

Mémento sur quelques rapports entre mémoire et linguistique

Dominique LEGALLOIS

Université de Caen Basse-Normandie

CRISCO, MODESCO

dominique.legallois@unicaen.fr

S'il y a bien un thème sur lequel des disciplines fort diverses peuvent se rencontrer, c'est bien celui de la mémoire, tant la notion participe à la structuration des sciences humaines et sociales, constitue l'objet d'analyses constantes en sciences de la nature, et se donne comme défi pour les sciences « techniques » telle que l'informatique. Le prix fort à payer pour une telle constance est évidemment une constellation polyphonique de définitions, de conceptions, qui tendent à rendre moins assurés les efforts de synthèse.

La mémoire est une fonction ou une faculté, celle de *conserver et de rappeler des états de conscience passés et ce qui s'y trouve associé*, nous dit Le Robert. Sans doute est-ce là la signification minimale qu'il faut avoir à l'esprit, pour aborder les deux champs dans lesquels la mémoire joue un rôle vital de structuration : l'individuel et le collectif. La mémoire individuelle est un objet biologique et psychologique, à la fois « répertoire » des traces mnésiques encodées et parcours : jusqu'à la restitution, en passant par la rétention et la consolidation. Ce même parcours vaut pour la mémoire collective, mais dans un lieu qui n'est plus le cerveau mais le social. Les récits de vie, les témoignages, les lieux de commémoration forment quelques-uns des genres discursifs ou des espaces symboliques dans lesquels la mémoire collective se constitue. La distinction entre mémoire individuelle et mémoire collective peut être vue comme le reflet d'une frontière épistémologique entre les disciplines que sont, d'une part, les sciences du vivant, et, de l'autre, les sciences de l'homme et les sciences sociales. D'un côté, principalement, la neuro-psycho-biologie, de l'autre, principalement, l'histoire et la sociologie. Cette distinction lexicalisée entre deux types de mémoire est manifestement légitime au niveau des processus en jeu : on ne saurait confondre circuits neuronaux et constitution culturelle. Cependant, elle ne tient pas au niveau du contenu : d'une part, il n'y a pas de souvenirs qui ne renverraient pas à du social, d'autre part, la médiation du langage – qu'il est difficile de mesurer dans l'encodage ou dans la restitution, mais qui est patente – fait que même la mémoire individuelle est sociale. Le grand théoricien de la mémoire collective, Maurice Halbwachs, allait beaucoup plus loin lorsqu'il attribuait aux dérèglements mnésiques ou intellectuels de l'individu, des causes éminemment sociales¹ : il

1. « Ce sont les autres qui font que les souvenirs nous reviennent » (M. Halbwachs, 1994 / 1925, V-VI).

s'agissait bien, dans cette perspective, et sur ce point, de faire de la science du vivant, une annexe de la science sociale.

Ce que nous voudrions indiquer ici, c'est que les effets de la distinction entre mémoire individuelle et mémoire collective sont particulièrement vivants en sciences du langage – et là sans doute plus qu'ailleurs. Pourtant, nous choisirons pour notre présentation un autre mode de division : non pas, la mémoire individuelle *vs* la mémoire collective (pour la simple raison que la démarcation entre les deux niveaux nécessiterait des justifications sans fin), mais une division selon les objets affectés par la mémoire, *i.e.* les événements discursifs, le contexte, les formes linguistiques assurant la cohésion, la compréhension des textes, le lexique et la grammaire.

Nous présentons donc, dans cet article, différents aspects des rapports qu'entretient la mémoire avec certains objets de la linguistique. Nous verrons que ces rapports ne sont pas homogènes, mais au contraire, riches de la diversité même de la discipline. Aussi, notre présentation sera nécessairement fragmentaire, incomplète, déséquilibrée et sûrement partielle. Certains points seront largement développés, alors que d'autres ne seront qu'évoqués. Ainsi, la première partie du travail - la mention du rôle de la mémoire individuelle dans l'élaboration des discours – sera-t-elle passée en revue, plutôt que discutée. En revanche, l'exposition de la notion de mémoire des textes sera illustrée par un type de travail particulier sur le rôle structurant de la répétition lexicale *dans* et *à travers* les textes. Mais surtout, nous nous concentrerons par la suite exclusivement sur l'efficacité de la mémoire lexico-grammaticale des sujets dans ses rapports avec la phraséologie – ce qui nous permettra d'étendre considérablement, en vertu de la corrélation mémoire/phraséologie, le spectre phraséologique dans la compétence – performance des locuteurs. Cette dernière partie trouvera une validité dans les travaux de la psycholinguiste britannique A. Wray, qui propose une nouvelle démarcation, entre la part de la créativité dans le langage et la part de la mémorisation.

Mémoire discursive des individus

Nous mentionnons dans cette partie, des travaux ayant pour objet direct ou indirect la mémoire individuelle des sujets. Trois domaines sont très succinctement passés en revue : 1) le rapport mémoire/conscience (la restitution dans un discours d'événements passés) ; 2) le rapport contexte/mémoire (pour le choix des formes anaphoriques) ; 3) le rapport mémoire/compréhension du texte (les mécanismes psycho-cognitifs).

Mémoire des événements

La mémoire n'est pas un objet linguistique. Il est cependant possible, à partir de certains énoncés, de constituer le type de mémoire en jeu dans la réminiscence des événements. Ce travail a été mené par l'un des linguistes les plus originaux de

son époque : W. Chafe. Très brièvement, Chafe (1973) s'intéresse au poids des adverbiaux temporels dans les phrases. Par exemple ²,

1 – *Steve fell in the SWIMMING pool*

2 – *Steve fell in the SWIMMING pool a couple minutes ago*

3 – *A couple MINUTES ago, Steve fell in the swimming pool*

sont les indices d'une remémoration d'un événement enregistré dans la mémoire à court terme (*surface memory*).

4 – *Steve fell in the SWIMMING pool yesterday*

5 – *YESTERDAY, Steve fell in the swimming pool*

sont les indices d'une remémoration d'un événement enregistré dans la mémoire à moyen terme (*shallow memory*).

6 – *Last CHRISTMAS, Steve fell in the SWIMMING pool*

est l'indice d'une remémoration d'un événement enregistré dans la mémoire à long terme (*deep memory*).

Bien que ces remarques paraissent au premier abord triviales, cette caractérisation (beaucoup plus complexe que ce que nous présentons ici) sert à analyser le rôle que joue la conscience dans la distribution de l'information et de sa mise en forme, en particulier selon ce que le locuteur présume de l'état des connaissances de l'interlocuteur. Ainsi, le choix des adverbes est déterminé en fonction de l'inscription de l'événement dans la mémoire. Un événement très ancien peut ainsi être exprimé dans un énoncé sans adverbe temporel, si, pour le locuteur cet événement reste particulièrement présent dans sa conscience. Chafe donne l'exemple réel et pathologique d'un dépressif qui énonce à des inconnus

7 – *my daughter died*

(donc, sans mention d'adverbe temporel) alors que la mort de sa fille remonte à plusieurs années. Il s'agit d'un témoignage extrême de la persistance, chez ce patient, d'une rétention continue. Autre exemple, moins tragique mais révélateur : lorsque quelqu'un prononce la phrase-cliché

8 – *je m'en souviens comme si c'était hier*

« *he evidently means that something from deep memory exhibits properties associated with shallow memory* » (Chafe, 1973 : 270).

Chafe analyse donc, à partir d'exemples très simples, les traces linguistiques de la constitution phénoménologique de la mémoire et ses rapports avec la conscience. Ce travail, que l'auteur poursuivra dans son étude passionnante *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing* (1994) constitue une voie encore peu empruntée en linguistique. Même si le corpus de Chafe est essentiellement composé de textes oraux, ce travail pourrait donner une assise formelle solide aux études littéraires ou philosophiques, généralement plus sensibles à cette thématique ³.

2. Les lettres capitales signalent les plus hauts degrés intonatifs.

3. Cf. par exemple L. Bougault (2007). Cf également, dans une perspective d'analyse de discours : S. Delvenne, C. Michaux et M. Dominicy (2005).

Les rapports entre mémoire et contexte

Dans un article de synthèse, G. Kleiber (1994) examine l'usage de la notion de contexte en linguistique. Dans une approche standard du contexte, celui-ci est conçu comme statique, venant soit préciser (enrichir le substrat sémantique) ou filtrer (sélectionner la bonne interprétation) le contenu linguistique de l'énoncé. En ce sens, la mémoire reste implicite dans la mesure où le type de linguistique en jeu suppose une conception non cognitive du contexte. Avec l'avènement des approches cognitives, le contexte est devenu dynamique, et en quelque sorte interne. Il n'est pas une réalité hors locuteur, mais une réalité cognitive, qui permet de repenser la nature objectivable du sens en termes d'intersubjectivité :

Contexte linguistique, situation extra-linguistique, connaissances générales se retrouvent tous traités mémoriellement : ils ont tous le statut de représentation interne, même s'ils se différencient quant à l'origine et au niveau de la représentation (mémoire courte, mémoire longue, etc.). Ainsi, pour prendre un exemple, le cas des déictiques est traité différemment entre une approche standard et une approche cognitive. Pour preuve, des sujets typiquement contextuels comme l'anaphore se retrouvent redéfinis et traités en des termes impliquant crucialement la mémoire. Une expression déictique ne sera plus non plus définie par le site de son référent, mais par le fait qu'elle introduit une entité non encore mémorisée dans le modèle contextuel. (Kleiber, 1994, 19).

On parle par exemple de *mémoire discursive* (Béguelin, 1988, Berrendonner, 1990) pour rendre compte, lors d'un échange, de l'accroissement progressif des connaissances partagées par les interlocuteurs. Cet accroissement/modification de la mémoire discursive par représentations schématiques successives, a des conséquences directes sur les types d'anaphores employées. Si l'antécédent est identifiable, c'est parce qu'il est présent dans la mémoire discursive.

La mémoire discursive, entendue dans cette acception, correspond à la fois à la mémoire de travail et à une mémoire plus étendue stockant l'interprétation des énoncés immédiatement précédents. Ce type de considérations indique déjà qu'il est difficile ici de parler encore de mémoire individuelle, puisqu'il s'agit bien, en raison du phénomène social qu'est le discours, de partager des connaissances et de construire en commun un espace discursif.

La mémoire dans la compréhension des textes

Une branche de la psychologie cognitive, depuis les travaux pionniers de F. Bartlett (1932) livre une abondante littérature sur les relations entre mémoire et compréhension des textes. Comprendre un discours oblige à construire, pour le sujet, des représentations cognitives de la signification du texte, représentations que la mémoire de travail structure. Mais en retour, se souvenir du texte dépend fortement de la constitution et des propriétés de ces représentations. Exposer les différents modèles qui traitent de ces problèmes dépasse nos compétences. Aussi pour une contribution récente et originale, nous renvoyons aux travaux de Denhière et Lemaire (2004) qui

complètent les théories de Kintsch (1998) par l'utilisation de la Sémantique Latente développée par Landauer *et al.* (1998) : des corpus représentatifs des connaissances d'une population (par exemple les enfants de 8-11 ans) sont transformés en espaces sémantiques permettant la représentation des unités linguistiques (mots, phrases, paragraphes, etc.). On peut ainsi comparer des productions textuelles aux résultats des simulations supposées représenter les connaissances en mémoire et les processus mis en jeu dans la compréhension et la production.

De la mémoire du texte à celle des sujets : l'intertexte

C'est bien sûr dans l'intertextualité que se mesure traditionnellement – depuis la découverte des travaux de Bakhtine – la mémoire des textes. L'intertextualité, soit l'ensemble des formulations récurrentes, l'ensemble des discours antérieurs, le déjà-dit. Le texte est conçu comme un lieu où transitent explicitement ou allusivement d'autres textes. Aussi la relation intertextuelle est-elle vue comme un réseau mémoriel des discours passés, même si on ne saurait lier trop étroitement intertextualité et mémoire de textes, puisque la relation peut être orientée vers les textes en devenir. Ainsi, le versant « pragmatique » de l'intertextualité, le dialogisme, met-il l'accent sur le phénomène d'anticipation, de responsivité, donc sur une orientation inverse.

De la mémoire intertextuelle des textes à la mémoire intertextuelle des locuteurs : l'hypothèse des réseaux lexicaux

La mémoire discursive des textes ne saurait être confondue avec la mémoire discursive des locuteurs. La mémoire discursive des textes est une propriété linguistique virtuelle, non cognitivement enracinée chez un sujet, que l'on peut stimuler par l'analyse lexicale. Considérons une remarque de J.-J. Courtine (lui-même se référant à Foucault) :

Pour Foucault [...] l'énoncé se définit, entre autres propriétés, par le fait de posséder un « domaine associé ». Celui-ci comprend notamment les formulations « auxquelles l'énoncé se réfère (implicitement ou non), soit pour les répéter, soit pour les modifier ou les adapter, soit pour s'y opposer, soit pour en parler à son tour ». On pouvait dès lors rapporter tout énoncé à un domaine de mémoire : il y figurait comme un élément dans une série, comme un « nœud dans un réseau ». On pouvait espérer alors combiner l'analyse linguistique de l'énonciation singulière, située et datée, d'une formation discursive, avec la profondeur historique d'un système de formations des énoncés ; tenter d'inscrire l'événement énonciatif sur le fond de la mémoire discursive, démêler le temps court et le temps long de l'espace des discursivités. (J.-J. Courtine, 1981 : 26-27).

Courtine proposera donc la notion de *mémoire discursive* (dans un sens évidemment différent de celui de Béguelin ou Berrendonner) pour désigner le rapport d'un texte à d'autres textes.

Dans deux de nos articles (D. Legallois 2004, 2006), nous avons montré que l'idée selon laquelle l'énoncé figure comme un élément dans une série, comme un nœud dans un réseau était « matériellement » justifiée. En effet, si on veut identifier les traces de cette mémoire discursive, l'observation de la composition lexicale des textes est fondamentale. En nous fondant sur le principe du linguiste anglais M. Hoey (1991) – principe (P) selon lequel

P : lorsqu'une phrase d'un texte écrit non narratif partage au moins trois co-occurrences lexicales avec d'autres phrases du même texte (quel que soit l'empan entre ces phrases), la suite constituée par ces phrases est cohérente, c'est-à-dire interprétable dans le contexte développé par le texte.

nous avons mis en évidence qu'un texte non narratif⁴ est composé d'une organisation linéaire évidente, mais aussi d'une organisation non linéaire, c'est-à-dire réticulaire. Autrement dit, le texte est constitué par des réseaux de phrases repérés par le partage d'items lexicaux. Soit le texte suivant :

1 Au mois de novembre 1990, le président George Bush affirmait que, dans un délai de six mois, l'Irak pouvait se doter de l'arme nucléaire. 2 Jusque-là les prévisions de la CIA étaient de l'ordre de cinq ans... 3 Cette menace contribua à mobiliser l'opinion internationale en faveur de la guerre. 4 Quelques mois plus tard, une mission des Nations unies dirigée par un Américain, M. David Kay, saisissait 45 000 pages de documents concernant le programme nucléaire du président Saddam Hussein : elle affirmait à son tour que Bagdad était seulement à douze ou dix-huit mois de la production de la bombe. 5 Troublante coïncidence, on apprenait au même moment que M. Kay, oubliant ses obligations de fonctionnaire de l'ONU, avait, durant son séjour en Irak, transmis directement des informations à son gouvernement.

6 Sans doute, dans le cadre du nouvel ordre international, ne percevait-il plus de différence entre Nations unies et États-Unis... 7 La presse de Bagdad, aux ordres bien évidemment, dénonçait M. Kay comme un « agent de la CIA ».

8 Aujourd'hui, c'est sur la base de ces mêmes 45 000 pages qu'une réunion secrète d'experts américains, russes, français et britanniques conclut dans un sens totalement différent : à la veille de la guerre, il fallait trois ans au moins et sans doute plus pour que l'Irak acquière l'arme nucléaire. 9 Le président Bush cherchait sans doute, par ses déclarations, à dissimuler ses propres responsabilités dans la construction de la machine de guerre irakienne.

10 Au printemps 1989, des experts avaient alerté l'administration américaine sur le programme nucléaire irakien, sans succès : l'alliance entre les deux pays était trop solide.

11 Dès le printemps 1982, l'administration Reagan - M. George Bush était alors vice-président – décidait de transmettre des informations militaires à M. Saddam Hussein, notamment par le canal de la CIA : les responsables américains se rangeaient aux côtés de Bagdad dans sa guerre d'agression contre l'Iran. 12 Les preuves de collusion sont innombrables et, depuis quelques semaines, un nouveau scandale ébranle Washington – on le nomme déjà « Iraqgate ». 13 Deux ans au moins avant le 2 août 1990, l'Irak

4. L'organisation des textes narratifs s'appuie sur une autre logique : la succession d'événements.

avait commencé à détourner l'aide alimentaire américaine pour acheter des armes, et M. George Bush et ses conseillers auraient tout fait (avant la crise du Koweït) pour le cacher à l'opinion et au Congrès.

14 Cynisme? Incompétence? Mauvais calculs?

15 Ces faits éclairent d'un jour nouveau les propos du président Bush sur la morale, le nouvel ordre international et la capacité des États-Unis à assumer la direction de la planète.

M. Bush, la bombe et le dictateur de Bagdad (Le monde diplomatique, juillet 1992, p. 23) Alain GRESH.

Il est possible, en suivant la propriété P, d'identifier les réseaux lexicaux constitutifs du texte :

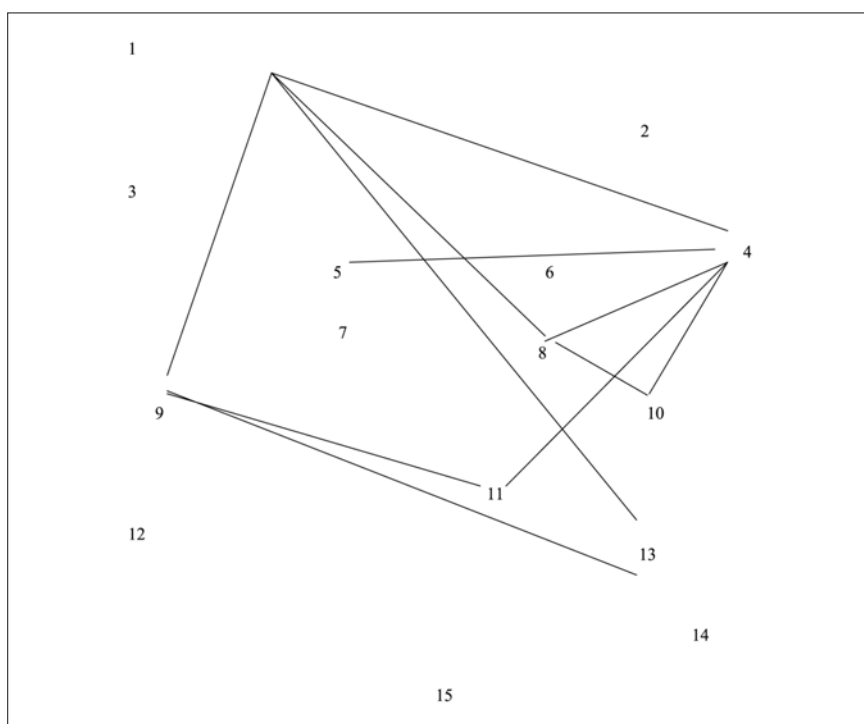


Fig. 1 : Réseaux lexicaux du texte *M. Bush, la bombe et le dictateur de Bagdad*

Le partage des occurrences est l'effet de la répétition lexicale. Il faut noter que la répétition se réalise par des « relations lexicales » diverses : répétition à l'identique, reprise anaphorique, relation synonymique ou antonymique. On peut par exemple donner le réseau suivant, constitué d'une suite cohérente de phrases non adjacentes :

1 Au mois de novembre 1990, le président George Bush affirmait que, dans un délai de six mois, l'Irak pouvait se doter de l'arme nucléaire

4 Quelques mois plus tard, une mission des Nations unies dirigée par un Américain, M. David Kay, saisissait 45000 pages de documents concernant le programme nucléaire

du président Saddam Hussein : elle affirmait à son tour que Bagdad était seulement à douze ou dix-huit mois de la production de la bombe

10 Au printemps 1989, des experts avaient alerté l'administration américaine sur le programme nucléaire irakien, sans succès : l'alliance entre les deux pays était trop solide

Ce réseau, parfaitement interprétable, forme un parcours dans le texte.

La propriété P s'applique aux simples textes. Mais le même principe a été appliqué à un niveau intertextuel, lorsque deux ou plusieurs textes appartiennent à un domaine associé. En ce cas, des appariements entre phrases de textes différents sont démontrés⁵, que ces textes soient en relations de synchronie (D. Legallois 2004, 2006) ou de diachronie (M. Hoey, 1995). Ces réseaux fictifs, virtuels (dans le sens où ils ne constituent pas des textes réels, mais des possibilités de textes), cette *hypertextualité externe*, serait le lieu possible d'une mémoire discursive des textes. Ainsi, à partir de deux articles journalistiques traitant de la même affaire, mais parus dans deux journaux différents (*L'Humanité* (H) du Lundi 12 janvier 2004 et *Libération* (L) – même date), il est possible d'appliquer le principe P pour apparier des phrases exogènes. Un exemple, parmi plusieurs possibilités :

H3-H4-L8

H3 Six personnes, interpellées mardi dernier à Vénissieux et à Auxerre et soupçonnées d'avoir tenté de préparer, en 2002, un attentat contre l'ambassade de Russie à Paris, sont présentées aujourd'hui à Paris au juge antiterroriste Jean-Louis Bruguière.

H4 Il s'agit, en premier lieu, de Chellali Benchellali, imam salafiste d'un lieu de prière sur le quartier de la Darnaise à Vénissieux, de sa femme et de son fils Hafed.

L8 Selon *le Monde* de dimanche, un des gardés à vue entendu cette semaine aurait confirmé que Ménad [BENCHELLALI] avait l'intention de fabriquer des produits « *hautement toxiques* » comme la « *ricine et l'acide botulique* ». Ici, la référence au journal *Le Monde* par *Libération*, ajoute, évidemment, à l'intertextualité.

Nous sommes sans doute dans un lieu limite : une mémoire des textes indépendante des locuteurs, en dehors d'une mémoire collective constituée. Cependant, puisque le principe P a une validité organisationnelle au niveau d'un seul texte, et puisqu'il s'applique également à l'intertexte, on pourrait le considérer comme ayant une validité au niveau de la mémoire discursive des locuteurs. Autrement dit, sans que nous soyons avertis du travail de la répétition et de la cohésion lexicale dans et au-delà du texte, nous sommes les réceptacles de paroles entremêlées, connectées. On pourrait ajouter que l'hypertexte externe, de nature virtuelle, existe grâce au réseau lexical interne du texte : ce sont en effet les phrases centrales d'un texte qui sont susceptibles d'être apparées à des phrases centrales d'autres textes.

L'intertextualité et la mémoire des textes ne se décrètent pas ; elles doivent être repérées par l'agencement lexical : chose que nous avons essayé de montrer dans

5. Nous ne pouvons présenter ici l'analyse intertextuelle entreprise dans nos articles. Nous renvoyons le lecteur à nos travaux et à ceux de M. Hoey pour un aperçu du partage de réseaux lexicaux entre textes différents.

cette partie. Ainsi, il existerait des virtualités d'appariements entre phrases de textes différents dont il « suffirait » aux locuteurs d'activer les nœuds pour élaborer une mémoire des connaissances discursives.

Mémoire lexico-grammaticale des locuteurs

Nous nous intéressons maintenant, et beaucoup plus longuement, à la place accordée à la mémoire des sujets dans l'élaboration des énoncés.

Règles combinatoires et lexique mental

Les théories syntaxiques ont problématisé les modes de génération des énoncés ; les locuteurs sont capables d'engendrer et de comprendre un nombre potentiellement infini de phrases différentes. Aussi, ce sont les processus de créativité qui ont retenu l'attention – la créativité étant de toute façon plus valorisée dans notre culture occidentale que la répétition. De ce fait, les grammaires qui ont intéressé les linguistiques sont celles qui s'inscrivent dans un paradigme rationaliste et non empirique, les données empiriques n'étant pas vues comme témoignant de l'essence même du langage, à savoir la production d'énoncés inédits. La tradition linguistique définit une langue comme un ensemble de mots (un lexique), et un ensemble de règles permettant la combinaison de ces mots en phrases. La perspective est donc générative : les énoncés sont construits/générés par des opérations élémentaires prescrites par des règles de combinaison. Ces règles ne sont pas redevables à des routines mémorisées, mais sont en quelque sorte implémentées dans la capacité cognitive des individus. Elles ne doivent donc rien à la mémoire, qui du reste est peu théorisée, et ne semble pas intervenir de façon déterminante dans les descriptions du fonctionnement modulaire des processus linguistiques. Ni le modèle fodorien de la modularité, ni les travaux de S. Pinker (1999) n'accordent à la mémoire un rôle essentiel.

Est considéré comme relevant de la mémoire, par contre, ce qui échappe aux opérations computationnelles : essentiellement les unités lexicales, qui ne peuvent évidemment être recréées à chaque moment langagier, et qui sont donc mémorisées dans le lexique mental pour être restituées en production. Il est en effet impossible de dire que nous produisons des mots de la même façon que nous produisons des énoncés : l'unité langagière minimale est bien sûr l'énoncé et non le mot. Nous sommes en partie responsables de la forme de nos énoncés mais en aucune façon de celle des mots du lexique. Il y a donc une relation préalable entre un mot et sa signification :

Fondamentalement, connaître un mot de la langue, c'est connaître la relation existante entre une forme et une signification. C'est savoir que la séquence sonore « table » réfère au meuble sur lequel je suis en train d'écrire ce texte tandis que la forme sonore « chaise » réfère au meuble sur lequel je suis assis pour écrire ce texte. Comment je sais cela ? Simple, parce que j'ai appris cette relation lors de l'acquisition de ma langue et que, d'une manière ou d'une autre, elle est condensée dans ma mémoire (J. Segui et L. Ferrand, 2000 : 24).

La production langagière, telle qu'elle est considérée par la *linguistique compositionnelle*⁶, suppose donc un lexique mémorisé et une connaissance implicite de règles de génération, qui ne relèvent pas de la mémoire, mais de la compétence linguistique des individus.

Le lexique mental est alors conçu comme un répertoire accueillant toute forme d'irrégularités : le lexique, mais aussi ce qui échappe à la rationalité du système – à savoir la phraséologie. Dans la perspective compositionnelle, les expressions idiomatiques ou phraséologiques, même si on les considère nombreuses, constituent des unités quelque peu à part, à isoler dans le répertoire du lexique mental. Ainsi, deux énoncés comme

9 – Paul a cassé sa pipe

et

10 – Paul a cassé le vases

ont, en ce qui concerne le prédicat, analysés de façon différente. Dans le second énoncé, le prédicat est le résultat d'une combinaison libre, alors que dans le premier, il est directement convoqué. Ainsi, « casser sa pipe » sera considéré comme un élément du lexique mental.

Phraséologie

Par ailleurs, c'est-à-dire essentiellement dans le domaine lexicologique, l'étude de la phraséologie, des expressions figées a connu un immense succès ; les articles et ouvrages sur ce sujet se sont multipliés, en même temps que la mesure du phénomène phraséologique devenait plus précis. Pourtant, d'une place importante, la phraséologie prend une place centrale. En effet, depuis une vingtaine d'années, la linguistique a entamé une révolution grâce au recours à l'informatique. La numérisation des textes et leur exploration par des outils et logiciels ont aidé le linguiste à élargir considérablement l'ensemble de ses observables, et plus encore, de s'affranchir des dangers de l'introspection. Ainsi, la linguistique de corpus est devenu une méthode d'investigation incontournable pour approcher ce qui relève de l'usage linguistique, déplaçant ainsi l'intérêt de la compétence pour privilégier celui de la performance. Ce sont alors, parmi la masse phénoménale des données, des régularités insoupçonnées qui sont apparues, des régularités d'une nature particulière : des régularités autonomes des règles *a priori* de constitution (de production). Autrement dit, se constitue peu à peu une linguistique échappant au rationalisme du cognitivisme orthodoxe et s'inscrivant dans l'empirisme de l'usage. Il est possible de montrer à partir d'un exemple simple, qu'un type d'énoncés difficilement catégorisable *a priori* comme phraséologique, est bel et bien un type d'unités non compositionnelles.

6. Tel est le terme que nous employons pour désigner l'ensemble des théories qui considèrent l'énoncé généré par des règles de combinaison.

Nous avons observé dans une année intégrale de *Libération* (1995), que la construction «GN ÊTRE DE SAVOIR sub. Interrogative indirecte» constitue indiscutablement une unité semi-figée.

En effet, 7,57 % des compléments infinitifs de la structure «GN ÊTRE DE INFINITIF»⁷ par exemple :

II – Notre objectif est de décrire les particularités de maturation des régions auditives corticales dans l'autisme (Internet)

se construisent avec *savoir* (113 occurrences) loin devant les autres verbes (*faire* 70, *mettre* 23, *créer* 20), *avoir* (85) et *être* (75) étant employés la plupart du temps comme auxiliaires. 7,57 peut sembler un pourcentage faible, mais au regard des autres co-occurents, *savoir* se détache sans discussion du peloton des infinitifs. Cet écart forme une prégnance tout à fait réelle et sensible à la suite de laquelle il est légitime de conclure à la «fixation» dans la compétence langagière de *N est de Savoir* + SUB. INTERROGATIVE. Qui plus est, on relève 64 «*question est de savoir*» (soit 56 % des emplois de «*est de savoir...*») et 90,14 % des emplois de *question* dans la construction), et 23 «*problème est de savoir*» (20, 35 % des emplois de «*est de savoir...*») et 31, 5 % des emplois de *problème* dans la construction). On considérera donc, au regard du corpus, que la *question est de savoir* + sub. interr. indirecte et le *problème est de savoir* + sub. interr. indirecte sont des unités semi-phraséologiques, admettant la variance (le GN peut par exemple être expansé, la subordonnée prendre des formes différentes).

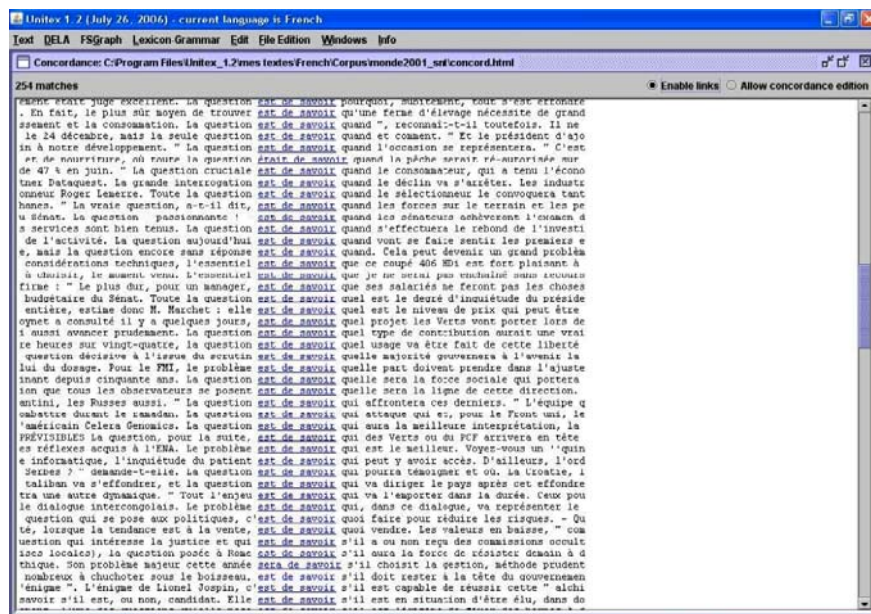


Fig. 2 : concordancier de «GN est de savoir»

7. Structure dite «spécificionnelle» qui se réalise, par exemple, dans : *l'essentiel est de participer, l'objectif est de réaliser des bénéfices, le problème est d'arriver à comprendre le raisonnement*, etc.

Dans notre article (Legallois et Gréa 2007), nous allons même plus loin en posant que la structure « GN ÊTRE DE INFINITIF » (donc, sans détermination lexicale) est elle-même une construction non compositionnelle, phraséologique, un tout préformé – à un niveau plus abstrait que celui de l'emploi avec *savoir*.

Il est évident que seule la fréquentation des corpus peut mettre en évidence la présence de telles constructions holistiques.

L'un des initiateurs de la linguistique de corpus a été le britannique John Sinclair. C'est à partir d'une observation fondamentale de cet auteur que nous avancerons les arguments en faveur d'une extension du domaine phraséologique à des unités autres que lexicales. Dans son livre de 1987, Sinclair dégageait deux grands principes :

the open-choice principle (principe du libre choix) : *this is a way of seeing language as the result of a very large number of complex choices. At each point where a unit is completed (a word or a phrase or a clause), a large range of choice opens up and the only restraint is grammaticality... Virtually all grammars are constructed on the open-choice principle* (Sinclair 1991 : 109-110).

the principle of idiom (principe d'idiomaticité) : *the principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments.* (Sinclair, 1991, 110).

Pour Sinclair, un des deux principes domine l'autre – ou plutôt est prioritaire par rapport à l'autre : le principe d'idiomaticité. Nous pouvons illustrer la portée du principe d'idiomaticité à partir d'un exemple : dans notre lecture du livre de Raymond Guérin, *Les Poulpes*, la phrase suivante – que nous restituons de mémoire – nous a arrêté :

12 – Il continuait à faire son numéro personnel devant les pontes du ministère

La pression du principe d'idiomaticité (ici sous une forme de collocation) fait que nous avons interprété pendant quelques secondes *numéro personnel* par *numéro de téléphone personnel*, avant de nous corriger. Ce petit exemple montre la « supériorité » du principe d'idiomaticité sur celui du libre choix. Autre exemple, emprunté à G. Kleiber (1994 : 10), ce titre du journal *L'Alsace* du 14/11/1992 :

13- La bonne partie du Finistère

La collocation *bonne partie* induit une interprétation qui ne correspond pas à l'intention de l'énonciateur puisque celle-ci était de désigner, par une description définie, *Bécassine*. Les jeux de mots ne sont pas sans exploiter ce principe qui montre que des unités polylexicales telles que les collocations sont interprétées à un coût cognitif inférieur à celui des unités combinées librement. Ce principe d'idiomaticité peut être étendu au-delà des seules collocations ou constructions figées ; tel était d'ailleurs l'objectif de Sinclair, qui allait dans le sens de B. Pottier quand celui-ci déclarait que

La mémoire doit être réintroduite à tout moment, et son rôle mis en évidence. Elle sous-tend les lexies (par cœur, charbon de bois, se la couler douce), les modules (qq'un donner qqch à qq'un), tout le phénomène de compréhension (la construction du sens chez le récepteur), etc. (B. Pottier, 1992 : 17).

Les modules de Pottier, qu'il nomme ailleurs *syntaxies*, sont des configurations syntaxiques, lexicalement sous-déterminées. Elles correspondent aux *constructions* de la *Construction Grammar* (cf. Fillmore et al. (1988), Goldberg (1995)) ou aux *Patterns* (cf. Hunston et Francis, 2000) de la linguistique contextualiste britannique (dans laquelle s'inscrit Sinclair)⁸. L'originalité de cette perspective est de considérer que ces configurations syntaxiques forment des tous, enregistrés par la mémoire comme le seraient des unités lexicales ou phraséologiques. Ainsi, pour reprendre l'exemple de B. Pottier, [qq'un donner qqch à qq'un] est un schème cohésif holistiquement produit et interprété. Il n'y aurait pas à proprement parler de combinaison de groupes, de sélection par le verbe d'arguments, mais directement instantiation de la syntaxie.

Toutes ces remarques vont dans le sens d'un article trop peu cité, celui de M.L. Moreau (1986) sur les « séquences préformées ». A partir d'un corpus oral, l'auteur montrait statistiques à l'appui, que des formes comme *i faut pas X, c'était pas X, il y a X, j'en sais rien*, etc. :

bien que susceptibles d'une analyse linguistique, ne devraient pas être identifiées comme des combinaisons du point de vue psycholinguistique : ces séquences fonctionneraient comme des entités globales et pourraient être stockées comme telles dans notre lexique interne (M.-L. Moreau 1986 : 154).

Les conséquences de ces perspectives sont importantes : la fonction de la mémoire linguistique aurait une importance beaucoup plus prononcée qu'on ne le croit généralement. Elle ne servirait pas seulement à stocker des mots, ou des collocations ou encore des expressions idiomatiques, mais aussi des combinaisons complexes de mots que l'on ne qualifierait pas, à première vue, d'unités figées. Plus encore, il semble possible d'étendre le champ de la mémoire à des constructions syntaxiques, lexicalement peu ou pas déterminées. Une partie des énoncés que nous « produisons » serait en fait remémorés. D. Bolinger avait déjà remarqué ce fait :

Speakers do at least as much remembering as they do putting together
(D. Bolinger, 1975 : 2).

L'idée que la remémoration participe à la production des énoncés, au même titre que le processus de génération, oblige à une réévaluation des qualités mémorielles souvent sous-estimées dans le paradigme chomskien ou traditionnel⁹. L'étude

8. Pour une présentation détaillée de ces perspectives, cf. D. Legallois (2005) et D. Legallois et J. François (2006).

9. C'est l'objectif de psychologues tels que N. Ellis (2005) ou P. Perruchet (2006) qui étudient le rôle de l'apprentissage implicite dans l'acquisition du langage, et l'importance des données mémorisées non consciemment.

de l'oral apporte des arguments significatifs pour la détermination du poids de la mémoire et des unités phraséologiques. B. Erman (2007) montre que la fréquence et la durée des pauses à l'oral informe de la taille des unités linguistiques mémorisées dans le lexique mental, puisque les pauses sont les indices de l'effort cognitif de convocation du lexique.

Nous terminons ce parcours par la présentation d'un modèle psycholinguistique récent, qui rend compte des propos tenus plus haut.

La phraséologie et le lexique mental : le modèle d'A. Wray

La psycholinguiste britannique Alison Wray a récemment proposé de nouvelles perspectives dans la description du lexique mental. Nous n'exposerons pas de façon exhaustive son modèle, mais nous indiquerons les points importants pour l'objet qui est le nôtre.

Si, comme nous l'avons vu, une des tendances en linguistique est de reconsidérer la place de la phraséologie dans la production langagière, il est logique que les psycholinguistes tiennent compte de cette nouvelle donne, pour éprouver ou reformuler les modèles cognitifs sous-jacents à l'activité linguistique. Ainsi, en proposant une lecture documentée de l'état de l'art, et en se fondant sur ses propres travaux, A. Wray énonce dans un premier temps une définition de ce qu'elle nomme *langage formulaire*, et qui englobe tous les faits phraséologiques :

a sequence, continuous or discontinuous of words or other elements, which is, or appears to be prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar (A. Wray, 2002 : 9)

Dans un deuxième temps, l'auteur recense, à partir de travaux antérieurs, les types d'emploi du langage formulaire. Nous adaptons ci-dessous le tableau synthétique donné par A. Wray.

Fonctions	Effets	Types	Exemples
Changer l'environnement physique et perceptif	Satisfaire des besoins physiques, émotionnels et cognitifs.	1 Injonction	<i>Pelouse interdite</i>
		2 Requêtes etc	<i>Je suis vraiment désolé</i>
		3 Excuses	
		4 Marqueurs de politesse	<i>Pourriez vous, s'il vous plaît,...</i>
Affirmer une identité propre	Être pris au sérieux	1 Raconter une histoire	<i>Tu ne croiras jamais ce qui m'est arrivé</i>
		2 Prendre la parole	<i>La première chose que je dirai...</i>
	3 Gestion des tours de parole	<i>À mon tour, maintenant, de te dire quelque chose</i>	

Affirmer son identité par rapport au groupe	Appartenir à un groupe	1 Chants 2 Formules rituelles	<i>On est les champions!</i> <i>Très bonne année à toi</i> <i>Nous sommes réunis aujourd'hui pour célébrer...</i> <i>À ta place, je ne ferai pas ça</i> <i>«je ne voudrais appartenir à aucun club qui aurait moi comme membre» (Groucho Marx)</i> <i>M. le Président...</i> <i>Je ne suis pas tout à fait certain que...</i> <i>Je te promets que...</i> <i>Je touche du bois</i> <i>Je vous salue Marie</i>
	Se placer dans la hiérarchie	3 Menace 4 Citation	
	Rituel	5 Formules d'adresse 6 Atténuation 7 Performatifs 8 Incantations 9 Prières	

Tab. 1 : Les emplois du langage formulaire selon Wray (2002 : 89)

À partir de ce tableau, il est possible de distinguer un ensemble de fonctions liées au langage formulaire (L.F.) :

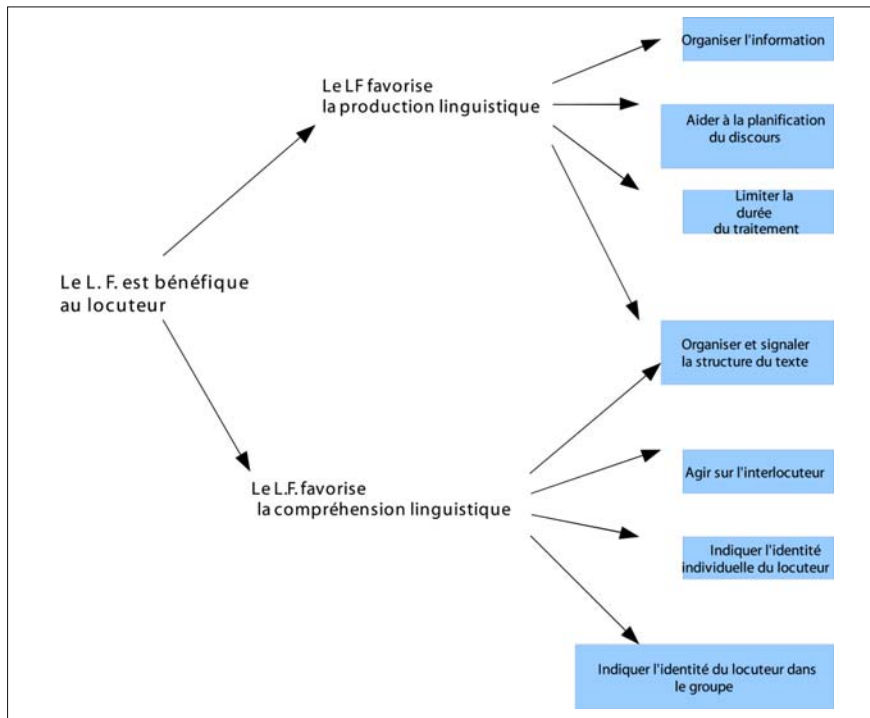


Fig. 3 : Les fonctions du langage formulaire (L.F.) d'après Wray (2002 : 97)

Reprenons succinctement les emplois du langage formulaire: avoir un accès rapide à l'information (formules mnémotechniques, textes appris par cœur, etc.), exprimer l'information avec fluidité, avoir des besoins physiques et émotionnels satisfaits rapidement, acquérir une information demandée, être compris facilement, être perçu comme un membre du groupe, etc. Le langage formulaire, selon A. Wray, réaliserait un but essentiel: la promotion de l'intérêt personnel du locuteur. Il s'agirait de la motivation fondamentale pour l'émergence et l'emploi des séquences phraséologiques, des routines discursives. Mais cette motivation fondamentale, à la fois psychologique et anthropologique, doit, pour le linguiste, être examinée dans la réalité langagière des locuteurs. Ainsi, son efficience peut être appréciée dans les parcours – les options – qui se posent aux sujets parlant. Le schéma suivant résume ces parcours, dont on peut identifier trois types: le locuteur fait l'usage d'énoncés « libres », c'est-à-dire non contraints – seules s'appliquent les règles combinatoires; le locuteur emploie des énoncés redevables autant au langage formulaire semi-figé qu'au langage produit par combinaison; le locuteur emploie des constructions figées et routinisées.

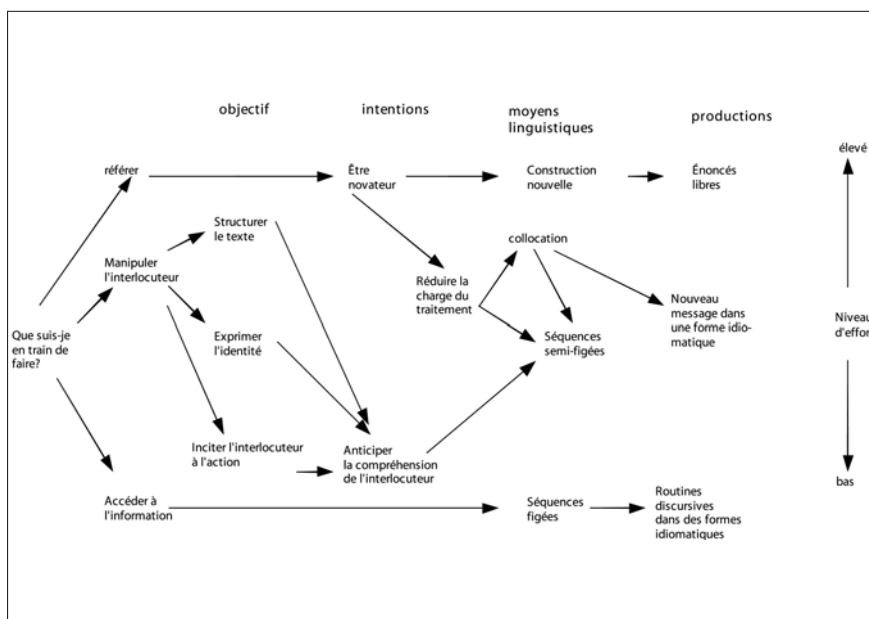


Fig. 4: Parcours productifs des énoncés selon A. Wray (2002 : 86)

Sans que nous puissions ici reproduire l'argumentation de l'auteur, c'est une reconsidération du lexique mental qui est alors proposée. Le lexique serait en fait hétérogène, et de ce fait, constitué de cinq grandes classes que mettent en évidence les travaux sur l'acquisition de la langue maternelle, d'une langue étrangère, ou encore sur la plasticité cérébrale des aphasiques qui compensent la perte d'un lexique par un autre lexique.

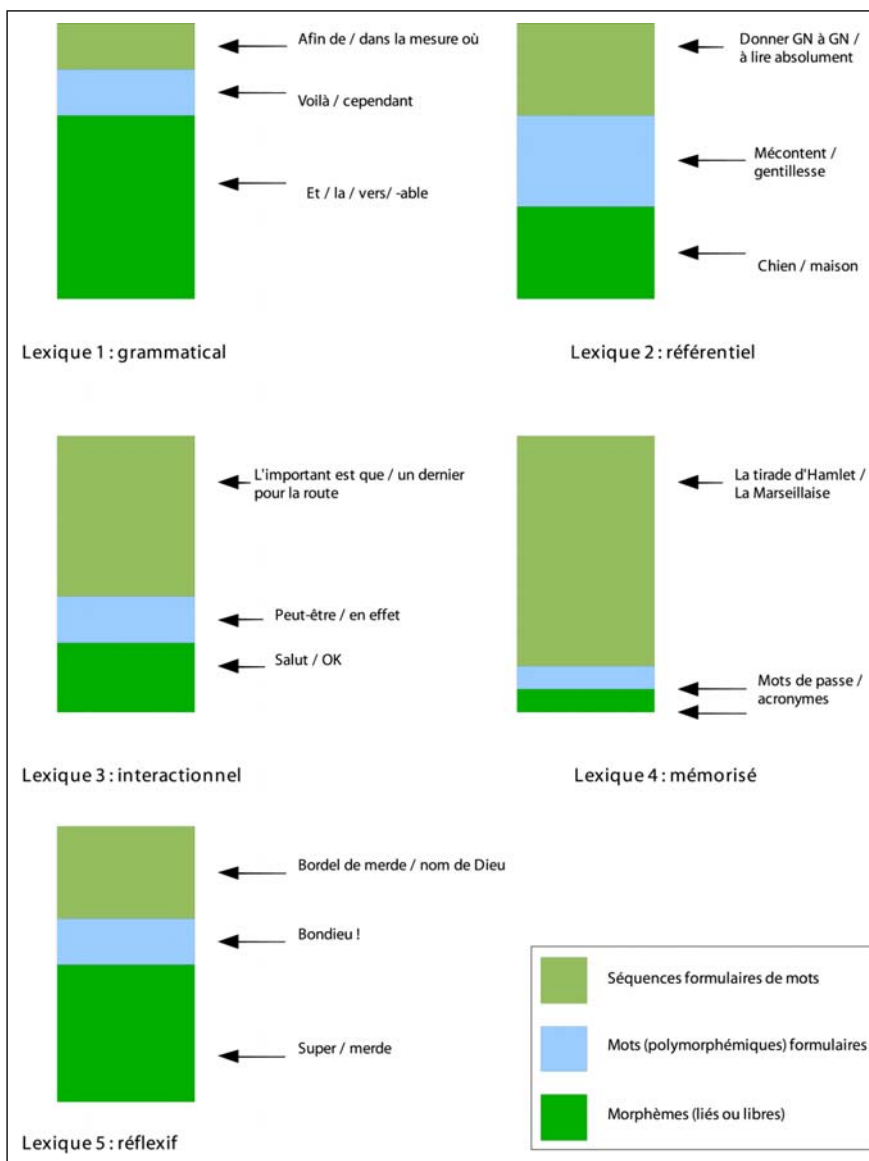


Fig. 5: Les cinq lexiques selon A. Wray (2002 : 263)

Pour l'auteur, la division en cinq lexiques n'est pas une facilité de représentation de l'hétérogénéité du lexique général. Au contraire, il s'agit de plusieurs lexiques, d'ailleurs localisés dans des hémisphères différents : les colonnes 1 et 2 sont associées à l'hémisphère gauche ; les colonnes 3 et 4 à l'hémisphère droite, et la colonne 5 au sous-cortex, région de l'émotion.

Le lexique grammatical (lexique 1) est nécessaire à la construction de phrases nouvelles, à la créativité linguistique, dans la mesure où il comprend le matériel grammatical. Le lexique référentiel (lexique 2) est celui des expressions référentielles (mono ou polymorphémiques), mais aussi des collocations. Le lexique interactionnel (lexique 3) est « spécialisé » dans le langage formulaire à l'œuvre dans les interactions. Le lexique mémorisé (lexique 4) est celui dans lequel sont enregistrées les productions linguistiques « apprises par cœur » (poèmes, chansons, prières, formules mnémotechniques, etc.). Enfin, le lexique réflexif – on pourrait dire *émotionnel* (lexique 5) recouvre les formules exprimées spontanément lors d'une réaction émotionnelle (jurons, exclamations, etc.). Chaque lexique est lui-même hétérogène quant à la nature des unités enregistrées : séquences polylexicales, mots morphologiquement construits, morphèmes liés ou libres. On notera que le lexique dit référentiel est particulièrement intéressant car il recouvre, outre les collocations et les mots lexicaux, des *constructions*, des *syntaxies* au sens de B. Pottier. Par exemple, le schéma (GN [verbe de donation] GN à GN).

Par ailleurs, les énoncés libres, mobilisant principalement le lexique 1, sont produits à un plus grand coût cognitif. Le lexique 5 est celui qui livre les expressions produites au plus faible coût. De plus, toute entité dans une colonne peut être manipulée indépendamment, mais seulement de façon holistique. Cependant, les composants peuvent être mémorisés séparément. Par exemple, « good morning » est enregistré dans la colonne 3 ; mais « good » et « morning » sont enregistrés de façon indépendante dans la colonne 2. De ce fait, l'enregistrement est holistique, car il est inutile pour le sujet de segmenter les mots s'il n'a aucune raison de le faire. Enfin, le modèle de A. Wray prévoit une certaine forme de récursivité : des items de différentes colonnes peuvent contribuer à l'expression d'une phrase nouvelle ; par exemple pour :

14- Could we just consider your bookshelf for a moment? *could we just consider DET N for a moment* appartient au lexique interactionnel. C'est une forme semi-figée. Mais le nom *bookshelf* appartient, lui, à la colonne 2 ; le mot outil *your* au lexique grammatical (colonne 1). Il y a donc interaction entre ces lexiques.

Quelle que soit la validité neuropsychologique que l'on accorde à ce modèle, il illustre parfaitement les résultats linguistiques venant parfois de traditions différentes : la grammaire de construction de Fillmore et de Goldberg, la linguistique fondée sur l'usage (usage-based) de Bybee (J. Bybee, 2007) la linguistique contextualiste de Sinclair. Toutes ces perspectives (et d'autres) plaident non seulement pour un continuum entre expressions figées et expressions libres, mais surtout pour la promotion des expressions holistiques (mots, énoncés, syntaxies) à une place fondamentale dans la production et la compréhension du langage. Ainsi, en considérant qu'une partie importante des structures d'énoncés est en partie lexicalisée, on conviendra qu'elle est stockée en mémoire dans le lexique mental, au même titre que les lexèmes et les grammèmes.

Conclusion

Dans ce parcours partial et partiel, nous avons essayé de proposer un *memento* rendant compte d'une part, de la place « habituelle » de la mémoire dans certaines préoccupations (psycho) linguistiques : compréhension des textes, accès au lexique mental ; d'autre part, d'une place plus originale : mémoire des événements, constitution de réseaux lexicaux, extension du domaine phraséologique. Nous pensons que les développements menés à la fois en linguistique de corpus – les corpus, *via* la mémoire des textes, permettent une modélisation de la mémoire des sujets – et en psychologie cognitive permettront dans un avenir proche d'appréhender mieux encore le phénomène mémoriel, en rééquilibrant le rapport entre ce qu'il est convenu d'appeler *l'esprit computationnel* et *l'esprit concordancier*. Si le propos du linguiste reste pauvre quant au fonctionnement effectif de la mémoire, il nous paraît en revanche tout à fait important pour rendre compte de son rôle.

Bibliographie

- BARTLETT F. C. (1932), *Remembering: A Study in Experimental and Social Psychology*, Cambridge, Cambridge University Press.
- BEGUELIN M.-J. (1988), « Anaphore, cataphore et mémoire discursive ». *Pratiques*, 57, p. 15-43.
- BERRENDONNER A. (1990), « Pour une macro-syntaxe », *Travaux de linguistique*, 21.
- BOLINGER D. (1976) « Meaning memory », *Forum Linguisticum*, vol. 1, n° 1, p. 1-14.
- BOUGAULT L. (2007), Le verbe *voir* dans le *Ravissement de Lol V. Stein* : construction d'une trace mnésique entre souvenir et fantasme, *Questions de Style* : « écriture et mémoire », Laure Himy-Piéri (éd.), <http://www.unicaen.fr/services/puc/revues/thl/questionsdestyle/>
- BYBEE J. (2006), « From usage to grammar : the mind's response to repetition », *Language* 82, p. 711-733.
- CHAFE W. (1973), « Language and Memory », *Language*, 49, p. 261-281.
- CHAFE W. (1994), *Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing*. Chicago, University of Chicago Press.
- COURTINE J.-J. (1981), « Analyse du discours politique : le discours communiste adressé aux chrétiens » *Langages*, 62, p. 5-125.
- DELVENNE S., MICHAUX C. et DOMINICY M. (2005), « Catégoriser l'impensable : la figure du « musulman » dans les témoignages de rescapés des camps nazis », *Oralités*, unité de recherche « sources audiovisuelles en histoire contemporaine », Université libre de Bruxelles.
- DENHIÈRE G., LEMAIRE, B. (2004), « A Computational Model of Children's Semantic Memory », *In Proceedings of the 26th Annual Meeting of the Cognitive Science Society* (CogSci'2004), p. 297-302.

- ELLIS, N. C. (2005), « At the interface : Dynamic interactions of explicit and implicit language knowledge », *Studies in Second Language Acquisition*, 27, p. 305-352.
- ERMAN B. (2007), « Cognitive processes as evidence of the idiom principle », *International Journal of Corpus Linguistics*, 12 : 1, p. 25-53.
- FILLMORE CJ, KAY P. et O'CONNOR M. C., (1988), « Regularity and idiomaticity in grammatical constructions : the case of let alone », *Language*, 64, p. 501-38.
- FOUCAULT M. (1969), *L'archéologie du savoir*, Paris, Gallimard.
- GOLDBERG A.E., (1995), *Constructions : A Construction Grammar Approach to Argument Structure*, Chicago, University of Chicago Press.
- HALBWACHS, M. (1994), *Les cadres sociaux de la mémoire*, Paris, Albin Michel.
- HOEY M. (1991), *Patterns of lexis in text*, Oxford, Oxford University Press.
- HOEY M. (1995), « The lexical nature of intertextuality : a preliminary study » in *Organisation in discourse*, Proceedings from the Turku Conference, Warvik, B. S-K. Tanskanen R. Hiltunen (eds), *Anglicana Turkuensia*, 14, p. 73-94.
- HUSTON S., FRANCIS G. (2000), *Pattern Grammar, a corpus-driven approach to the lexical grammar of English*, Amsterdam/Philadelphia, John Benjamins.
- KAY P. et FILLMORE CJ., (1999), « Grammatical constructions and linguistic generalizations : The What's X doing Y? construction », *Language*, 75, p. 1-33.
- KINTSCH, W. (1998), *Comprehension! : A paradigm for cognition*. Cambridge, Cambridge University Press.
- KLEIBER, G. (1994), « Contexte, interprétation et mémoire approche standard vs approche cognitive », *Langue Française*, 103, p. 9-22.
- LANDAUER Thomas K., FOLTZ Peter W., LAHAM Darrell (1998), « An introduction to Latent Semantic Analysis », *Discourse Processes*, 25, p. 259-284.
- LEGALLOIS D. (2004), « Cohésion lexicale et réseaux phrastiques dans la construction du texte expositif » in *L'unité texte*, Actes du colloque international de Chypre, mars 2004, S. Porhiel & D. Klinger (éds), Association perspectives.
- LEGALLOIS D., (2005), « Du bon usage des expressions idiomatiques dans l'argumentation de deux modèles anglo-saxons : la Grammaire de Construction et la Grammaire des Patterns », *Cahiers de l'Institut Linguistique de Louvain*, 31, 2-4, p. 109-127.
- LEGALLOIS D. (2006) « L'hypertextualité comme constitution de la mémoire discursive », *Pratiques*, 129-130, p. 139-156.
- LEGALLOIS D. & FRANÇOIS J., (2006), « Autour des grammaires de construction et de patterns », *Cahier du CRISCO*, 21 (prépublication accessible en ligne <http://elsapi.unicaen.fr/cahiers/>).
- MOREAU, M.-L. (1986), « Les séquences préformées : entre les combinaisons libres et les idiomatismes. Le cas de la négation avec et sans ne », *Le Français moderne*, 54, p. 137-60.
- PERRUCHET, P. (2006), « Discordes autour de l'acquisition du langage », *Les dossiers de La Recherche*, 22, p. 18-22.
- PINKER S. (1999), *L'instinct du langage*, Paris, O. Jacob.

SÉMINAIRES

- SEGUI J. et FERRAND L., (2000), *Leçons de parole*, Paris, O. Jacob
- SINCLAIR J., (1987), *Corpus, concordance, collocation*, Oxford, Oxford University Press.
- Wray (2002), *Formulaic language and the lexicon*, Cambridge, Cambridge University Press.