

## Entre les mots possibles et les mots existants : les unités lexicales à faible probabilité d'actualisation

Danielle Corbin

Texte initialement paru dans *Sillexicales* 1, pp. 79-90. Nous remercions Pierre Corbin d'avoir accepté que ce texte soit de nouveau mis à la disposition de la communauté scientifique, et d'avoir rédigé la présente note.

*Les recherches de Danielle Corbin sur le lexique construit, qui remontent au début des années 1970, se sont inscrites d'emblée dans le paradigme génératif qui pénétrait depuis peu la linguistique française par des introductions vulgarisatrices et des traductions de textes fondateurs. À la suite de François Dell, elle fut parmi les premiers, pour le domaine français, à transférer en morphologie dérivationnelle, dont l'autonomie n'allait pas de soi, la problématique, ancrée en syntaxe, de la prédictivité de règles, en l'appliquant à la description des procédés de construction d'unités lexicales. Ne pas inférer ces règles de la seule prise en compte des observables, généralement filtrés par les dictionnaires, pour atteindre la compétence lexicale par delà ses manifestations aléatoires fut donc un maître mot traversant toute son œuvre depuis ses premiers articles, déterminant ses prises de position contre toute « morphologie de l'évidence » et suscitant dans l'ensemble de ses productions une profusion de formes non observées dont la plausibilité était évaluée positivement ou négativement.*

*L'article ici republié, un des derniers de son auteur puisque paru initialement trois ans avant son décès, s'ouvre sur une reformulation de cette discrimination méthodologique – « La distinction entre les mots possibles et les mots existants est généralement utilisée » – qui donne lieu à un examen de la polyréférence des deux termes de cette dichotomie dans la littérature spécialisée, mais pas à leur mise en cause frontale, ce qui a de quoi surprendre au regard à la fois du fait qu'ils ne sont pas sans concurrents (lexique réel / virtuel, de fait / de droit...) et des prises de distance antérieures de la morphologue à leur égard, et notamment à celui de mots existants, retenu pourtant dans le titre de la contribution qu'on va lire ou relire, en écho à l'intitulé du colloque qui la suscita (« Mots possibles et mots existants »). Reformulée avec les mots acceptés par son auteur, la question sur laquelle se penche ce texte est la suivante : pourquoi certains mots "possibles" non seulement n'"existent" pas mais n'ont pas une forte probabilité d'"exister" ? De là découle une typologie de causes, linguistiques ou mondaines, de non-actualisation d'items virtuels qui, au sein des « lacunes lexicales », ont vocation à expliquer celles qui peuvent être considérées comme « non-accidentelles ».*

*Avec le recul, le propos de Danielle Corbin trouve un renforcement dans les faiblesses mêmes, plus perceptibles aujourd'hui que naguère, de l'étalement empirique de ses raisonnements. Par le recours aux moteurs de recherche, l'attestation non problématique de divers items qu'elle réputait difficilement actualisables est aisée à observer – les lecteurs les trouveront par eux-mêmes –, ce qui fragilise, voire rend caducs, certains pans de l'argumentation développée. Mais cela ne fait que confirmer que les linguistes, quelles que soient l'étendue de leur documentation et la défiance méthodologique dont ils font preuve, sont nécessairement portés à sous-estimer l'acceptable linguistique : c'était le message même de Danielle Corbin.*

*La présente réédition reproduit le texte original de l'article sans modifications autres que de forme et d'actualisation de la bibliographie, les références « (à paraître a) » et « (à paraître b) » d'articles de Danielle Corbin ayant été remplacées respectivement par « (1997a) » et « (1997b) ». Le titre de son ouvrage « (en préparation) », à ce jour toujours inédit, a été conservé dans sa formulation originale (La construction des unités lexicales en français. Introduction méthodique à la morphologie dérivationnelle) sans être remplacé par celui qui fut finalement retenu (Le lexique construit. Méthodologie d'analyse), pour conserver le témoignage de celui envisagé par l'auteur au moment de la rédaction de l'article ici republié. Enfin, les deux articles évoqués au § 1. de celui-ci par la formulation « G. Dal et B. Fradin, ici même », tous les deux publiés dans le même volume d'actes que l'article qui y réfère (Corbin D., Fradin B., Habert B., Kerleroux F. & Plénat M. éds, Mots possibles et mots existants. Forum de morphologie (1<sup>ères</sup> rencontres). Actes du colloque de Villeneuve d'Ascq (28-29 avril 1997), Silexicales 1, U.R.A. 382 du C.N.R.S. (SILEX) / Université de Lille III), ont pour intitulés et références respectifs « Tous les mots existants sont-ils possibles ? » (pp. 91-100) et « Les mots-valises : une forme productive d'existants impossibles ? » (pp. 101-110). Ils venaient, dans le recueil, immédiatement à la suite de celui de Danielle Corbin, qui figurait aux pp. 79-89.*

## Introduction

La distinction entre les mots possibles et les mots existants est généralement utilisée, mais la définition des deux ensembles est loin de faire l'objet d'un consensus : *mots existants* peut renvoyer aux mots répertoriés dans les dictionnaires, ou à ceux dont on peut trouver une attestation, qu'ils soient répertoriés ou non dans les dictionnaires, ou encore à ceux que croit connaître l'énonciateur ; *mots possibles* renvoie généralement à l'ensemble complémentaire du premier, qui comporte les mots construits conformément à la grammaire mais non répertoriés dans les dictionnaires ou non attestés, à moins qu'il ne renvoie à l'ensemble entier des mots construits conformément à la grammaire, indépendamment de la question de savoir s'ils sont répertoriés ou non. Quelle que soit la façon de le définir, chacun de ces deux ensembles en délimite apparemment un autre en négatif, celui des *mots inexistantes* et celui des *mots impossibles*. Il sera question ici d'une classe particulière de mots possibles non-existants : ceux qui, bien que construits conformément à la grammaire, ne peuvent pas être répertoriés lexicographiquement ni même, quelquefois, être simplement attestés. Après une tentative de clarification de la notion de lacune lexicale qui me permettra de circonscrire mon sujet, je passerai en revue un certain nombre de types différents de ce que j'appelle des *unités lexicales à faible probabilité d'actualisation*.

### 1. Typologie des lacunes lexicales

Je me propose de retravailler ici la notion de lacune lexicale, dans laquelle j'avais déjà tenté de mettre un peu d'ordre dans (1975) et (1987). Par un déplacement du critère traditionnel de définition des lacunes de la seule non-attestation à la conformité ou la non-conformité grammaticale, je distinguais alors les lacunes accidentelles (mots bien formés) et les lacunes systématiques (mots mal formés). Cette distinction binaire me paraît aujourd'hui insuffisante : je voudrais montrer que certaines lacunes dites *accidentelles* sont néanmoins motivées, ce qui passe par un réexamen des notions de mot existant et de mot possible et de leurs limites.

La notion de mot existant implique une instance d'actualisation qui permette d'authentifier le mot. Cette instance peut être, au choix, les dictionnaires, les productions langagières primaires ou encore le stock lexical que croit connaître le linguiste en situation d'effectuer une classification de ce point de vue. Par rapport à cette instance, l'existence d'un mot est une notion discrète. Pour des raisons de commodité pratique (seuls les dictionnaires offrent un accès objectivable aux unités actualisées, même s'ils ne peuvent répertorier que celles qui figurent dans les documents à partir desquels ils sont constitués), j'entendrai ici par *mot existant* un mot répertorié dans l'un ou l'autre des grands dictionnaires actuels de référence, à savoir, pour l'essentiel, le *Grand Robert de la langue française (GRLF)* et le *Trésor de la langue française (TLF)*. J'appellerai *lexique actuel* l'ensemble – nécessairement partiellement aléatoire<sup>1</sup> – des mots existants ainsi définis. Cet ensemble n'a aucune valeur théorique ; il constitue seulement un repère par rapport auquel il paraît raisonnable de parler de *lacunes* et qui peut servir, par exemple, à mesurer la rentabilité actuelle d'un procédé de construction. L'ensemble des lacunes, quel que soit leur statut, forme le *lexique non-actuel*.

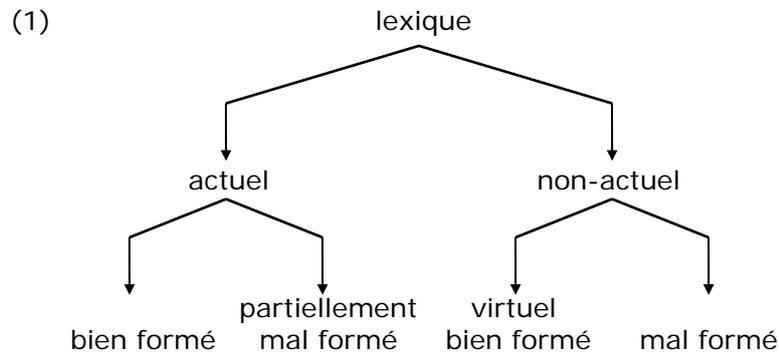
La notion de mot possible implique, quant à elle, une norme de référence par rapport à laquelle évaluer la conformité du mot : d'un point de vue morphologique, c'est la grammaire de chaque procédé particulier de construction qui définit la conformité attendue des propriétés formelles, structurelles et sémantiques des mots construits. Par rapport à la classe ainsi définie à partir d'un principe organisateur de nature linguistique, la notion de lacune acquiert une valeur non plus documentaire mais argumentative, et l'on peut s'attendre à ce que la notion de mot possible ne soit pas une notion discrète. G. Dal et B. Fradin, ici-même, traitent chacun d'une classe hybride de mots apparemment impossibles au regard de la grammaire mais auxquels leur existence confère au moins

---

<sup>1</sup> On le sait, aussi étendue soit la nomenclature d'un dictionnaire, elle ne peut pas être complète : le stock de mots répertoriés dans les dictionnaires varie selon la période dont le dictionnaire prétend rendre compte, selon le type de dictionnaire, sa taille, le public visé, sans parler des contingences internes au dictionnaire lui-même (cf. Corbin D. & P. (1989)).

une légitimité de données observables, c'est-à-dire de matériaux à partir desquels les linguistes sont en droit de formuler des hypothèses leur permettant d'évaluer le degré de possibilité de telle ou telle réalisation non-existante.

La combinaison logique des deux critères hétérogènes d'actualisation et de bonne formation aboutit donc à la typologie suivante :



Le lexique actuel contient les mots actualisés, dont la plupart sont supposés bien formés mais qui peuvent comporter dans leurs rangs des individus au moins partiellement monstrueux. Les mots du lexique non-actuel peuvent être bien formés (lacunes accidentelles, qui constituent le *lexique virtuel*) ou mal formés (lacunes motivées grammaticalement) <sup>2</sup>.

Quelle que soit la façon dont on détermine le lexique actuel et quelle que soit l'extension que l'on donne à la classe des mots possibles (qu'on la restreigne au lexique virtuel ou pas), une telle typologie rencontre deux sortes de limites, l'une méthodologique, l'autre théorique.

D'un point de vue méthodologique, le descripteur du lexique est confronté, comme tout linguiste, à la complexité des critères qui fondent la distinction entre les lacunes accidentelles et les lacunes motivées et à la difficulté de leur mise en œuvre. En effet, une fois balayée l'erreur – répandue, on le verra – qui consiste à identifier les mots impossibles aux mots inexistantes ou supposés tels, l'évaluation du degré de conformité grammaticale d'une unité lexicale suppose la connaissance et la mise en application hiérarchisée de l'ensemble complexe des facteurs, phonologiques, catégoriels, structurels, combinatoires, sémantiques, qui conditionnent la bonne formation des unités.

La limitation théorique de la typologie proposée est plus intéressante parce que moins connue. L'opposition entre mots possibles et mots impossibles, telle qu'elle a été définie et articulée sur la notion de mots existants, ne fait intervenir que des conditions particulières de bonne formation mais ignore les conditions générales qui pèsent sur l'utilisabilité et l'actualisation des unités lexicales. Or, étant donné la double nature de la notion d'unité lexicale (dénomination <sup>3</sup> et unité de langue), il n'y a aucune raison, en droit, de supposer que toutes celles qui sont possibles au regard de la grammaire soient actualisables.

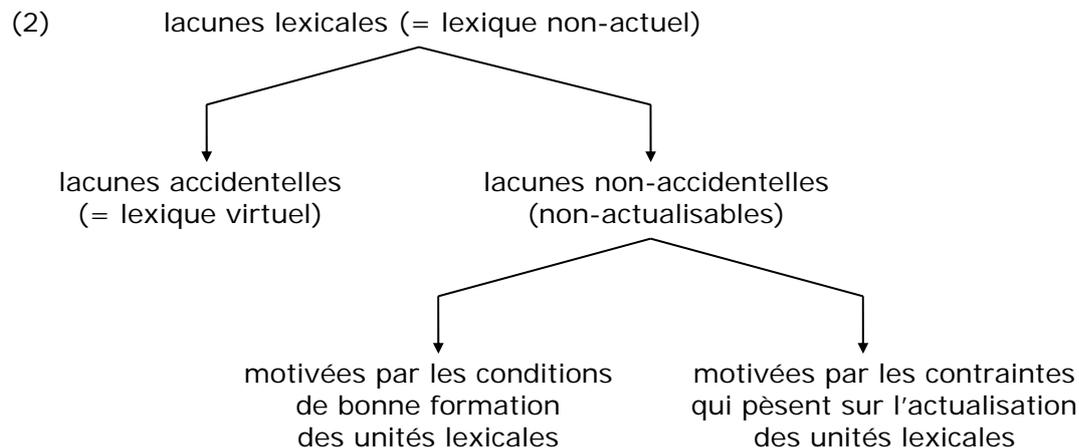
Je propose de définir une unité lexicale comme une séquence linguistique associée ou associable de façon stable, hors contexte, à un sens et porteuse d'une catégorie qui l'autorise à occuper dans des énoncés une position syntaxiquement atomique. Une unité lexicale matérialise l'intersection entre les modes de filtrage culturel et conceptuel (critère définitionnel de type sémiotique) et les modes d'organisation structurelle (critère définitionnel de type catégoriel) propres à une langue <sup>4</sup>. Construite ou non, une séquence

<sup>2</sup> On peut tenter de croiser cette typologie et l'opposition que propose Milner (1989) entre « le possible de langue » et le « possible matériel » : dans le domaine lexical, le « possible de langue » recouvre la plupart des unités du lexique actuel et toutes celles du lexique virtuel, et le « possible matériel » le lexique actuel dans son ensemble, y compris les monstres actualisés.

<sup>3</sup> Seules m'intéressent ici les unités lexicales à sens référentiel, et non les "mots grammaticaux". *Dénomination (name)* est à entendre au sens large, non limité à la catégorie du nom (*noun*).

<sup>4</sup> On trouvera une argumentation et un commentaire de cette définition dans Corbin (1996, 1997a, 1997b).

linguistique n'est utilisable comme unité lexicale, donc actualisable dans des énoncés, et partant répertorié dans des dictionnaires, que si elle peut avoir un correspondant référentiel. Pour qu'une séquence linguistique construite soit actualisable, il ne suffit pas qu'elle soit conforme à la grammaire ; il faut encore qu'elle rencontre les conditions de son actualisation en tant qu'unité lexicale. Entre les mots possibles et les mots existants, on peut donc faire l'hypothèse qu'il y a des mots possibles qui ne peuvent pas exister. L'ajout de ce type d'unités aux lacunes accidentelles et aux lacunes motivées grammaticalement permet de proposer la typologie suivante de la notion de lacune lexicale (ou du lexique non-actuel) :



## 2. La notion d'unité lexicale à faible probabilité d'actualisation

Deux grands types de facteurs sont susceptibles d'interdire ou de limiter l'actualisation en tant qu'unités lexicales de séquences construites grammaticalement possibles : des facteurs liés aux propriétés propres de ces séquences, et des facteurs liés à l'adéquation et/ou à la pertinence de leur utilisation en tant qu'unités lexicales. Je soutiendrai que certaines de ces unités, bien que difficilement actualisables, constituent néanmoins des étapes intermédiaires légitimes dans l'analyse morphologique (§§ 2.1., 2.2.1.), et qu'il serait dommageable de mettre sur le même plan les conditions qui pèsent sur la bonne formation des unités et les facteurs de non-actualisation évoqués.

### 2.1. Les unités lexicales dont les propriétés linguistiques limitent les probabilités d'actualisation

Nous avons fait l'hypothèse, dans D. & P. Corbin (1991), que le suffixe *-ier/-ière* sert à construire fondamentalement des adjectifs, et que les noms en *-ier/-ière* sont les résultats d'une conversion de ces adjectifs. Cette hypothèse peut paraître contre-intuitive (il est difficile d'imaginer d'emblée un adjectif de forme *pommier*, *serrurier* ou *chéquier*) et contrefactuelle, puisque, sur environ 1 500 entrées en *-ier/-ière* répertoriées dans le *GRLF*, plus de 70% sont catégorisées seulement comme noms et moins de 5% seulement comme adjectifs, le reste étant doublement catégorisé. Certes, les catégorisations lexicographiques sont souvent peu fiables, notamment parce qu'elles sont indexées sur les contextes dans lesquels les mots sont illustrés, mais, en l'occurrence, l'écart est suffisamment significatif pour mériter une explication.

L'hypothèse que je voudrais avancer est que la non-actualisation de la majorité des adjectifs en *-ier* (et donc leur absence des dictionnaires) ainsi que leur aspect contre-intuitif s'expliquent par le fait que les propriétés sémantiques de ces adjectifs limitent fortement le choix des noms recteurs susceptibles de leur offrir un ancrage syntaxique. Le suffixe *-ier* ayant la particularité sémantique de donner à voir les référents des noms auxquels il peut s'appliquer comme ayant une utilité pratique d'un point de vue anthropocentrique, le sens construit des adjectifs en *-ier* fait de ceux-ci des moyens de sous-catégorisation explicite d'une catégorie référentielle (celle d'un nom recteur) à partir des propriétés utilitaires des référents de leur nom de base : par exemple, *arbre*

*fruitier* renvoie à une sous-catégorie d'arbres qui se caractérisent par le fait qu'ils produisent des fruits comestibles par l'homme, *betterave sucrière* à une variété de betterave utilisée pour la fabrication de sucre. Les séquences constituées d'un nom et d'un adjectif suffixé par *-ier* constituent donc souvent des dénominations descriptives de sous-catégories référentielles. Il s'ensuit qu'un adjectif en *-ier* ne peut être actualisé dans un énoncé que s'il peut sous-catégoriser de façon pertinente le référent d'un nom recteur en fonction des propriétés utilitaires du référent du nom sur lequel il est construit, ce qui suppose que le nom recteur remplisse deux conditions : être d'un niveau hiérarchique compatible avec la sous-catégorisation qu'introduit l'adjectif<sup>5</sup>, et être sémantiquement compatible avec ce type de sous-catégorisation. Si, par exemple, il n'existe qu'un hyperonyme d'un niveau hiérarchique trop élevé par rapport à l'adjectif, ou si la compatibilité d'un hyperonyme de niveau adéquat est bloquée sémantiquement, la séquence, et donc l'adjectif, a peu de chances d'apparaître. Soit par exemple ??*arbre pommier* : la production de pommes peut servir à sous-catégoriser les arbres fruitiers, et non les arbres en général ; le nom recteur est donc d'un niveau trop élevé<sup>6</sup>. Mais ??*fruitier pommier* n'est pas meilleur : comme les pommes sont simplement des fruits (et non, par exemple, des fruits plus comestibles que d'autres, ou qui auraient un type d'utilité remarquable) et que les arbres dits *fruitiers* tirent déjà leur dénomination de leur utilité anthropocentrique, la relation utilitaire introduite par l'adjectif en *-ier* est redondante avec celle introduite par *fruitier* ; en revanche *palmier dattier* ne pose aucun problème, parce que la production de dattes est un moyen de sous-catégoriser les palmiers du point de vue de leur utilité anthropocentrique (*palme* et *datte* ne sont pas en relation d'inclusion de classe). On peut faire la même analyse pour les adjectifs qui pourraient servir à sous-catégoriser des types d'activités : un adjectif comme *serrurier* ne pourrait apparaître que dans des contextes comme *industrie serrurière* (la fabrication des serrures est une branche de l'activité industrielle) mais ni dans ??*vente serrurière* (la vente des serrures n'est pas une branche d'activité institutionnalisée en tant que telle) ni dans ??*travailleur serrurier* (il y a redondance entre le sens de *travailleur* et celui de *serrurier* : les serruriers tirent leur dénomination du fait qu'ils travaillent sur les serrures)<sup>7</sup>. Étant donné le sens construit de l'adjectif et les contraintes mentionnées, les noms susceptibles de servir d'ancrages syntaxiques à un adjectif de forme *chéquier* sont rares ; les moins mauvais paraissent être *paiement*, *transactions* :

(3) ?*La prochaine généralisation de la carte à puce sonne le glas des paiements chéquiers*

Les adjectifs suffixés par *-ier* sont conformes à la grammaire, mais difficilement actualisables. Ceux qui le sont sans problème (*financier*, *laitier*, *pétrolier*, *ratier*) sont ceux qui, à cause des propriétés du référent de leur base, peuvent servir à construire des dénominations complexes descriptives adéquates (*marché financier*<sup>8</sup>, *vache laitière*, *industrie pétrolière*, *chien ratier*). Contrairement aux adjectifs en *-ier*, les noms convertis sur ces adjectifs n'ont aucun mal à être actualisés. La différence tient au fait qu'un adjectif et un nom n'ont ni la même autonomie référentielle ni, par conséquent, la même autonomie syntaxique (cf. Wierzbicka (1986), Riegel (1993)) : alors qu'un adjectif exprime une propriété d'une catégorie référentielle, un nom dénomme une catégorie. S'il est difficile de trouver une catégorie d'un niveau hiérarchique adéquat pour être sous-catégorisée spécifiquement par un adjectif suffixé par *-ier*, notre mode d'appréhension de

<sup>5</sup> Cf. les recherches sur les « niveaux » dans la sémantique du prototype (rapportées par exemple dans Kleiber (1990 : 78-98)).

<sup>6</sup> La seule attestation que nous ayons trouvée d'une séquence [*arbre* + adjectif suffixé par *-ier* sur un nom référant à un fruit] est *arbre canellier*, datée de 1575 par les dictionnaires.

<sup>7</sup> *Ouvrier serrurier* serait interprété, quant à lui, comme une dénomination hiérarchique des serruriers, et non comme une dénomination distinctive des ouvriers.

<sup>8</sup> L'adjectif qui figure dans cette dénomination est construit sur *finance*. Il existe un autre adjectif de même forme construit sur le nom *financier*, que l'on trouve dans *sauce financière* (cf. D. & P. Corbin (1991 : 96-97)).

la réalité autorise au contraire la constitution (et la dénomination) d'un grand nombre de catégories dont le principe organisateur est l'utilité pratique d'un point de vue anthropocentrique <sup>9</sup>.

## 2.2. Les unités lexicales dont des facteurs d'adéquation ou de pertinence limitent les probabilités d'actualisation

L'inadéquation ou la non-pertinence référentielle <sup>10</sup> de certaines expressions linguistiques construites conformément à la grammaire du lexique peut tenir à quatre types de facteurs : des facteurs pragmatiques, sémiotiques, symboliques et conventionnels.

2.2.1. Des raisons pragmatiques, relatives aux normes implicites de la communication linguistique et à l'effet que peuvent provoquer les mots sur les acteurs de l'énonciation, peuvent être à l'origine de la non-actualisation de certains mots construits possibles. Par exemple, alors que les adjectifs cités en (4b) appartiennent au lexique actuel, ceux cités en (4a) ne sont pas actualisés :

- |                         |  |
|-------------------------|--|
| (4) a <i>chauffable</i> | b <i>inchauffable</i> (Béchade (1992 : 134))             |
| ° <i>démodable</i>      | <i>indémodable</i> ( <i>ibid.</i> )                      |
| ° <i>usable</i>         | <i>inusable</i> (Grevisse (1993 : 233))                  |
| ° <i>violable</i>       | <i>inviolable</i> (Denis & Sancier-Chateau (1994 : 317)) |

Dans les références citées entre parenthèses, l'inexistence des adjectifs de (4a) est utilisée comme argument pour analyser parasyntétiquement les adjectifs de (4b) : comme ni l'adjectif non préfixé (par exemple °*usable*) ni le verbe préfixé (par exemple \**inus(er)*) ne font partie du lexique actuel, on en déduit que *in-* et *-able* se sont appliqués simultanément sur la base verbale non préfixée (par exemple *us(er)*) <sup>11</sup>. Or, les adjectifs de (4a) sont conformes à la grammaire, les propriétés auxquelles leur sens leur permet de renvoyer sont concevables, mais des raisons pragmatiques font qu'il y a peu de chances que ces mots soient actualisés dans les énoncés, et donc répertoriés dans les dictionnaires. Par exemple, étant donné le sens d'*us(er)*, la propriété "pouvoir être usé" s'applique de façon privilégiée à des catégories d'objets concrets fabriqués (par exemple des semelles, des moquettes, etc.) aux propriétés intrinsèques essentielles <sup>12</sup> desquelles l'usabilité appartient <sup>13</sup> ; son actualisation sous la forme de l'adjectif °*usable* d'une part n'aurait guère dans ce cas de pertinence informative, d'autre part dévaloriserait les référents des noms que l'adjectif prédiquerait. Il en irait autrement si la propriété était appliquée par exemple à une roche : dans le contexte d'une description géologique, l'adjectif °*usable* retrouverait une pertinence (on pourrait dire par exemple qu'une roche est plus usable qu'une autre par l'érosion). Cet exemple conjoint deux facteurs qui, dans les autres cas, peuvent se trouver plus ou moins disjoints : le caractère essentiel, définitionnel ou présupposé normal de la propriété à laquelle renvoie l'adjectif non préfixé (il est présupposé normal qu'une maison soit chauffable), et l'inadéquation référentielle

<sup>9</sup> Notons, en passant, que cette observation argumente contre les traitements qui font des adjectifs nominalisés des résidus de groupes nominaux dont le nom aurait été effacé (cf. Marchand (1969 : 361), Guilbert (1971 : XXXVIII), Tamba-Mecz (1980 : 122), Riegel (1985 : 193), etc.).

<sup>10</sup> Cf. Sperber & Wilson (1989).

<sup>11</sup> Denis & Sancier-Chateau (1994 : 317) analysent *inviolable* parasyntétiquement à partir de *viol* et non de *viol(er)*, sans que cette analyse soit justifiée ni justifiable.

<sup>12</sup> Anscombe (1994 : 302) définit les propriétés intrinsèques essentielles de la façon suivante : « Une propriété [intrinsèque, *i.e.* constitutive] P est essentielle pour une entité I si elle est définitoire de la classe des I », et il utilise ces propriétés pour expliquer certaines lacunes d'adjectifs en *-able* non préfixés : par exemple « \**muable/immuable*. Le fait que l'on n'ait pas l'adjectif traduit le simple fait que rien ne dure et que tout change » (p. 305). Je ne le suis toutefois pas en ce qui concerne le marquage de *muable* : si l'astérisque marque une expression agrammaticale, alors il ne convient pas pour °*muable*, qui est un être linguistique dont l'actualisation est improbable mais pas un monstre.

<sup>13</sup> En conséquence, dire qu'une semelle est *inusable* revient à dire qu'elle s'use moins vite que ce à quoi on pourrait s'attendre, et pas qu'elle ne peut pas être usée.

des énoncés dans lesquels l'adjectif pourrait figurer (il serait anticommercial de dire qu'une robe peut se démoder, et contraire aux attentes qu'un secret puisse être violé). On peut d'ailleurs remarquer que cette inadéquation n'est pas forcément la même pour les noms de propriété correspondants<sup>14</sup> : dans un test comparatif, il sera plus naturel de dire qu'on a mesuré la plus ou moins grande usabilité d'un produit que d'affecter la propriété directement au produit par l'intermédiaire d'une prédication adjectivale.

Comme dans l'exemple précédent, de telles unités lexicales difficilement actualisables sont néanmoins des étapes tout à fait admissibles dans des dérivations morphologiques : les adjectifs listés en (4a) sont les bases auxquelles s'appliquent le préfixe *in-* pour construire les adjectifs listés en (4b), et le suffixe *-ité* pour construire les noms de propriété correspondants. Si tel n'était pas le cas, l'analyse linguistique se condamnerait à être *ad hoc*, comme l'est l'analyse parasynthétique.

2.2.2. Il peut aussi arriver que la forme de certaines unités lexicales soit inappropriée à la dénomination des référents que leur sens construit leur permettrait *a priori* d'atteindre. Ce sont alors des facteurs que l'on peut qualifier de *sémiotiques* qui empêchent ces unités d'être actualisées, facteurs particulièrement peu visibles dans la mesure où ce sont des connotations formelles, et non la forme elle-même, qui sont en cause. J'illustrerai ce cas de figure par l'exemple des composés savants de structure [ [ X ]<sub>N</sub> [ Y ]<sub>V</sub> ]<sub>A</sub>, que leur sens construit autorise, *a priori*, à qualifier un nom référant à un être humain ou à un objet (et, par conversion, à dénommer ce référent) : *sémaphore* réfère ainsi à un dispositif transmetteur de signes, que ce soit dans la marine, dans les chemins de fer ou en informatique, *lampadophore* à un porteur de flambeaux antique. La forme "savante" de ces composés convient à leur application à des référents scientifiques, techniques ou culturels, mais les empêche d'être appropriés à la dénomination de référents appartenant à l'univers quotidien commun<sup>15</sup>. Par exemple, un composé comme *galactophore* a un sens construit, glosable par "qui (trans)porte du lait", qui lui permet de servir à la dénomination de toute entité, humaine ou non, susceptible d'avoir de façon saillante la propriété ou la fonction de porter ou de transporter du lait. Mais *galactophore* n'est utilisé que pour dénommer les canaux qui conduisent le lait secrété par les glandes mammaires vers le mamelon, et ne conviendrait pas, sauf avec effet burlesque, pour dénommer par exemple le livreur qui, dans la culture anglo-saxonne, dépose le lait à domicile.

2.2.3. J'illustrerai conjointement le cas des unités inactualisables pour des raisons symboliques ou conventionnelles par des exemples empruntés à Gross (1988). Dans un cadre théorique où les unités lexicales comportant plusieurs mots – qu'il appelle *mots composés*<sup>16</sup> –, sont considérées comme issues du figement, plus ou moins rigide selon les cas, de séquences syntaxiques, G. Gross propose un certain nombre de critères permettant de mesurer le degré de figement des « noms composés ». Parmi ces critères, il retient la restriction des commutations possibles à l'intérieur de la séquence complexe (ce qu'il appelle « rupture d'une série distributionnelle ») : les expressions *bras droit* et *mardi gras* sont ainsi considérées comme figées (p. 66), et acquièrent de ce fait un statut lexical, parce que, dans la première, l'adjectif, dans la deuxième, le nom ne sont pas remplaçables par d'autres éléments d'un paradigme attendu. En contraste, les séquences « *\*bras gauche, \*mercredi gras* » sont, selon G. Gross, exclues du lexique. Comme la perspective de G. Gross est celle de la constitution de dictionnaires électroniques susceptibles de servir par exemple à la traduction automatique, la conclusion qu'il tire de cette observation est que

---

<sup>14</sup> Merci à P. Corbin pour cette remarque.

<sup>15</sup> Comme me l'a fait remarquer F. Kerleroux, un moyen d'améliorer l'adaptation des mots de forme "savante" à une dénomination familière est l'apocope : cf. *mégalo[mane]*, *claustr[o]phobe*, *phono[graphie]*, etc.

<sup>16</sup> Je préfère la dénomination *unités polylexématiques*, qui n'engage pas le traitement morphologique ou syntaxique de ces unités : cf. Corbin (1992) et (1997a).

« plus la série est restreinte par rapport à son extension maximale moins il y a lieu de calculer le sens à partir des éléments » (*ibid.*)

Autrement dit, le sens de *bras droit* et celui de *mardi gras* ne seraient pas calculables compositionnellement<sup>17</sup> parce que l'adjectif *droit* et le nom *mardi* seraient seuls à pouvoir occuper la place qu'ils occupent.

L'analyse de G. Gross pose à mon sens deux problèmes : on peut montrer d'une part que le sens de ces séquences est calculable compositionnellement, d'autre part que l'exclusion de *bras gauche* et de *mercredi gras*, stigmatisée par l'astérisque, est due à des raisons respectivement symboliques et conventionnelles, suffisantes pour empêcher l'actualisation de ces séquences en tant qu'unités lexicales mais insuffisantes pour identifier ces séquences à des monstres lexicaux.

L'utilisation du nom d'une partie du corps ou d'un attribut caractéristique pour référer à un être humain est un procédé régulier : on peut référer par exemple à un enfant par *bâton de vieillesse*, à un complice par *œil* (cf. *l'œil de Moscou*) ou *oreille*, à un exécutant par *bras armé*, etc. L'emploi de *bras droit* pour référer à une personne indispensable n'a donc rien d'exceptionnel. Comme c'est l'utilité de ces personnes qui est mise en vedette, le choix de *bras* est également motivé. Enfin, si la langue a retenu *bras droit* et non *bras gauche* pour désigner une personne dont le rôle social est valorisé, c'est que la valeur symbolique associée aux propriétés auxquelles renvoient les adjectifs *droit* et *gauche* n'est pas équivalente : la localisation droite correspond à une valorisation<sup>18</sup>, la gauche généralement à une dévalorisation. Cette différence trouve un écho dans le fait que chacun des deux adjectifs renvoie par ailleurs à une propriété axiologique, comme le montrent les cooccurrences privilégiées suivantes :

- (5)a *un caractère droit*  
*aller dans le droit chemin*
- b *un geste gauche*  
*un style gauche*  
*se lever du pied gauche*

D'une part le sens de *bras droit* – et celui d'un éventuel °*bras gauche* – peut donc se calculer à partir de ses éléments constituants, même si sa référence précise n'est pas entièrement prédictible. D'autre part °*bras gauche* pourrait éventuellement convenir pour nommer de façon stable, par exemple, la classe des deuxièmes vice-présidents dans une université. Il est clair que si cette dénomination a peu de chances d'être retenue, c'est parce que la valeur symbolique qu'elle connote dévaloriserait son référent.

Quant au choix de *mardi* plutôt que *mercredi* pour désigner le dernier jour où, dans la religion catholique, on a le droit de manger gras avant le Carême, il est imposé par le calcul des dates rituelles : Pâques tombant nécessairement un dimanche et le début du Carême étant situé quarante-six jours avant Pâques, comme une semaine a sept jours, le dernier jour avant le Carême tombe nécessairement un mardi, et le premier jour du Carême un mercredi (d'où *le mercredi des Cendres*). Par conséquent, d'une part le sens de *mardi gras* est calculable (comme l'adjectif *gras* renvoie à une propriété incompatible avec le nom d'un jour, on infère qu'il qualifie par hypallage le jour où quelque chose de

---

<sup>17</sup> Conformément aux habitudes linguistiques, j'entends par *sens calculable compositionnellement* ou *sens compositionnel* le sens d'une séquence complexe (que ce soit ou non une unité lexicale) quand il est calculable à partir du sens des constituants formellement identifiables dans la structure de cette séquence et conformément aux modes combinatoires réguliers pour cette structure. Selon le rôle que joue l'adjectif, le sens compositionnel d'une structure [N Adj.] lui permet normalement de renvoyer soit à la même catégorie que celle à laquelle réfère le N, caractérisée par la propriété à laquelle renvoie l'adjectif (c'est le cas de *bras droit* dans la phrase *Son bras droit est plus musclé que le gauche* ou de *mardi gris* dans la séquence *Par un mardi gris d'hiver...*), soit à une catégorie que le référent de l'adjectif permet de subordonner à celle à laquelle réfère le nom (c'est le cas de *piano droit* (comparé à *piano à queue*) ou de *angle droit* (comparé à *angle obtus*, *angle plat*)).

<sup>18</sup> Dans les représentations émanant de la religion catholique, les justes sont placés à la droite de Dieu.

saillant (la nourriture, donc) peut recevoir ce qualificatif), d'autre part ce sont des raisons conventionnelles complètement indépendantes de la langue qui excluent ° *mercredi gras* du lexique actuel.

### 2.3. Il y a deux types d'unités lexicales sans référent

Comme le montrent les exemples développés, la mise en évidence d'une classe d'unités lexicales possibles mais difficilement actualisables est cruciale pour ne pas confondre inactualisable et impossible, c'est-à-dire pour éviter d'affecter à la grammaire ce qui relève de l'utilisation référentielle des expressions linguistiques. Un exemple final permettra de consolider la différence entre une impossibilité linguistique et une inadéquation référentielle. En contraste avec ° *usable* ou ° *bras gauche*, qui ne violent aucune condition de bonne formation grammaticale, examinons l'exemple de *retuer*, qui, à première vue, partage avec les expressions qui précèdent le fait de ne pas avoir de référent. Notons d'abord qu'il n'est impossible de trouver un référent à ce verbe que lorsque celui-ci renvoie à une occurrence de procès où l'agent et le patient sont identiques à ceux du premier procès présumé par le préfixe *re-*, comme le montre le contraste entre (6a) et (6b) :

- (6)a *Il n'a quand même pas retué une vieille femme ?*  
b \**Il n'a quand même pas retué cette vieille femme ?*

Dans (6a), qui peut se dire d'un assassin en série, l'itération porte sur le procès générique *tuer une vieille femme*. Dans (6b), au contraire, le déictique impose une interprétation occurrence de procès, interdite pour des raisons inscrites dans le sens même du verbe *tuer* : alors que le procès *lire* par exemple peut être appliqué plusieurs fois à la même classe de livres (*Pour la première fois depuis dix ans, j'ai relu un roman policier*) ou à la même occurrence de livre (*J'ai relu Madame Bovary*), on ne peut pas tuer deux fois la même personne<sup>19</sup>. C'est pourquoi le caractère inactualisable de *retuer* (dans l'interprétation (6b)) n'est pas, à mon sens, du même ordre que celui de ° *usable* ou ° *bras gauche* : celui de *retuer*, dans cette interprétation, ressortit à une interdiction sémantique, celui de ° *usable* et ° *bras gauche* à l'inadéquation entre les propriétés de ces unités bien formées et les conventions qui régissent l'utilisation référentielle des unités lexicales.

### 3. Conclusion

L'existence d'une classe d'unités lexicales grammaticalement bien formées mais difficilement actualisables parce qu'incapables de trouver un contexte d'actualisation ou un référent<sup>20</sup> argumente en faveur d'une certaine indépendance de la langue par rapport à la cognition : le sens construit des unités lexicales construites est différent de leur capacité à référer, et ne conditionne qu'approximativement celle-ci. Contrairement à ce qu'affirme Langacker (1987 : 5), le sens n'est pas seulement un « phénomène cognitif ». Le sens est un phénomène linguistique, la conceptualisation un processus cognitif, et l'utilisation d'une unité lexicale pour référer consiste en une tentative de mise en relation de ce que la langue et de ce que la conceptualisation autorisent respectivement, la tentative pouvant échouer. La place me manque<sup>21</sup> pour étayer cette hypothèse par l'étude des bases de mots construits à l'aide de certains affixes : beaucoup d'entre elles n'ont que l'apparence d'unités lexicales (ce sont des bases possibles inactualisables) sans en avoir toutes les propriétés qui leur permettraient d'être actualisées en tant que telles (cf. Corbin & Temple (1994)). Loin d'être un phénomène marginal, cette classe d'unités permet donc de poser des problèmes fondamentaux (statut des lacunes lexicales, rapport

<sup>19</sup> Sauf à changer les règles de l'"univers de croyance", et à prendre pour victimes des "morts-vivants".

<sup>20</sup> Ce qu'on pourrait nommer en anglais des "unspeakable words", pour plagier le titre de Banfield (1982). Merci à F. Kerleroux pour cette référence.

<sup>21</sup> Cette contribution se trouvera développée dans le chapitre 2 de Corbin (en préparation).

entre sens, dénomination et référence, statut des étapes morphologiques non attestées, statut des bases dans les mots construits) qui n'auront été ici qu'esquissés <sup>22</sup>.

## BIBLIOGRAPHIE

- ANSCOMBRE J.-C. (1994), « L'insoutenable légèreté morphologique du préfixe négatif *in-* dans la formation d'adjectifs », in P. Attal éd., *La négation. Actes du colloque de Paris X-Nanterre, 12-13-14 novembre 1992*, Paris, Université Paris X-Nanterre, Numéro spécial de *Linx*, pp. 299-321.
- BANFIELD A. (1982), *Unspeakable sentences*, New York, Routledge & Kegan Paul.
- BÉCHADE H. (1992), *Phonétique et morphologie du français moderne et contemporain*, Paris, Presses Universitaires de France.
- CORBIN D. (1975), « La notion de néologisme et ses rapports avec l'enseignement du lexique », *BREF nouvelle formule* 4, Paris, Librairie Larousse, pp. 41-57.
- CORBIN D. (1987), *Morphologie dérivationnelle et structuration du lexique*, 2 vol., Tübingen, Max Niemeyer Verlag ; 2<sup>e</sup> éd., Villeneuve d'Ascq, Presses Universitaires de Lille, 1991.
- CORBIN D. (1992), « Hypothèses sur les frontières de la composition nominale », *Cahiers de grammaire* 17, pp. 25-55.
- CORBIN D. (1996), « Décrire un affixe dans un dictionnaire », in G. Kleiber & M. Riegel eds, *Les formes du sens. Études de linguistique française, médiévale et générale offertes à Robert Martin à l'occasion de ses 60 ans*, Louvain-La Neuve / Paris, Duculot, pp. 79-94.
- CORBIN D. (1997a), « Locutions, composés, unités polylexématiques : lexicalisation et mode de construction », in Martins-Baltar M. éd., *La locution, entre langue et usages*, Fontenay / Saint-Cloud, ENS Éditions, pp. 53-101.
- CORBIN D. (1997b), « La représentation d'une "famille" de mots dans le *Dictionnaire dérivationnel du français* et ses corrélats théoriques, méthodologiques et descriptifs », *Recherches linguistiques de Vincennes* 26, pp. 5-37 + Supplément, pp. I-VIII.
- CORBIN D. (en préparation), *La construction des unités lexicales en français. Introduction méthodique à la morphologie dérivationnelle*.
- CORBIN D. & P. (1989), « Sélection et description des dérivés et des composés dans le dictionnaire monolingue », in F.J. Hausmann, O. Reichmann, H.E. Wiegand & L. Zgusta Hrsg., *Wörterbücher. Dictionaries. Dictionnaires. [...] Encyclopédie internationale de lexicographie*, Berlin, Walter de Gruyter, t. 1, art. 89, pp. 937-946.
- CORBIN D. & P. (1991), « Un traitement unifié du suffixe *-ier(e)* », *Lexique* 10, pp. 61-145.
- CORBIN D. & TEMPLE M. (1994), « Le monde des mots et des sens construits : catégories sémantiques, catégories référentielles », *Cahiers de lexicologie* 65, pp. 213-236.
- DENIS D. & SANCIER-CHATEAU A. (1994), *Grammaire du français*, Librairie Générale Française, Le Livre de Poche.
- GREVISSE M. (1993), *Le bon usage. Grammaire française*, refondue par A. Goosse, 13<sup>e</sup> éd. revue, Louvain-la-Neuve, Duculot.
- GRLF = REY A. dir. (1985), *Le Grand Robert de la langue française. Dictionnaire alphabétique et analogique de la langue française* de Paul Robert, 2<sup>e</sup> éd. entièrement revue et enrichie, 9 vol., Paris, Le Robert.
- GROSS G. (1988), « Degré de figement des noms composés », *Langages* 90, pp. 57-72.

---

<sup>22</sup> Merci à tous ceux qui ont participé à l'élaboration de ces pages : à Pierre, pour sa relecture toujours critique et vigilante, à François, pour l'aide au graphisme, à Stéphane, pour l'ambiance musicale.

- GUILBERT L. (1971), « De la formation des unités lexicales », in *Grand Larousse de la langue française*, Paris, Librairie Larousse, t. 1, pp. IX-LXXXI.
- KLEIBER G. (1990), *La sémantique du prototype*, Paris, Presses Universitaires de France.
- LANGACKER R.W. (1987), *Foundations of Cognitive Grammar*, Stanford, Stanford University Press.
- MARCHAND H. (1969), *The Categories and Types of Present-Day English Word-Formation*, 2<sup>e</sup> éd., München, C. H. Beck Verlag.
- MILNER J.-C. (1989), *Introduction à une science du langage*, Paris, Éditions du Seuil.
- RIEGEL M. (1985), *L'adjectif attribut*, Paris, Presses Universitaires de France.
- RIEGEL M. (1993), « Grammaire et référence : à propos du statut sémantique de l'adjectif qualificatif », *L'information grammaticale* 58, juin 1993, pp. 5-10.
- SPERBER D. & WILSON D. (1989), *La pertinence. Communication et cognition*, trad. franç., Paris, Les Éditions de Minuit.
- TAMBA-MECZ I. (1980), « Sur quelques propriétés de l'adjectif de relation », *Travaux de linguistique et de littérature* XVIII.1, pp. 119-132.
- TLF = *Trésor de la langue française. Dictionnaire de la langue du XIX<sup>e</sup> et du XX<sup>e</sup> siècle (1789-1960)*, 16 vol., Paris, Éditions du Centre National de la Recherche Scientifique (t. 1-10) / Gallimard (depuis le t. 11), 1971-1994.
- WIERZBICKA A. (1986), « What's in a noun? (or: How do nouns differ in meaning from adjectives?) », *Studies in Language* 10.2, pp. 353-389.