Publié dans *Énoncer – l'ordre informatif dans les langues*, dir. Pierre Cotte, Martine Dalmas et Hélène Wlodarczyk, Coll. Sémantiques, L'Harmattan, Paris.

Centres d'intérêt et ordres communicatifs

André WLODARCZYK

Université Charles de Gaulle (Lille 3) & Équipe de recherche « Forme – Discours – Cognition »

1 La linéarisation des énoncés: du parallèle au séquentiel

Étant donné que l'information à communiquer ne peut l'être que sous forme linéaire (donc, du point de vue logique, séquentielle), en supposant que sa représentation dans la mémoire est de nature parallèle, le locuteur est forcé de *faire des choix*. Par exemple, au cas où le locuteur doit linéariser les deux processus parallèles P || Q, il a deux possibilités:

- soit (1) représenter en premier le processus P par la séquence S1 (S1 || P), puis représenter le processus Q par la séquence S2 (S2 || Q) de sorte que S1 précède S2

- soit (2) représenter en premier le processus Q en tant que séquence S1 1 (S1 || Q), puis présenter le processus P en tant que séquence S2 (S2 || P) de sorte que S1 précède S2.

On peut supposer que cette *contrainte de choix* est à l'origine de la nécessité d'établir un ordre de priorités : (S1 || P) > (S2 || Q) ou (S1 || Q) > (S2 || P). Cet état de choses n'est qu'en partie comparable avec les principes de la logique qui nous invite à faire abstraction de l'ordre de notation des relations.

La difficulté principale dans l'analyse et la synthèse des énoncés provient de ce que, malgré leur forme linéaire, leurs différents constituants sont en fait le résultat de processus **parallèles**, notamment ceux qui concernent l'information elle-même (sémantique) et ceux qui concernent la métainformation (pragmatique métainformationnelle). La notion classique de *sujet* dans les langues européennes cause tant de problèmes aux linguistes formalisants (qu'ils soient structuralistes ou génératives) parce que le sujet est un véritable nœud où l'on trouve à la fois une composante sémantique et une composante pragmatique. L'important est de ne pas confondre ces

composantes bien que les expressions linguistiques les expriment de manière synthétique (ce qui résulte d'opérations diverses).

Dans les langues naturelles, l'énoncé élémentaire est le résultat de deux opérations que j'appelle *formation* et *composition*[1]. La *formation* de l'énoncé concerne l'ensemble des relations sémantiques souvent appelées prédicatives[2] en linguistique. Selon nous, elle correspond à la situation sémantique dont seuls certains éléments sont exprimés. En effet, bien que le locuteur possède l'image de la situation sémantique (actants sémantiques définis comme participants jouant des rôles sémantiques, ancrages spatio-temporels, moyens et buts ainsi que toutes sortes de repérages par rapports aux quatre espaces topologiques présentés plus bas) il n'exprime jamais tout. Il va sans dire qu'afin de construire une description *métalinguistique* de la situation sémantique, nous emprunterons à la logique la notion de prédication (elle-même passée à la logique à partir de la grammaire traditionnelle) comme forme de représentation de nos connaissances. Cependant, nous n'établissons pas d'équation entre la représentation formelle de la prédication et la réalité linguistique qui est censée exprimer la situation sémantique, ceci en raison de la partialité des messages linguistiques par rapport aux intentions des locuteurs. Cette démarche nous permet de reprendre le terme prédication au sens linguistique dans ce que nous appelons la *composition*. Cette dernière ajoute à la *formation* une présentation de nature considérée comme « informative » (qui est, en fait, métainformative[3]). La *composition* consiste notamment, pour le locuteur, à choisir parmi les participants de la situation sémantique un *centre d'intérêt*[4] et à attribuer une valeur métainformative aux

---

[1] J'ai emprunté les termes de formation et composition à Mizutani 1991. Le terme de composition me semble bien adapté pour désigner des opérations qui se laissent modéliser par les espaces abstraits tels que les définit la topologie (Wlodarczyk, A, 1999).

[2] En effet, bien que les théories sémantiques prennent le plus souvent comme modèle la logique des prédicats du premier ordre où pourtant ces derniers sont de nature syntaxique (formelle), en linguistique, on procède sans toujours préciser ce que les relations prédicatives ont de *syntaxique* (forme de la relation et de ses arguments ou places vus comme sujet et compléments régis) ou de *sémantique* (il s'agit de la situation et de ses participants avec les rôles qu'ils jouent dans la situation).

[3] Du moment où tout message linguistique contient de l'information de niveau 1 quel que soit son caractère (rationnel, émotif, modal etc.), les informations de niveau 2, qui présentent les informations de niveau 1, sont de nature métainformative.

[4] La Théorie du centrage (Centering Theory), récemment élaborée en linguistique computationnelle (Grosz, B. J., Joshi, A. K., Weinstein S., 1995), jette

différentes parties de l'énoncé ou à l'énoncé dans son ensemble. Les opérations que je décris comme *composition* (ou aussi *composition protocolaire* pour insister sur l'importance de l'ordre linéaire) sont, pour certaines, connues sous le nom de « divison actuelle de l'énoncé » dans la tradition pragoise qui fut à l'origine de l'étude de la structure métainformative de l'énoncé (Mathesius 1939).

Pour modéliser les fonctions linguistiques, nous proposons de définir différents modules qui interagissent dans la production et la compréhension du discours. Nous en parlerons ici en termes d'*espaces*, en nous inspirant de la topologie. Nous en dégageons quatre :

a. l'Espace de la Parole (locution où interagissent les participants de l'acte locutif à un temps t et dans un lieu l),

b. l'Espace de la Sémantique (formation des expressions où il est question des situations sémantiques et de leurs acteurs ou rôles participatifs),

c. l'Espace de la Pragmatique (composition protocolaire des énoncés où le locuteur attribue la valeur modale de la vérité à la situation sémantique et où il structure la présentation de l'information, cf. les centres d'intérêt),

d. l'Espace de la Cognition (connaissances communes et encyclopédiques, croyances et autres connaissances liées à l'expérience individuelle).

On peut penser que l'interaction de ces quatre espaces est réalisée en parallèle[5] car en chaque point d'un énoncé interviennent à la fois les structurations *locutive, sémantique, pragmatique* et *cognitive*. D'autre part, comme le discours se développe dans le temps, la mise en place des contenus successifs (parmi lesquels nous distinguerons ci-après les « centres d'intérêts » comme les plus saillants) se réalise de manière dynamique. Cependant ni la linguistique formelle ni même la linguistique informatique n'ont encore mis au point de modèles capables de prendre en compte le caractère dynamique de la production des énoncés dans le discours. Nous espérons que les linguistes ou les informaticiens eux-mêmes pourront trouver des suggestions utiles sur cette question en s'inspirant de la programmation des processus parallèles en informatique (cf. Wlodarczyk, A., 2003b). Tout en restant linguistes, il est intéressant de remarquer que, par exemple, la valence verbale (à la différence

---

une lumière nouvelle sur certains concepts développés depuis longtemps dans le cadre de la perspective fonctionnelle de l'Ecole de Prague. La théorie du centrage repose notamment sur la notion de "centre d'attention", élaborée en psychologie et appliquée à l'analyse du discours.

[5] Pour plus de précisions sur la relation entre les processus séquentiels et parallèles, cf. Wlodarczyk, A, 2003b.

de la prédication en logique) est par sa nature même hétérogène, réunissant les résultats de « fonctions » variées, à la fois de type sémantique (rôles) que pragmatique (centres d'intérêt). C'est notamment pour cette raison que le modèle prédicat-arguments (tel qu'il est défini dans la logique élémentaire) ne rend que partiellement compte de la valence verbale des langues naturelles.

2. Identification des centres d'intérêt des messages linguistiques

Linéariser des symboles (c.-à-d.: les rendre séquentiels) suivant les résultats d'opérations parallèles, c'est construire des objets linguistiques (de niveaux variés de complexité) connus comme *unités du discours*. Les linguistes s'accordent souvent à considérer ces unités comme structurées et exprimant (« représentant ») les niveaux respectifs. Ainsi, les différents types d'expressions constituant le texte (discours) s'organisent en structures qui s'emboîtent les unes dans les autres (Wlodarczyk, A, 2001):

1. lexique : unités linguistiques élémentaires,

2. expression : composition d'unités linguistiques élémentaires produites avec l'intention d'informer et construites à partir de configurations d'instances du lexique; les expressions se subdivisent en deux classes:

2.1. énoncé non étendu : expression simple (contenant un ou plusieurs centres d'intérêt souvent obligatoires; par exemple: le Sujet et l'Objet) ne pouvant être réduite,

2.2. énoncé étendu: expression composée (contenant un ou plusieurs centres d'intérêt facultatifs; par exemple: le Topique et le Focus).

3. texte/dialogue : ensemble ordonné d'expressions linguistiques (contenant un ou plusieurs centres d'intérêt; par exemple: les Thèmes généraux et les Thèmes particuliers).

TYPE D'EXPRESSION	**CENTRE D'INTERET**	
	Global	**Local**
Énoncé non étendu	SUJET	OBJET
Énoncé étendu	TOPIQUE	FOCUS
Texte / dialogue	THEME GENERAL	THEME PARTICULIER

Tableau 2. Les Pivots du discours

Il va de soi que les expressions reflètent un certain nombre de régularités que j'appelle **syntaxe** dont le rôle consiste uniquement à mettre en *ordre* de

manière *séquentielle* un certain nombre de contenus[6], eux-mêmes étant organisés dans l'esprit des locuteurs de manière *parallèle*. Dans ce cas, la syntaxe peut être conçue comme résultat d'un compromis entre les deux types *d'ordres : séquentiel* et *parallèle*. Ce qui, croyons-nous, structure les expressions linguistiques avant toute chose, c'est ce que nous appelons *l'identification des centres d'intérêt* : en effet, le locuteur en construisant son message n'a pas d'autre *choix* que de repérer un ou plusieurs centres d'intérêt à l'intérieur des quatre topologies ou *espaces linguistiques*. Cette opération d'identification est fondamentale pour la mise en forme du message autour d'un pivot que ce soit sous la forme d'un sujet, d'un objet, d'un topique ou d'un focus.

3. La validation métainformative de l'énoncé

Le choix des centres d'interêt est lié à la validation métainformative attribuée par le locuteur aux différentes parties de l'énoncé : ce que la théorie pragoise appelle l'opposition du 'donné' et 'nouveau'. Mais contrairement à cette dernière théorie, nous aimerions souligner que *tous* les énoncés ne sont pas divisés en *donné* et *nouveau*, comme certains exégètes de Mathesius ont pu le penser. De plus, au cours de nos recherches, les concepts de *donné* et *nouveau* se sont avérés trop approximatifs (pas assez précis) pour décrire les divers centres d'intérêt. Notamment, pour expliquer « l'actualisation des éléments de l'énoncé » d'abord dans les groupes nominal (Wlodarczyk, A. [1979, 1996]) et verbal (Wlodarczyk H. [1979, 1988]), ensuite dans la corrélation entre les catégories nominales et verbales (Wlodarczyk A. & H. [1997]), nous avons élaboré deux échelles de valeurs opposées[7] (Tableau 1) qui précisent les notions du *donné* et du *nouveau*.

[6] Nous ne partageons pas l'idée « générativiste » selon laquelle les contenus sont ordonnés de façon à évoquer uniquement des relations de voisinage immédiat.

[7] Soulignons que ces deux échelles permettent de ne pas réduire la validité des énoncés linguistiques à l'opposition des deux seules valeurs du vrai et du faux de la logique.

information "donnée" S	5	Générique
	4	Général
	3	Habituel
	2	Potentiel
	1	Anaphorique
zone d'ambiguïté	0	*indécidable*
information "nouvelle" P	1	Cataphorique
	2	Actuel
	3	Occasionnel
	4	Particulier
	5	Spécifique

Tableau 1. La dualité graduée de la validité métainformative de l'énoncé

Dans (Wlodarczyk, A. [1999 et 2003a]), nous avons adopté une approche topologique car elle nous permet de désigner ces notions par les symboles ∑ et ∏ qui sont bien définis et ne sont pas encore chargés de connotations en linguistique. La dualité entre les espaces ∑ et ∏ a été utilisée pour la première fois en linguistique par Igarashi J. [1993] dans sa théorie de l'arrière-plan contextuel de l'aspect et de la quantification. Nous avons fait le rapprochement entre notre théorie de la validation des énoncés et la théorie d'Igarashi pour la raison suivante: Igarashi s'est aperçue que ce qu'elle avait découvert pour l'aspect pouvait également expliquer dans une certaine mesure les quantificateurs. Les symboles topologiques ∑ et ∏ — mettant en rapport les catégories nominales avec les catégories verbales — seraient donc des opérateurs plus abstraits que l'aspect ou les quantificateurs. Cette démarche est en fait complémentaire par rapport au couple traditionnel du *donné* et du *nouveau* car, malgré la rigueur mathématique de la topologie, cette dernière ne peut, à l'heure actuelle, être évoquée en linguistique qu'à titre métaphorique ou en se référant aux espaces subconceptuels de la cognition (cf. Gärdenfors P. [2000]). Notamment le caractère **dual** des espaces ∑ et ∏ reflète bien le type d'opposition entre nos deux échelles de valeurs, ce qui contribue réellement à une meilleure compréhension des relations entre les notions traditionnelles de *donné* et *nouveau*.

En ce qui concerne la définition formelle de la nature graduelle des domaines de validation ∑ et ∏, nous avons envisagé d'utiliser l'opération logique de *remplacement (*Wlodarczyk, A. [1999]) et évoqué[8)] la possibilité d'utiliser les concepts de la logique modale (le *nécessaire* et le *possible*) accompagnés de connecteurs temporels (*jusqu'à présent* et *dorénavant*) pour définir les deux valeurs opposées du *générique* et du *spécifique*.

Il est facile de s'apercevoir également que la binarité notionnelle que nous avons évoquée se trouve aussi à la base de la définition des *quantificateurs* logiques (*existentiel* et *universel*) mais il convient de souligner que, dans ce dernier cas, il s'agit de la détermination du nom et non pas de la modification du verbe. Cette remarque rejoint celle de [Geach P. T. 1979] selon qui la théorie platonicienne de la prédication (d'abord reprise puis rejetée par Aristote) établissait une relation hétérogène entre le sujet et le prédicat. Geach insiste sur le fait qu'Aristote a abandonné à tort ce point de vue. Le fait de renouer à notre tour avec l'idée du sujet et du prédicat en les pensant comme des objets hétérogènes nous permet de revoir également ce qui oppose la *détermination* (du nom) à la *modification* (du verbe) au sens large. Ainsi, pour analyser la validation des énoncés, il est nécessaire d'étudier les rapports inter-catégoriels: entre les verbes (aspect, modalité etc) et les noms (détermination, nombre etc) cf. [Wlodarczyk, A. & H., 1998]. Cependant, malgré l'hétérogénéité évoquée ci-dessus, certaines langues ont développé l'accord morphologique entre le sujet et le verbe. De plus, nous considérons que la relation Sujet-Prédicat fonde la *validation métainformative* (interprétation dans un contexte discursif et cognitif) sans laquelle un énoncé ne serait pas « complet ». Nous présentons plus loin les raisons pour lesquelles le rapport entre le sujet et le prédicat est en même temps de nature *syntaxique* que *parataxique*[9].

Notre cadre théorique nous permet de comprendre que les langues peuvent (1) présenter l'*information* sous forme d'énoncés non étendus soit comme *entièrement donnée*, le sujet et le prédicat appartenant à l'espace ∑, soit comme *entièrement nouvelle,* le sujet et le prédicat appartenant à l'espace ∏, ou (2) repérer tel ou tel *centre d'intérêt* comme appartenant à l'espace ∑ (topique) ou comme appartenant à l'espace ∏ (focus) en lui ajoutant un commentaire opposé, c-à-d dont la valeur *méta*informative est du type inverse

---

[8] Lors de notre communication au Xe Congrès Romano-Slave à Cracovie en septembre 2000.

[9] Nous proposons d'expliquer le sujet en adoptant la notion de *congruence* (Wlodarczyk A. 2003c). Ce concept, formalisé par Pogonowski J. (1991), permet à travers un ensemble d'isomorphismes locaux de rendre compte des aspects homogènes (parties accommodées) de la relation sujet-prédicat.

(énoncés étendus). Ainsi, les concepts de *topique* et *focus* que nous employons diffèrent considérablement des concepts de *thème* et *rhème* issus de la tradition pragoise en ce qu'ils sont définis comme « centres d'intérêt de l'énoncé étendu » et non simplement comme les parties donnée et nouvelle de l'énoncé.

4. Le sujet et l'objet comme 'centres d'intérêt' de l'énoncé non étendu

La syntaxe des langues résulte de la combinaison des opérations sémantiques, pragmatiques (ou protocolaires), locutives, cognitives etc. Ce que nous appelons « protocole » dicte le choix d'une stratégie de selection et de présentation des rôles participatifs d'un énoncé. Les segments de l'énoncé qui sont distingués (mis en valeur) par différents moyens linguistiques expriment ce que nous avons appelé *centres d'intérêt*. Dans l'énoncé non étendu, le sujet. n'est pas l'un quelconque des arguments du prédicat mais un argument *privilégié* dont dépend toute la construction de l'énoncé. La grammaire traditionnnelle distingue entre le sujet et l'objet sans pourtant proposer de définitions formelles de ces notions. Selon nous, le sujet[10] ne peut être défini que de façon complexe pour tenir compte de sa nature double: indiquer un participant de la situation sémantique et faire de ce participant le centre d'intérêt global de l'énoncé.

Nous avons proposé de recourir à la topologie (Wlodarczyk, 2003a) pour donner une représentation formelle de la notion de sujet. En topologie[11], l'espace est défini comme la paire (X, O) composée d'un ensemble X et d'un ensemble O des sous-ensembles de l'ensemble X. Le sujet peut être défini comme un espace *sélectionné* et mis en rapport avec un autre espace. Cette représentation formelle retient ce que contenait la conception traditionnelle du sujet en tant que « propos » ou « ce dont on parle » et du prédicat servant de formateur de phrase comme « ce qu'on dit sur le sujet ».

En logique, le prédicat est en effet conçu comme une relation généralisée car pouvant concerner plusieurs arguments (mais qui ne sont pas hiérarchisés les uns par rapport aux autres). Il est d'usage d'emprunter à la logique élémentaire (du 1er ordre) la définition du prédicat pour représenter (donc exprimer de façon métalinguistique) les relations linguistiques de tout genre (syntaxiques, sémantiques, pragmatiques etc.). Cependant, ce qui semble plus

[10] Pour une discussion plus approfondie sur la notion de sujet cf. [Wlodarczyk A. 2003a et 2003c].

[11] Structure définie sur un ensemble E par la donnée de parties de E appelées « ouverts » et satisfaisant à certains axiomes. («Le Petit Larousse»)

judicieux en sciences du langage, c'est de réserver la notion de prédicat au seul niveau protocolaire (stratégique) car (a) c'est ce niveau qui reflète les relations entre les divers *centres d'intérêt* et (b) cela est conforme au sens étymologique de la notion de prédicat. Notre proposition consiste donc à ajouter – pour décrire les langues humaines – un niveau métainformationnel où le sujet est défini comme une *métainformation* (c'est à dire : information concernant une autre information*)* distinguée au moyen d'un sélecteur extérieur à la *situation sémantique*[12] (cette dernière correspond à l'information). En aucun cas, nous ne devons utiliser les conventions de la logique formelle comme des schémas qui représenteraient directement les structures linguistiques.

Le fait de privilégier (= s'intéresser davantage à) l'un des participants de la *situation sémantique* crée une *métainformation*. Pour cette raison, ce que nous avons évoqué comme "sélecteur extérieur" (par rapport à la sémantique) exprime la relation entre les deux niveaux d'analyse : sémantique et pragmatique. L'expression simple (*énoncé non étendu*), en plus d'une information sémantique, contient ainsi un protocole métainformationnel qui lui est appliqué. Cette mise en relation des deux niveaux d'analyse permet la validation des énoncés par rapport à la situation de discours et aux connaissances. C'est notamment au moyen de cette validation (métainformationnelle) que l'information acquiert un statut «donné» ou «nouveau».

Quelles sont les situations sémantiques qui peuvent comporter plus d'un centre d'intérêt ? Il va sans dire que, selon que le verbe est intransitif ou transitif, le locuteur peut utiliser un nombre différent de groupes nominaux parmi lesquels il donne priorité à certains d'entre eux Pour exprimer ses choix protocolaires, tout en se conformant aux règles de la formation et de la composition, le locuteur construit des énoncés pouvant comporter un ou plusieurs centres d'intérêts suivant un ordre bien défini, sans toutefois que des permutations séquentielles soient totalement exclues.

Du fait de la linéarité fondamentale du discours, certains centres d'intérêts ont une portée globale (maximale), d'autres ont une portée locale (minimale). Dans la perspective topologique que nous avons adoptée, on doit s'attendre à ce que les centres d'intérêt puissent s'emboîter. Ainsi, en présence

---

[12] Il n'est pas impossible, pour représenter les *relations sémantiques* (comme toutes autres relations, d'ailleurs), d'utiliser les prédicats logiques (augmentés ou non des rôles de leurs arguments comme cela se pratique par exemple en logique attributive) ou d'autres notions appartenant à toute autre logique ; par exemple à la logique combinatoire. Mais il convient de garder à l'esprit que cette utilisation de la logique ne peut être considérée comme révélatrice des analogies directes (problème d'adéquation) entre les structures linguistiques et leurs modèles.

de deux centres d'intérêt, pour le deuxième d'entre eux, deux cas de figures sont possibles dans une énumération ordonnée: (a) soit il est plus global que l'autre et (b) soit il est plus local que l'autre. Il semble plus naturel que le global précède le local. En effet, les notions de SUJET et d'OBJET ne peuvent être définies que par rapport à la segmentation des énoncés car elles sont spécialement liées au caractère séquentiel (linéaire) du discours par opposition au parallélisme des représentations mentales. Pour cette raison, nos définitions de ces notions s'appuient sur les concepts des centres d'intérêt et de leurs interdépendances sans correspondance avec les rôles sémantiques.

Quant aux rapports entre les deux niveaux, le centre d'intérêt global peut sélectionner différents rôles sémantiques parmi lesquels l'agent et le patient sont certes les plus fréquents mais pas les seuls possibles. Ce sont les règles de la composition des énoncés qui sont « responsables » de l'ordre des syntagmes ainsi que des correspondances entre les centres d'intérêts (métainformations) et les diverses parties des situations sémantuiques qu'expriment les énoncés. L'exemple #1 montre le cas le plus fréquent où le contre d'intérêt global (« sujet ») correspond à l'agent sémantique et s'accorde avec le prédicat dans une construction active .

#(1) Les parents élèvent leurs enfants.

Une construction active peut aussi comporter un centre d'intérêt local qui est dans la position syntaxique de complément régi (patient ou autre rôle sémantique) mais tout objet n'est pas un centre d'intérêt local.

Quand la présentation de l'information met le patient en position de centre d'intérêt global nous avons une construction passive.

#(2) Les enfants sont élevés par leurs parents.

Les langues possèdent aussi des moyens lexicaux (converses lexicales, cf. Apresjan 1974) permettant de faire varier la présentation d'un rôle sémantique comme centre d'intérêt local ou global. Les converses lexicales sont des expressions sémantiquement proches permettant de représenter la situation de plusieurs manières. En voici des exemples en français :

#(3) Pierre prête de l'argent à Paul. ⇔ Paul emprunte de l'argent à Pierre.

#(4) Marie vend sa voiture à Jeanne. ⇔ Jeanne achète la voiture de Marie.

C'est la cohérence du texte qui décide du choix de l'une des diathèses ou de l'une des converses : par exemple, pour ce qui est des énoncés (3), si le texte a pour centre d'intérêt (thème principal) « Pierre », on choisira l'énoncé à verbe « prêter » qui permet de placer « Pierre » en position de sujet, centre d'intérêt global de l'énoncé non étendu.

5. Le topique et le focus dans l'énoncé étendu

le TOPIQUE et le FOCUS sont également définis comme des centres d'intérêt distingués dans l'énoncé mais, en outre, cette mise en valeur se réalise lorsqu'il y a un **contraste** : quand l'information contenue dans le centre d'intérêt appartient au domaine de l'opérateur ∑ et le commentaire appartient au domaine de l'opérateur ∏, nous avons affaire à un énoncé à TOPIQUE ; inversement, quand le centre d'intérêt appartient au domaine de l'opérateur ∏ et le commentaire au domaine de l'opérateur ∑, nous avons affaire à un énoncé à FOCUS.

T = S) & P : Sujet ∑ & prédicat ∏ **CONTR (S - P)**

#(5) (Il y a) le satellite X03 (qui) a été détruit.

(F = S) & P : Sujet ∏ & prédicat ∑ **CONTR (P - S)**

#(6) C'est le satellite X03 qui a été détruit.

L'opposition se manifeste dans les langues naturelles par le caractère – on pourrait dire – « généralisant » de l'opération de topicalisation et au contraire « particularisant » de la focalisation. Les expressions linguistiques marquant la topicalisation ou la focalisation prennent respectivement les valeurs des deux échelles S et P tout en appartenant à diverses catégories dans différentes langues (modalité, aspect, détermination). Dans les langues à articles, cela peut aller de pair avec l'emploi dans les syntagmes nominaux topicalisés de l'article défini en valeur anaphorique ou générique[13]:

(#7) Rivière avait écouté, avec ennui, parler de bilans, d'assurances, et surtout d'opinion publique : « **L'opinion publique** … ripostait-il, on la gouverne ! » (Saint-Exupéry, Vol de nuit)

Nous avons proposé (Wlodarczyk, A., 1979, 1999) de représenter les relations entre les parties de l'énoncé validées en ∑ ou ∏ par un carré sémiologique (Fig. 1).

[13] Ces valeurs sont décrites notamment par [Guillaume, G. 1975].

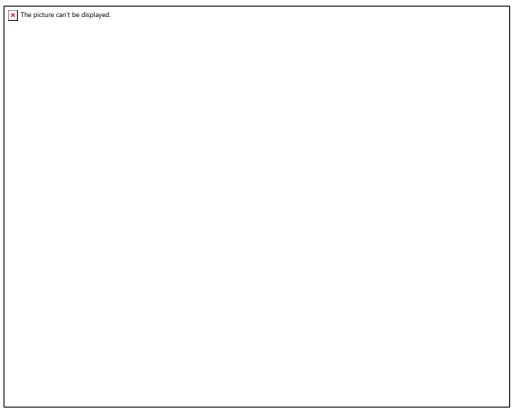


Figure 1. Carré sémiologique de la composition des énoncés

La figure 1 représente le cas des énoncés où le groupe nominal sujet est en même temps topique ou focus. Il va sans dire que d'autres constituants[14] de l'énoncé sont susceptibles d'être en position de topique ou de focus.

L'une des meilleures représentations formelles des centres d'intérêt de l'énoncé étendu (topique et focus) qui jette une lumière complémentaire sur ces notions semble être la métaphore des ensembles. Il s'agit notamment de l'indication de l'*identité* de l'entité choisie comme centre d'intérêt par rapport aux autres membres du même ensemble (quel que soit le statut de ces derniers : éléments ou sous-ensembles). Le topique permet de repérer un élément d'un ensemble par rapport aux autres éléments de cet ensemble qui partagent les mêmes propriétés en le mettant sur le même plan qu'eux, le focus au contraire sélectionne un élément pour l'opposer aux autres en insistant sur ce qui le distinguent des autres.

#(8) Jean et Marie dansent bien. Pierre, lui aussi, danse très bien.

(9) Jean et Marie dansent bien. Mais c'est Pierre qui danse le mieux.

Dans les exemples #(8) et (9), les relations ensemblistes éclairent l'interprétation des centres d'intérêt. En #(8), Pierre fait partie d l'ensemble des humains qui ont la capacité de danser. Dans le premier énoncé, il n'est pas question de lui. En décidant de dire quelque chose de pierre, le locuteur pose d'abord Pierre comme une information donnée du fait qu'il le compare aux deux autres danseurs,puis il introduit du nouveau en affirmant son excellente habileté à la danse. En #9, l'accent est mis sur Pierrecomme unique élément

[14] La topicalisation ou focalisation du groupe verbal (et non d'un groupe nominal) est cependant sans doute soumise à des conditions spécifiques qui mériteraient une étude séparée. Nos remarques sont fondées sur la topicalisation ou focalisation des groupes nominaux.

du même ensemble que Jean et Marie possédant à un aussi haut degré la propriété qui définit les autres éléments de l'ensemble.

#(10) Jean et Marie dansent bien. Quant à Pierre, il ne sait pas danser du tout.

En #10, du fait de la négation qui établit un contraste avec l'énoncé précédent, Pierre est identifié et choisi comme un être qui n'appartient pas à l'ensemble de ceux qui dansent bien : il est exclu de l'ensemble ou (dans une autre formulation) il est repéré comme appartenant à l'ensemble complémentaire.

On peut dire, de manière imagée, que le topique est une mise en relief diffuse, tandis que le focus est une mise en relief compacte.

6. La dualité homotopique entre le Topique et le Focus

C'est plus particulièrement le concept d'homotopie emprunté à la topologie que nous avons utilisé pour formaliser les notions de topique et focus (Wlodarczyk, A., 2003a). L'homotopie est une déformation continue des espaces ouverts qui s'opère dynamiquement par la projection d'un sous-espace sur un autre sous-espace. Le topique est défini comme la projection d'un point sur un ensemble, et par opposition, le focus consiste en la projection d'un ensemble sur un point (fig.3). Un espace ramené par homotopie à un seul point est dit contractile et est appelé *contraction*: le FOCUS peut être représenté par cet espace *contractile*. L'homotopie inverse ou *expansion* construit un espace qui définit le TOPIQUE. Les homotopies du TOPIQUE et du FOCUS représentent le mouvement de l'attention respectivement de l'un vers le multiple et du multiple vers l'un.

La notion d'homotopie permet de saisir l'aspect dynamique de la communication verbale, c'est-à-dire de formaliser l'intuition que le centre d'intérêt n'est pas un point ou espace fixe mais bien une opération construisant de façon dynamique deux types d'espaces inverses : l'un qui est le centre d'intérêt centrifuge (topique), l'autre qui est le centre d'intérêt centripète (focus). C'est en raison de ce caractère dynamique que nous utilisons non seulement les termes de topique et focus mais aussi ceux de topicalisation et focalisation pour désigner les opérations qui les produisent. En revanche, il convient d'insister sur le fait que nous les concevons comme des opérations cognitives qui s'appliquent à des représentations de la situation sémantique et non comme des « transformations » grammaticales qui s'appliqueraient à des énoncés linguistiques déjà construits.

De plus, expansion et contraction sont des déformations *fortes*, c-à-d. induisant des espaces qui résultent de la projection d'eux-mêmes sur eux-mêmes, l'un d'eux étant nécessairement un point. L'identité peut être conçue

comme un cas particulier d'homotopie, et notamment comme une déformation continue forte d'un espace, nous l'utiliserons comme propriété des centres d'intérêt de l'information. Donc, le concept d'homotopie paraît pertinent également parce que le topique et le focus jouent dans l'énoncé le rôle d'identificateurs : ils servent littéralement à identifier (et attirer l'attention sur) un centre d'intérêt dont il sera question dans la suite de l'énoncé (dans le commentaire).

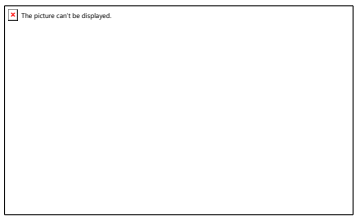


Fig. 3 Expansion et Contraction des espaces du Topique et du Focus

7. Entre la Syntaxe et la Parataxe

Les projections homotopiques permettent de représenter le fait que le choix d'un topique ou d'un focus ne peut s'expliquer uniquement par des opérations syntagmatiques mais doit faire appel à l'axe paradigmatique. Nous proposons ici le terme de *parataxe* dans une acception différente de celle qu'il a dans la syntaxe traditionnelle, entendant par parataxe tout ce qui – dans la chaîne parlée – se trouve à la fois dans une relation syntagmatique et en même temps renvoie à des éléments *in absentia*. On trouve parmi les marqueurs de la parataxe des adverbes paradigmatisants[15] tels que *aussi, même*. Un énoncé comportant l'un de ces adverbes ne peut être compris que si l'on rapporte l'élément *in praesentia* sur lequel porte l'adverbe en question, à un ensemble *in absentia* auquel il appartient et auquel le locuteur et l'allocutaire ont accès grâce à un savoir partagé.

Le sujet lui-même relève de la parataxe dans la mesure où il se définit à la fois dans sa relation syntagmatique au prédicat et en tant que centre d'intérêt global de l'énoncé non étendu. En tant que centre d'intérêt, il est la trace d'une opération de choix qui se produit (préalablement à la formulation de l'énoncé) à l'intérieur d'un ensemble possible d'autres centres d'intérêt que sont tous les participants de la situation sémantique considérée.

De même, le topique et le focus, tout en étant constituants d'une structure syntagmatique, font référence à un ensemble d'éléments non présents dans la chaîne. Selon la métapore ensembliste, le topique consiste à envisager

[15] Terme de Nølke,H., 1983.

l'élément qu'il désigne comme l'un des centres d'intérêt parmi d'autres possibles (envisagés comme un ensemble de référence). C'est ce que l'homotopie représente comme projection d'un point sur un espace (expansion).

#11) La mise en page des articles (parmi les autres tâches possibles telles que la rédaction, la traduction, la correction, etc) je m'en charge.

Le focus au contraire présente (selon la vision ensembliste) l'élément choisi comme le seul possible parmi un ensemble. Cela est représenté par l'homotopie de la contraction : projection d'un espace sur un point.

#(12) C'est la mise en page dont je me charge (et non les autres tâches telles que la rédaction, la traduction, la correction, etc).

L'énoncé qui comporte un topique ou un focus (ou les deux) est étendu à la fois parce que ces centres d'intérêt viennent s'ajouter ou se superposer à ceux que comporte déjà l'énoncé non étendu (sujet et objet) et à la fois parce qu'il oppose une partie d'information de type S à une partie d'information de type P□ou l'inverse. Ces propriétés sont marquées par des procédés syntaxiques adaptés pour le topique et le focus et différents selon les langues. L'autonomie du topique par rapport au reste de l'énoncé est assurée dans la syntaxe des langues par l'ordre des mots, l'intonation et, parfois, des morphèmes démarcatifs. En effet, le topique est le plus souvent projeté à l'avant de l'énoncé:

#(13) **Ce bouton** je l'ai trouvé par terre / Il est en nacre (Desnos, Corps et biens, Vieille clameur)

Mais il existe aussi, dans des langues diverses, la possibilité de rejeter le topique à l'arrière de l'énoncé[16] :

#(14) Elles existaient quelque part **ces terres en paix**, douces sous leurs grandes ombres de lune. (Saint-Exupéry, Vol de nuit)

Il peut même être introduit par une expression qui le soustrait à la structure syntaxique du reste de l'énoncé:

#(15) **Quant aux grosses bêtes**, je ne crains rien. J'ai mes griffes. (Saint-Exupéry, Le Petit Prince)

Le focus, quant à lui, est distingué du reste de l'énoncé par un accent fort (plus fort que celui des autres mots ou syntagmes du même énoncé) et l'intonation descendante. Dans de nombreuses langues, la présentation neutre

[16] On le nomme parfois antitopique (Lambrecht 1994).

de l'information consiste à mettre à l'avant de l'énoncé les éléments d'information ancienne et à garder pour la fin l'information nouvelle. La mise en valeur de l'information nouvelle comme focus peut être réalisée seulement par une intonation emphatique sans déplacement d'ordre, mais elle est souvent réalisée également grâce au déplacement de cette information vers l'avant de l'énoncé.

Dans la focalisation, l'intonation descendante doit accompagner le déplacement d'un constituant à l'avant de l'énoncé car l'intonation montante d'un constituant situé à l'avant de l'énoncé caractérise le topique. Dans les langues à déclinaison, le déplacement se réalise sans perdre l'information sur la relation syntaxique et sémantique de l'élément déplacé par rapport au verbe. Nous prenons l'allemand comme exemple de ce type de langues.

#(16) Wahrlich, **mit andern Augen**, meine Brüder, werde ich mir dann meine Verlorenen suchen, **mit einer andern Liebe** werde ich euch dann lieben.

En vérité, c'est avec d'autres yeux que je chercherai alors mes amis perdus, c'est d'un autre amour que je vous aimerai. (Nietzsche, Also sprach Zarathustra, Bilingue Aubier Flammarion 1969, trad. G. Blanquis, p.182-183)

Certes, dans l'énoncé original #(16), le groupe nominal focalisé est en même temps un groupe prépositionnel dans lequel la préposition concourt avec les désinences casuelles à indiquer la relation au groupe verbal. La traduction française, cependant, ne peut se contenter d'un simple déplacement vers l'avant même avec la préposition et doit recourir à la clivée.

Même dans les langues qui marquent les relations syntaxiques par des morphèmes casuels ou des particules agglutinantes, le focus peut être marqué non seulement par sa place et un accent fort mais aussi souligné par une particule, un adverbe ou une construction clivée.

#(17) Aber **gerade dem Helden** ist das Schöne aller Dinge Schwerstes…

Mais c'est précisément au héros qu'il est le plus difficile d'atteindre au Beau… (ibid, p..256-257)

Dans les langues pour lesquelles l'ordre des mots sert à indiquer les relations syntaxiques entre le verbe et les groupes nominaux (français, anglais etc.), la mise à l'avant d'un constituant focalisé (même s'il s'agit d'un syntagme nominal introduit par une préposition) s'accompagne d'une construction clivée comme on l'observe dans les traductions de Nietzsche ci-dessus #(16) et #(17) ou dans l'exemple suivant :

#(18) Je le sauve de la peur[…] Si je l'écoute, si je le plains, si je prends au sérieux son aventure, il croira revenir d'un pays de mystère, et **c'est du mystère** seul que l'on a peur. (Saint-Exupéry, Vol de nuit)

Le caractère « paradigmatisant » et « contrastif » du focus est bien illustré par l'énoncé (18) : le mystère est opposé implicitement à d'autres causes possibles de la peur et il est présenté comme la partie nouvelle de l'information (type P) opposée à la partie ancienne « on a peur » (de type S). Cet énoncé montre aussi que « nouveau » ne veut pas dire introduit pour la première fois dans le discours car le mot « mystère » est introduit comme déterminant de « pays » immédiatement avant d'être introduit comme focus par la clivée. Ce qui est nouveau, ce n'est pas le mystère en tant que concept mais en tant que cause de la peur. Il en va ainsi dans tout énoncé : un élément de l'énoncé n'est pas donné ou nouveau de manière toujours identique dans l'univers des connaissances mais de manière variable en fonction du rôle qu'il joue en un point précis du texte.

Il a été remarqué que les langues peuvent utiliser des énoncés comportant à la fois un topique et un focus. Dans la plupart des langues que nous connaissons, la linéarisation semble imposer de placer le topique avant le focus quand on a affaire dans un même énoncé à un topique antéposé et à un focus antéposé.

#(16) Nous voyons au point de départ une expérience très particulière et en somme assez grossière ; au point d'arrivée, une loi tout à fait générale, tout à fait précise, et dont nous regardons la certitude comme absolue. **Cette certitude, c'est nous qui** la lui avons conférée pour ainsi dire librement, en la regardant comme une convention. (Henri Poincaré, La Science et l'hypothèse)

#(17) Ie, o-ryori wa watakushi ga shimasu kara… (Non, puisque la cuisine, c'est moi qui la fais...)

Cette mise en ordre linéaire du topique avant le focus pourrait s'expliquer par le fait que l'ordre linéaire place d'abord ce qui est global (le topique) avant ce qui est local (le focus).

Cependant, cet ordre peut être modifié par le fait de rejeter le topique en fin d'énoncé :

#(18) Non, puisque c'est moi qui la fais, la cuisine…

Dès qu'on a ffaire à un topique postposé, il apparaît non seulement après le focus mais après l'énoncé tout entier.

6. Conclusion

L'ordre dans lequel un locuteur choisit de communiquer une information dépend de la partie de cette information qu'il identifie comme centre d'intérêt. Cet ordre consiste à ajouter linéairement un commentaire à la suite d'un centre d'intérêt souvent placé au début, mais aussi à d'autres endroits de l'énoncé (ce qui complique évidemment les choses pour les linguistes). La complexité des ordres communicatifs dans les langues humaines est due à ce qu'il existe plusieurs centres d'intérêt différents (détachés ou non du reste de l'énoncé, de même nature métainformative que lui ou non). D'autre part, certaines langues, dites à syntaxe positionnelle, utilisent l'ordre linéaire non seulement pour la présentation de l'information (métainformation - pragmatique) mais aussi pour représenter les relations sémantiques d'une situation (information - sémantique).

La linguistique traditionnelle a mis en évidence un nombre de correspondances plus ou moins fréquentes entre le sens et la forme. Ainsi, par exemple, le fait que le Sujet corresponde le plus souvent à l'Agent ou à l'Expérient. Le caractère hybride de tels faits (traces des processus parallèles dont ils sont issus) est généralement attribué au manque de rigueur scientifique de la linguistique traditionnelle par les tenants de la linguistique structurale, soucieux de séparer ce qui relève de différents niveaux du langage.

Faut-il cependant rappeler que l'idée saussurienne du Langage défini comme « système » n'est en fait qu'une hypothèse. Malgré une quarantaine d'années de recherches en traduction automatique et en TAL dans le monde (USA, URSS puis Russie, Japon, UE etc.), il n'a toujours pas été possible de décrire une langue de manière "systématique". Tout porte à croire que, en plus de la cohérence constatée, les langues présentent des fragments où "tout ne se tient pas" aussi bien qu'on l'espérait. En effet, les structures syntaxiques (de même que les structures morpho-syntaxiques, y compris la valence des verbes) sont de nature hétérogène. Par exemple, tout comme il serait vain de chercher à déterminer le rôle sémantique qui correspondrait au sujet des prédications statiques («Jean est grand »), il ne serait pas très constructif de poser l'existence de rôles sémantiques (en position de Sujet) dans certains types de prédication où l'agentivité pose problème («Les prix montent »). Pour échapper à la contradiction, il suffit d'admettre que certains sujets (tout comme les sujets des prédications statiques ou certains centres d'intérêt de niveau supérieur: topique ou focus) ne correspondent strictement à aucun rôle sémantique.

Pour conclure, l'étude des centres d'intérêt doit tenir compte de l'imbrication des faits linguistiques (de différente nature: phonologique, morphologique, syntaxique, sémantique et pragmatique) de telle sorte que leur présence dans l'expression linéaire ne peut être expliquée que par la

reconnaissance de structures hétérogènes. Autrement dit, afin de saisir la régularité des langues (« cachée » par la linéarité du discours), il est impérieux de garder constamment en mémoire la nature « symbolique » des unités linguistiques. Cela implique qu'il est nécessaire de dépasser l'inventorisation des faits observés sans aucune représentation formelle. Ferdinand de Saussure n'a-t-il pas insisté sur le fait que le langage n'est pas une « substance » mais une « forme » ?

7. Bibliographie

APRESJAN, Jurij D. (1974) *Leksiceskaja semantyka. Sinonimiceskije sredstva jazyka*, Moskva.

GEACH, P. T. (1971) "Names and Predicables", in «Semiotics in Poland 1894-1969», PWN – Polish Scientific Publishers, Warsaw, p. 240-246.

GÄRDENFORS, P. (2000) Conceptual Spaces – The Geometry of Thought, The M.I.T. Press, Cambridge, Massachussets, London, England, p. 307.

GROSZ, B. J., JOSHI A.K., WEINSTEIN S. (1995) "Centering: a framework for Modeling the Local Coherence of Discourse", *Computational Linguistics*, Vol. 21, Nr 2, 1995, p. 203-226.

GUILLAUME, Gustave (1975) *Le problème de l'article et sa solution dans la langue française,* réédition, Nizet, Les Presses de l'Université Laval, Québec.

IGARASHI, Jitsuko (1993) *Haikei-shori moderu ni okeru bun-asupekuto no ∑∏□hyôgen (∑∏-Expressions of Sentential Aspects in a Contextual Language Processing Model)*, Researches of the Electrotechnical Laboratory (ETL), 950, Tsukuba, p. 93.

LAMBRECHT, Knut (1994) *Information Structure and Sentence Form. Topic, focus, and the mental representations of the discourse referents*, Cambridge University Press, Cambridge, p. 388.

MATHESIUS, Vilhelm (1939) „O tak zvanem aktualnim cleneni v vetnem" (A propos de ce qu'on appelle la division actuelle de la phrase », *Slovo a slovesnost* 5, p. 171-175.

MIZUTANI, Shizuo (1991) *Description Systématique de la Grammaire Japonaise*, trad. du japonais par Reiko Shimamori et André Wlodarczyk, dans «Travaux de linguistique japonaise», Vol. IX, Université de Paris VII, Paris, p. 198

NØLKE, Henning (1983) "Les adverbes paradigmatisants : fonction et analyse", *Revue Romane* Numéro Spécial 23. Copenhague : Akademisk Forlag

SOWA, John F. (2000) *Knowledge Representation – Logical, Philosophical, and Computational Foundations*, Brookes/cole, USA.

POGONOWSKI, Jerzy (1991) *Hierarchiczne analizy języka* (Analyses hiérarchiques du langage), en polonais, Wyd. Naukowe UAM, Seria Językozawstwo Nr 14 Poznań, (p. 64)

WLODARCZYK, André (1979) "La linguistique et ses applications aux études japonaises en France", Actes du colloque scientifique franco-japonais (Paris - 1979), Paris (réimprimé dans *Le Japon vu depuis la France*, p. 195-208, Maison Franco-Japonaise, Tokyo - 1981).

(1996) “La Thématisation en japonais et l'article français”, in *Semantyka a konfrontacja językowa* (Sémantique et Linguistique Contrastive), red. Koseska-Toszewa V. & Rytel-Kuc D., Editions Slawistyczny Ośrodek Wydawniczy (SOW), Varsovie, p. 131-140.

(1998) “The Proper Treatment of the Japanese “wa” and “ga” Particles, *Proceedings of the International Workshop on Human Interface Technology 1998* (IWHIT ‘98) - Aizu-Wakamatsu, Japon, p. 63-70.

(1999) “La validation informative des énoncés et la quantification linguistique”, *Etudes cognitives / Studia kognitywne III*, SOW, PAN, Warszawa, p. 121-133.

(2001) « On Identifying 'Centres of Interest' in the Japanese Discourse », *International conference, Japanese Phenomenon: views from Europe*, Russian Association of Japanologists, Japan Foundation, Moscow, p. 90-105.

(2003a) “Les Homotopies du topique et du focus”, *Ordre et distinction dans la langue et le discours*, Actes ducolloque international de Metz (18,18 et 20 mars 1999, publiés par Bernard Combettes, Catherine Schnedecker, Anne Theissen, Paris, Honoré Champion éditeur, p.513-526.

(2003b) “Linguistique dynamique: évolution du discours dans le temps”, *Etudes linguistiques romano-slaves offertes à Stanisław Karolak*, Oficyna Wydawnicza Edukacja, Cracovie 2003, p. 497-510.

(2003c) “Generalising the Theory of *wa* and *ga* Particles (An Attempt of the Formal Reconstruction of a Functional Approach) ”, in *Paris Lectures on Japanese Linguistics*, invited for the New Millenium and edited by Wlodarczyk A., Kurosio Shuppan, Tokyo.

WLODARCZYK, André & WLODARCZYK, Hélène (1998) “Graded Informative Content of Linguistic Messages”, CD-ROM: *Proceedings of the 16th International Congress of Linguists.* Pergamon, Oxford, Paper N° 0431