

Pour citer cette notice :

Berrendonner (A.), 2023, « L'opposition grammatical / lexical », in *Encyclopédie Grammaticale du Français*, en ligne : <http://encyclogram.fr>
DOI ; <https://nakala.fr/10.34847/nkl.466233b1>

ναὶ ναί, οὐδ' οὐ
τὸ δὲ περισσὸν τούτων ἐκ τοῦ πονηροῦ ἐστίν.
Si c'est oui, c'est oui, si c'est non, c'est non.
Tout le reste vient du démon. [Mat. V, 37]

1. DÉLIMITATION DU DOMAINE.

1.1. La thèse 'L/G'.

1.1.1. La distinction entre unités *lexicales* vs *grammaticales* (L/G) est un lieu commun en linguistique. On la trouve exposée dans les manuels (Lyons 1970 : 334 ; Arrivé & al. 1986 : 377 ; Gardes-Tamine 1988/1 : 43), dans les grammaires de référence (Riegel & al. 2009 : 536 ; Wilmet 1998 : 42), et dans maints travaux théoriques, où elle apparaît sous des formulations diverses :

- *mots pleins* vs *mots outils* (Brunot 1922 : 5) ;
- *mots autonomes* vs *mots grammaticaux* (Meillet 1948 : 131) ;
- *mots pleins* vs *mots vides* (Tesnière 1959 : 53) ;
- *sémantèmes* vs *ligaments* (Bally 1965 : 113) ;
- *mot lexical* vs *mot fonctionnel* (Abeillé 1993 : 159) ;
- *catégories lexicales* vs *fonctionnelles* (Pollock 1998 : 29) ;
- *lexèmes* vs *grammèmes* (Pottier 2000 : 97)...
- *vocabulary* vs *syntacticon* (Emonds 2000)

On tient ordinairement pour lexicaux les membres des catégories dites majeures (Noms, Verbes, Adjectifs, Adverbes), et pour grammaticaux les autres : Déterminants, Conjonctions, affixes de toutes sortes, etc. (Fradin 2003 : 89). On admet en outre qu'il existe des unités de statuts intermédiaires (Verbes auxiliaires et copules, Prépositions...), ce qui n'est pas incompatible avec l'existence d'un axe d'opposition bipolaire L/G, mais suppose simplement qu'il y a plus de deux positions sur cet axe. Il est souvent assimilé à un continuum (Halliday & Hasan 1976 : 5 ; Victorri 1999 : 85).

1.1.2. La pertinence de l'opposition L/G est notamment le postulat principal sur lequel repose la notion de *grammaticalisation*, qui a donné lieu à un grand nombre de travaux ces dernières décennies. On nomme ainsi, depuis Meillet (1948 : 133) le processus diachronique par lequel un

élément passe au fil du temps du statut d'unité lexicale à celui d'unité grammaticale (ou plus grammaticale, l'opposition étant conçue comme graduelle) :

Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical, or from a less grammatical to a more grammatical status. (Kuryłowicz 1965 : 52).

Bien évidemment,

Ce type de description du changement linguistique suppose une certaine conception de la langue. Il nécessite que l'on admette les [...] traits suivants comme inhérents à la langue : 1. La langue comporte deux sortes de mots : lexicaux (les lexèmes) et grammaticaux. » (Marchello-Nizia 2006 :19)

La présente notice a pour but d'inventorier les diverses façons de définir l'opposition L/G, et de présenter une synthèse critique des conceptions qui les sous-tendent.

1.2. La notion de lexique.

Au préalable, il convient de préciser comment on doit entendre les termes de l'opposition. Le terme de *lexique* est en effet ambigu, et dénote au moins trois réalités différentes.

1.2.1. Lexique^(A).

En un premier sens, le lexique est l'ensemble des unités significatives minimales (morphèmes), qui sont les atomes de la morpho-syntaxe :

Le stock complet des morphèmes d'une langue constitue son lexique. (Bloomfield 1933=1970 : 154)

Entendu ainsi, le lexique^(A) ne contient pas d'assemblages syntagmatiques : les mots construits et les mots fléchis, qui sont composés de plusieurs morphèmes, n'en font pas partie. D'autre part, c'est un ensemble énumérable, et à peu près non expansible : les morphèmes d'une langue sont en nombre fini, et il n'existe pas de moyen régulier d'en créer de nouveaux (cela ne peut se faire que par l'emprunt à d'autres langues¹, par siglaison, troncation, ou par figement d'expressions poly-morphémiques (univerbation), qui sont des accidents imprédictibles).

1.2.2. Lexique^(B).

Dans l'usage ordinaire des grammairiens, cependant, le terme de lexique désigne autre chose : c'est l'ensemble des *mots*. On appelle communément ainsi certains segments de texte délimités dans l'écriture par des espaces, et composés la plupart d'une base (ou *radical*) suivie de certains

¹ Il est à noter que les emprunts se font au bénéfice des catégories dites grammaticales aussi bien que lexicales. C'est ainsi p. ex. que le paradigme des suffixes formateurs de noms d'action s'est enrichi récemment d'un morphème */-ing/* des plus productifs, et qu'un morphème */because/* est venu s'ajouter à la classe des prépositions :

- (i) *caravaning, couponning, cataloguing, démerding, réunining, vacancing, forcing, truanding...* (web)
- (ii) (a) J'habite place de la Nation, et je dors aussi avec des boules Quiès, *because* bagnoles ! (cerf)
(b) jamais réussi à lire en profondeur dans les transports en commun, *because* manque de concentration. (cerf)

affixes particuliers (de temps-aspect, personne, nombre, genre). Parmi ces mots-occurrences, ceux qui sont formés sur une même base sont considérés comme des ‘instances’ (Fradin 2003 : 85) ou *flexions* d’un même mot-type. C’est celui-ci, appelé *lexème*, *item lexical*, *lexie* ou *lemme* (dans la pratique des dictionnaires), qui est tenu pour élément du lexique^(B). On le représente conventionnellement par l’infinitif dans le cas des verbes, et par le masculin singulier dans le cas des noms et adjectifs. On doit donc se demander quelle est la propriété empirique commune aux mots et à eux seuls qui justifie les lemmatisations usuelles, autrement dit sur quel critère est fondée la notion de mot, et par suite l’appartenance au lexique^(B).

1.2.2.1. Ce n’est évidemment pas sur un critère de complexité, puisque, contrairement au lexique^(A), le lexique^(B) contient à la fois des unités minimales (*ménage*, *utile*, *mal...*) et syntagmatiques (*dé-ménage-ment*, *in-util-is-able*, *mal-de-mer*, *porte-plume...*)

Remarque.

Il existe au demeurant des divergences quant à la façon de décrire la structure interne des mots.

- Certains linguistes les analysent en morphèmes. Pour eux, un mot peut être constitué d’un seul morphème (*chat*, *beau*, *chou*) ou de plusieurs morphèmes concaténés (*chat-ière*, *beau-té*, *chou-fleur*, *essuie-glace...*) (Corbin 1987 : 163 ; Apothéloz 2002 : 11).

- D’autres linguistes rejettent la notion de morphème au profit de celle de mot en tant qu’unité significative de base, et traitent les affixes comme de simples caractéristiques formelles inhérentes aux mots, dont elles reflètent des traits sémantiques ou catégoriels (Fradin 2003 ; Bonami 2014).

Ces divergences entretiennent une certaine équivoque sur l’extension du lexique^(B) (on ne sait pas toujours si on doit l’entendre comme incluant ou non les morphèmes), et par conséquent sur l’existence d’une éventuelle zone de recouvrement entre lexique^(B) et lexique^(A).

1.2.2.2. Selon une certaine approche psycho- et neurolinguistique, ce qui distinguerait les mots éléments du lexique^(B), c’est qu’ils font l’objet d’un traitement mémoriel : on les décrit comme des unités ritualisées et ‘encapsulées’, stockées dans la mémoire à long terme des locuteurs (Babin 1998 : 3), et que ceux-ci peuvent encoder et décoder ‘en bloc’, sans devoir nécessairement procéder par composition à partir de leurs éléments constitutifs. Cette définition du ‘lexique mental’ « mémorisé de façon stable dans le cerveau » (Glessgen 2011 : 421) pose cependant problème, car elle repose sur une caractéristique qui n’est pas propre aux mots. Sont aussi stockés en mémoire toutes sortes de syntagmes plus ou moins routinisés (*regain d’intérêt*, *développement durable*, *bête comme ses pieds*, *être en situation de handicap...*) et d’énoncés ‘formulaïques’ (proverbes, maximes, slogans, etc. Voir chez Kleiber (1994 : 209) l’analyse des proverbes comme « dénominations », c’est-à-dire comme « associations mémorisées », « résultats d’une habitude associative »). Or, ces expressions ne sont pas considérées comme faisant partie du lexique^(B). Dans sa délimitation usuelle, celui-ci ne s’identifie donc pas à la classe des unités sujettes à traitement mémoriel, mais à un sous-ensemble non précisé de celle-ci².

² Ce vague extensionnel s’explique en partie par le caractère graduel du processus de ritualisation : il présente des stades plus ou moins avancés, qui se traduisent par divers degrés de figement des expressions linguistiques (Bolly 2010), formant un continuum sur lequel il est difficile de placer un

1.2.2.3. Selon une autre définition courante, le lexique^(B) serait l'ensemble des faits de langue *idiosyncratiques*, dont l'existence et la valeur ne sont pas prédictibles par des règles, et qu'un modèle de la langue ne peut donc qu'énumérer un par un :

Nous proposons que l'ensemble LEX [...] soit la liste de toutes les lexies idiosyncratiques, c'est-à-dire la liste de toutes les lexies qui ont certaines propriétés qui ne peuvent pas être déduites des propriétés des éléments qui les composent [...]. C'est précisément parce que les propriétés d'une lexie idiosyncratique ne peuvent pas toutes être déduites des propriétés des éléments qui la composent, que cette lexie doit être mémorisée individuellement par les locuteurs. (Dell 1979)

Le lexique dénote l'ensemble des unités considérées comme des signes basiques dans une langue, parce qu'ils ne se déduisent de rien d'autre. (Fradin 2003 : 181)

Dans le lexique^(B) ainsi conçu, il semble justifié d'intégrer non seulement tous les morphèmes, qui sont autant de cas particuliers (Bloomfield 1933=1970 : 257), mais aussi les mots construits et fléchis, qui présentent de nombreuses irrégularités : lacunes dans les dérivations (**mangement*), signifiés non compositionnels (*dormition*), formations anormales (*contraceptuel*, *sachez*, *enverrai*, etc.).

Cependant, il existe aussi de nombreux syntagmes affectés d'irrégularités semblables. Certains dispositifs syntaxiques présentent des lacunes, comme p. ex. la construction [sujet *être* attribut] (1) ; certaines constructions ont un signifié non compositionnel (2) ; et d'autres présentent un ordre des mots qui peut sembler anormal, comme p. ex. (3) :

- (1) (a) le coupable, c'est Max / Max est le coupable
(b) le plus fort, c'est qu'il le croyait / *qu'il le croyait est le plus fort
- (2) (a) il s'est fait soigner = 'il a fait qu'on le soigne' vs
(b) il s'est fait agresser ≠ 'il a fait qu'on l'agresse'
- (3) (a) C'est tout moi qui fais dans cette maison (oral)
(b) On a beau eu chauffer, c'est irrécupérable. (web)
(c) De qui pourrait donc-t-il s'inspirer ? (web)³

Or, ces constructions ne sont pas pour autant traitées comme des mots et versées dans le lexique^(B). Celui-ci n'inclut donc qu'une partie des idiosyncrasies observables, sans qu'il soit précisé laquelle.

Par ailleurs, l'identification d'un fait de langue comme idiosyncratique ne tient pas seulement à sa nature, mais aussi au type de modélisation pratiqué : tel fait qui échappe aux règles d'un modèle peut apparaître régulier à la lumière d'un autre, plus puissant ou plus souple. La délimitation du lexique^(B) comme ensemble d'idiosyncrasies ne dépend donc pas seulement des données empiriques, mais aussi des postulats latents sur lesquels sont fondés les modèles, ce qui n'est pas sans comporter une part d'arbitraire. À la limite, cela peut même conduire à abolir la

seuil. Voir p. ex. la notion de *synapsie* dont Benveniste se sert pour caractériser un statut intermédiaire entre les mots composés et les syntagmes qui sont des « groupements occasionnels » (1974 : 171). Néanmoins, il existe parmi les expressions figées des phrases entières (*À bon chat, bon rat ; Les femmes et les enfants d'abord...*), et celles-ci ne sont jamais intégrées au lexique^(B) :

³ Nombre approximatif d'occurrences sur Google :
a beau eu = 18'000 / *a eu beau* = 163'000
donc-t-il = 80'900 / *t-il donc* = 1'110'000

distinction entre lexique et grammaire. Si p. ex. on considère, avec certaines grammaires constructionnelles, que toute construction est une unité non compositionnelle,

« dont certains aspects formels ou sémantiques ne sont pas strictement prévisibles à partir des éléments constitutifs de cette construction ou d'une autre construction identifiée » (Carlier & Prévost 2021 : 11)

il s'ensuit en toute rigueur que tous les assemblages syntagmatiques, étant traités comme des idiosyncrasies, se trouvent absorbés dans le lexique^(B), si bien que celui-ci devient coextensif à la langue entière... La notion perd alors tout intérêt.

1.2.2.4. En fin de compte, le lexique^(B) apparaît comme une collection vague, un rassemblement d'unités dont l'extension ne repose pas sur une propriété caractéristique claire et distincte. Cela vient de ce que la notion de mot, qui découle de nos habitudes d'écriture (séparation par des blancs) et des pratiques de lemmatisation usuelles, est un outil de catégorisation pratique (Berrendonner & (Reichler-)Béguelin 1989), et non un concept scientifique dûment défini. D'où sa récusation par de nombreux linguistes :

Il faut donc s'affranchir de la notion incertaine de mot. (Bally 1965 : 288)

C'est derrière l'écran du mot qu'apparaissent bien souvent les traits réellement fondamentaux du langage humain. (Martinet 1966 : 53)

La notion préthéorique de mot semble aussi commode que trompeuse, et la linguistique gagnera peut-être à s'en priver. (Rastier 1990)

Le « mot » [...] serait en fait très embarrassant à définir dans un savoir scientifique. (Blanche-Benveniste 1993 : 139)

Le mot ne constitue pas une unité phonologique et ne possède aucun statut théorique ou formel défini. (Laks 2005 : 103)

[We] do not have a good answer to the question of how to define the notion of word in a clear and consistent way that accords with our intuitions and with conventional practice, despite decades of research that has tried to address the issue. (Haspelmath 2011 : 2)

En dépit de ces avertissements, la plupart des modèles linguistiques postulent l'existence d'un lexique^(B) composé de mots, dont la délimitation est prudemment laissée dans le vague. Et c'est sur cette base (énoncés = suites de mots) que se font la plupart des analyses syntaxiques (voir p. ex. GGF, p. xxx : « Les mots sont les unités minimales de la syntaxe »).

1.2.3. Lexique^(C).

Enfin, admettre l'opposition L/G, c'est postuler une partition du lexique global, qu'on le conçoive comme lexique^(A) ou comme lexique^(B), en deux sortes d'éléments : 'grammaticaux' vs 'lexicaux'. D'où une troisième acception, lexique^(C), dans laquelle le terme est utilisé pour désigner un sous-ensemble particulier du lexique global, excluant les affixes, les mots clitiques, les conjonctions et autres 'mots outils'. Si l'on conçoit l'opposition L/G comme graduelle, le lexique^(C) est un ensemble flou au sens mathématique du terme [Kaufmann 1977], c'est-à-dire que la fonction d'appartenance à cet ensemble peut prendre des valeurs intermédiaires entre l'appartenance pleine et entière et la non-appartenance. Mais ce flou se double généralement d'un caractère vague, hérité du lexique^(B), faute de critères précis permettant de savoir où passe

exactement la frontière entre les assemblages qui sont considérés comme des mots situés sur l'axe L/G, et ceux qui sont traités comme des syntagmes.

1.2.4. Les propriétés distinctives par lesquelles on a cherché à caractériser le lexique^(C) sont de quatre sortes :

- Critères « systémiques » (= place dans l'économie générale du système langue)
- Critères syntaxiques
- Critères sémantiques
- Critères phonologico-prosodiques

Pour juger de la pertinence de l'opposition L/G, il nous faut donc examiner ces critères un à un.

2. CRITÈRES « SYSTÉMIQUES ».

2.1. Classes ouvertes / classes fermées.

Entre unités lexicales [+L] vs grammaticales [+G], la différence la plus souvent mentionnée est d'ordre quantitatif : les premières appartiennent, dit-on, à des classes ouvertes, et les secondes à des classes fermées (*i.a.* Martinet 1960 : 119 ; Lyons 1980 : 21 ; Riegel & al. 2009 : 536 ; Talmy 1995 : 140 ; Fradin 2003 : 89 ; Emonds 2000 : 5). Cette différence peut être envisagée sous deux aspects.

2.1.1. Inégalité d'effectifs.

Les éléments des paradigmes [+G] apparaissent en général peu nombreux comparativement à ceux des paradigmes [+L].

Les monèmes lexicaux sont ceux qui appartiennent à des inventaires illimités. Les monèmes grammaticaux sont ceux qui alternent, dans les positions considérées, avec un nombre relativement réduit d'autres monèmes. (Martinet 1960 : 119)⁴

Cependant, les différences de taille observées entre classes dépendent à l'évidence de la façon dont celles-ci sont définies. Or, le classement qui est communément pratiqué prend pour base les parties du discours traditionnelles, qui présentent des degrés de finesse très inégaux. Les catégories dites 'majeures' (N, V, Adj, Adv) sont en fait des archi-catégories dont chacune subsume toute une série de classes de distribution différentes (> Notice 'Catégories'), tandis que les catégories 'grammaticales' (articles, pronoms clitiques, affixes...) coïncident généralement avec une classe de distribution plus précisément définie et relativement homogène. En procédant à des dénombrements sur cette base, on compare donc ce qui n'est pas comparable.

Si en revanche on raisonne partout sur des classes de distribution précises, les écarts d'effectifs apparaissent bien moins flagrants. Parmi les morphèmes réputés [+L], certains appartiennent à des séries limitées. Exemples :

⁴ Martinet utilise le terme de *monème* au lieu de celui de *morphème* pour désigner les unités significatives minimales de la langue.

- Les verbes régissant à la fois un ablatif et un locatif valenciels (= compléments respectivement proportionnels aux pronoms *en* et *y*) sont peu nombreux. *Dicovalence* (Eynde & Mertens 2010) ne mentionne que {*arriver, descendre, s'écouler, partir, rentrer, venir*}, à quoi il faut ajouter leurs dérivés {*redescendre, repartir, revenir*}.
- Les verbes exclusivement impersonnels se comptent sur les doigts de la main : {*falloir, y avoir, s'agir*}. Si on y ajoute les verbes 'atmosphériques' {*pleuvoir, neiger, tonner...*}, en négligeant le fait qu'ils ont des emplois figurés à sujets personnels, cela n'allonge guère la liste.
- L'inventaire des adjectifs qualificatifs dont l'antéposition au nom est le placement par défaut se limite à une vingtaine d'items (> Notice 'Place de l'Adjectif épithète' § 411).

À l'inverse, certaines classes considérées comme grammaticales contiennent de très nombreux éléments. Celle des numéraux cardinaux, qu'on range souvent parmi les déterminants du nom (Arrivé & al. 1986 : 426 ; Riegel & al. 2009 : 297 ; Monneret 2020 : 116), compte des centaines de milliers d'éléments engendrables par concaténation à partir d'une vingtaine de morphèmes de base (*un, deux, trois... neuf cent quatre-vingt dix-neuf mille neuf cent quatre-vingt dix-neuf*)⁵.

Ce n'est donc pas sur le simple dénombrement des éléments d'une catégorie que l'on peut fonder rationnellement sa qualification usuelle [+L] ou [+G].

2.1.2. Créativité.

2.1.2.1. Par classes ouvertes, on entend d'autre part des classes expansibles, qui peuvent être augmentées de nouveaux membres et sont donc le lieu d'une certaine créativité. Celle-ci est de nature combinatoire. Elle tient à l'existence d'opérateurs applicables à des mots pour construire d'autres mots : suffixes et préfixes dérivationnels (*bon-té, norm-al, dé-faire, in-mange-able...*) ou opérateurs binaires qui à partir de deux lexèmes en composent un troisième, comme la simple concaténation (*chou-fleur, poisson-chat...*) ou le morphème /et/ formateur de composés 'dvandva' (*père et mère, nuit et jour, sain et sauf, aller et venir...*)⁶. Une classe ouverte est une classe de distribution qui comprend des unités poly-morphématiques construites au moyen de tels opérateurs. Cette notion de créativité morphologique appelle plusieurs remarques.

2.1.2.2. Selon certains linguistes, elle ferait du lexique^(C) un ensemble infini :

Le nombre des mots d'une langue qui possède des règles de construction de mots est infini. (Corbin 1987 : 23)

Le stock des unités lexicales, en revanche, est potentiellement infini dans la mesure où les locuteurs peuvent créer de nouveaux membres. (Fradin 2003 : 89)

Cette thèse est cependant contestable, car l'usage récursif des opérateurs constructeurs de mots est en fait très limité, voire le plus souvent exclu. La récursivité affixale directe est rarement

⁵ À vrai dire, certaines grammaires préfèrent ranger les numéraux parmi les adjectifs (GGF : 588), ce qui est sans doute distributionnellement plus juste. Si on adopte cette analyse, ils sont un exemple de classe lexicale qui, tout en contenant un grand nombre de mots construits, est strictement finie (au-delà de 999'999, on dénombre au moyen de substantifs : *un million de N*, et non **mille mille N*).

⁶ Ces composés par coordination binaire sont particulièrement fréquents en sanscrit, ce qui leur a valu le qualificatif de *dvandva* (= 'couple') dans les grammaires de cette langue.

pratiquée (Corbin 1987 : 496), et limitée de fait à une seule ré-application (*re-refaire*, *anti-antigénique*, *archi-super-chouette*). Quant à la récursivité médiata, elle permet quelques allers-et-retours du genre *orner* > *orne-ment* > *ornement-er* > *ornement-ation*, mais ces formations se limitent à une seule ré-application d'un même type de translatif (les chances de rencontrer **ornementationisation* ou **ornementationnement* semblent nulles). Le nombre des morphèmes bases étant fini, et les opérations de dérivation n'étant pas récursives *ad libitum*⁷, rien ne permet donc d'affirmer que leur produit soit à proprement parler un ensemble infini. Certes, la possibilité de composer plusieurs opérateurs de dérivation les uns sur les autres fait que le nombre des mots construits possibles est très élevé et défie tout inventaire exhaustif ; mais il n'en est pas moins fini. Et par conséquent, les classes dites 'ouvertes' sont elles aussi des ensembles finis. L'impression qu'elles peuvent s'accroître indéfiniment peut être récusée comme une illusion, due à notre incapacité d'en recenser tous les membres (mots possibles) à partir des données observables (mots attestés, vocabulaire actif).

2.1.2.3. D'autre part, un examen attentif montre que la plupart des catégories [+G] comprennent des éléments poly-morphématiques formés par composition. Ainsi :

- Il existe des composés *dvandva* à base de pronoms clitiques, de déterminants du nom, de suffixes ou de préfixes⁸ :

- (4) (a) *Je et tu* sommes libres ou pas d'aller acheter des sandwiches ici ou là. (web)
 (b) mes chats *il et elle* sont tres tres beau et belle [*sic*, web]
 (c) *Elles et ils* ont participé à ce numéro. (*M le Magazine du Monde*)
- (5) Je vais presque tous les jours voir *le et la* libraires d'Outremont. (Robin, frantext)
- (6) Quelqu'un m'aperçut : M. Germain de Saint-Germain soi-même, entouré d'une petite cour d'*admirateurs et trices* en blue-jeans. (L. Malet)
- (7) (a) La période *pré et post*-opérateur (web)
 (b) Les militants *pro- et anti*-avortement (web)

- Ce qu'on appelle 'coalescence' d'affixes n'est autre que la formation d'un composé par concaténation de deux morphèmes grammaticaux, donc un processus créatif. Selon Kilani-Schoch & Dressler (2004) p. ex., les formes de conditionnel comme *parle-r-ai-t* ont des significations non compositionnelles, qui ne peuvent pas résulter de la simple combinaison des deux affixes indépendants de futur /R/ et d'imparfait /ε/. On peut donc penser que ces affixes se sont solidarisés pour composer un nouveau morphème /Rε/, porteur d'un signifié modal spécifique. On peut, de la même manière, analyser les noms déverbaux d'instruments *lavette*, *dînette*, *roulette*, *épuisette*... comme formés au moyen d'un suffixe /-εt(ə)/ résultant de la

⁷ Du moins si l'on admet que l'usage limité des opérations récursives est dû à des contraintes d'optimalité qui font partie de la compétence linguistique des sujets parlants, et qu'on se refuse à mettre ces limitations au compte de facteurs extra-linguistiques (capacité limitée de la mémoire à court terme, etc.) comme le fait Chomsky pour s'en débarrasser (1965 = 1971 : 24 *sqq.*).

⁸ On traite généralement ces groupes comme des coordinations syntaxiques ordinaires, mais cela ne rend pas compte de leurs particularités : structure binaire et non n-aire, non commutativité (**tu et je* est inattesté), lacunes dans la formation (**je et on*, **la et cette*), signifié parfois non compositionnel (*deux ou trois* ≅ <un petit nombre indéterminé>). Ces propriétés s'expliquent si on y voit des mots composés, mais on doit alors admettre que les catégories [+G] peuvent contenir des mots construits.

coalescence des affixes /-ε(t)/ diminutif et /-ə/ de féminin. (Cette analyse évite d'avoir à supposer l'existence de formes intermédiaires inattestées du genre *lav-et, *épuis-et, comme le fait Corbin (1987 : 662 sqq.)). L'histoire du français donne par ailleurs maints exemples de mots grammaticaux formés par coalescence, comme les déterminants *ledit* ou *quelque(s)*, les conjonctions *parce que*, *lorsque*, *quoique*, le modalisateur *est-ce que*, etc.

La créativité morphologique n'est donc pas le propre des catégories lexicales. En fait, quasiment tous les paradigmes d'unités significatives, clitiques et affixes compris, apparaissent virtuellement expansibles par composition (voir aussi la Note ¹ *supra*). Ne semblent faire exception que deux ou trois classes d'affixes flexionnels (désinences de genre, de nombre, de personne) qui ne contiennent qu'un tout petit nombre de morphèmes, voire un seul couple oppositif (pluriel /-z/ vs /ø/ singulier), et aucune forme construite (?⁹). Ce sont ces paradigmes « figés » (Riegel & al. 2009 : 895) qui sont toujours cités lorsqu'il s'agit d'illustrer le statut [+G].

2.1.2.4. Si les paradigmes [+L] et [+G] ne se différencient pas quant à la formation d'éléments composés, il n'en va cependant pas de même de la créativité par dérivation (Amiot 2021 : 1894). Il existe en effet clairement deux sortes de classes de distribution : les unes contiennent des bases de dérivation et des formes dérivées en plus ou moins grand nombre (verbes, noms, adjectifs), tandis que les autres ne contiennent ni bases de dérivation ni dérivés (pronoms clitiques, déterminants, affixes...). Le couple classes ouvertes vs classes fermées apparaît donc en dernière analyse comme l'appréhension quantitative d'une différence fonctionnelle entre deux types de catégories : celles qui sont le domaine et/ou le codomaine d'opérateurs de dérivation vs celles qui ne le sont pas. Cette distinction est le principal critère sur lequel on puisse fonder empiriquement l'opposition L/G (§ 63). On notera cependant que la définir ainsi interdit de la considérer comme un axe graduel : ou bien une classe contient des dérivés, ou bien elle n'en contient pas, mais il ne peut pas y avoir de statuts intermédiaires.

2.2. *Fréquence.*

2.2.1. Une autre propriété souvent invoquée pour fonder l'opposition L/G est la fréquence d'emploi des unités :

Les marqueurs grammaticaux s'emploient plus fréquemment que les lexèmes pleins. (Lamiroy 1999 : 37)

Mais cette idée se heurte à de multiples démentis empiriques.

• En réalité, beaucoup de morphèmes grammaticaux ont une fréquence très basse. Dans le corpus oral CTFP (Blanche-Benveniste & al. 2002) (env. 65'000 mots), on trouve p. ex. :

⁹ Encore que certaines pratiques d'écriture 'inclusive' puissent être vues comme des tentatives d'acclimater un troisième genre 'neutre' ayant pour signifiants des affixes composés par concaténation, p. ex. : *écol+[ier-ère]*, *act+[eur-rice]*, *c+[eux-elles]*. Voir aussi l'usage militant du pronom clitique *iel* (= composé *il+elle*) censé neutraliser l'opposition masculin/féminin. Même si ces innovations sont essentiellement orthographiques et n'ont guère de chances d'être ratifiées par l'ensemble de la communauté parlante, elles n'en montrent pas moins que les catégories affixales ne sont pas plus que les autres à l'abri de toute créativité spontanée.

- morphème de passé simple : 16 occurrences (concentrées dans deux textes) ;
- déterminants *quelques* : 17 occ. ; *plusieurs* : 12 occ. ;
- suffixes *-eraie* = 1 occ. ; *-esque* = 1 occ ; *-issime, -âtre* = 0 occ.

Ces fréquences sont bien inférieures à celles de morphèmes [+L] comme *temps* (101 occ.), *jour* (84 occ.) ou *grand* (83 occ.).

• Par ailleurs, il n'est pas rare que des morphèmes appartenant à la même classe de distribution aient des fréquences très inégales. Ainsi, dans un écrit narratif de 125 000 mots (journal d'un civil 1939-1945), on dénombre 1 035 occurrences de *il(s)* contre seulement 2 occurrences de *tu*. Il est donc clair que l'appartenance à la catégorie des pronoms clitiques sujets n'est pas en soi un gage de fréquence élevée. On observe dans d'autres corpus des écarts convergents, quoique plus ou moins marqués selon le type d'interaction. P. ex. à l'oral :

OFROM :	<i>il(s)</i> = 14 622 occ.	<i>tu</i> = 6 126 occ.
CRFP :	9 162	1 513
CLAPI :	8 146	4 950

Si l'on fonde l'opposition L/G sur la fréquence relative, on doit donc considérer que *il(s)* est plus grammatical que *tu*. Cela revient à dire que la position d'une unité sur l'axe L/G est indépendante de la catégorie syntaxique à laquelle elle appartient, et qu'il n'est donc pas pertinent d'attribuer aux catégories des valeurs sur cet axe. Le critère de fréquence entre ainsi en contradiction avec la définition de l'opposition L/G en termes de classes ouvertes vs fermées.

• Autre exemple : une comparaison des prépositions *avec* et *sans* dans un corpus varié de 10 M de mots a montré que *avec* représente 70% des occurrences, contre 30% pour *sans* (Clech 2005). Cette proportion s'accroît dans les petites annonces (pour des raisons argumentatives : il est avantageux de souligner qu'un objet va avec certains accessoires, mais pas qu'il en est dépourvu), et elle s'inverse en poésie (le manque, ça fait mélancolique). Si l'on se fie au critère de fréquence, on doit donc admettre que la position respective de ces prépositions sur l'axe L/G change selon le type de texte où elles sont employées, sans qu'on puisse mettre cela au compte d'un phénomène classique de variation sémantique ou syntaxique.

2.2.2. En fait, il semble que la fréquence d'une unité dépende principalement (i) de son signifié individuel et (ii) du genre de discours dans lequel elle est employée ; autrement dit, de sa pertinence communicative en contexte. Et si certains morphèmes grammaticaux apparaissent particulièrement fréquents (*de, à, le...*) dans tous les types de textes, cela ne tient pas à leur appartenance catégorielle, mais plutôt à leur sémantisme générique ou basique, qui en fait des unités passe-partout. Les écarts de fréquence entre unités 'lexicales' s'expliquent d'ailleurs de la même façon. Cf. p. ex. les *verba dicendi* dans le corpus oral OFROM :

<i>dire</i> :	6'446 occurrences
<i>raconter</i> :	189
<i>affirmer</i> :	8
<i>déclarer</i> :	5

Si *dire* apparaît de beaucoup le plus fréquent, cela tient évidemment à son signifié qui constitue le 'niveau de base' du paradigme des *verba dicendi* (au sens de Rosch & al. (1976), v. Kleiber 1990 : 83), et se trouve approprié à un plus grand nombre de situations. Faut-il en conclure que

tout terme de base est par nature plus grammatical que ses hyponymes et hypéronymes ? Ce serait là une façon inédite d'envisager l'opposition L/G.

Finalement, la conclusion qui s'impose est qu'il n'y a pas de corrélation nécessaire entre le classement usuel d'une unité comme [+L] ou [+G] et sa fréquence d'emploi, si bien que celle-ci ne peut pas servir de justification à celui-là.

3. CRITÈRES SYNTAXIQUES.

3.1. Relation de dépendance.

3.1.1. Un critère souvent invoqué à l'appui de l'opposition L/G est le degré d'autonomie syntaxique. L'idée que les unités grammaticales sont d'emploi moins libre que les unités lexicales a reçu des formulations multiples et variées :

	+L	+G
Meillet 1948	mot autonome	élément grammatical, groupé avec un autre d'une manière qui tend à devenir fixe
Hopper & Traugott 1993	<i>more independent</i>	<i>less independent</i>
Haspelmath 1998	<i>loose structure independent words</i>	<i>tight structure reducing autonomy, loss of autonomy less freedom in the use of expression</i>
Diewald 1997	<i>autonomous form</i>	<i>non-independent form</i>
Traugott 1999		<i>increased bonding syntactic scope reduction</i>
Pagliuca 1994	<i>unbound items</i>	<i>affixes entirely dependant on their hosts</i>
Marchello-Nizia 2006	mot autonome à haut degré de liberté	position fixe

Et le processus de grammaticalisation est décrit comme une perte de liberté combinatoire :

Grammaticalization of a linguistic sign is a process in which it loses in autonomy by becoming subject to constraints of the linguistic system. (Lehmann 2004 : 155).

Grammaticalization comprises [...] the development of function words from content words, the development of affixes from function words. [...] These changes can be understood as resulting from the gradual loss of autonomy of linguistic signs. (Haspelmath 1998 : 52)

3.1.2. Sous la notion d'autonomie syntaxique, on met généralement le fait que certaines unités ne peuvent être employées qu'avec un acolyte dont elles impliquent la cooccurrence, tandis que d'autres unités ne semblent comporter aucune dépendance du même genre. Cette différence est à la base de la distinction bloomfieldienne entre formes 'liées' vs 'libres' :

Une forme linguistique qui n'est jamais prononcée seule est une forme *liée* ; toutes les autres sont des formes *libres*. (Bloomfield 1933=1970 : 152)

On l'invoque souvent pour définir l'opposition L/G. Les unités grammaticales sont réputées astreintes à cooccurrence avec une unité 'hôte', voire décrites comme inséparables de celle-ci (Marchello-Nizia 2006 : 41) ou 'obligatoires' (Diewald 2004), tandis que les unités lexicales

passent pour indemnes de telles contraintes. On en veut pour preuve qu'elles peuvent être énoncées en isolation :

Since grammatical elements have secondary prominence, they depend on host elements in relation to which they are secondary (in *chair-s*, for instance, the plural *-s* is dependent in relation to *chair*). This provides a rationale for distinguishing between uncontroversial cases of lexical items (such as *chair*) and grammatical ones (such as plural *-s*), but can also throw light on cases that are not equally obvious. For instance, it provides a rationale for distinguishing between lexical and grammatical prepositions (*of* comes to stand as grammaticalized congener of the lexical preposition *off*) and between grammatical and lexical pronouns in French (*me* vs *moi*).

In certain cases, lexical items can be used as potential full utterances. Utterances like *fire*, *tomorrow* or *run* may be used without being combined with other linguistic expressions. This is not the case for grammatical expressions. In order to be operational, they need an additional combinatory procedure to be usable by the speaker. The 'dependency' property introduced above as a characteristic of grammatical expressions translate into the property 'obligatory procedural linkage' with a lexical 'host'. (Boye & Harder 2020 : 137-140)

3.1.3. Il s'avère malheureusement que l'indépendance vs dépendance syntaxique ainsi comprise ne coïncide nullement avec l'opposition L/G telle qu'on l'admet communément.

- En français p. ex., tous les verbes sont des morphèmes liés. Ils ne s'emploient que conjugués, c'est-à-dire accompagnés d'une désinence de personne, d'infinitif ou de participe avec laquelle ils sont en relation d'implication réciproque. Il en va de même des Noms et des Adjectifs, qui ne s'emploient qu'en cooccurrence avec des affixes de nombre et/ou de genre. Si l'on s'en remet au critère de la dépendance syntaxique, tous les lexèmes soumis à flexion devraient donc être rangés parmi les morphèmes grammaticaux¹⁰.

Remarque

Il est intéressant de constater que les affixes flexionnels sont unanimement décrits comme des formes liées, tandis que les bases fléchies, bien qu'elles soient avec eux dans le même rapport de dépendance réciproque, sont traitées implicitement comme des formes libres. L'usage qui est fait de l'opposition formes liées / formes libres apparaît ainsi biaisé par les pratiques lexicographiques traditionnelles : la lemmatisation, qui consiste à traiter les formes fléchies comme des avatars de leur base, revient à ériger celle-ci en unité indépendante de ses flexifs.

- À l'inverse, des morphèmes considérés ordinairement comme [+G] peuvent à l'occasion constituer un énoncé à eux seuls, et devraient par conséquent être classés comme lexicaux :

- (8) Bouboule [...], la paluche tendue par-dessus le zinc, lance simplement : – *Re... !* (Simonin, frantext)

¹⁰ On ne saurait objecter que les marques de flexion sont souvent absentes, car ces absences commutent avec des morphèmes ou des allomorphes pleins, et ont donc la valeur de morphèmes zéro. Cf. par exemple à l'impératif : [*parl-ons*], [*all-ons*] = 1^{ère} personne vs [*parl-Ø*], [*v-a*] = 2^{ème} personne. Si on se refuse à admettre l'existence de morphèmes zéro, cela n'arrange rien : on doit alors considérer qu'un verbe est tantôt une forme libre [*parl*], tantôt une forme liée [*parl-ons*], et passe donc du statut d'unité lexicale à celui d'unité grammaticale au cours de sa conjugaison...

- (9) – Vous êtes né le 20 décembre 1920 ?
 – Oui Monsieur.
 – À ?
 – À Paris. (Genet, frantext)
- (10) – Tu es entré ?
 – Ben oui, pour voir.
 – Et ?
 – Et rien. (Chalandon, frantext)¹¹

• Quant à la supposée inséparabilité des morphèmes clitiques, elle est démentie par toutes sortes d'insertions parenthétiques :

- (11) (a) *Je* (plus que probablement) l'avais désactivée. (e-mail)
 (b) Irina est mariée avec *un*, comme on dit, musicologue. (web)
 (c) Vous allez faire un stage au service de presse *de*, tenez-vous bien, Citroën ! (Manœuvre, frantext)

On doit donc abandonner l'idée que l'opposition usuelle L/G puisse être justifiée par le jeu des relations syntaxiques de dépendance.

3.1.4. En outre, le fait d'être lié à des voisins d'un certain type équivaut à une forte limitation d'emploi. Si les unités grammaticales se distinguent par leur absence d'autonomie, elles devraient donc présenter des distributions plus étroites que les unités lexicales, et le processus de grammaticalisation d'un élément devrait se manifester par une restriction de sa distribution. Or, divers diachroniciens ont fait observer l'inverse : la grammaticalisation d'un mot comporte le plus souvent (ou commence par) l'apparition d'occurrences métaphoriques (Claudi & Heine 1986, Heine 1993) ou métonymiques (GGHF : 29). Ces tropes consistant à employer une unité dans des contextes insolites en violation de ses restrictions sélectives ordinaires, on doit en conclure que le statut [+G] va de pair avec une distribution plus large.

La désémantisation d'un élément linguistique entraîne un changement du point de vue de la distribution, qui peut également être considéré comme un élargissement plutôt que comme un appauvrissement (Heine 1993 : 54). (Lamiroy 1999 : 35)

Comment perçoit-on en effet qu'un phénomène de grammaticalisation est en cours ? Par le changement dans les emplois du lexème concerné, c'est-à-dire par la modification du type de constructions dans lesquelles il peut entrer, et qui dans la plupart des cas s'élargit. (Marchello-Nizia 2006 : 22)

On peut en donner pour exemple certains verbes qui sélectionnaient originellement un sujet [+ animé] (*menacer, faire mine, risquer...*) et qui, en devenant des auxiliaires d'aspect

¹¹ On pourrait arguer que les énoncés (8-10) ne sont pas formés d'un unique morphème, mais sont des constructions elliptiques composées de ce morphème et d'un constituant \emptyset dont il dépend. Le fait que (9) et (10) s'interprètent comme des questions partielles portant sur un constituant non dit (\cong À <quel endroit ?>, Et <quoi ?>) va en ce sens. Cependant, les énoncés constitués d'un simple item lexical sont dans le même cas ; ils peuvent eux aussi être analysés comme elliptiques. Cf.

- Il fait beau ?
 - *Superbe*, dit Laura. (Aventin, frantext) \cong <Il fait> superbe

Que l'emploi en isolation soit une preuve d'indépendance syntaxique ou le résultat d'une ellipse, il n'est donc pas une aptitude réservée aux unités [+L], et qui les différencierait des unités [+G].

prospectif, ont acquis la liberté de se combiner avec n'importe quel type de sujets, y compris impersonnels (*La mer risque d'être mauvaise ; il menace de pleuvoir ; il fait mine de pleuvoir*). Le cas le plus spectaculaire est celui du verbe *penser*, qui a gagné au XVII^e siècle des emplois d'auxiliaire avec sujets quelconques, au sens de *faillir*¹² :

- (12) (a) Une pierre qui tomba pensa le tuer. (Acad.)
(b) L'ambition et les jalousies qui se mirent parmi eux les pensèrent perdre. (Bossuet)
(c) Il a un je ne sais quoi dans le regard qui a pensé m'arrêter. (Marivaux)

L'idée que l'axe d'opposition L/G coïnciderait avec une échelle de latitude distributionnelle <distribution large/étroite> ne semble donc pas s'accorder avec ce type de faits diachroniques.

Remarque.

Décrire la grammaticalisation à la fois comme perte d'autonomie et comme élargissement d'emploi peut donc sembler contradictoire. À vrai dire, les phénomènes diachroniques de transcatégorisation comportent souvent à la fois un élargissement distributionnel relativement à certains facteurs contextuels et une restriction relativement à d'autres, ce qui rend très difficile la mesure comparative des latitudes combinatoires avant et après. La 'grammaticalisation' du verbe *aller* en auxiliaire de futur, par exemple, a eu pour effet d'une part d'élargir sa distribution aux sujets inanimés et aux infinitifs impersonnels (*les roses vont fleurir, il va pleuvoir*), et d'autre part d'exclure sa construction avec des compléments locatifs (**fleurir, les roses y vont*) ainsi que sa combinaison avec certains flexifs de temps-aspect (morphème de passé simple, auxiliaire de parfait : **il alla pleuvoir, *il est allé pleuvoir*).

3.2. L'opposition L/G comme postulat.

3.2.1. Une autre façon de fonder en syntaxe l'opposition L/G est d'en faire une conséquence, dans l'ordre paradigmatique, d'un postulat théorique d'*autonomie de la syntaxe*. Selon ce postulat, professé notamment par Chomsky dès ses premiers travaux (1957=1969 : 17), la composition des énoncés obéit à deux sortes de contraintes combinatoires. Les unes portent sur leur structure formelle ; elles relèvent de la syntaxe, qui est supposée constituer dans la langue un sous-système de règles autonome. Les autres sont vues comme de simples conditions d'interprétabilité ; elles sont affaire de sémantique ou de logique, mais pas de syntaxe. Cette distinction entre *grammaticalité* (bonne formation syntaxique) vs *appropriété* (du sens) est à la base de la plupart des modélisations grammaticales actuelles (GGF : XXVI ; Riegel & al. 2009 : 33).

Parmi les propriétés qui caractérisent les éléments du lexique^(A), elle conduit à distinguer deux sortes de traits, syntaxiques vs purement sémantiques :

- On tient classiquement pour syntaxiques les traits qui entrent dans la formulation des règles de placement séquentiel, d'accord et de régime : catégorie grammaticale, nombre, personne, animation, genre, cas, temps-aspect, modalité, etc.

¹² Reste à comprendre pourquoi ces emplois ont disparu après le XVIII^e siècle.

- Les autres traits sont non pertinents syntaxiquement. Ils consistent en différences sémantiques subalternes dont dépend seulement l'interprétabilité des énoncés. Est souvent donné en exemple le trait [+ liquide] que le verbe *boire* exige normalement de ses compléments. Cf., à construction syntaxique identique :

- (13) (a) boire de l'eau / du pétrole / dix litres de bière... = interprétable et cohérent
 (b) boire du courage / ma voiture / dix mètres d'altitude = incohérent ou ininterprétable

Sur la distinction entre ces deux types de traits est fondée une partition du lexique^(A) en deux sous-ensembles. L'un a pour éléments les morphèmes grammaticaux, qui ne possèdent que des traits syntaxiques. L'autre comprend les morphèmes lexicaux, qui possèdent en outre des traits purement sémantiques. Emonds (2000, 2013) les nomme respectivement *Syntacticon* vs *Dictionary* :

The grammatical morphemes [...] are then exactly those morphemes *whose only features are syntactic F*; they have no purely semantic features.
 The Syntacticon is the set of lexical items in a language *that lack purely semantic features f*.
 The "Dictionary" is then defined as the Lexicon less the Syntacticon. That is, every Dictionary entry has some purely semantic feature *f* not used in syntax.

Selon cette version de l'opposition L/G, les membres des paradigmes grammaticaux ne s'opposent donc les uns aux autres que par des traits syntaxiques, *i.e.* ont tous un comportement syntaxique différent, tandis que deux lexèmes peuvent avoir la même distribution syntaxique et ne différer que par des traits sémantiques (cf. p. ex. *voiture, auto, bagnole*).

The non lexical categories are called functional or closed class categories.
 Any two members of closed class categories differ by at least one syntactic feature used in grammar, that is, no two closed class items should have identical syntax. (Emonds 2019)

3.2.2. Remarques.

(i) 'Syntaxique' ne veut pas dire dépourvu de pertinence sémantique : les traits syntaxiques sont tenus pour interprétables, au moins dans certaines de leurs occurrences (Pollock 1998 : 116). Il s'ensuit que les morphèmes grammaticaux ne sont nullement des unités vides de sens.

(ii) Ainsi conçue, l'opposition L/G est sans rapport avec la différence entre distributions larges vs étroites (*contra* § 314). Tout dépend en effet de la nature des traits responsables des restrictions distributionnelles. Ainsi, beaucoup de termes appartenant à des vocabulaires de spécialité ou à des nomenclatures techniques ont une distribution très étroite, du fait qu'ils ne sont prédicables que sur une classe de SN bien particulière : le verbe *faseyer* ne se dit que des voiles de bateau, les adjectifs *bai, alezan* ne peuvent qualifier que des chevaux, *isocèle* ne s'applique qu'aux figures géométriques, etc. Mais les traits [+ voile] ou [+ cheval] n'étant pas syntaxiques, ce genre de restrictions ne doit pas être tenu pour la marque d'un statut [+G] (voir cependant le § 324 *infra*).

(iii) Cette façon de définir l'opposition L/G entraîne certains reclassements par rapport à la tradition. L'assimilation des catégories majeures au lexique^(C) se trouve notamment remise en cause :

Importantly, it is perfectly possible that lexical head morphemes (of category N, V, A or P) can lack any purely semantic feature, and thus be part of the Syntacticon. [...] That is, the lexical categories N, V, A, P contain closed subsets of grammatical items (Emonds 2013).

Emonds en donne pour exemple les divers types de verbes auxiliaires, qui ne portent que des traits syntaxiques de temps-aspect ou de modalité. On pourrait y ajouter le ‘pro-verbe’ *faire* (Eriksson 2006), et les ‘noms généraux’ comme *fait* (Huyghe 2018) ou *chose* (Blanche-Benveniste 1986), qui dans certains de leurs emplois peuvent être plausiblement décrits comme de purs opérateurs syntaxiques. Admettre ainsi que les catégories majeures incluent des sous-classes grammaticales ‘fermées’ prévient dans une certaine mesure l’objection adressée *supra* § 211 à la notion usuelle de ‘classes ouvertes’.

(iv) La distinction entre traits syntaxiques vs sémantiques n’implique pas nécessairement une partition binaire du lexique^(A), du genre *dictionary* / *syntacticon*. Elle est compatible avec une conception plus diversifiée de l’opposition L/G, fondée sur certaines conjonctions de traits.

Exemple.

Dans la grammaire du GARS (Groupe Aixois de Recherche en Syntaxe), les pronoms personnels accentués {*moi, toi, lui, elle(s), eux, nous, vous*} et les mots interrogatifs indéfinis {*qui, quoi, lequel, où...*} sont traités de ‘semi-lexicaux’ (Blanche-Benveniste 1990 : 84). Le critère sur lequel repose cette qualification n’est pas explicitement déclaré, mais semble résider dans la conjonction de deux propriétés :

- (i) ces unités occupent les mêmes positions que les SN ‘lexicaux’ ;
- (ii) ce sont des ‘proformes’, c’est-à-dire des expressions référentielles dépourvues de traits purement sémantiques.

Ces morphèmes diffèrent donc à la fois, par (i), des morphèmes grammaticaux que sont les pronoms clitiques, et par (ii), des unités lexicales (SN) qui expriment des valeurs sémantiques particulières. Leur caractérisation de ‘semi-lexicaux’ apparaît ainsi comme une façon de raffiner l’articulation de l’axe L/G, sans pour autant remettre en question la dichotomie traits syntaxiques / traits sémantiques. (Cela suppose toutefois que le signifié des proformes soit exclusivement imputable à des traits syntaxiques, ce qui prête à discussion).

3.2.3. Comme tout postulat, la distinction entre traits syntaxiques vs sémantiques encourt le reproche d’être fondée sur des décisions arbitraires. Étant donné p. ex. que le verbe *boire* sélectionne un complément marqué à la fois [+ accusatif] (SN directs, clitique *le*), [– animé] et [+ liquide], pourquoi les deux premiers traits sont-ils tenus pour syntaxiques, et le troisième pour purement sémantique ? Aucun argument empirique n’est avancé en réponse à cette question.

Ce traitement différentiel est vraisemblablement motivé par une difficulté d’analyse. Les traits réputés syntaxiques sont relativement faciles à définir, car ils entrent dans des oppositions peu diversifiées, souvent binaires voire privatives. En outre, ils s’instancient isolément dans certains morphèmes (flexifs de nombre, genre, personne, temps-aspect, etc.), et cette inscription dans des signifiants distincts est un garant de leur pertinence codique. Les traits sémantiques, au contraire, sont malaisés à caractériser précisément : en l’état actuel des connaissances, on ne saurait dire quel est au juste l’axe d’opposition dont le trait [+liquide] est le terme, ni énumérer les items dans lesquels il est présent. Ce trait demeure donc suspect de n’être qu’un étiquetage

intuitif, dont la pertinence descriptive n'est pas avérée. Le mettre hors syntaxe apparaît alors comme une mesure de « précaution pratique » (Blanche-Benveniste 1981 : 64), dictée par les capacités d'analyse du grammairien plutôt que par les propriétés des constructions à analyser : sont réputés syntaxiques les traits que l'on sait définir, et non-syntaxiques tous les autres... Fondée sur ce principe, la division des unités de la langue en lexicales *vs* grammaticales n'est au mieux qu'un classement heuristique provisoire, voué à fluctuer au gré de l'ingéniosité des grammairiens, et dont la pertinence descriptive n'est pas plus assurée que celle des traits sémantiques qu'il vise à mettre hors-jeu.

3.2.4. Certains chercheurs récusent d'ailleurs cette partition L/G comme faisant obstacle à des généralisations possibles. Ainsi, G. Gross (1994, 2008), constatant que les prédicats signifiés par les verbes polysémiques dépendent des traits sémantico-référentiels portés par leurs arguments, cf.

(14) suivre un chemin	≅ parcourir, longer
suivre un cours	≅ fréquenter, assister à
suivre un conseil	≅ se conformer à...

en conclut que des traits comme [+voie], [+enseignement], [+injonction]... caractérisent des sous-classes de noms qui ont pour propriété distinctive de sélectionner certains prédicats compatibles. Ces traits de sens permettent, autrement dit, de définir parmi les noms des classes de distribution plus fines que les sous-catégories fondées sur des traits syntaxiques comme [+/- animé], [+/- massif], etc. Aucune différence, sinon de granularité classificatoire, n'est donc à faire entre traits syntaxiques et sémantiques, et par suite, l'opposition L/G est sans fondement :

Le lexique ne peut pas être séparé de la syntaxe, c'est-à-dire de la combinatoire des mots. La sémantique n'est pas autonome non plus : elle est le résultat des éléments lexicaux organisés d'une façon déterminée (distribution). (Gross 2008)

3.3. Au total, ce n'est donc pas de la syntaxe qu'il faut attendre une définition satisfaisante de l'opposition L/G. Les relations de dépendance ne peuvent pas lui tenir lieu de justification empirique, et la faire découler d'une clôture théorique a priori du domaine de la syntaxe revient à prendre le champ de vision limité des grammairiens pour une propriété de la langue elle-même, ce qui n'est pas le bon moyen d'en appréhender la structure paradigmatique immanente.

4. CRITÈRES SÉMANTIQUES.

C'est sans doute au plan du contenu que les tentatives de définir l'opposition L/G ont été les plus luxuriantes. On peut les classer sous trois grands thèmes.

4.1. Charge informationnelle.

Les définitions les plus communes reflètent l'intuition que les morphèmes grammaticaux ont un signifié plus pauvre en information que les morphèmes lexicaux. Cette idée a reçu des formulations diverses.

4.1.1. Version naïve.

Elle consiste à voir dans les morphèmes [+G] des ‘mots vides’, contrairement aux mots ‘pleins’ qui composent le lexique^(C). Ce propos est bien attesté sur les sites Internet de vulgarisation ‘linguistique’ :

Tous les morphèmes grammaticaux sont des morphèmes liés (ils ne signifient rien eux-mêmes), qui représentent uniquement une catégorie grammaticale. En se liant avec un morphème lexical (qui constitue la base, le radical), les deux ensemble donnent un mot. (<https://www.docsity.com/it/la-morphologie/2428749/>)

Or, le débat n’est pas clôt (*sic*) quant au statut exact des auxiliaires. Sont-ils de réelles marques grammaticales ? N’ont-ils pas plutôt une teneur lexicale puisqu’ils ont gardé un sens en eux-mêmes, contrairement aux morphèmes grammaticaux (terminaisons *-ons, ez, etc.*) ? (https://fr.wikipedia.org/wiki/Aspect_grammatical)

La notion de mot ou morphème vide (= signe qui ne signifie pas), étant contradictoire en soi, ne mérite pas qu’on s’y attarde, du moins lorsqu’elle est prise au pied de la lettre. Elle est simplement l’aveu d’une impuissance à décrire le sens des morphèmes dits ‘grammaticaux’¹³.

4.1.2. Version lexicographique.

Elle repose sur l’affirmation que les signifiés des mots grammaticaux ne peuvent pas être définis *per genus proximum et differentiam specificam*, c’est-à-dire caractérisés par inclusion dans un genre super-ordonné¹⁴. Les unités [+G] partageraient cette propriété avec les lexèmes qui ont pour sens une notion primitive inanalysable (comme p. ex. les noms de couleur *vert, bleu...* ou les hypéronymes évaluatifs *bien, mal*), ce qui serait la preuve qu’ils ont un contenu sémique incomplexe :

Les mots grammaticaux (*car, que, etc.*) ont, comme les primitifs lexicaux, [...] un contenu pauvre, quasiment réduit à leur fonction. C’est la raison pour laquelle ils ne peuvent être définis par le système sémantique de l’inclusion... (Lehmann & Martin-Berthet 1998 : 21)

Si l’on s’en tient au principe que la charge sémique d’une unité est faite de différences oppositives, on doit cependant récuser l’idée que tous les morphèmes grammaticaux auraient un signifié réduit à un seul trait primitif, et impossible à définir ‘par inclusion’. Certes, certains

¹³ Les citations ci-dessus sont caractéristiques d’une tendance répandue à dénier toute signification aux affixes, notamment flexionnels. Cette cécité morpho-sémantique est sans doute due à la prégnance du format du mot, institué en monade significative par l’apprentissage scolaire : seuls sont perçus comme ayant un sens les lemmes qui figurent en entrées dans les dictionnaires.

¹⁴ Cette méthode de définition, couramment pratiquée dans les dictionnaires, repose sur l’idée que chaque nom dénote une classe d’objets, et qu’il existe une hiérarchie d’inclusions entre ces classes. L’espèce d’objets que dénote un nom peut donc être définie (i) par le genre qui l’inclut et (ii) par ce qui la différencie des autres espèces du même genre. Ex. :

chaise = *siège* + *à dossier et généralement sans bras* (TLFi)

où *siège* = le genre incluant, et *avec dossier, sans bras* = la différence propre à la sous-classe des chaises. Dans les sémantiques structurales, ces composants du sens sont traités comme des traits oppositifs, appelés *sèmes*. Le signifié de certains mots apparaît cependant indécomposable, et réduit à un seul sème élémentaire. Pour plus de détails, consulter Lyons (1970 : 347 *sqq.*) ou Rastier 1987.

morphèmes [+G] sont vraisemblablement mono-sémiques, comme le suffixe de pluriel /-z/, qui ne s'oppose qu'à son absence. Mais il en est d'autres pour lesquels une décomposition en sèmes, fondée sur les oppositions paradigmatiques dans lesquelles ils entrent, est tout à fait possible. Ainsi, lorsque Benveniste (1966 : 252) paraphrase *je* par '*la personne qui énonce la présente instance de discours contenant je*', il ne fait que formuler en bon français une décomposition sémique du genre {[+animé], [+personne], [+singulier], [+locuteur]}¹⁵, dont il découle entre autres que *je* est un hyponyme de *nous* (non marqué en nombre), lui-même hyponyme de *on* (non marqué comme personne). Il apparaît donc parfaitement possible de définir ces pronoms par 'inclusion' les uns dans les autres. On pourrait en faire de même avec les affixes verbaux de temps-aspect ou les relateurs de cause {*car, parce que, puisque, à cause que, vu que...*}. Simplement, il ne faut pas confondre ce type de définition avec la possibilité de l'exprimer toujours sous forme de paraphrases dictionnaires comme celles qui servent pour les noms ou pour les verbes.

4.1.3. Version diachronique.

4.1.3.1. Elle réside dans la thèse maintes fois redite que la grammaticalisation d'une unité s'accompagne d'une perte de contenu sémantique. On décrit celle-ci tantôt comme le passage du statut de mot *principal* à celui de mot *accessoire*, tantôt de façon imagée comme 'décoloration' (*bleaching*), tantôt plus techniquement comme *abstraction métaphorique* :

Le mot accessoire finit par ne plus être qu'un élément privé de sens propre, joint à un mot principal pour en marquer le rôle grammatical. (Meillet 1948 : 139)

Quant au paramètre sémantique, la grammaticalisation implique une désémantisation (Damourette-&-Pichon 1911-36) ou, pour calquer le terme de Lehmann (1982), une « javellisation » (*semantic bleaching*), c'est-à-dire que l'entrée lexicale se vide progressivement de son sens plein pour acquérir en revanche un sens fonctionnel, grammatical. [...] La désémantisation implique bel et bien un rétrécissement sémantique dans la mesure où il y a perte de traits sémantiques originaux. (Lamiroy 1999 : 35)

One of the key concepts of grammaticalization theory is desemanticization, whereby a given lexical unit loses in semantic complexity and tends to be reduced essentially to one basic semantic element. [...] The claim made in this paper is that it is metaphorical language use which is responsible for desemanticization, and hence for the rise of grammar. Once a given lexeme is made the vehicle of a metaphor, its meaning tends to be drastically reduced, and only those features will survive which are transferred to the metaphorical topic – the same what happens in the case of desemanticization. (Claudi & Heine 1986 : 313)

4.1.3.2. Ce scénario de désémantisation est fondé sur une conception référentialiste du sens, qui consiste à prendre pour des traits sémantiques pertinents les caractéristiques attribuées aux référents par la physique du sens commun. Cela conduit à privilégier comme ayant un sens 'plein' les acceptions concrètes et anthropocentriques, « grounded in the socio-physical world of reference » (Traugott 1999). Mesurés à leur aune, les emplois 'figurés' qui s'en écartent sont considérés comme plus abstraits et plus pauvres en sèmes¹⁶. Il en va de même des mots

¹⁵ Pour Benveniste, *personne* = participant à l'interlocution (1966 : 228).

¹⁶ Voir p. ex. Meillet (1948 : 136) :

Soit par exemple le mot *pied* ; employé isolément, il désigne une partie du corps humain très définie, de forme très spéciale ; groupé avec le nom d'un objet, dans des expressions comme *le*

grammaticaux, nés de métaphores, et auxquels on impute la dénotation de réalités « typically not accessible to physical experience » (Claudi & Heine 1986 : 300).

Il existe cependant des sémantiques moins rudimentaires, qui envisagent les signifiés comme de pures valeurs oppositives indépendamment de la référence (Saussure 2002 : 74 ; Le Guern 2003 : 32), ou bien qui décrivent les diverses interprétations référentielles d'une unité comme autant de spécifications apportées en contexte à une valeur de base sous-spécifiée (voir p. ex. la notion de *signifié de puissance* chez Guillaume (1969 : 246)). Si on souscrit à ce genre de sémantique, la charge sémiologique des occurrences 'propres' et 'figurées' est la même :

Il n'y a pas de différence entre le sens propre et le sens figuré des mots (ou : les mots n'ont pas plus de sens figuré que de sens propre) parce que leur sens est éminemment négatif. (Saussure 2002 : 72, manuscrit daté de 1890).

et le signifié d'un lexème, réduit aux traits invariants qu'il a dans toutes ses occurrences, apparaît tout aussi abstrait et dépouillé que celui des mots grammaticaux.

4.1.3.3. Des linguistes de diverses orientations ont ainsi été amenés à dénoncer comme abusive l'équation « sens abstrait = sens appauvri » :

Nous ne sommes pas d'accord avec cette approche du sens en termes de désémantisation. D'une certaine façon, on pourrait même dire qu'en se grammaticalisant ou en se pragmatiquisant, une unité devient souvent sémantiquement plus complexe. Cela est particulièrement frappant dans le cas des MD (= marqueurs discursifs) qui présentent souvent des sens très difficiles à décrire, non pas parce qu'ils sont peu complexes, mais parce qu'ils sont hautement abstraits. Du coup, nous rejetons la terminologie relative aux MD dont la connotation serait celle de « mots vides ». (Dostie 2004)

Dire que les mots grammaticaux « ont un contenu pauvre, quasiment réduit à leur fonction » (Lehmann & Martin-Berthet, 1998 : 21) soit apparaît comme une généralisation abusive, soit nécessite une redéfinition de la catégorie susvisée. (Cusin-Berche 1999 : 19)

Un des critères les plus usuels est le caractère « concret » ou « abstrait » du sens, l'évolution étant supposée se faire du « concret » à l'« abstrait ». Nous n'insisterons pas sur l'ambiguïté de ces termes, hérités d'une philosophie désuète. (Benveniste 1966 : 298)

It is by no means obvious, for example, that an affix conveys less information than a full form or that an element with the status of a functional category conveys less information than one with the status of a lexical category. [...] *Have* in the sense of 'perfect aspect' does not convey less information than *have* in the sense of 'possess'; it simply conveys different information. (Newmeyer 2001 : 198)

4.1.4. Version cognitive.

C'est l'idée que le contenu des mots grammaticaux se réduit à quelques traits conceptuels abstraits, parce qu'ils remplissent une fonction de formatage des informations signifiées : au plan syntagmatique, ils indiquent les 'grandes lignes' de la scène représentée par un énoncé, alors que les mots lexicaux en spécifient les détails ; au plan paradigmatique, ils expriment un cadre de rubriques cognitives stéréotypé, dans lequel viennent se mouler les représentations très spécifiées et foisonnantes véhiculées par les mots lexicaux :

pied d'une table, d'une chaise, d'une lampe ou le pied d'une montagne, le mot perd sa valeur concrète tout entière, et il n'en reste plus qu'un élément abstrait : partie d'un objet qui supporte et est en contact avec une surface portante.

The guiding principle in languages is that the more general meanings are expressed through the grammar, and the more specific meanings through the vocabulary. [Halliday & Hasan 1976 : 5)¹⁷

On voit alors ce qui distingue le grammatical du lexical : est grammatical ce qui sert dans la construction de la scène, et lexical ce qui sert, au-delà de cette construction, à évoquer les rapports avec l'univers socio-culturel et physico-anthropologique qui structure le mode d'existence des sujets parlants. (Victorri 1999 : 104)

[Dans une phrase,] The grammatical elements are more numerous, and their specifications seem more spare and simpler, and more structural in function. Together, their specifications seem to establish the main delineations of the scene organization and communicative setting of the CR [= cognitive representation] evoked by the sentence. The lexical elements are fewer in number, but their specifications are greater in quantity and complexity, and function more to contribute content than structure. The lexical specifications are greater in three ways: compared to a grammatical specification, each has (a) more total information, (b) greater intricacy of information, and (c) more different types of information together. Taken together, their specifications comprise most of the conceptual content of the CR scene that is evoked by the sentence. (Talmy 1995 : 146)

Why are grammatical elements – say, such prepositions – not a large and open class marking indefinitely many distinctions? One may speculate that the cognitive function of such classification lies in unifying contentful material within a single conceptual system and in rendering it manipulable – i.e. amenable to transmission, storage and processing – and that its absence would render content an intractable agglomeration. (*ibid.*: 167)

À cette théorie, on peut objecter, à la suite de Pottier (1987 : 156), qu'il arrive à des morphèmes lexicaux et grammaticaux d'exprimer une même notion structurante, et donc que les statuts [+L] vs [+G] ne coïncident pas toujours avec des pertinences cognitives différentes. Ex. :

Notion	Expressions [+L]	Expressions [+G]
<être humain>	<i>une personne</i>	<i>quelqu'un, quiconque</i>
modalité <◇>	<i>possible, peut-être</i>	morphème /Rε/ de conditionnel
<accompli>	<i>ancien</i>	préfixe <i>ex-</i> , morphème de participe passé
<futur>	<i>futur, avenir</i>	morphème /R/ de futur
<négation>	<i>faux</i>	<i>ne, pas, non</i> , préfixe <i>in-</i>
<diminutif>	<i>petit</i>	suffixes <i>-et, -ette...</i>
<sexe ♀>	<i>femme, femelle</i>	suffixes <i>-e, -esse...</i>

Bien sûr, les expressions mises en regard dans ce tableau ne sont pas des synonymes exacts, ce qui explique et justifie leur coexistence dans la langue. On peut néanmoins soutenir qu'elles ont le même degré d'abstraction et la même pertinence cognitive, c'est-à-dire contribuent à structurer de la même façon les représentations des 'scènes' communiquées.

¹⁷ Halliday & Hasan précisent toutefois que les unités [+L] et [+G] contribuent les unes comme les autres à assurer la fonction cohésive dans les discours, et que la frontière entre les deux types d'unités n'est pas nette. Cela les conduit à associer systématiquement lexicale et grammaire sous le nom de codage *lexicogrammatical* des significations, laissant ainsi entendre que la distinction L/G n'a qu'une importance secondaire à leurs yeux.

4.1.5. Au total, de quelque façon qu'on la prenne, la thèse que les mots grammaticaux exprimeraient par nature moins d'information que les lexèmes ne peut donc pas être retenue.

4.2. Sens +/- référentiel.

4.2.1. Une autre tentative de fonder l'opposition L/G sur la sémantique consiste à décrire les mots et morphèmes [+G] comme porteurs d'informations non référentielles, contrairement aux lexèmes [+L] :

Parmi les unités linguistiques, on distingue deux catégories, qui s'opposent par leur façon de faire sens : d'une part, il y a les mots [...] qui, même en dehors de tout emploi dans un énoncé, évoquent une réalité. [...] D'autres mots [...] ne réfèrent pas. (Mortureux 1997)

C'est cette thèse, notamment, que Tesnière exprime en termes de 'mots pleins' vs 'vides' :

Les mots pleins sont ceux qui sont chargés d'une fonction sémantique, c'est-à-dire ceux dont la forme est associée directement à une idée, qu'elle a pour fonction de représenter et d'évoquer. Ainsi fr. *cheval*, all. *Pferd*, angl. *horse*, lat. *equus*, etc. sont des mots pleins, parce que leur forme [...] suffit à évoquer l'idée d'un cheval.

Les mots vides sont ceux qui ne sont pas chargés d'une fonction sémantique. Ce sont de simples outils grammaticaux dont le rôle est uniquement d'indiquer, de préciser ou de transformer la catégorie des mots pleins et de régler leurs rapports entre eux. Ainsi, fr. *le*, all. *das*, angl. *the*, etc. sont des mots vides, parce qu'ils n'évoquent rien en eux-mêmes et qu'ils ne servent qu'à déterminer les catégories dont relèvent les mots *cheval* (substantif, masculin, singulier), *Pferd* (substantif, neutre, singulier, nominatif), *horse* (substantif, singulier). (Tesnière, 1969 : 53)

4.2.2. Cette conception de la référence comme évocation d'une réalité ou d'une 'idée' de chose est extrêmement rudimentaire, et se voit infirmée par la moindre analyse un peu précise. Les nombreux travaux portant sur la sémantique des déterminants ont montré qu'ils étaient, quoi qu'en dise Tesnière, porteurs d'instructions essentiellement référentielles (repérages déictiques, quantifications). Et il est impossible de soutenir que les pronoms clitiques qui désignent des personnes (*je*, *nous*), les désinences verbales qui désignent des moments du temps (*parlait/era*), les indices qui désignent des lieux (*-ci*, *-là*) et les prépositions spatiales (*sur*, *devant*) ne référerait pas, ou dénoteraient des réalités « not accessible to physical experience ». En sens inverse, certains sémanticiens ont fait observer que bien des lexèmes ne sont pas susceptibles de recevoir une définition référentielle en termes de conditions de vérité. C'est l'impossibilité de décrire ainsi le sens de mots comme *courage*, *prudent*, *sage* ou *vertu* qui a donné naissance aux sémantiques ascriptivistes et argumentatives, selon lesquelles les significations manifestées par la langue ne sont pas des descriptions du monde (Anscombe & Ducrot 1983 : 169).

Bref, la distinction [+/- référentiel] plaquée sur l'opposition L/G apparaît mal dégrossie, et ne saurait tenir lieu de critère de définition décidable. Il est en outre à noter qu'un tel critère serait incompatible avec une conception graduelle de l'opposition : ou bien un signe réfère à quelque chose, ou bien il ne réfère pas, mais on ne saurait concevoir que la fonction dénotative s'exerce à divers degrés, et qu'un signe puisse « passer d'un sens plus référentiel à un sens moins référentiel » (Lamiroy 1999 : 35).

4.3. Sens +/- fonctionnel.

4.3.1. Cette caractérisation, qui n'est pas toujours clairement distinguée de la précédente, s'exprime à travers plusieurs variantes terminologiques.

- L'une est la distinction entre mots *catégorématiques* vs *syncatégorématiques* :

Les catégories lexicales ont plutôt une fonction catégorématique alors que les catégories grammaticales ont plutôt une fonction syncatégorématique. (Fradin 2003 : 89)

Ces termes sont empruntés aux logiques médiévales¹⁸, qui nomment *catégorèmes* les signes aptes à exprimer par eux-mêmes le sujet ou le prédicat d'une proposition, et *syncatégorèmes* tous les autres signes, auxquels on ne trouve pas d'autre fonction que celle d'ajouter quelque précision à un catégorème :

Les termes syncatégorématiques [...] n'ont pas de signification définie et bien déterminée, et ne signifient pas des choses distinctes de celles que signifient les catégorèmes. Au contraire [...], un syncatégorème ne signifie rien à proprement parler, mais ajouté à un autre terme, il fait que celui-ci signifie quelque chose ou représente (*supponit pro*) une ou plusieurs choses sur un mode déterminé ; ou bien il exerce quelque fonction envers le catégorème. (Ockham, *Summa logicae*, I, 4) [Exemple ¹⁹]

- La même distinction est parfois formulée, à la suite de Marty (Fryba-Reber 2006), en opposant les mots *autosémantiques*, qui expriment par eux-mêmes une représentation (*selbstbedeutend*), aux mots *syntagmatiques*, qui ne signifient qu'associés à d'autres mots (*bloss mitbedeutend*) :

Bien qu'appartenant au lexique, les auxiliaires sont moins *autosémantiques* (Heine 1991 : 28) que les entrées lexicales pleines : alors que les entrées lexicales pleines ont leur propre contenu sémantique, les auxiliaires auront tendance à constituer un sens en combinaison avec un autre élément. (Lamiroy 1999 : 38)

- On trouve encore l'opposition L/G revêtue des qualificatifs *descriptif* vs *instructionnel* :

Certains mots dits *vides* n'ont pas de sens par eux-mêmes : il s'agit surtout des mots grammaticaux (prépositions, conjonctions, déterminants, pronoms) opposés aux mots lexicaux ou mots pleins [...]. Ceux-ci ont un sens plus autonome et une référence : ils renvoient à un objet ou à une notion définissable. [...]

Dans une terminologie plus récente, on oppose *sens instructionnel* (consistant plutôt en « instructions » guidant la mise au point du sens des mots dans la phrase) et *sens référentiel* ou *descriptif* (consistant à décrire les propriétés d'un référent). (Lehmann & Martin-Berthet 2018)

La notion de signifié *instructionnel* a été introduite par Ducrot (1980 : 12) pour décrire des unités linguistiques dont le sens contient un paramètre à instancier à partir des informations fournies par la situation de discours, et qui donnent en quelque sorte l'instruction d'inférer la valeur de ce paramètre. Par extension, le terme d'*instructionnel* s'est vu ensuite appliqué à des signifiés contenant une variable quelconque, p. ex. aux affixes dérivationnels, qui expriment des

¹⁸ Sur les origines antiques de ces notions, voir Lallot 2003.

¹⁹ Exemple donné par Ockham (*ibid.*) : les noms *homo* et *animal*, qui peuvent être les termes (= sujet et prédicat) d'une proposition (*homo est animal*), sont des catégorèmes, tandis que le déterminant *omnis*, qui ne désigne rien en soi mais s'associe à un nom pour désigner quelque chose (*omnis homo est animal*), est un syncatégorème.

fonctions entre classes de prédicats (V-*eur* = ‘agent qui V’ ; V-*oir* = ‘instrument qui sert à V’ ; A-*issime* = ‘qui est A au plus haut degré’...) :

[Les unités du lexique se classent d’après] la nature de leur sens, prioritairement descriptif ou instructionnel : si les propriétés sémantiques de l’unité décrivent une catégorie conceptuelle, elles ont un sens descriptif, si elles donnent à voir d’une certaine façon un référent, elles ont un sens instructionnel (sur ce point, cf. Kleiber 1999, chap. II). Les affixes sont des unités infralexicales à sens instructionnel. (Corbin 1999 : 70)

4.3.2. Ce que décrivent identiquement les qualifications de *syncatégorématique*, *syntagmatique* ou *instructionnel* (et peut-être aussi celle plus vague de sens *fonctionnel*), c’est le fait que certains signes ne sont interprétables qu’en association avec d’autres. On a affaire en quelque sorte à un critère de dépendance appliqué au plan du contenu. Il conduit à distinguer dans le lexique^(B) deux sortes d’éléments. Les uns, sémantiquement auto-suffisants, ont un sens assimilable, en termes de mathématique, à une constante. Les autres expriment des fonctions sur ces constantes : leur contenu comporte une variable qui doit recevoir sa valeur d’un signe cooccurrent exprimant une constante.

4.3.3. Appliqué de manière systématique, ce critère conduit cependant à exclure du lexique^(C) une bonne partie de ce qu’on y met d’habitude. Ainsi, on sait depuis Beauzée que de nombreux lexèmes signifient une relation qui ne s’entend que complétée par un terme, autrement dit une fonction d’une variable :

Les mots qui ont une signification relative, exigent un complément, dès qu’il faut déterminer l’idée générale de la relation par celle d’un terme conséquent : & tels sont plusieurs noms appellatifs, plusieurs adjectifs, quelques adverbes, tous les verbes actifs relatifs & quelques autres, & toutes les prépositions. Exemples de noms relatifs : *le fondateur de Rome* [...], *le mari de Lucrèce*, etc. Exemples d’adjectifs relatifs : *nécessaire à la vie*, *digne de louange* [...] Exemples de verbes relatifs : *aimer Dieu*, *craindre sa justice*, *aller à la ville* [...]. Exemples d’adverbes relatifs : *relativement à vos intérêts*, *indépendamment des circonstances*, etc. (Beauzée, 1756)

Selon une tradition convergente qui remonte à Frege, tous les verbes sont des syncatégorèmes, car ils ont pour sens des fonctions ‘insaturées’, dont les places d’argument doivent être remplies par des noms d’objets pour constituer des propositions interprétables (Desclés 2009). Et certains sémanticiens, constatant que le sens d’un adjectif épithète dépend fortement du nom qu’il modifie (cf. *un bon vin* / *un bon couteau* / *un bon camarade*...), en ont conclu que les adjectifs sont eux aussi des mots syncatégorématiques (Goes 2021). En somme, les supposées constantes n’étant elles-mêmes que des fonctions, la non-autonomie sémantique s’avère être le lot commun, et ne peut plus servir de discriminant à l’appui d’une opposition L/G.

4.4. Au terme de cette revue des tentatives faites pour fonder en sémantique l’opposition L/G, on ne peut que conclure à leur échec :

Ce qu’il faut souligner ici, c’est qu’il ne semble pas y avoir de différence essentielle entre le type de sens qui est associé aux unités lexicales et le type de sens qui est associé aux unités grammaticales, là où la distinction peut être faite entre ces deux classes d’éléments. [Lyons 1968=1970 : 336]

5. CRITÈRES MORPHO-PHONOLOGIQUES.

On doit encore mentionner, pour mémoire, certaines saisies de l'opposition L/G à partir de caractéristiques purement formelles.

5.1. *Longueur syllabique.*

Riegel & al. signalent que :

contrairement aux morphèmes lexicaux indifféremment courts ou longs (*an, dé, portail, maréchal, mathématique*, etc.), les morphèmes grammaticaux sont le plus souvent très courts et monosyllabiques (*-s, -ons, à, si*, etc.). (2009 : 536)

Comme on ne dispose pas de comptages précis permettant de dénombrer les morphèmes en fonction de leur longueur, il n'est pas possible de décider de manière sûre s'il existe ou non une corrélation entre celle-ci et les statuts L/G usuels. Il est toutefois évident qu'une bonne partie des morphèmes [+L] sont monosyllabiques : Sauvageot (1972) estime les 'mots primaires' du type *eau, tour, dos, pied, feuille* à 20% du total des monosyllabes de la langue. Et il ne manque pas non plus, à l'inverse, de morphèmes [+G] pluri-syllabiques, p. ex. parmi les préfixes {*anti, contra, pluri, multi, intra, inter, supra, infra, juxta...*).

5.2. *Accentuation.*

Une autre propriété formelle parfois mise en rapport avec l'opposition L/G est l'autonomie accentuelle. La plupart des théories de l'accent admettent en effet une distinction entre deux types de morphèmes : les uns, « accentogènes », peuvent librement porter un accent sur la finale, ce qui leur permet de constituer à eux seuls un groupe accentuel ; les autres, clitiques, ne le peuvent pas, et doivent nécessairement s'intégrer au même groupe accentuel qu'un mot accentogène adjacent. Cette distinction a pu être identifiée à l'opposition L/G :

Il est également d'usage de distinguer entre mots accentogènes et mots non accentogènes. À la différence des premiers (noms propres, substantifs, adjectifs, etc.), les seconds (en particulier les clitiques) ne peuvent en principe pas recevoir un accent. En règle générale, il a été admis que les mots pourvus d'un contenu sémantique sont accentogènes, alors que les mots fonctionnels ne possèdent pas cette propriété. (Di Cristo 2013 : 5)

Di Cristo avertit cependant que « cette affirmation doit être révisée lorsqu'on étudie la prosodie du langage en usage ». De fait, la plupart des morphèmes clitiques peuvent à l'occasion former un groupe accentuel autonome (occurrences frappées d'un accent focalisateur, isolées par une insertion parenthétique, ou constituant à elles seules un énoncé elliptique, v. ex. (9-11) *supra*). D'autre part, si l'on range au nombre des morphèmes [+G] les connecteurs, les marqueurs discursifs, les verbes copules et les auxiliaires, il y a parmi eux des items régulièrement accentogènes (*donc, pouvoir, eh ben...*).

5.3. À vrai dire, la longueur syllabique et l'autonomie accentuelle, lorsqu'elles sont évoquées à l'appui de l'opposition L/G, sont présentées comme de simples éléments de signalement qui valent pour les cas prototypiques, mais ne prétendent pas constituer des critères de définition à portée générale. C'est ainsi qu'il faut les prendre.

6. BILAN.

6.1. *Une opposition vague.*

Au terme de cet inventaire, le moins qu'on puisse dire est que la pertinence de l'opposition L/G apparaît problématique.

6.1.1. Les critères sur lesquels on a tenté de la fonder procèdent ou bien d'idées reçues qui ne résistent pas à un examen attentif des faits, ou bien de présupposés théoriques qui encourent des objections. Certains d'entre eux ne correspondent à aucune réalité (mots 'pleins' vs 'vides'). En outre, il arrive qu'ils se contredisent (§ 222, 314).

6.1.2. Ces critères, si on les applique avec précision, ne coïncident généralement pas avec la partition L/G communément admise : des deux côtés, il y a des classes finies dont certaines sont peu nombreuses ; des deux côtés figurent des unités fréquentes et d'autres non ; des deux côtés se trouvent des formes syntaxiquement liées, des termes syncatégorématiques, des signifiés cognitivement structurants, des monosyllabes et des segments accentuables... On a affaire à des traits de signalement qui, additionnés, permettent en gros de contraster deux sortes d'unités prototypiques (ou l'idée qu'on en a) : d'une part les noms concrets, pris pour modèles des lexèmes, et d'autre part les affixes flexionnels, tenus pour parangons des grammèmes. Mais ni chacun de ces traits pris individuellement, ni leur conjonction, ne constituent une définition de l'opposition L/G qui soit généralisable à la totalité du lexique^(A ou B), et qui en induise une division conforme aux pratiques usuelles.

6.1.3. Postuler une échelle sur laquelle les éléments du lexique^(A ou B) se répartiraient selon le degré auquel ils vérifient les divers critères n'est pas une réponse satisfaisante à la déficience de ceux-ci. D'une part en effet, certains de ces critères ne sont pas graduables (§§ 2124, 422). D'autre part, faire cette hypothèse, c'est se mettre en devoir de déterminer quelle position occupe chaque unité sur l'axe L/G. Or, personne n'est en mesure de donner une réponse à cette question, ni en valeurs absolues ni en valeurs relatives. Comme le souligne Prévost :

Il est parfois difficile de décider si une forme est plus grammaticale (fonctionnelle, pour reprendre le terme de Haspelmath) qu'une autre. [...] Il n'en serait pas moins souhaitable, d'un point de vue « théorique », de parvenir à une échelle consensuelle des degrés de grammaticalité. Il est en effet gênant de définir la grammaticalisation comme l'acquisition par une forme linguistique d'un statut plus grammatical, sans pouvoir décider pour l'ensemble des formes lesquelles sont plus grammaticales que les autres. (2003 : 152)

Dans l'état actuel des connaissances, définir l'opposition L/G comme un continuum revient simplement à reconnaître l'existence, entre les catégories prototypiques qui en sont les pôles, d'un résidu de catégories tierces dont le degré de lexicalité / grammaticalité ne peut être déterminé :

On a l'impression (et là encore, cette situation est fréquente en linguistique) d'avoir affaire à un continuum, avec deux pôles qui s'opposent clairement, et une série de cas intermédiaires pour lesquels toute coupure franche comporte une part inévitable d'arbitraire. (Victorri 1999 : 85).

En d'autres termes, les cas dits 'intermédiaires' sont des unités auxquelles le couple de notions L/G ne peut pas s'appliquer de manière décidable, et qui montrent donc son absence de généralité.

6.1.4. On a là un exemple caractéristique d'importation dans le discours grammatical de *notions de catégorisation pratique* (Berrendonner & (Reichler)-Béguelin 1989) directement empruntées au métalangage commun, et qui sont en quête d'une définition scientifique a posteriori, qu'on ne parvient pas à leur trouver. Les notions de ce type sont destinées à penser les pratiques langagières courantes du groupe social (cf. en l'occurrence le couple dictionnaire + grammaire comme outillage pour apprendre les langues). Elles sont fondées sur des prototypes saillants, et non sur toute la diversité des données observables ; elles sont multicritères et définies par un lot de propriétés non équivalentes, ce qui les rend malléables et donc adaptables à tout fait nouveau rencontré. Cette plasticité en fait des outils commodes pour gérer les usages quotidiens de la langue, mais a une contrepartie fâcheuse : ce sont des notions non consistantes qui, si on les utilise comme moyens de modélisation, importent avec elles dans le discours grammatical des contradictions et des défauts de généralité. De telles notions, employées à des fins descriptives, constituent un outil de catégorisation de premier abord, préalable à la mise en œuvre de différenciations plus précises et méthodologiquement plus solides (Gomila 2013). Il semble qu'avec le couple L/G, on en soit resté à ce stade préscientifique, celui d'une 'linguistique provisoire' (Milner 1989 : 65).

6.2. Conséquences.

6.2.1. Il n'est donc pas étonnant que divers linguistes aient été amenés à récuser la pertinence de l'opposition L/G, et à l'abandonner.

- Pour Saussure,

Quantité de rapports exprimés dans certaines langues par des cas ou des prépositions sont rendus dans d'autres par des composés [...], ou par des dérivés [...] ou enfin par des mots simples. [...] On voit donc qu'au point de vue de la fonction, le fait lexicologique peut se confondre avec le fait syntaxique. [...] En résumé, les divisions traditionnelles de la grammaire peuvent avoir leur utilité pratique, mais ne correspondent pas à des divisions naturelles et ne sont unies par aucun lien logique. (CLG/E 187, 2144-2149)

- Martinet, après avoir admis en 1960 une opposition entre *morphèmes* (= pour lui, signes grammaticaux) et *lexèmes* (1960 § 19), fait en 1985 son *mea culpa* :

On a eu, depuis, maintes occasions de regretter d'avoir, sur ce point, cédé à la pression d'une certaine tradition [...] Dans la remise en question des fondements de l'analyse linguistique qu'entraîne nécessairement un traitement fonctionnel, il ne saurait être question de poser au départ une distinction entre fait grammatical et fait lexical. (1985 : 28-29)

- Ont aussi été conçues des théories de la grammaire qui ignorent tout simplement la dichotomie L/G, et lui substituent une paradigmatique générale fondée sur d'autres critères. C'est le cas des grammaires catégorielles, pour lesquelles le stock des unités linguistiques se divise en opérateurs vs opérands absolus. L'ensemble des opérateurs subsume à la fois des morphèmes et des assemblages syntagmatiques de tous rangs, ce qui permet de se débarrasser de la notion de lexique^(B) et du vague qu'elle traîne avec elle. L'assignation aux éléments du lexique^(A) des

statuts d'opérateur ou d'opérande absolu repose sur des postulats ou des hypothèses d'analyse explicites, dont la vertu généralisante peut être discutée par confrontation avec les données empiriques.

- Une autre façon de rejeter l'opposition L/G réside dans la thèse que toute unité linguistique possède à la fois des propriétés grammaticales et des propriétés lexicales :

Un mot comme *chien*, par exemple, présente à la fois des aspects grammaticaux et lexicaux. En tant que nom comptable (caractéristique qu'il partage avec des mots comme *chat*, *table*, *arbre*, etc.), il relève du sous-système grammatical. Et c'est uniquement en tant que lexème possédant un contenu sémantique qui l'oppose aux autres noms comptables qu'il fait partie du sous-système lexical. (Victorri 2004 : 77)

Cette idée, originaire des grammaires cognitives (Langacker 1987 : 3), revient à déplacer le couple de notions *lexical / grammatical* du plan extensionnel au plan intensionnel. Ces termes ne servent plus à opposer deux classes de signes, mais deux sortes de propriétés (combinatoires vs représentationnelles) que possèdent toutes les unités linguistiques, ainsi d'ailleurs que des « éléments plus abstraits, tels que les catégories grammaticales (parties du discours), les sous-catégories [...], les fonctions syntaxiques (sujet, objet, etc.) et les constructions syntaxiques (l'ordre des mots notamment) » (Victorri, *ibid.*). Cette conception, dont ont hérité les grammaires constructionnelles, revient à dénier au couple L/G toute pertinence classificatoire : on ne peut fonder sur lui une partition du lexique global.

6.2.2. Une autre conséquence mérite d'être soulignée : la notion de *grammaticalisation*, qui présuppose la validité de l'opposition L/G, est fondée sur une primitive vague et dépourvue de réelle pertinence descriptive. Elle est donc elle-même un outil descriptif vague et d'une pertinence contestable (ce qui explique la prolifération des discussions byzantines portant sur sa définition et son application aux faits). Si elle permet de décrire à gros traits des changements diachroniques envisagés à longue distance et dont les points de départ et d'arrivée peuvent être caractérisés en termes de parties du discours prototypiques, l'ériger en scénario général des phénomènes de transcatégorisation ne peut se faire qu'au prix d'approximations ou de décisions arbitraires, étant donné que souvent, les degrés de lexicalité / grammaticalité des unités en jeu ne sont tout simplement pas décidables (lorsqu'un subordonnant évolue en coordonnant ou en 'connecteur', lequel des deux usages est le plus grammatical ? (Fruyt, 2008)). Pour une critique, voir entre autres Campbell (2001), Newmeyer (2001), Guérin (2007), Béguelin (2010, 2014).

6.2.3. Au total, si on a le souci d'éviter autant que possible le vague et les inconsistances dans les grammaires, il serait donc prudent :

- de réserver le terme de *lexique* pour désigner l'ensemble des unités significatives minimales de la langue (lexique^(A)), seule réalité paradigmatique claire et distincte ;
- de s'affranchir de la notion vague de lexique^(B), et par conséquent de révoquer en doute la notion de *mot* avec qui elle va de pair. (Cela conduit du même coup à remettre en question la frontière institutionnelle entre morphologie et syntaxe comme deux composantes étanches de la grammaire).
- de reléguer l'opposition lexical / grammatical au magasin des accessoires préscolaires périmés, et de chercher pour la remplacer des critères de classement plus pertinents.

- d'en tirer les conséquences sur le manque de fermeté de la notion de grammaticalisation utilisée pour décrire les phénomènes diachroniques de transcatégorisation.

6.3. Étude à faire.

Partis à la recherche d'une réalité empirique qui corresponde au couple L/G et qui lui confère une pertinence descriptive, nous n'avons rencontré qu'un seul fait de langue qui soit caractérisable avec précision et mérite donc d'être pris en compte dans une théorie de la grammaire. Ce sont les phénomènes de dérivation morphologique (§ 2124), au regard desquels il semble bien que le lexique^(A) se divise en deux sous-ensembles. D'une part, certains morphèmes appartiennent à des classes de distribution qui contiennent aussi des assemblages construits par dérivation, p. ex. *forêt* ∈ {*forêt*, *sapin-eraie*, *hêtr-aie*, *clair-ière*, *plant-ation...*}, etc. Ces classes sont celles que l'on regroupe ordinairement sous le nom de 'catégories majeures'²⁰. D'autres morphèmes, au contraire, appartiennent à des classes de distribution qui ne contiennent pas de dérivés (mais éventuellement des assemblages construits par d'autres moyens). Cette dichotomie n'encourt pas le reproche d'être inconsistante ou vague, pour autant que la classe des opérateurs de dérivation ait été clairement délimitée au préalable, ce qui ne semble pas hors de portée. Elle est tout ce qui reste de l'opposition L/G, et mieux vaudrait, pour éviter toute équivoque sur le terme de *lexique*, lui trouver un autre nom (catégories *dérivantes* vs *non dérivantes* ?). Elle ne coïncide ni avec une différence d'abstraction des signifiés exprimés, ni avec la distinction entre expressions ouvertes vs expressions saturées, ni avec une disparité fondamentale de comportements syntaxiques (dépendances, flexions), ni avec quelque autre différence fonctionnelle discernable. Expliquer cette division du travail entre deux classes de signes dans le système de la langue française reste donc une question à résoudre, et qu'il serait notamment intéressant d'envisager sous l'angle typologique.

7. RÉFÉRENCES BIBLIOGRAPHIQUES IMPORTANTES.

Il n'existe à ma connaissance aucune étude qui soit consacrée spécifiquement à l'opposition L/G, et qui puisse être recommandée en tant qu'ouvrage de référence. Il n'existe pas non plus d'étude représentative d'un point de vue critique sur cette opposition.

8. OUVRAGES CITÉS.

Abeillé (A.), 1993, *Les nouvelles syntaxes. Grammaires d'unification et analyse du français*, Paris, A. Colin.

Amiot (D.), 2021, « Procédés morphologiques de création lexicale », in *GGHF*, 1894-1927.

Anscombe (J.-C.) & Ducrot (O.), 1983, *L'argumentation dans la langue*, Paris, Mardaga.

Apothéloz (D.), 2002, *La construction du lexique français*, Paris, Ophrys.

²⁰ Certaines des classes de distribution rangées sous les catégories majeures N, V, Adj contiennent à la fois des dérivés et des éléments à partir desquels sont construits des dérivés. La catégorie des adverbes s'en distingue cependant : elle comprend des dérivés, mais ses membres ne peuvent apparemment pas être bases de dérivation (Amiot 2021 : 1894).

- Arrivé (M.), Gadet (F.) & Galmiche (M.), 1986, *La grammaire d'aujourd'hui. Guide alphabétique de linguistique française*, Paris, Flammarion.
- Babin (J.-P.), 1998, *Lexique mental et morphologie lexicale*, Berne, P. Lang.
- Bally (Ch.), 1965, *Linguistique générale et linguistique française*, Bern, Francke.
- Beauzée (N.), 1756, Article *Régime* in *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers*, Paris.
- Béguelin (M.-J.), 2010, « Le statut des *identités diachroniques* dans la théorie saussurienne : une critique anticipée du concept de *grammaticalisation* ». in *Le projet de Ferdinand de Saussure*, Bronckart, J.-P., C. Bota & E. Bulea eds., Genève, Droz, 237-267.
- Béguelin (M.-J.), 2014, « Deux points de vue sur le changement linguistique », *Langages* 196, 13-36.
- Benveniste (É.), 1966, *Problèmes de linguistique générale I*, Paris, Gallimard.
- Benveniste (É.), 1974, *Problèmes de linguistique générale II*, Paris, Gallimard.
- Berrendonner (A.) & (Reichler-)Béguelin, 1989, « Décalages : les niveaux de l'analyse linguistique », *Langue française*, n° 81, 99-125.
- Blanche-Benveniste (C.), 1981, « La complémentation verbale : valence, rection et associés », *Recherches sur le Français Parlé*, n° 3, 57-98.
- Blanche-Benveniste (C.), 1986, « Une chose dans la syntaxe verbale », *Recherches sur le Français Parlé*, n° 7, 141-168.
- Blanche-Benveniste (C.), 1990, *Le français parlé, études grammaticales*, Paris, Éditions du CNRS.
- Blanche-Benveniste (C.), 1993, « Les unités : langue écrite, langue orale », *Proceedings of the workshop on orality vs literacy*, C. Pontecorvo & C. Blanche-Benveniste eds., Siena.
- Blanche-Benveniste (C.), Rouget (C.) & Sabio (F.), 2002, *Choix de textes de français parlé*, Paris, Champion.
- Bloomfield (L.), 1933=1970, *Le langage*, Paris, Payot.
- Bolly (C.), 2010, « Flou phraséologique, quasi-grammaticalisation et pseudo-marqueurs de discours : un no-mans-land entre syntaxe et discours ? », *Linx* 62-63, 11-38.
- Bonami (O.), 2014, *La structure fine des paradigmes de flexion. Études de morphologie descriptive, théorique et formelle*, Paris, Univ. Paris-7, en ligne : <https://theses.hal.science/tel-01531860>
- Boye (K.) & Harder (P.), 2020, « Dual processing in a functional-cognitive theory of grammar and its neurocognitive basis », in *Grammar and Cognition*, A. Haselow & G. Kaltenböck eds., Amsterdam, Benjamins, 133-155.
- Brunot (F.), 1922, *La pensée et la langue*, Paris, Masson.
- Bybee (J.-L.), Pagliuca (W.) & Perkins (R.-D.), 1991, « Back to the Future », in *Approaches to Grammaticalization*, E. Traugott & B. Heine eds., Amsterdam, Benjamins.

- Campbell (L.), 2001, « What's wrong with grammaticalization ? » *Language Sciences* n° 23, 113-161.
- Carlier (A.) & Prévost (S.), 2021, « Constructions, constructionnalisation et changement linguistique », *Langue française* n° 209, 9-22.
- Chomsky (N.), 1957=1969, *Structures syntaxiques*, Paris, Seuil.
- Chomsky (N.), 1965=1971, *Aspects de la théorie syntaxique*, Paris, Seuil.
- Claudi (U.) & Heine (B.), 1986, « On the metaphorical base of grammar », *Studies in Language* 10-2, 297-335.
- Clech (S.), 2005, « Étude distributionnelle de *avec* et *sans* dans un corpus de 10 millions de mots : syntaxe, lexicque, sémantique », in P. Vergely (éd.) : *Rôle et place des corpus en linguistique*, Toulouse, univ. Le Mirail.
- CLG/E = Saussure (F. de), 1915, *Cours de linguistique générale*, Ch. Bally, A. Sechehaye & A. Riedlinger édés, Paris, Payot.
- Corbin (D.), 1987, *Morphologie dérivationnelle et structuration du lexique*, Lille, PUL.
- Corbin (D.) 1999, « Pour une théorie sémantique de la catégorisation affixale ». *Faits de langue* n° 14, 55-77.
- Cusin-Berche (F.), 1999, « Le lexique en mouvement : création lexicale et production sémantique », *Langages* 136, 5-26.
- Dell (F.), 1979, « La morphologie dérivationnelle du français et l'organisation de la composante lexicale en grammaire générative », *Revue Romane* 14/2.
- Desclés (J.-P.), 2009, « Le concept d'opérateur en linguistique », *Histoire, Épistémologie, Langage* n° 31/1, 75-98.
- Di Cristo (A.), 2013, *La prosodie de la parole*, Bruxelles, De Boeck.
- Diewald (G.), 1997, *Grammatikalisierung. Eine Einführung in Sein und Werden grammatischer Formen*, Tübingen, Niemeyer.
- Diewald (G.), 2004, « Pragmaticalization defined as grammaticalization of discourse functions », *Linguistics* 49-2, 365-390.
- Dostie (G.), 2004, *Pragmaticalisation et marqueurs discursifs*, Bruxelles, De Boeck..
- Ducrot (O.), 1980, *Les mots du discours*, Paris, Minuit.
- Emonds (J.), 2000, *Lexicon and grammar : the English Syntacticon*, Berlin, Mouton-de Gruyter.
- Emonds (J.), 2013, « Permissive Subordination : a Language-Particular Grammatical Option » *Studia Universitatis Babeş-Bolyai Philologia* (LVIII, 4, 2013). 111-129.
- Eriksson (O.), 2006, « Un cas de suppléance verbale en français comme illustration d'une méthode d'analyse linguistique », *Revue belge de philologie et d'histoire*, tome 84/3, 913-928.
- Eynde (K. van den) & Mertens (P.), 2010, *Dicovalence*, En ligne : <https://www.ortolang.fr/market/lexicons/dicovalence/v1>
- Fradin (B.), 2003, *Nouvelles approches de la morphologie*, Paris, PUF.

- Fruyt (M.), 2008, « Adverbes latins, grammaticalisation et lexicalisation », in *Adverbes et évolution linguistique en latin*, M. Fruyt & S. van Laer édés., Paris, L'harmattan, 49-66.
- Fryba-Reber (A.), 2006, « La postérité de Brentano et la linguistique suisse : Anton Marty (1847-1914) », *Le genre humain* n° 45-46, 491-522.
- Gardes-Tamine (J.), 1988, *La Grammaire (I)*, Paris, A. Colin.
- GGF = Abeillé (A.), Godard (D.), Delaveau (A.) & Gautier (A.), 2021, *Grande Grammaire du Français*, Paris, Actes sud.
- GGHF = Marchello-Nizia (C.), Combettes (B.), Prévost (S.) & Scheer (T.) (édés.), 2021, *Grande grammaire historique du français*, Berlin, De Gruyter.
- Glessgen (M.-D.), 2011, « Le statut épistémologique du lexème », *Revue de Linguistique Romane* 75, 391-468. En ligne : www.rose.uzh.ch/dam/jcr:e634a625-301f-4ac3-beb0-3d15e41262c4/D64_2011_Lexeme.pdf
- Goes (J.), « L'adjectif, une partie du discours éminemment syncatégorématique », *Kalbotyra* n° 74, 72-87.
- Gomila (C.), 2013, « Le petit mot de la classe : entre catégorisation pratique et classification grammaticale », in *Enseigner la grammaire*, O. Bertrand & I. Schaffner édés., Paris, Éditions de l'école polytechnique, 145-158.
- Gross (G.), 2008, « Le modèle des classes d'objets », in *Modèles syntaxiques. La syntaxe à l'aube du XXIe siècle*, D. Van Raemdonck éd., Bruxelles, P. Lang.
- Guérin (F.), 2007, « La grammaticalisation : théorie ou épiphénomène ? », *Contextos* XXV-XXVI, 211-232. En ligne : <http://www.revistacontextos.es/Contextos%2049-52/Guerin.pdf>
- Guillaume (G.), 1969, *Langage et science du langage*, Paris, Nizet.
- Halliday (M.) & Hasan (R.), 1976, *Cohesion in English*, London, Longman.
- Haspelmath (M.), 1998, « Does grammaticalization need reanalysis ? » *Studies in Language* 22, 49-85.
- Haspelmath (M.), 2011, « The indeterminacy of word segmentation and the nature of morphology and syntax », *Folia linguistica* 45, 1.
- Heine (B.), 1993, *Auxiliaries: Cognitive Forces and Grammaticalization*, New-York, Oxford University Press.
- Hopper (P.-J.) & Traugott (E.-C.), 1993, *Grammaticalization*, Cambridge, Cambridge University Press.
- Huyghe (R.), 2018, « Généralité sémantique et portage propositionnel : le cas de *fait* », *Langue française* n° 198, 35-50.
- Kaufmann (A.), 1977, *Introduction à la théorie des sous-ensembles flous*, Paris, Masson.
- Kilani-Schoch (M.) & Dressler (W.), 2004, « Compositionnalité et iconicité dans la flexion du verbe français. Une approche de morphologie naturelle », *Verbum* XXVI, 421-437.
- Kleiber (G.), 1990, *La sémantique du prototype*, Paris, PUF.

- Kleiber (G.), 1994, *Nominales. Essais de sémantique référentielle*, Paris, A. Colin.
- Kuryłowicz (J.), 1965, « The evolution of grammatical categories », *Diogenes* 51, 55-71 repris dans 1975, *Esquisses linguistiques*, vol II, Munich, Finck.
- Laks (B.), 2005, « La liaison et l'illusion », *Langages* 158, 101-125.
- Lallot (J.), 2003, « À propos des syncatégorèmes : consignification et signification adjacente dans la tradition logico-grammaticale grecque », *Histoire épistémologie langage* 25/2, 9-32.
- Lamiroy (B.), 1999, « Auxiliaires, langues romanes et grammaticalisation », *Langages* n° 135, 33-45.
- Langacker (R.), 1987, *Foundations of cognitive Grammar*, Stanford, Stanford University Press.
- Le Guern (M.), 2003, *Les deux logiques du langage*, Paris, Champion.
- Lehmann (C.), 2004, « Theory and method in grammaticalization », *Zeitschrift für germanistische Linguistik*, 32(2), 152-187.
- Lehmann (A.) & Martin-Berthet (F.), 2018, *Lexicologie: sémantique, morphologie et lexicographie*, Paris, A. Colin.
- Lyons (J.), 1970, *Linguistique générale, Introduction à la linguistique théorique*, Paris, Larousse.
- Lyons (J.), 1980, *Sémantique linguistique*, Paris, Larousse.
- Marchello-Nizia (C.), 2006, *Grammaticalisation et changement linguistique*, Bruxelles, De Boeck.
- Martinet (A.), 1960=1966, *Éléments de linguistique générale*, Paris, A. Colin.,
- Martinet (A.), 1985, *Syntaxe générale*, Paris, A. Colin.
- Meillet (A.), 1948, *Linguistique historique et linguistique générale*, Paris, Champion.
- Milner (J.-C.), 1989, *Introduction à une science du langage*, Paris, Seuil.
- Monneret (P.), 2020, *Grammaire du français, Terminologie grammaticale*, Paris, Ministère de l'éducation nationale, en ligne : https://www.ac-paris.fr/portail/upload/docs/application/pdf/2022-03/terminologie-grammaticale_version21.pdf
- Mortureux (M.-F.), 1997, *La lexicologie entre langue et discours*, Paris, Sedes.
- Newmeyer (F.-J.), 2001, « Deconstructing grammaticalization » *Language Sciences* 23, 187-229.
- Nølke (H.), 1993, *Le regard du locuteur*, I, Paris, Kimé.
- Pagliuca (W.), 1994, *Perspectives on grammaticalization*, Amsterdam, Benjamins.
- Pollock (J.-Y.), 1998, *Langage et cognition, Introduction au programme minimaliste de la grammaire générative*, Paris, PUF.
- Pottier (B.), 1987, *Théorie et analyse linguistique*, Paris, Hachette.

- Pottier (B.), 2000, *Représentations mentales et catégorisations linguistiques*, Louvain-Paris, Peeters.
- Prévost (S.), 2003, « La grammaticalisation : unidirectionnalité et statut », *Le Français Moderne* 2 (71), 144-166. En ligne : <https://halshs.archives-ouvertes.fr/halshs-00087734>
- Prévost (S.), 2006, « Grammaticalisation, lexicalisation et dégrammaticalisation : des relations complexes », *Cahiers de praxématique* n° 46, 121-139.
- Rastier (F.), 1987, *Sémantique interprétative*, Paris, PUF.
- Riegel (M.), Pellat (J.-C.) & Rioul (R.), 2009, *Grammaire méthodique du français*, Paris, PUF.
- Saussure (F. de), 2002, *Écrits de linguistique générale*, S. Bouquet & R. Engler éd., Paris, Gallimard.
- Sauvageot (A.), 1972, *Français écrit, Français parlé*, Paris, Larousse.
- Talmy (L.), 1995, « The relation of Grammar to Cognition », *Acta Romanica Basiliensia* n° 3, 139-173.
- Tesnière (L.), 1959, *Éléments de syntaxe structurale*, Paris, Klincksieck.
- Traugott (E.-C.), 1999, « The rhetoric of counter-expectation in semantic change : a study in subjectification », in *Historical Semantics and Cognition*, A. Blank & P. Koch éd., Berlin, De Gruyter, 177-196.
- Victorri (B.), 1999, « Le sens grammatical », *Langages* 136, 85-105.
- Victorri (B.), 2004, « Les grammaires cognitives », in *La linguistique cognitive*, C. Fuchs éd., Paris, Ophrys, 73-98.