

Quelques problèmes fondamentaux dans l'analyse du lexique

Madeleine Mathiot

Volume 18, Number 1-2, mars 1973

Actes du deuxième colloque international de linguistique et de traduction. Montréal, 4-7 octobre 1972

URI: <https://id.erudit.org/iderudit/003642ar>

DOI: <https://doi.org/10.7202/003642ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

Cite this article

Mathiot, M. (1973). Quelques problèmes fondamentaux dans l'analyse du lexique. *Meta*, 18(1-2), 19–33. <https://doi.org/10.7202/003642ar>

Quelques problèmes fondamentaux dans l'analyse du lexique

Les problèmes dont je parlerai dans cet essai sont des problèmes qui concernent la lexicologie et non la lexicographie. Par lexicologie j'entends la connaissance abstraite du lexique, alors que par lexicographie j'entends l'application de cette connaissance à la compilation de dictionnaires. Pour moi la connaissance abstraite du lexique relève directement de la connaissance générale du langage. Les problèmes lexicologiques font donc partie intégrale de la linguistique et c'est en tant que linguiste, et non lexicographe, que je voudrais les aborder.

L'expérience lexicographique, autant celle que j'ai moi-même péniblement acquise au cours de la compilation d'un seul dictionnaire¹, que celle beaucoup plus approfondie des lexicographes de métier, n'est pertinente ici que dans la mesure où elle met en cause des questions d'ordre lexicologique². Cette expérience constitue une toile de fond de la discussion. Elle ne sera pas directement mentionnée. Parmi les problèmes présentés par la lexicologie, je propose de distinguer ceux qui sont d'ordre théorique de ceux qui sont d'ordre méthodologique.

Les problèmes d'ordre théorique ont trait à la nature des faits lexicaux. Les problèmes d'ordre méthodologique ont trait à l'investigation systématique du lexique. Ces deux types de problèmes sont étroitement reliés l'un à l'autre : la façon dont l'investigation du lexique est poursuivie dépend en grande partie de la façon dont les faits lexicaux sont conçus.

Deux problèmes théoriques me paraissent cruciaux pour l'investigation du lexique. Le premier concerne la nature du lexique; le second concerne la nature des unités lexicales.

Pour traiter de la nature du lexique, je propose de distinguer entre 1) les critères qui entrent dans sa définition et 2) le rôle qu'il joue dans le langage. L'existence du lexique est unanimement reconnue par les linguistes mais pour la plupart d'entre eux c'est une existence sans autonomie : le lexique fait partie de la grammaire. Pour les néo-bloomfieldiens, le lexique est un appendice de la grammaire proprement dite, laquelle comprend une description des sons du langage (phonologie) et une description des formes signifiantes du langage (morphologie)

1. Voir mon *Dictionary of Papago Usage*, Research Center for the Language Sciences, University of Indiana (sous presse).

2. Les problèmes soulevés par la lexicographie ont été traités par Josette Rey-Debove dans son article sur « Le domaine du dictionnaire », dans « La lexicographie », *Langages*, vol. 19, 1970.

et syntaxe). Pour les transformationalistes, le lexique fait partie intégrante de la grammaire. C'est l'une des composantes de la grammaire avec la phonologie, la syntaxe et la sémantique³.

En vue de comparer comment ces deux écoles conçoivent la nature du lexique je propose de distinguer entre inventaire des formes incluses dans le lexique, c'est-à-dire inventaire d'unités lexicales, et information donnée sur ces formes. Le choix des formes qui constituent l'inventaire lexical révèle les critères qui entrent dans la définition du lexique. Dans le cas des néo-bloomfieldiens et des transformationalistes, ces critères différencient le lexique du *reste* de la grammaire.

Les informations qui accompagnent les unités lexicales révèlent la place qu'occupe le lexique dans le langage. Dans le cas des néo-bloomfieldiens et des transformationalistes, ces informations révèlent la place qu'occupe le lexique par rapport au *reste* de la grammaire. Les critères utilisés pour la sélection des unités lexicales semblent être identiques chez les néo-bloomfieldiens et chez les transformationalistes. Le lexique est une liste de tous les morphèmes du langage et de tous les groupes de morphèmes dont la signification ne peut être prédite en fonction de la signification individuelle des morphèmes qui constituent le groupe⁴.

Il est intéressant de noter que cette conception de l'inventaire lexical est précisément celle de Bloomfield. Rappelons que pour Bloomfield la notion d'inventaire s'oppose à celle de système de règles. Un inventaire est constitué par toutes les formes du langage qui ne se prêtent pas à la formulation de règles. Il semblerait donc que les transformationalistes s'accordent avec les néo-bloomfieldiens pour choisir la notion d'irrégularité comme critère distinctif du lexique. Le lexique est conçu comme une liste de tout ce qui est fondamentalement irrégulier dans le langage, alors que le reste de la grammaire rend compte des régularités du langage.

Si l'on considère maintenant le type d'information qui accompagne les unités lexicales en vue de spécifier la place qu'occupe le lexique par rapport au reste de la grammaire, on constate des différences profondes entre néo-bloomfieldiens et transformationalistes. Dans les deux cas on peut distinguer trois types d'informations : 1) information phonologique et morphophonémique ; 2) information morphosyntaxique ; 3) information sémantique.

3. Notons que chez les néo-bloomfieldiens aussi bien que chez les transformationalistes le lexique n'est pas toujours traité en un seul bloc. Chez les néo-bloomfieldiens il peut y avoir deux inventaires, l'un constitué par une liste des morphèmes inflectionnels et dérivatifs, l'autre constitué par une liste de « mots » ou un seul inventaire constitué par ces trois types d'unités. Chez les transformationalistes on peut également avoir ou bien deux inventaires, c'est-à-dire un lexique associé avec la syntaxe et la phonologie et un dictionnaire associé avec la sémantique, ou bien un seul inventaire constitué par les unités et les informations considérées pertinentes. Dans la discussion ces différences de traitement du lexique ne seront pas considérées, mon but étant de savoir dans leurs grandes lignes les caractéristiques de ces deux écoles.

4. Le point de vue des néo-bloomfieldiens est représenté par Charles Hockett (*A Course in Modern Linguistics*, Londres, MacMillan, 1958, p. 172) : « *If we are to be consistent in our use of the definition (of idioms), we are forced to grant every morpheme idiomatic status save when it is occurring as a constituent of a larger idiom, since a morpheme has no structure from which its meaning could be deduced.* » Le point de vue des transformationalistes est représenté par Bruce Fraser (« *Idioms within a Transformational Grammar* », *Foundations of Language*, vol. 6, n° 1, 1970, p. 22) : « *Shall regard an idiom as a constituent or series of constituents for which the semantic interpretation is not a compositional function of the formatives of which it is composed.* »

Chez les néo-bloomfieldiens, l'information sémantique se limite à une traduction dans le cas des mots et à une caractérisation sémantique superficielle dans le cas des morphèmes inflectionnels et dérivatifs. Il est à noter que cette information sémantique ne se rapporte à aucune partie du reste de la grammaire puisque l'analyse sémantique est exclue du domaine de la linguistique. Dans le cas des mots, la traduction fonctionne comme une étiquette qui relie une forme linguistique donnée à la connaissance du monde des sujets parlants.

Dans le cas des morphèmes inflectionnels et dérivatifs, la catégorisation sémantique fonctionne uniquement comme un code et elle pourrait tout aussi bien être un chiffre ou tout autre symbole, puisque son seul but est identificatif⁵ : elle permet d'aller chercher dans le reste de la grammaire la description grammaticale qui a trait à un morphème donné.

En ce qui concerne les autres types d'informations qui accompagnent les unités lexicales chez les néo-bloomfieldiens, on peut distinguer entre information implicite et information explicite. L'information phonologique est présente de façon implicite seulement puisque c'est le système adopté pour la notation des sons du langage qui la transmet. Quant à l'information morphophonémique et à l'information morphosyntaxique, elles sont clairement explicitées : telle forme est un allomorphe de telle autre forme ; tel morphème est un suffixe ; tel mot est un nom, etc.

Il est important de noter que ces informations ne sont jamais nouvelles par rapport à celles présentées dans la description grammaticale. Comme dans le cas de la caractérisation sémantique des morphèmes inflectionnels et dérivatifs mentionnés plus haut, ces informations n'ont qu'un seul but, celui de servir de code. Mais dans ce dernier cas, le code relie les formes constituant l'inventaire lexical aux règles qui ont pu être formulées à leur sujet.

Dans le cadre néo-bloomfieldien, le lexique n'est donc que la matérialisation des règles formulées dans le reste de la grammaire. De plus cette matérialisation n'est pas conçue comme intéressante du point de vue théorique puisqu'elle n'est pas sensée être elle-même soumise à des règles. Le lexique, une fois son existence reconnue, ne joue donc aucun rôle dans la théorie néo-bloomfieldienne⁶.

Chez les transformationalistes, l'information qui accompagne les unités lexicales a pour but non seulement de déclencher le fonctionnement des règles qui constituent les trois autres composantes de la grammaire (phonologie, syntaxe et sémantique) mais aussi de les faire fonctionner correctement⁷. Deux types d'informations peuvent donc être distingués : 1) l'information servant de code reliant les

5. Cela a été fait en particulier par les élèves de Voegelin. Paul Garvin (« Kutenai III : Morpheme Distributions, Prefix, Theme, Suffix », *IJAL*, vol. 14, n° 3, 1948, p. 171-187) utilise un système de chiffres indiquant la position qu'occupe chaque morphème à l'intérieur du mot.

6. Cela ne veut pas dire que la connaissance du lexique ne soit pas considérée comme utile pour l'analyse de certains faits linguistiques. Dans le cas des études de linguistique historique le lexique joue un rôle important. Les seuls néo-bloomfieldiens qui aient fait un effort vrai — même si c'est à une petite échelle — pour traiter le lexique, sont les élèves de Mary Haas dont l'intérêt pour l'étude historique comparative des langues amérindiennes est bien connu.

7. Cf. l'article de Bruce Fraser cité à la note 4.

formes constituant l'inventaire lexical aux règles qui constituent les trois autres composantes ; 2) l'information servant de complément aux règles formulées dans les trois autres composantes. Il est à noter que ce type d'information inclut à son tour non seulement des informations idiosyncratiques, mais aussi des règles non productives.

Le deuxième type d'information est ce qui distingue le lexique des transformationalistes du lexique des néo-bloomfieldiens. Il est manifeste que le lexique occupe une place très importante dans la machine à gérer le langage appelée grammaire par les transformationalistes : il déclenche le fonctionnement correct des règles et il donne l'inventaire des manifestations possibles de ces règles. Il est loin d'être clairement établi cependant quelle est exactement cette place, car les transformationalistes, quand ils considèrent ce problème, sont loin d'être d'accord.

Ni la conception du lexique des néo-bloomfieldiens ni celle des transformationalistes ne me paraît satisfaisante et cela pour la même raison : il n'est pas tenu compte du rôle que joue le lexique dans le langage vu comme instrument de communication⁸. À mon sens ce rôle constitue la caractéristique essentielle du lexique.

La définition du lexique que je propose est la suivante : le lexique est le système de référence dénomminative aux phénomènes distingués par une culture donnée⁹. Il est important de noter que ni la notion de référence ni celle de catégories fondamentales de phénomènes distingués par une culture donnée ne sont en elles-mêmes distinctives. La notion de référence s'applique à toutes les unités linguistiques ayant une signification dans la mesure où l'on peut dire qu'avoir une signification c'est se référer à quelque chose (le dénoter et/ou le connoter).

La notion de phénomènes distingués par une culture donnée peut s'appliquer à tout ce qui est signifié, par exemple à ce qui est signifié par des catégories grammaticales comme genre, nombre, aspect, quand elles sont présentes dans un langage donné ; elles indiquent des phénomènes distingués par la culture en question.

Ce qu'il y a d'essentiel dans la définition du lexique que je propose c'est la notion de référence dénomminative. Celle-ci implique l'existence d'unités linguistiques se distinguant d'autres unités linguistiques par le fait qu'elles dénomment, c'est-à-dire qu'elles fonctionnent comme des étiquettes dont on se sert pour se référer à quelque chose. Les phénomènes qui sont dénommés par ces unités se distinguent de tous les autres par le fait qu'ils sont dénommés.

La notion de référence dénomminative définit donc la fonction du lexique que je propose d'appeler la « fonction dénomminative ». Cette fonction oppose radicalement le lexique à la grammaire. Par grammaire, je veux parler de ce que tous les

8. Le même reproche peut être adressé à Halliday. Dans son article « Lexis as a Linguistic Level » (*In Memory of J.R. Firth*, Bazell, Catford, Halliday et Robins, édit., Londres, Longmans, 1966, p. 148-162), il caractérise le lexique grâce aux *collocational properties of lexical items*, ce qui n'est pas à mes yeux un critère fonctionnel.

9. Pour une formulation un peu différente, voir mon article « The Place of the Dictionary in Linguistic Description », *Language*, n° 43, 1967, p. 708.

linguistes semblent s'accorder à considérer comme des notions grammaticales, par exemple les notions de morphème, mot, racine, suffixe aussi bien dérivatif qu'inflexionnel, nom, verbe, sujet, objet, etc. Bien que je ne puisse offrir de définition qui rende compte de ces notions d'une manière satisfaisante¹⁰, il semble que ce qui les caractérise c'est qu'elles n'ont pas trait à la fonction dénomminative telle que je l'ai définie plus haut.

Vu le peu de faveur de la position adoptée ici, je crois qu'il n'est pas inutile d'insister sur le fait que proposer une définition du lexique qui se base sur des notions indépendantes de la grammaire, c'est considérer que lexique et grammaire constituent deux dimensions linguistiques autonomes, chaque dimension ayant sa structure propre, c'est-à-dire des unités, des relations et des propriétés qui lui sont caractéristiques¹¹. Cela revient à dire que la description du lexique comme celle de la grammaire devra comporter et un inventaire de formes (les unités) et un système de règles (ayant trait aux relations et aux propriétés de ces unités).

En ce qui concerne la place qu'occupe le lexique par rapport à la grammaire dans le cadre de référence qui est le mien, comme le dit Hymes au sujet des fonctions du langage, c'est quelque chose qu'il faut découvrir empiriquement et non postuler. Les deux systèmes sont beaucoup trop compliqués pour que la façon dont ils interagissent soit immédiatement apparente.

Les travaux pilotes que j'ai faits sur le lexique permettent dès maintenant une observation importante : il semble que les relations entre lexique et grammaire ne soient pas réciproques. Aussi toutes les unités lexicales ont un statut grammatical mais pas *vice versa*. Beaucoup de mots-outils, comme les particules ou les auxiliaires, n'ont pas automatiquement de statut lexical.

Le second problème théorique qui me paraît crucial pour l'investigation du lexique concerne la matière des unités lexicales. Les unités lexicales étant des unités linguistiques, deux caractéristiques générales s'appliquent à elles : 1) ce sont des entités à deux faces ; 2) il y en a de plusieurs types.

Pour ma part, la conception saussurienne des unités linguistiques comme étant des entités à deux faces a les conséquences suivantes : d'un côté, la notion d'entité requiert que les unités linguistiques aient une manifestation dans la substance linguistique ; de l'autre côté, la notion des deux faces des unités linguistiques (c'est-à-dire leur substance linguistique et leur signification) requiert que ces deux aspects ne soient pas dissociés l'un de l'autre. Ainsi, dire que les unités lexicales comme les autres unités linguistiques sont manifestées par une substance linguistique élimine la possibilité que ces unités soient des concepts semblables aux tagmèmes de Pike puisque les tagmèmes sont des relations (entre *slot* et *fillers*) et non des entités. De plus, dire que les unités lexicales ont comme les autres unités

10. Dans l'article cité précédemment (note 9) j'avais défini la grammaire comme *structure of discourse*. J'abandonne ici cette définition car elle implique que le lexique n'a pas de dimension syntagmatique.

11. Pour une discussion sinon de preuves à l'appui de cette position tout au moins de faits qui semblent indiquer son bien-fondé, voir mon article « Grammatical Problems in Lexicography : Grammatical versus Lexical Status », *International Conference on Lexicography in English*, New York, 5-7 juin 1972, New York Academy of Sciences (sous presse).

linguistiques et une substance linguistique et une signification qui sont inséparables l'une de l'autre, élimine la possibilité que ces unités soient des concepts ayant trait à la substance linguistique isolée de la substance linguistique.

Si pour l'analyser en détail le linguiste concentre son attention sur l'un ou l'autre des deux aspects des unités linguistiques, son étude ne doit pas aboutir, à mon avis, à l'élaboration de structures séparées. Aussi l'étude de la signification des unités lexicales ne doit-elle pas amener à l'élaboration d'une structure sémantique avec des unités qui lui sont propres, comme les sémèmes. Les concepts de structure sémantique et de sémèmes n'ont pas de place dans le cadre de référence que j'utilise ¹².

En ce qui concerne la différenciation des unités lexicales, je propose de la faire en fonction des quatre types d'unités linguistiques qui découlent de ma conception du système linguistique. Ces quatre types d'unités entrent dans les deux oppositions suivantes : 1) unités minimales s'opposant aux unités fusionnées ; 2) unités comportementielles s'opposant aux unités analytiques. Ces deux oppositions ne s'excluent pas mutuellement. Il y a donc : 1) des unités comportementielles minimales ; 2) des unités comportementielles fusionnées ; 3) des unités analytiques minimales et 4) des unités analytiques fusionnées.

Les notions d'unités minimales s'opposant aux unités fusionnées sont basées sur une conception gestaltiste du langage. La chaîne parlée n'est pas conçue comme une succession d'éléments qui s'égrènent comme les perles d'un collier ou pour reprendre l'image de Hockett comme « *the individual bricks... in a row of bricks* ¹³ ». Elle est caractérisée par la fusion des unités élémentaires dans des unités plus complexes.

À chaque niveau fonctionnel (phonologie, morphologie et lexique) les unités élémentaires sont considérées comme minimales et les unités plus complexes sont considérées comme fusionnées. Toute analyse dont le but est de déterminer les unités d'un système donné devra donc spécifier non seulement les unités minimales de ce système mais aussi les unités fusionnées.

La notion d'unités comportementielles s'opposant aux unités analytiques est basée, d'un côté, sur une conception du langage comme fait social et, de l'autre côté, sur une conception de l'analyse des faits sociaux comme devant tenir compte de leur réalité psychologique. Dans ce cadre, il est tenu comme indispensable que les critères utilisés au cours de l'analyse reposent sur l'observation du comportement des êtres humains, comportement qui manifeste le système que l'on veut comprendre. On arrive au système sous-jacent en observant et en manipulant le type de comportement des êtres humains qui manifeste ce système.

12. Je préfère parler d'un domaine sémantique qui à mon avis s'oppose à un domaine cognitif. Pour une discussion de ce point de vue, voir mon article « The Semantic and Cognitive Domains of Language », dans Paul Garvin, éd., *Cognition : A Multiple View*, Spartan Books, 1972, p. 251. Je retiens par contre le concept de traits sémantiques distinctifs comme utile pour l'analyse sémantique : ce sont les variables — ou dimensions — que l'on peut inférer en observant les contrastes formés par les divers aspects du sens des unités lexicales.

13. Cf. Hockett, *A Manual of Phonology*, Baltimore, 1955, p. 43. Ce point de vue est développé dans un article que j'ai écrit avec Paul Garvin, intitulé « Fused Units in Phonemics », dans Paul Garvin, éd., *On Linguistic Method Selected Papers*, La Haye, Mouton & Co., 1964, tout spécialement p. 12-13.

Les unités comportementielles sont les unités que l'on peut observer directement dans le comportement des sujets. Ce sont des unités dont les limites n'ont pas besoin d'être établies par l'analyse (par exemple : un poème, une réunion). Les unités analytiques sont les unités dont les limites doivent être établies par l'analyse (par exemple : les morphèmes).

Tout effort qui tend à déterminer les unités d'un système donné peut se résumer ainsi : on utilise les unités comportementielles qui sont les mieux à même de mener à la détermination des unités analytiques, d'abord les unités analytiques minimales, ensuite les unités analytiques fusionnées.

Les problèmes d'ordre méthodologique que je voudrais examiner ici ont trait à la délimitation des unités lexicales. Vu les considérations théoriques présentées en détail plus haut dans le cadre de référence qui est le mien, les unités lexicales doivent être délimitées grâce à l'utilisation de critères qui sont propres au lexique c'est-à-dire qui ont trait à la fonction dénomminative), les critères grammaticaux étant considérés comme non pertinents¹⁴. L'observation du comportement qui manifeste le système lexical permet d'établir quels sont les critères propres au lexique.

Ces critères sont utilisés pour spécifier quelles sont les unités comportementielles du lexique les mieux à même de mener à la détermination des unités analytiques du lexique (d'abord les unités analytiques minimales, ensuite les unités analytiques fusionnées). Comme je l'ai déjà proposé ailleurs¹⁵, parmi les types de comportements qui manifestent le système lexical, celui qui est directement associé à la fonction du lexique est le comportement dénomminatif (*naming behavior*). C'est l'action de dénommer les phénomènes différenciés dans une culture donnée.

Les unités qui manifestent ce genre de comportement, je les appelle des unités dénomminatives (*naming units*) : ce sont les manifestations linguistiques d'actes dénomminatifs individuels. Des exemples intéressants de comportement dénomminatif et d'unités dénomminatives furent observés à Québec le 28 novembre 1971 au cours d'un programme de télévision sur les termes culinaires intitulé « Le français d'aujourd'hui ». Dans ces exemples, le locuteur, un chef cuisinier, signale son

14. Comme je l'ai souligné ailleurs (cf. l'article mentionné à la note 11), les unités lexicales ne sont pas identiques aux mots d'un langage donné. « *The immediate consequence of postulating the independence of the lexicon from the grammar is that grammatical criteria are not believed to be in any way useful in ascertaining lexical units. Thus lexical units are not conceived of as the « words » of the language, the word being a concept that is regarded as applying only to grammatical units of a certain order. Although many lexical units do consist of single words each, the lack of one-to-one correspondence between lexical units and words manifests itself in two very important ways : 1) many lexical units consist of more than one word each ; 2) many words do not have lexical status ; they only have grammatical status. The first manifestation of a discrepancy between grammar and lexicon has recently been recognized by most lexicologists and lexicographers. The tendency, however, is to play it down by limiting its occurrence to frozen expressions. The second manifestation has been observed long ago : it is at the bottom of the distinction between function words and content words. The tendency in this case among both lexicologists and lexicographers is to disregard this distinction completely because it goes against the age-old conception of the lexicon as the inventory of all the words of the language.* »

15. « The Place of the Dictionary in Linguistic Description », *Language*, n° 43, 1967, p. 710-711.

propre comportement dénominatif par l'emploi d'expressions comme « on va dire », « ça se dit », « on peut parler de », « le terme c'est ». La convention orthographique des guillemets identifie dans le texte écrit les unités dénominatives. Les exemples sont les suivants :

... quand vous saupoudrez de la farine *les cuisiniers diront* « *singer* »...
 ... *on va dire* « verser une larme de cognac »...
 ... *ça se dit* (*ça s'appelle*) « les faire suer »...
 ... *le terme c'est* « les faire rissoler »...
 ... *on va dire* « une prise »...
 ... *on peut parler de* « la pincée »...
 ... *on va dire par exemple* « un soupçon »...

Afin de tirer parti du comportement dénominatif comme source d'information pour la définition des unités lexicales, je crois important de distinguer entre 1) la manifestation linguistique de l'acte dénominatif (les unités dénominatives étant un aspect de cette manifestation) et 2) ce qui est désigné par l'acte dénominatif, c'est-à-dire les phénomènes dénommés. Quand le comportement dénominatif est spontané, comme c'est le cas dans les exemples cités plus haut, les unités dénominatives sont données spontanément par le sujet parlant et le problème non résolu c'est celui de déterminer quelles catégories de phénomènes elles dénomment¹⁶.

Le problème est inversé aussitôt que l'on veut susciter le comportement dénominatif afin d'éliciter des unités dénominatives. Dans ce cas pour obtenir des unités dénominatives, on a besoin de partir ou bien de phénomènes qui soient clairement identifiables (par exemple : « comment appelle-t-on cela ? ») ou bien des catégories de phénomènes qui soient connues (par exemple : « comment appelez-vous ce que vous faites en ce moment ? »). Puisque les phénomènes clairement identifiables sont limités à un seul type, les phénomènes concrets, si l'on veut éliciter des unités dénominatives se référant à des types variés de phénomènes, il devient nécessaire de développer une procédure qui spécifie quelles sont les catégories de phénomènes dénommés dans la culture étudiée.

On peut donc distinguer deux approches dans la manipulation du comportement dénominatif : une approche directe et une approche indirecte. L'approche directe consiste à placer le sujet parlant dans une situation où il donne spontanément les noms de phénomènes qu'il distingue. C'est une situation où le sujet parlant enseigne une terminologie. Cette terminologie peut ou non être technique, quand le sujet parlant s'adresse ou bien à un enfant de sa propre culture ou bien à un étranger. Par contre, il semble qu'elle soit toujours technique quand le sujet parlant s'adresse à des adultes de sa propre culture (témoin : les termes culinaires présentés par le chef cuisinier à son auditoire québécois).

L'approche indirecte consiste à éliciter les noms que les sujets parlants donnent aux phénomènes qu'ils distinguent grâce à la connaissance du système de

16. Les catégories de phénomènes dénommés dans les exemples cités plus haut semblent être des actions concrètes (« *singer* », « les faire rissoler ») et des quantités concrètes (« une pincée », « un soupçon »). Il est intéressant de noter que cela contredit la croyance très répandue parmi les lexicologues (Josette Rey-Debove est un exemple) que ce que j'appelle le comportement dénominatif n'a trait qu'à un seul genre de référence, la référence aux entités (les noms-noms étant les unités lexicales par excellence).

classification qu'ils emploient pour distinguer ces phénomènes. C'est cette dernière approche que je voudrais examiner ici en détail ¹⁷.

Étant donné les considérations précédentes, on peut dire que la manipulation indirecte du comportement dénominatif inclut deux tâches successives : 1) la détermination des phénomènes dénommés dans une culture donnée ; 2) l'analyse de la manifestation linguistique de l'acte dénominatif.

La première tâche peut être précisée grâce à deux questions : a) quels sont les phénomènes distingués par la culture ? b) tous les phénomènes distingués par la culture sont-ils dénommés ?

La deuxième tâche peut, elle aussi, être précisée grâce à deux questions : a) comment procéder à l'élicitation d'unités dénomminatives individuelles ? b) comment, en partant de ces unités dénomminatives individuelles, arriver à la détermination des unités analytiques du lexique (unités minimales et unités fusionnées).

On notera que 1b et 2a) sont très voisins, puisque dans le cas des phénomènes distingués par la culture et dénommés on devra pouvoir éliciter des unités dénomminatives se référant à ces phénomènes. Par conséquent, les trois tâches spécifiées par les questions 1a), 2a) et 2b) peuvent être considérées comme constituant les étapes principales de la délimitation des unités lexicales. Ces trois étapes sont les suivantes : 1) détermination des phénomènes distingués par une culture donnée (à la fois catégories de phénomènes et phénomènes individuels) ; 2) détermination des unités comportementielles du lexique : *unités dénomminatives*, c'est-à-dire cas individuels de référence dénomminative aux phénomènes distingués par la culture en question ; 3) détermination des *unités analytiques* du lexique : a) *termes*, c'est-à-dire familles d'unités dénomminatives (incluant des variations grammaticales et sémantiques non distinctives) ; b) structure lexicale interne des termes : *lexèmes* (unités minimales), *groupes de lexèmes* (unités fusionnées du premier ordre) et agglomérations de *lexèmes* (unités fusionnées du deuxième ordre) ¹⁸.

La notion de phénomènes distingués par une culture donnée est cruciale pour l'étude du lexique pour la raison suivante : les catégories de phénomènes et les phénomènes individuels une fois spécifiés fournissent la base de comparaison nécessaire à l'observation de la façon en apparence extrêmement variée et changeante dont les sujets parlants se réfèrent à ces phénomènes dans les divers dialectes fonctionnels et sociaux. L'observation de cette variation permet de spécifier ce qui caractérise la référence dénomminative par comparaison avec les autres modes de référence aux phénomènes distingués par la culture étudiée.

Comme je l'ai suggéré ailleurs ¹⁹, deux critères semblent caractériser la

17. L'hypothèse de base est que le système de classification des phénomènes distingués par une culture donnée comporte un nombre fini de variables. Si cette hypothèse est justifiée, la connaissance de ce système de classification offre un moyen de décrire de façon exhaustive le système lexical bien que celui-ci soit un système ouvert, c'est-à-dire contienne un inventaire presque illimité et toujours changeant d'unités.

18. « The Place of Dictionary in Linguistic Description », *Language*, n° 43, 1967, p. 717-721.

19. Voir « Review of *La lexicographie* », édit. par Josette Rey-Debove, *Language* (à paraître). « These 2 criteria can serve to identify naming behavior in two types of use of the lexicon, the technical and non-technical, as follows : in both the technical and non-

référence dénomminative : 1) la conscience que les sujets parlants ont d'employer des « noms » et 2) la stabilité avec laquelle certaines expressions sont employées pour se référer à des phénomènes donnés.

Dans ce qui suit je me bornerai à la considération des problèmes présentés par la première étape dans la détermination des unités lexicales. Rappelons la question qui s'y pose : comment procéder à la détermination des phénomènes distingués dans une culture donnée ? Les travaux pilotes que j'ai faits en vue de répondre à cette question suivent deux approches bien différentes. Ici encore on peut distinguer entre approche directe et approche indirecte.

Les deux hypothèses de base de l'approche directe s'inspirent du système de classification proposé par Aristote (entités, actions, événements et qualités). La première hypothèse propose que les catégories de phénomènes distingués dans une culture donnée soient dénommées par des expressions classificatrices en usage parmi les membres de cette culture. La deuxième hypothèse propose qu'il n'y ait qu'un petit nombre de ces catégories et par conséquent qu'il y ait un petit nombre d'expressions classificatrices opérant à un niveau d'abstraction élevé dans la culture étudiée.

Le tout dans l'approche directe est donc d'obtenir des expressions classificatrices de types variés, mais à un même niveau d'abstraction et non des expressions classificatrices d'un même type, mais à des niveaux d'abstraction inférieurs (comme par exemple, « animal sauvage », « animal domestique », « chien », etc.). La méthode employée pour obtenir ces expressions classificatrices consiste à présenter aux sujets parlants des expressions délimitées de manière impressionniste et choisies comme ayant trait à des concepts en apparence aussi différents les uns des autres que possible.

Un essai d'application de cette méthode²⁰ avec comme sujets parlants des Américains choisis au hasard, mais ne faisant pas partie du corps enseignant, révèle trois faits importants s'ils sont généralisables :

1) Les catégories de phénomènes dont les sujets parlants sont conscients semblent être à un niveau d'abstraction peu élevé. Par conséquent, les catégories que l'on obtient par l'élicitation directe sont extrêmement nombreuses et les relations entre ces catégories sont loin d'être flagrantes. Le système de catégorisation des phénomènes que l'on veut découvrir est donc extrêmement complexe et non relativement simple comme la catégorisation proposée par Aristote semblait l'indiquer.

2) Les catégories de phénomènes dont les sujets sont conscients ne semblent pas toujours être dénommées. On peut conclure par conséquent que le système de classification des phénomènes n'est pas manifesté de façon homogène : il n'est qu'en partie dénommé par des expressions classificatrices qui, de plus, opèrent à des niveaux d'abstraction différents.

technical use of the lexicon naming behavior is characterized by a high degree of stability of lexical units. The technical differs from the non-technical use of the lexicon in the degree of awareness that native speakers have of naming behavior : a high degree in technical use, low degree in non-technical use. »

20. Dans ce cas les travaux pilotes ont été faits par Revere Rerkins.

3) La délimitation des expressions à présenter aux sujets parlants pour qu'ils les classifient offre de grandes difficultés aussitôt que l'on sort du domaine des expressions se référant à des faits concrets.

Ces observations tendent à prouver que la méthode directe ne semble pas indiquée pour la découverte du système de catégorisation des phénomènes distingués par une culture donnée. L'approche indirecte que j'ai investiguée est basée sur l'utilisation d'un comportement culturel qui révèle une classification des phénomènes distingués par une culture donnée sans qu'il y ait lieu de se soucier si les catégories sont ou non dénommées ²¹.

Le comportement culturel que j'ai utilisé est constitué par les questions stéréotypées que posent les membres d'une culture avec les réponses qui leur sont données. Par questions stéréotypées, je veux dire celles qui jouent un rôle important dans la culture étudiée et non pas celles qui, bien qu'étant grammaticalement « possibles », ne sont pas employées dans la vie quotidienne ²². Une double hypothèse est à la base de cette approche :

1) Les questions étant des demandes d'information, l'information contenue dans les questions stéréotypées que posent les membres d'une culture et dans les réponses qu'ils donnent à ces questions, révèle la façon dont ils catégorisent le monde qui les entoure. Plus précisément, les questions stéréotypées que posent les membres d'une culture révèlent les catégories de phénomènes qu'ils distinguent. Les réponses données à ces questions révèlent les phénomènes individuels qu'ils distinguent ;

2) Les phénomènes distingués dans les questions et leurs réponses ont une forte chance d'être dénommés. Les phénomènes individuels ont plus de chance d'être dénommés que les catégories de phénomènes.

L'utilisation des questions et de leurs réponses pour extraire le système de classification des phénomènes distingués par une culture donnée qu'elles révèlent implique deux spécifications préalables : 1) une distinction entre vraies questions et pseudo-questions ; et 2) une distinction entre vraies réponses et pseudo-réponses.

Les vraies questions sont des demandes d'information. Les phrases qui ont la forme d'une question, mais qui ne sont pas des demandes d'information, sont des pseudo-questions. Un exemple de pseudo-question est : *Who do you think Sam ?* « Pour qui me prenez-vous ? » Comme l'indique la paraphrase explicative :

21. Notons que l'hypothèse de base était identique à celle des ethnoscientistes. Mais comme ils ont fini par le signaler eux-mêmes (cf. Dennis Breedlove et Peter Raven, « Covert Categories and Folk Taxonomies Breut Berlin », *American Anthropologist*, 70, 1968, p. 290-299), l'absence de dénominations n'indique pas forcément un manque de différenciation dans le système conceptuel des sujets parlants.

22. Notons que les questions et leurs réponses peuvent aussi être utilisées à des fins grammaticales (cf. Paul Garvin, « Behavioral Tests in Linguistics », dans Paul Garvin, éd., *Method and Theory in Linguistics*, La Haye, Mouton & Co., 1972). Notons également que l'utilisation des questions stéréotypées et des réponses qui leur sont faites est à la base des travaux des ethnoscientistes. Ces derniers utilisent les questions afin d'obtenir des informations. Ils n'analysent pas les questions elles-mêmes comme je propose de le faire. Un effort semblable au mien a été fait par Harvey Sarles (cf. son article, « An Examination of the Question-Response System in Language », *Semiotica*, vol. II, n° 1, 1970, tout spécialement p. 84-85).

« je ne suis pas ce que vous avez l'air de croire que je suis », cette phrase n'est pas une demande d'information.

Un autre type de pseudo-question est constitué par les demandes de répétition de l'information, demandes dues au fait qu'on n'a pas compris le message la première fois. En voici un exemple : *You didn't find what ?*

Les vraies réponses sont celles qui fournissent l'information requise par la question posée. Les réponses qui indiquent ou bien un refus de fournir l'information requise par la question (par exemple : *None of your business*) ou bien une incapacité de fournir cette information (par exemple : *don't know*), sont des pseudo-réponses.

Les travaux pilotes que j'ai faits sur le système des questions et réponses de l'anglais des États-Unis suggèrent qu'elles ont trait à quatre types principaux d'information : 1) explications ; 2) directives ; 3) décisions impliquant un choix entre oui et non ; 4) informations substantives.

Les explications sont manifestées par des questions du type *Why X ?* ou *How come X ?* et des réponses du type (*Because*) *Y ?* Les directives sont manifestées par des questions du type *How do you do X ?* ou *How does one get to X ?* et des réponses du type (*first*) *A, (then) B, etc.* Les décisions impliquant un choix entre oui et non sont manifestées par des questions ayant trait à l'existence d'une certaine condition (par exemple : *Are you going out tonight ? Do you want another drink ?*) et des réponses allant de l'affirmative à la négative (par exemple : *may be, I may, I might, I think so, of course, etc.*). Les informations substantives sont manifestées par les nombreuses questions qui ne sont ni des demandes d'explications ni des demandes de directives ou des demandes de décisions et les réponses à ces questions. Des exemples de questions substantives et de leurs réponses sont :

Q.1 *What did you have for lunch today ?*
R.1 *Turkey.*

Q.2 *What did you do last night ?*
R.2 *I watched TV.*

Q.3 *What did you do at school today ?*
R.3 *We read.*

Q.4 *What's your name ?*
R.4 ...

Q.5 *How tall are you ?*
R.5 *5.6.*

Q.6 *How old are you ?*
R.6 *6 and a half.*

Q.7 *When did you get back ?*
R.7 *Last night.*

Q.8 *How long have you been waiting ?*
R.8 *About 5 minutes.*

Q.9 *At what time did you get back ?*
R.9 *At 5 P.M.*

Notons que les informations substantives peuvent être manifestées par des questions et des réponses semblables à celles qui manifestent des décisions impliquant un choix entre oui et non.

Aussi une façon coutumière de s'enquérir de l'état civil d'une personne et d'y répondre, c'est-à-dire de faire une demande d'information substantive et de donner des informations substantives est la suivante :

Q.10 *Is she married ?*

R.10a) *No, she is an old maid (i.e. she never married).*

b) *No, she is single (she has not yet married).*

c) *No, she is a widow (she is no longer married because her husband died).*

d) *No, she is divorced (she is no longer married).*

e) *No, she has a lover
she has a companion
she lives with a guy* } (*she is not quite married*).

f) *Yes, she is (married).*

Un autre exemple du même genre :

Q.11 *Is he working ?*

R.11a) *No, he is a student.*

b) *No, he is retired.*

c) *No, he is unemployed on welfare.*

d) *No, he is independently wealthy.*

e) *Yes, he is (working).*

Si l'on compare les quatre types d'informations distingués plus haut (explications, directives, décisions et informations substantives) encore en vue d'établir leur relation avec la notion de phénomènes distingués par une culture donnée, on remarque que : 1) les explications et les directives sont constituées par des informations substantives ; 2) les informations substantives peuvent être constituées par plusieurs types d'informations substantives ; 3) les décisions semblent être indépendantes des informations substantives. Les deux types d'informations qui sont indépendants l'un de l'autre sont donc les décisions et les informations substantives.

Il est possible, mais non évident, que les décisions puissent être considérées comme des catégories de phénomènes — je préfère laisser cette question sans réponse pour le moment. Par contre, il est évident que les informations substantives sont ou bien des catégories de phénomènes ou bien des phénomènes individuels distingués par une culture donnée.

Puisque les informations substantives sont si nombreuses, le problème immédiat qui se présente est celui de les grouper d'une façon arbitraire. Un critère de groupement est flagrant : la relation entre réponses à une question donnée et cette même question. Les phénomènes individuels manifestés par les réponses à une question donnée appartiennent à la catégorie de phénomènes manifestés par la question.

Dans le cas de la catégorie de phénomènes manifestés par les questions, le problème est beaucoup plus complexe : il y a trop de questions substantives et par conséquent trop de catégories de phénomènes pour qu'on ne cherche pas à les grouper à un niveau d'abstraction plus élevé. Mais en dehors de certains cas, comme les notions temporelles et les notions de bien, les critères de groupement ne sont pas flagrants²³. Il faut donc les découvrir. La poursuite d'une

23. Je ne trouve pas flagrant, par exemple, que le *what* de la Q.1 et celui de *what's your name ?* aient quelque chose en commun. Dans le premier cas il s'agit de quelque chose que l'on mange. Dans le second cas il n'est même pas clairement établi de quoi il s'agit. Pour la plupart des sujets parlants un nom propre est un nom propre et n'a rien de commun avec un morceau de poulet ou un bifteck aux pommes frites.

telle tâche est en dehors des limites de cet exposé. Mon but était ici seulement de spécifier la notion de phénomènes distingués par une culture donnée afin d'insister sur son importance pour l'étude du lexique.

L'utilisation des questions stéréotypées et de leurs réponses, pour spécifier quels sont les phénomènes distingués par une culture donnée, offre à mon avis trois avantages extrêmement importants pour l'étude du lexique :

- 1) C'est un comportement extrêmement fréquent que l'on peut par conséquent observer et manipuler d'une façon naturelle ²⁴.
- 2) Les questions stéréotypées et leurs réponses fournissent un critère sans ambiguïté pour différencier entre les cas d'homonymie et les cas de polysémie. Des formes linguistiques semblables, mais qui constituent des réponses à des questions manifestant des catégories différentes de phénomènes, sont des cas d'homonymie. Par exemple, les formes *he writes* et *he is writing a letter* sont des cas d'homonymie comme le montrent les questions pour lesquelles ces formes pouvaient constituer des réponses.

Q.1 *What does he do for a living ?*

R.1 *He writes* (il est écrivain).

Q.2 *What is he doing ?*

R.2 *He is writing a letter* (il est en train d'écrire une lettre).

- 3) Les questions stéréotypées et leurs réponses permettent de spécifier quels sont les éléments d'un langage donné qui n'ont pas de statut lexical mais seulement un statut grammatical. Comme j'ai essayé de le montrer en détail ailleurs ²⁵ : « *When the utterances of a given language are analysed in terms of whether or not all the words or word sequences that compose them may constitute possible answers to substantive questions, the answer is negative : not all of them necessarily do. The words or word sequences which cannot constitute minimal answers to substantive questions are considered to have no lexical status. They have only grammatical status.* »

Pour conclure cet essai, je voudrais mentionner deux conséquences importantes qui découlent de la nécessité théorique de considérer le lexique comme autonome de la grammaire, c'est-à-dire comme constituant un système séparé du système lexical. La première conséquence a trait à la nécessité de spécifier en détail les relations qui existent entre le système lexical et le système grammatical. Proposer que l'investigation du système lexical se fasse sans recourir à des critères grammaticaux n'implique pas qu'il n'y ait aucune relation entre lexique et grammaire. Je crois au contraire que la compréhension profonde de ces relations constituerait un apport théorique très important pour l'étude du langage. La deuxième conséquence a trait au système lexical lui-même. Dire que le lexique est un système implique l'existence d'unités de relations et de propriétés plus complexes et plus variées que celles dont il a été question dans cet essai (les

24. Mary Black a fait des observations très intéressantes sur le comportement des Ojibwas dans son article « Ojibwa Questioning Etiquette and Use of Ambiguity », *Special Issue of Studies in Linguistics*, 1972.

25. Voir note 11.

variétés étant les lexèmes, les groupes de lexèmes et les agglomérations de lexèmes). Quelles pourraient être ces unités, ces relations et ces propriétés ?

Pour en avoir une idée, il me semble qu'il faudrait comprendre mieux quel est le rôle du lexique dans les nombreux types de comportements linguistiques que l'on peut différencier dans une culture donnée (conversations, narrations, sermons, etc.). Je veux dire par là que le rôle que joue le lexique dans le langage vu comme instrument de communication est sans doute multiple. Le lexique a aussi d'autres fonctions en dehors de sa fonction dénomminative. La compréhension de ces autres fonctions permettrait d'établir quels sont les divers niveaux fonctionnels du lexique et par conséquent de rendre compte du système lexical dans toute sa complexité.

MADELEINE MATHIOT