

Henri Boyer

Sociolinguistique : faire *corpus* de toute(s) voix ?

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Henri Boyer, « Sociolinguistique : faire *corpus* de toute(s) voix ? », *Mots. Les langages du politique* [En ligne], 69 | 2002, mis en ligne le 14 mai 2008. URL : <http://mots.revues.org/10553>

DOI : en cours d'attribution

Éditeur : ENS Éditions
<http://mots.revues.org>
<http://www.revues.org>

Document accessible en ligne sur : <http://mots.revues.org/10553>

Ce document est le fac-similé de l'édition papier.

© ENS Éditions

Henri BOYER^o

Sociolinguistique : faire *corpus* de toute(s) voix ?¹

Lorsqu'on observe les discours tenus (en France) au sein des sciences du langage, depuis maintenant plus d'un quart de siècle, sur la notion-concept de *corpus*, on ne peut qu'être frappé par l'évolution de son statut théorique et méthodologique et de la nature de la chose à laquelle elle fait référence. En première approximation, on peut dire que si dans le discours structuraliste des années 1960-1970, les corpus étaient des objets clos et homogènes pour les besoins de la démonstration, le discours sociolinguistique des années 1970-1990 pose la nécessité de corpus ouverts et forcément hétérogènes, si l'on prend au sérieux les pré-supposés laboviens.

Ainsi pour A. Greimas (1966, p 143), « un corpus, pour être bien constitué, doit satisfaire à trois conditions : *être représentatif, exhaustif et homogène* ». Et ce point de vue semble parfaitement représentatif des discours tenus sur les corpus dans années 1960. C'est à propos de la condition d'*homogénéité* que la sociolinguistique réagit le plus clairement, à l'instar de J.-B. Marcellesi et B. Gardin (1974, p. 240) qui s'interrogent ainsi :

Homogénéité ? C'est peut-être en linguistique sociale le concept le plus difficile à utiliser puisque l'étude des contrastes discursifs exclut l'homogénéité.

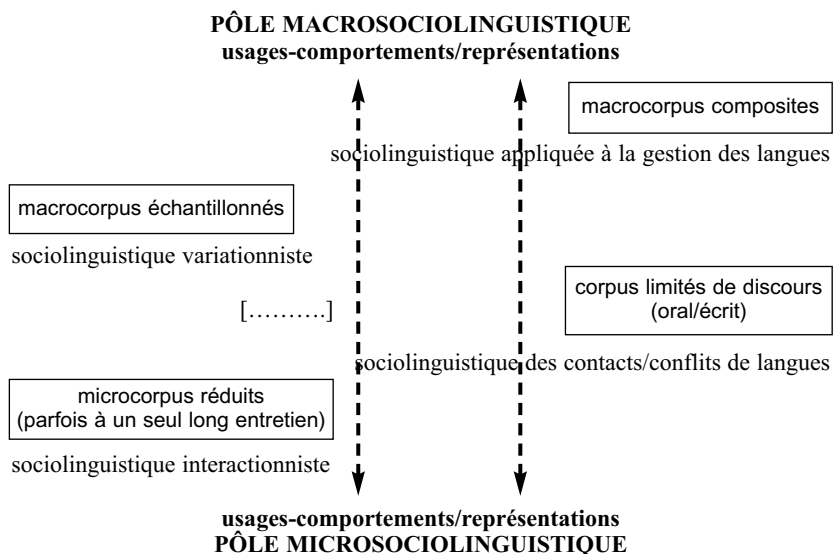
Ou encore de J. Boutet (citée dans D. Lehmann, 1985, p. 15) :

^o ARSer/DIPRALANG, Université Montpellier 3, Route de Mende, 34199 Montpellier Cedex 5.

1. Cet article et le suivant de J.-P. Honoré proviennent d'une journée d'étude organisée le 30 novembre 2001 à Montpellier par la Société des Amis de Mots, en collaboration avec les équipes de recherche DIPRALANG et PRAXILING. La journée portait sur la diversité des approches des corpus en analyse de discours.

La façon dont les structuralistes définissent les propriétés des corpus [...] est bien révélatrice des précautions prises au départ afin d'évacuer ce qui risque de compromettre l'élaboration d'un système linguistique uniforme : le corpus doit être homogène et représentatif. Représentatif de quoi, si ce n'est de l'idée que se font a priori les linguistes du système linguistique à établir. Quant à l'homogénéité, l'ensemble des travaux sur la variation linguistique a montré que ce n'est pas une propriété évidente des langues.

En fait, la linguistique structurale travaillait à partir d'un usage normé, alors que la sociolinguistique prend en compte *des* usages concurrents. Mais au sein même de la démarche sociolinguistique, force est de constater que le statut du corpus a considérablement évolué depuis l'émergence et l'expansion du *pôle macrosociolinguistique*, de la sociolinguistique variationniste en premier lieu, jusqu'au développement du *pôle microsociolinguistique*, plus particulièrement de la sociolinguistique interactionniste, selon la figuration schématique ci-après (H. Boyer, 2001, p. 17) :



Pour F. Gadet (2000, p. 72), faisant référence à Franceschini, on est passé « d'une phase objectivante qu'est le paradigme variationniste, à une période interactive, sous l'influence de l'ethnométhodologie, de l'ethnographie de la communication et de l'anthropologie sociale ».

Cependant, une curiosité métalinguistique mérite d'être relevée : dans l'ouvrage *Sociolinguistique. Concepts de base*, ne figure comme sens du

concept qui nous occupe ici, en particulier à l'entrée *corpus* (M.-L. Moreau (dir.), 1997, p. 102) que celui qui a cours en sociolinguistique appliquée (concernant les politiques linguistiques), en référence à la conceptualisation de H. Kloss : *corpus* (dans le cadre d'une politique qui vise l'aménagement intra-linguistique de la langue) est opposé à *statut* (la politique vise précisément le statut social/extra-linguistique de la langue). Serait-ce parce que le concept de *corpus* au sens de « collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques et *extra-linguistiques* explicites pour servir d'échantillon d'*emplois déterminés d'une langue* » (B. Habert, 2000, p. 13) est constitutif du paradigme sociolinguistique ? Ou s'agit-il d'un parti pris rédactionnel ? Je pencherais plutôt pour la première interprétation, qu'on pourrait aisément vérifier ailleurs, me semble-t-il, car tout indique que la préoccupation principale, dans le discours des sociolinguistes, n'est pas de théoriser l'exigence et la nature des corpus, mais d'exposer les procédures et les conditions du recueil des productions constitutives des corpus². On en trouve un remarquable exemple dans P. Singuy (1996), à propos d'une enquête labovienne sur les représentations du français parlé en Pays de Vand (Suisse romande). Mais on en trouve bien d'autres exemples comme dans L.-J. Calvet et P. Dumont (dir., 1999), où, entre autres, J. Bres s'intéresse à « L'entretien et ses techniques » (J. Bres, 1999), B. Maurer aux « Méthodes d'enquêtes [...] effectivement employées aujourd'hui en sociolinguistique (B. Maurer, 1999) », tandis que Ch. Deprez traite des « enquêtes "micro" » et livre au passage quelques recommandations pratiques (Ch. Deprez, 1999, p. 92) :

L'entretien dure souvent plus d'une heure. Il doit être enregistré et minutieusement transcrit *in extenso* en indiquant bien les pauses, les hésitations, les bredouillis, les lapsus, toutes ces « scories » et ces « ratés » de la communication qui prennent un sens dans l'analyse de discours que le linguiste fera ensuite à partir de cette transcription.

Cependant, on trouve également des remarques concernant le corpus et son volume, en particulier chez Labov, parfois même un peu décalées. Ainsi dans W. Labov 1976, le sociolinguiste prêche à la fois la modération (p. 283) : « On s'aperçoit que [...] les structures fondamentales de la stratification par classe se dégagent d'échantillons aussi restreints que vingt-cinq locuteurs », alors qu'ailleurs (p. 288), l'observation n'a pas

2. Corpus qui peuvent être de nature diverse, comme : entretiens non directs, ou semi-directifs transcrits ; ensembles de documents écrits, imprimés ou manuscrits ; ensembles de productions audio-visuelles ; questionnaires ; produits d'observations directes ; produits de séances de groupes ; ensembles composites, intégrant plusieurs des ensembles précédemment cités...

la même teneur : « Notre approche initiale de la communauté linguistique est dirigée par la nécessité d'obtenir *un volume important* de discours naturel correctement enregistré » (c'est moi qui souligne).

De plus en plus, lorsqu'il est question de prôner des « méthodes susceptibles de tenir compte de l'échange interactif », il n'est pas rare qu'on s'interroge « sur la position du chercheur » et qu'on en vienne à considérer que « sur le terrain, l'expert ne saurait être le maître » (F. Gadet 2000, p. 72-73).

Mais la grande nouveauté, en ce qui concerne notre problématique, c'est à n'en pas douter la montée en puissance de la sociolinguistique interactionniste dont il a déjà été question ici, et d'une « recherche [...] qui se caractérise [...] par le souci de rendre compte de la réalité, telle qu'elle est (re)construite et interprétée par les sujets eux-mêmes » (G. Lüdi et B. Py, 1995, p. 95 ; voir également X.-P. Rodríguez Yáñez, 1997), ainsi que l'envisage par exemple B. Py dans un texte récent : « Une Mexicaine à Neuchâtel », où il cherche à « établir [...] un inventaire de thèmes affleurant dans un document unique » : pour lui, les récits et impressions recueillis dans l'entretien avec l'enquêtée sont un « effort de représentation et de théorisation effectué par le sujet lui-même au moyen de ses ressources personnelles et en collaboration avec l'enquêteur ». Les « interprétations [de l'enquêteur] se situent donc à un second degré : elles prennent comme objet un premier niveau d'interprétation, celui de l'acteur social lui-même » (B. Py, 2000, p. 71).

Avec ce type de perspective sociolinguistique, le corpus a perdu, en hétérogénéité et en représentativité, ce qu'il a gagné en intensité représentationnelle et en exemplarité : G. Lüdi et B. Py (1995, p. 26) posent ainsi une « *dialectique du particulier et du général* : les propos tenus par un migrant peuvent avoir une signification pour la migration interne suisse dans son ensemble, même s'ils ne sont énoncés que par un individu unique ». C'est dire si le *corpus* revêt aux yeux d'une certaine sociolinguistique une valeur qui ne doit rien à son volume.

Bibliographie

- Bilger M. (dir.), 2000, *Linguistique sur corpus*, Presses Universitaires de Perpignan.
- Boyer H., 2001, *Introduction à la sociolinguistique*, Dunod.

- Bres J., 1999, « L'entretien et ses techniques », dans L.-J. Calvet et P. Dumont (dir.), *L'enquête sociolinguistique*, L'Harmattan.
- Deprez Ch., 1999, « Les enquêtes "micro". Pratiques et transmissions familiales des langues d'origine dans l'immigration en France », dans L.-J. Calvet, P. Dumont (dir.), *L'enquête sociolinguistique*, L'Harmattan.
- Gadet F., 2000, « Derrière les problèmes méthodologiques du recueil des données », dans M. Bilger (dir.), *Linguistique sur corpus*, Presses Universitaires de Perpignan.
- Greimas A. J., 1966, *Sémantique structurale*, Larousse.
- Habert B., 2000, « Des corpus représentatifs: de quoi, pour quoi, comment ? », dans M. Bilger (dir.), *Linguistique sur corpus*, Presses universitaires de Perpignan.
- Labov W., 1976, *Sociolinguistique*, Minuit.
- Lehmann D., 1985, « Des objectifs de la recherche aux objets concrets de l'analyse », *Cahiers du français des années quatre-vingts*, 1.
- Lüdi G., Py B., 1995, *Changement de langage et langage du changement*, Lausanne, L'Age d'Homme.
- Marcellesi J.-B. et Gardin B., 1974, *Introduction à la sociolinguistique*, Larousse.
- Maurer B., 1999, « Quelles méthodes d'enquête sont effectivement employées aujourd'hui en sociolinguistique », dans L.-J. Calvet et P. Dumont (dir.), *L'enquête sociolinguistique*, L'Harmattan.
- Moreau M.-L., (dir.), 1997, *Sociolinguistique. Concepts de base*, Mardaga.
- Py B., 2000, « Une Mexicaine à Neuchâtel », *Bulletin suisse de linguistique appliquée*, 71.
- Rodriguez Yañez X.-P., 1997, « Aléas théoriques et méthodologiques dans l'étude du bilinguisme », dans H. Boyer (dir.), *Plurilinguisme : « contact » ou « conflit » de langues*, L'Harmattan, coll. « Sociolinguistique ».
- Singy P., 1996, *L'image du français en Suisse romande*, L'Harmattan.