

MARIE-MADELEINE BERTUCCI

**Quelle place et quel statut pour les œuvres littéraires dans
les corpus du français contemporain des cités ?**

*The article concerns the modalities of use by literature of the contemporary French of cités and the way sociolinguists can use these sources to build a corpus of linguistic usage of this type. The question is how to take into account the forms proposed in these works and if it is necessary to treat them in the same way as the forms collected in natural context and which one can find in oral corpora. At first, we shall try to show under what conditions these works can be considered as sources for sociolinguistics and within what limits. Thus, we shall ask the question whether text sociolinguistics is possible at all. Secondly, the article will come up with a hypothesis on the basis of the study of an extract of Magyd Cherfi's *Ma part de Gaulois* (2016) according to which such a source needs to be analysed with criteria other than those of sociolinguistics in a narrow sense, and requires diverse methodologies that are necessary for the understanding of the multilingual identities and the otherness embodied in these linguistic practices.*

L'article porte sur les modalités d'utilisation par la littérature du français contemporain des cités et sur l'usage que les sociolinguistes peuvent faire de ce type de sources pour la construction d'un corpus de pratiques langagières des cités. On peut donc se demander comment prendre en compte les formes proposées dans ces œuvres et s'il faut les situer sur le même plan que les formes recueillies en *contexte naturel* (Cappeau, Gadet, 2007) et figurant dans les corpus oraux. Dans un premier temps, on tentera de montrer à quelles conditions ces œuvres peuvent être considérées comme des sources pour la sociolinguistique et dans quelles limites. On posera donc la question de savoir si une sociolinguistique du texte est possible. Dans un second temps, l'article fera l'hypothèse à partir de l'étude d'un extrait de corpus, issu du récit de Magyd Cherfi, *Ma part de Gaulois* (2016), qu'une telle source nécessite d'être appréhendée avec d'autres critères que ceux de la sociolinguistique stricto sensu et qu'elle nécessite de croiser des démarches et des méthodologies diverses.

I. Une sociolinguistique du texte est-elle possible ?

I.1 Passer des données orales collectées sur le terrain au texte.

L'importance pour la sociolinguistique des *données orales* issues de *contextes naturels* a été soulignée à propos des grands corpus :

Une double « évidence » est aujourd'hui partagée par la majorité des linguistes : il faut travailler sur l'oral, et le faire sur des corpus, les plus grands possibles. Il serait ainsi urgent de constituer de vastes corpus de données orales (Cappeau, Gadet, 2007 : 99).

Le caractère problématique de la notion de corpus et au-delà de grand corpus pour les sociolinguistes a été souligné, surtout dans la perspective d'une opposition corpus/terrain.

Qu'est-ce qu'un sociolinguiste peut attendre des corpus ? Les frontières des disciplines des sciences humaines étant floues, et en particulier les entours de la sociolinguistique, on peut y mettre des objets que d'autres feraient relever de l'analyse de discours, où la notion de corpus (ou d'archive) est fondatrice. (Cappeau, Gadet, 2007 : 100).

On insistera donc sur la nécessité de construire une réflexion épistémologique sur l'extension du champ de la sociolinguistique et notamment sur la question du terrain et des corpus (Blanchet, 2007 ; Robillard, 2007) pour légitimer l'utilisation de ces sources et surtout pour préciser les modalités de leur utilisation au regard de l'objectif poursuivi ici.

Dans ces conditions, admettre qu'on peut s'appuyer sur des occurrences extraites d'œuvres littéraires suppose de s'interroger au préalable sur le type et la nature des corpus qui en découlent. En effet, certains sociolinguistes considèrent que de telles données ne sont pas représentatives des pratiques langagières visées au même titre que les données orales recueillies en contexte naturel car elles sont extraites de sources écrites, littéraires de surcroît et de ce fait ambiguës. En effet, elles peuvent être considérées comme des données écrites oralisées, mimant dans un but d'expressivité rhétorique ou stylistique des occurrences observables dans un contexte naturel. De plus il s'agit d'un oral donné comme déjà rectifié (Larthomas, 1980) et en conséquence échappant aux aléas des interactions. La position adoptée par Paul Cappeau et Françoise Gadet supra ne semble donc pas conférer aux textes littéraires le statut de source légitime sur laquelle reposeraient des recherches en sociolinguistique.

On peut se demander néanmoins si on ne doit pas accorder aux données issues d'œuvres littéraires un statut spécifique, dont la nature serait *extra-linguistique* (Mayaffre, 2005 : 2), qui justifierait de les retenir (ibid.).

On pourra évoquer les raisons suivantes. On admettra d'une part que le corpus est un « *lieu problématique de rencontre entre la langue et la société ; l'analyste jouant le rôle conscient et déterminant d'entremetteur* » (Ibid.) et d'autre part que la recherche porte moins sur « le mot ou la phrase » que sur « le texte » (Ibid.). On fera donc l'hypothèse qu'il est possible de dessiner les contours d'une sociolinguistique du texte au même titre que certains chercheurs (Bakhtine 1978 ; Rastier, 2001 ; Adam, 1990, 1999), ont revendiqué d'inscrire leurs travaux dans une « linguistique du texte » (Mayaffre, 2005 : 2). Dès lors les corpus textuels peuvent intéresser la sociolinguistique et ce d'autant plus que la notion de contexte occupe une place de premier plan dans ce champ des sciences du langage, en considérant que si « *la signification se construit en langue [...] il n'y a de sens [...] qu'en contexte* » (Ibid.).

I.2 La question de la structuration interne des corpus : homogénéité versus hétérogénéité

On s'interrogera maintenant sur la façon dont se structurent les corpus et sur les relations que chaque unité du corpus entretient avec les autres. On fera l'hypothèse qu'un corpus n'est pas qu'une entité statique, qu'il n'est pas qu'un « réceptacle » (Mayaffre, 2005 : 3), qu'il est au contraire une structure dynamique et que chaque unité du corpus est en interrelation avec les autres. On a pu écrire à ce sujet que :

Tout texte placé dans un corpus en reçoit des déterminations sémantiques, et modifie potentiellement le sens de chacun des textes qui le composent (Rastier, 2001 : 92 cité par Mayaffre, 2005 : 3).

On peut supposer que l'intégration de textes littéraires dans les corpus des pratiques langagières des cités et leur confrontation à des données orales recueillies en contexte naturel se justifie si on admet que le corpus est un « lieu d'invention » (Mayaffre, 2005 : 3), qu'il est par nature un processus « dynamique » ; « réflexif » (Ibid.), voire une « unité multiplexe » (Morin, 1990) et non pas un ensemble amorphe. Entrant dans une relation dialectique avec les autres textes du corpus, il ne se limite pas à une simple fonction de « validation » (Mayaffre, 2005 : Ibid.) d'une hypothèse de travail car il est aussi « un lieu théorique » (Ibid.). Cette proposition valide et légitime l'hétérogénéité inhérente à

un corpus de ce type, notamment en terme de référentialité. La référentialité des corpus où figurent des textes littéraires est rendue problématique par leur littérarité. On fera donc l'hypothèse qu'un tel corpus n'est pas référentiel comme peuvent l'être par exemple les corpus politiques (Mayaffre, 2005 : 4) ou qu'en tout cas sa référentialité est sujette à caution et peut être critiquée, sauf si l'on admet l'existence d'une dimension discursive du monde social.

I.3 La dimension discursive du monde social

On empruntera à l'histoire culturelle la notion de « dimension discursive du monde social » (Schaub, 2001 : 981), qui permettra de souligner « *l'emprise des schèmes cognitifs et expressifs du récit sur l'écriture de l'histoire* » (Ibid.)¹. L'anthropologie, la sociologie narrative, les cultural studies anglo-saxons ont des orientations analogues. Ces disciplines ont en commun de s'intéresser au récit dont la nature protéiforme est suffisamment souple pour s'adapter à des projets aussi distincts. Ainsi de même que les historiens sont très attentifs « *aux modes d'énonciation des phénomènes historiques* » (Schaub, 2001 : 982), les sociolinguistes peuvent gagner à prendre en compte le type d'énonciation des textes littéraires concernés ainsi que les contraintes génériques pesant sur cette dernière. Une telle démarche pourrait contribuer à conférer aux formes narratives le statut de source et à authentifier les pratiques langagières issues de ces textes au même titre que des pratiques langagières authentiques observables dans des interactions orales contextualisées. La question est de savoir à quelles conditions.

II. À quelles conditions l'œuvre littéraire peut-elle constituer une source ? Comment construire une analyse pluridisciplinaire de tels corpus ?

II.1 Les conditions de l'analyse pluridisciplinaire

La difficulté est de savoir quelles relations il convient de construire au-delà des sciences du langage, avec d'autres disciplines (Cappeau, Gadet, 2007 : 100), notamment la littérature, l'histoire, l'anthropologie, la sociologie, les cultural studies... afin de considérer les œuvres littéraires comme des sources à

¹ Voir le projet collaboratif (2013-2015) « Raconter la vie » ayant pour objet d'écrire « Le roman vrai de la société d'aujourd'hui » de Pierre Rosanvallon et la collection du même nom au Seuil fermée en 2017 ; ou l'ouvrage intitulé *L'Histoire est une littérature contemporaine. Manifeste pour les sciences sociales* (Jablonka, 2014).

part entière. Une solution possible est d'admettre leur littérarité, à savoir qu'elles constituent des formes marquées stylistiquement, qui s'inscrivent dans un projet d'écriture bien déterminé, doté parfois d'une dimension autonymique, comme l'indiquent les commentaires métalinguistiques du narrateur de *Ma part de Gaulois* : « Il parlait comme tous les Manouches à la première personne du pluriel » (2016 : 94) ; « C'est intéressant, c'est du langage et il en vaut un autre » (2016 : 96). Ces commentaires concernent la morphologie verbale, et en particulier les désinences : « Tu tournons l'dos et y te plante » (p. 91) ; « Vous f'zons des photos » (p. 92). La morphologie possède dans ce cas une valeur mimétique qui contribue à la création d'un effet de réel². Cette valeur mimétique relève de la fonction « mathésique » (Adam, Petitjean, 1995) du récit réaliste, laquelle permet de « disposer à l'intérieur du récit les savoirs de l'auteur » (Adam, Petitjean, 1995 : 26). Néanmoins, l'identification du type de récit auquel renvoie ce texte connaît un certain caractère d'indécidabilité. On fera l'hypothèse qu'il constitue une autofiction. La mention *récit* figure en exergue et l'auteur ne dit nulle part qu'il s'agit d'une autobiographie. L'autofiction se situe « entre deux pactes » de lecture (Delangue, 2014 : 138). Magyd Cherfi a recours à « la forme littéraire pour parler de sa vie »³ (Ibid.). Il « romance sa vie, ce n'est donc pas une autobiographie à proprement parler, mais bien une autofiction » (Ibid.). Cette indécidabilité générique met en question le caractère authentique des données recueillies dans *Ma part de Gaulois*, sauf si on admet de critiquer la notion d'authenticité du fait de l'hétérogénéité d'un tel corpus. On reprendra ce point en conclusion en proposant des pistes pour l'utilisation de ce type de sources.

II.2 Le corpus

Le corpus est constitué de relevés issus du récit de Magyd Cherfi, *Ma part de Gaulois* (2016).

² On citera également l'occurrence du pronom démonstratif « Pas comme *célass* » (ceux-là p. 92).

³ Magyd Cherfi a été le parolier et le chanteur du groupe toulousain de rap, *Zebda*, avant de commencer une carrière de chanteur en solo. Il a vécu à Toulouse dans la cité des Isards qu'il évoque dans son récit. Voir sur ce point l'article en ligne « Le livre de Magyd Cherfi, l'ex-leader de Zebda, a du mal à passer dans son quartier d'enfance » in : Rédaction Europe1.fr (31-10-2016), <http://www.europe1.fr/culture/le-livre-de-magyd-cherfi-lex-leader-de-zebda-a-du-mal-a-passer-dans-son-quartier-denfance-2887310> (page consultée le 15.04.18).

Les occurrences du texte peuvent être classées pour le lexique à partir des catégories élaborées par J.-P. Goudaillier (2001), qui sont les archaïsmes, les faits d'argot classique, les termes d'origine tzigane ou à consonance tzigane (finale en ave). Les termes relevés ne figurent pas en revanche dans cet ouvrage. On les trouve dans le dictionnaire d'Aristide Bruant et de Léon de Bercy, *L'argot au XX^e siècle*.

Archaïsmes	<i>Le gueux</i> ⁴ (p. 100)
Argot classique	« Je crois ne pas avoir <i>moufté</i> ⁵ ... » (p. 37) « Tu l'as pas <i>affranchi</i> ⁶ , le poète » (p. 68)
Termes d'origine tzigane ou à consonance tzigane	« Je peux pas les <i>dicave</i> (respirer ⁷) » (p. 75) « je l'ai déjà <i>guintchave</i> (sollicitée ⁸) » (p. 91)

On relève également des faits de contacts de langue(s), qui permettent d'évoquer la dimension plurilingue et pluriculturelle de ces pratiques, qui se manifeste par une incrustation d'énoncés ou de termes à consonance arabe et kabyle ou encore tzigane, selon l'auteur, avec un effet de traduction et de reproduction des accents résultant d'un jeu stylistique sur les graphies et d'effets de traduction comme l'indiquent les occurrences suivantes.

Incrustation d'énoncés à consonance arabe et kabyle	« <i>Ti va à la sucritti, ti monti troi mitaj et ti donn li cachi di disser</i> ». « (Traduction.) Tu vas à la Sécu, tu montes trois étages et tu donnes le casier judiciaire ! » ; « " <i>Li cachi di disser</i> ", m'a fallu quinze ans pour comprendre qu'il s'agissait du "casier judiciaire" ! » (p. 23). « Que Dieu te brûle la langue ! (En kabyle bien sûr, <i>Aké seurgh reubi ilsikh</i> ! ») (p. 79).
Incrustation d'un terme à consonance tzigane	Ce que t'as <i>tchouré</i> ⁹ la veille (p. 94).

⁴ Voir *L'argot au XX^e siècle*.

⁵ Ibid.

⁶ Ibid.

⁷ La traduction est donnée par l'auteur. Le dictionnaire de Jean-Pierre Goudaillier indique le sens de « regarder, voir » ; *Tout l'argot des banlieues. Le dictionnaire de la zone en 2600 définitions* également. Les deux dictionnaires attestent de l'origine tzigane du terme.

⁸ La traduction est de l'auteur. Le terme n'est pas présent dans les deux dictionnaires consultés.

On observe enfin d'autres faits de marquage stylistique qui sur le plan sociolinguistique renvoient à des traits perçus comme caractéristiques des cités urbaines sensibles et fortement altéritaires. Il s'agit d'éléments renvoyant à l'ethnicité.

L'ethnicité se manifeste par des effets de variation sur les dénominations des différentes catégories ethniques. Le substantif *Arabe* semble fonctionner comme un collectif non marqué : « *Nous les Arabes* » (p. 93), qui s'oppose à des variantes relevant de variétés non normées de français, argotiques pour *bicots* (p. 91) ; *bougnoules* (p. 112), identifiées comme caractéristiques du français des cités pour *Reubeu* (verlan de beur, lui-même verlan de arabe ; p. 96). Pour les Gitans, on trouve *gitan* (p. 92) ; *manouche* (p. 94). Pour tous les non Gitans, on note le générique *raclos* : « *vous, les raclos* » (p. 92). Le juron *sa race* : « *Sa race, là si j'y vais c'est la lache* » (honte ; p. 91) constitue un marqueur d'ethnicité emblématique de par sa valeur générique totalisante.

La densité de ces occurrences augmente au cours de dialogues où sont présents des locuteurs censés y avoir recours de par le statut dans le récit que leur confère le narrateur, comme on le voit dans le dialogue entre le narrateur et le locuteur gitan (p. 90-95) notamment. Ce dialogue compte une fréquence plus importante de termes marqués que le reste du texte. On relève ainsi *daronne* (p. 91) ; *lové* (verlan de vélo ; p. 91) ; *guintchave* (sollicitée ; p. 91) ; *lache* (honte ; p. 91) ; *poucave* (trahir ; p. 95) ; *schmitts* (policiers ; p. 94-95), *dicave* (regarde ; p. 98). Il a pour fonction de rendre vraisemblable le récit et il compte au nombre des « *procédures d'accréditement du monde représenté* » (Adam, 1985 : 14) tant « *au niveau des exigences du vraisemblable* », qu'« *au niveau de la dimension symbolique* » (Ibid.) en ce qui concerne la « *rhétorique du sens* » visée (Ibid. : 15).

Pour conclure la présentation de cet extrait de corpus, on dira que le français des cités, dans ce texte, participe de la création d'un effet de réel et qu'il s'inscrit dans l'ordre de la mimesis (Auerbach, 1946). C'est la raison pour laquelle on fera l'hypothèse que les informations sur le français des cités extraites de ce type de corpus doivent être complétées par des données orales et qu'elles ne peuvent se substituer à celles qui sont recueillies sur le terrain pour

⁹ On fera l'hypothèse que *tchouré* est un dérivé du romani *chourav* (Goudaillier, 2001 : 19). *Chouraver*, aligné sur la morphologie de l'infinitif des verbes du premier groupe en -er est attesté dans *Le Petit Robert 2015*. Cette liste est ouverte et admet l'entrée de verbes nouveaux.

les raisons qu'on a tenté de mettre en évidence. Néanmoins, on plaidera pour les conserver, en les situant sur un autre plan plus « global » (Mayaffre, 2005 : 6) et en particulier celui d'une « sémiotique des cultures » (Rastier, 2001, cité par Mayaffre, 2005 : 6). Ceci suppose d'admettre que le corpus de pratiques langagières extrait des textes littéraires renvoie à une réalité sociale, qui nécessite pour être comprise de faire appel à d'autres disciplines.

Les occurrences observables dans l'autofiction de Magyd Cherfi sont précieuses pour la mise en évidence et la compréhension des identités plurielles et plurilingues. Ce type de récit permet de représenter l'altérité et de l'attester. Il affirme, de ce fait, ce que Bakhtine appelle la relation exotopique (1978), laquelle rend possible de poser en tant que telle l'existence de l'Autre. L'autofiction, en donnant à entendre l'identité narrative du sujet fait émerger la conscience de soi et par l'énonciation constitue ce que Ricœur appelle la « médiation existentielle entre soi et le monde » (1990 : 178). Les pratiques langagières représentées ici peuvent être analysées, dans la perspective d'une sémiotique globale des cultures relevant des médiations existentielles évoquées plus haut. C'est en ce sens et en faisant l'hypothèse que la notion de dimension discursive du monde social est pertinente pour cette approche, qu'on pourra les considérer comme des sources représentatives des pratiques langagières de locuteurs du français contemporain des cités.

Bibliographie

- ADAM Jean-Michel (1985), *Le texte narratif*, Paris, Nathan.
- ADAM Jean-Michel, PETITJEAN André (1995), *Le texte descriptif*, Paris, Nathan (1^e éd. 1989).
- AUERBACH Erich, (1946) (traduction française (1968)), *Mimesis, la représentation de la réalité dans la littérature occidentale*, Paris, Gallimard.
- BAKHTINE Mikhaïl (1978), *Esthétique et théorie du roman*, Paris, Gallimard.
- BLANCHET Philippe (2007), « Sur le statut ethnologique de la notion de corpus », in : *La mise en œuvre des langues dans l'interaction* (Michèle Auzanneau), Paris, L'Harmattan, p. 341-352.
- BRUANT Aristide, BERCY Léon de (1901), , *L'argot au XX^e siècle*. Édition inversée et raisonnée du dictionnaire français-argot (1901 et 1905) par Denis Delaplace, Paris, Classiques Garnier. (Rééd. 2009)
- CAPPEAU Paul, GADET Françoise (2007), « L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale » in : *Revue française*

de linguistique appliquée (Volume XII), p. 99-110.

<https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2007-1-page-99.htm>

CHERFI Magyd (2016), *Ma part de Gaulois*, Paris, Actes Sud.

DELANGUE Henri (2014), « Autobiographie ou autofiction chez Amélie Nothomb ? » in : *Cedille*, 10, p. 129-141.

<https://cedille.webs.uill.es/10/08delangue.pdf>

GOUDAILLIER Jean-Pierre (2001), *Comment tu tchatches. Dictionnaire du français contemporain des cités*, Paris, Maisonneuve et Larose. (1^e éd. 1997)

JABLONKA Ivan (2014), *L'Histoire est une littérature contemporaine. Manifeste pour les sciences sociales*, Paris, Seuil.

LARTHOMAS Pierre (1980), *Le langage dramatique*, Paris, PUF (1^e éd. 1972).

MAYAFFRE Damon (2005), « Les corpus politiques : objet, méthode et contenu. Introduction » in : *Corpus*, [En ligne], 4. <http://corpus.revues.org/292>

MORIN Edgar (1990), *Introduction à la pensée complexe*, Paris, ESF.

RASTIER François (2001), *Arts et sciences du texte*, Paris, PUF.

RICEUR Paul (1990), *Soi-même comme un autre*, Paris, Seuil.

ROBILLARD Didier de (2007), « Lyophiliser ou interpréter l'ornithorynque ? Corpus, métonymie, métaphore », in : *La mise en œuvre des langues dans l'interaction* (Michèle Auzanneau), Paris, L'Harmattan, p. 353-370.

SCHAUB Jean-Frédéric (2001), « Une histoire culturelle comme histoire politique (note critique) », in : *Annales. Histoire, Sciences Sociales*, 2001/4, p. 981-997. <https://www.cairn.info/revue-annales-2001-4-page-981.htm>

Sociologie narrative : www.sociologienarrative.com

TENGOUR Abdelkarim (2013), *Tout l'argot des banlieues. Le dictionnaire de la zone en 2600 définitions*, Paris, Éditions de l'Opportun.

MARIE-MADELEINE BERTUCCI

Université de Cergy-Pontoise

Courriel : marie-madeleine.bertucci@u-cergy.fr