

## Le CNRS et les débuts de la traduction automatique en France

Jacqueline Léon

---



### Édition électronique

URL : <https://journals.openedition.org/histoire-cnrs/3461>

DOI : [10.4000/histoire-cnrs.3461](https://doi.org/10.4000/histoire-cnrs.3461)

ISSN : 1955-2408

### Éditeur

CNRS Éditions

### Édition imprimée

Date de publication : 5 mai 2002

ISBN : 978-2-271-05926-0

ISSN : 1298-9800

### Référence électronique

Jacqueline Léon, « Le CNRS et les débuts de la traduction automatique en France », *La revue pour l'histoire du CNRS* [En ligne], 6 | 2002, mis en ligne le 05 juillet 2007, consulté le 20 mai 2021. URL : <http://journals.openedition.org/histoire-cnrs/3461> ; DOI : <https://doi.org/10.4000/histoire-cnrs.3461>

---

Ce document a été généré automatiquement le 20 mai 2021.

Comité pour l'histoire du CNRS

---

# Le CNRS et les débuts de la traduction automatique en France

Jacqueline Léon

---

## NOTE DE L'ÉDITEUR

L'auteur tient à remercier André Lentin, Yves Gentilhomme, René Moreau, Guy Bourquin, Bernard Pottier, Antoine Culioli, Maurice Gross, Sophie Fisher et Claude Del Vigna pour les informations qu'ils lui ont communiquées ou les archives qu'ils ont bien voulu lui confier. Ils ne peuvent bien sûr être tenus pour responsables des erreurs ou omissions qui figureraient dans cet article.

- 1 C'est en 1959-1960 que se concrétise l'intérêt des Français pour la traduction automatique. Le CNRS y participe pleinement en créant en décembre 1959 le CETA (Centre d'études pour la traduction automatique) au sein de l'Institut Blaise-Pascal, avec deux sections, l'une à Paris, le CETAP, dirigé par Aimé Sestier, l'autre à Grenoble, le CETAG, dirigé par Bernard Vauquois. La création du CETA est précédée par celle de l'ATALA (l'Association pour l'étude et le développement de la traduction automatique et de la linguistique appliquée) en septembre 1959. Elle sera suivie par celle du groupe de traduction automatique de la faculté des lettres de Nancy, en mai 1960, dirigé par Bernard Pottier et Guy Bourquin.
- 2 La traduction automatique est donc un des premiers bénéficiaires de la politique gaullienne de renforcement de la recherche initiée à la fin des années 1950<sup>1</sup>. Pourtant, il faut noter que l'intérêt des Français à son égard est relativement tardif, puisque les premières recherches sur la faisabilité de traductions à l'aide d'une calculatrice électronique datent de 1948 en Grande-Bretagne et aux États-Unis, et de 1954 en Union soviétique. Aussi, dans cet article, qui retrace l'histoire de la création par le CNRS d'un laboratoire de traduction automatique, tenterons-nous d'élucider deux questions. La première consiste à savoir pourquoi la France et le CNRS ne se sont intéressés que tardivement à la traduction automatique. La seconde concerne les raisons qui ont mené

les Français à être quasiment les seuls au monde à poursuivre ces recherches après la parution du rapport de l'ALPAC (*Automatic Language Processing Advisory Committee*) en 1966<sup>2</sup>.

La traduction automatique : un enjeu scientifique, technique et économique sans précédent<sup>3</sup>

- 3 L'idée d'automatiser des traductions à l'aide de machines électroniques naît en 1946, dès l'apparition des premières calculatrices électroniques, lorsque le britannique A. D. Booth, électronicien au *Birbeck College* de Londres, sollicite auprès de W. Weaver, mathématicien, cybernéticien et vice-président de la fondation Rockefeller, l'aide financière américaine pour construire le premier ordinateur britannique. Les premières expériences engagées en 1948 par A. D. Booth et R. H. Richens, relayées en 1949 par le *Memorandum* de W. Weaver, suscitèrent des recherches dans plusieurs universités américaines. Celles-ci furent rapidement évaluées par un premier rapport de Y. Bar-Hillel en 1951, suivi d'un colloque en 1952.
- 4 En janvier 1954 eut lieu à New York la première démonstration sur ordinateur, une machine IBM 701, qui déclencha une accélération des recherches. Il s'agissait de la traduction de russe en anglais de phrases utilisant un vocabulaire de 250 mots et six règles de syntaxe mises au point par la *Georgetown University*. Bien que très limitée, cette démonstration fut montée en épingle par la presse et fit grande impression sur le public et certains scientifiques.
- 5 Les premières expériences de traduction automatique étaient placées sous la pression d'une demande sociale croissante en traductions scientifiques et techniques, ce dont témoigne un rapport de l'Unesco de 1953. Cette pression, sous forme de demande de l'État soviétique en traductions multilingues, conduisit dès 1931 l'ingénieur soviétique Smirnov-Troianski à élaborer le premier projet de machine à traduire<sup>4</sup>. Aux Etats-Unis, en 1954, la demande se trouva accrue par les possibilités qu'offraient les calculatrices électroniques. Outre des intérêts strictement militaires et politiques liés à la période de guerre froide, la traduction des articles russes, notamment dans le domaine spatial, était jugée indispensable par les scientifiques américains. La démonstration IBM-Georgetown suscita le financement de nouveaux groupes aux Etats-Unis et en Grande-Bretagne grâce à des subventions de la NSF (*National Science Foundation*) et de la CIA.
- 6 Les recherches en traduction automatique prennent alors une dimension mondiale<sup>5</sup>. En 1958, on dénombre une douzaine de groupes de traduction automatique aux Etats-Unis. Les Soviétiques s'engagent dans les recherches dès 1954, après que D. Y. Panov, directeur de l'Institut de mécanique de précision et de calcul mécanique de l'académie des sciences de Moscou, eut assisté à la démonstration de Georgetown-IBM. La recherche en traduction automatique est approuvée par le 20<sup>e</sup> congrès du PCUS en 1956 et prend d'emblée une très grande ampleur : dès 1958, le premier congrès de traduction automatique, organisé à Moscou, rassemble quelque 340 participants issus de 79 institutions différentes. De nombreux pays s'engagent ensuite dans ces recherches : le Japon en 1956, la Tchécoslovaquie en 1957, la Chine en 1958-59, l'Italie et la France en 1959, le Mexique en 1960, la Belgique en 1961. Les recherches restent peu développées en République fédérale d'Allemagne, en Suède et en Finlande.
- 7 Partout ailleurs, les investissements en personnes et en moyens financiers furent considérables. Ils ouvrirent parfois la voie à un certain nombre de recherches dont l'objectif consistait avant tout à produire des réalisations automatisées très coûteuses

pour des résultats difficilement évaluables. L'alarme est donnée dès 1958 par Y. Bar-Hillel, chargé à nouveau de l'évaluation des recherches. Son rapport de 1960 aboutit à la création de l'ALPAC en 1964. L'argumentation du rapport<sup>6</sup>, fondée sur des statistiques concernant le nombre de traducteurs par rapport aux besoins en traduction, visait à montrer l'inutilité de la traduction automatique. Seule restait nécessaire la fabrication d'outils d'aide à la traduction. Même si ce rapport fut très critiqué, souffrant notamment de la partialité de certains de ses rédacteurs, dont A. G. Oettinger d'Harvard et D. Hays de la *Rand Corporation*, engagés tous deux dans le courant de linguistique computationnelle qui, aux Etats-Unis, tira son épingle du jeu en bénéficiant du discrédit jeté sur la traduction automatique, les subventions furent arrêtées, et les groupes de recherche démantelés. La traduction automatique n'était plus une discipline ni même un domaine de recherche autonome. Partout dans le monde, les domaines de recherches se diversifièrent en intégrant la traduction automatique sous couvert d'autres thèmes : automatique documentaire, linguistique mathématique ou computationnelle, intelligence artificielle, avant de renaître une décennie plus tard dans un contexte scientifique, social et politique bien différent. Une des conséquences les plus frappantes de cette mise à l'index brutale de la traduction automatique comme objet scientifique à part entière a été d'astreindre les travaux ultérieurs à s'auto-évaluer de façon régulière, voire à s'inscrire dans une légitimité historique.

De la traduction automatique à l'automatisation de la traduction : parcours historique

Première période (1948-1960) : idées et expérimentations

Les premières expériences de traduction automatique ont été engagées en 1948 en Grande-Bretagne et relayées en 1949 par le *Memorandum* de W. Weaver qui a suscité des recherches dans plusieurs universités américaines.

En janvier 1954 eut lieu à New York la première démonstration sur ordinateur qui déclenche une accélération des recherches, qui prennent bientôt une dimension mondiale. Successivement, les Soviétiques, les Japonais, les Italiens, les Français et les Belges s'engagent dans ces recherches. Les recherches restent peu développées en RFA, en Suède et en Finlande. Les traductions concernent essentiellement le russe, pour des raisons militaires et politiques liées à la période de la guerre froide, mais aussi pour des raisons scientifiques. Partout, les investissements en personnes et en moyens financiers sont considérables. Ils ouvrent parfois la voie à un certain nombre de recherches dont l'objectif consiste avant tout à produire des réalisations automatisées opérationnelles très coûteuses pour des résultats difficilement évaluables.

Pourtant, cette première période est foisonnante d'idées : méthodes probabilistes, langues, intermédiaires sémantiques, méthodes empiriques sur corpus, analyseurs syntaxiques automatiques. Apparaissent aussi les premiers outils pour le traitement automatique des langues, comme les dictionnaires électroniques. La plupart des idées à la fois théoriques et méthodologiques qui présideront au développement du traitement automatique des langues jusqu'à nos jours apparaissent pendant cette première période.

Deuxième période (1960-1966) : l'analyse syntaxique

Dans la période 1960-66, c'est l'analyse syntaxique qui est mise en avant comme la seule voie de recherche possible pour faire avancer la traduction automatique. En conséquence, la production en séries de traductions est reléguée à des temps futurs, ce qui va conduire, à terme, à la suppression des recherches en traduction automatique au profit de la seule linguistique computationnelle. Cette exclusivité de l'analyse syntaxique tient au développement des grammaires formelles, notamment de la grammaire catégorielle de Y. Bar-Hillel (1953) et de la grammaire générative de Chomsky (1955), facilement programmables dans des parseurs, mais

aussi à la disparition des méthodes empiriques et probabilistes qui connurent plusieurs adversaires, dont tout particulièrement Y. Bar-Hillel et Chomsky. Quant aux méthodes sémantiques par langues intermédiaires, mises au point par des non-Américains, elles furent laissées à l'écart. Ainsi, lorsque le rapport de l'ALPAC met fin au financement des recherches en traduction automatique aux Etats-Unis en 1966, en promouvant à la place la linguistique computationnelle, le terrain est tout préparé.

Troisième période (1966-1980) : la survie et la « force brute » des grands systèmes

Partout dans le monde, les recherches subirent un coup d'arrêt. Les systèmes qui ont survécu ou qui sont créés, ailleurs qu'aux Etats-Unis, pendant la quinzaine d'années ayant succédé à l'ALPAC, obéissent à des critères très différents de ceux de la période précédente. Il reste quelques systèmes de recherche (tel Ariane, développé par le GETA à partir de 1971, ou Susy, élaboré à l'université de Sarrebruck) qui ont surmonté la crise parce qu'ils étaient soutenus par de grands organismes de recherche, tel le CNRS, préoccupé seulement de façon secondaire par les aspects économiques de la traduction automatique. C'est le cas également des systèmes qui répondent à une volonté politique de construire des systèmes multilingues comme Eurotra (1977-1994) pour la communauté européenne. Aucun de ces différents systèmes n'a abouti à une industrialisation.

Le deuxième groupe concerne les systèmes industrialisés qui « marchent ». Ce sont des systèmes qui répondent à un besoin interne de traductions, et qui, non destinées à être publiées, peuvent se contenter d'être de qualité médiocre. Le cas de Systran est exemplaire. A la fois le plus industrialisé et le plus diffusé, encore actuellement, il n'est fondé sur aucun modèle linguistique cohérent. Descendant d'un système mis au point dans les années 50 à la *Georgetown University*, il a été un des rares à avoir dépassé le stade de l'expérimentation et à fournir des traductions brutes, dans un domaine très spécialisé, à usage interne d'une entreprise, d'où son surnom de « force brute ».

Le tournant japonais et l'automatisation de la communication (1980-1990)

Le début des années 80 marque un tournant important. Le développement des micro-ordinateurs et des traitements de texte, leur usage démocratisé et la mondialisation de la consommation favorise l'apparition d'une nouvelle étape pour la traduction automatique, celle de la commercialisation. Il s'agit maintenant de traduire des modes d'emploi, des descriptifs de produits ou de machines, conditions indispensables pour leur commercialisation à l'échelon mondial.

Le tournant est impulsé par les Japonais en 1982, avec l'annonce du projet « 5<sup>e</sup> génération » à l'ICOT (*Institute for New Generation Computer Technology*) et soutenu par le très puissant MITI (*Japan's Ministry of International Trade and Industry*). L'enjeu ne concerne pas simplement des intérêts économiques à court terme mais il s'agit d'un pari à long terme sur une société future, fondée sur l'information. Se développent des systèmes interactifs de traduction automatique assistée par des humains, et des stations de travail de traduction assistée par ordinateur (TAO), offrant aux traducteurs des ensembles d'outils rendus utilisables par la micro-informatique et les logiciels de bureautique.

Avec les années 1980, la traduction automatique n'est plus confinée à une utilisation « maison » mais fait partie de l'ingénierie linguistique, au même titre que les dictionnaires électroniques, les bases de données terminologiques et la génération de texte, qui constitue son application privilégiée notamment dans le domaine de la production de textes multilingues.

Depuis 1990, le retour des méthodes empiriques

Les années 1990 voient le renouveau de l'intérêt pour la traduction automatique aux États-Unis. Celui-ci est d'abord d'ordre économique : nécessité de traduire la documentation des concurrents japonais, baisse des coûts de l'informatique et hausse des coûts de la traduction surtout pour les langues à alphabets non latins comme le japonais. Ce renouveau reste toutefois limité et se caractérise par la mise au point de systèmes utilisant des stratégies et des techniques mixtes. Des modules

« intelligents » et des modules probabilistes viennent compléter les modules d'analyse linguistique plus traditionnels. Mais ces systèmes sont loin de pouvoir être commercialisés et le clivage entre stations d'aide à la traduction et systèmes de traduction s'accroît.

Mais surtout, une des caractéristiques des années 1990, c'est le retour des approches empiriques et du traitement statistique de grands corpus. Ce renouveau, suscité par le succès des méthodes stochastiques dans le traitement du signal et la reconnaissance de la parole dans les années 1970, a bénéficié de la mise à disposition de corpus de données textuelles importants grâce aux nouvelles possibilités de traitement des ordinateurs et aux efforts de normalisation des textes. Plusieurs méthodes de traduction automatique sont fondées sur les mémoires de traduction utilisant des corpus bilingues alignés, tels le corpus anglais-français des actes parlementaires canadiens, les *Canadian Hansards*. L'alignement, c'est-à-dire la mise en correspondance des phrases, est effectué à partir de critères exclusivement statistiques utilisant le fait que les phrases longues dans la langue source ont tendance à être traduites par des phrases longues dans la langue cible et que les phrases courtes ont tendance à être traduites par des phrases courtes.

A l'heure actuelle, on constate une certaine déconvenue vis-à-vis de la traduction automatique. Certains déplorent le peu de progrès accomplis depuis les années 1960, en constatant que les systèmes se sont beaucoup améliorés sur le plan de la vitesse mais peu sur le plan de la qualité, et que la machine produit des traductions de niveau bien inférieur à celles produites par des traducteurs humains. Un second constat d'échec concerne l'incapacité de la linguistique à produire des systèmes performants de traduction automatique. Tout au plus les systèmes fondés sur des modèles linguistiques puissants servent-ils de bancs d'essai aux théories linguistiques computationnelles. Que penser également de ces nouveaux systèmes, fondés sur des méthodes probabilistes, qui se prévalent de ne pas utiliser de savoir linguistique ? On aboutit actuellement à un double paradoxe : on n'exige plus des systèmes de traduction automatique, fondés linguistiquement et qui ont reconquis leur légitimité dans la linguistique computationnelle, qu'ils fournissent des traductions ; quant aux stations d'aide aux traducteurs, elles n'utilisent les systèmes de traduction automatique, quand ils existent, que comme un outil parmi d'autres. Autrement dit, on en arrive à se demander si l'automatisation de la traduction a encore réellement besoin de la traduction automatique.

Le retard français L'informatique<sup>7</sup> en France à la fin des années 1950

- 8 Dans les années 1950, on note une nette prédominance du calcul analogique sur le calcul numérique, qui en était encore à ses balbutiements. La France est très en retard. En 1954, il n'existait aucune machine numérique en France, alors que la Grande-Bretagne en possédait deux et l'Allemagne une.
- 9 Ce retard tient à un ensemble de raisons. Henri Boucher<sup>8</sup> raconte que, de 1945 à 1951, l'informatique n'était aux États-Unis qu'une simple expérimentation dont la renommée n'avait pas atteint la France. Dans les cours de l'École polytechnique, de l'ENSGM, ou de Sup. Aero (écoles formant les militaires), aucune allusion n'était faite à ces recherches. Aussi A. Sestier, qui prendra la direction du CETAP, fut-il un des rares, dans l'armement, à s'intéresser au calcul numérique.
- 10 Les diverses compagnies accoutumées à travailler avec la Défense sur les problèmes d'électronique et de mécanique de précision refusèrent toutes de prendre un risque technologique en se lançant dans la construction d'une machine française. Se reconnaissant incompetentes tant en architecture qu'en composants, elles proposèrent de s'associer à des sociétés américaines par le biais de licences. Toutefois, la Défense, particulièrement soucieuse du secret de ses études, ne donna pas suite à ces

propositions. La seule société qui ait accepté de mener des études technologiques sans recourir à une compétence étrangère a été IBM France dont les prix étaient par ailleurs inférieurs à toutes les propositions des sociétés françaises. Alors que l'armée et l'industrie françaises ne manifestèrent qu'un intérêt frileux pour les calculatrices électroniques, et que l'Université fit preuve d'une très grande hostilité *a priori* à l'égard des machines et du traitement de l'information, le CNRS fut le seul, par le biais de l'Institut Blaise-Pascal, à se lancer dans un projet de construction de machine électronique. Hélas, ce projet n'aboutira pas. Il s'agissait de la fameuse « affaire Couffignal » et de la « non-construction en France du premier ordinateur » étudiée par Girolamo Ramunni<sup>9</sup>.

L'Institut Blaise-Pascal<sup>10</sup>

- 11 Afin de faire face aux besoins des laboratoires de recherche français en matière de calcul et pour développer la construction de calculateurs électroniques français, le CNRS crée en 1946 l'institut Blaise-Pascal (IBP) sous la supervision de Joseph Perès, alors directeur adjoint pour les sciences au CNRS (de 1945 à 1951). L'IBP regroupe deux laboratoires existants : le Laboratoire de calcul analogique, dirigé par Lucien Malavard, et le Laboratoire de calcul mécanique, dirigé par Louis Couffignal.
- 12 En 1947, le CNRS confie à L. Couffignal le soin de construire la première calculatrice électronique française mais, en raison d'une erreur de conception, la machine n'est pas techniquement viable et le projet est abandonné en 1952. Après ces précieuses années perdues, le CNRS décide finalement, en 1955, d'acheter une machine anglaise, une Elliott 402, pour équiper l'Institut Blaise-Pascal. La première machine française est donc anglaise.
- 13 À son arrivée comme directeur du CNRS en 1957 (il le restera jusqu'en 1962), Jean Coulomb met fin aux fonctions de L. Couffignal qui est remplacé par René de Possel (1905-1974). C'est à l'IBP que sont créés le Centre d'étude pour la traduction automatique (CETA), en décembre 1959, puis la Section d'automatique documentaire (SAD), le 20 décembre 1960, sous la direction de Jean-Claude Gardin.

Les machines et les hommes

- 14 En 1959, il n'existe que huit machines en tout et pour tout pour le CNRS et l'Université. L'année 1959 marque toutefois un tournant. On assiste alors à un bond décisif dans le développement de l'équipement des universités en calculateurs électroniques grâce à la décision, soutenue par une politique gaullienne favorisant la recherche, de développer prioritairement trois grands centres informatiques à Paris, Grenoble, et Toulouse. C'est donc là où sont les machines, Paris et Grenoble, que sont créés les centres de traduction automatique.
- 15 Selon le projet préparatoire au rapport de conjoncture 1960 du CNRS, le retard français en informatique ne tient pas seulement au manque de matériel. Il tient aussi à une pénurie générale de personnel spécialisé. Il est difficile d'intéresser et de former les ingénieurs au calcul numérique. Le désintérêt est lié au fait que le calcul électronique, considéré comme une technique subalterne, n'est pris au sérieux ni par les universitaires ni par les entreprises. Par ailleurs, il faut assurer la reconversion à la programmation des calculateurs travaillant sur des machines de bureau. Dans le cas du CNRS, précise ce rapport, la situation est particulièrement grave : il est devenu impossible de recruter des personnes qualifiées auxquelles l'industrie privée offre des situations beaucoup plus avantageuses.

### La cybernétique

- 16 On peut invoquer un autre élément pour expliquer le retard français en matière de traduction automatique. On a vu que c'est le *Mémoire* de W. Weaver, rédigé en 1949, qui a suscité les premières expériences de traduction automatique aux États-Unis et en Grande-Bretagne. Or W. Weaver est aussi vice-président de la fondation Rockefeller qui a financé tous les colloques internationaux du CNRS (38 plus précisément), de 1945 à 1955. Un de ces colloques aurait pu traiter de la traduction automatique et susciter un intérêt pour ce domaine en France. Il s'agissait d'un colloque organisé par L. Couffignal à l'IBP, en janvier 1951, et intitulé « Les machines à calculer et la pensée humaine ». Les discussions ont porté sur la cybernétique et sur les machines à calculer numérique<sup>11</sup> mais, malgré la présence de A. D. Booth, un des pionniers britanniques de la traduction automatique, cette question n'a pas été abordée et ce domaine ne semble pas avoir trouvé d'écho en France.

### La linguistique au CNRS en 1959

- 17 Les linguistes français ne manifestent pas d'intérêt spécifique pour les langages formels, qui, aux États-Unis, sont développés parallèlement aux expériences de traduction automatique, souvent au sein même de ses centres. La linguistique, représentée au CNRS par deux sections du Comité national, « linguistique et philologie classique » et « linguistique et philologie non classique », est orientée essentiellement vers la philologie, l'enseignement des langues et les prescriptions normatives<sup>12</sup>.

Émile Delavenay, l'Unesco et l'ATALALA préhistoire de la traduction automatique en France : le rôle d'É. Delavenay et de l'Unesco

- 18 A la fin des années 1950, la direction du CNRS n'est pas sans avoir une idée, au moins sommaire, sur les expériences en traduction automatique, du moins sur celles qui sont effectuées aux États-Unis. Il semble qu'elle en ait été informée en 1956, lorsque Gaston Dupouy, directeur du CNRS (de 1950 à 1957), est convié, par l'intermédiaire d'un militaire, le général Brison, à une réunion internationale « sur les machines à traduction », organisée par Léon Dostert, responsable du groupe de traduction automatique de *Georgetown University*, le 20 octobre 1956 au MIT. Il en informe le linguiste Michel Lejeune (1907-2000)<sup>13</sup>, directeur adjoint pour les lettres et les sciences humaines au CNRS (de 1955 à 1963), lequel ne peut pas s'y rendre.
- 19 Un an après, en octobre 1957, l'attention de M. Lejeune est de nouveau attirée sur la traduction automatique par Henri Laugier, professeur à la faculté des sciences, qui lui transmet une lettre exposant l'intérêt de monter un groupe de ce type en France. Cette lettre est signée d'É. Delavenay, directeur du service des documents et des publications de l'Unesco.
- 20 Né en 1905, angliciste de formation et ancien normalien<sup>14</sup>, É. Delavenay, par ses fonctions jusqu'en 1950 de responsable des services de traduction et d'édition à l'Organisation des nations unies (ONU) à New York, a été amené à s'intéresser de très près aux problèmes de traduction sur le plan international. Il n'est donc pas étonnant qu'il ait été l'instigateur de la traduction automatique en France.
- 21 Finalement, en novembre 1957, M. Lejeune et É. Delavenay se rencontrent. Entre-temps, M. Lejeune a fait la connaissance d'Erwin Reifler, sinologue et responsable du groupe de traduction automatique de l'université de Washington, lors du 8<sup>e</sup> congrès de linguistique d'Oslo. À l'issue de cette entrevue, É. Delavenay convoque en février et mars 1958 deux réunions sur la « machine à traduire » à l'Institut de linguistique de la

Sorbonne en collaboration avec André Martinet. À cette réunion, présidée par Benveniste, assistent M. Lejeune, les mathématiciens Georges Guilbaud, Benoît Mandelbrot, Marcel-Paul Schützenberger, Clemens Heller de la VI<sup>e</sup> section de l'École pratique des hautes études, et Louis Ziéglé. Dans ses mémoires<sup>15</sup>, É. Delavenay évoque la faible réceptivité des linguistes, et des universitaires en général, à l'idée de fabriquer une machine à traduire en France, et ces réunions ne semblent pas avoir abouti à un projet concret.

- 22 Parallèlement, autour de E. Delavenay, se constitue un groupe de travail sur la traduction automatique qui se tient au courant des travaux américains, britanniques et russes. Ce groupe prend le nom de « groupe international d'études sur la traduction automatique » et se réunit régulièrement à l'Unesco. Il sera à l'origine de l'ATALA.
- 23 Bien que les activités de ce groupe ne soient pas directement liées aux fonctions d'É. Delavenay à l'Unesco, le rôle de cette institution n'est pas négligeable. Les problèmes de traductions scientifiques et techniques l'intéressaient de très près. L'Unesco a ainsi produit deux rapports sur la question. Le premier, datant de 1949<sup>16</sup>, est consacré aux dictionnaires terminologiques spécialisés ; le second, datant de 1953, dresse un état des lieux des besoins croissants en traductions scientifiques et techniques dans le monde, faisant état du manque de formation des traducteurs et du coût trop grand des traductions.
- 24 Par ailleurs, la création de l'ATALA, en septembre 1959, a bénéficié du succès du premier congrès de l'IFIP (*International Federation for Information Processing*) organisé par l'Unesco, à Paris, en juin 1959, où sont présentées de nombreuses communications sur la traduction automatique. Enfin, c'est grâce à É. Delavenay et Michael Corbe, tous deux de l'Unesco, qu'A. Sestier, futur directeur du CETAP, prend connaissance des recherches en traduction automatique à l'étranger.
- 25 É. Delavenay continue à jouer un rôle important dans le développement de cette discipline en France. Il est en relation avec de nombreux groupes anglais et américains : le groupe de Locke au MIT, le groupe de Reifler à l'université de Washington, et le groupe de Cambridge en Grande-Bretagne. Président fondateur de l'ATALA jusqu'en 1965, auteur d'un *Que sais-je ?* intitulé *La Machine à traduire*, et d'une bibliographie sur la traduction automatique<sup>17</sup>, il fera partie du Conseil scientifique du CETA et de la section 22 « Linguistique générale, langues modernes et littérature comparée » du Comité national.

L'ATALA<sup>18</sup> et le centre Favard

- 26 Les membres fondateurs de l'ATALA, présidée par É. Delavenay, sont des mathématiciens, des linguistes, des ingénieurs, des traducteurs, des documentalistes. L'ATALA se considère moins comme un organisme de recherches que comme un forum de discussion sur la traduction automatique, sur la documentation automatique et la linguistique appliquée ; ce dont témoigne sa revue créée en avril 1960, la *Traduction automatique*, fabriquée alternativement par les compagnies Bull et IBM, un des lieux de pénétration en France de la linguistique américaine et des langages formels.
- 27 Signalons également, en mars 1960, la création du séminaire de linguistique quantitative dirigé par Jean Favard, à l'Institut Henri-Poincaré, qui constitue un des lieux importants de confrontation entre linguistique et mathématiques. Un groupe fondé par le ministère de la Défense après l'aventure de Suez pour étudier le codage des messages à l'aide de modèles statistiques est à l'origine de ce séminaire. Certains de ses

membres ont travaillé pour le chiffre pendant la guerre; c'est le cas de René Moreau, alors capitaine de gendarmerie et qui deviendra le directeur du centre scientifique d'IBM. Dans ce séminaire est diffusé un enseignement de linguistique pour mathématiciens (André Martinet et Jean Dubois) et un enseignement pour non-spécialistes de mathématiques, logique mathématique et théorie de l'information (R. Moreau et Daniel Héroult). En 1962-1963, un cours supplémentaire sur la théorie des langages est assuré par J. Pitrat et Maurice Cross.

#### Création du CETA

- 28 En 1959, l'intérêt du CNRS pour les recherches en traduction automatique se confirme. Dans le rapport de conjoncture de 1959 du CNRS, elle figure parmi les sujets prioritaires que se donne la commission de mathématiques générales et appliquées, présidée par le doyen Pérès, en collaboration avec la commission de physique théorique et de théorie des probabilités. Celle-ci fait état d'une demande en traduction automatique du russe, en documentation automatique, en analyse numérique et en automatisme.
- 29 À cette fin, il faut équiper en calculateurs électroniques deux ou trois super-centres, Paris, Grenoble et Toulouse, avec des Gamma 60 ou des IBM 704, centres auxquels est confiée l'étude des problèmes prioritaires.
- 30 Par ailleurs, la Direction de l'armement, alertée par le général Grossin, alors directeur du SDEC (Service de documentation et de contre-espionnage), commence à s'intéresser à la traduction automatique et se montre disposée à y consacrer une partie de son budget de recherche.
- 31 Comme pour la plupart des commanditaires américains, les objectifs sont en partie liés à l'armée et au contre-espionnage. Et l'enjeu de la traduction automatique consiste essentiellement dans la traduction du russe d'articles scientifiques et techniques soviétiques<sup>19</sup>.
- 32 Le CETA est finalement créé le 24 décembre 1959 par une convention entre le CNRS, la DEFA (Direction des études et fabrications d'armement) et le CASDEN (Comité d'action scientifique de défense du centre d'exploitation scientifique et technique) du ministère de la Défense<sup>20</sup>). Coulomb représente le CNRS, le général Guérin, le CASDEN, et le général Sorlet, la DEFA. Il est créé, sous la supervision du doyen Pérès, au sein du Laboratoire de calcul numérique de l'Institut Blaise-Pascal du CNRS. Il a pour mission « l'étude et la conception d'une méthode pour la traduction automatique notamment du russe en français et l'étude de l'organisation générale d'une machine pour cette fin ».
- 33 La création du CETA à l'Institut Blaise-Pascal associe donc étroitement la traduction automatique au développement du calcul numérique et des calculateurs électroniques ainsi qu'à la documentation automatique. Traduction automatique et documentation automatique sont toutes deux sous le contrôle des mathématiques appliquées et dépendent de la section « Mécanique générale et mathématiques appliquées » du Comité national. Même si les équipes sont composées de spécialistes des diverses disciplines intéressées, les mathématiciens appliqués considèrent qu'ils sont les seuls à pouvoir animer la recherche, dans la mesure où ils sont les seuls capables d'assurer le suivi des questions de programmation et de structure des machines.
- 34 Le directeur du laboratoire de calcul numérique, R. de Possel, et son sous-directeur, André Lentin, par ailleurs membre fondateur de l'ATALA, joueront un rôle important

dans l'interaction entre mathématiques appliquées, langages formels et linguistique. Ils sont tous deux membres du conseil scientifique du CETA.

Le CETAP

- 35 L'armée est, à plusieurs titres, partie prenante dans l'affaire. Outre l'aide du CASDEN, qui assure pour moitié le financement de l'ensemble du CETA, le CETAP bénéficie des locaux du Laboratoire central de l'armement (LCA) au Fort de Montrouge à Arcueil. Son directeur, A. Sestier, né en 1920, ancien élève de l'École polytechnique, est ingénieur militaire en chef, chef de la section des machines à calculer du LCA. L'armée est très présente dans le comité de direction du CETA: outre des représentants du CASDEN comme le général Guérin, et du LCA comme le général Sorlet, le comité de direction comprend l'ingénieur général Ollier représentant le SDEC<sup>21</sup>.
- 36 Les ingénieurs et le personnel technico-administratif appartiennent au LCA Les deux ingénieurs à plein temps du CETAP sont Lucien Dupuis, ingénieur civil à la DEFA, qui par ailleurs connaît le russe, et M. Gross, né en 1934, ancien élève de l'École polytechnique et ingénieur de l'École nationale supérieure de l'armement. Celui-ci est attaché au service des machines à calculer du LCA.
- 37 Les linguistes, sept slavistes et deux germanistes, sont rémunérés par le CNRS. Cependant, à l'exception d'Yves Gentilhomme, les russisants engagés par le CETAP ne sont pas des chercheurs, ce sont des agrégés de russe et parfois même de simples locuteurs natifs du russe. On dispose sur ce point du témoignage d'une des linguistes russes engagés au CETAP sous la direction d'Y. Gentilhomme. Arrivée en France à 18 ans en 1957 avec un « baccalauréat » russe qui comprenait une légère formation paramilitaire, elle est affectée à la traduction de listes de termes militaires.

Le CETAG

- 38 Grenoble, on l'a dit, est l'un des principaux lieux d'implantation de l'informatique en France. C'est à Grenoble qu'est fondée l'AFCAL (Association française de calcul) en 1957, et Kuntzmann y crée la revue Chiffres en 1958.
- 39 Il n'est donc pas étonnant que la seconde section du CETA soit créée à Grenoble. Et c'est au laboratoire de calcul de la faculté des sciences, avec le soutien du professeur Kuntzmann qu'est créé le CETAG sous la direction de Bernard Vauquois.
- 40 B. Vauquois (1929-1985) est d'abord chercheur au CNRS de 1952 à 1958 en radioastronomie au Laboratoire d'astronomie à l'Institut d'astrophysique de l'observatoire de Meudon, après avoir effectué des études de mathématiques, de physique et d'astronomie. Dès 1957, son programme de recherche porte aussi sur les méthodes appliquées à la physique du point de vue des calculateurs électroniques et il enseigne la programmation aux physiciens. Ce double intérêt pour l'astrophysique et les calculateurs électroniques se reflète dans le sujet de thèse et dans celui pour la thèse complémentaire en sciences physiques qu'il soutient en 1958: (i) étude de la composante lentement variable du rayonnement radioélectrique solaire, (ii) étude de la théorie des machines (machine de Turing et von Neumann). Il est nommé maître de conférences en mathématiques appliquées à la faculté des sciences de Grenoble en octobre 1958.
- 41 Le groupe de réflexion sur la traduction automatique, qui deviendra le CETAG, est composé de mathématiciens et d'ingénieurs mais manque cruellement de linguistes, notamment de russisants. B. Vauquois, dans le rapport pour le conseil scientifique du

20 février 1960, souligne le gros atout du groupe : l'accès à la machine Gamma et à tambour magnétique.

- 42 Après une demande spécifique en personnels linguistes effectuée par B. Vauquois en mai 1960 auprès de la direction du CNRS, le CETAG comprend, à la fin de la même année, quatre spécialistes de russe, une technicienne germaniste, et une technicienne spécialiste de japonais. En ce qui concerne les informaticiens, le directeur de la programmation, l'ingénieur Jean Veyrunes, disparu prématurément en 1966, a sous ses ordres deux programmeurs. Le groupe compte également deux ingénieurs mathématiciens préparant un doctorat de mathématiques appliquées.

Le rapport Sestier

- 43 C'est A. Sestier, le directeur du CETAP qui a rédigé, en date du 23 novembre 1959, le rapport « Comment doit être organisé à l'échelle française l'effort pour la traduction automatique », rapport qui va servir de base à la définition des orientations de cette discipline en France. C'est une des raisons pour lesquelles les deux groupes, bien que recevant des crédits équivalents, ne sont pas à égalité au départ. De plus, le CETAP a pris de l'avance puisqu'en tant que groupe du LCA, il travaille sur l'allemand depuis 1958 alors que le CETAG vient tout juste de commencer les recherches en octobre 1959. Ce déséquilibre, on va le voir, sera à la source d'un certain nombre de malentendus entre les deux centres.

Un centre de traduction en séries

- 44 Selon le rapport Sestier, le CETA doit poursuivre un objectif pratique. D'ici un délai d'environ cinq ans, le CETA a pour vocation d'être à la fois une usine de production en série de versions françaises de russe scientifique et technique, un centre de recherches appliquées où des équipes constituées prépareront la traduction automatique d'autres langues, et un centre de recherches fondamentales en linguistique sans objectif immédiat de traduction automatique, mais avec l'espoir que les résultats de celle-ci, complétés par des études indépendantes, pourront être utiles à des linguistes pour élaborer ou contrôler des théories.
- 45 Pour A. Sestier, la production en série de traductions est donc au premier plan et le CETA devra offrir un certain nombre de prestations : traductions grossières et études sur l'indexation et l'extraction automatique susceptibles d'assurer à terme, d'ici cinq à dix ans, l'autofinancement des recherches<sup>22</sup>. A. Sestier propose d'ailleurs le nom de « Centre d'études et d'exécution de traductions automatiques » qui souligne sa vocation de réponse à la demande sociale.
- 46 A. Sestier conclut son rapport par une mise en garde concernant le risque de gaspiller les efforts et les crédits en ne faisant que redécouvrir ce qui a déjà été trouvé aux États-Unis et en Union soviétique. Il faut noter que le rapport de conjoncture du CNRS (1959-60) reprend à son compte l'idée que, dans un délai de cinq à dix ans, on aura dépassé le stade de la recherche pour produire des traductions en série.

Technologie linguistique et ingénieurs linguistes

- 47 La méthode préconisée par le rapport Sestier est axée sur l'analyse de la langue source, c'est-à-dire le russe. Une lettre de A. Sestier du 18 octobre 1960 à Éric de Grolier, directeur du Centre français d'échange et de documentation technique à Milan, expose la méthode qu'il préconise : la traduction automatique peut se passer de théorie parfaite ; ce qu'il lui faut, c'est une « technologie linguistique ». Comme É. Delavenay<sup>23</sup>, A. Sestier parle de machine à traduire française.

- 48 Cette technologie linguistique doit être le fondement scientifique de toute traduction automatique raisonnablement correcte, même si elle ne vise pas l'idéal bien lointain de la FAHQMT (*Full Automatic High Quality Machine Translation*) préconisée au début des années 1950 par certains chercheurs américains. Les recherches doivent s'attacher dans un premier temps, moins à construire la morphologie, la syntaxe ou la sémantique de telle ou telle langue particulière, qu'à dégager la structure que doivent avoir ces disciplines, en tant que destinées à servir de base à la traduction automatique; et cela par des amorces d'études comparatives de diverses langues.
- 49 Dans son rapport, A. Sestier reprend les propositions qu'il a faites dans son article, « La traduction automatique des textes écrits scientifiques et techniques d'un langage dans un autre », publié en 1959.
- 50 Les outils habituels de la traduction manuelle, dictionnaires, grammaires et syntaxes sous leur forme usuelle, sont inadéquats. La traduction entre langues naturelles doit s'inspirer de la traduction entre langages artificiels et des méthodes de programmation automatique, et la construction d'une grammaire formelle de la langue source est nécessaire. Comme aucun procédé n'a jusqu'à présent permis de construire une telle grammaire pour aucune langue naturelle et qu'on ne peut attendre qu'elle soit construite pour commencer les recherches en traduction automatique, force est de rechercher des procédés plus rapides pour construire les grammaires spécifiques nécessaires.
- 51 C'est l'objectif de la technologie linguistique. Celle-ci consiste à inventorier et à classer des faits linguistiques, syntaxiques, morphologiques et lexicaux, à partir d'un corpus de textes à traduire. L'exécution des tâches matérielles exigées par ce travail de classification ne peut être menée à bien dans des délais raisonnables qu'avec l'aide de calculateurs automatiques. De plus, cette méthode de technologie linguistique demande une organisation particulière du travail qui suppose trois groupes d'intervenants : les théoriciens (les linguistes savants), les expérimentateurs (les linguistes exécutants) et les consultants (les traducteurs scientifiques et techniques).
- 52 Cette ingénierie linguistique avant l'heure trouvera un écho dans les rapports des sections de linguistique du CNRS (rapport de conjoncture du CNRS de 1963-1964) qui déclarent que la traduction automatique, comme la documentation automatique, arrivée au stade de l'exploitation rentable, doit faire appel à des « ingénieurs linguistes » comme on dit des « ingénieurs chimistes ». Le rapport signale toutefois que le problème principal, pour le développement de la linguistique en France, reste la pénurie de collaborateurs qui soient pourvus d'une formation linguistique « fondamentale »,  
La répartition des tâches : un problème de machines ?
- 53 Le rapport Sestier prévoit une répartition des tâches entre les deux équipes, parisienne et grenobloise, répartition fondée officiellement sur les différences de matériel et les recherches entreprises avant le 1<sup>er</sup> janvier 1960. Selon cette répartition, Grenoble est chargée de la morphologie, considérée comme plus facile et plus rapide à réaliser, et Paris, dont l'expérience est plus grande, se charge de la syntaxe. Puisque sa tâche est « simple », l'équipe de Grenoble décide de s'attaquer aussi aux problèmes de polysémie lexicale.
- 54 Cette décision est décrétée provisoire et fluctuante dès la première réunion du conseil scientifique le 20 février 1960. Martinet et Benveniste, qui en font partie, critiquent

vivement cette division entre morphologie et syntaxe. Celle-ci n'est pas pertinente, disent-ils, dès lors qu'il s'agit de comparer deux structures : il y aurait intérêt à partir d'une solution moins graphique et plus linguistique, c'est-à-dire de l'autonomie du syntagme, nominal et verbal.

- 55 Cette belle répartition va de plus se trouver très rapidement contrariée. Les prétextes en sont les différences de machines autant que les différences de méthodes.
- 56 Selon A. Sestier, le groupe de Grenoble développe un système morphologique adaptable uniquement sur machine binaire, donc strictement inutilisable pour le CETAP qui dispose d'une machine décimale de taille mémoire réduite. Par ailleurs, le système du CETAG lui paraît inutilement compliqué. Les Parisiens décident donc d'affronter directement l'analyse morphologique. A. Sestier insiste toutefois sur l'harmonisation indispensable des deux systèmes d'analyse morphologique développés indépendamment par les deux équipes.
- 57 Pour leur part, dans leur projet du 12 décembre 1960, les membres du CETAG montrent leur intention de faire aussi des recherches en syntaxe sur la traduction russe-français, japonais-français et allemand-français, en adoptant principalement le modèle que Sydney Lamb a développé au sein du groupe de l'université de Berkeley.
- 58 Sur la question des calculateurs utilisés, il y a un déséquilibre certain entre les équipes. Les programmes de traduction automatique exigent des machines à grande capacité mémoire. L'avantage est cette fois-ci au CETAG qui ne dispose au début que d'une machine Bull, binaire-décimale à mémoire-tambour de grande capacité, mais qui pourra dès la fin 1960 utiliser le GAMMA 60 et l'IBM 7090 du Centre de calcul de Grenoble, qui comptent parmi les machines les plus puissantes de l'époque.
- 59 Le CETAP ne peut utiliser qu'une IBM 650, machine décimale à 2 000 mots avec bandes magnétiques 727. A. Sestier adresse différents rapports au CNRS sur l'insuffisance des machines. Plusieurs solutions sont envisagées. Le comité de direction du 17 décembre 1960 décide d'examiner les possibilités offertes par la machine Mark II mise au point par l'*US Air Force* en collaboration avec le groupe de traduction automatique de Washington. Une autre solution consiste à envisager la construction d'une machine à traduire spécifiquement française en donnant des directives aux constructeurs IBM et Bull.
- 60 Dans un rapport de novembre 1960, A. Sestier préconise l'utilisation de Mark II, machine spécifiquement dédiée à la traduction automatique, intégrant une mémoire photocopique<sup>24</sup> de grande capacité à un calculateur IBM classique. B. Vauquois, dans un rapport de janvier 1961, se montre moins enthousiaste à l'égard de la machine Mark II qui, certes, présente l'avantage d'une consultation rapide de dictionnaires, mais qui est très peu adaptée à l'analyse syntaxique et sémantique, dans la mesure où elle ne peut dépasser le stade de l'étude de mots consécutifs qu'au prix d'énormes difficultés alors que le calculateur classique offre des possibilités immédiates.
- 61 Il ne semble pas que le projet d'utilisation de Mark II ait abouti, malgré la visite qu'ont effectuée A. Sestier et B. Vauquois au centre de l'*US Air Force* à Rome dans l'État de New York, lors de leur mission aux États-Unis de mai-juin 1961. Même si la solution adoptée provisoirement a été de renforcer les machines existantes (la IBM 650 a été complétée en mai 1961 par des disques magnétiques 355 à très grande capacité mais d'accès relativement lent), l'idée de construire une machine à traduire française n'est pas abandonnée.

## La dissolution du CETAP

- 62 L'abandon du travail en linguistique fondamentale, puis de l'idée qu'il suffit d'une technologie linguistique pour réaliser des dispositifs fonctionnels de traduction automatique, va progressivement conduire A. Sestier à démissionner et à dissoudre le CETAP en octobre 1962.

## L'abandon de la recherche en syntaxe

- 63 La position d'A. Sestier se radicalise lorsqu'en automne 1960, le CETAP décide d'arrêter les recherches sur la syntaxe « à la main », c'est-à-dire selon la méthode traditionnelle des linguistes, en raison de son rendement nécessairement faible, pour se consacrer à la mise au point d'outillage informatique pour linguistes. Par ailleurs, tous les groupes étrangers travaillant sur la syntaxe, selon lui, se heurtent aux mêmes difficultés théoriques, quant à la validité des modèles utilisés. « À la réflexion, rien de bien étonnant à cela : la linguistique est une science aussi ancienne que les mathématiques (4 000 ans au moins). Si les spécialistes les plus chevronnés continuent à discuter de ses principes mêmes, c'est sans doute qu'en quelques mois, on ne peut, avec une équipe de jeunes chercheurs tout juste sortie de sa phase d'organisation, faire plus que des générations de savants, même en entreprenant ces recherches avec des idées de mathématicien ou d'ingénieur, bref, en essayant d'examiner les faits linguistiques sous un angle neuf. C'est ce que prévoiaient, dès le départ, les maîtres de la linguistique française qui nous ont encouragés et conseillés, non sans quelque secret amusement de notre folle entreprise. »

## Les linguistes français qui s'intéressent à la traduction automatique

- 64 Qui sont ces maîtres de la linguistique française mentionnés par A. Sestier ? Les linguistes français sont présents dans les instances d'évaluation du CETA mais ne participent pas directement aux recherches. Outre Martinet, directeur de recherche d'Y. Gentilhomme, et Benveniste, le conseil scientifique du CETA comprenait en 1960 M. Lejeune, Marcel Cohen, Georges Gougenheim, Bernard Quemada et deux slavistes, Jean Train et Marc Vey. Plus tard, Jean Fourquet en fera partie. Au CNRS, la traduction automatique est mentionnée, dans les années 1963-1964, comme appartenant à la linguistique appliquée et est encouragée à ce titre. Alors que le cercle de la linguistique officielle, la SLP (Société de linguistique de Paris), se méfie de cette « linguistique pour ingénieurs » et n'ouvre les pages de son bulletin qu'à quelques comptes rendus de Georges Mounin sur l'état de la traduction automatique aux États-Unis et à ceux des slavistes, comme René L'Hermitte sur la situation en Union soviétique et dans les pays de l'Est, c'est au sein de l'ATALA que se retrouvent les linguistes qui s'intéressent à la traduction automatique. G. Gougenheim et M. Cohen y sont particulièrement actifs. A. Martinet, A. Culioli, B. Pottier, David Cohen et G. Mounin sont aussi très tôt membres de l'ATALA. La traduction automatique permettait, selon eux, de poser de véritables questions linguistiques.
- 65 Notons que la linguistique en France est aussi marquée pendant cette période par la mécanisation de la lexicologie qui en assure le renouveau et que la traduction automatique n'est pas le seul lieu de confrontation entre linguistique et informatique. En 1959 est créé à Besançon le Laboratoire d'analyse lexicologique sous la direction de B. Quemada. Le plan définitif de l'organisation d'un *Trésor* ou grand *Dictionnaire général historique de la langue française* est approuvé par la section de linguistique générale et des études littéraires en novembre 1960, servant de base à la création du *Trésor de la langue française* (TLF).

### Le groupe de traduction automatique de Nancy

- 66 Le groupe de Nancy est toutefois le seul groupe de traduction automatique dirigé par des linguistes. Il est créé lors d'une réunion organisée, le 25 mai 1960, à la faculté des lettres de Nancy par le doyen Schneider et A. Culioli, alors professeur dans cette même faculté, à laquelle participent, outre B. Pottier et G. Bourquin, É. Delavenay, le président de l'ATALA, A. Martinet, et M. Legras, directeur de l'Institut de calcul automatique de l'université qui met à la disposition du groupe une machine IBM 650.
- 67 Au départ, le groupe comprend deux linguistes, B. Pottier, professeur à la faculté des lettres de Strasbourg, et G. Bourquin, chargé d'enseignement à la faculté des lettres de Nancy. Avec le soutien du doyen Schneider et du recteur Imbs, le groupe de Nancy obtient d'être examiné par le CNRS lors du comité de direction du CETA dès janvier 1962 ; en juillet 63, il reçoit deux collaborateurs techniques supplémentaires. Le groupe de Nancy se donne comme langue source l'anglais et comme langues cibles le français et l'espagnol, choix, dit-il, qui se justifie par un intérêt pratique, la traduction d'ouvrages scientifiques. Sa méthode se présente comme un moyen terme entre la méthode empirique (la traduction mot à mot) et la formalisation mathématique, avec comme préoccupation essentielle l'étude des problèmes linguistiques posés par la traduction automatique.
- La démission d'A. Sestier
- 68 Plusieurs événements accélèrent le processus de démission d'A. Sestier. L'année 1962 correspond au moment où le financement du CETA par l'armée se trouve compromis par le remplacement du CASDEN par le CASD, dépourvu de ressources propres. C. de Gaulle avait alors un compte à régler avec les militaires et les a dépossédés de la recherche au profit de civils. On est en effet à la fin de la guerre d'Algérie et le putsch d'Alger est loin d'être oublié<sup>25</sup>.
- 69 Malgré un dossier circonstancié et l'appui du CNRS, la demande de subvention effectuée par A. Sestier auprès de l'OTAN<sup>26</sup> n'aboutit pas. La DRME (Délégation ministérielle pour l'armement, Direction des recherches et moyens d'essai) prend le relais mais pour un quart seulement du financement, le CNRS continuant à financer le CETA à cinquante pour-cent pour un budget total d'environ un million de francs. Le CNRS fait alors appel à un certain nombre de ministères afin de financer les vingt-cinq pour-cent restants. Lors d'une réunion regroupant en mai 1962 la direction du CNRS et les représentants des différents ministères concernés, ceux-ci se déclarent favorables aux recherches en traduction automatique, mais exposent leurs difficultés à disposer des crédits demandés. Quatre-vingt-dix mille francs seulement seront réunis à l'issue de cette réunion; le SDEC se montrant le plus généreux en accordant cinquante mille francs.
- 70 À l'occasion de ces difficultés de financement, on voit apparaître, dans les réunions du CNRS chargées de l'évaluation du CETA, la question de la demande sociale en traductions. Bien que constituant un élément important du rapport Sestier, préalable à la création du CETA et figurant dans ses statuts au travers de la création d'un centre de production de traductions devant assurer l'autofinancement du centre, ce point était pourtant jusqu'alors peu présent dans les discussions. C'est lors d'une réunion de juin 1962 qu'est évoquée la pertinence de limiter les recherches à la traduction du russe vers le français et à celle des textes techniques et scientifiques. Mais cela n'intéresse pas nécessairement certains ministères, par exemple, le ministère de la Culture. À l'issue de la réunion, ces deux orientations sont toutefois conservées, notamment en

raison de l'intérêt que présente la traduction des textes techniques et scientifiques pour les pays en voie de développement.

#### Ambiguïtés pour les humains/ambiguïtés pour la machine

Les mots, les expressions, les syntagmes et les propositions exprimées dans une langue naturelle ont tendance à être interprétables de plusieurs façons différentes. Malgré cela, on remarque que les êtres humains ne butent presque jamais sur des ambiguïtés qu'ils ne perçoivent même que rarement (et difficilement) dans un contexte donné. Pour la machine, en revanche, tous les cas d'homonymies et de polysémie lexicale ou syntaxique donnent lieu à des ambiguïtés. C'est un des problèmes les plus difficiles de l'analyse automatique du langage.

Si l'on considère les deux exemples suivants, célèbres cas d'école pour le traitement automatique des langues, deux interprétations sont possibles pour chaque phrase :

le pilote ferme la porte

le (art., pron.) pilote (nom, verbe) ferme (verbe, nom, adj.) la (art., pron., nom) porte (nom, verbe)

interprétation 1 : le (art.) pilote (nom) ferme (verbe) la (art.) porte (nom)

interprétation 2 : le (art.) pilote (nom) ferme (adj.) la (pron.) porte (verbe)

la belle ferme le voile

la (art., pron.) belle (nom, adj.) ferme (verbe, adj.) le (art., pron.) voile (verbe, nom)

interprétation 1 : la (art.) belle (nom) ferme (verbe) le (art.) voile (nom)

interprétation 2 : la (art.) belle (adj.) ferme (nom) le (pron.) voile (verbe)

Pour ce type de phrases, le contexte et/ou la prosodie permet de lever l'ambiguïté qui est rarement perçue par les humains. Pour la machine, la seule solution consiste parfois à fournir les deux interprétations.

#### Ambiguïtés de parenthésages

Soit l'exemple suivant, utilisé par Maurice Gross dans son rapport critique sur la traduction automatique : la directrice de banque enrhumée.

Deux interprétations sont possibles :

((la directrice de banque) enrhumée)

(la directrice (de banque enrhumée))

Pour lever l'ambiguïté, M. Gross préconise une analyse linguistique qui opère la distinction animé/inanimé ne permettant le regroupement de « enrhumée » qu'avec un humain à savoir « la directrice ».

#### Ambiguïtés de sens commun

Exemple extrait du fameux rapport de Y. Bar-Hillel (1960) contre la traduction automatique : *the pen was in the box*

*the box was in the pen*

La seconde phrase est possible si l'on sait que *pen* en anglais signifie à la fois crayon et parc à bébé, et qu'on peut la rencontrer dans une phrase telle que: *Little John was looking for his toy box. Finally he found it. The box was in the pen. John was very happy.*

Selon Y. Bar-Hillel, un ordinateur est incapable de résoudre cette ambiguïté qui suppose des connaissances de sens commun, notamment sur les tailles relatives d'un crayon et d'un parc à bébé.

#### Ambiguïté de référence des pronoms

Les suites suivantes montrent que pour trouver la référence des pronoms *they*, la machine doit faire des inférences du type :

si quelqu'un est assassiné, il est mort, il faut l'enterrer

si quelqu'un est assassiné, il y a un assassin, il faut le retrouver.

The men murdered the women. They were caught three days later.

The men murdered the women. They were buried three days later.

Comme pour l'exemple précédent, la machine doit avoir des connaissances de sens commun. Des systèmes de représentation des connaissances ont été mis au point dans le cadre de l'intelligence artificielle pour tenter de résoudre ces problèmes. A l'heure actuelle, ils sont souvent utilisés dans les systèmes de traduction automatique.

Les *garden-paths sentences* ou phrases buissonnières

Ces phrases sont aussi des cas d'école. Elles sont ambiguës pour les humains comme pour l'analyse syntaxique automatique. Dans le processus d'analyse, le parseur se heurte à une impossibilité de choix et doit revenir en arrière :

the boat floated clown the river sank

le bateau sombra en descendant la rivière

the horse raced past the barn fell

le cheval tomba en passant près de la grange

Le même phénomène existe aussi en français :

sur le miroir de l'eau stagnait une couleur pourpre

quand il pleure beaucoup le regarde

Voir : M. Gross, « Notes sur certains aspects des recherches en linguistique au CNRS », rapport au directeur général du CNRS du 26 septembre 1967 et Y. Bar-Hillel, 1960, « The present Status of Automatic Translation of Languages », in *Advances in Computers*, vol.1, EC. Alt ed. Academic Press, N.Y., London, 1960, pp. 91-141.

- 71 Limitation des crédits et discussions sur la demande sociale en traduction automatique soulèvent la question cruciale des délais dans lesquels il sera possible d'effectuer les premières expériences. A. Sestier et B. Vauquois ne sont pas d'accord sur ce point : trois ans suffisent pour B. Vauquois, alors que pour A. Sestier, rien n'est possible avant huit ans. Celui-ci laisse entendre qu'il ne pourra pas, avec des délais si courts, continuer à assumer la direction du CETAP. Le directeur du CNRS assure que le CNRS est prêt à jouer le jeu et à continuer de financer le CETA pour moitié et que l'on trouvera les crédits nécessaires de façon que l'affaire se poursuive au moins jusqu'en 1970, date butoir d'A. Sestier.
- 72 Un autre facteur enfin est à prendre en compte dans la démission d'A. Sestier : la publication du rapport Bar-Hillel. Les Français semblent n'avoir eu que tardivement<sup>27</sup> connaissance du rapport Bar-Hillel, publié en 1960, à la suite d'une mission d'évaluation des centres de traduction automatique, commanditée en 1958 par la *National Science Foundation*. Ce rapport présentait de façon très négative la plupart des activités et des résultats des centres de traduction automatique dans le monde et surtout aux États-Unis. Lors du congrès de l'IFIP à Munich en août 1962, où il organisait un symposium intitulé « *Modern techniques of language translation* », A. Sestier le mentionne comme une chose assez récente, mais la façon dont il en parle laisse prévoir sa décision<sup>28</sup>. La direction du CNRS n'ignore pas les positions d'Y. Bar-Hillel sur la traduction automatique. Dans un courrier du 10 octobre 1962, J. Coulomb attire l'attention de M. Lejeune sur le texte d'Y. Bar-Hillel dans les actes du congrès de l'IFIP que vient de lui envoyer A. Sestier.
- 73 Finalement, en octobre-1962, A. Sestier donne sa démission du CETAP. Dans le rapport d'activité du CNRS 1961-1962, les raisons de cette démission sont exposées de la façon suivante : « La construction d'une syntaxe pour une langue naturelle est un problème de recherche fondamentale inséparable de celui de la sémantique; ce problème a été abordé prématurément et exigerait, pour être résolu, de nombreuses recherches préliminaires (de psychologie et de théorie des automates, notamment) auxquelles la structure actuelle du CETAP ne se prête pas. »
- 74 Après la dissolution du CETAP, certains de ses membres, comme M. Gross et Y. Gentilhomme, rejoignent le Laboratoire de calcul numérique de l'Institut Blaise-Pascal. Quant à Jacques Perriault, il rejoint la SAD (Section d'analyse documentaire) dirigée par J.-C. Gardin. Le groupe de Nancy devient, en 1969, le CRAL (Centre de recherche et d'applications linguistiques de Nancy), associé au CNRS, qui abandonne les travaux en traduction automatique proprement dits pour des travaux plus généraux en traitement

automatique des langues. Ainsi, une des sections du CRAL se consacre au traitement automatique des textes sémitiques, à la paléographie automatique et à l'analyse documentaire automatique.

- 75 Ne reste donc plus que le groupe de Grenoble, qui, tout en continuant à être rattaché à l'Institut Blaise-Pascal, devient laboratoire propre du CNRS en 1963, en gardant le nom de CETA<sup>29</sup>. À partir de fin 1962, son directeur, B. Vauquois, monte en puissance dans les instances représentatives. En 1963, il est membre de la section 22 du Comité national : « Linguistique générale, langues moderne et littérature comparée ». Puis, en 1969, il sera membre de la section 28 : « Linguistique générale, langues et littératures étrangères ». Dès 1965, il est vice-président de l'ATALA. IL en sera le président de 1966 à 1971. Il sera un des fondateurs, également en 1965, de l'ICCL (*International Committee on Computational Linguistics*).

#### Conclusion

- 76 Une des spécificités de l'histoire du Centre de Grenoble, c'est qu'il ne semble pas avoir été ébranlé par la crise de 1965, contrecoup de la crise américaine, qui affectera profondément l'ATALA. Le déclin de la traduction automatique, amorcé depuis 1960 aux États-Unis suite à la publication du rapport Bar-Hillel, devient inéluctable avec la publication du rapport de l'ALPAC en 1966 qui mit fin aux subventions des recherches en traduction automatique aux États-Unis et pratiquement dans le monde entier.
- 77 Le CNRS accuse la crise de façon modérée et ne cesse de soutenir le CETA. Selon le rapport de conjoncture du CNRS de 1963-1964, les sections de linguistique du Comité national maintiennent leur soutien prudent à la traduction automatique. La section 23, « Linguistique française et études littéraires », mentionne dans ses recommandations l'intérêt de la linguistique appliquée, dont fait partie la traduction automatique.
- 78 La section de mathématiques appliquées parle d'une phase de reflux, mais précise que les recherches en cours en traduction automatique doivent être poursuivies avec toutes les ressources actuelles, et, qu'à condition d'être menées en profondeur, elles conduiront un jour à des résultats intéressants. La section ne semble pas affectée par la démission d'A. Sestier et encourage la collaboration entre linguistes et logiciens pour mener à bien une « étude minutieuse et patiente des langues naturelles ». Par ailleurs, l'idée de construire une machine dédiée à la traduction est loin d'être abandonnée. La section préconise des études de logique de circuits et d'organes de machine adaptées à des fins de traduction, de même que des études de programmation dans le domaine non numérique visant à la création et à la compilation de langages adaptés.
- 79 Concernant plus spécifiquement les activités du CETA, le conseil scientifique de décembre 1966 recommande de diversifier les domaines d'applications des méthodes élaborées par le centre. La traduction automatique n'est plus à considérer comme le but premier et doit apparaître comme un produit dérivé d'études linguistiques et logiques plus générales.
- 80 Donc, malgré la crise, le CETA tient bon. Ses crédits sont augmentés et ses locaux agrandis<sup>30</sup>. R. de Possel demande, au nom du conseil scientifique réuni le 15 décembre 1966, les félicitations pour le travail accompli par B. Vauquois. Enfin, la seconde conférence internationale sur le traitement automatique des langues se tient en août 1967 au CETA qui y présente sa première expérience de traduction sur ordinateur (le premier test aura lieu en juin 1967).

- 81 Toutefois, il faut pondérer cette impression de prospérité : le CETA se porte très bien en 1967, mais moins bien en 1971<sup>31</sup>, date à laquelle il sera privé d'une partie de ses forces, chercheurs et moyens, et perd son statut privilégié de laboratoire propre du CNRS. Cela dit, sa reconversion se fait progressivement<sup>32</sup>, et non brutalement comme aux États-Unis. À cet égard, on peut invoquer le fait que l'instance d'évaluation et le bailleur de fonds qu'est le CNRS assure une certaine stabilité aux équipes en raison de ses caractéristiques propres : présence des intéressés dans les instances d'évaluation, façon dont est prise en compte la demande sociale, évaluation des modèles théoriques et évaluation des résultats. Dans le cas de la traduction automatique, qui doit viser avant tout une production industrielle, les choses sont éminemment ambiguës. On peut d'ailleurs supposer que cette stabilité a permis au CETA, devenu GETA en 1971, de prendre un second souffle et de réaliser dès le début des années 1970 Ariane-78, un système multiniveaux fondé sur la méthode de transfert, considéré comme un des plus importants et des plus influents de l'époque.
- 82 Avec la publication du rapport de l'ALPAC en 1966, la traduction automatique n'a plus eu droit de cité aux États-Unis jusqu'au début des années 1990. Elle ne sera définitivement pas cette science du langage appliquée qu'elle aspirait à devenir et le mythe de la traduction de très bonne qualité (équivalente à la traduction humaine) est bel et bien révolu.
- 83 Pourtant, la traduction automatique a survécu. Elle est, à l'heure actuelle, un secteur important du traitement automatique des langues. Le domaine s'est considérablement modifié au fil des décennies, tant par le développement technologique que par la transformation de la demande en traductions suscitée par la mondialisation de la consommation. Le succès des outils d'aide à la traduction semble avoir finalement permis de conjurer le spectre de l'ALPAC. Mais ce succès passe par un double paradoxe : on n'exige plus des systèmes de traduction automatique de fournir des traductions ; quant aux stations d'aide aux traducteurs, elles n'utilisent les systèmes de traduction automatique, quand ils existent, que comme un outil parmi d'autres. Autrement dit, on en arrive à se demander si l'automatisation de la traduction a encore réellement besoin de la traduction automatique.<sup>33</sup>

---

## BIBLIOGRAPHIE

En plus des ouvrages et rapports cités en notes, l'auteur a consulté :

- Fonds documentaire du CNRS de Gif-sur-Yvette : les rapports d'activité annuels du CNRS (1950-1975) et les rapports de conjoncture du CNRS (1959-1974).
- Fonds d'archives du CNRS de Gif-sur-Yvette : dossiers scientifiques CNRS 910024 DPC.
- Fonds d'archives du Centre des archives contemporaines de Fontainebleau : archives historiques F 780309 (53) et (54).
- Archives de l'ATALA.

## - Rapports :

A. Sestier, « L'automatisation de l'analyse syntaxique », document du CETAP, note n° 13, 19 octobre 1961.

Unesco, *Report on scientific and technical translating and related problems* [WS/073.119], 1953.

B. Vauquois, « Étude sur le choix du matériel destiné à la traduction automatique », 13 janvier 1961.

- *Cahiers pour l'histoire du CNRS* (1939-1989).

- J.-L. Chiss et Christian Puech (dir.), « La linguistique comme discipline en France », *Langue française*, n° 117, février 1998.

A. Sestier, « La traduction automatique des textes écrits scientifiques et techniques d'un langage dans un autre », in *Ingénieurs et techniciens*, partie I et II, 1959.

## NOTES

1. Cf. *La Revue pour l'histoire du CNRS*, n° 1, novembre 1999.

2. Rapport qui interrompt les recherches pour une période allant d'une dizaine d'années en Europe jusqu'à près de vingt-cinq ans aux États-Unis.

3. Sur les enjeux de la traduction automatique de la fin des années 1940 à nos jours, voir J. Léon, « La traduction automatique I : les premières tentatives jusqu'au rapport ALPAC », et « La traduction automatique II : développements récents » *Handbücher zur Sprachund Kommunikationswissenschaft*, Berlin, Éd. Walter de Gruyter and co., volume 3, Histoire des sciences du langage (à paraître).

4. Reposant sur une consultation de dictionnaires, une pré et post-édition, et une langue intermédiaire, cette machine était présentée comme une méthode économique de traduction multilingue dans la mesure où chaque traducteur n'était tenu de connaître qu'une seule langue.

5. Sur les débuts de la traduction automatique aux États-Unis, en Grande-Bretagne et en URSS, voir W. I. Hutchins, *Machine translation, part, present, future*, Ellis Horwood Ltd, 1986; J. Léon, « Les débuts de la traduction automatique en France (1959-1968) : à contretemps ? », *Modèles linguistiques*, tome XIX, fascicule 2, p. 55-86, 1998 ; « La mécanisation du dictionnaire dans les premières expériences de traduction automatique (1948-1960) », *History of Linguistics* 1996, vol. II, D. Cram, A. Linn, E. Nowak (dir.), p. 331-340, John Benjamins Publishing Company, 1999, et « Traduction automatique et formalisation du langage. Les tentatives du Cambridge Language Research Unit (1955-1960) », in *The History of Linguistics and Grammatical Praxis*, P. Desmet, L. Jooker, P. Schmitter, P. Swiggers (dir.), Louvain-Paris, Peeters, 2000, p. 369-394 ; S. Archaimbault et J. Léon, « La langue intermédiaire dans la traduction automatique en URSS (1954-1960). Filiations et modèles », *Histoire-Épistémologie-Langage*, 1997, p. 19-2 et 105-132.

6. « Language and machines. Computers in translation and linguistics », *A report by the Automatic Language Processing Advisory Committee (ALPAC), National Academy of Sciences, National Research Council*, 1966.

7. On ne parlait évidemment pas d'informatique à l'époque, le terme n'étant apparu qu'en 1962 sous la plume de P. Dreyfus (*Dictionnaire des mots nouveaux*, P. Gilbert, 1972),

formé à partir de information et automatique. Il ne fut réellement en usage qu'à la fin des années 1960.

8. Henri Boucher, « L'informatique dans la défense », actes du colloque sur l'histoire de l'informatique en France, Grenoble, mai 1988, p. 83-102.

9. G. Ramunni, *La physique du calcul. Histoire de l'ordinateur*, Paris, Hachette, 1989, 287 p. ; « La non-construction du premier calculateur électronique au CNRS (I) et (II) », *Micro-bulletin* nos 42, décembre 1991 et 43-44-45, février-mars 1992.

10. Concernant l'histoire de l'Institut Blaise-Pascal, voir Pierre Mounier-Kuhn, « L'Institut Blaise-Pascal du CNRS de 1946 à 1969 », actes du colloque *De Bourbaki à la Machine à Lire – Journée d'hommage à R. de Possel*, ENS-Ulm, 16 novembre 1994, p. 13-30.

11. G. Ramunni, « La non-construction... », *op. cit.*

12. Sur la situation de la linguistique en France au début des années 1960, voir, entre autres, l'article de J. -C. Chevalier et P. Encrevé, « La création de revues dans les années 1960. Matériaux pour l'histoire récente de la linguistique en France », *Langue française*, n° 63, et le numéro de *Langue française* consacré à « La linguistique comme discipline en France » (n° 117).

13. M. Lejeune est spécialiste de grec ancien des langues celtiques italiques, et s'est beaucoup intéressé à l'histoire des systèmes d'écriture.

14. J. Coulomb, M. Lejeune et É. Delavenay ont été, la même époque, élève à l'École normale supérieure Ulm.

15. É. Delavenay, *Témoignage. D'un village savoyard au village mondial*, diffusion Édisud, Aix-en-Provence, 1992.

16. J. E. Holmstrom, « Rapport sur les dictionnaires scientifique et techniques multilingues », Paris, Unesco, 1951.

17. É. Delavenay et K. Delavenay, *Bibliographie de la traduction automatique*, Gravenhage, Mouton and co, 1960.

18. Sur la création de l'ATALA et sur la biographie de ses membres fondateurs, voir A. Lentin « Naissance et premiers pas de l'ATALA : quelques souvenirs et quelques réflexions », *Traitement automatique des langues*, 1992, vol. 33, n° 1-2, p. 7-24, et J. Léon, « De la traduction automatique à la linguistique computationnelle. Contribution à une chronologie des années 1959-1965 », *Traitement automatique des langues*, 1992, vol. 33, n° 1-2, p. 25-44.

19. Outre les enjeux purement politiques et militaires, renforcés dans un contexte de guerre froide, qui motivaient les besoins en traduction du russe, il existait également une très grande rivalité scientifique entre Américains et Soviétiques qui stimulait la production de machines à traduire. Les Américains étaient persuadés que les Soviétiques étaient très en avance sur eux sur le plan scientifique, non seulement dans le domaine de l'aérospatiale (le premier Spoutnik date de 1957) mais aussi dans le domaine de la traduction automatique. Cette opinion était partagée par les Français, si l'on en juge par la première version d'une demande de subvention adressée par A. Sestier à l'OTAN en juin 1961 qui se termine par le constat que, bien que l'effort des Soviétiques en matière de traduction automatique soit, de l'avis même des experts américains, encore plus grand que celui des Etats-Unis, on n'a, depuis 1958, que peu de renseignements sur les travaux russes. Une des hypothèses est que « les Soviétiques préparent à l'Occident une surprise spectaculaire, telle, peut-être que l'édition de traductions machine de livres soviétiques ».

20. Le CASDEN était déjà en charge de la traduction d'articles scientifiques russes.

21. Le comité de direction, présidé par J. Coulomb, directeur du CNRS, comprenait également le doyen Pérès, É. Delavenay, R. de Possel, M. Lejeune, Raymond Cheradame directeur des études de l'École polytechnique, A. Sestier et B. Vauquois.
22. R. de Possel va s'attaquer à un des problèmes en amont de la traduction automatique, à savoir la lecture automatique des caractères. Cf. actes du colloque *De Bourbaki à la Machine à Lire – Journée d'hommage à R. de Possel*, ENS-Ulm, 16 novembre 1994.
23. É. Delavenay, *La Machine à traduire*, Paris, PUF, « Que sais-je? », 1960.
24. Il s'agit d'un système permettant d'enregistrer 33 millions de bits sur quatre disques de verre revêtus d'une émulsion photographique, sous forme de carrés rouges, verts, blancs ou noirs, lus à l'aide de tubes à rayons cathodiques et de photomultiplicateurs.
25. Une grande partie des officiers engagés dans des activités de recherche ont alors démissionné de l'armée pour rejoindre l'industrie. Nous remercions René Moreau pour cette information.
26. A. Sestier, « Justification d'une demande de subvention à l'OTAN », version 1 de juin 1961 ; version 2 du 8 octobre 1961. Il est probable que cette demande a souffert de la détérioration des relations de la France avec l'OTAN. On sait que Ch. de Gaulle manifesta une vive opposition, en décembre 1962, aux accords conclus à Nassau par J. F. Kennedy et Macmillan sur le projet de force multilatérale. Cette opposition aboutira en mars 1966 à un double retrait : des forces françaises soumises à l'autorité de l'OTAN et des forces étrangères qui se trouvaient sur le sol français dans le cadre de l'alliance.
27. Pourtant É. Delavenay possédait la première version de ce rapport datant de février 1959 (cf. archives de l'ATALA)
28. Nous remercions André Lentin pour cette information.
29. Sur l'histoire des réalisations du CETA, voir l'article de Christian Boitet qui prit la direction du GETA, après la disparition de B. Vauquois en 1985, « TA et TAO à Grenoble... 32 ans déjà ! » *TAL*, vol. 33, n° 1-2, 1992, p. 45-84.
30. Avec un budget d'un million de francs et 20 collaborateurs en 1963, il atteint les 35 personnes début 1965. Ce budget de 1 million sera augmenté en 1966. En 1964, un budget spécial est voté pour la construction du bâtiment du CETA sur le campus de Saint-Martin-d'Hères, bâtiment qui sera inauguré en août 1967.
31. Sur ce point et l'impact du rapport de M. Cross au directeur général du CNRS, « Notes sur certains aspects des recherches en linguistique au CNRS », 26 septembre 1967, voir J. Léon, « Les débuts de la traduction automatique en France (1959-1968) : à contretemps ? », *Modèles linguistiques*, tome XIX, fascicule 2, 1998, p. 55-86.
32. Actuellement, le GETA, toujours sous la direction de C. Boitet, poursuit ses travaux sur la traduction assistée par ordinateur, dans le laboratoire CLIPS de l'Institut d'informatique et de mathématiques appliquées de Grenoble.
33. Ce point est argumenté dans J. Léon, « De la traduction automatique à l'automatisation de la traduction : parcours historique », in *Bulag*, n° 25, 2000, p. 5-21.

---

## INDEX

**Mots-clés** : Traduction automatique, ATALA, CETA, CETAP, CETAG

## AUTEUR

### JACQUELINE LÉON

Jacqueline Léon est chargée de recherche au CNRS au Laboratoire d'histoire des théories linguistiques.