

Utilisation électronique de la Banque de mots

Jean-Paul Vinay

Volume 16, numéro 1-2, mars 1971

Actes du colloque international de linguistique et de traduction.
Montréal, 30 septembre - 3 octobre 1970

URI : <https://id.erudit.org/iderudit/004568ar>

DOI : <https://doi.org/10.7202/004568ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)

1492-1421 (numérique)

[Découvrir la revue](#)

Citer cet article

Vinay, J.-P. (1971). Utilisation électronique de la Banque de mots. *Meta*, 16(1-2), 95–104. <https://doi.org/10.7202/004568ar>

Utilisation électronique de la Banque de mots*

Puisque le présent colloque coïncide heureusement avec la création officielle d'une *Banque de terminologie* à Montréal, je voudrais passer brièvement en revue les principaux aspects de sa consultation future par les traducteurs canadiens. Et puisque cette consultation doit en principe s'effectuer électroniquement, comme je l'avais prévu dès 1965 au Colloque de Stanley House, je me tournerai essentiellement vers l'ordinateur pour poser les bases d'un programme d'exploitation¹.

Par une coïncidence qui n'en est sans doute pas une, je relève dans *le Devoir* du 18 septembre 1970 un article reproduisant la conférence prononcée par M. Jean-Pierre Brassard à l'occasion du premier salon canadien de l'ordinateur. L'auteur évoque les multiples possibilités offertes par une centralisation de l'information dans une banque d'information (*Data Bank*), exploitée en commun à partir d'un ordinateur central utilisant un ensemble défini de périgrammes².

* Le présent article reflète une étape de la recherche effectuée en vue du Colloque de Montréal (octobre 1970). Depuis cette date, les travaux se sont poursuivis à l'Université de Victoria grâce à une subvention de la Faculty of Arts & Science de cette institution. L'auteur tient à exprimer sa gratitude aux membres du Cambridge Language Research Unit, et en particulier à Mme M. Masterman-Braithwaite, au Dr E. W. Bastin et au professeur D. Shillan, qui ont été consultés à plusieurs reprises sur l'ensemble du projet (Colloques de Cambridge (mai 1970); du Havre (juillet 1970); de Victoria (septembre 1970)). Mme Masterman a bien voulu relire le texte de la communication présentée à Montréal et je lui suis très reconnaissant de ses suggestions et de ses critiques. Mes remerciements vont également à M. Bruce Kallio, de la société IBM Canada, qui a été détaché auprès de mon département pour collaborer au travail de mise en œuvre du programme esquissé ci-après.

Le texte de cette communication a été présenté sans vouloir anticiper sur les résultats de la recherche. Un deuxième article, en collaboration avec Mme Masterman, est en préparation; il donnera des détails précis sur le programme actuellement à l'essai à Victoria; l'ordinateur utilisé est un IBM 360/44; celui-ci doit être changé au début de 1972 pour un IBM 370/145.

1. L'expression « utilisation électronique » s'oppose ici à « traitement mécanographique »; ces deux aspects sont en somme ceux qui distinguent les travaux du Trésor de la langue française (Nancy) de ceux du Laboratoire de lexicographie de Besançon. Le terme « électronique » ne doit pas suggérer un exposé entrant dans le détail du traitement de l'information dans l'ordinateur, encore moins un article traitant du matériel (*hardware*). Il est intéressant de noter que le rapport de la conférence de Mannheim (1968), intitulé *Maschinelle Sprachverarbeitung*, emploie également l'expression *Elektronische Lexicographie* et *Elektronisches Wörterbuch* (p. 29 et 31).
2. C'est-à-dire *software*, pour ne pas dire *programmoïde*, *périlogie*, *logistique*, *intellectronique* ou même... *softouère*! La terminologie employée dans le présent article reflète l'usage d'IBM France, tel qu'il se trouve dans l'édition 1970 de *Terminologie du traitement de*

M. Brassard développe alors les principes de ce qu'il appelle une banque centrale d'information (BCI), tout entière orientée vers les besoins de l'industrie. L'auteur va très loin dans le sens de l'automatisation, puisqu'il envisage que, vers 1980, « l'entreprise pourra être totalement informatisée, tant dans ses structures et son fonctionnement que dans sa politique ». Dans le même temps, je recevais une note de la bibliothèque universitaire de Victoria, qui fait déjà largement appel à l'ordinateur pour cataloguer ses collections. Le conservateur commente la politique actuelle d'automatisation du catalogue et trace la voie aux futurs bibliothécaires : « *The [present] system... provides for a wide variety of data not readily obtainable through manual methods, and is designed to be compatible with recommended North American standards for the transmission of data... to central bibliographic banks which will hopefully come into being in the next decade.* » Ce n'est plus une coïncidence, c'est une convergence.

QUE PEUT-ON ATTENDRE D'UNE BANQUE DE TERMINOLOGIE ?

En ce qui concerne le domaine de la lexicographie automatique, certains travaux sont déjà en cours pour faire d'une banque de terminologie (BT) une réalité à la fois pratique et rentable. C'est ainsi que le Centre de recherches IBM Thomas J. Watson, à Yorktown Heights, N.Y., étudie actuellement un ensemble programmatique comprenant les points suivants : « *Text editing, formatting, dictionary lookup procedures and general outline utilities conducted upon one of [our] 360-67 systems under the timesharing operating system TSS/360* »³. Il existe en Europe des centres qui pratiquent déjà la « traduction électronique » (*computer-aided*), tels celui des forces armées de la République fédérale allemande (Mannheim) et celui de la Communauté européenne charbon-acier (Luxembourg). MM. Krollmann et Bachrach nous parlent dans le présent numéro, de leurs organisations respectives. La BT de Montréal s'insère donc dans un cadre dont les grandes lignes sont dès maintenant tracées ; son originalité pourra résider dans les rapprochements qu'elle permettra entre l'anglais et le français, dans le caractère souvent pratique et quotidien de ses données, peut-être enfin dans la périlogie qui sera la clef de son fonctionnement.

Si le principe d'une centralisation est déjà fort ancien, puisqu'il a présidé à la création des centres de documentation (CD) qui existaient dans la plupart des grandes officines de traduction, il existe cependant une différence très nette entre CD et BT. Le premier de ces deux organismes est essentiellement un lieu physique, où les fiches de terminologie s'accumulent dans les tiroirs, prêtes à être compulsées manuellement par les usagers. Une bibliothèque de référence accompagne généralement ce fichier. Une BT se distinguera à la fois par l'ampleur

Pinformation (Paris, NF2-0001-1, 214 p.). Notons au passage que le français ne distingue pas entre *memory* et *storage*, tous deux rendus par *mémoire* ; *core* est la *mémoire centrale*. *On-line* est rendu par *en direct* ; *time sharing* par *utilisation collective*. IBM a rendu *batch* par *lot* et parle donc de « traitement par lots ». Enfin, les *bytes* sont des *multipléts* ; on distinguera, le cas échéant, entre le *quartet* (4-bit byte), l'*hexet* (6-bit byte) et l'*octet* (8-bit byte). Le terme *bit*, au contraire, se rend tout simplement par *bit*.

3. Voir à ce sujet E. O. Lippmann et W. J. Plath, « Time-sharing and Computer-aided Translation », *The Finite String*, VII, 8 (1970). Ce système doit faire l'objet d'un article dans le numéro de février 1971 des *I.E.E.E. Transactions on Engineering Writing and Speech* [note de E. O. Lippmann].

de sa documentation et la rapidité avec laquelle on pourra consulter cette dernière, même à distance (télécopie, téléphone). Comme le note M. Brassard, « l'ère des ordinateurs individuels sera bientôt révolue. Un nombre croissant d'utilisateurs se brancheront désormais sur un grand ordinateur collectif, relié à une BCI. » Les dimensions de cet ordinateur seront fixées en fonction des besoins, et il est possible qu'il faille attendre de plus amples ressources financières pour assurer à la lexicographie les moyens électroniques envisagés pour répondre aux besoins de l'industrie et du commerce. Mais ordinateur il y aura, car la différence essentielle entre les deux types d'organismes évoqués plus haut est que le CD est artisanal et la BT, électronique.

M. Brassard énumère, en conclusion de son article, cinq principes que nous pouvons reprendre à notre compte, en les appliquant au domaine qui nous intéresse ⁴.

1) La BT assurera la comptabilité des éléments d'information. Elle permettra de déceler les divergences terminologiques ; facilitera la normalisation du vocabulaire et de l'usage entre les services d'une même organisation ; vérifiera que les termes rapprochés correspondent bien à une même définition, à un même code, voire à un même schéma.

Commentaires : Cette comptabilité est toujours difficile à atteindre à partir de glossaires préparés dans différents centres ; la centralisation devrait la faciliter. Quant à la normalisation, elle suppose que la BT sera dotée d'un organisme de contrôle semblable à l'O.L.F. de Québec, ou le C.I.L.F. de Paris, ou tout au moins qu'elle appliquera les recommandations de ces organismes. En ce qui concerne les rapprochements lexicographiques, l'automatisation assurera facilement la réversibilité des données lexicales, rarement possible dans le cas des dictionnaires actuels.

2) La BT éliminera les doubles emplois dans la recherche terminologique. Sur ce point, la centralisation et l'automatisation des fichiers se traduiront certainement par une économie d'efforts et de moyens financiers.

Commentaires : Il faudra sans doute prévoir plusieurs dépôts de la documentation pour parer aux accidents toujours possibles ; mais par contre les usagers pourront se contenter d'un minimum d'outils de travail, à condition d'être reliés à la BT par téléphone ou, mieux, par terminal. Il faudra étudier les avantages respectifs de certains matériels d'accès ⁵ ; adoptera-t-on les cartes perforées ou les lecteurs automatiques (*optical scanners*) permettant d'emmagasiner directement l'information ? Devra-t-on opter pour un traitement par lots ou par un système travaillant en direct selon une utilisation collective de l'ordinateur ? Nous évoquerons plus loin la question de la rentabilité d'une BT.

3) La BT facilitera les échanges entre les usagers. Non seulement une information pourra être transmise à ceux qui en ont besoin et ne la possèdent pas, mais inversement le fichier central pourra s'enrichir des trouvailles lexicographiques de centres périphériques. Le fichier pourra ainsi être tenu à jour d'une façon permanente, alors qu'un dictionnaire ordinaire est en retard sur l'évolution de la langue dès le jour de sa parution.

Commentaires : M. Brassard ajoute ici « échange facile... dans la mesure où les usagers emploient la même définition pour un même élément ». Cette condition est sans doute plus

4. D'autres points ont été soulignés dans un article de présentation de la BT paru dans *Forum*, Université de Montréal, vol. V (5 octobre 1970).

5. C'est-à-dire *hardware*, pour ne pas dire *mécanoïde*, *mécanaire*, *mécanectronique* ou même ... *hardouère* ! Voir note du CILF en date de janvier 1970, qui met cette question à l'étude.

facile à remplir dans le cas qui nous occupe que dans celui d'une BCI orientée vers le commerce et l'industrie. L'une des principales préoccupations du CLRU est précisément d'établir des règles de synonymie applicables à des domaines pratiques où les divergences lexicales sont parfois énormes. E. W. Bastin en cite de nombreux exemples : tel coude en cuivre peut prendre, selon les constructeurs, les noms suivants : *wall plate fitting, wallplate elbow, lug elbow, flanged connection, bend with back plate, bibtap with flange, female wall elbow, blackplate elbow, etc.*⁶.

4) La BT facilitera le développement des systèmes et de la programmation, en raison des descriptions précises des éléments de données.

Commentaires : Ce point avait été souligné lors du Colloque de Stanley House. Depuis, la bibliographie s'est considérablement enrichie et le choix des modes opératoires les mieux adaptés au travail du traducteur est très vaste. La revue *Finite String* notamment en donne chaque mois des détails précieux.

5) Enfin la BT s'oriente vers un système de questions et réponses, exigeant une liaison en direct avec l'ordinateur⁷. M. Brassard prévoit cette liaison à plus longue échéance, mais des premières réalisations sont d'ores et déjà acquises dans le domaine qui nous occupe, par exemple au centre de Yorktown Heights.

Commentaires : Cette question sera traitée principalement dans un deuxième article, où les composantes sémantiques des définitions seront examinées à la lumière de ce dialogue *usager/machine*, qui me paraît être l'une des possibilités les plus prometteuses de la BT.

QUELS SERONT LES USAGERS DE LA BANQUE DE TERMINOLOGIE ?

Les programmes de traitement de l'information contenue dans une BT dépendront évidemment de la clientèle et de ses besoins particuliers. Il en sera de même pour le contenu du fichier central. Examinons donc rapidement quelques types probables d'usagers.

1) Pour les *techniciens* (qui seront peut-être surtout des rédacteurs), la BT devra se prêter au fractionnement de l'information. Le fichier devra prévoir un certain nombre de rubriques, elles-mêmes subdivisées selon les applications théoriques ou pratiques qu'elles peuvent comporter. Pour définir et classer ces rubriques, il sera sans doute utile de s'inspirer des grands systèmes de classification en usage à l'heure actuelle, et éventuellement de les améliorer. Les anglophones possèdent le système de classification de Roget, susceptible d'un grand nombre d'applications pratiques ; le Centre national de la recherche scientifique

6. E. W. Bastin et B. Adamczewski, *Commodity Coding* (« The Storekeeper Problem »), Cambridge Language Research Unit, avril 1970.

7. C'est ce qui ressort nettement de l'article de Lippmann et Plath cité à la note 3 : « *The advent of time-sharing in electronic data processing has opened up promising new avenues for providing computational aids to the human translator. In a time-sharing environment, many people communicate with a computer simultaneously via terminals which look and operate much like regular typewriters. Since such terminals are linked to the computer by telephone lines, a user has considerable flexibility with respect to his choice of work location, provided that he has access both to a portable terminal and an ordinary telephone set. Compared to batch-processing a time-sharing system offers considerably greater potential for enhancing the productivity of human translators.* »

8. Le plan de classification du CNRS (*Bulletin signalétique*, vol. XXXII, 1971) a été remanié pour faciliter l'application de l'automatisation à la recherche bibliographique. Voilà encore une Banque d'information qui se développe. Quant au SYNTOL, c'est essentiellement un instrument de classification qui s'adapte à l'élaboration de résumés automatiques. Il pourrait utilement servir à la confection d'un système de questions et réponses basé sur les paradigmes sémantiques. C'est ce qu'indique J. Noël à propos de la consultation automatique du code, dans *Mélanges Fohalle*, Liège, 1970. Pour le détail, consulter J. C. Gardin

(Paris) vient de publier un nouveau *Plan de classification (1971)* qui pourrait être un bon point de départ ; on pourrait aussi s'inspirer des catégories du système SYNTOL⁸. Ce n'est pas lieu ici d'entrer dans les détails ; soulignons que l'avantage principal d'une classification logique sera de permettre aux rédacteurs de fiches de définir un *code* permettant de caractériser chaque entrée. Or, un code de ce genre est aussi une plaque tournante qui pourra aiguiller l'ordinateur vers la traduction de chaque entrée ; en d'autres termes, le code peut être la base d'une langue neutre (*interlingua*) bâtie à peu de frais — mais n'oublions pas qu'il s'agit de traductions de segments lexicaux et non de textes continus. Jack Mills, ardent défenseur de la classification décimale universelle, suggérait en 1964 une telle possibilité : « *We could go from an English term to its UDC member and from that to its subject description in other languages*⁹. » Mais les problèmes posés par la classification des livres sont évidemment plus simples que ceux qui nous préoccupent ici.

2) Pour les *non-techniciens* (c'est-à-dire la majorité des traducteurs), la BT devra fournir des renseignements encyclopédiques supplémentaires. Parmi ceux-ci, il faut placer au premier rang les *définitions* et les *exemples* (contextes), éventuellement des *images*. Le projet de l'Université de Victoria, dans sa première partie, comprend la possibilité d'inclure trois définitions, munies de leurs références bibliographiques, et des exemples accompagnés de leurs références. Ces renseignements peuvent être fournis aussi bien pour la fiche française que pour la fiche anglaise correspondante. Il est d'ailleurs intéressant de pouvoir confronter ainsi deux types de définitions ; cette comparaison est généralement très suggestive sur le plan de la traduction et de la stylistique comparée. Nous supposons bien entendu que ces définitions sont rédigées d'une façon originale et indépendante, pour chaque langue.

3) Pour les *monolingues*, la BT devra fournir des renseignements complets dans chaque langue. Il y aura, en fait, deux dictionnaires parallèles comportant des points d'accrochage permettant le passage de l'un à l'autre. Les renseignements fournis pourront différer suivant les exigences linguistiques ou stylistiques des langues en présence. Cependant, puisque la BT est surtout créée en fonction de la traduction, les fiches monolingues porteront aussi des indications dans l'autre langue.

4) Il est probable, en effet, que tous les usagers seront *bilingues* au moins dans un certain sens et qu'ils voudront avant tout utiliser l'information contenue dans le fichier central pour trouver des *traductions* que ne leur fournissent pas les dictionnaires habituels. Placée à Montréal, la BT pourra, et devra sans doute,

et F. Lévy, *l'Automatisation des recherches documentaires. Un modèle général : le SYNTOL*, Paris, 1964. G. Salton en donne une description au chapitre 6 de son livre *Automatic Information Organization and Retrieval*, New York, McGraw-Hill, 1968.

9. Jack Mills, *The Universal Decimal Classification*, Graduate School of Library Services, Rutgers, The State University, 1964. Cette analogie a été soulignée depuis assez longtemps déjà, et avec raison : cf. M. Masterman, R. M. Needham et K. Sparck Jones, « *The Analogy between Machine Translation and Library Retrieval* », *International Conference on Scientific Information*, Washington (D.C.), 1958. Un exposé très utile des méthodes de recherche en bibliographie automatique a été donné par J. W. Perry, A. Kent et M. M. Berry, dans un manuel intitulé *Machine Literature Searching*, Western Reserve University Press, N.Y., 1956.

indiquer les variantes régionales de certains termes et parfois trancher en faveur de tel ou tel usage : français canadien, français d'Europe, anglais canadien, anglais américain, anglais d'Europe. Ces préoccupations sont d'ailleurs celles qui inspirent plusieurs centres de terminologie, comme le montrent bien les fiches et feuillets publiés par le Comité de linguistique de Radio-Canada.

QUE DEMANDERA-T-ON À LA BANQUE DE TERMINOLOGIE ?

Les programmes de la BT dépendront évidemment aussi du type d'information que les usagers voudront recueillir, plus exactement de l'attitude de ces usagers à l'égard de cette information. Il semble que l'on puisse distinguer deux cas généraux :

1) *L'utilisateur connaît un maillon de la chaîne.* Autrement dit, il cherche la traduction ou la définition d'un terme *qu'il connaît dans une langue*, ou la description d'une machine *dont il connaît le nom*. Ce sera là sans doute, tout au moins au début, la position la plus fréquente des usagers de la BT. L'ordinateur doit pouvoir fournir, sur demande, et *rapidement*, pour un domaine particulier, dans un contexte donné, une information parfois considérable, susceptible d'appeler par renvois d'autres renseignements pertinents : synonymes, antonymes, parties d'un ensemble, variations d'usage, etc. Dans cette hypothèse, la BT fonctionne comme un grand dictionnaire encyclopédique qui serait fractionné en rubriques spécialisées pour la commodité des usagers.

Cette utilisation pourrait être appelée « globale » ; en effet, elle ne suppose pas de recherches heuristiques de la part du programme, sinon pour extraire de la fiche consultée toutes les possibilités d'information qui lui ont été fournies au préalable. Elle répond parfaitement aux objectifs 1, 2 et 3 cités plus haut. Elle ne diffère pas sensiblement de la consultation ordinaire d'un dictionnaire, sinon qu'elle se fait électroniquement (cela ne veut pas forcément dire « rapidement », dans le cas d'un traitement par lots) et à distance. Le gain par rapport à la recherche manuelle dans un fichier se place au niveau de sous-programmes, exploitant les liens logiques ou formels qui existent entre les fiches¹⁰. Cette exploitation permet de déborder largement du cadre des dictionnaires classiques, même de ceux qui, comme le *Dictionnaire du français contemporain* (Larousse), veulent « effectuer des regroupements autour des termes de base » et montrer « l'interdépendance des sens et des formes » (Avant-propos, p. III). C'est une utilisation essentiellement pratique, qui tire parti des travaux antérieurs de la lexicographie « statique », mais c'est aussi dans une certaine mesure une impasse qu'il faudra éviter si l'on veut automatiser la recherche lexicale à partir d'un texte ou à partir de composantes sémantiques. Josette Rey-Debove a bien exprimé ce dilemme dans son article qui sert d'introduction au numéro 19 de *Langages*, lorsqu'elle définit les démarches du lexicographe : « [On peut] envisager le dictionnaire d'un double point de vue. D'abord d'un point de vue

10. Il sera, par exemple, possible de mettre automatiquement le dictionnaire à jour ; de corriger des erreurs d'impression ou des fautes de frappe au moment de l'entrée de l'information (cf. E. J. Galli et H. Yamada, « An Automatic Dictionary and the Verification of Machine-readable Texts », *IBM Systems Journal*, VI, 3 (1967)).

traditionnel, comme une *praxis* amenant la production d'assertions plus ou moins justes, c'est-à-dire plus ou moins proches des théories linguistiques... L'autre point de vue, plus nouveau, consiste à considérer le dictionnaire comme une *production*, un texte original ayant fonction de message, et exprimant la pensée métalinguistique propre à une société ¹¹.»

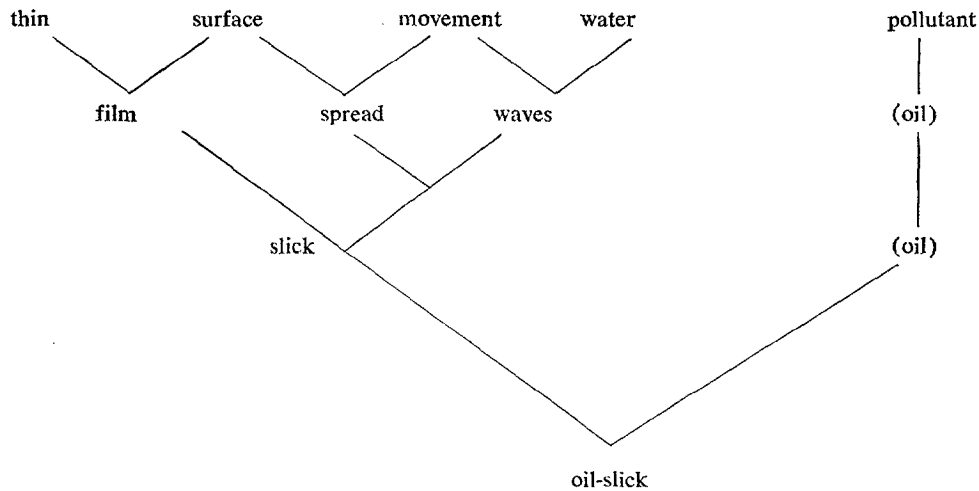
2) *L'utilisateur préfère une « approche analytique »* s'appuyant sur le rapprochement d'éléments composants pour suggérer des solutions nouvelles. Cette analyse exige des recherches systémiques beaucoup plus proches de celles qui préoccupent actuellement les « abstrauteurs » de l'équipe du SYNTOL et les « sémanticiens » comme Noël, Perry, Leech, Lyons, Masterman et Salton. C'est dans ce sens que travaillent les membres du Cambridge Language Research Group. Comme l'écrivent Jane Collier et Andrew Barbour : « *Not only has computerisation replaced manual handling, but very often it has been the case that there has been a feedback in the sense that development of computer routines has suggested methods of handling the data which previously were not necessarily the most obvious ones* ¹². »

Cette approche est envisagée dans le deuxième temps du programme de Victoria. Elle modifiera considérablement la fiche idéale utilisée actuellement pour la recherche globale. Celle-ci comporte pour le moment quinze éléments d'information, dont certains sont susceptibles de subdivision :

-
1. Entrée 2. Exposant (1, 2, etc.) 3. Qualité (fiche principale, fiche secondaire) 4. Renseignements grammaticaux sur l'entrée 5. Domaine lexicologique 6. Code de classification 7. Définition(s) 8. Références bibliographiques de ces définitions 9. Exemple(s) 10. Références bibliographiques de ces exemples 11. Notes sur l'usage du terme (variantes orthographiques, niveau de langue, etc.) 12. Traduction(s) 13. Références justifiant ces traductions 14. « Descripteurs » ou éléments de définition 15. Renvois (synonymes, antonymes, sous-ensembles sémantiques, etc.)
-

L'approche analytique a comme conséquence d'élargir considérablement les rubriques 14 et 15 du modèle ci-dessus, pour considérer en détail les composantes de nature à figurer dans les matrices sémantiques ; ces composantes devront pouvoir s'articuler selon des rapports logiques, hyponymiques ou causatifs (les *dependency trees* de Salton) permettant une analyse hiérarchisée, « *a hierarchical arrangement of words and word stems [which] can be used both for information identification and for retrieval purposes* ¹³ ». E. W. Bastin, du CLRU, utilise une démarche hiérarchisée partant de noyaux sémantiques (des *root semanticals*) pour arriver à un terme particulier, aboutissement d'une chaîne de rapports logiques. Voici un exemple de cette démarche :

-
11. Dans le même numéro, on relève un intéressant article de J. Darbelnet sur « Dictionnaires bilingues et lexicologie différentielle ».
12. J. Collier et A. Barbour, *Pilot Project for Computerising Coding on the Family Expenditure Survey*, Report to the Office of Population Censuses and Surveys, Cambridge Language Research Unit, novembre 1970.
13. G. Salton, *Automatic Information Organization and Retrieval*, chap. 2.



Que devrait-on porter sur les fiches de documentation? Il ressort des considérations ci-dessus que la BT ne pourra fonctionner utilement que si on lui fournit des données exactes, comparables et documentées. Il existe à l'heure actuelle plusieurs centres de documentation qui publient des fiches, souvent fort bien faites, mais dont la rédaction varie beaucoup d'un centre à l'autre. On pense notamment aux fiches de la C.I.L. (Montréal) et à celles du Bureau des traductions (Ottawa), généralement très simples, mais pouvant comporter des définitions; à celles de Radio-Canada (cf. ci-contre), beaucoup plus étoffées et permettant de remplir facilement les cadres de la fiche idéale évoquée plus haut; à celles publiées par le Service de linguistique des Chemins de fer nationaux du Canada, véritables feuillets encyclopédiques illustrés; à celles de l'Office de la langue française (Québec), sous forme de vocabulaires demandant à être dépouillés pour prendre une forme acceptable par la machine. Toutes ces fiches fourniront au lexicographe de service une excellente documentation qu'il lui faudra digérer, normaliser et éditer; on a vu l'importance fondamentale des *définitions*; cette importance grandira encore si la BT doit un jour devenir analytique. Il faudra alors une véritable équipe de lexicographes éprouvés (dans les deux sens du mot) pour fournir des aliments à ce moloch électronique. Au moment où la BT prend son essor à Montréal, qu'il nous soit permis d'insister sur l'urgence qu'il y a à définir avec précision les différentes rubriques de la fiche, afin d'assurer une moisson rapide et immédiatement utilisable de documents bilingues.

Le domaine choisi par l'équipe de Victoria pour faire une démonstration du premier volet de son programme de documentation automatique porte sur le vocabulaire de l'informatique. Outre l'excellent glossaire de Radio-Canada, nous avons utilisé celui publié récemment par IBM-France; le code s'inspire du système de classification préparé pour la revue *Computing Reviews* (« Communications of the ACM », 11 janvier 1968). La sortie des différents programmes de cette première partie est prévue pour le printemps 1971. Nous espérons alors que les

HAPPENING

203

© RADIO-CANADA

1968

Définition : Sorte de pièce improvisée où les acteurs brodent sur un thème donné et à laquelle le public est appelé à participer.

Ex. : L'objet du "happening", c'est la participation totale au drame qui est joué sur la scène et qui bientôt gagne la salle. J.-J. Lebel, Express, 17-5-66.

Traduction proposée : IMPROMPTU

Observation : Les tenants du "happening" trouveront sans doute bien faible la désignation "impromptu". Il y a peut-être loin de L'Impromptu de Versailles aux manifestations psychédéliques des "happenings" de New York ou de San Francisco, mais, dans l'un ou l'autre cas, l'élément capital reste l'improvisation.

Une fiche de Radio-Canada

traducteurs voudront bien les utiliser pour se rendre compte de leur efficacité et étudier le moyen de les améliorer.

Nous avons tenu à présenter dès maintenant l'état de nos recherches afin de bien poser les problèmes, même s'ils ne peuvent être tous résolus à l'heure actuelle. Je trouve une justification de cet examen préalable dans une remarque de Lippmann et Plath : « *What is still critically lacking... is the sort of experience that can only come from having professional translators use such [systems] over a period of time while developing the requisite data bases and dictionary files*¹⁴. »

JEAN-PAUL VINAY

BIBLIOGRAPHIE

La bibliographie du sujet est évidemment très vaste, et changeante. Nous n'avons pu qu'indiquer quelques titres importants au passage. En voici quelques autres, qui se rapportent directement aux préoccupations du présent article.

BÉLY, N., A. BORILLO, N. SIOT-DECAUVILLE et J. VIRBEL, *Procédures d'analyse sémantique appliquées à la documentation scientifique*, Paris, Gauthier-Villars, 1970.

[B. V.], *Kolloquium Maschinelle Sprachverarbeitung* (Mannheim, 1968), Bundesministerium der Verteidigung, Bonn, 1969 (en particulier, articles de Krollmann, Ungehauer, Storz, Schuck).

CLAS, A., « La Banque de terminologie », *META*, XIV, 4 (1969), p. 191-194.

HARRIS, B. et al., *Echantillon d'un texte traité par consultation automatique d'un dictionnaire comme aide au traducteur humain*, Université de Montréal, rapport polycopié, janvier 1970.

HERZOG, R., « Méthodes de compilation et de recherche automatique appliquées aux dictionnaires techniques » [en allemand], *Beiträge zur Linguistik und Informationsverarbeitung* (Munich), XVIII (1970), p. 26-40.

NOËL, J., *Dictionary Making for Computation*, rapport, University of Victoria, mai 1969.

14. « Time-sharing and Computer-aided Translation », *The Finite String*, VII, 8 (1970), p. 4.

- OETTINGER, A. G., *A Study for the Design of an Automatic Dictionary*, thèse de Ph.D., Harvard University, 1954.
- SCHANK, R. C., *A Conceptual Dependency Representation for a Computer-oriented Semantics*, Computer Science Department, School of Humanities and Science, Stanford University, 1969.
- STOCK, C., *Program and Dictionary Maintenance under Changing Systems*, rapport, Saskatoon, University of Saskatchewan, 1970.
- WALKER, Donald E., *Computational Linguistic Techniques in an On-Line System for Textual Analysis* (MITRE Project 1232), Information System Language Studies, n° 22, juillet 1969, the MITRE Corporation, Bedford (Mass.).
- WINKLER, U., « AUTOQUEST. Ein Verfahren der automatischen Abfragengenerierung », *Nachrichten für Dokumentation*, XX, 1 (1969), p. 16-20.