

AVANT-PROPOS

La question de la validité réside au cœur de la démarche scientifique. Par voie de spécification, elle intéresse donc pleinement les sciences du langage et, plus particulièrement, les études ayant pour objet le lexique. L'analyse lexicale relève en conséquence d'une épistémologie générale qui, dans ses traits fondamentaux, caractérise l'ensemble des méthodes et des raisonnements constitutifs d'un *savoir rationnel*.

Mais certains traits de structure des grandeurs lexicales contribuent à spécialiser les principes généraux de validité qui légifèrent la pratique scientifique, et même, contribuent à élargir la problématique de la validité. Rappelons en effet qu'un des traits les plus fondamentaux des entités lexicales est leur ambivalence relationnelle. Ainsi — en admettant, par pure commodité, l'opposition schématique entre grammaire et lexique — à la différence des unités grammaticales qui relèvent exclusivement d'un réseau de relations internes au système de la langue, les éléments du lexique engagent des rapports sur deux plans distincts. D'une part, au plan *interne*, un vocable est caractérisé par sa position relationnelle au sein du système stratifié que compose le lexique d'une langue. Cette position qualifie le *sens* du vocable. D'autre part, au plan *externe*, un vocable est défini par des relations d'*échange*, ou relations de *signification*, établies entre lui et des occurrences d'objets et de situations relevant d'un ordre empirique extérieur à la langue. On aura reconnu les deux composantes de la notion saussurienne de "valeur".

Dans le champ de l'analyse et des traitements du lexique, les principales figures que prend la question de la validité se laissent alors déduire de cette configuration structurale qui conjugue le plan des formes internes du lexique et celui des substances empiriques externes, suivant le mode d'une *incorporation* par laquelle s'accomplit la catégorisation linguistique des univers d'expérience. Pour l'essentiel, ces figures, qui se recouvrent largement, sont au nombre de trois.

En premier lieu, se pose la question de la validité de la description des formes internes du lexique. Il s'agit, entre autres, de questions portant sur la délimitation des unités, sur leur composition et sur les types de connexions fonctionnelles qu'elles entretiennent. La seconde grande question concerne la validité des qualifications de la substance du contenu en tant qu'elle se trouve ordonnée suivant les formes du lexique. Il s'agit

là du problème plus général de la description sémantique. La troisième question, enfin, est celle de la *légitimité* des connexions établies entre les plans (et leurs éléments) de la forme et de la substance — connexions produisant des catégories et des identités de sens toujours réajustées et renégociées par les partenaires en parole. Cette dernière question, qui ouvre au vaste problème de l'acceptation, du renouvellement et du *bon droit* des usages du lexique, présente elle-même de multiples facettes, dont les principales concernent la recension (délimitation et enregistrement) des unités lexicales en fonction de domaines d'expérience ou de situations d'usage, et la reconnaissance des catégorisations que ces unités réalisent.

Les contributions que rassemble ce volume présentent autant d'approches visant ce même objet : le lexique. Chacune d'elles l'aborde suivant des formes descriptives et des finalités pratiques particulières, et, mettant ainsi l'accent sur telle ou telle caractéristique structurale du lexique, rencontre la question de la validité/légitimité sous une ou plusieurs facettes privilégiées.

Dans un premier article, **F. Henry** traite les questions de la validité et de la légitimité des analyses dans la pratique lexicographique. D'un côté, un dictionnaire doit fournir au lecteur une description fondée des unités lexicales. Aussi, le travail lexicographique va suivre quelques grands principes sur lesquels la linguistique contemporaine est parvenue à un certain consensus (informations combinatoires et morphosyntaxiques, informations paradigmatiques, type et structure des définitions, phraséologie...). Il s'agit donc de principes conçus indépendamment d'une production matérielle et socio-culturelle particulière, et qui visent une authentique objectivation des grandeurs lexicales. Il en est de même pour d'autres critères, généraux (intégration, cohérence, exhaustivité...) ou plus particuliers (contraintes philologiques, technique des collocations...), qui guident le travail des rédacteurs et dont F. Henry donne une description minutieuse. Mais la confection d'un dictionnaire est aussi tributaire d'un ensemble de contraintes *externes*. Ainsi, par exemple, l'existence d'une tradition lexicographique, les positions et fonctions socio-culturelles du dictionnaire, les limitations matérielles, le public visé, sont autant de contraintes qui motivent certains choix rédactionnels (nomenclature, catégories d'informations...) et qui conduisent à fournir au lecteur une description ainsi légitimée (et légitimante) des mots de sa langue.

Alors que les dictionnaires traditionnels, destinés exclusivement à des êtres humains, tolèrent un certain degré d'implicite ou certaines lacunes que la compétence du lecteur est à même de combler, les dictionnaires électroniques, conçus à des fins de traitement automatique ou pour assister le lecteur dans certaines tâches coûteuses en temps, rendent nécessaires l'explicitation des informations linguistiques. Dans sa contribution, **D. Le Pesant** expose comment le concept de "classe d'objets", introduit

par G. Gross, qui permet de produire une description intégrée et extensive des grandeurs lexicales, satisfait aux contraintes fonctionnelles d'un dictionnaire électronique. Le concept de "classe d'objets" est ainsi doublement fondé. Dans la perspective d'une analyse linguistique, il permet d'affiner considérablement les analyses syntactico-sémantiques — ce qui établit sa validité. Chaque classe d'objets (définie en extension) est caractérisée par un ensemble précis de contraintes syntaxiques, de restrictions distributionnelles (notamment en faisant appel à des opérateurs "généraux" et "appropriés"), de latitudes transformationnelles... et les unités rassemblées dans une même classe d'objets partagent des propriétés sémantiques communes. Les classes d'objets permettent ainsi l'intégration d'informations de différents types dans une structure ordonnée, unitaire et explicite. Les données décrites et organisées suivant le schéma des "classes d'objets" satisfont alors aux conditions architecturales des dispositifs informatiques qui, en retour, légitiment au plan opératoire un tel schéma descriptif.

B. Fradin aborde pour sa part la question de la détermination des unités lexicales indépendamment de toute fin pratique (rédactionnelle ou informatique). Sa contribution vise à reprendre à la lumière des développements récents en morphosyntaxe certains problèmes centraux, notamment ceux concernant le choix, la délimitation et le format d'encodage des unités lexicales. Après avoir discuté des principes généraux de délimitation des grandeurs lexicales et des critères motivant l'enregistrement des unités atomiques de langue dans le lexique, B. Fradin présente deux méthodologies qui articulent différents critères d'analyse (distribution, grammaticalité, compositionnalité sémantique...) aux plans syntagmatiques et paradigmatiques en vue d'une délimitation et d'une qualification rigoureuses des unités lexicales. Ces procédures permettent d'établir une distinction, à un premier niveau, entre les mots-forme et les expressions dépendantes, et, à un second niveau, entre différentes identités morphologiques (lexèmes, affixes et clitiques). Il s'agit ainsi, en prenant appui sur des méthodes explicites et détaillées, de fonder méthodologiquement l'existence empirique d'entités lexicales. Enfin, B. Fradin aborde le problème de la typologie des unités lexicales, et tout particulièrement, celui du statut des formes complexes produites par les différents procédés morphologiques.

Bien qu'essentiellement consacrée aux fondements théoriques de la description morpholexicale, la contribution de B. Fradin n'est pas sans rencontrer cependant des préoccupations plus finalisées. Ainsi, par exemple, le problème du stockage des unités (choix des catégories descriptives, spécification des rapports lexème/forme de mot...) concerne bien évidemment tout système de traitement de données langagières, notamment dans l'optique d'une optimisation de la taille des bases. C'est justement de ce problème, et des nombreuses questions afférentes, que traite l'article de **J.-M. Pierrel** et de **L. Romary**.

Après avoir rappelé la position centrale du lexique dans les systèmes actuels de traitement de données en langue naturelle, J.-M. Pierrel et L. Romary examinent quelles ressources lexicales (unités et types d'informations) doivent être mises en œuvre dans l'objectif d'un dialogue homme-machine finalisé. Dans un tel cadre applicatif — et, en quelque sorte, à l'instar du lexicographe (cf. article de F. Henry) qui doit rechercher un équilibre entre les exigences de l'analyse lexicologique et les finalités pragmatiques de l'ouvrage *dictionnaire* — la base lexicale d'un système de dialogue doit satisfaire, d'un côté, à un ensemble de contraintes applicatives spécifiques et, de l'autre, aux exigences génériques du système linguistique. Or, il se trouve que nombre de contraintes spécifiques, qui conditionnent l'efficacité et l'adéquation de l'interface de dialogue, débouchent sur des déterminations qui ne sont pas toujours compatibles avec les fonctionnements généraux du système de la langue. Ainsi en est-il, par exemple, des productions dérivationnelles, du vocabulaire restreint, de la caractérisation pragmatique et référentielle des contenus lexicaux ou encore des qualifications syntactico-sémantiques. Comme les lexiques d'applications différentes (et les informations qu'ils contiennent) ne sont généralement pas superposables, la validité de telles ressources lexicales semble difficile à établir par voie de généralisation. Il s'ensuit que le choix du lexique et de ses déterminations ne peut être que légitimé dans l'usage. Dans la dernière partie de leur contribution, J.-M. Pierrel et L. Romary réfléchissent en conséquence sur la structure d'un "métalexique" générique. Il s'agirait d'un système d'accès à l'ensemble des données textuelles et lexicales disponibles sous format électronique et préalablement normalisées. Un tel système, dont le principe rejoint les exigences philologiques exposées par F. Henry, permettrait alors d'initier, d'alimenter et de valider les descriptions lexicales.

Le problème de la spécialisation sémantique et référentielle des unités lexicales prend un tour particulier dans le champ de la terminologie. C'est dans cette composante du lexique — que J. Humbley nous invite à examiner — qu'apparaît le plus clairement le caractère d'acte volontaire de la légitimation : plus que dans le reste de la langue, la légitimation en terminologie est affaire de convention, de consensus auxquels arrivent les usagers prioritaires, les spécialistes d'un domaine — même si les modalités de cette légitimation varient en fonction des grands secteurs terminologiques (scientifique, technique et juridique). Mais la pratique terminologique, qui se fonde sur des méthodes et produit des résultats, rencontre aussi la question de la validité de ses descriptions. C'est dans cette perspective que J. Humbley examine la notion d' "arbre du domaine" qui sert à la fois d'outil de construction de terminologie et d'outil de vérification et de validation. Mais l' "arbre du domaine" est une notion controversée car la validité du découpage en domaines de spécialité

cloisonnés semble dépassée dans le contexte de la production moderne. Une autre critique, plus fondamentale, adressée à la terminologie vient des sociolinguistes qui reprochent la distinction faite entre termes et mots. De fait, dans leur usage, les termes semblent bien plus proches (par leur phraséologie, par leur polysémie...) des mots que des concepts. Un autre aspect de la légitimation en terminologie est celui de l'aménagement linguistique, et à ce titre, l'article de J. Humbley examine, d'une part, les configurations et les circonstances politiques, culturelles et sociales qui favorisent l'implantation d'un terme et, d'autre part, la nécessité et l'efficacité du recours à des procédés connus de la néologie ou aux relations de sens en langue.

L'article de **D. Piotrowski**, enfin, fait écho à la difficulté méthodologique que soulève la notion d' "arbre du domaine" — celui-ci constituant un outil de description *et* de validation. Cette configuration de bouclage qui affecte l'ensemble des sciences du langage, fait des théories linguistiques des théories non réfutables. Dans sa contribution, D. Piotrowski propose un traitement de cette difficulté et, à cette fin, commence par la situer dans le cadre d'une épistémologie générale. Le problème de la connexion constituante entre théories et données d'observation est en effet un problème *classique* de l'épistémologie. Une lecture de l'approche poppérienne permet de montrer que la sortie du cercle vicieux, où tombe une théorie T_1 , procède de l'existence d'un système de qualifications T_2 *croisant* les données visées par la théorie T_1 . La cadre T_2 constitue alors une base d'évaluation autonome et consistante des descriptions apportées par T_1 . Or il se trouve que les théories linguistiques ne disposent pas de grilles théoriques externes *croisant* leurs données empiriques. Il s'ensuit que la validité des théories linguistiques doit reposer sur un principe *affaibli*, qui consiste à exiger des édifices théoriques qu'ils rendent compte non pas des propriétés de leurs objets, mais des modalités fonctionnelles qui produisent ces objets.

Je voudrais, pour clore cette présentation, remercier les contributeurs qui ont nourri ce volume de leur expérience, de leurs problématiques et qui nous ont aussi fait part des doutes et des difficultés qu'ils rencontrent dans leur activité scientifique. Je remercie aussi la direction de *Sémiotiques* qui a bien voulu nous confier la responsabilité de ce numéro, et F. de Chalonge pour son efficace travail rédactionnel.

David Piotrowski

