



HAL
open science

Compte rendu d'ouvrage : Language, Cohesion and Form

Jacqueline Léon

► **To cite this version:**

Jacqueline Léon. Compte rendu d'ouvrage : Language, Cohesion and Form. Histoire Epistémologie Langage, 2007. halshs-00295133

HAL Id: halshs-00295133

<https://shs.hal.science/halshs-00295133v1>

Submitted on 11 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

retrouver immédiatement les très nombreux exemples cités par l'auteur à l'appui de ses démonstrations et, par ce biais, de vérifier constamment son argumentaire. Ceci est important car la typologie comparative se doit de reposer sur une analyse précise, fine et justifiée de chaque langue. L'auteur a su rassembler des données particulièrement étendues, appartenant à des langues très diverses et géographiquement dispersées sur le cinq continents. Qu'il nous soit néanmoins permis de signaler une imprécision graphique à propos des deux exemples 9a. et 9b. (p. 580-581) de l'italien : dans les deux cas, on trouve « Demandò » qui est traduit aussi bien dans 9a que dans 9b par « il demanda » tant au niveau de l'analyse juxtalinéaire qu'à celui de la traduction française. La traduction en français est exacte, mais le texte en italien est « domandò » passé simple du verbe « domandare » (fr. 'demander') et non pas « demandò » qui est le passé simple du verbe « demandare » (fr. 'déférer (en justice)').

En conclusion, nous sommes en présence d'un ouvrage à la fois scientifique et pédagogique. Scientifique, en raison de la documentation considérable que l'auteur a su rassembler et systématiser, mais aussi de ses analyses, ses démonstrations et ses réflexions théoriques. Pédagogique, en raison de ses choix thématiques et des multiples mises en garde d'ordre méthodologique. Ces qualités confèrent à cet ouvrage une valeur d'instrument de travail particulière dans le domaine de la typologie actuelle.

Emilio Bonvini
Laboratoire LLACAN



Masterman Margaret,
Language, Cohesion and Form,
édition, introduction et commentaires de
Yorick Wilks, Cambridge, Cambridge
University Press, 2005, coll.: Studies in
natural language processing, ISBN 978-0-
521-45489-6.

Cet ouvrage est l'édition de onze textes,

dont six inédits, de Margaret Masterman (1910-1986), philosophe et logicienne, qui a créé et dirigé, de 1955 jusqu'à la fin des années 1970, le *Cambridge Language Research Unit* (CLRU), groupe pionnier britannique de traduction automatique. L'éditeur, Yorick Wilks (né en 1939), spécialiste d'intelligence artificielle et de sémantique lexicale computationnelle, fut un des membres de ce groupe et le légataire testamentaire de l'œuvre de Margaret Masterman. Celle-ci n'a jamais publié d'ouvrage mais a laissé un nombre considérable de documents inédits sur des aspects pionniers, voire parfois en avance de plusieurs décennies, dans les domaines de l'intelligence artificielle, la traduction automatique, le traitement automatique des langues et la documentation automatique. La publication de ces textes présente un grand intérêt pour l'histoire des sciences du langage ainsi que, du fait de la grande actualité des questions abordées, pour les acteurs des domaines concernés.

Dans son introduction et en commentaire des inédits, Yorick Wilks discute la pertinence des options théoriques de Margaret Masterman et leur devenir dans les courants de l'intelligence artificielle et du traitement automatique du langage à base de sémantique lexicale qui en sont issus¹. Les cinq parties regroupant les onze articles écrits entre 1954 et 1980 sont autant de jalons de l'œuvre de Margaret Masterman : i) la première partie « les formes de bases du langage » comprend les textes « Words » et « Classification, concept-formation and language » préalablement publiés dans des revues philosophiques, ainsi qu'un texte sur les techniques mathématiques (fans et treillis) préconisées pour l'étude scientifique du langage en opposition aux formalismes logiques de Carnap, Bar-Hillel ou

¹ Il faut noter qu'un des inédits (chap. 6) fait l'objet d'un commentaire d'un autre membre du CLRU, Karen Spärck Jones (1935-2007), spécialiste récemment disparue de documentation et classification automatique.

Chomsky²; ii) la seconde partie est consacrée au thesaurus comme interlangue sémantique et méthode pour la traduction automatique ; iii) la troisième partie traite des expériences de traduction automatique ; iv) la quatrième partie est consacrée aux unités (syntagmes et « breath groups ») de traitement automatique des textes ; v) enfin la cinquième partie comprend un article critique sur Kuhn, paru dans l'hommage à Richard Braithwaite, philosophe des sciences et époux de M. Masterman.

Ce regroupement d'articles atteste l'originalité du CLRU : le fait qu'il ait été un des rares groupes à avoir privilégié une approche sémantique (non logique) de l'automatisation de la traduction à une époque où la priorité de la syntaxe faisait la quasi-unanimité, ainsi que la centralité des options philosophiques dans leur approche du traitement automatique du langage. Enfin, il faut souligner le caractère exceptionnel de ce groupe très peu ancré dans l'institution, mais qui a réussi à survivre près de quinze ans en menant des projets ambitieux difficiles et coûteux, grâce à la tenacité et à l'imagination de sa directrice, de même qu'à la diversité et la complémentarité de ses membres³.

L'introduction de Wilks permet d'éclairer un certain nombre de contradictions dans les positions défendues par Masterman. Ainsi, pour définir les unités sémantiques de la méthode de traduction automatique par thesaurus, Masterman adopte les primitives sémantiques *a priori* (les

« naked ideas ») de l'interlangue « Nude » à prétention universelle élaborée par le botaniste Richens, membre du CLRU⁴. Ces primitives entrent en contradiction avec la conception héritée de Wittgenstein - dont Masterman fut l'élève - du sens des mots infiniment extensible et déterminé par l'usage. Selon Wilks, Masterman résout cette contradiction en attribuant aux primitives sémantiques un rôle fonctionnel pour la traduction automatique - désambiguïsation et unités de l'interlangue - tout en exigeant qu'elles trouvent leur justification empirique à partir de leur distribution dans les textes.

Wilks se montre plus critique quant au choix des « breath-groups » comme unités pour le traitement automatique des textes. Une des convictions de Masterman sur le traitement automatique du langage, dit-il, est qu'il doit refléter la cohérence du langage, à savoir sa redondance, notion qu'elle emprunte à la théorie de l'information. Elle donnait ainsi beaucoup d'importance à la redondance des rythmes, accents et groupes de souffle dans la poésie et la rhétorique, ce qui l'avait conduite à tenir ces éléments prosodiques comme délimitant des unités dans le texte pour le traitement automatique. Ces « breath-groups » auraient été empruntés à un thérapeute yougoslave du nom de Guberina, spécialiste des sourds, qui prétendait que même les sourds profonds étaient sensibles aux rythmes marqués et répétés dans le langage (cf. chapitre 9 « Commentary on the Guberina hypothesis »)⁵ Selon Wilks, ni les « patterns »

² Certains inédits font partie des huit essais dédiés à Yehoshua Bar-Hillel (CLRU 1959) en réponse aux critiques acerbes que celui-ci avait adressées au groupe britannique dans son fameux rapport dont la première version date de février 1959.

³ Ont fait partie du CLRU des mathématiciens, des physiciens, des botanistes, des philosophes et des linguistes, dont M.A.K. Halliday, Martin Kay et Yorick Wilks. La consultation des documents non publiés du groupe (indexés ML) permet de se faire une idée du fonctionnement éminemment collectif du groupe (Archives et documentation sur l'histoire de la traduction automatique et du traitement automatique du langage (1954-1975))

⁴ Richard H. Richens (1919-1984) est un véritable pionnier de la traduction automatique. Dès 1948, il élabore une méthode avec Donald Booth de Birbeck College, qu'il présente lors du premier colloque sur la traduction automatique organisé par Bar-Hillel en 1952 au MIT.

⁵ Petar Guberina (1913-2005) élabore avec Paul Rivenc une méthode audio-visuelle structurale globale à Saint-Cloud et rencontre probablement MAK Halliday au centre pour l'apprentissage des langues créé par Bernard Quemada à Besançon en 1954 (Chevalier 2006).

d'accents, ni les « breath groups » n'ont pu être validés empiriquement à partir de critères de surface et ne peuvent être tenus pour des unités acceptables du langage.

Sur ce point, comme sur d'autres, on pourrait envisager une autre filiation théorique en plus de l'héritage wittgensteinien privilégié par Wilks. Le terme de « breath-group » a été forgé par Henry Sweet comme unité possible de langue. De plus, l'importance accordée aux textes, à la prosodie, dans la délimitation des unités, à leur définition comme « chunks » ou syntagmes plutôt que comme mots, indiquent une filiation à la tradition linguistique britannique fortement ancrée dans la phonétique, à laquelle appartiennent la London School et son chef de file John Rupert Firth (1890-1960)

Même le rôle central accordé par Masterman au contexte et à l'usage dans la détermination du sens pourrait être attribué autant à l'influence firthienne qu'à celle de Wittgenstein⁶. Il faut préciser que, bien qu'ayant assisté à la première réunion du CLRU en août 1955, Firth s'est très peu intéressé à la traduction automatique. De plus, dans son article de 1956 « Linguistic analysis and translation » il a sévèrement critiqué les « naked ideas » de Richens et l'idée de sens *a priori* qu'elles véhiculent. Toutefois un de ses élèves, MAK Halliday (né en 1925) était un membre très actif du CLRU, et la référence que fait Margaret Masterman à la London School atteste ses affinités avec ce courant linguistique et son attachement à l'empirisme (cf. chapitre 10 « semantic algorithms » p.269).

Jacqueline Léon,
Laboratoire d'histoire des théories
linguistiques,
CNRS, Université Paris 7 Diderot

Références

Machine Translation, vol.3, n°1 et 2 1956

⁶ Firth se réfère à Wittgenstein pour sa définition du sens des mots par l'usage, et à Malinowski pour sa notion de « context of situation ».

Chevalier J-Cl., 2006, *Combats pour la linguistique*, de Martinet à Kristeva, Lyon : ENS Editions.

Firth J.R., 1968 [1956], « Linguistic analysis and translation » in Palmer F.R. (ed.), *Selected papers of J.R. Firth (1952-59)* Longmans, Green and co ltd, London and Harlow: 74-83.

Léon J., 2004, « *The inkstand was in the pen* and other stories. The controversy between Bar-Hillel and the Cambridge Language Research Unit about language formalization and machine translation » *Bulletin of the Henry Sweet Society* n°42 :4-10.

Léon J., 2007, « From universal languages to intermediary languages in Machine Translation : the work of the Cambridge Language Research Unit (1955-1970) » *History of Linguistics 2002* (Eduardo Guimaraes & Diana Luz Pessoa de Barros eds), Amsterdam & Philadelphia : John Benjamins Publishing Company :123-132.

Sparck Jones Karen, 2000, « Richard H. Richens. Translation in the Nude » in Hutchins W.J.(ed.) *Early years in Machine Translation* Amsterdam, Philadelphia : Benjamins : 263-278

Wilks Y., 2000, « Margaret Masterman » in Hutchins W.J.(ed.) *Early years in Machine Translation* Amsterdam, Philadelphia : Benjamins :279-297

CLRU, 1959, *Essays on and in Machine Translation by the Cambridge Language Research Unit* Margaret Masterman, A.F. Parker-Rhodes, Karen Sparck Jones, M.Kay, E.B. May, R.M. Needham, E.W. Bastin, C. Wordley, F.H. Ellis, R. McKinnon Wood. [document non publié, Archives et documentation sur l'histoire de la traduction automatique et du traitement automatique du langage (1954-1975)]