

Théorie de l'information, information et linguistes français dans les années 1960. Un exemple de transfert entre mathématiques et sciences du langage.

Jacqueline Léon

Laboratoire d'Histoire des Théories Linguistiques
UMR7597, CNRS, Université Paris 7 Denis-Diderot
jleon@linguist.jussieu.fr

1 Introduction

Alors que l'impact de la théorie de l'information dans les diverses sciences a été étudiée de façon approfondie (voir entre autres Aurox 1990, Segal 2003), la réception de cette théorie en linguistique est loin d'avoir fait l'objet d'une étude systématique, notamment dans le domaine français. Dans notre article, nous nous intéresserons à l'importation de la théorie de l'information dans l'œuvre de trois linguistes, Emile Benveniste (1902-1976), André Martinet (1908-1999) et Jean Dubois (né en 1920) qui se sont explicitement intéressés à cette nouvelle théorie dans les années 1960, et lui ont fait une part plus ou moins grande dans leurs ouvrages de linguistique générale.

Le rôle de passeur de Roman Jakobson (1896-1982), avec la systématisation de l'analyse des traits distinctifs en phonologie et la mise en place du schéma et des fonctions de communication, est indéniable. Nous attacherons toutefois à montrer qu'il existe une filiation spécifique de la théorie de l'information dans la linguistique française des années 1960, en partie distincte de l'influence de Jakobson et empruntant d'autres itinéraires parfois complémentaires.

On a déjà montré que les statistiques de vocabulaire, au moment de l'informatisation du langage et des premiers travaux en traduction automatique, a présidé au renouvellement de la lexicologie au sein de la linguistique structurale française des années 1960, en permettant de concevoir de nouvelles unités linguistiques, les unités lexicales composées (Léon 1998, Léon 2004). Ainsi, les trois linguistes que nous allons considérer ont aussi été des acteurs importants du renouveau de la lexicologie par la création d'unités lexicales composées : Benveniste et ses synapsies, Martinet et ses synthèmes et Dubois et ses unités complexes au sein de classes d'équivalence. Dans cet article, nous montrerons que ce sont également les études statistiques appliquées aux mots dans les textes qui ont contribué à l'introduction en linguistique de la théorie de l'information.

Afin d'appréhender les enjeux pour la linguistique des années 1960 de l'attrait pour cette théorie, nous nous appuyerons sur l'examen du terme *information* tel qu'il est apparu dans les ouvrages de ces linguistes. De façon plus générale, nous nous demanderons s'il y a lieu de parler, dans ce cas précis, de transfert de concepts et de méthodes entre mathématiques et sciences du langage¹.

2 Le terme *information*

2.1 Hypothèses

Le rapport entre termes, notions et concepts est certes complexe et l'étude de la réception d'une théorie à partir de l'examen d'un mot particulier n'est possible que si certaines conditions sont remplies : il faut identifier le moment d'importation du terme, savoir si celui-ci n'a pas déjà été utilisé dans le champ disciplinaire où il est importé, en l'occurrence les sciences du langage ; il faut repérer si le terme est

importé seul ou avec un ensemble de termes (une constellation de termes) ; enfin il faut parvenir à déterminer la façon dont s'effectue la migration du terme d'une discipline à une autre.

Le terme *information* semble répondre à ces exigences:

1) Avec la parution de l'ouvrage de Shannon et Weaver *The Mathematical Theory of Communication* en 1949, on peut dater de façon précise la diffusion de la théorie de l'information vers les autres disciplines. Rappelons que c'est à partir des années 1920 qu'en physique, en statistique et en télécommunications, que commence à émerger l'information comme notion scientifique et technique dans un sens bien différent de son usage commun. Plus tard, à partir de la seconde guerre mondiale, des ingénieurs, physiciens et mathématiciens appellent *information* la grandeur qui caractérise différents modes de communication, qui s'apparente à l'entropie de la thermodynamique et qui permet de déterminer la capacité de stockage ou de traitement des ordinateurs, à condition de renoncer à toute dimension sémantique du terme *information*. Cette notion mathématique d'information reste au plus près de son sens étymologique, à savoir la mise en forme du signal. Sur le plan des télécommunications, il s'agissait pour la Bell Telephone Company d'évaluer les conditions les plus économiques qui permettent de faire passer le plus grand nombre possible de communications sur un seul fil, c'est-à-dire de transmettre des mots que les gens reconnaissent. Ces travaux trouvèrent leur aboutissement avec la publication des articles de Shannon en 1948 suivi de l'ouvrage de Shannon et Weaver en 1949.

2) Le traçage de sa réception en linguistique paraît relativement aisée : outre les titres d'articles faisant explicitement référence à la théorie et les compte-rendus parus dans les revues, on peut repérer l'apparition dans les ouvrages de linguistique de termes entièrement nouveaux ou qui se chargent d'un sens spécifique.

2) La théorie de l'information est la partie quantitative de la théorie de la communication, qui, elle-même, fait partie de la cybernétique². Parmi la constellation de termes apparus, certains sont issus de la théorie de la communication comme *code, codage, décodage, transcodage, message, communication, émetteur, récepteur, locuteur, interlocuteur, émission, réception, transmission, signal, signaux* ; d'autres appartiennent à la cybernétique comme *feedback* ou *contrôle*. Enfin, certains termes réfèrent davantage à la partie quantitative comme *information, quantité d'information, probabilités, chaînes de Markov, redondance, bruits, entropie, coût ou rendement...*

3) Parmi l'ensemble de ces termes, on peut faire l'hypothèse que certains ont un statut particulier comme *code* (Fehr 2003) ou *information* dans la mesure où ils marquent une étape particulière pour les sciences du langage. On part en effet du constat qu'*information* n'existait pas auparavant comme métaterme de la linguistique et était très peu utilisé par les linguistes dans son sens du langage ordinaire.

4) On peut s'interroger sur le statut de l'usage de ces nouveaux termes : s'agit-il de l'introduction de concepts transformant la théorie, de métatermes affectant le langage de description, ou bien d'une utilisation approximative et/ou métaphorique ?

5) De façon corrolaire, on s'intéressera donc à la stabilité de l'usage du terme *information* dans les théories linguistiques. Comment s'y articulent sens ordinaire et sens technique ? Au sein d'une même théorie, est-il stable ou bien revêt-il plusieurs sens ? S'intègre-t-il de façon permanente ou temporaire ?

2.2 Usage «ordinaire» versus usage technique du terme information

Selon le *Robert historique*, le sens usuel d'*information* renvoie depuis le début du XXe siècle soit à l'action de rassembler des renseignements ou des connaissances, soit à cet ensemble lui-même.

(1) Le mot se spécialise dans la seconde moitié du XIXe siècle désignant l'action de prendre des renseignements (*voyage d'information*, 1867). Le sens aujourd'hui le plus usuel apparaît sous la IIIe République avec le développement de la presse: il s'agit alors de l'*information* que l'on porte à la connaissance d'un public (1886, Zola); d'où au début du XXe siècle, les acceptions "ensemble des *informations*" et "action d'*informer* les publics".

C'est d'ailleurs le sens de l'unique occurrence d'*information* dans le *Cours de linguistique générale*. Celle-ci apparaît dans le chapitre sur la linguistique diachronique : « L'histoire seule peut nous renseigner... Mais dès que l'*information historique* est en défaut, il est bien difficile de déterminer ce qui est agglutination et ce qui relève de l'analogie » (Saussure, CLG :245).

Face à cet usage 'ordinaire', le sens technique d'*information* présente les caractéristiques suivantes (voir Shannon et Weaver 1949, Moreau 1964, Auroux 1990, Segal 2003):

- L'information est une mesure à caractère statistique visant à quantifier ce qu'apporte un message à un récepteur, et à analyser les facteurs qui peuvent affecter sa transmission. Le message, pour convoier de l'information, doit représenter un choix parmi les messages possibles.

- L'information fournie par un événement est donc fonction de sa probabilité (P) et se mesure selon la formule : $I = \log Px$. Plus la probabilité d'un événement est grande, moins il nous apporte d'information. Ainsi, si l'événement est certain on n'a besoin d'aucune information pour le déterminer ; il ne nous apporte aucune information. L'information se mesure en unités appelées *binary digits* ou *bits*.

- Lorsqu'il y a bruit dans le canal de transmission, le message est affecté par des distorsions et des erreurs. Le bruit augmente donc l'incertitude et devrait augmenter l'information. Mais, comme le précise Shannon, cette incertitude liée à l'augmentation des erreurs devient indésirable.

- Le bruit peut être compensé par la redondance. En envoyant l'information sous une forme redondante la probabilité des erreurs peut être réduite de sorte qu'il y a un avantage à ne pas utiliser un processus de codage qui élimine toute redondance, car la redondance restante permet de lutter contre le bruit. Ainsi, dans le domaine des télécommunications, on a intérêt à élever la redondance pour s'assurer d'une bonne transmission. Shannon recommande d'ailleurs que la redondance de la langue anglaise, de l'ordre de 50%, ne doit pas être éliminée dans la transmission télégraphique. Toutefois, plus les données sont structurées (donc redondantes), moins elles contiennent d'information au sens de Shannon, ce qui va à l'encontre du sens commun du mot 'information'.

Dans ce modèle, l'information est une notion non sémantique. C'est une quantité abstraite qui qualifie le message indépendamment de sa signification. La signification du message, son contenu, ne sont pas considérés comme un élément pertinent : ce qui est transmis, c'est une forme et non un sens.

3 Les premiers passeurs : Jakobson et Mandelbrot

En 1951, Benoît Mandelbrot (né en 1924) publiait deux compte-rendus à l'Académie des sciences sur la théorie de l'information intitulés *Mécanique statistique et théorie de l'information* et visant à la généralisation de la loi de Zipf (1902-1950) concernant le traitement statistique du vocabulaire dans les textes. Ces compte-rendus seront repris en 1954 sous forme d'un article destiné aux linguistes dans la revue *Word*. Celui-ci, intitulé *Structure formelle des textes et communication*, est le premier article en français appliquant les notions mathématiques de la théorie de l'information à des objets linguistiques. Il faut noter que la publication de ces compte-rendus précèdent de peu la publication du rapport de Jakobson, Fant et Halle au MIT en 1952, *Preliminaries to Speech Analysis*, ayant pour objectif d'appliquer la théorie de l'information à l'analyse des traits distinctifs en phonologie. En France, malgré cette parution quasi simultanée, la communauté des linguistes s'est fait davantage l'écho des travaux de Mandelbrot et de leur importance pour les statistiques lexicales. Quant aux travaux de Jakobson, ils ont été connus en France de façon très diverse et très variable dans le temps. Alors que certains ont eu accès dès leur parution aux travaux du début des années 1950 en phonologie, d'autres ont connu en premier les travaux ultérieurs sur les fonctions de communication au moment de leur traduction en français dans les *Essais de linguistique générale*, à savoir seulement en 1963. Cette diversité d'accès aux premiers travaux d'introduction de la théorie de l'information en linguistique explique en partie la singularité française de cette réception.

3.1 Jakobson, la théorie de l'information et les linguistes français

Il n'est pas le lieu ici d'étudier en détail le rôle de passeur de Jakobson ni sa conception de la théorie de l'information. Ses premiers travaux d'application de la théorie de l'information à la phonologie (Jakobson et al. 1952, Cherry et al. 1953, Jakobson and Halle 1956) ont été lus très tardivement en France. Rappelons que les *Essais de linguistique générale*, parus en 1963, ne comprennent pas ces travaux. Seul, Martinet, à New York jusqu'en 1955 et encore très ami avec Jakobson qui l'avait introduit dans les milieux universitaires américains, cite les *Preliminaries* dès leur parution (Martinet, 1952). Mandelbrot, quant à lui, cite sans véritablement les commenter les *Preliminaries* dans son article de Word de 1954 (probablement parce que la phonologie n'est pas son objet).

Dans *Preliminaries to Speech Analysis. The Distinctive Features and their Correlate*, Jakobson, Fant et Halle se proposent d'effectuer le traitement mathématique de l'information portée par les traits distinctifs dans un message et d'étudier leur pouvoir informationnel dans un code linguistique donné. Ils empruntent à Shannon son schéma de communication avec code, encodeur/ locuteur et auditeur/ décodeur et transmission d'information: un trait distinctif est reconnu par le récepteur s'il appartient au code commun à lui et à l'émetteur, s'il est transmis correctement et s'il atteint le récepteur. Ils montrent l'intérêt de la redondance qui augmente la fiabilité de la communication parlée en la rendant résistante aux différentes sources de distortion. Enfin, ils empruntent à la théorie de l'information son analyse en termes binaires des processus de communication :

(2) Information Theory uses a sequence of binary selections as the most reasonable basis for the analysis of the various communication processes. It is an operational device imposed by the investigator upon the subject matter for pragmatic reasons. In the special case of speech, however, such a set of binary selections is inherent in the communication process itself as a constraint imposed by the code on the participants in the speech event, who could be spoken of as the encoder and the decoder. (Jakobson et al., 1952 :9).

Le fait de traiter le message oral et le code sous-jacent en unités discrètes binaires est emprunté à Markov (1856-1922) via Shannon (1916-2001). Pour les auteurs, les traits distinctifs ont un caractère universel. Ceux-ci sont limités à douze, cependant aucune des langues ne contient l'ensemble de ces traits. Leur incompatibilité ou leur cooccurrence à l'intérieur d'une même langue et d'un même phonème est déterminée par des lois d'implication universellement valides ou bien une grande probabilité statistique. Cette probabilité permet d'éliminer les traits hautement probables et de diminuer la redondance

Enfin les auteurs privilégient les traits acoustiques aux traits articulatoires qui ne transmettent pas d'information directe au récepteur :

(3) In decoding the message received (A), the listener operates with the perceptual data (B) which are obtained from the ear responses (C) to the acoustical stimuli (D) produced by the articulatory organs of the speaker (E). The closer we are in our investigation to the destination of the message (ie its perception by the receiver), the more accurately we can gauge the information conveyed by its sound shape. This determines the operational hierarchy of levels of decreasing pertinence : perceptual, aural, acoustical, and articulatory (the latter carrying no direct information to the receiver). (Jakobson et al., 1952 :12).

Dans le texte de *Language* 1953, Cherry, Halle et Jakobson appliquent la théorie de l'information à la description phonémique des langues et en particulier à un corpus de conversations russes enregistrées. Il s'agit de déterminer le nombre de traits distinctifs (11 en l'occurrence) nécessaires à l'auditeur pour différencier les unités de sens minimale du code, à savoir les morphemes et leur combinaison en mots, sans l'aide du contexte. A partir de là toutes les autres différences phonétiques des morphèmes et des mots seront prédictibles et considérées comme redondantes. Ces articles seront suivis du petit essai *Fundamentals of Language* signé de Halle et Jakobson et paru en 1956, et de cinq articles parus en 1953, 1959 et 1960 (traduits dans les *Essais de linguistique générale*) où Jakobson développera les fonctions communicatives du langage.

3.2 Benoît Mandelbrot et la réception de la théorie de l'information dans la communauté linguistique française

Les scientifiques français s'intéressent à la cybernétique et à la théorie de l'information très tôt. En 1947, le mathématicien Szolem Mandelbrojt (1899-1983) invite Norbert Wiener (1894-1964) à un colloque à Nancy. Celui-ci se voit offrir la possibilité de publier en France un ouvrage sur le caractère unificateur de la cybernétique. C'est ainsi que *Cybernetics* fut publié en 1948 conjointement par les Editions Hermann à Paris et par les MIT Press et John Wiley & Sons aux Etats-Unis.

En 1949, Léon Brillouin (1889-1969), membre de l'Ecole des Hautes études de New York depuis 1941, promeut la théorie de l'information en physique. En 1950, Louis de Broglie (1892-1987) organise une série de conférences ayant pour titre "Cybernétique. Théorie du signal et de l'information" qui conduisit à la reconnaissance de la théorie de l'information comme science autonome. Dans sa thèse publiée en 1953, Marcel-Paul Schützenberger (1920-1996) insiste sur le caractère unificateur pour les sciences de la théorie de l'information³.

Dans ce contexte, le mathématicien Benoît Mandelbrot joue un rôle clé dans l'introduction de la théorie de l'information dans la linguistique en France. Polytechnicien, neveu de Szolem Mandelbrojt, il complètera sa formation aux Etats-Unis (Californie en 1948-49, Princeton et MIT en 1951-53), et est chargé de relire les épreuves de l'ouvrage *Cybernetics* de Wiener. Il soutient sa thèse d'état en 1952 sur l'interaction entre théorie des jeux et théorie de l'information. Il montre que la thermodynamique comme les propriétés statistiques du langage peuvent être expliquées comme les résultats de jeux entre nature et émetteur. Mandelbrot fait partie, aux côtés de Shannon, Wiener et Jakobson, du comité de lecture de la revue *Information and Control*, fondée en 1958 par Brillouin, Peter Elias et Colin Cherry (co-auteur avec Jakobson et Halle de l'article de *Language* de 1953).

Dans son article de *Word* paru en 1954 "Structure formelle des textes et communication", première publication en français sur l'application de la théorie de l'information à des objets linguistiques, Mandelbrot critique la loi d'Estoup et Zipf qui présuppose que les mots-formes (formes fléchies) auraient des propriétés intrinsèques dans un texte. Il propose à la place une théorie canonique, fondée sur la théorie de l'information et les formes vides (de sens) dont la loi de Zipf ne serait qu'un cas particulier. Cet article a ouvert la voie à nombre de travaux en statistiques lexicales, plus ou moins inspirés de Shannon et Weaver et de Zipf, prenant le relais des études de vocabulaire très développées en France dans les décennies précédentes⁴. A partir de 1954, Pierre Guiraud (1912-1983), dans une version très contestée par les statisticiens, dont Mandelbrot lui-même, poursuit les travaux de statistiques lexicales en diffusant la théorie de l'information auprès des linguistes et en promouvant le terme de *statistique linguistique* (Guiraud 1954, 1960).

Le Bulletin de la Société de Linguistique de Paris est un bon observatoire de cette diffusion auprès des linguistes. Les compte-rendus d'ouvrage sur les statistiques lexicales se multiplient dans les années 1950, de même que les exposés au sein de la Société elle-même et les articles publiés dans le *BSL*: compte-rendus de Mandelbrot par Guiraud, et de Guiraud par Mandelbrot, compte-rendu de Herdan (1956) par Mandelbrot. Benveniste et Martinet ont tous deux rédigé dans le *BSL* des compte-rendus d'ouvrages sur l'application de la théorie de l'information par des linguistes ou psychologues américains. Dans le même tome (T.53, 1957-58), un compte-rendu très critique par Benveniste de l'ouvrage de Joshua Whatmough *Language. A Modern Synthesis* (1956) est suivi de deux compte-rendus de Martinet : l'un assez négatif de *Language and Communication* de G-A. Miller (1951, tr. française 1956), l'autre plus élogieux de l'ouvrage de Vitold Belevitch *Langage des machines et langage humain* (1956) que Martinet considère comme une bonne introduction à la théorie de l'information à destination des linguistes. La contribution, plus tardive, de Dubois consiste en un compte-rendu de la parution en français des *Essais de Linguistique générale* de Jakobson dans le *Français Moderne* (1964).

Benveniste, Martinet et Dubois comptent parmi les linguistes français, hors ceux bien sûr qui se sont consacrés aux statistiques lexicales, à s'être intéressés à la théorie de l'information au point de tenter de l'intégrer dans leurs travaux.

4 Benveniste : une extériorité prudente

Chacun des deux tomes des *Problèmes de Linguistique générale (PLG)* comporte une partie intitulée 'Communication'. Dans les index figurent un certain nombre de mots appartenant à la constellation des termes du domaine: *information, codage, communication, cybernétique* et *signal*, dans le tome 1. Le tome 2 semble plus pauvre, on y trouve seulement *théorie de l'information, communication* et *non-redondance*.

Ces termes ne sont pas très fréquents dans le texte. Si l'on s'en tient à *information*, il n'y a que 7 occurrences dans le tome 1, et 4 dans le tome 2 dont 3 *théorie de l'information*. Le terme *information* est toujours utilisé dans son acception ordinaire, comme par exemple dans les trois extraits suivants, où il s'agit d'une *information* ou de *l'information* + déterminant, et jamais dans le sens technique d'une quantité abstraite:

(4) Ici *l'information* du N.E.D. est en défaut (PLG, T1 :342).

Il [l'homme parlant] veut lui [interlocuteur] transmettre un élément de connaissance, ou obtenir de lui une information, ou lui intimer un ordre. Ce sont les trois fonctions interhumaines du discours qui s'impriment dans les trois modalités de l'unité de phrase, chacune correspondant à une attitude du locuteur (PLG, T1 :130).

En décrivant, il y a quelques années, les formes subjectives de l'énonciation linguistique, nous indiquions sommairement la différence entre je jure, qui est un acte, et il jure, qui n'est qu'une information (PLG, T1 :270).

Quant à la *théorie de l'information*, Benveniste se contente de citer cette nouvelle théorie sans que sa mise en œuvre théorique soit réellement envisagée. Comme la traduction automatique, la théorie de l'information offre de nouvelles techniques qui s'avèrent probablement très prometteuses pour les sciences du langage, mais dans un horizon qui reste indéterminé. Ainsi, en conclusion du chapitre I-2, intitulé *Coup d'œil sur le développement de la linguistique*, Benveniste aborde les nouvelles techniques permettant le développement de la symbolisation du langage dont *les théories de l'information*⁵ :

(5) Par ailleurs on sait que les descriptions formelles des langues ont une utilité directe pour la construction des machines logiques aptes à effectuer des traductions, et inversement on peut espérer des *théories de l'information* quelque clarté sur la manière dont la pensée est codée dans le langage (PLG, T1 :30-31).

Benveniste fait un parallèle entre phrase et mots comme unités de la sémantique, et message et unités de codage pour la théorie de l'information (ici à prendre au sens de schéma de communication) dont il postule une organisation syntagmatique. On sait que, pour Benveniste, une phrase est un énoncé de caractère nécessairement sémantique considéré dans son « emploi instantané, spontané, personnel » (PLG, T2 :232) :

(6) Que l'idée ne trouve forme que dans un agencement syntagmatique, c'est là une condition première, inhérente au langage... il [le linguiste] peut seulement conjecturer que cette condition toujours nécessaire reflète une nécessité de notre organisation cérébrale. On retrouve dans les modèles construits par *la théorie de l'information* la même relation entre le message et les unités probables du codage. (PLG, T2 :226)

Plus haut il indique que le mot est « l'unité nécessaire du codage de la pensée » (PLG, T2 :225). C'est donc une interprétation très personnelle que Benveniste propose de la théorie de l'information puisque ses unités de codage sont des unités sémantiques qui s'avèrent tout à fait en contradiction avec les unités d'information de Shannon et Weaver, quantités dépourvues de sens.

Ainsi, bien que Benveniste nous indique la voie en proposant dans son texte *Genèse du terme 'scientifique'*, d'examiner la genèse d'une discipline scientifique à partir des termes (dont le terme *information* d'ailleurs), lui-même ne semble pas faire grand cas de l'information :

(7) Nous tenons donc l'apparition ou la transformation des termes essentiels d'une science pour des événements majeurs de son évolution. ... Les termes instructifs sont ceux qui s'attachent à un concept neuf désigné à partir d'une notion théorique (civilisation, évolution, transformisme, *information*, etc.), mais aussi bien ceux qui, dérivés d'une notion antérieure, y ajoutent une détermination nouvelle (PLG, T2 :247-248).

Il tient en effet la (ou les) théorie(s) de l'information à une distance prudente et ses références à ce nouveau domaine sont très limitées. Il ne prête aucune attention à l'aspect quantitatif de l'information et fait au contraire glisser les unités de codage vers des unités de sens. Il est probablement plus intéressé par les questions posées par le schéma de communication de Jakobson et l'hypothèse d'interlocution qu'elle suppose⁶. L'extériorité de la nouvelle théorie par rapport à la linguistique reste donc entière.

5 Martinet : pour une conception fonctionnelle de l'information

Martinet, nous l'avons dit, est sans doute le seul linguiste français à avoir lu les *Preliminaries* de Jakobson et al., c'est-à-dire la première étude de phonologie appliquant la théorie de l'information. Toutefois, bien qu'il cite ce texte, Martinet ne reprend pas à son compte cette utilisation de la théorie de l'information lorsqu'il élabore sa phonologie diachronique.

Dans son *Économie des changements phonétiques*, publié en 1955, et qui synthétise les travaux phonologiques de ses années américaines, Martinet prétend s'inscrire dans la filiation de Troubetzkoy (1890-1938) contre Jakobson. Selon lui, l'analyse des traits distinctifs est présente à l'état latent dans l'œuvre de Troubetzkoy (Martinet, 1955 :67, note 8 et 1957-58 :75). Par ailleurs, il consacre plusieurs pages à une critique sévère de l'apriorisme universaliste et binariste développé par Jakobson et al. (1952, 1953) qui dévoie selon lui les positions de Troubetzkoy. Le binarisme procède par affirmations de caractère général et cherche à faire entrer toute réalité phonologique dans des cadres préétablis et identiques pour toutes les langues.

(8)... ce qui rend la position binariste absolument inacceptable en matière diachronique, c'est l'élimination arbitraire, comme « redondantes », de caractéristiques phoniques résultant d'évolutions qui ont changé les rapports à l'intérieur du système, ce qui aboutirait à poser que ces changements sont nuls et non venus. le point de vue diachronique exige un beaucoup plus vif souci de la réalité phonétique que celui qui est de mise lorsqu'on s'escrime à réduire au minimum le nombre des traits distinctifs (Martinet, 1955 :76).

... l'apriorisme qui consiste à préciser les traits pertinents, moins en s'inspirant du système de la langue à l'étude, que par référence à un schéma préétabli dont on postule la valeur universelle (Martinet, 1957-58 :75).

Par ailleurs, il s'oppose à la détermination des phonèmes par des moyens acoustiques préconisée par Jakobson et al. et préfère, en suivant Troubetzkoy, les données articulatoires arguant que l'analyse spectrographique ne fait que confirmer l'explication que fournissent immédiatement les données articulatoires.

Il ne discute pas spécifiquement les emprunts faits par Jakobson et ses co-auteurs à la théorie de l'information. Ceux-ci sont dénigrés et qualifiés de pur habillage physico-mathématique destiné à séduire les esprits :

(9) ... cependant dépouillée de l'appareil physico-mathématique dont on l'a revêtue pour l'agrément des esprits avides de rigueur abstraite, cette théorie apparaît plus comme une vue de l'esprit que comme un effort pour coordonner les résultats d'observations préalables (Martinet, 1955 :73)

Toutefois, le terme *information* apparaît deux fois dans l'*Economie* (p.140) associé au pouvoir distinctif. Le pouvoir distinctif, liée à une fréquence faible limite l'incertitude et compense une dépense d'énergie parfois plus forte, comme dans le cas de la prononciation d'une gémée. *Information* est bien employé ici dans le sens de Jakobson tel que repris à la théorie de l'information, sans que d'ailleurs celui-ci soit cité à cet endroit du texte.

(10) Je puis donc dire qu'en français /-kt-/ est beaucoup plus riche d'*information* que /-t-/ et que du point de vue de l'économie de la langue, la distinction supplémentaire apportée par /-k-/ vaut bien le surcroît de travail musculaire qu'il implique ... Si nous passons maintenant à /-atta-/ , nous pourrions dire que, là où cette combinaison se rencontre cinq fois sur cent fois /-ata-/ , le surcroît de travail que représente la prononciation du /-t-/ implosif, qui distingue /-atta-/ de /-ata-/ vaut largement la peine, puisqu'il a pour effet de limiter l'incertitude dans la proportion de 100 à 5. la gémée a donc un grand pouvoir distinctif. Dans la langue où l'on rencontre /-atta-/ 80 fois pour 100 exemples de /-ata-/ et 10 exemples de /-akta-/ , /-tt-/ représentera une dépense d'énergie de même ordre que celle qui est nécessaire pour articuler /-kt-/ mais pour un *pouvoir d'information* beaucoup plus bas (Martinet, 1955 :140).

On voit ici opérer la tendance contradictoire interne à la notion d'économie, signalée par Verleyen et Swiggers (2006 :176), entre la satisfaction des besoins de la communication qui exige la préservation de la quantité maximale d'information, à savoir un nombre maximal d'oppositions fonctionnelles, et celle de l'inertie qui tend à utiliser un nombre restreint d'unités plus fréquentes.

Dans l'extrait (10), on voit comment l'information contribue au principe du moindre effort que Martinet emprunte à Zipf. L'homme cherche, dans toutes ses activités y compris ses activités langagières, à minimiser l'effort (travail musculaire, dépense d'énergie) nécessaire par rapport au but à atteindre. Selon Martinet, le principe de Zipf réalise mieux la synthèse entre besoins de communication et inertie humaine que le principe d'économie élaboré par Paul Passy dans le cadre de sa théorie fonctionnaliste des changements phonétiques⁷. Enfin, il préconise de mesurer l'importance fonctionnelle d'une opposition phonologique ou *rendement fonctionnel* à l'aide de statistiques dans les textes. En cela aussi, il s'est inspiré de Zipf qui, le premier, a jeté les bases d'une phonologie fonctionnelle qui ne soit pas purement descriptive mais fondée statistiquement.

Cet intérêt pour les aspects quantitatifs de l'information se retrouve dans les *Éléments de linguistique générale*, parus en 1960, dont une grande partie du chapitre 6 *L'évolution des langues* est consacrée à la théorie de l'information⁸. Dans cet ouvrage, surtout dans le § 6.9 intitulé *La théorie de l'information et le linguiste*, Martinet prétend introduire ses vues personnelles sur la théorie de l'information sans aucune référence aux travaux antérieurs d'application de la théorie de l'information à la linguistique, notamment ceux de Jakobson⁹. D'ailleurs, la notion d'information est dissociée de celle de traits distinctifs (et donc de l'apport jakobsonien) : ceux-ci n'apparaissent plus que de façon sporadique (10 occurrences de *traits distinctifs* en tout dans les *Éléments* dont aucune dans le chapitre 6). Les références concernant la théorie de l'information sont les exposés « relativement simples » de Guiraud à la SLP en 1954 et l'ouvrage de Belevitch dont il a fait le compte-rendu dans le BSL (Martinet, 1957b).

Selon Martinet, les formules mises au point par les ingénieurs des télécommunications pour résoudre leurs problèmes ne sont pas adaptées aux problèmes spécifiques des linguistes. Ce qui intéresse le linguiste, dit-il, c'est de savoir dans quelle probabilité la variation de certains facteurs vont entraîner celle de certains autres.

(11) Ces variables sont le *nombre* des unités entre lesquelles le locuteur choisit à un point de l'énoncé, la probabilité des unités ramenée à leur *fréquence*, le *coût* de chaque unité qui comporte, outre l'énergie nécessaire à sa réalisation, ce qu'on pourrait appeler les frais d'emménagement dans la mémoire, et enfin l'*information* qu'apporte chaque unité (Martinet, 1960 : 182).

Bien qu'il utilise également d'autres termes de la constellation comme *code*, *communication*, *coût*, *fréquence*, *message* ou *redondance*, le terme *information* est un des plus fréquents: on en trouve 59 occurrences dans l'ensemble du texte, dont 48 dans la partie 6-III. A l'exception de quelques occurrences du terme dans son sens ordinaire, Martinet utilise *information* au sens technique, soit dans un syntagme, *quantité d'information* ou *coût de l'information*, soit en collocation avec des verbes ou des nominalisations de verbes de quantité (*mesurer*, *augmenter*, *doubler*, *varier*, *élever*, *dépasser*, *accroissement*), ou encore avec des adjectifs ou des adverbes exprimant la quantité (*grande information*, *information bien supérieure*, *faible information*, *trop d'information*, *autant d'information*, *masse*

d'information, dose d'information, densité d'information, richesse informationnelle). Le verbe *apporter* occupe une place à part avec dix occurrences en collocation avec *information*.

Suivant en cela la théorie de l'information, Martinet considère qu'il existe un rapport constant et inverse entre la fréquence d'une unité et l'information qu'elle apporte. Plus une unité est probable, moins elle est informative. Ce qu'il ajoute, c'est une définition de l'information en termes de moindre effort à savoir d'efficacité par rapport à l'énergie consommée à chaque utilisation d'une unité.

Outre la phonologie, il applique la notion d'information à la littérature et à d'autres unités linguistiques, comme les monèmes ou les formes lexicales. Martinet semble utiliser les chaînes de Markov pour traiter ces séquences d'unités, d'une façon très proche de celle de Harris (1955). Il rapporte d'ailleurs dans ses *Mémoires* (Martinet, 1993 :71), que l'idée selon laquelle la limite des mots pourrait s'établir sur la base de la probabilité des éléments successifs lui est venue lors d'une discussion avec Harris (1909-1992) à New York au début des années 1950. Dans cette utilisation, il précise que l'information n'est en rien une entité sémantique :

(12) Si j'entends /il a p.../, /p/ n'a pas de signification en lui-même, mais il est doué d'information dans ce sens qu'il exclut toutes sortes d'énoncés possibles, comme *il a donné, il a bougé*. Si à l'énoncé tronqué s'ajoute /r/ (/il a pr.../), l'incertitude est de nouveau réduite puisque sont exclus *il a payé, il a poussé*, etc., ce qui indique que /r/ est doué d'information. L'information n'est donc pas un attribut de la signification, puisque des unités non significatives comme /p/ et /r/ y participent. (Martinet, 1960 :177-78)

C'est au moment où il aborde le rapport entre information et littérature (§6-18) que Martinet utilise le syntagme *contenu informationnel* qui apparaît d'ailleurs cinq fois. Par un choix d'unités lexicales nouvelles et en réduisant la redondance, l'auteur pourra augmenter le contenu informationnel d'un texte afin de retenir l'attention du lecteur. Ce qui est redondant, ce sont les collocations attendues, et ce qu'on attend du poète ce sont des collocations rares.

(13) L'auteur pourra se contenter de présenter, dans les termes les plus directs, des événements, réels ou imaginaires, assez exceptionnels pour que la densité informationnelle du récit retienne l'attention. Il pourra aussi, par un choix original des unités linguistiques, élever le contenu informationnel de son texte et le doser exactement. Ceci le dispensera d'aller, à chaque instant, chercher l'inattendu dans les péripéties du récit. (Martinet 1960 :192).

Avec le terme *contenu informationnel*, l'information semble perdre son caractère quantitatif au profit d'une dimension sémantique : *contenu informationnel* fait en effet penser à *contenu de sens*. A noter que le terme *contenu d'information*, sous quelle que forme et langue que ce soit, n'existe pas dans l'ouvrage de Shannon et Weaver.

En fait, on peut faire l'hypothèse que *contenu d'information* vient de Guiraud, une des sources revendiquées par Martinet. Guiraud applique la loi de Zipf et la théorie de l'information aux statistiques de vocabulaire dans des études stylistiques. Comme on le voit dans (14), Guiraud assimile de façon un peu rapide fréquence (des signes) et contenu d'information, alors qu'on a vu que l'information est une mesure abstraite dépendant de choix, donc une mesure probabiliste et pas seulement statistique.

(14) En fait l'équation de Zipf est susceptible d'une double interprétation. Pour les uns (dont je suis) la fréquence des signes définit leur *contenu d'information* (au sens que Shannon donne à ce terme) et n'est que le produit de l'économie de la communication. Pour Mandelbrot le langage est un système moléculaire, assimilable à une masse de gaz et soumis à des lois similaires ... Mais autant l'hypothèse d'une économie de la communication en terme de *contenu d'information* est riche et pratique pour le linguiste, autant il répugne de s'aventurer sur le terrain de la thermodynamique où l'entraîne Mandelbrot... (Guiraud 1957:24).

Comme on le voit, cette position oppose Guiraud à Mandelbrot qui, quant à lui, reste fidèle à la définition de l'information comme entropie négative : « La quantité précisément nécessaire à l'établissement de l'information correspond exactement à la notion thermo-dynamique d'entropie » (Shannon et Weaver [1975 : 42]).

Il faut savoir que l'utilisation approximative des méthodes statistiques par Guiraud a été très critiquée par les mathématiciens, que ce soit Mandelbrot (1954b) ou Moreau (1964), et que leur différend est aussi de nature linguistique. Celui-ci divisera les tenants de la linguistique statistique pendant des décennies, et plus récemment la linguistique de corpus. Pour Guiraud, il existe une fréquence intrinsèque des mots dans la langue, ce que conteste Mandelbrot. On a vu (cf. §3.2. ci-dessus) que c'est d'ailleurs une des critiques que Mandelbrot adresse à Zipf. On notera que la position de Martinet n'est pas très explicite sur ce point.

6 Dubois : une version harrisso-jakobsonienne

Curieusement, Dubois, bien qu'ayant commencé ses travaux de linguiste en lexicographie avec sa thèse parue en 1962 sur *Le vocabulaire politique et social en France de 1869 à 1972*, ne s'est pas intéressé aux aspects quantitatifs du vocabulaire, pourtant très en vogue en France dans les années 1950-60 et impulsés par les travaux de Mandelbrot et de Guiraud. Dans sa thèse, Dubois met en oeuvre la méthode distributionnelle de Harris, dont on peut dire qu'il est l'introduit en France, et son utilisation en analyse du discours¹⁰. Son intérêt pour la théorie de l'information apparaît dans des textes postérieurs, d'abord dans son article sur l'aspect et le temps paru en 1964 dans *le Français Moderne*, puis dans sa *Grammaire structurale du français* en trois tomes parus respectivement en 1965, 1967 et 1969.

Comme pour les autres auteurs, *information* n'apparaît pas seul mais avec une cohorte de termes de la constellation, comme on le voit dans le tableau (15):

(15)

	1964 (FM)	1965 (GS 1)	1967 (GS 2)	1969 (GS 3)
Information-, informer	0	33	3	1
Communication	8	20	27	5
Cod-	24	43	7	9
message	9	24	13	6
Emetteur, locuteur, interlocuteur	5	13	29	12
récepteur	1	5	2	6
Redondan-	2	35	14	1
Bruit	0	10	1	0

Comme on le voit, c'est surtout dans le tome 1 de 1965 qu'*information* apparaît massivement pour quasiment disparaître dans les tomes 2 et 3, alors que les autres termes (dont on n'a fait figurer ici que les plus fréquents) restent encore plus ou moins utilisés.

Si l'on examine les formes suivantes apparaissant dans le tome 1 :

(16)

- la théorie de l'information : 1
- la quantité d'information : 4
- l'information : 3

- SN + l'information (conservation/perte de l'information, coût de l'information) : 5
- SN + information (perte d'information, les éléments d'information perdus) : 2

(17)

- une information : 1
- une information +det (nouvelle) : 1
- deux informations: 2
- SN + une information (conservation d'une information) : 1
- l'information + det (précédente, première, donnée, féminin, de genre, de nombre): 9
- les informations + det (nécessaires et constantes) : 1

Les occurrences des expressions rassemblées en (16) réfèrent bien à la théorie de l'information. Dans son introduction, Dubois précise qu'il a pour objectif une analyse distributionnelle des marques de genre et de nombre du nom et du pronom, qui se veut complémentaire de l'analyse distributionnelle de Harris. C'est en termes d'information, redondance et bruit qu'il se propose d'analyser les contraintes croisées entre marques de genres et de nombres pour montrer que ce sont des règles et non des exceptions en français. Alors que, dit Dubois - faisant sans doute référence à la notion de *rendement fonctionnel* de Martinet - les linguistes ont surtout emprunté à la théorie de l'information la notion de coût, il se propose, quant à lui, de mettre en œuvre celles de redondance et de bruit. La seule occurrence du terme *théorie de l'information* apparaît dans cette exposition du problème.

La notion de quantité d'information (de même que celles de conservation, perte et coût de l'information) est bien celle de la théorie de l'information : elle est calculée en termes de probabilités et en termes de fonction inverse : plus la probabilité d'apparition d'une marque est grande, plus la quantité d'information est faible. Ainsi la quantité d'information apportée par le code écrit est faible, puisque celui-ci est très redondant :

(18) Les marques du code parlé et du code écrit ne sont donc pas équiprobables : la *quantité d'information* apportée par une marque redondante dans le code parlé est plus grande que celle apportée dans le code écrit, puisque la probabilité de la rencontrer après la marque initiale est moindre. La *quantité d'information* présentée en français écrit par la seconde marque est très faible, puisqu'il est très rare que le code écrit ne présente pas de redondance. (Il n'en est évidemment pas de même pour le code parlé) (1965 : 21).

Cette utilisation quantitative et shannonienne de l'information par Dubois est probablement issue de la lecture de Mandelbrot (1954a) figurant dans sa bibliographie mais aussi de celle de Harris, dont les principaux travaux des années 1950 sont cités, et de son utilisation des chaînes de Markov¹¹.

Si maintenant on regarde les expressions dans la liste (17) où *information* est déterminé (*une information, la première information, l'information du pluriel, l'information féminin, l'information de genre, celle de nombre, les /ces deux informations*), on est face à un tout autre usage du terme :

(19) On peut se demander ce que deviennent *ces deux informations* lorsqu'elles sont cumulées, c'est-à-dire lorsque les énoncés au pluriel sont aussi porteurs de *l'information féminin* (1965 :82).
... il arrive que *l'information de genre* se maintienne, tandis que celle de nombre s'efface. ... *l'information féminin* a disparu des deux phrases (1965 :83).

Cet emploi d'*information* fait glisser le terme vers *contenu* ou *sens*. Toutefois il ne s'agit pas ici de l'information au sens ordinaire, c'est-à-dire au sens de renseignement, connaissance, acte de rechercher des connaissances, ou ensemble de connaissances existantes. On peut alors se demander pourquoi Dubois utilise le terme *information* dans ce cas précis. Alors qu'il désignait une quantité abstraite, calculée à l'aide de probabilités et excluant le sens, le terme *information* définit maintenant une signification grammaticale, d'ordre morpho-syntaxique, donnée par les marques.

Pour répondre à cette question, il faut se rappeler que Dubois est familier des *Essais de Linguistique Générale* de Jakobson, dont il a fait un compte-rendu dans un numéro du *Français Moderne* en 1964, numéro dans lequel il publie également un article où il applique les théories de Jakobson aux problèmes

du temps et de l'aspect. Entre autres choses, Dubois emprunte à Jakobson l'idée que toutes les formes grammaticales (en particulier l'aspect et le temps) sont marquées dans le code de façon variable selon les langues. Aussi cherche-t-il les structures formelles dans lesquelles sont traduites dans un code donné les oppositions sémantiques d'accompli/ non accompli, antériorité / non-antériorité, et postériorité / non postériorité qui définissent ce que l'on appelle l'aspect et le temps ; on doit en effet constater qu'il existe des langues qui sont appelées à traduire non ces trois oppositions, mais parfois seulement deux d'entre elles, le développement de la postériorité / non postériorité se faisant quelquefois tardivement.

Or ces idées sont développées chez Jakobson dans le chapitre des *Essais* intitulé *La notion de signification grammaticale selon Boas* (1963 :197-206). Et c'est à Boas (1858-1942) qu'il emprunte l'idée de *grammatical meaning*. Afin que la démonstration soit plus convainquante, examinons la version anglaise :

(20) The choice of a grammatical form by the speaker presents the listener with a definite number of *bits of information*. The compulsory character of this kind of *information* for any verbal exchange within a given speech community and the considerable difference between the *grammatical information* conveyed by diverse languages were fully realized by Franz Boas, thanks to his astonishing grasp of the manifold semantic patterns of the linguistic world... (Jakobson, 1959 : 490).
 It was clear to Boas that any difference of grammatical categories carries *semantic information*. (Jakobson, 1959 : 493)

Pour Boas, les marques morpho-syntaxiques sont porteuses de signification (*grammatical meaning*, ou signification grammaticale) de sorte que l'information elle-même devient sémantique. Il est tout à fait étonnant que Jakobson, qui, comme on l'a vu, fut co-auteur avec des mathématiciens de plusieurs textes appliquant la théorie de l'information à la phonologie, utilise dans le même paragraphe le terme *information* au sens mathématique (*bits of information*) et au sens sémantique (*grammatical information*). Dans ce texte, il semble avoir s'être approprié le terme et l'utiliser à sa guise sans chercher à rester au plus près de son sens mathématique. Cette attitude est cohérente avec le fait que Jakobson considère que les ingénieurs de télécommunications et les linguistes doivent collaborer à la construction d'une théorie de l'information sans qu'il y ait prééminence des premiers sur les seconds ; la théorie de l'information n'appartient pas seulement aux ingénieurs et aux mathématiciens, mais aussi aux linguistes et les rapports entre linguistique et théorie de l'information (dans sa version probabiliste et shannonienne) ne sont pas à sens unique. Les mathématiciens peuvent aussi bénéficier de l'étude des langues susceptible de susciter le développement de nouvelles théories.

(21) Quelle est donc exactement la relation entre la théorie de la communication et la linguistique ? Ya-t-il peut-être des conflits entre ces deux modes d'approche ? En aucune façon. Il est un fait que la linguistique et les recherches des ingénieurs convergent, du point de vue de leur destination. Mais alors de quel ordre est exactement l'utilité de la théorie de la communication pour la linguistique, et vice-versa ? Il faut reconnaître que, sous certains aspects, les problèmes de l'échange de l'information ont trouvé chez les ingénieurs une formulation plus exacte et moins ambiguë, un contrôle plus efficace des techniques utilisées, de même que des possibilités de quantification prometteuses. D'un autre côté, l'expérience immense accumulée par les linguistes relativement au langage et à sa structure leur permet de mettre au jour les faiblesses des ingénieurs quand ils s'attaquent au matériel linguistique (Jakobson, 1963 :28).

Pour revenir à Dubois, en empruntant les significations grammaticales de Boas-Jakobson, il donne à l'information un sens sémantique tout à fait opposé au sens quantitatif de la théorie de Shannon et Weaver. Toutefois ce n'est pas le seul emprunt qu'il fait à l'interprétation jakobsonienne de la théorie de l'information. Il fait également sienne la façon dont Jakobson s'est approprié la notion de redondance :

(22) La notion de redondance, empruntée par la théorie de la communication à une branche de la linguistique, la rhétorique, a acquis une place importante dans le développement de cette théorie et a été audacieusement redéfinie comme équivalant à

‘un moins l’entropie négative’ ; sous cet aspect, elle a fait sa rentrée dans la linguistique actuelle, pour en devenir un des thèmes centraux. (Jakobson, 1963 :89)

Dubois inscrit la redondance au cœur de son étude des marques du français, et l’utilise encore de façon très active dans le tome 2 de la *Grammaire structurale* (voir tableau 15), contrairement à la notion d’information qui disparaît.

Alors que chez Jakobson, la migration des termes s’est véritablement accompagnée d’une appropriation de la théorie de l’information¹², on pourrait dire que chez Dubois cet emprunt jakobsonien, de seconde main en quelque sorte, peut paraître contradictoire avec l’usage distributionnaliste et quantitatif qui cohabite dans ses textes. Il faut toutefois reconnaître que la théorie de l’information acquiert avec Dubois un pouvoir heuristique qui lui permet de penser la dimension sémantique de certaines catégories linguistiques.

7 Conclusion

Avec nos trois auteurs, nous assistons à trois modes de réception en linguistique générale d’un concept appartenant à une théorie non-linguistique. À côté de la prudente mise à distance de Benveniste, qui s’empare de *code* et de *message* mais se préoccupe peu d’*information*, on rencontre chez Martinet un mode d’importation plus complexe. C’est au moment où il fait le moins référence au terme *information*, dans *Economie des changements phonétiques*, que l’usage qu’il en fait est le plus proche de celui de Jakobson et de sa systématisation des traits distinctifs. Bien que cet usage soit accompagné d’une très vive critique de l’apriorisme binariste de Jakobson et de son utilisation des traits acoustiques, et que Martinet revendique davantage l’influence de Troubetzkoy que celle de Jakobson, il est certain que la notion de pouvoir d’information, lié au pouvoir distinctif qui intervient dans sa définition du concept d’économie, est inspiré par la conception jakobsonienne. Il est à noter toutefois que sa notion d’économie en phonologie diachronique doit tout autant aux travaux de Zipf et son principe de moindre effort qu’à ceux de l’Ecole de Prague. Dans les *Éléments de linguistique générale*, il consacre un long développement à sa propre vision de l’information, ramenée à l’idée d’énergie et de coût. On notera aussi sa conception probabiliste issue partiellement de l’utilisation de Harris des chaînes de Markov et son utilisation du terme *contenu d’information* inspirée par les travaux en stylistique de Guiraud. Dubois, quant à lui, met en œuvre de façon heuristique les différentes facettes de l’interprétation jakobsonienne tout en étant également attiré par l’utilisation qu’en fait Harris. Il se trouve ainsi tiraillé entre deux options contradictoires, d’une part une conception de l’information fonctionnaliste comportant une dimension sémantique, d’autre part une approche distributionnaliste qui reste au plus près de l’utilisation probabiliste de chaînes de Markov mais qui devrait exclure le sens.

Il est intéressant de noter que ce n’est pas le sens ordinaire du terme ‘information’ qui provoque l’instabilité de la notion. Cet usage cohabite de façon assez étanche avec le sens technique. Il apparaît dans des environnements distincts comme dans « obtenir de l’interlocuteur une *information* » ou « le descripteur vérifie son *information* ». C’est au contraire le sens technique qui fluctue entre l’information quantitative et non sémantique de Shannon mise en œuvre souvent avec une utilisation distributionnelle des chaînes de Markov, une information fonctionnelle du moindre coût largement issue des travaux psychologico-statistiques de Zipf, une information sémantique inspirée de Jakobson dans sa lecture de Boas, un contenu d’information des unités linguistiques issu des études stylistiques du vocabulaire. Ces lectures croisées d’inspirations multiples : la diffusion en France à travers la SLP, la traduction automatique, les interprétations de la théorie de l’information par des chercheurs d’horizons aussi divers que le mathématicien Mandelbrot, le stylisticien Guiraud, les linguistes fonctionnalistes comme Jakobson ou néo-bloomfieldiens comme Harris, ou bien encore le psychologue anti-behaviouriste Miller, ont configuré de façon multiple la réception de la théorie de l’information par les linguistes français des années 1960. Cette diversité de positions et d’influences a probablement contribué à la multiplicité des structuralismes (cf. Puech, à paraître) manifeste y compris au sein même de la linguistique française de cette même période.

Références Bibliographiques

Sources primaires :

- Benveniste, E. (1957). Compte-rendu de l'ouvrage de Joshua Whatmough *Language. A Modern Synthesis*, 1956, London : Secker & Warburg. *BSL T.53*, 18-22.
- Benveniste, E. (1966/1974). *Problèmes de linguistique générale*, 2 tomes Paris : Seuil.
- Bulletin de la Société de Linguistique : tome 43 (1942-1946) à tome 65 (1970).
- Cherry E.C., Halle M. & Jakobson R. (1953) Toward the Logical Description of Languages in their Phonemic Aspect. *Language* 29, 34-46.
- Dubois, J. (1964). La traduction de l'aspect et du temps dans le code français (structure du verbe). *Le Français moderne* 32 -1, 1-26.
- Dubois, J. (1965). *Grammaire structurale du français, nom et pronom*. Paris : Larousse.
- Dubois, J. (1967). *Grammaire structurale du français, le verbe*. Paris : Larousse.
- Dubois, J. (1969). *Grammaire structurale du français, la phrase et les transformations*. Paris : Larousse.
- Guiraud P. (1954). *Les Caractères statistiques du vocabulaire*. Paris : P.U.F.
- Guiraud, P. (1957). Compte-rendu de Apostel L, B. Mandelbrot et A. Morf Logique, 1957, *Langage et théorie de l'information*. *BSL T.53*, 23-25.
- Guiraud P. (1960). *Problèmes et méthodes de la statistique linguistique*. Paris : P.U.F.
- Harris, Z.S. (1955). From phoneme to morpheme. *Language* 31, 190-222.
- Jakobson R., Fant C. G. M., Halle M. (1952). *Preliminaries to Speech Analysis. The Distinctive Features and their Correlates* MIT Acoustic Laboratory, Technical report n°13.
- Jakobson, R., Halle M. (1956). *Fundamentals of Language*. The Hague : Mouton.
- Jakobson, R. (1959). Boas' view of grammatical meaning. *Selected Writings, T.2*, 489-496
- Jakobson, R. (1963). *Essais de linguistique générale*. Paris : Les Editions de Minuit.
- Mandelbrot, B. (1951). Mécanique statistique et théorie de l'information. *Compte rendus des séances de l'Académie des Sciences* t.232, 30 avril 1951, 1638-40, 2003-2005.
- Mandelbrot, B. (1954a). Structure formelle des textes et communication. *Word* 10-3, 1-27.
- Mandelbrot, B. (1954b). Compte-rendu de l'ouvrage Pierre Guiraud 1954 *Les caractères statistiques du vocabulaire. Essai de méthodologie*. Paris : PUF. *BSL T.50*, 16-21.
- Martinet A. (1952). Function, Structure and Sound Change. *Word* 8-1, 1-32.
- Martinet, A. (1955). *Économie des changements phonétiques*. Berne : Francke.
- Martinet, A. (1957a). Compte-rendu de George A. Miller *Langage et communication* PUF 1956 (tr.fr.). *BSL T.53*, 25-26.
- Martinet, A. (1957b). Compte-rendu de Vitold Belevitch *Langage des machines et langage humain* Paris : Hermann 1956. *BSL T.53*, 27-29.
- Martinet, A. (1957-58). Substance phonique et traits distinctifs *BSL T53* :72-85.
- Martinet, A. (1960). *Eléments de linguistique générale*. Paris : Armand Colin.
- Miller, G.A. (1951). *Language and Communication*. New York : CGraw-Hill. (tr. fr. 1956 *Langage et communication* Paris : PUF)
- Moreau, R. (1964). La théorie de l'information. Ses limites dans l'explication linguistique. *BSL T.59*, xix-xxiii.
- Saussure, F. de [1931, 1972] *Cours de linguistique générale* (ed. Tullio de Mauro). Paris : Payot.

- Shannon, C.E. et W. Weaver. (1949). *The Mathematical Theory of Communication*. Urbana-Champaign : Univ. of Illinois Press. (trad. fr. 1975 *La Théorie mathématique de la communication*. Paris :CEPL).
- Whatmough, J. (1952). Natural Selection in language. *Scientific American* 186, 82-86.
- Zipf G. K. (1949). *Human Behavior and the Principle of Least Effort*. Cambridge Mass : Addison-Wesley.

Sources secondaires :

- Auroux, S. (ed.) (1990) Encyclopédie philosophique universelle, Tome 2 : Les notions. Paris :PUF.
- Chevalier J-C. avec P. Encrevé (2006). *Combats pour la linguistique, de Martinet à Kristeva*. Lyon : ENS Editions.
- Fehr, J. (2003). Interceptions et interférences: la notion de "code" entre cryptologie, télécommunications et les sciences du langage. Selected papers from the Eighth International Conference on the History of the Language Sciences, Sylvain Auroux (ed.) Amsterdam & Philadelphia : John Benjamins Publishing Company, 363-372.
- Léon J. (1998). Les débuts de la traduction automatique en France (1959-1968): à contretemps? *Modèles Linguistiques* 9-2, 55-86.
- Léon J. (2004). Lexies, synapsies, synthèmes: le renouveau des études lexicales en France au début des années 1960 *History of Linguistics in Texts and Concepts. Geschichte der Sprachwissenschaft in Texten und Konzeptionen*, Gerda Hassler (ed.) Münster: Nodus Publikationen vol I, 405-418.
- Mindell D., Segal J., Gerovitch S. (2003). From communication engineering to communications science : Cybernetics and Information Theory in the United States, France and the Soviet Union. *Science and Ideology: A Comparative History*, Mark Walker (ed.), London : Routledge, 66-96.
- Puech C. (ed.) (à paraître) Les structuralismes linguistiques. Problèmes d'historiographie comparée, *Les dossiers d'HEL* n°3.
- Rey, A. (1992). *Dictionnaire historique de la langue française*. Paris : Dictionnaires Le Robert.
- Segal, J. (2003). *Le Zéro et le Un. Histoire de la notion scientifique d'information au 20ème siècle*. Paris : Editions Syllepse.
- Verleyen S. (2007). Le fonctionnalisme entre système linguistique et sujet parlant : Jakobson et Troubetzkoy face à Martinet. *Cahiers Ferdinand de Saussure* 60, 163-188.
- Verleyen S., Swiggers P. (2006). Causalité et conditionnement dans le fonctionnalisme diachronique. *Folia Linguistica Historica* 27/1-2, 171-195.

¹ Ce projet de recherche est développé en relation avec le programme d'ANR DETCOL (Développement et exploitation textuelle d'un corpus d'œuvres linguistiques) dirigé par Bernard Colombat. Dans ce cadre, seront examinées dans des études ultérieures la réception de la théorie de l'information chez d'autres linguistes.

² Sur les rapports entre cybernétique, théorie de la communication et théorie de l'information, voir Segal (2003).

³ Pour une étude comparée de la réception de la théorie de l'information dans les sciences, voir Segal (2003) et Mindell et al. (2003).

⁴ A noter que Shannon lui-même propose d'utiliser les analyses statistiques de Zipf à des fins linguistiques. Se fondant sur l'idée de la redondance de la langue anglaise, il montre qu'il est possible de calculer l'entropie d'une langue à partir d'études statistiques sur la fréquence avec laquelle les éléments constitutifs d'une phrase sont sélectionnés.

⁵ L'allusion au codage de la pensée dans le langage suggère que Benveniste rassemble sous le pluriel 'les théories de l'information' à la fois la cybernétique, la théorie de la communication et la théorie de l'information. Ce point reste à explorer.

⁶ Ce point devrait faire l'objet d'une étude spécifique.

⁷ Dans son étude sur Martinet et l'École de Prague, Verleyen (2007) insiste davantage sur l'influence de Zipf comme psychologue plutôt que comme statisticien. Il montre que Jakobson et Troubetzkoy conçoivent la langue en diachronie comme une totalité organique dans laquelle l'influence des locuteurs est relativement faible, alors qu'au contraire Martinet envisage la systématisme du changement en termes de propriétés du sujet parlant. Ce qui explique, dit Verleyen, le recours de Martinet au psychologue Zipf, dont le principe du moindre effort tente de rendre compte du comportement humain en général.

⁸ On ne sait si Martinet connaissait les travaux du philologue de Harvard Joshua Whatmough parus dans le *Scientific American* en 1952 ; celui-ci tente d'expliquer l'évolution des langues par la notion de variation sélective fondée sur une formule définissant la quantité d'information.

⁹ Dans la bibliographie sommaire présente dans l'édition de 1970, l'ouvrage de Jakobson et Halle (1956) n'apparaît que comme référence aux conceptions binaristes et aprioristes de la phonologie, critiquées par Martinet.

¹⁰ Chevalier (2006) souligne bien le rôle de passeur de Dubois dans l'introduction des théories linguistiques américaines en France dans les années 1960.

¹¹ Alors que, dans sa thèse, Dubois cite l'ouvrage de G.A. Miller *Language and Communication* (1951, tr.fr.1956), qui est une introduction anti-behaviouriste à la théorie de l'information destinée aux étudiants de psychologie, il ne reprend pas cette référence dans ses travaux ultérieurs.

¹² Nous ne prétendons pas épuiser ici la relation de Jakobson avec la théorie mathématique de la communication qui est bien plus vaste, notamment en ce qui concerne le domaine de l'interlocution et du dialogisme.