

**Traduction automatique et modes d'intégration
dans les sciences du langage.
Le cas des premiers travaux
britanniques et russes**

Jacqueline LÉON

Université Paris Diderot – CNRS

Résumé:

Dans cet article, je propose d'examiner comment la traduction automatique, technologie de guerre froide conçue aux États-Unis en dehors de la linguistique, a été intégrée dans les sciences du langage. On examinera les premières expériences menées dans les années 1950 en Grande-Bretagne et en URSS en montrant que, dans les deux cas, cette intégration s'est effectuée grâce à un ancrage solide dans les traditions linguistiques et culturelles, et sur le long terme. Ont été développées des méthodes de traduction automatique par langue intermédiaire fondées sur des théories sémantiques propres à chaque tradition, très différentes des méthodes américaines centrées sur l'analyse syntaxique et les langages formels.

Mots-clés: traduction automatique, horizon de rétrospection, horizon de projection, tradition britannique, tradition russe, langues universelles, langues intermédiaires

INTRODUCTION

La traduction automatique (*infra* TA), apparue à la fin des années 1940, est issue des sciences de la guerre¹, à l'interaction des sciences et de l'ingénierie, qui ont connu un essor particulier aux États-Unis. La TA a été développée en tant que technologie à haute valeur stratégique destinée à fournir des traductions en série, au départ par les États-Unis, puis par les trois autres pays, Grande-Bretagne, URSS et France, tenus pour les «vainqueurs» de la Seconde guerre mondiale, possédant l'arme atomique et impliqués au premier chef dans la guerre froide. Bien que son objectif ait été de traduire des langues dans d'autres langues, la TA a été conçue en dehors de la linguistique et a dû être intégrée dans un second temps par les sciences du langage. Cette intégration, qu'on peut qualifier de tournant de l'automatisation pour les sciences du langage, s'est effectuée de façon diverse selon les traditions intellectuelles, culturelles et linguistiques. Dans cet article, j'examinerai plus particulièrement les traditions britannique et russe qui, tout aussi curieux que cela puisse paraître, peuvent être comparées sur le plan des «modes d'intégration» de la TA. Pour appréhender ce tournant, je vais m'appuyer sur les notions d'*horizon de rétrospection*² et de *projection*³ introduites par Sylvain Auroux⁴, en développant deux parties: a) la mise en place de la TA aux États-Unis comme nouvel horizon de rétrospection pour les sciences du langage; b) une étude comparative de l'intégration de ce nouvel horizon dans les traditions britannique et russe.

¹ Dahan, Pestre (éds), 2004.

² «Parce que c'est un phénomène social (et non biologique), l'activité scientifique est impossible sans transmission du savoir, sans le fonctionnement institutionnel des formations, du cumul des connaissances (par exemple sans bibliothèque), et sans la mémoire individuelle. Le texte scientifique comprend essentiellement deux types d'expressions. Les unes font référence au seul domaine des phénomènes, manipulés à l'aide de concepts appartenant à la connaissance commune [...] les autres font référence à d'autres travaux, par définition antérieurs. Nommons *horizon de rétrospection*, cet ensemble de références [...] la structure de l'horizon de rétrospection est une cause dans la production momentanée de la recherche [...] Mais à l'inverse, la structure du système scientifique détermine celle des horizons de rétrospection» (Auroux 1987, p. 29).

³ «Parce qu'il est limité, l'acte de savoir possède par définition une épaisseur temporelle, un *horizon de rétrospection*, aussi bien qu'un *horizon de projection*. Le savoir (ou les instances qui le mettent en œuvre) ne détruit pas son passé, comme on le croit souvent à tort, il l'organise, le choisit, l'oublie, l'imagine ou l'idéalise, de la même façon qu'il anticipe son avenir en le rêvant tandis qu'il le construit. Sans mémoire et sans projet, il n'y a tout simplement pas de savoir» (Auroux 1995, p. 49).

⁴ Auroux 1987 et 1995.

1. CONSTRUCTION D'UN HORIZON DE RÉTROSPECTION POUR LA TA COMME TECHNOLOGIE DE GUERRE

Deux personnalités et trois rapports ont impulsé et jalonné les débuts de la TA, en délimitant une périodisation en trois temps, un début, un milieu, une fin, sur une période courte de 17 ans (1949-1966). Il s'agit de Warren Weaver et de Yehoshua Bar-Hillel, et pour les rapports du *Memorandum* de Weaver (1949), du rapport Bar-Hillel (1960) et du rapport de l'ALPAC (1966).

1.1. LE MEMORANDUM DE WEAVER (1949)

Au début l'histoire est américaine, et la figure de Warren Weaver (1894-1978) y joue un rôle central. Mathématicien, il dirige, pendant la Seconde guerre mondiale au MIT, la section D2 «Fire Control» du tout puissant *Office of Scientific Research and Development*, qui est au cœur du dispositif militaro-scientifique américain. Après la guerre, Weaver retourne à la *Rockefeller Foundation* et fait partie de nombre d'institutions gouvernementales d'organisation et de financement de la recherche, civiles comme la NSF (*National Science Foundation*) et l'*American Association for the Advancement of Science*, ou militaire comme la *Naval Research Advisor Committee*. De plus, il est un des membres fondateurs de la *Rand Corporation*, premier *think tank* américain, fondée par l'*US Air Force* et regroupant des logiciens et des mathématiciens.

La grande affaire de l'après-guerre, c'est la guerre froide, où, dans la compétition scientifique et militaire avec les Soviétiques, la traduction des textes scientifiques russes en anglais occupe pour les Américains une place cruciale. La TA correspondait, pour Weaver, au type d'objectif défini par la *Rand*, dans le sillage de la «culture de guerre»⁵: interpénétration des sciences fondamentales et de l'ingénierie; automatisation d'un travail jusqu'alors spécifiquement humain; définition d'un objectif unique, en l'occurrence la production en série de traductions scientifiques et techniques; mise en œuvre de moyens humains et techniques colossaux pour réaliser l'objectif en privilégiant les solutions formelles par rapport aux solutions humaines. En diffusant son *Memorandum* «*Machine Translation*» en 1949, Weaver impulse des recherches en TA au sein de la *Rand Corporation* et dans trois universités américaines, dont le *Research Laboratory of Electronics* du MIT où Bar-Hillel va être embauché en 1951.

⁵ Dahan, Pestre (éds), 2004.

1.2. UNE LINGUISTIQUE POUR INGÉNIEUR

Dans le cadre de la culture scientifique de guerre, la linguistique n'a pas de place. La TA est une technologie destinée à produire des traductions en série et économiquement rentables où les questions linguistiques sont considérées comme non prioritaires au regard des problèmes liés à l'architecture et à la limitation de la puissance des machines, limitation de mémoire, de vitesse et de stockage. En créant une linguistique pour la machine [*machine translation linguistics*], les ingénieurs vont d'ailleurs dénier à la linguistique toute légitimité dans l'automatisation de la traduction, en créant des objets linguistiquement artificiels, notamment en morphologie pour la confection de dictionnaires électroniques⁶. On assiste ainsi à la négation de l'horizon de rétrospection des néo-bloomfieldiens qui avaient consacré plusieurs décennies à décrire la morphophonologie des langues.

1.3. RAPPORT BAR-HILLEL ET RAPPORT DE L'ALPAC

Malgré le succès médiatique de la première démonstration sur ordinateur qui a eu lieu en 1954 à New York⁷, et qui fut suivie de la création de nouveaux centres de TA, notamment des premiers centres en Grande-Bretagne et en URSS, les bailleurs de fonds, NSF et CIA, commencent à douter de l'efficacité de cette nouvelle technologie, dont les résultats s'avèrent très médiocres.

Bar-Hillel (1915-1975) va jouer alors un rôle considérable. Philosophe des sciences et logicien, il est nommé premier chercheur à plein temps en TA au MIT en 1951, et sera chargé en 1958 par la NSF d'établir un rapport sur la TA. Il enquêtera sur la vingtaine d'équipes existant dans le monde: dix centres aux États-Unis, sept centres en URSS, deux centres en Grande-Bretagne, un centre en Italie. L'argumentation de Bar-Hillel, dans son rapport publié en 1960, aboutit à un jugement sans appel. La «Fully Automatic High Quality Translation»⁸, préconisée par plusieurs groupes de TA, est une illusion et un objectif déraisonnable même pour les textes scientifiques. Malgré les investissements considérables, très peu de problèmes linguistiques ont été résolus. Le seul objectif «raisonnable» pour la commercialisation de la traduction scientifique de bonne qualité, c'est la traduction humaine assistée par ordinateur.

Le relais institutionnel est pris par la mise en place du comité ALPAC (*Automatic Language Processing Advisory Committee*) dont le rapport, publié en 1966, aura des conséquences importantes certes sur le

⁶ Léon 1999.

⁷ Celle-ci, qui consistait dans la traduction par la machine de quelques phrases mettant en jeu un dictionnaire de quelques mots et quelques règles de syntaxe, a connu un grand succès.

⁸ 'Traduction entièrement automatisée de très bonne qualité'.

devenir de la TA, mais aussi sur le statut et la réorganisation de la linguistique aux États-Unis. L'ALPAC, en provoquant l'arrêt des financements des groupes de TA aux États-Unis et dans le monde en général, clôt cette période de 17 ans des débuts de la TA. Toutefois, en décrétant la légitimité de la linguistique computationnelle comme «nouvelle linguistique», il joue également un rôle important dans l'histoire des sciences du langage. La linguistique computationnelle, fondée sur l'interaction entre langages formels à base logico-mathématique, analyse syntaxique et programmation marque le tournant de l'automatisation des sciences du langage.

Bar-Hillel joue là encore un rôle essentiel. Dès 1952, il préconise le développement d'une syntaxe opérationnelle pour la TA à base logico-mathématique, à partir des travaux du logicien Kazimierz Ajdukiewicz et des grammaires transformationnelles de Zellig Sabbetai Harris, dans le sillage des travaux de Rudolf Carnap. Il faut noter que c'est grâce à cet environnement, le groupe de TA au MIT où il fut recruté en tant qu'élève de Harris, et les travaux en analyse syntaxique automatique de Bar-Hillel, que Noam Chomsky a commencé à travailler sur la grammaire générative et transformationnelle. Le rapport de l'ALPAC s'appuie d'ailleurs sur le programme chomskyen pour promouvoir la linguistique computationnelle. Le noyau dur et scientifique de la linguistique computationnelle, c'est l'analyse syntaxique automatique.

1.4. UN NOUVEL HORIZON DE RÉTROSPECTION

La TA n'est donc pas créée *ex nihilo*, elle a son propre horizon de rétrospection. C'est celui de son maître d'œuvre, Weaver, homme-clé des sciences de la guerre, et c'est aussi celui de Bar-Hillel, instaurateur de l'analyse syntaxique automatique. Pour Weaver, ce sont la théorie de l'information, la logique, les statistiques, les probabilités, le calcul numérique et la cryptographie qui vont servir d'horizon de rétrospection à la constitution de la TA comme première application non numérique des calculateurs électroniques. Bar-Hillel ancre la TA dans la première mathématisation du langage, la logique mathématique des années 1930 et l'analyse syntaxique.

Cet horizon de rétrospection va s'intégrer à la linguistique américaine de deux façons. La TA ravive chez les néo-bloomfieldiens, bien que de façon relativement brève, un intérêt pour la traduction⁹. Parallèlement, la possibilité d'automatisation conduit à un développement de la formalisation des procédures de description chez les bloomfieldiens, notamment Charles F. Hockett et Zellig Harris. Les représentations arborescentes deviennent dynamiques et on voit apparaître les premières grammaires génératives préfigurant les travaux de Chomsky des années 1955-1957¹⁰.

⁹ Archaimbault, Léon (à paraître).

¹⁰ Léon (à paraître [b]).

Les tentatives de formalisation, analyse syntaxique de Bar-Hillel et représentations hiérarchiques dynamiques chez Hockett et Harris, sont ancrées dans la première mathématisation du langage. En intégrant la possibilité d'automatisation, une nouvelle étape de cette mathématisation va s'accomplir, celle de l'automatisation-mathématisation de la grammaire. La linguistique computationnelle et la grammaire générative deviennent alors les horizons de projection des sciences du langage aux États-Unis¹¹.

2. LES PREMIÈRES EXPÉRIENCES DE TA EN GRANDE-BRETAGNE ET EN URSS: UNE INTÉGRATION SUR LE TEMPS LONG

Les premières expériences de TA en Grande-Bretagne et en URSS présentent plusieurs similitudes. Acteurs majeurs de la guerre froide, les Britanniques et les Russes sont en effet les premiers, dans le sillage des États-Unis et de la première démonstration sur ordinateur de 1954, à s'engager dans des expériences de TA. Même si les groupes russes et britanniques ont mené des expérimentations en TA de façon parallèle et souvent dans la méconnaissance totale des travaux des uns et des autres¹², ils ont comme point commun d'avoir entrepris ces expérimentations en les inscrivant dans leurs horizons de rétrospection respectifs, et en arrimant les horizons de projection sur le temps long. En donnant la priorité au transfert de sens dans le processus de traduction et non à la logique mathématique et à la syntaxe comme les Américains, ils ont été amenés à développer des méthodes pour la TA de langues intermédiaires à base sémantique. Ce qui différencie ces deux modèles, outre le contexte politique, c'est la familiarité avec la première mathématisation. Avec les travaux de Bertrand Russell, Alan Turing, Ludwig Wittgenstein (entre autres), les Britanniques sont très au fait de la première mathématisation qui composait l'essentiel de l'horizon de rétrospection de Weaver et de la TA. Ils ont leur propre école de théorie de l'information avec E. Colin Cherry, Dennis Gabor et David McKay. Pour les Soviétiques, c'est au travers des débats sur la cybernétique que ces questions sont abordées. Même si Norbert Wiener (1894-1964)¹³ reconnaît les travaux pionniers de certains mathématiciens russes – il cite Andrej Kolmogorov (1903-1987), Nikolaj Krylov

¹¹ Outre la TA, il faut prendre en considération l'interaction entre théorie de l'information et phonétique acoustique, également issue des sciences de la guerre, dans la construction de nouveaux horizons de rétrospection et de projection pour les sciences du langage. Ainsi, les travaux de Roman Jakobson, Gunnar Fant, Morris Halle et E. Colin Cherry sur la théorie des traits distinctifs ont joué un rôle essentiel dans la mathématisation des sciences du langage de la seconde moitié du XX^{ème} siècle (Léon [à paraître (a)]).

¹² On notera que les Russes connaissaient les travaux du CLRU (Mel'čuk 1961; Žolkovskij 1961), l'inverse n'étant pas vrai.

¹³ Wiener 1948.

(1879-1955) et Nikolaj Bogoljubov (1909-1992)¹⁴ – il est difficile de cerner actuellement l'ancrage des Russes dans la première mathématisation du langage¹⁵.

2.1. LES PIONNIERS:

L'ESPÉRANTO COMME LANGUE INTERMÉDIAIRE

2.1.1. LA MACHINE DE SMIRNOV-TROJANSKIJ (1931, 1947)

Étrangement, les tout premiers pionniers de la machine à traduire sont d'une part un Russe, d'autre part deux Britanniques dont aucun n'est linguiste. Ce sont des inventeurs qui, bien qu'à des époques différentes, se sont appuyés sur l'espéranto comme langue intermédiaire pour imaginer une machine à traduire.

Le projet de machine à traduire conçu par l'ingénieur Petr Smirnov-Trojanskij (1894-1950) en 1931, et rédigé en 1947 sous le titre *De la machine à traduire, construite sur la base d'une méthode monolingue*, a été publié en 1959 par les pionniers de la TA en URSS, Izabella Bel'skaja, Dmitrij Panov et Lev Korolev. En reprenant ce projet des années 1930-1940, l'objectif des chercheurs soviétiques était de tenter d'instaurer les Russes comme pionniers de la TA, avant les Américains, à l'instar du sputnik lancé en 1957.

Le tableau ci-dessous (p. 56) illustre la méthode de traduction élaborée par Smirnov-Trojanskij¹⁶:

¹⁴ «Let it be remarked parenthetically that some of my speculations in this direction attach themselves to the earlier work of Kolmogoroff in Russia, although a considerable part of my work was done before my attention was called to the work of the Russian school. [Kolmogoroff A.N. 1941 "Interpolation und Extrapolation von stationären zufälligen Folgen" Bull. Ac. Sciences USSR ser. maths 5 3-14]» (*ibid.*, p. 11).

¹⁵ Ce point, abordé dans Mindell, Segal, Gerovitch 2003, mériterait d'être approfondi.

¹⁶ Archambault, Léon 1997, p. 115.

A	A1	(B1 + C1)		(B + C)	
Das Bild	Das Bild - o	картина - o	Le tableau - o	картина	Le tableau
der Welt	die Welt de - o	мир de - o	le monde de - o	мира	du monde
zeigt,	zeigen - as	показывать - as	montrer - as	показывает,	montre
wie	wie	как	comment	как	comment
die Materie	die Materie - o	материя - o	la matière - o	материя	la matière
sich bewegt	sich bewegen - as	двигаться - as	se mouvoir - as	движется,	se meut
wie	wie	как	comment	как	comment
die Materie	die Materie - o	материя - o	la matière - o	материя	la matière
denkt	denken - as	мыслить - as	penser - as	мыслит	pense

Comme base d'automatisation de la traduction, Smirnov-Trojanskij postule une structure commune à toutes les langues, au delà de leur diversité; entre la langue-source et la langue-cible, il suggère de passer par un texte intermédiaire, qu'il dit *logique*, qui consiste en une suite de mots apparaissant sous leur forme grammaticale de base, comme autant d'entrées de dictionnaire qui seront ensuite traduites. Dans le système de Smirnov-Trojanskij, un traducteur obtient, à partir d'un texte en langue source A, un texte en langage *logique* A₁ que la machine traduira automatiquement en langage *logique* B₁, grâce à un dictionnaire qui traduit automatiquement les entrées. Le texte B₁ sera ensuite arrangé manuellement par un rédacteur en langue cible B.

Certaines formes de ce langage logique: forme + marque de reconnaissance du verbe (as) et du nom (o) sont des emprunts à l'espéranto.

Pour Smirnov-Trojanskij, le grand intérêt de cette méthode de traduction réside dans le fait qu'il s'agit d'une méthode *monolingue*, chaque traducteur pouvant ne connaître qu'une langue et une seule: il met ainsi le travail de traduction à la portée de tous les citoyens, puisque «ce travail ressemble aux exercices des écoliers du cours élémentaire qui décomposent un texte de leur langue natale en parties de propositions et en parties de discours»¹⁷. En outre, les avantages économiques de la méthode sont nombreux, notamment la possibilité de faire appel à des traducteurs non qualifiés.

¹⁷ Cf. in Panov (éd.), 1959, p. 8.

2.1.2. LE *PIDGIN ENGLISH* DE RICHENS ET BOOTH (1948)

Pour Andrew D. Booth (1918-2009) et Richard H. Richens (1919-1984), une langue est une suite de symboles représentant des idées. La traduction est une opération de substitution d'une langue à une autre pour exprimer le même ensemble d'idées. La procédure consiste à décomposer les mots du texte source en unités sémantiques (racines et affixes séparés par des *), chacune représentant une idée (1), puis à opérer une traduction mot à mot en appliquant un dictionnaire bilingue. Le résultat est une suite de mots et de directives grammaticales ressemblant à une forme stéréotypée de *Pidgin English* (2) contenant des directives grammaticales issues de l'analyse de la langue source (3). L'astérisque signale le lieu d'une décomposition morphologique, résultat d'une post-édition. Celle-ci doit être réécrite en anglais standard (4).

(1) texte français (source)

Il n'est pas étonnant de constater que les hormone*s de croissance ag*issent sur certain*es espèce*s, alors qu'elles sont in*opér*antes sur d'autre*s, si l'on song*e à la grand*e spécificité de ces substance*s.

(2) résultat: *Pidgin English*

v not is not / step astonish v of establish v that / which ? v hormone m of growth act m on certain m species m, then that / which ? v not operate m on of other m if v one dream / consider z to v great v specificity of those substance m.

(3) directives grammaticales utilisées dans l'exemple

m: pluriel (multiple, pluriel ou duel)	v vide (dénué de signification) z non spécifié
--	---

(4) résultat du post-editing

It is not surprising to learn that growth hormones may act on certain species while having no effect on others, when one remembers the narrow specificity of these substances¹⁸.

Dans le cas russe, comme dans le cas britannique, ces approches, qui peuvent paraître naïves, témoignent de l'importance d'un horizon de rétrospection commun qui peut ici se confondre avec une tradition, dont font partie, dans les années 1930-1950, les langues internationales, en particulier l'espéranto. Mais les similitudes s'arrêtent là. Les modèles de langues intermédiaires qui en sont issus sont très différents. Les contextes, eux aussi, sont différents. Smirnov-Trojanskij conçoit sa machine à traduire bien avant l'apparition des ordinateurs dans un contexte de planification

¹⁸ Richens, Booth 1955, p. 36-37.

multilingue, dont la figure de proue est Ernest Drezen dans les années 1930¹⁹. L'objectif est de traduire toutes les langues de l'Union.

Le début de la TA en Grande-Bretagne commence avec la rencontre de Booth, directeur du centre informatique du Birbeck College de l'Université de Londres avec Weaver. Celui-ci promet de doter l'Université d'un ordinateur électronique à la condition que des expérimentations non numériques, en particulier de TA, y soient effectuées. Comme signalé plus haut, ni Booth ni Richens n'est linguiste, ni même mathématicien. Le premier est cristallographe, le second est biologiste, spécialiste de la génétique des plantes. Ce sont des «inventeurs» sans aucun ancrage ni dans les sciences de la guerre, ni dans la première mathématisation du langage.

2.2. LEXIQUE, CONTEXTE ET THÉSAURUS: LES EXPÉRIENCES DU CAMBRIDGE LANGUAGE RESEARCH UNIT²⁰

Le *Cambridge Language Research Unit* (CLRU) ne comporte pas d'ingénieurs, et n'est pas soumis, comme les Américains, à la pression stratégique-économique de fournir des traductions en série économiquement rentables. Il est dirigé par une philosophe, Margaret Masterman (1910-1986), élève de Wittgenstein, et comprend des linguistes (Martin Kay et Michael Alexander Kirkwood Halliday) et des mathématiciens (le statisticien Arthur Frederick Parker-Rhodes, les informaticiens Yorick Wilks et Karen Spark-Jones). La présence en son sein de Richens, qui, comme on l'a vu, a commencé des expérimentations de TA dès 1948, contribue probablement au fait que le groupe ait été reconnu très tôt par les Américains. En particulier, Richens signe avec Booth un chapitre du premier ouvrage collectif de TA publié au MIT²¹. Le CLRU participe dès 1956 à un colloque organisé par le MIT et, la même année, reçoit un financement de la NSF.

Malgré leur proximité avec les Américains, les membres du CLRU n'adoptent pas une approche logico-syntaxique. La langue intermédiaire à base sémantique qu'ils élaborent se situe à la croisée de deux courants appartenant à la tradition britannique, et constitutifs de l'horizon de rétrospection de ces premières expérimentations: les langues universelles du XVII^{ème} siècle et les théories contextuelles du sens des années 1930 (Bronisław Kasper Malinowski et le second Wittgenstein). On peut distinguer trois projets successifs de langues intermédiaires développés par les membres du CLRU entre 1955 et 1968: *Nude*, *Nude II*, et les *templates*.

¹⁹ Archaimbault 2001.

²⁰ Pour un exposé détaillé de ces travaux, cf. Léon 2004 et 2007.

²¹ Locke, Booth (eds.), 1955.

2.2.1. LA LANGUE INTERMÉDIAIRE NUDE

Pour élaborer sa langue intermédiaire, Richens²² s'inspire des projets de langues universelles britanniques du XVII^{ème} siècle, notamment l'*Ars Signorum* de George Dalgarno (1661) et l'*Essay towards a Real Character and a Philosophical Language* de John Wilkins (1668). Il faut préciser qu'un des descendants de Wilkins est le botaniste Carl von Linné dont Richens, spécialiste des plantes, connaissait probablement les travaux.

Cette langue est conçue comme un réseau sémantique de cinquante «idées nues» [*naked ideas*] où les particularités structurales (et surtout morphologiques) de la langue source sont supprimées, d'où son nom de *Nude*. Le réseau sémantique, dit Richens, est ce qui est invariant durant la traduction²³. Ces idées nues sont structurées en *chunks* (ou unités sémantiques). Ce sont des formules composées de primitives sémantiques et d'opérateurs syntaxiques.

Par exemple²⁴:

.Pz	one
Xp'CL	seed
Xp'CL.Pz	one seed

Six primitives sont utilisées dans cet exemple²⁵:

p	plant	X	part, component
z	negation, opposite, contrary	C	causation, influence
P	plurality, group, number	L	living, alive

ainsi que deux opérateurs syntaxiques:

Le point [.] marque une relation monadique: un adjectif ou un verbe intransitif sont marqués par l'attente d'une relation avec un nom²⁶.
Les apostrophes ['] sont utilisées comme des parenthèses à l'intérieur d'un mot.

L'interlingua «algébrique» de Richens, conçue comme un réseau de primitives sémantiques, est très proche d'une caractéristique universelle. Comme Wilkins, Richens pose le problème de la compositionnalité des particules verbales de l'anglais²⁷. Le sens d'un mot est une composition d'éléments de sens:

²² Richens 1955.

²³ «The semantic net thus represents what is invariant during translation» (Richens 1956, p. 24).

²⁴ Richens 1955, p. 4-5.

²⁵ *Ibid.*, p. 9.

²⁶ L'opérateur [.] marque une relation dyadique, par exemple un verbe transitif marqué par l'attente d'un sujet et d'un objet. *Nude* a une syntaxe qui a pu être considérée comme une préfiguration de la grammaire des cas: un verbe transitif est marqué pour un sujet et un objet.

²⁷ Cram 1994.

seed (graine) = Xp'CL [X partie + p plante / C cause L vie]

Toutefois, aucune procédure automatique d'extraction des réseaux sémantiques à partir des textes de base n'est envisageable à partir de *Nude*, et celle-ci ne sera pas réellement exploitée pour la TA.

2.2.2. NUDE II: UNE LANGUE INTERMÉDIAIRE PAR THÉSAURUS

À partir de *Nude*, Masterman va élaborer un second projet de langue intermédiaire qui s'inspire tout à la fois de *Nude*, de la notion de *thesaurus* et des théories wittgensteiniennes sur la définition du sens d'un mot par ses contextes d'usage. Si contradictoire que puisse paraître cette tentative d'allier une conception du sens *a priori* – que comporte tout projet de langue universelle – avec une conception du sens par l'usage, elle eut des retombées intéressantes.

Une des composantes de la nouvelle langue intermédiaire est le *Thesaurus*, paru en 1852, de Peter Mark Roget, considéré à raison par les historiens de la linguistique comme un des continuateurs de Wilkins. Masterman remplace les cinquante primitives de *Nude* par les *heads* du *Thesaurus* et améliore la syntaxe.

Les primitives sont des mots de l'anglais. La syntaxe de *Nude II* consiste en deux connecteurs et une convention de parenthésage²⁸.

[:] relie deux éléments dont la relation est celle d'ajout à l'élément principal;

[/] est un connecteur verbal non commutatif représentant la relation de sujet à verbe ou de verbe à objet;

de vraies parenthèses remplacent les apostrophes de Richens, et regroupent les primitives par deux.

Les primitives de *Nude II* utilisées dans l'exemple sont les suivantes:

	NUDE ELEMENT	APPROXIMATING AREA OF MEANING	EXAMPLE
6	CAUSE	Causative actions	Cause / (have / sign) (say)
13	HAVE	Pertain «of»	Cause / (nohave / life) (kill)
33	MAN	Human Kingdom	(part: folk): man (member of family)
41	SIGN	Symbol (any sort)	Cause / (have / sign) (speak)

Speak
cause / (have / sign)

he says
man / (cause / (have / sign))

Speaker
man: (cause / (have / sign))

²⁸ L'exemple suivant est extrait de Masterman 1959, p. 62-63.

L'organisation des *heads* forme un réseau sémantique de structure non hiérarchique (qui a donné lieu à une formalisation en treillis). Les *heads* doivent avoir un sens sans pour autant être des mots appartenant à une langue donnée. Masterman, en tant qu'élève de Wittgenstein, partage sa conception du sens des mots par l'usage et refuse d'envisager les primitives de *Nude II* comme des concepts universaux *a priori*, ce qu'ils étaient dans la version de Richens. De même, elle est absolument hostile à toute hypothèse cognitive considérant les primitives comme des éléments d'un langage de la pensée, comme par exemple le *Mentalese* de Jerry Alan Fodor²⁹. Pour Masterman, c'est le thésaurus qui permet de faire le lien entre les deux conceptions contradictoires du sens (langue universelle / usage). La méthode de langue intermédiaire par thésaurus est, selon Masterman, directement inspirée des intuitions de Wittgenstein³⁰, en ce que le sens d'un mot est défini à partir de ses contextes d'usage. L'hypothèse fondamentale qui étaye la faisabilité d'un thésaurus (et donc d'une langue intermédiaire par thésaurus) est que, bien que l'ensemble des usages possibles des mots dans une langue soit infini, le nombre de situations extralinguistiques primaires nécessaires pour communiquer est fini. Cette hypothèse a des conséquences importantes pour la TA. La traduction, comme la communication, n'est possible que si les deux populations et les deux cultures correspondant aux langues envisagées dans la traduction partagent un stock commun de contextes extralinguistiques même si elles sont très différentes. Cette conception implique que la traduction entre deux langues, c'est également la traduction entre deux cultures.

2.2.3. LES TEMPLATES DE WILKS

Le troisième projet est développé par un des plus jeunes membres du CLRU et élève de Masterman, Yorik Wilks (né en 1939), qui adapte le langage *Nude* à un projet de résolution d'ambiguïtés sémantiques dans un texte³¹. La notion d'*ambiguïté* admise par tout expérimenteur de TA – puis plus tard de traitement automatique des langues en général – consiste à assimiler mot ambigu sémantiquement et mot polysémique, autrement dit à considérer que tous les sens possibles de ce mot sont les sens attribués par le dictionnaire. Wilks, au contraire, dans la lignée de Wittgenstein, de John Rupert Firth et de Masterman, définit les ambiguïtés sémantiques par rapport à leur contexte d'usage, en l'occurrence un texte. Pour résoudre ces ambiguïtés, il met au point un système de représentation sémantique des textes à l'aide de *templates* qui doivent capturer l'essentiel de l'information dans les textes. Les *templates*, sortes de patrons du sens d'une proposition ou d'une phrase simple, sont fabriqués à l'aide de for-

²⁹ Fodor 1975.

³⁰ Wittgenstein 1953.

³¹ Wilks 1968.

mules très proches des formules de *Nude*. Ces formules permettent d'encoder les différents sens d'un mot à l'aide de 53 primitives, dont 45 sont communes avec les *archiheads* de la méthode par thésaurus. Grâce à une procédure de *pattern matching*, les différentes représentations des mots vont être confrontées à celles du texte, et comme pour le thésaurus, si deux représentations de deux mots contiennent les mêmes primitives, elles permettront de résoudre l'ambiguïté, ou du moins de proposer un sens préféré (Wilks parlera plus tard de sémantique préférentielle).

Les travaux de Wilks ont le mérite d'introduire les études sur les primitives sémantiques, conçues dans le cadre de la TA, dans le débat tout nouveau à la fin des années 1960 de la sémantique non référentielle et de l'intelligence artificielle. Ils constituent pour une part l'horizon de projection des premières tentatives de TA, ouvrant sur des développements internationalisés de la linguistique, tout en gardant une spécificité britannique héritée de Henry Sweet, Wittgenstein et Firth d'un travail sur les textes plutôt que sur des phrases isolées³².

2.3. CYBERNÉTIQUE ET TA EN UNION SOVIÉTIQUE DANS LES ANNÉES 1950

Alors qu'aux États-Unis, la TA est clairement un héritage de la cybernétique et de la théorie de l'information, en URSS, leur relation n'est pas aussi immédiate mais tout aussi certaine. Il faut préciser que la cybernétique a connu une fortune contrastée en URSS³³. L'ouvrage de Wiener, *Cybernetics* paru en 1948, fait l'objet en URSS, dès 1952, d'une campagne hostile, la cybernétique étant considérée comme une science bourgeoise. Après la mort de Staline, on peut noter un changement d'attitude de la communauté scientifique soviétique à l'égard de la cybernétique. En novembre 1954, l'exposé «Qu'est-ce que la cybernétique?» d'Arnold Kolman devant l'Académie des sciences sociales du Comité central du Parti communiste, sera publié en RDA puis dans *La Pensée* et dans *Behavioral Science*. Ce changement peut aussi être imputé au fait que, grâce à ses écrits critiques sur la société américaine, Wiener n'est plus considéré comme un «scientifique impérialiste»; notamment, dans la deuxième édition de son ouvrage *The Human Use of Human Being*³⁴, il s'en prend au maccarthysme et milite contre l'utilisation de l'arme atomique. Enfin, la cybernétique et la rétroaction, à l'œuvre dans les systèmes de défense anti-aérienne mis au point pendant la Seconde guerre mondiale, est devenue

³² Les travaux fonctionnalistes de Halliday, qui faisait partie du CLRU, sont aussi issus des travaux sur le thésaurus. Pour Halliday, la théorisation des applications demeure intrinsèquement liée à la linguistique théorique. Cette position, manifeste dans son engagement dans la TA, l'enseignement des langues et plus tard la *Corpus Linguistics*, se trouve en complète continuité avec l'empirisme britannique hérité de Sweet et de Firth.

³³ Mindell, Segal, Gerovitch 2003.

³⁴ Wiener 1950.

nécessaire pour les militaires soviétiques, et un rapport confidentiel rédigé par le vice-amiral russe Aksel' Berg semble jouer un rôle décisif. En 1958, l'ouvrage de Wiener est traduit en russe. À partir de 1961, la cybernétique n'est plus une science bourgeoise mais une théorie à promouvoir. Ainsi Nikita Khrouchtchev et Walter Ulbricht s'y réfèrent dans les congrès tenus respectivement par leurs partis.

L'engagement des Soviétiques dans la TA est contemporain du retour en grâce de la cybernétique. C'est en effet à partir de 1954, date de la présentation de Kolman, et de la première démonstration de TA sur ordinateur à New York, que les premiers travaux en TA sont envisagés. En 1956, une résolution du XX^{ème} Congrès du PCUS promeut de façon explicite la TA comme projet scientifique prioritaire. C'est en 1958, date de la traduction de l'ouvrage de Wiener, que sont créées deux revues où seront publiés des articles sur la TA, *Problemy kibernetiki* [Problèmes de cybernétique] et *Mašinnyj perevod i prikladnaja lingvistika* [Traduction automatique et linguistique appliquée].

Toutefois, contrairement aux États-Unis, la TA n'occupe pas, en URSS, une place stratégique dans la guerre froide. Il y a peu de financement et les ordinateurs sont réservés aux services secrets ou à des fins strictement militaires. Les chercheurs qui sont engagés dans la TA appartiennent à des disciplines spéculatives comme les mathématiques ou la linguistique mathématique, considérées comme peu «dangereuses» idéologiquement. À ce titre, ils bénéficient d'une certaine tranquillité, et leur situation, toutes choses égales par ailleurs, est comparable à celle des Britanniques, eux aussi peu soumis aux pressions économiques et stratégiques.

Ainsi, la demande des autorités est moins de traduire l'anglais vers le russe que de traduire le russe dans les autres langues de l'Union. La TA est conçue comme la première étape d'un programme plus général de traitement de l'information, de son extraction et de son stockage à partir de textes écrits dans différentes langues, tel qu'éditer, faire des résumés, fournir des références bibliographiques, etc.

Dans leurs premiers travaux, les chercheurs russes se positionnent vis-à-vis des travaux américains, pionniers dans le domaine et référence obligatoire en ces temps de guerre froide et de compétition³⁵. Ils critiquent la conception de la traduction héritée de la théorie de l'information et préconisée par Weaver, de traiter les langues comme des codes et la traduction comme un décodage. Au contraire, la traduction se définissant comme un transfert de sens d'une langue à l'autre, ne peut se traiter comme un changement de code, c'est-à-dire de formes uniquement. Cette

³⁵ Les travaux respectifs sont relativement connus des deux côtés du rideau de fer. Un des pionniers de la TA, Panov, assiste à la première traduction sur ordinateur à New York. Les travaux russes en TA ou en documentation automatique sont traduits systématiquement par un service américain, le *Joint Publication Research Service (JPRS)* dès 1956.

position les conduit à développer une approche sémantique de la TA, au détriment de l'analyse syntaxique, dominante aux États-Unis, et à donner plus d'importance à la synthèse en langue cible, d'où l'élaboration de langues intermédiaires.

2.4. MODÈLES DE LANGUE INTERMÉDIAIRE³⁶

Plusieurs modèles ont servi pour l'élaboration de langues intermédiaires: le russe, les langues internationales, l'idée de protolangue, tous inscrits sur le long terme dans la tradition culturelle ou linguistique russe. D. Panov³⁷ écarte les langues artificielles, telles que les langues internationales qu'il tient pour incapables de transmettre la richesse des langues naturelles. Il écarte également l'idée de construire une langue spécifique adaptée à la traduction et propose le russe comme langue intermédiaire. Les arguments en faveur du russe sont économiques et linguistiques, le russe étant présenté comme comportant des propriétés spécifiques comme la non-ambiguïté morphologique du verbe permettant d'identifier exactement la forme d'un verbe même hors contexte.

Nikolaj Andreev (1920-1997) dirige à l'Université de Leningrad le Laboratoire expérimental de traduction par machine rattaché à l'Institut de mathématiques et de mécanique. Son modèle de langue intermédiaire est issu de l'idée de langue univoque d'information promue par Drezen³⁸. Sa méthode de construction repose sur un traitement statistique des invariants linguistiques et sur une conception des langues comme faits sociaux. Ainsi, ne font partie de la langue intermédiaire que les structures suffisamment distribuées dans les langues naturelles examinées. La présence ou absence de ces structures se trouve de plus pondérée par le nombre de locuteurs et le nombre de textes déjà traduits. Cette méthode fondée sur des données statistiques (nombre de structures communes, nombre de locuteurs, nombre de textes traduits) rappelle les méthodes qui, au XIX^{ème} siècle, présidaient au choix d'une langue internationale *a posteriori* ainsi qu'une tendance bien connue en URSS à la planification chiffrée. Par ailleurs, au travers du terme de *hiéroglyphes* par lequel Andreev nomme les éléments de base de sa langue intermédiaire, on reconnaît un des vieux mythes de construction d'une langue parfaite des images, de même que certains projets de pasigraphie du XIX^{ème} siècle. Andreev propose de calculer pour chaque langue un indice de congruence à partir du nombre de hiéroglyphes concordants (le hiéroglyphe de la langue d'entrée est le même que celui de la langue de sortie) et non concordants qu'elle comporte, et de la pondération qui lui est affectée.

Soit la phrase anglaise suivante et sa traduction en hiéroglyphes:

³⁶ Cf. Archaimbault, Léon 1997.

³⁷ Panov 1956.

³⁸ Andreev 1957.

<i>the</i>	<i>sun</i>	<i>is</i>	<i>a</i>	<i>yellow</i>	<i>star</i>
φ5τ4	δ1	φ1 φ8 τ5	φ6 τ4	δ2τ2	δ3 τ1

Les hiéroglyphes sémantiques δ correspondent aux unités lexicales: par exemple δ1 est le hiéroglyphe sémantique de *sun*.

Les hiéroglyphes formels φ correspondent aux informations grammaticales: φ5 est le hiéroglyphe formel de l'article défini.

Les hiéroglyphes tectoniques traduisent l'ordre des mots: φ4 est le hiéroglyphe tectonique indiquant que l'article est suivi de N³⁹.

Igor' Mel'čuk (né en 1932), hispanisant et francisant, travaille à l'Institut de linguistique de l'Académie des Sciences de l'URSS. Il développe en 1956 avec la mathématicienne Ol'ga Kulagina (1932-2005) un système de TA français-russe, avant de s'intéresser à un algorithme de traduction russe-hongrois pour lequel il concevra l'idée d'une langue intermédiaire⁴⁰. Le hongrois présente une particularité dans le processus de traduction: alors que dans la traduction de langues comme le russe, l'anglais, le français ou l'allemand, il est possible de considérer la différence d'ordre des mots comme négligeable, ce n'est pas possible dans la traduction hongrois-russe. La solution consiste soit à concevoir une procédure spéciale qui rende compte de ce changement pour chaque paire de langues, soit à envisager une méthode dissociant analyse et synthèse de façon à neutraliser le problème de transfert de l'ordre des mots. Cette méthode, c'est la langue intermédiaire.

Les travaux de Mel'čuk sont inspirés par ceux de Vjačeslav Ivanov (né en 1929), qui fut son directeur de thèse. Ivanov propose de réinvestir dans la réflexion sur la langue intermédiaire le travail théorique de la linguistique historique et comparée. Ce travail s'inscrit dans l'héritage direct de Jan Baudouin de Courtenay (1845-1929) par les références explicites que fait Ivanov à la théorie de la mixité de toutes les langues. L'idée en est que les méthodes de comparaison des faits de langue sont largement éprouvées, mais qu'une métathéorie capable de décrire la structure formelle de la comparaison, une *théorie des relations des systèmes linguistiques* reste à construire. Celle-ci s'appuierait sur une langue abstraite, nommée protolangue, qui ne serait pas représentative d'un état antérieur d'une langue donnée, mais qui en constituerait le système abstrait, au sens où celle-ci recèlerait toutes les potentialités communes à un groupe de langues, que ces potentialités soient exprimées ou non. C'est cette protolangue que propose Ivanov comme langue intermédiaire pour la TA.

³⁹ Andreev 1957, p. 62.

⁴⁰ Kulagina, Mel'čuk 1967.

Mel'čuk conçoit la traduction automatique comme le transfert de sens d'une langue à une autre en s'appuyant sur les invariants linguistiques. Il s'agit alors de savoir comment traiter les traits variants dans la langue intermédiaire. C'est en termes de significations grammaticales (opposées aux significations lexicales), empruntées à Franz Boas via un article de 1959 de Jakobson «Boas' View of Grammatical Meaning» que Mel'čuk traite les variations. Toutefois, il ne considère pas les significations grammaticales comme des universaux (il utilisera le terme de *flexions* dans son *Cours de morphologie générale*), et exploite le fait que celles-ci ne sont pas identiques au regard de la traduction. Il va les utiliser pour un traitement distinct des langues sources et cibles. Une bonne partie des significations grammaticales (informations morpho-syntaxiques), singulières pour chaque langue et utiles lors de l'analyse du texte en langue source, ne sont plus nécessaires lors de la phase de traduction proprement dite. Ainsi il n'est pas nécessaire de connaître directement le genre, le nombre, le cas d'un adjectif russe quand on le traduit en français ou en allemand. Ce qui est important c'est d'avoir trouvé, lors de la phase d'analyse, le substantif avec lequel il s'accorde. Il en va autrement des indicateurs grammaticaux de significations lexicales, servant au repérage référentiel des éléments lexicaux (nombre des noms, temps et mode des verbes, etc.) qui, eux, sont nécessaires à la traduction.

La langue intermédiaire ainsi construite ne peut être une langue ni entièrement artificielle ni entièrement naturelle. Elle est fondée sur l'extraction de propriétés de langues. Pour la TA, Mel'čuk en vient à privilégier la synthèse. Commencer par l'analyse, dit-il, est trop difficile, dans la mesure où, le sens étant un élément fondamental de la traduction, l'analyse suppose que l'on maîtrise la compréhension d'un texte. Ce renversement sera un des traits caractéristiques du modèle qu'il développera dans les années 1970, avec Aleksandr Žolkovskij et Jurij Apresjan à partir de ses travaux sur la langue intermédiaire: le modèle sens → texte de préférence au modèle texte → sens⁴¹. Le modèle sens-texte est ainsi la partie la plus visible de l'horizon de projection des travaux de TA des années 1950 en Union soviétique. Par ailleurs Mel'čuk se considère comme un typologue et sa conception des significations grammaticales non universelles est investie dans son *Cours de morphologie générale*⁴².

⁴¹ Mel'čuk, Žolkovskij 1971; Apresjan, Žolkovskij, Mel'čuk 1973.

⁴² Mel'čuk 1993-2000.

CONCLUSION

Les deux cas examinés, Grande-Bretagne et Russie, d'un mode d'intégration de la TA dans les sciences du langage s'inscrivent dans la continuité de la tradition culturelle sur le long terme. Continuité ne veut pas dire linéarité. Dans le cas britannique, il s'agit d'aller puiser dans la tradition des projets de langue universelle du XVII^{ème} siècle repris au XIX^{ème} siècle, aussi bien que dans les théories contextualistes du sens contemporaines des premières expériences de TA. Ont également présidé à ces premières expériences menées par des philosophes, linguistes et mathématiciens, l'attachement britannique à l'interaction entre théorie et visées pratiques. Dans le cas de l'URSS, la planification linguistique est aussi ancrée dans le long terme, une forte tradition de planification linguistique était apparue dès le XVII^{ème} siècle de Pierre le Grand, ravivée à la fin du XIX^{ème} siècle et dans les années 1930, avec les projets de langues internationales. La tradition de linguistique comparée, de Baudouin de Courtenay à Ivanov, a été reprise par Mel'čuk dans son modèle sens-texte. Dans l'horizon de rétrospection des pionniers de la TA en URSS, il faut aussi inclure une conception des langues comme systèmes complexes et subtils permettant d'exprimer des pensées, ancrés dans l'histoire et les pratiques sociales. Dans les deux traditions, traduction entre langues implique traduction entre cultures. La possibilité d'automatisation a servi de catalyseur et a fait naître de nouveaux modèles pour les sciences du langage qui, pour les traditions concernées⁴³, se sont inscrits dans le long terme.

© Jacqueline Léon

⁴³ Pour les Français, au contraire, l'expérience est brutale puisque rien dans l'horizon de rétrospection des sciences du langage en particulier et des sciences en général ne les préparait à l'intégration de la TA. Cela va nécessiter des opérateurs, personnalités et institutions de tous ordres, pour assurer le passage et l'intégration de l'automatisation-mathématisation dans les sciences du langage (cf. Léon 2010).

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANDREEV Nikolaj Dmitrievič, 1957: «Mašinnyj perevod i problema jazyka-posrednika», in *Voprosy jazykoznanija*, 1957, № 5, p. 117-121 [Traduction automatique et le problème de la langue-intermédiaire] [traduction anglaise: «Machine Translation and the Problem of an Intermediary Language», JPRS/DC-68⁴⁴, p. 58-67]
- APRESJAN Juri D., ŽOLKOVSKIJ Aleksandr K., MEL'ČUK Igor A., 1973: «Materials for an Explanatory-Combinatory Dictionary of Modern Russian» in Kiefer F. (ed.), *Trends in Soviet Theoretical Linguistics*, vol. 18. Boston: Dordrecht-Holland, p. 411-438
- ARCHAIMBAULT Sylvie, 2001: «Les approches normatives en Russie», in Koerner K., Auroux S., Niederehe H.J., Versteegh K. (eds.), *History of the Language Sciences*, vol. 18. Berlin: Walter de Gruyter, p. 901-907
- ARCHAIMBAULT Sylvie, LÉON Jacqueline, 1997: «La langue intermédiaire dans la Traduction Automatique en URSS (1954-1960). Filiations et modèles», in *Histoire. Épistémologie. Langage*, 1997, vol. 19, № 2, p. 105-132
- , (à paraître): «Machine Translation: Theoretical and Practical Shift within American and Russian Linguistics», in *Proceedings of the International Conference «Shifting Paradigms: How Translation Transforms the Humanities» (University of Illinois at Urbana Champaign, October 2010)*
- AUROUX Sylvain, 1987: «Histoire des sciences et entropie des systèmes scientifiques. Les horizons de retrospection», in Schmitter P. (éd.), *Zur Theorie und Methode der Geschichtsschreibung der Linguistik*. Tübingen: Narr, p. 20-42
- , 1995: «L'histoire des sciences du langage et le paradoxe historiographique», in *Le Gré des langues*, 1995, № 8, p. 40-63
- BAR-HILLEL Yehoshua, 1960: «The Present Status of Automatic Translation of Languages», in Alt F.C. (ed.), *Advances in Computers*, vol. 1. New York – London: Academic Press, p. 91-141
- BOOTH Andrew D. (ed.), 1967: *Machine Translation*. Amsterdam: North Holland Publishing Company

⁴⁴ Plusieurs articles parus en russe dans des revues russes, ont été traduits en anglais par un centre de traduction américain des années 1960 le JPRS (*Joint Publications Research Service*) (cf. aussi la note 35). Certains de ces documents traduits se trouvent dans les archives sur l'histoire de la traduction automatique et du traitement automatique du langage (1954-1975), CNRS, Université Paris Diderot.

- CRAM David, 1994: «Collection and Classification: Universal Language Schemes and the Development of Seventeenth Century Lexicography», in Blaicher G., Glaser B. (éds), *Anglistentag 1993 Eichstätt Proceedings*. Tübingen: Max Niemeyer Verlag, p. 59-69
- DAHAN Amy, PESTRE Dominique (éds), 2004: *Les sciences pour la guerre (1940-1960)*. Paris: Éditions de l'ÉHÉSS
- FODOR Jerry Alan, 1975: *The Language of Thought*. New York: Thomas Y. Crowell Company, Inc.
- KULAGINA Olga S., MEL'ČUK Igor A., 1967: «Automatic Translation: Some Theoretical Aspects and the Design of a Translation System», in Booth (ed.), 1967, p. 137-173
- LÉON Jacqueline, 1999: «La mécanisation du dictionnaire dans les premières expériences de traduction automatique (1948-1960)», in Cram D., Linn A., Nowak E. (eds.), *History of Linguistics 1996*, vol. II. Amsterdam: John Benjamins Publishing Company, p. 331-340
- , 2004: «*The inkstand was in the pen* and Other Stories. The Controversy between Bar-Hillel and the Cambridge Language Research Unit about Language Formalization and Machine Translation», in *The Bulletin of the Henry Sweet Society for the History of Linguistic Ideas*, 2004, № 42, p. 4-10
- , 2007: «From Universal Languages to Intermediary Languages in Machine Translation: The Work of the Cambridge Language Research Unit (1955-1970)», in Guimaraes E., Luz Pessoa de Barros D. (eds.), *History of Linguistics 2002*. Amsterdam – Philadelphia: John Benjamins Publishing Company, p. 123-132
- , 2010: «Automatisation-mathématisation de la linguistique en France dans les années 1960. Un cas de réception externe», in Neveu F., Muni-Toke V., Durand J., Kingler T., Mondada L., Prévost S. (éds), *Actes du 2^{ème} Congrès mondial de linguistique française, La Nouvelle-Orléans, CMLF10*. Paris: Institut de la Langue Française
- , (à paraître [a]): «La théorie des traits distinctifs de Jakobson: transferts et convergences entre mathématiques, ingénierie et linguistique», in Archaimbault S., Fournier J.-M., Raby V. (éds), *Hommage à Sylvain Auroux*
- , (à paraître [b]): «Traduction, procédures, formalisation. Le tournant de l'automatisation de la linguistique structurale américaine», in *Actes du colloque du Cerlico, mai 2011, Orléans*
- LOCKE William N., BOOTH Andrew D. (eds.), 1955: *Machine Translation of Languages, 14 Essays*. New York: MIT – John Wiley & son
- MASTERMAN Margaret, 1959: «What Is a Thesaurus?» [CLRU, ML 90 (archives sur l'histoire de la traduction automatique et du traitement automatique du langage [1954-1975], CNRS, Université Paris Diderot)]

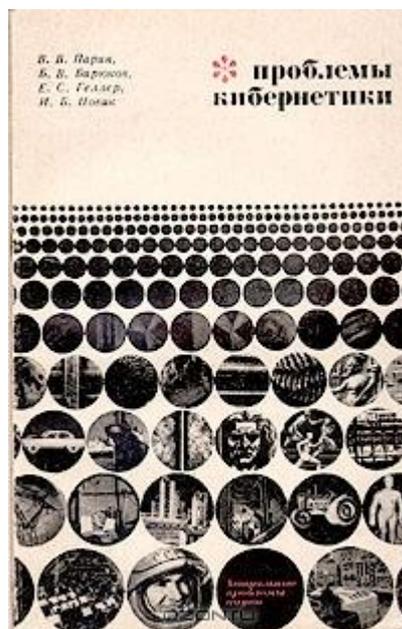
- MEL'ČUK Igor A., 1961: «Some Problems of MT Abroad», in *Reports at the Conference on Information Processing, MT and Automatic Text Reading*. [Moscow:] Academy of Sciences, Institute of Scientific Information, 1961, № 6, p. 1-44 [JPRS⁴⁵ 13135, mars 1962, p. 1-75]
- , 1993-2000: *Cours de morphologie générale (théorique et descriptive)*. Montréal – Paris: Presses de l'Université de Montréal – CNRS Éditions
- MEL'ČUK Igor A., ŽOLKOVSKIJ Aleksandr K., 1971: «Construction d'un modèle actif de la langue sens-texte», in *Documents de linguistique quantitative*, 1971, № 10: *La sémantique en URSS*, p. 11-50
- MINDELL David, SEGAL Jérôme, GEROVITCH Slava, 2003: «Cybernetics and Information Theory in the United States, France and the Soviet Union», in Walker M. (ed.), *Science and Ideology: A Comparative History*. London: Routledge, p. 66-95
- PANOV Dmitrij Jur'evič, 1956: *Avtomatičeskij perevod*. Moskva: Izdatel'stvo AN SSSR [Traduction automatique] [traduction anglaise: *Automatic Translation*. Oxford – New York: Pergamon Press Inc, 1960]
- (éd.), 1959: *Perevodnaja mašina P.P. Trojanskogo*. Moskva: Izdatel'stvo AN SSSR [CASDN № T/R 136-1059⁴⁶ (archives sur l'histoire de la traduction automatique et du traitement automatique du langage [1954-1975], CNRS, Université Paris Diderot)] [La machine à traduire de P.P. Trojanskij]
- RICHENS Richard H., 1955: «General Program for Mechanical Translation between Any Two Languages via an Algebraic Interlingua» [CLRU, ML5 (archives sur l'histoire de la traduction automatique et du traitement automatique du langage [1954-1975], CNRS, Université Paris Diderot)]
- , 1956: «Preprogramming for Mechanical Translation», in *Mechanical Translation*, 1956, vol. 3, № 1, p. 20-27
- RICHENS Richard H., BOOTH Andrew, 1955: «Some Methods of Mechanized Translation», in Locke, Booth (eds.), 1955, p. 24-46
- WEAVER Warren, 1949 [1955]: «Translation», in Locke, Booth (eds.), 1955, p. 15-23
- WIENER Norbert, 1948: *Cybernetics or Control and Communication in the Animal and the Machine*. Paris – Cambridge, MA – New York: Librairie Hermann & co – The MIT Press – John Wiley
- , 1950: *The Human Use of Human Beings*. Boston: Houghton Mifflin
- WILKS Yorick, 1968: «On Line Semantic Analysis of English Texts», in *Mechanical Translation*, 1968, vol. 11, № 3-4, p. 59-72

⁴⁵ Cf. la note 44.

⁴⁶ Ce document a été traduit en français par le Comité d'Action Scientifique de Défense Nationale (CASDN).

-
- WITTGENSTEIN Ludwig, 1953: *Philosophical Investigations*. New York: Basil Blackwell
 - ŽOLKOVSKIJ Aleksandr K., 1961: «Essays on and in MT by the Cambridge Research Unit, Cambridge, England, June 1959», in *Mašinnyj perevod i prikladnaja lingvistika*, 1961, № 5, p. 81-89 [JPRS 13761⁴⁷, mai 1962, p. 102-115]
 - [SANS AUTEUR], 1966: *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

⁴⁷ Cf. la note 44.



V.V. Parin, B.V. Birjukov, E.S. Geller, I.B. Novik *Problemy kibernetiki*.
Moskva: Znanie, 1969 [Problèmes de cybernétique]