

Semantique et Mecanographie¹

by Georges Mounin, C.N.R.S., Paris

La sémantique est un domaine où on a peu progressé. Les travaux de Gardin proposent une analyse sémantique neuve. Pour des utilisations mécanographiques, elle substitue aux noms du langage ordinaire des noms de code, non-arbitraires par rapport aux choses nommées, fondés sur les traits définitoires des choses. Ces travaux offrent une méthode pour traiter de sémantique sans recourir aux langues naturelles et mettent en évidence des unités sémantiques minima, plus petites que le mot, grâce auxquelles on peut structurer certains vocabulaires techniques.

1. Malgré le nombre et l'intérêt des travaux qui la concernent depuis deux ou trois décades (ceux de Trier, ou de Zinsli, de Matoré, de Quemada, par exemple, ceux de Guiraud), malgré des traités à jour comme celui d'Ullmann, on reconnaît généralement que la sémantique est le domaine linguistique où l'on a le moins avancé depuis trente ou quarante ans. Rien de comparable ici, en effet, aux grandes constructions systématiques de la phonologie, ou de la linguistique descriptive formelle ou structurale, ou distributionnelle. Rien qui ressemble à la mise en lumière d'éléments constituants de la sémantique (en tant que "système des significations"), constituants qui soient comparables aux phonèmes, aux morphèmes ou aux monèmes, quant à leur importance fonctionnelle dans le système. Rien de comparable à la mise en évidence d'éléments derniers aussi fondamentaux que celle des *traits pertinents* en phonologie.

2. Non que l'on ne cherche pas dans ces directions. La critique, que Bloomfield a faite si longuement, de la notion de *signification* (de "sens", de *meaning*) était une longue invite à constituer scientifiquement la sémantique sur une base béhavioriste. Les efforts de Hjelmslev pour atteindre un statut scientifique de la *substance du contenu* [sémantique] des signes linguistiques, et ses efforts pour isoler des *figures de contenu* qui seraient à la sémantique ce que les phonèmes sont à la phonologie, — ce seraient les atomes sémantiques par excellence, — sont une autre invite: à prendre, cette fois, les problèmes sémantiques fondamentaux par le biais de la logique formelle. C'est dans le même sens que vont les efforts de Prieto pour isoler des *traits pertinents* de contenu sémantique.

Malgré la séparation, devenue traditionnelle, et de rigueur, entre linguistique et psychologie, entre langage et logique, entre l'étude des mécanismes de la langue et l'étude des mécanismes de la pensée (séparation qui était une réaction tardive contre la vieille grammaire logique, ou logique grammaticale, des siècles antérieurs), les logiciens des mathématiques, — les Russell, les Tarski, les Carnap, — ont proposé des théories et des terminologies sémantiques qui pénètrent peu à peu en linguistique (par le canal de recherches sémiologiques comme celles de Charles Morris, de Hjelmslev ou de Borgstrøm, ou de Sørensen). Cette interpénétration

nouvelle (hérétique par rapport à la pensée linguistique des années 1920-1930) de la logique et de la linguistique doit être accueillie avec précaution, mais avec beaucoup d'attention aussi. En effet, peut-être aidera-t-elle à vaincre la difficulté fondamentale de l'analyse sémantique, aperçue et formulée nettement par André Martinet: "la difficulté qu'on éprouve à manipuler la réalité sémantique sans le secours d'une réalité concrète correspondante, phonique ou graphique" — la difficulté qui vient du fait que "nous ne disposons pas des ressources terminologiques qui pourraient nous permettre de traiter avec quelque rigueur des faits sémantiques *indépendamment* de leurs supports formels"². C'est-à-dire: comment procéder à l'analyse sémantique sans le secours du langage ordinaire (situation dangereuse parce qu'ainsi le langage est à la fois l'objet de l'investigation sémantique, et l'instrument de cette investigation, ce qui fait qu'on doit accepter de se servir de lui comme instrument avant de savoir scientifiquement ce qu'il est comme objet, donc ce qu'il vaut comme instrument).

3. Les travaux poursuivis, et publiés, jusqu'ici sur les machines à traduire, concernant le traitement mécanographique du vocabulaire apportent moins de choses qu'on n'en aurait attendues. Au vrai, constituer des dictionnaires automatiques électroniques n'implique pas d'analyses sémantiques en profondeur, mais uniquement le classement formel, puis l'enregistrement de toutes les formes ambiguës (mots à déclinaisons; mots à sens multiples; "groupes de mots" du type *pain, pain de sucre, pain d'épice, pain à cacheter*; formes idiomatiques). Ce problème est beaucoup plus technique que linguistique. Les dictionnaires existants ont déjà fait le gros du travail d'inventaire et de regroupement des acceptions. Le vrai problème est celui de la mémoire ou des mémoires électroniques, — à mesure que la technologie fournit des mémoires plus vastes, les problèmes sémantiques devenant des problèmes d'enregistrement pur et simple. Et si la "mémoire" de Gilbert King, qui peut inscrire sur un disque de verre trente millions de signes binaires, tient ses promesses, tous les problèmes lexicaux se trouveront *ipso facto* résolus: "Plus n'est besoin, commente Emile Delavenay, de séparer les bases invariantes des mots d'avec leurs flexions ou affixes,

chaque forme de chaque mot peut être inscrite dans la mémoire . . . Un idiotisme qui se conjugue peut figurer dans cette mémoire sous chacune de ses formes, avec sa traduction pour chacune.³³ En effet, trente millions de signes binaires peuvent coder probablement quelque deux ou trois millions de mots ou formes linguistiques ordinaires, c'est-à-dire, vraisemblablement toute l'étendue d'une langue riche.

4. Les travaux mécanographiques de Jean-Claude Gardin sont foncièrement différents. Cette différence peut être exprimée de la manière suivante: les recherches sur les machines à traduire tendent à substituer, pour les manipuler, des symboles alpha-numériques *arbitraires* aux signes du langage ordinaire. On crée ainsi des signes substitutifs, des numéros de code *arbitraires par rapport au sens des mots* qu'ils codent. Le fait que le mot russe *periodicheskogo* soit codé dans le dictionnaire automatique électronique d'Oettinger au moyen du numéro 00A-0449 n'implique entre le chiffre et le mot nul rapport intrinsèque. Le chiffre indique uniquement la place où trouver le mot (puis sa traduction) dans la mémoire. Cette place n'est pas liée au sens du mot, mais à l'ordre adopté par le dictionnaire (ordre alphabétique des mots, ou ordre numérique des chiffres, etc. . . .). Les symboles avec lesquels opère J. C. Gardin sont au contraire des symboles qui décrivent le contenu sémantique des signes qu'ils codent.

Le point de départ de Gardin, c'était en effet le besoin d'organiser le classement de documentations données, sur la base de leur contenu sémantique⁴: comment construire les règles d'un inventaire d'objets archéologiques, de manière à couvrir tous les caractères de tous les objets considérés, que ce soient des outils de métal à l'âge du bronze depuis les Balkans jusqu'à l'Indus, ou des formes de vases en poterie, ou des éléments décoratifs géométriques sur ces vases, ou des ornements variés sur des monnaies grecques, ou des motifs de sceaux orientaux.

5. Le premier pas de Gardin, qui concerne directement le linguiste, c'est de ne pas utiliser, pour ce classement, le vocabulaire du langage ordinaire, c'est-à-dire le véhicule ordinaire des contenus sémantiques: "On a délibérément renoncé, dit-il, à nommer les outils, ou certains de leurs aspects fragmentaires, à l'aide des termes du langage courant. En effet, les frontières sémantiques entre ces termes sont généralement imprécises; tel est, par exemple, le cas des mots *serpes*, *faucilles*, *couteaux courbes*, qui désignent des outils souvent mal différenciés, d'un groupe à l'autre"⁵. Les codes constitués doivent toujours "fournir une manière d'exprimer, par le moyen d'un ensemble relativement limité de traits élémentaires non-ambigus, un très grand nombre de caractères, intriqués les uns dans les autres, les objets à décrire et classer, qui ont des noms trop vagues ou qui n'ont pas de noms du tout dans l'usage ordinaire"⁶.

6. Comment va-t-il donc procéder pour obtenir une analyse sémantique des objets, *qui soit indépendante de leurs noms dans les langues ordinaires*? Il constitue pour chaque sorte d'objets le code des symboles qui

noteront la présence ou l'absence de tous les *traits distinctifs* du type d'objet à décrire et classer. Le codage est donc précédé d'une analyse sémantique, destinée à établir le cadre de toutes les descriptions d'objets de ce type, c'est-à-dire le cadre exhaustif où tiendra la définition de chaque objet. Par exemple, pour constituer le fichier mécanographique de l'outillage en métal à l'âge du bronze, des Balkans à l'Indus, — qui contient plus de 4.000 fiches à l'Institut Français d'Archéologie de Beyrouth, Gardin s'est constitué le cadre suivant:

- A. Forme de la partie fonctionnelle de l'outil (15 types)
- B. Mode d'emmanchement (15 types)
- C. Dimensions (6 cotes)
- D. Section des faces de l'outil (20 types)
- E. Section des côtés de l'outil (15 types)
- F. Contour de la partie fonctionnelle, côté supérieur (35 types)
- G. Contour, côté inférieur (35 types)
- H. Liaison du corps de l'outil et du tranchant (10 types)
- I. Forme du tranchant (10 types)
- J. Section du trou d'emmanchement (20 types)
- K. Section longitudinale du talon de l'outil (20 types)
- L. Coupe longitudinale du talon de l'outil (20 types)
- M.&N. Talons, soies et appendices divers (20 types)
- O. Profil du bord supérieur de la douille (10 types)
- P. Profil du bord inférieur de la douille (10 types)
- Q. Flancs de la douille (10 types)
- R. Nervures et arêtes (50 types)
- S, T & U. Particularités diverses (210 types)
- V. Décoration.

En quoi ce travail intéresse-t-il le linguiste? D'abord il aboutit à des définitions mécanographiques (en code) de chaque objet, qui sont une nouvelle nomination (mécanographique) des objets, — véritable *nom* grâce auquel "les ambiguïtés de la terminologie courante se trouvent pour la plupart résolues par la précision des caractères entre lesquels le choix doit s'opérer . . . D'autre part, le repérage même des traits distinctifs de l'outil est facilité par l'existence du cadre analytique exhaustif que constitue le code"⁷. (Il faut remarquer que ces traits distinctifs indiqués ici, au nombre de 22, sous forme de termes du langage ordinaire, ont en fait une définition, — métrique ou graphique, — indépendante du langage ordinaire). Voici par exemple le *nom* mécanographique de l'outil originaire d'Agha Evlar (talyche persan) de niveau Kourgan No. 2, de date supposée 1450-1350 (Schaeffer), et de dimension 16, 6cm: A 13 — B 34 — C 36 — D 12 — E 23 — F 123 — G 1258 — H 24 — I 138 — J 13 — K 12 — L 12 — O 12 — P 126 — Q 247 — T 167 — V 257.

Cette nouvelle nomination mécanographique peut être considérée, d'abord, comme une définition référen-

tielle (au sens des logiciens), car elle définit l'outil d'abord en le montrant, c'est-à-dire en permettant comme un numéro d'inventaire ou de classement, d'aller en chercher la reproduction (dans une fiche classée, dans un catalogue) ou l'échantillon (dans une vitrine). Mais elle est aussi autre chose, une définition opérationnelle: "ainsi, on peut extraire du fichier, *en une seule opération mécanique*, l'une ou l'autre des collections suivantes: toutes les haches à moignons, les haches à moignons à côtés concaves, les haches à moignons à côtés concaves et talon droit, les haches à moignons arrondis, côtés concaves et talon droit, etc."⁸.

Ces noms mécanographiques ont une valeur *sémantique* opérationnelle parce qu'ils contiennent leur propre analyse sémantique, leur propre définition: ce sont à la fois, sous la même forme graphique, des signes dénotant certains objets, et les définitions de ces signes. Par rapport aux signes linguistiques que sont les mots des langues ordinaires, ces noms mécanographiques offrent cette différence de n'être pas des signes *arbitraires*. Cette différence entraîne à son tour une autre: tandis que les signes linguistiques ne sont pas susceptibles d'enregistrer toutes les gradations distinctives des éléments sémantiques qui constituent la définition des objets qu'ils désignent (à part ces acceptions qui restent trop vagues: pot, pichet, cruche, jarre, amphore, etc. . . .), les *mots mécanographiques* de Gardin possèdent cette valeur classificatoire intrinsèque. Si l'on prend le code des *traits distinctifs* au moyen duquel il analyse les formes de vases⁹, on trouve (définis graphiquement ou métriquement) les *traits pertinents* ("descriptive features") suivants: base, corps, col, anse, bec. Le "corps" est lui-même analysé en deux demi-profilés définis par six termes géométriques: droit, concave, convexe, divergent, convergent, parallèle. La liaison entre les deux demi-profilés est analysée par trois termes: courbe, angle, ressaut. Le résultat de ces analyses des traits sémantiques susceptibles de définir et de classer les formes de vases est celui-ci: "Le nombre de formes qui peuvent être différenciées de cette manière s'élève à 12.150, par emploi de huit termes dans chaque cas, choisis dans un total de 27. En fait, parmi ces 27 termes, 11 reviennent deux fois ... de sorte que le nombre total de traits descriptifs s'élève réellement à 16 seulement"⁹.

Gardin, procédant toujours de la même manière, a constitué un code susceptible de décrire et de nommer d'une manière classificatoire tous les ornements et combinaisons d'ornements géométriques rencontrés sur des vases. Au moyen de 20 signes élémentaires seulement, (dont chacun symbolise un élément d'ornement, point, droite, courbe, spirale, etc. . . .) et de 30 signes combinatoires (indiquant l'arrangement géométrique des éléments), le code peut nommer 600 ornements primaires, puis 18.000 ornements secondaires, puis 500.000 ornements tertiaires: il pourrait nommer 15 millions d'ornements du quatrième degré—de telle sorte que cette nomination définisse dans chaque cas "la spécificité d'un objet quelconque"¹⁰.

8. Gardin lui-même a bien senti l'analogie des unités

qu'il appelle "traits distinctifs" (*distinctive features*) ou "traits descriptifs" (*descriptive features*)¹¹, — unités qui découpent le "nom mécanographique" — avec les unités minima de la linguistique structurale, les phonèmes. Il voit bien que les codes qu'il a construits découpent dans les noms des objets archéologiques—serpe, faucille, ou bien jarre, amphore, etc. . . des unités sémantiques plus petites que les noms eux-mêmes: "Les codes [. . .] substituent à l'anarchie des *apparences macroscopiques* [les noms arbitraires des objets dans les langues naturelles] un système fait à partir d'un petit nombre d'*unités micrographiques* [les traits distinctifs]"¹². Mais il marque aussitôt les limites de cette analogie entre ses traits distinctifs et des phonèmes.

En fait, l'intuition de Gardin semble juste si on la reporte au domaine qui lui convient: non pas la phonologie, la linguistique structurale formelle, mais au contraire la *sémantique structurale*. Les signes de Gardin ne sont pas arbitraires par rapport au sens des mots qu'ils codent, bien qu'on puisse croire à première vue le contraire. Qu'il s'agisse des signes alpha-numériques du dictionnaire automatique d'Oettinger (00A-0449, etc. . . .), ou des éléments constituant un "nom mécanographique" chez Gardin, (A 13-B 14-C 36 etc. . . .), les représentations semblent d'abord arbitraires comme dans n'importe quel signe: il n'y a aucun rapport intrinsèque entre la représentation formelle A 13 et son contenu sémantique: telle forme de la partie fonctionnelle de l'outil. Mais la ressemblance des deux sortes de signes alpha-numériques s'arrête là.

Les signes alpha-numériques chez Oettinger (qu'on prend ici comme exemple, mais dont il ne s'agit pas de faire la critique en les opposant à ceux de Gardin, puisqu'ils n'ont pas le même objet), les signes chez Oettinger, donc, sont en fait des nombres ordinaux, dont l'analyse formelle ne donnerait aucun renseignement de nature sémantique concernant le mot qu'ils codent: seulement sa place dans le code. Les "noms mécanographiques" de Gardin sont bien formés par des unités minima dont le choix est arbitraire (A, B, C, D, etc. . . .), mais ces unités minima formelles dénotent de façon bi-univoque des unités minima sémantiques: les signes complexes de Gardin ne sont donc pas des signes arbitraires, leur analyse formelle nous renseigne sur le contenu sémantique de chaque "unité plus petite" qui les compose; ainsi que, par addition (Gardin appelle ses "noms mécanographiques" des *summarizing names*)¹³, — par addition, donc, sur le contenu sémantique du signe entier.

Pour illustrer encore mieux l'originalité des codes construits par Gardin, on pourrait dire que les signes qu'il construit ressemblent aux mots d'un hypothétique langage naturel qui a fait souvent parler de lui, depuis le Platon du *Cratyle*, en passant par le Président de Brosses et Court de Gébelin, jusqu'à Jespersen¹⁴, Whorf¹⁵ et Zellig S. Harris:¹⁶ un langage naturel où les phonèmes ne seraient pas arbitraires, où chaque phonème aurait un sens (comme celui qu'on croit sous-jacent à certaines séries anglaises: *th*. contiendrait une signifi-

cation démonstrative, *wh.* une valeur interrogative, *sl.* une valeur de glissement, *gl.* une idée de lumière, etc. . .).

De telles langues naturelles, même s'il en a jamais existé, n'existent plus. L'expérience de Gardin montre qu'on peut en construire d'artificielles où le système des formes calque rigoureusement le système des significations: chaque élément de signe (A, B, C, D, etc. . .) constitue à la fois une unité formelle minimum *et* une unité sémantique minimum. Contrairement à ce qui se passe dans les langues naturelles, les unités minima ont une face signifiante et une face signifiée, le système des signifiants reflète donc un système isomorphe des signifiés. C'est là que réside pour la recherche sémantique l'intérêt des travaux de Gardin: s'il aboutit aux résultats qu'on vient de formuler, c'est parce qu'en réalité son système des signifiants (le code) est construit a posteriori sur un système de signifiés dont il calque étroitement l'organisation: or cette organisation du système des signifiés, est une Systématique, fondée sur des critères sémantiques: le classement scientifique des formes d'outils en bronze, ou de récipients. Cherchant une parenté "phonémique" à ses unités minima (traits distinctifs, ou traits descriptifs), Gardin semble

avoir trouvé, plutôt que des unités minima distinctives non-signifiantes (comme sont les phonèmes), des espèces d'unités minima signifiantes, —des espèces d'unités minima de contenu sémantique. Ou, tout au moins, une des méthodes pour mettre en évidence de telle unités.

9. Outre qu'ils font invinciblement penser aux spéculations de Descartes et de Leibniz, de Wilkins et de Dalgarno, sur la classification rationnelle en sémantique (qu'ils appelaient une Combinatoire, une Caractéristique), les travaux de Gardin sont intéressants (du point de vue du linguiste) parce qu'ils proposent une analyse sémantique originale, qui répond aux desiderata formulés dans ce domaine par André Martinet: manipuler la réalité sémantique sans le secours des signes linguistiques. Sans doute est-il trop tôt pour affirmer que cette recherche pourrait fournir une solution des problèmes théoriques posés par l'analyse de la *substance du contenu* en unités minima de contenu, — qu'on les appelle *figures de contenu* comme Hjelmslev, ou *traits pertinents de contenu*, comme Prieto. Mais elle jette sur ces problèmes une lumière dont il est sûrement profitable de tenir compte.

Received February 2, 1960

References

¹ "La mécanographie consiste, on le sait, à exprimer les éléments caractéristiques de la matière étudiée, quelle qu'elle soit, par des symboles, —lettres ou chiffres, — transcrits à l'aide de diverses combinaisons de *positions* perforées, sur des cartes d'un modèle particulier". J. C. Gardin, v. *sub.* No. 5.

² Martinet, A. — "Arbitraire linguistique et double articulation" dans: *Cahiers F. de Saussure*, N. 15 (1957), p. 107.

³ Delavenay, E. — "La traduction automatique des langues, état présent de la recherche", Bad Godesberg, IIIe Congrès International des Traducteurs, 29 Juillet 1959. (V. *Acta*).

⁴ J. C. Gardin. — "Problèmes de la documentation", dans: *Diogène*, N. 11 (1955) pp. 107, 124.

⁵ J. C. Gardin. — *Le fichier mécanographique de l'outillage*, Beyrouth: I. F. A., 1956, p. 3.

⁶ J. C. Gardin. — "On the coding of geometrical shapes and other representations, with reference to archeological documents", dans: *Preprints of papers for the international conference on scientific information*, 1958-AREA 5, pp. 75-87.

⁷ J. C. Gardin. — *Le fichier mécanographique*, p. 13.

⁸ Id., *ibid.*, p. 15.

⁹ J. C. Gardin. — "On the Coding . . .", p. 81-84.

¹⁰ Id., *ibid.*, pp. 77, 79.

¹¹ J. C. Gardin. — "Four codes for the description of artifacts: An essay in Archeological technique and theory", dans: *American Anthropologist*, 60, 2 (April 1958) pp. 335, 357. Notamment pp. 336, 351, 353.

¹² Id., *ibid.*, p. 351.

¹³ Id., *ibid.*, p. 353.

¹⁴ Jespersen, O. — *Language*, London, 1922, Ch. XX, pp. 396 et ss.

¹⁵ Whorf, B. L. — *Language, thought and reality*, New York, 1956, notamment pp. 11, 13, 25, et 27.

¹⁶ Harris, Z. S. — *Methods in structural linguistics*, Chicago, 1951, pp. 187-188 (note 66) et pp. 192-193.