tradition aristotélicienne. Dès lors, tout phénomène complexe est conçu comme une combinaison d'unités et la description scientifique elle-même comme une analyse: par diverses formes de compositionnalité, cette position suppose la détermination du local sur le global.

Les problèmes d'identification des unités doivent être abordés en fonction de la dualité des problématiques: par exemple, la phrase est une segmentation logique (Benveniste la définit comme propositionnelle, *catégorématique*), la période une segmentation physiologique et/ou émotionnelle. Au-delà de la période, dont l'empan est sans doute mesuré par nos capacités motrices et respiratoires, le texte ne présente pas de signifiant identifiable par des procédures de segmentation, sinon les démarcations fortes, comme les pauses longues ou les changements de chapitre. C'est une raison fondamentale pour échapper au modèle du signe: les unités sémantiques textuelles n'ont pas de signifiants isolables comme des parties du discours; elles sont constituées par des connexions de signifiés des paliers inférieurs de la période, du syntagme, de la sémie. Ces connexions ne constituent pas un réseau uniforme: certaines sont mises en saillance, valorisées, modalisées.

L'opposition entre les conceptions logico-grammaticale et rhétorique / herméneutique de l'interprétation se concrétise enfin par des différences dans les régimes temporel et aspectuel des processus productifs et interprétatifs. À la régularité distributionnelle et itérative des intervalles égaux du temps logico-grammatical s'opposent dans le temps rhétorique / herméneutique les alternances du ponctuel et du duratif, du perfectif et de l'imperfectif.

1.2. *Redéfinir le signe comme passage.* – Il convient de proposer une redéfinition du signe qui s'accorde avec la problématique textuelle. L'unité, quelle que soit sa taille et son palier de description, peut être redéfinie comme un passage: or un passage n'a pas de bornes fixes et dépend évidemment du point de vue qui a déterminé sa sélection.

Définir le signe comme passage, c'est élaborer une définition purement relationnelle et donc contextuelle. Puisque la parole commande la langue, le signe est d'abord un "segment de parole"³: au plan du signifiant, c'est un extrait – entre deux blancs, s'il s'agit d'une chaîne de caractères; entre deux pauses ou ponctuations, s'il s'agit par exemple d'une période. L'extrait peut renvoyer aux étendues connexes, par exemple par des règles d'isophonie, ou de concordance de morphèmes: ce sont des *cooccurrents* expressifs.

Au plan du signifié, le passage est un *fragment* qui pointe vers ses contextes gauche et droit, proche et lointain. Cela vaut pour le contenu de la lexie comme pour celui du syntagme ou de la période, de la section, etc. Les méthodes statistiques de la linguistique de corpus permettent aujourd'hui de qualifier les unités contextuelles, qui sont elles aussi des passages de taille variable définis comme corrélats sémantiques:

³ Remarquablement, Saussure emploie l'expression *signe de parole* (cf. *ELG*, p. 265) mais non *signe de langue*.

Plan du contenu

$$\langle \text{corrélat}_1 \rangle \langle \text{corrélat}_n \rangle \supset \textit{fragment} \subset \langle \text{corrélat}_1 \rangle \langle \text{corrélat}_n \rangle$$

$$\langle \text{cooccurrent}_1 \rangle \langle \text{cooccurrent}_n \rangle \supset \textit{extrait} \subset \langle \text{cooccurrent}_1 \rangle \langle \text{cooccurrent}_n \rangle$$
Plan de l'expression

Figure 1: Le passage et ses contextes

La redéfinition du signe comme passage permet, nous le verrons, de considérer les unités textuelles comme des moments stabilisés dans des séries de transformations textuelles et intertextuelles, rapportées aux discours, champs génériques et genres.

1.3. *Paliers d'analyse*. – Faute de compositionnalité du sens, la problématique logico-grammaticale s'applique mal aux textes: les procédures de segmentation utilisant des balises sont utiles pour traiter de l'expression, mais sans plus. Aussi, la sémantique des textes a dû (re)définir d'autres formes d'unités et de relations qui en sont indépendantes: isotopies, thèmes et topoï, motifs et fonctions dialectiques, etc. (cf. l'auteur, 1989, 2001). Les isotopies sont des fonds sémantiques, les thèmes et topoï des formes qui peuvent être décrites comme des *molécules sémiques*, petits réseaux sémantiques dont les nœuds sont des sèmes et les liens des cas⁴. Voici par exemple la molécule sémique récurrente dans le vers d'Éluard, *Allume l'aube dans la source*:

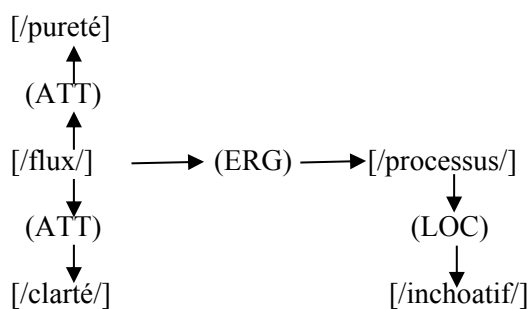


Figure 2: La molécule sémique d'une forme textuelle

N.B.: ATT abrège attributif, ERG, ergatif, LOC, locatif. Les liens du graphe sont entre parenthèses, les nœuds entre crochets. Cette représentation par graphe sémantique est adaptée des graphes conceptuels de Sowa, 1984.

⁴ Nous développons ici des rudiments présentés auparavant (l'auteur et coll., 1994, § La morphosémantique, ch. V.5, p. 130 sq.; 1996², Postface, Les formes sémantiques, p. 282 sq.; 1997, § 4, De l'ordre herméneutique aux formes sémantiques, pp. 134-142).

La même molécule est ici l'objet de trois lexicalisations successives, *allume*, *aube* et *source*, qui laissent apparaître des différences de statut des sèmes: par exemple, /pureté/ est inhérent à 'source', mais afférent à 'allume' dans ce contexte (cf. l'auteur, 1991, p. 203). Cet exemple permet de souligner quatre points qui ne sont pas sans portée sémiotique: (i) La représentation des formes sémantiques par des molécules sémiques est évidemment une simplification: les sèmes ne sont pas des atomes, mais des moments de parcours interprétatifs. (ii) Une forme sémantique, ici représentée comme une molécule sémique, peut être appariée à diverses formes expressives (ici les lexèmes *aube*, *allume* et *source*). (iii) De l'appariement d'une forme sémantique et d'une forme expressive résulte une forme textuelle. La stabilisation de formes textuelles est cruciale pour rendre compte de la sémosis. Bien qu'elles soient composées d'unités microsémantiques (sèmes et relations actantielles), la manifestation des formes sémantiques s'étend souvent sur des sections supérieures au paragraphe⁵: elles sont ainsi le lieu de médiation entre micro- et macro-sémantique et revêtent ainsi une fonction cruciale pour la production et l'interprétation du texte. (iv) Les formes textuelles ne satisfont pas à l'idéal d'un appariement bi-univoque entre signifié et signifiant: c'est pourquoi, au plan méthodologique, nous adoptons dans cette étude une perspective onomasiologique élargie qui procède des formes sémantiques aux formes expressives.

2. Formes textuelles et sémantique des parcours

On s'attachera dans cette section à questionner les rapports entre formes et fonds sémantiques au palier textuel, les transformations systématiques entre formes, les relations entre parties régulières et parties singulières des formes, enfin leurs indices de connectivité. Chaque fragment peut en effet manifester une part d'un fond sémantique, une section de forme sémantique ou une forme entière. En outre, chaque corrélat d'un fragment peut être un autre fragment, chaque cooccurrent d'un extrait un autre extrait.

2.1. *Forces et formes*. – Décrire les structures sémantiques comme des ensembles de relations statiques a certes le mérite de la simplicité et de l'économie: on peut typer ces relations comme des relations logiques, ce qui a valu leur succès aux premières descriptions par traits sémantiques – employées de fait par les partisans comme par les adversaires de la sémantique structurale. Comme la description des formes n'a de valeur que si l'on peut rendre compte de leur évolution, la morphosémantique a pour objectif de rendre compte des formes en termes de forces. Les deux aspects, force et

⁵ D'après nos évaluations, la manifestation d'une molécule sémique peut s'étendre à un espace d'environ 300 mots, soit une page standard, mais un mot peut suffire, comme on le voit ici.

forme, sont complémentaires: une force s'éprouve et se mesure par les déformations qu'elle induit; une forme stabilisée résulte d'un équilibre toujours momentané de forces. Le déséquilibre des forces conduit au déplacement des points critiques et à la déformation concomitante des sections normales de la forme.

On distingue en morphodynamique les points réguliers et les points singuliers. Comme une forme est reconnue par ses points singuliers plutôt que par ses points réguliers, certains des rapports que par analogie avec la perception l'on caractérise comme des rapports forme / fond peuvent être décrits ou reformulés comme des rapports entre sections régulières et sections singulières de la forme. Par exemple, au palier textuel, nous avons décrit les isotopies comme des produits de la loi gestaltiste de bonne continuation: elles définissent des portions régulières de formes textuelles, et apparaissent ainsi comme des fonds sémantiques. En revanche, les allotopies sont des points singuliers et certains tropes introduisent des discontinuités qualitatives par rupture d'isotopie⁶. Enfin, le rapport entre sèmes spécifiques et génériques peut se comprendre comme un rapport entre points réguliers et points singuliers: les molécules sémiques se rattachent par leurs sèmes génériques à des fonds et assurent leur continuité avec ces fonds. Les points singuliers des formes assurent leur saillance.

2.2. *Relations entre formes et fonds.* – L'étude des formes sémantiques se prolonge ainsi dans trois directions: liens entre fonds, dans le cas par exemple des genres qui comportent plusieurs isotopies génériques, comme la parabole; liens entre formes; et surtout liens des formes aux fonds, cruciaux pour l'étude de la perception sémantique.

En général, un fond sémantique ne se réduit pas à une seule isotopie et consiste en un faisceau d'isotopies. De ce fait, il n'est pas homogène mais comprend naturellement des irrégularités (pour un faisceau, les différences entre les isotopies, les ruptures ponctuelles et les disparitions d'isotopie); ces légères hétérogénéités permettent au demeurant de le percevoir.

Toutes les formes ne sont pas à égalité: outre qu'elles se stabilisent et se démembrent dans le cours du texte, *a fortiori* en diachronie, elles connaissent entre elles des inégalités qualitatives. Certaines formes doivent leur prééminence à leur valorisation: par exemple dans l'œuvre de Spinoza, la Nature est la forme sémantique qui semble dominer toutes les autres; ou encore, dans *Madame Bovary*, Emma; chez Bossuet, la Providence. Dans les textes théoriques, ces formes prééminentes sont les concepts qui subsument les autres; dans les récits, les agonistes qui résument une classe d'acteurs

⁶ Sans le lier à un quelconque degré zéro, nous reformulons ainsi le problème de l'écart qui préoccupe traditionnellement la stylistique.

(comme Don Juan). Elles revêtent ainsi le statut de *parangons* ou formes typiques⁷.

Enfin, certaines lexies concourent à l'établissement de fonds sémantiques (isotopies); d'autres lexicalisent de façon synthétique des formes sémantiques (molécules sémiques); d'autres enfin participent à la construction de formes qu'elles lexicalisent partiellement et manifestent ainsi des *contours* de ces formes. Soit en bref:

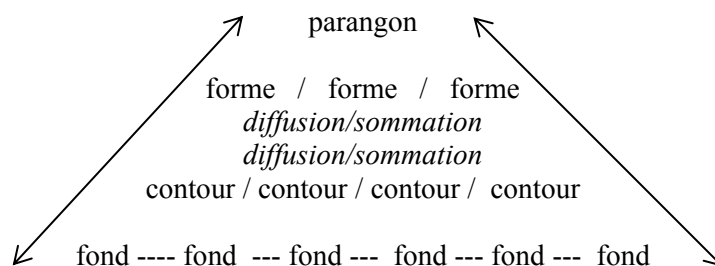


Figure 3: Médiations entre formes et fonds sémantiques

Par la médiation des contours, les formes et les fonds entrent ainsi au cours du texte dans des relations de *diffusion* des formes en fonds ou de *sommation* des fonds en formes (voir *infra*, § 3).

Plus qu'à des passages d'un fond à un autre ou d'une forme à une autre, les parcours interprétatifs entre fonds ou entre formes s'apparentent à la perception de formes ambiguës: ainsi, la métaphore fait percevoir simultanément deux fonds sémantiques, d'où l'effet anagogique qui lui est souvent attribué; une hypallage ou une syllepse font percevoir simultanément deux formes ou deux parties de formes, dans une ambiguïté qui rappelle les classiques illusions visuelles du canard-lapin ou de la duègne-ingénue.

2.3. *Le rapport entre local et global.* – Dans la mesure où elle échappe à l'atomisme de la tradition grammaticale, la conception morphosémantique du texte permet de détailler le concept de parcours interprétatif, puisque la (re)construction du sens consiste ainsi à restituer des formes et les fonds sur lesquels elles se détachent. Cela engage à concevoir le rapport du global au local d'une façon moins simpliste et moins statique que celle qui unit l'élément à l'ensemble ou la partie au tout. L'accès du global au local, dans

⁷ Cf. l'auteur 1991, p. 198 sq. Ces trois instances (fond, formes normales, parangons) se distinguent par des degrés de saillance, et l'on pourrait ainsi reverser à une théorie des formes sémantiques l'opposition saillance-prégnance empruntée aux études de perception animale par Thom et Petitot.

la mémorisation par exemple, reste en effet médiatisé par la reconnaissance et le parcours de formes sémantiques.

Le fond sémantique dominant a une fonction d'unification, et semble une émanation du tout dans la mesure où il définit la "matière" ou le "sujet" du texte. Le rapport entre tout et partie n'est pas un rapport d'abstraction qui érigerait le "sujet" du texte en type représenté par une "macroproposition" (selon Van Dijk). Alors qu'un fond homogène instaure une globalité du texte, les formes sémantiques la transforment par leurs évolutions en totalité progressive. L'homogénéité relative du fond sémantique et l'évolution des formes font de la textualité un réseau unique et complexe susceptible de divers parcours dont la cohérence postulée définit l'objectivité du sens.

3. Des formes à la textualité

3.1. *Les formes sémantiques et leurs dynamiques.* – Le sens d'un texte ne se déduit pas d'une suite de propositions, mais résulte du parcours de formes sémantiques⁸ liées à des formes expressives, les formes peuvent faire l'objet de diverses descriptions: par exemple, rapporté aux diverses composantes sémantiques, une molécule sémique peut être décrite comme thème, comme acteur, comme but ou source d'un point de vue modal, comme moment dans la linéarité du texte.

Si la description statique peut convenir à certaines applications, en didactique par exemple, une description plus fine doit restituer l'aspect dynamique de la production et de l'interprétation des textes. La première étape consiste à décrire les dynamiques des fonds et des formes: par exemple, la construction des molécules sémiques, leur évolution, et leur dissolution éventuelle. Ces dynamiques et leurs optimisations relatives sont paramétrées différemment selon les genres et les discours, car les formes et les fonds sont constitués et reconnus en fonction de ces diverses normes. En outre, comme les régimes de production et d'interprétation associés aux genres et aux discours guident le parcours des fonds et des formes, la sémantique des textes doit adapter ses descriptions à ces régimes.

La perception des fonds sémantiques semble liée à des rythmes, et celle des formes à des contours dont sur le plan de l'expression les contours prosodiques peuvent présenter un analogue. Si les fonds sémantiques sont constitués par des isotopies, en général produites par la récurrence de traits génériques, la temporalisation de ces récurrences est assurée par des rythmes sémantiques⁹. Comme les fonds sémantiques semblent des suites de points

⁸ On retrouve même dans la compréhension de textes des problèmes analogues à ceux que pose la reconnaissance de formes bruitées ou incomplètes.

⁹ On sait le rôle fondamental des rythmes dans la perception: ils ont un effet de facilitation à court terme, dont le corrélat linguistique est la création de zones de pertinence. Ils rendent ainsi compte pour une part de la *présomption d'isotopie* qui permet d'actualiser les sèmes.

réguliers et comme les formes sont discrétisées par leurs points singuliers, le parcours productif ou interprétatif de ces formes et de ces fonds suppose un rythme, cellule de base de toute action: il détermine les segments réguliers des formes, aux extrémités desquels se détachent les points singuliers.

Au palier macrosémantique, des inégalités qualitatives marquent des lieux ou moments remarquables que l'on pourrait appeler des points *nodaux sémantiques*, définis par leur haut degré de connectivité: les plus faciles à isoler sont des mots qui connectent plusieurs isotopies génériques, ou des répliques qui transforment la structure narrative. Au-delà d'une concaténation de symboles, on peut concevoir ainsi le texte comme un *cours d'action* sémiotique, scandé par des moments critiques. Le genre codifie la conduite de cette action, mais ce qu'on pourrait appeler le *ductus* particularise un énonciateur et permettrait de caractériser le *style sémantique* par des rythmes et des tracés particuliers des contours de formes.

3.2. *Les transformations.* – Retenons qu'une forme est une *famille de transformations*: la forme et la métamorphose sont deux moments d'un même processus: cette transformation est diachroniquement orientée dans un temps du texte et de la tradition. Selon les composantes mises en jeu, il faut distinguer les transformations thématiques, dialectiques (narratives), dialogiques (modales, selon les "points de vue" et les "positions de parole"), tactiques (positionnelles). Nous nommerons *métamorphismes* l'ensemble de ces transformations. Nous y incluons les *méréomorphismes*, définis comme des relations entre parties du texte qui présentent de manière compacte et locale des formes amplifiées ailleurs de manière globale et diffuse. Par exemple, des configurations codifiées comme la description initiale, la parabole, le rêve annonciateur, sont transposées dans la suite du texte par d'autres formes plus étendues. Les méréomorphismes traduisent en linguistique textuelle des phénomènes de *solidarité d'échelle*.

Outre les métamorphismes, qui intéressent les formes sémantiques, il faut tenir compte enfin des *transpositions* qui intéressent les changements du *fond* sur lesquels elles sont perçues: toute forme est en effet définie par sa transposabilité (d'où les antiques théories des formes pures).

La génération d'un texte consiste en une série de métamorphismes et de transpositions, qu'on peut mettre en évidence à l'oral par l'étude des reformulations, à l'écrit par celle des brouillons. Son interprétation consiste pour une bonne part dans l'identification et l'évaluation des métamorphismes: par exemple, le sens d'un récit est articulé par des transformations thématiques et dialectiques.

La théorie des métamorphismes intéresse également les tropes, qui loin d'être des écarts par rapport à une norme, participent des parcours de transformation, à des moments remarquables des parcours génétiques et interpré-

tatifs; les plus discutés correspondent sans doute à des *points critiques*.¹⁰ Ils assument alors quatre fonctions générales, selon qu'ils modifient les fonds sémantiques, les formes sémantiques ou les relations entre formes et fonds: (i) Rupture de fonds sémantiques (allotropies) et connexion de fonds sémantiques (polyisotropies génériques). (ii) Rupture ou modification de formes sémantiques: ces transformations s'opèrent par addition ou délétion de traits sémantiques des molécules sémiques. (iii) Modification réciproque de formes sémantiques par allotopies spécifiques (antithèses) ou métathèses sémantiques (ex. l'hypallage). (iv) Modification des rapports entre formes et fonds: toute transposition d'une forme sur un autre fond modifie cette forme, d'où par exemple les remaniements sémiques induits par les métaphores: ainsi, la blonde chevelure de la Duchesse de Guermantes se métamorphose-t-elle dans les frangipanes que l'on mange à la sortie de la messe de Combray, où elle apparaît inoubliablement. Si l'exemple de transposition le plus illustre reste la métaphore, une syllepse ou une antanaclase par exemple peuvent en articuler également si le contexte y conduit. La métonymie, qui est une classe de transformations plutôt qu'une figure, est également un exemple de métamorphisme.

L'évolution des formes dans le texte est marquée par diverses inégalités qui dessinent des parties: on peut distinguer par exemple les *démarcations* de l'expression (comme le début et la fin d'une section), ou les *tournants* sémantiques (thématiques, dialogiques, dialectiques). Ainsi les introductions (thématiques, dialectiques ou dialogiques), les ruptures (thématiques, dialectiques ou dialogiques), les sommations (thématiques) ou les basculements (dialectiques), enfin les révélations (dialogiques).

3.3. *Diffusion et sommation*. – Les formes et les fonds sont interdépendants et l'on ne peut transporter les formes d'un fond à l'autre sans qu'elles se modifient dans cette transposition. En effet, les fonds et les formes sont entre eux dans un rapport de diffusion et de sommation: un fond est une forme "oubliée", au sens où elle a perdu sa saillance, et les passages de l'avant-plan à l'arrière-plan se comprennent ainsi.

La *diffusion sémantique* rend compte des phénomènes d'isotopie locale: comme tout trait sémantique est actualisé et *a fortiori* propagé à partir et en fonction du contexte immédiat et lointain, les contextes manifestent des redondances locales multiples. Par exemple, dans un corpus de romans français, on trouvera diverses formes du verbe *tordre* dans les contextes du mot *grimace*.

Les phénomènes de diffusion jouent un rôle important au palier textuel. Par exemple, dans *Hérodias* de Flaubert, la première phrase commence par

¹⁰ Les points critiques d'une forme sont soit les sèmes instables, soit les sèmes saillants, donc caractérisants, nécessaires à son identification à travers ses transpositions (le rouge du Chaperon rouge).

ces mots: “La citadelle de Machaerous se dressait à l’Orient de la mer Morte sur un pic de basalte, ayant la forme d’un cône.” Dans le contexte, on peut analyser: (i) ‘Orient’: /solaire/, /inchoatif/, /ascendant/, (/vital/); (ii) ‘Mer morte’: /maritime/, /terminatif/, /descendant/, /funèbre/. Dans la dernière description du panorama sur le désert de Judée, quatre pages plus loin¹¹, les sèmes de la molécule ‘Mer Morte’ sont repris pour constituer un fond: les sèmes /maritime/, /descendant/, /funèbre/ alternent et se combinent: ‘grands flots’ /m/ ‘pétrifiés’ /f/, ‘gouffres’ /m/ /d/, ‘noirs’ /f/, ‘falaises /m/’, ‘profondeur’ /d/ des abîmes /f/ /d/’, ensevelis /f//d/ ‘plus bas’ /d/ que les ‘rivages’ /m/ sous /d/ les ‘eaux’ /m/ ‘pesantes’ /d/. D’une forme initiale, la molécule sémique de ‘Mer Morte’ est ainsi devenue un fond: les liens casuels entre ses sèmes ne sont pas conservés et ses sèmes sont manifestés de façon récurrente et rhapsodique. Cette forme sémantique réapparaîtra de façon partielle dans le cours du récit: ultérieurement, les malédictions du Baptiste sont par exemple comparées à des flots¹².

Dans les moments de diffusion, la suspension des liens casuels a pour effet de dissoudre la forme initiale; mais ils restent paradigmatiquement présentes et remotivables dans les moments de sommation.

Des mouvements de diffusion et de sommation organisent aussi le plan de l’expression. Ainsi, chez un auteur aussi peu cryptique que Racine, Valérie Beaudouin a pu prouver par des méthodes statistiques que les phonèmes du nom du personnage principal, surtout quand il est éponyme, se trouvent significativement diffusés dans l’ensemble du texte (Beaudouin, 2002, §8.3.2). Pour ce qui concerne la sommation phonétique, on relève aisément des exemples comme cette phrase de Baudelaire (cf. Starobinski, 1971, p. 158; je garde sa notation):

Je sentis ma gorge serrée par la main terrible de l’hystérie
 HY S TERI (HYSTÉRIE)

Les mouvements de diffusion et sommation constituent des inégalités qualitatives, par transformation réciproque des groupements de points singuliers en séries de points réguliers. Du point de vue psycholinguistique, ces mouvements correspondent vraisemblablement à des moments de préparation du geste énonciatif, puis à leurs effets de récence dans la mémoire immédiate du cours d’action.

¹¹ “Tous ces monts autour de lui, comme des étages de grands flots pétrifiés, les gouffres noirs sur les flancs des falaises, l’immensité du ciel bleu, l’éclat violent du jour, la profondeur des abîmes le troublaient, – et une désolation l’envahissait au spectacle du désert, qui figure, dans le bouleversement de ses terrains, des amphithéâtres et des palais abattus. Le vent chaud apportait, avec l’odeur du soufre, comme l’exhalaison des villes maudites, ensevelies plus bas que les rivages sous les eaux pesantes.”

¹² Sans doute par la médiation du cliché *flots d’injures*, resté implicite même dans les brouillons, alors que *pluie d’injures* figure dans la version finale.

3.4. *L'unité contenu-expression et la sémiologie textuelle.* – Manifeste dans les parcours énonciatifs et interprétatifs entre signes, l'interdépendance des deux plans du langage intéresse évidemment les paliers de complexité supérieurs, comme au plan de l'expression les schémas prosodiques ou métriques et au plan du contenu les reprises ou transformations de molécules sémiologiques (*topoi*, acteurs, fonctions narratives). Cependant, faute de correspondance assurée entre les formes sémantiques et les formes expressives, on trouve des acteurs ou des thèmes sans nom déterminé, ou dont la désignation évolue (par exemple, l'*étudiant* au premier chapitre du *Père Goriot* devient *Rastignac* aux suivants).

L'unité entre contenu et expression s'établit par les parcours interprétatifs conformes au genre: par exemple, dans un corpus romanesque, Évelyne Bourion a pu confirmer la corrélation entre des noms de sentiments et les ponctuations dans les contextes où ces noms apparaissent. L'appariement entre formes sémantiques et formes expressives assume ainsi un rôle crucial dans la sémiologie au palier textuel.

Les fonds sémantiques et expressifs obéissent à la loi perceptive de bonne continuation à l'œuvre dans les isotopies et les isophonies – en comprenant par là aussi les phénomènes rythmiques qui assurent des répétitions sans se réduire à la répétition du même. En revanche, les formes sémantiques et expressives sont construites par des processus de discrétisation: en activant notamment des sèmes et des phèmes qui sont des points singuliers, ils constituent des contours.

La discrétisation résulte de deux processus, l'activation et l'inhibition. Les travaux sur les cooccurrences décèlent des activations réciproques localisées. Comme le sens est fait de différences, faire sens c'est discrétiser, c'est-à-dire limiter voire éliminer les stimuli environnants, comme le rappelle dans le domaine de la perception l'omniprésence de l'inhibition latérale.

L'interprétation restitue ces parcours de discrétisation que l'activité artistique systématiser par la stylisation (par exemple, la ligne dans le dessin accentue les séparations des aplats perceptibles). Ainsi, pour paraphraser Saussure, les deux chaos apparents des sons et des sens, en s'unissant, donnent un ordre. Et l'investigation en corpus permet alors d'identifier des phénomènes de sémiologie textuelle, qui n'est aucunement donnée par les chaînes de caractères qui constituent le document.

4. Pour un agenda

4.1. *Le décloisonnement de la limite phrastique.* – Privilégiés par la tradition logico-grammaticale, les mots par excellence sont les substantifs, réputés représenter des substances et assurer ainsi l'ancrage ontologique du langage. En revanche, le passage n'a pas d'ancrage dans une référence extralinguistique – et la question de la vérité n'a pas de pertinence en linguistique puisqu'elle a toujours relevé de la métaphysique. Il nous faut cependant y reve-

nir, car les lectures benvenistiennes de Saussure se sont efforcées de rétablir le caractère dirimant de la prétendue limite phrastique¹³. Le précellence de la prédication reste toujours fondée sur l'idéal du jugement vrai, c'est-à-dire, sauf cas de tautologie, sur la représentation d'un "état de choses". Aussi la question de la prédication nous permettra d'illustrer l'opposition entre ontologie et praxéologie.

Saussure appelle *syntagme* tout ce qui relève de la syntagmatique, laquelle inclut tous les faits de syntaxe. Mais l'empan d'un syntagme peut parfaitement dépasser la phrase, car la phrase n'est qu'un "type de syntagme". Ainsi, à l'inverse de Benveniste, Saussure n'accorde aucune prééminence à la phrase, mais apporte cependant sur ce palier d'organisation un point de vue nouveau: il ne reprend pas la conception prédicative de la phrase, mais la compare à une activité productive. C'est en effet au palier de la phrase que se pose le plus clairement le problème de la praxéologie, si l'on considère la parole comme un cours d'action. Adoptant de manière elliptique mais claire un point de vue praxéologique, Saussure éprouve le besoin de différencier une phrase d'un rite – alors même que personne, à ma connaissance, hors de la théologie sacramentelle¹⁴, n'avait même songé à les rapprocher: "Un rite, une messe, n'est pas comparable du tout à la phrase, puisque ce n'est que la répétition d'une suite d'actes. Phrase est comparable à l'activité du compositeur de musique (et pas à celle de l'exécuteur)" (2002, Nouvelles notes Item). Dans ce passage, la phrase est bien considérée comme une suite d'actes (et non une proposition ou une représentation), et c'est seulement sur le caractère instituant de ces actes que la réflexion se porte: la phrase est comparée non (ou non seulement) à un *objet* culturel, mais à une *activité* créatrice.

Une contradiction se fait jour ainsi, entre la phrase, ordinairement considérée comme unité syntaxique close, et le passage (le "syntagme" saussurien) défini comme activité compositionnelle créatrice. Cela évoque certes la générativité chomskienne, mais pour Chomsky, la générativité est tout entière gouvernée par des règles opérant sur des unités données et préexistantes, alors que la syntagmatique saussurienne crée sans cesse de nouvelles grandeurs. Parmi elles, des formes qui ne sont que des moments dans des séries de transformations: c'est ainsi que la théorie des métamorphismes argumente pour une conception transformationnelle du texte (qui ne postule aucune transformation de structures profondes qui engendraient des structures de surface). L'activité propre de l'élaboration du texte (écrit et oral) est redoublée par l'activité interprétative: les unités linguisti-

¹³ Cf. Benveniste postulant la «prédication ontologique» du syntagme nominal et affirmant: «Il n'y a pas de niveau linguistique au-delà du niveau catégorématique» (1966, p. 129; cette position est rappelée et soutenue par Jean-Michel Adam et les principaux tenants de l'Analyse du discours, comme Maingueneau).

¹⁴ Cf. l'excellent ouvrage d'Irène Rosier, 1995, *La parole comme acte*, Paris, Vrin.

ques sont ainsi restituées ou reconstituées par l'interprète, comme le montrent les phénomènes de défigement, qui accroissent le nombre des lexies: par exemple, dans *Sarkozy monte au créneau*, *monte au créneau* compte pour une lexie unique, tandis que dans *Bayard monte au créneau*, cette expression recèle trois lexies. Ainsi, loin d'être donnée, toute unité syntagmatique dépend-elle du point de vue qui l'établit, en fonction de normes textuelles qui n'ont jamais été décrites comme des règles. Ainsi les unités linguistiques de toute taille sont-elles le résultat d'une sémosis, et non son point de départ.

4.2. *Grammaire de la période et activité compositionnelle*. – Alors même que toute la sémantique logique et une bonne part de la sémantique cognitive réduit les textes à des enchaînements de propositions, beaucoup reste à faire pour relier l'en-deça et l'au-delà de la période, pour lever l'obstacle épistémologique qu'élève la surimposition de la proposition logique à l'analyse des périodes et des unités textuelles de toute grandeur.

Les mouvements inventoriés par le Groupe de Fribourg (2013), comme l'identification, la préparation, l'action témoignent d'une conception praxéologique de l'activité constructive dans la période. Ils pourraient parfaitement être réanalysés dans un cadre morpholinguistique de construction de formes et de fonds sémantiques et expressifs. Pour en rester à la sémantique, prenons l'exemple cité p. 202:

Les Égyptiens leurs dieux *c'est les chats*

Les Égyptiens est un "cadrage" qui détermine un fond isotopique (l'Égypte – antique, ce qui est inféré à partir de 'dieux'). *Leurs dieux* introduit un élément de forme (1), *c'est les chats* un autre (2), qui sont combinés par le présentatif à valeur équative. Forme et fond sont ici liés par le possessif *leurs*. Soit, schématiquement, dans la notation propre aux graphes sémantiques:

FORME [*Leurs dieux* (1)] <---(ATTRIBUTIF *c'est*)---> [*les chats* (2)]
 |
 FOND *Les Égyptiens* -----

Les nouveaux observables mis en évidence par la syntaxe de l'oral (et d'abord les études du français parlé dirigée par la regrettée Claire Blanche-Benveniste) ont permis de dépasser le logicisme prédicatif. Ce tournant empirique s'est accompagné d'un essor théorique qui permet de restituer la place de la prosodie et de montrer comment l'intonation, notamment, relie les formes sémantiques aux formes expressives.

4.3. *Des cooccurrences aux formes et fonds textuels.* – Les nouveaux observables mis en évidence par la linguistique de corpus appellent également des remaniements théoriques. Ils intéressent notamment l’interprétation des sorties logicielles, qui ne sont jamais par elles-mêmes des résultats, mais des indices à interpréter. Or, on peut les caractériser en fonction de leur statut ultérieur, dans la perspective d’une analyse morphosémantique. Ainsi, les chaînes de caractères qui ont des fréquences élevées traduisent des phénomènes “de fond” et sont uniformément réparties dans les textes. Cette hypothèse a été confirmée par la thèse de Perlerin sur l’isotopie météorologique dans les bulletins d’informations boursières.

Mais les méthodes probabilistes ne sont pas seulement fréquentistes en ce sens; des tests comme celui de l’écart réduit permettent de faire ressortir des indices bayésiens: locaux, et susceptibles d’une significativité élevée, ils correspondent souvent à des points singuliers des formes sémantiques, et nous avons souvent: par exemple, dans *Toine*, de Maupassant, le mot *dix* (anormalement fréquent) permet de relier des formes analogues sur différentes isotopies comme le hameau de *dix* maisons sur l’isotopie géographique, les *dix* poussins sur l’isotopie animale, ce qui complète une analyse “manuelle” antérieure (l’auteur, 1989, II, ch. 3).

Dans le cas où le test de l’écart réduit souligne des mots sémantiquement reliés parce qu’ils lexicalisent des éléments d’une même forme (thème ou acteur, par exemple), ils forment cependant un “nuage” de chaînes de caractères, qui peut s’étendre sur trois cents mots (environ une page). Ce “nuage” n’a rien d’une unité au sens grammatical du terme: il est discontinu, non-borné, et son étendue comme ses composants varient avec le seuillage que l’on impose au test d’écart réduit. Il manifeste pourtant une unité textuelle, d’un contenu (la molécule sémique du thème ou de l’acteur) et d’une expression (les chaînes de caractères soulignées et qualifiées par l’interprétation).

Mieux encore, les manifestations d’unités textuelles peuvent se chevaucher, et il est fréquent par exemple qu’un thème commence alors que le précédent n’a pas fini de se manifester.

Enfin, les tests probabilistes, quand ils sont employés comme auxiliaires de la méthodologie comparative en linguistique de corpus, permettent de mettre en évidence des mots absents (ou *nullax*): par exemple, le mot *homme* est absent des textes racistes. Cela reste inconcevable pour les conceptions positivistes du texte véhiculées par les grammaires, notamment formelles. Or sans même évoquer les censures ou refoulements, les textes oraux ou écrits ne signifient pas moins par ce qu’ils disent que par ce qu’ils taisent; quiconque a travaillé sur le discours politique le sait bien.

Tout cela peut présager une redéfinition de la notion même d’unité, déjà initiée par Saussure puis Hjelmslev, de manière à pouvoir rendre compte de la complexité des textes tant sur le plan de l’expression que sur celui du contenu.

Bibliographie

- Blanche-Benveniste, Claire (1998). *Le français parlé, études grammaticales*. Paris: Editions du CNRS.
- Beaudouin, Valérie (2002). *Rythme et mètre du vers classique. Corneille et Racine*. Paris: Champion.
- Benveniste, Émile (1966). *Problèmes de linguistique générale*. Paris: Seuil.
- Charaudeau, Patrice (1992). *Grammaire du sens et de l'expression*. Paris: Hachette.
- Chomsky, Noam (1969). *Structures syntaxiques*. Paris: Seuil.
- Floettum, Kjersti et Rastier, François (éds.) (2003). *Academic Discourse – Multidisciplinary Approaches*. Oslo: Novus.
- Loiseau, Sylvain (2003). Philosophical discourse from autonomy to engagement: Deleuze commentator of Spinoza, in Kjersti Flottum et François Rastier (éds.) *Academic Discourse – Multidisciplinary Approaches*. Oslo: Novus.
- Groupe de Fribourg (2013). *Grammaire de la période*. Berne: Lang.
- Perlerin, Vincent (2004). *Sémantique légère pour le document*, Thèse d'informatique. Caen: Université de Caen.
- Petitot, Jean (1996). Les modèles morphodynamiques en perception visuelle. *Visio*, I, pp. 65-73.
- Pinker, Steven (1999). *Words and Rules. The Ingredients of Language*. New York: Harper Perennial.
- Rastier, François (1971). *Idéologie et théorie des signes*. La Haye: Mouton.
- Rastier, François (1996²). *Sémantique interprétative*. Paris: PUF.
- Rastier, François (1989). *Sens et textualité*. Paris: Hachette. Rééd. <http://www.revue-texto.net>
- Rastier, François (1991). *Sémantique et recherches cognitives*. Paris: PUF.
- Rastier, François (1994). *Sémantique pour l'analyse*, avec la collaboration de Marc Cavazza et Anne Abeillé. Paris: Masson.
- Rastier, François (1997). Stratégies génétiques et destruction des sources – L'exemple d'Hérodiade, in Eric Le Calvez et Marie-Claude Canova-Green. (éds.) *Texte(s) et intertexte(s)*. Amsterdam-Atlanta: Rodopi, pp. 193-218.
- Rastier, François (2000). De la sémantique cognitive à la sémantique diachronique: les valeurs et l'évolution des classes lexicales, in Jacques François (éd.) *Théories contemporaines du changement sémantique, Mémoires de la société de linguistique de Paris*, IX. Louvain: Peeters, pp. 135-164.
- Rastier, François (2001). *Arts et sciences du texte*. Paris: PUF.
- Rastier, François (2003). Semantic approaches of theoretical texts, in Kjersti Flottum et François Rastier (éds.) *Academic Discourse – Multidisciplinary Approaches*. Oslo: Novus.
- Rastier, François (2011). *La mesure et le grain – Sémantique de corpus*. Paris: Champion.

- Saussure, Ferdinand de (2002). *Écrits de linguistique générale*. Paris: Gallimard [éd. Simon Bouquet et Rudolf Engler].
- Sowa, John. 1984. *Conceptual Structures*. New York: Addison Wesley.
- Starobinski, Jean (1971). *Les mots sous les mots*. Paris: Gallimard.
- Valette, Mathieu (2003). Conceptualisation and Evolution of Concepts. The example of French Linguist Gustave Guillaume, in Kjersti Flottum et François Rastier (éds.) *Academic Discourse – Multidisciplinary Approaches*. Oslo: Novus.
- Wetherill, Philip. M. (éd.) (1988). Flaubert: *Trois contes*. Paris: Garnier.