HAL open science

”Du corpus au terrain”

Marie-Madeleine Bertucci

To cite this version:

Marie-Madeleine Bertucci. ”Du corpus au terrain”. Journée d’étude des doctorants et des jeunes docteurs en sciences du langage - sociolinguistique et didactique –, Marie-Madeleine Bertucci, avec le soutien de l’institut d’études avancées de l’université de Cergy-Pontoise, laboratoire LT2D, Jun 2019, Cergy-Pontoise, France. hal-03704418

HAL Id: hal-03704418

https://cyu.hal.science/hal-03704418v1

Submitted on 29 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Marie-Madeleine Bertucci
Université de Cergy-Pontoise
EA 7518 LT2D

Communication à la journée d'étude[1] des doctorants et des jeunes docteurs en sciences du langage - sociolinguistique et didactique – du 26 juin 2019

-

Maison Internationale de la Recherche de CY Cergy Paris université

Du corpus au terrain

La question est de savoir si le corpus est une entité statique ou dynamique, s'il se suffit à lui-même ou s'il a besoin d'un arrière-plan qui le documente. Dans un premier temps, on se demandera, à cet effet, si le corpus est un ensemble de données textuelles (ou non) et dans un second temps si le corpus peut être défini comme un lieu empirique, un contexte. Dans un troisième temps, on conclura cette approche en précisant les enjeux de la notion de corpus pour la sociolinguistique, compte tenu du fait que la sociolinguistique est une discipline de terrain, ce qui confère un statut secondaire au corpus par rapport au terrain et conduit à redéfinir la notion de données dans la perspective d'une « construction interprétative du monde social » (Blanchet, 2007) réflexive et altéritaire.

I. Le corpus : un ensemble de données textuelles (ou non) ?

Le mor corpus est polysémique comme on va tenter de le montrer en partant de l'ensemble des définitions de ce terme proposées par Ortolang[2] dans le cadre du CNRTL.

Définitions du terme corpus proposée par le CNRTL[3].

En philologie et sciences humaines, le terme évoque un :

I.1 Un recueil

« *PHILOL., SC. HUM.* Recueil réunissant ou se proposant de réunir, en vue de leur étude

---

[1] Colloque financé par l'Institut d'études avancées de l'université de Cergy-Pontoise. This workshop is sponsored by the Institute for Advanced Studies, University of Cergy-Pontoise.

[2] Ortolang est un acronyme de Outils et Ressources pour un Traitement Optimisé de la LANGue. ORTOLANG est un équipement d'excellence validé dans le cadre des investissements d'avenir. Son but est de proposer une infrastructure en réseau offrant un réservoir de données (corpus, lexiques, dictionnaires, etc.) et d'outils sur la langue et son traitement clairement disponibles et documentés.
Le Centre national de ressources textuelles et lexicales (CNRTL) a été créé en 2005 par le CNRS, le CNRTL fédère au sein d'un portail unique, un ensemble de ressources linguistiques informatisées et d'outils de traitement de la langue. Le CNRTL intègre le recensement, la documentation (métadonnées), la normalisation, l'archivage, l'enrichissement et la diffusion des ressources.

[3] Rubrique lexicologie. http://www.cnrtl.fr/definition/corpus

scientifique, la totalité des documents disponibles d'un genre donné, par exemple épigraphiques, littéraires, etc.
L'épigraphie fut sans doute la première, et en tout cas la plus heureuse, à réaliser des « corpus » et à donner l'exemple (*Hist. et ses méthodes,*1961, p. 473) :
1. Il serait vivement souhaitable qu'un « **corpus** » de catalogues des bibliothèques médiévales fût publié, reprenant les éditions anciennes d'une façon scientifique... *La Civilisation écrite,*1939, p. 5402.
♦ *P. anal.* Répertoire scientifique :
2. 1893. La tendance actuelle est aux recueils généraux. Citons : en Allemagne, **Corpus** *des monnaies de Dacie, de Thrace, de Macédoine,* ouvrage collectif entrepris par l'Académie des Sciences de Berlin ; en France, *Recueil général des monnaies grecques d'Asie mineure* patronné par l'Académie des Inscriptions. *La Civilisation écrite,* 1939, p. 4816.

– *Spécialement*
1. *DR.* Recueil, collection du droit romain. *Le corpus juris,* p. abrév., *le corpus.*
2. *ÉLECTRON., INFORM.* Ensemble de données exploitables dans une expérience d'analyse ou de recherche automatique d'informations. *Perforation de corpus.* »

I.2 Un ensemble de données structurées en fonction d'un principe d'organisation

Le *Dictionnaire de linguistique* de Dubois et al propose la définition suivante :

> [Un corpus pour] un parler donné réunira d'abord ou fera réunir un grand nombre d'énoncés produits librement ou sur incitation des enquêteurs. Beaucoup de ces énoncés pourront fort bien n'avoir aucun rapport avec la recherche et ne contenir aucun des mots qui intéressent le linguiste. La totalité des énoncés recueillis est l'univers. A partir de l'univers des énoncés réunis, disons en vrac, le linguiste trie les énoncés qu'il va soumettre à l'analyse : dans le cas qui nous intéresse ce pourra être l'ensemble des phrases, ou groupes de phrases, comprenant des mots présentant tel trait phonétique ou bien une terminaison ou une origine étrangère. Ce sont uniquement ces segments d'énoncés qui seront soumis à l'analyse et qui constitueront le corpus. On pourra aussi, sur des bases statistiques, délimiter soit dans l'univers, soit dans le corpus, des passages qui seront soumis à une analyse quantitative : par exemple, une page toutes les dix pages ; les pages ainsi retenues constituent un échantillon du texte. Par hypothèse, on considérera comme échantillon toute partie représentative du tout. Le corpus peut évidemment, si le chercheur le juge utile ou nécessaire, être constitué par l'univers d'énoncés tout entier. De même une analyse quantitative pourra fort bien se passer d'échantillonnage. Le corpus lui-même ne peut pas être considéré comme constituant la langue (il reflète le caractère de la situation artificielle dans laquelle il a été produit et enregistré), mais seulement comme un échantillon de la langue. Le corpus doit être représentatif, c'est-à-dire qu'il doit illustrer toute la gamme des caractéristiques structurelles. On pourrait penser que les difficultés sont levées si un corpus est exhaustif, c'est-à-dire s'il réunit tous les textes produits. En réalité, le nombre d'énoncés possibles étant indéfini, il n'y a pas d'exhaustivité véritable et, en outre, de grandes quantités de données inutiles ne peuvent que compliquer la recherche en l'alourdissant. Le linguiste doit, donc chercher à obtenir un corpus réellement représentatif et écarter tout ce qui peut rendre son corpus non représentatif (méthode d'enquête choisie, […], préjugé sur la langue, etc.), en veillant à éviter tout ce qui conduit à un artefact* (1973 : 129).

Cette acception comprend le corpus comme l'ensemble des énoncés réunis dans une perspective de description d'un fait de langue. Elle peut néanmoins être transposée dans des champs des sciences du langage dont la finalité descriptiviste est moins affirmée comme l'analyse du discours par exemple. Ainsi D. Mayaffre écrit en citant F. Rastier que « l'objet

pertinent du sémanticien n'est donc pas le mot ou la phrase – « artefact de grammairien » (Rastier, 2001 : 30) mais le texte » (Mayaffre, 2005 : 2). Rastier et Mayaffre s'opposent ainsi aux approches traditionnelles de la linguistique de corpus, notamment anglosaxonnes. On renvoie ici aux définitions élaborées par John Sinclair :

> A collection of naturally occurring language text, chosen to characterize a state or variety of a language[4] (Sinclair, 1991).

> A corpus is a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research[5] (Sinclair, 2005).

La définition que propose pour la linguistique le CNRTL est un écho aux précédentes. En effet, dans ce cadre le corpus renvoie à un :

> Ensemble de textes établi selon un principe de documentation exhaustive, un critère thématique ou exemplaire en vue de leur étude linguistique. *Le corpus des textes parus d'un journal, d'une revue; un corpus littéraire; le corpus du vocabulaire français*.

Le passage en revue de ces définitions conduit à considérer que, pour la linguistique, le corpus n'est pas un recueil ou une archive mais qu'il est d'abord un ensemble de données relatives à la / aux langues ou variétés de langues. Cet ensemble est régi par un principe organisateur. Comme on va le voir dans les lignes qui suivent, le corpus n'est pas une archive.

II. Le corpus n'est pas une archive

Rastier (2002 : 3) souligne que « l'archive contient des documents accessibles. Elle n'est pas un corpus, parce qu'elle n'est pas constituée pour une recherche déterminée ». D. Mayaffre à ce sujet écrit que :

> Rastier, dans une démarche différentielle ou contrastive (qu'il préconise d'ailleurs pour la sémantique ; […] distingue *corpus* et *archives*, *corpus* et *document*, *corpus* et *œuvre*, *corpus* et *web*, ou encore de manière problématisée *corpus* et *texte* (Mayaffre, 2011 : 2).

On se limitera à traiter ici de l'opposition corpus / archive(s).

II.1 La distinction entre *corpus* et *archives*

On distinguera deux niveaux dans la définition de l'/ les archive(s), celle du *Petit Robert* et une autre plus spécialisée, empruntée au code du patrimoine.

II.1.1 L'acception commune de la notion d'archive

[4] Un ensemble empirique de textes en langage naturel choisis pour caractériser une langue ou une variété de langue (M.-M. Bertucci traduction).

[5] Un corpus est un ensemble d'occurrences textuelles d'une langue ou d'une variété de langue, présentées sous un format électronique, choisies en fonction de critères externes pour la pertinence de leur représentativité et leur aptitude à constituer une source productrice de données pour la recherche en linguistique (M.-M. Bertucci traduction).

Le *Petit Robert 2015* définit l'archive ainsi : « 1 au pluriel comme un « ensemble de documents anciens rassemblés et classés à des fins historiques » ; 2 au singulier comme « un document conservé aux archives » ; 3. Par extension comme le « lieu où les archives sont conservées, déposées ».

La notion d'archive s'inscrit dans une logique de conservation de documents anciens, de patrimonialisation de ces derniers dans une perspective historique. La définition du *Code du Patrimoine* confère une amplitude plus large à la notion que le *Petit Robert 2015* dans la mesure où elle l'élargit à l'ensemble des documents quels qu'ils soient sans introduire la restriction que construit le *Petit Robert 2015*.

II.1.2 La définition de l'archive selon le *Code du Patrimoine*. La notion de donnée

Le Code du patrimoine définit les archives comme

> l'ensemble des documents, y compris les données, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité (Guyon[6], 2016).

Remarque : il faut comprendre la notion de données au sens de données électroniques qu'on va aborder maintenant. Il convient de noter que le support n'a pas d'impact sur la qualification des archives. Le régime juridique des documents est le même quel que soit le support. La dématérialisation d'un document n'a pas d'incidence sur son statut juridique :

> Le terme de données a été explicitement introduit par la loi LCAP en 2016. Cette insertion permet de réaffirmer que le cadre règlementaire de la gestion des archives demeure et doit demeurer le même, quel que soit le support des documents. Qu'il soit papier ou électronique, un document est porteur d'un ensemble d'informations déterminées par un contexte et un environnement de production. En aucun cas le changement de support n'interfère sur le contenu informationnel : la dématérialisation d'un acte par exemple n'entraîne pas modification des informations que contient cet acte. Par ailleurs, la représentation formelle des informations (sous la forme d'un document, de données dans une base de données ou un flux XML par exemple) ne peut pas être et ne doit pas être un critère pour introduire une distinction, notamment dans le régime juridique, entre les documents (sous- entendu papier) et les données (sous-entendu électroniques). C'est bien tout l'enjeu de l'introduction formelle du terme « données » dans la définition des archives. Il est d'ailleurs intéressant de noter que la CADA (Commission d'accès aux documents administratifs) n'introduit pas non plus de distinction entre les supports ou la forme des informations pour qualifier un document administratif : les documents administratifs sont tous les documents produits par l'administration (Guyon, 2016 : 1).

II.1.3 La polysémie de la notion de données

Le sens du mot *données* qui vient d'être abordé supra ne constitue pas la seule acception et il ne coïncide pas avec le sens que *données* possède en sciences du langage.

Le *Petit Robert* 2015 définit les *données d'une science* comme « ce qui est admis, connu ou reconnu, et qui sert de base à un raisonnement, de point de départ pour une recherche ». Le PR 2015 propose comme synonymes « point de départ, élément, renseignement ».

---

[6] In « Le concept d'archives : d'une définition à l'autre. » 2016. Prépublication. Document de travail accessible dans <hal-01401138>.

Le dictionnaire de Dubois et al. propose comme sens celui de faits langagiers *empiriques* ou *data* pour désigner les *actes de parole* qui constituent ces données (1973 : 167).

Dans une perspective plus générale, on retiendra la définition que l'OCDE[7] donne de la notion :

> Dans le cadre de ces *Principes et Lignes directrices*, les « données de la recherche » sont définies comme des enregistrements factuels (chiffres, textes, images et sons), qui sont utilisés comme sources principales pour la recherche scientifique et sont généralement reconnus par la communauté scientifique comme nécessaires pour valider des résultats de recherche. Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche.
> Ce terme ne s'applique pas aux éléments suivants : carnets de laboratoire, analyses préliminaires et projets de documents scientifiques, programmes de travaux futurs, examens par les pairs, communications personnelles avec des collègues et objets matériels (par exemple, les échantillons de laboratoire, les souches bactériennes et les animaux de laboratoire […]). L'accès à tous ces produits ou résultats de la recherche est régi par d'autres considérations que celles abordées ici. Ces Principes et Lignes directrices portent essentiellement sur les données de la recherche sur support numérique exploitable sur ordinateur (2007 : 18).

Les approches de la notion de corpus qui viennent d'être abordées paraissent présenter le corpus comme un espace statique parce qu'elles théorisent peu la notion de corpus pour elle-même. Cette vision a été critiquée par certains théoriciens de la notion qui le définissent comme un lieu empirique, un contexte, un espace de la construction du sens à condition qu'il existe une réflexivité du corpus (Mayaffre, 2002) *dans une conception philologique-herméneutique* (Rastier, 2002 : 1) qui prenne en compte les interactions qui se construisent dans le corpus d'un texte à l'autre (Ibid.).

Rastier peut alors définir le corpus comme :

> Un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés : (i) de manière théorique réflexive en tenant compte des discours et des genres, et (ii) de manière pratique en vue d'une gamme d'applications (2002 : 2).

Le corpus doit être construit en fonction de critères linguistiques et tendre *idéalement*[8] à l'homogénéité et à la représentativité. De ce point de vue, Rastier précise qu'un corpus n'est pas une *banque textuelle*, laquelle « peut regrouper des textes numériques de statuts divers » (2002 : 2).

Dans ces conditions, le corpus devient l'espace de la construction du sens. Les corpus sont « essentiels parce que non pas seulement « déterminants » mais « constituants » (Rastier, 2011 : 242 cité par Mayaffre, 2011 : 2).

II. 2 Le corpus comme espace de la construction du sens dans la perspective d'une linguistique de corpus dont le texte est l'unité

[7] Organisation de Coopération et de Développement Économiques.
[8] C'est nous qui soulignons.

On fera l'hypothèse qu'un corpus n'est pas une entité statique, qu'il n'est pas qu'un « réceptacle » (Mayaffre, 2005 : 3), mais qu'il est au contraire une structure dynamique et que chaque unité du corpus est en interrelation avec les autres. FR 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).

Linguistique de corpus et linguistique du texte

Au-delà, Damon Mayaffre souligne que FR « semble avoir trouvé dans les corpus le lieu empirique ou la définition pratique du contexte élargi » (2011 : 2) qui lui permettent d'affirmer les conditions de possibilité d'une linguistique du texte. En effet, Rastier insiste sur le fait que « tout texte doit être rapporté à un corpus pour être interprété » (Rastier, 2011 : 25 cité par Mayaffre, 2011 : 2) et que si « le morphème est l'unité élémentaire, le texte est pour une linguistique évoluée l'unité minimale et le corpus l'ensemble dans lequel cette unité prend son sens. (Rastier, 2011 : 33 cité par Mayaffre, 2011 : 2).

Pour Rastier, « Un texte ne peut se lire que dans un corpus, qu'il soit implicite, comme en général dans les études littéraires, ou explicite, comme en linguistique de corpus ». (Rastier, 2011 : 64 cité par Mayaffre, 2011 : 2). Les corpus « apparaissent avant tout, de manière dynamique, comme la condition de l'interprétation, […] le lieu empirique » (Rastier, 2011 : 80 cité par Mayaffre, 2011 : 2).

On conclura cette section en soulignant que le corpus apparaît comme une structure fondamentale dans les champs des sciences du langage qui ont été abordés jusqu'ici. On insistera sur la nécessité de considérer que « construire un corpus n'est pas une succession de gestes anodins » (Gadet, 2007 : 20). Ce travail suppose de mettre en place une « traçabilité » des données (Ibid.) et de « mesurer les implications de chaque décision » (Ibid.) de manière à garantir la *cohérence* et la *pertinence* du corpus, son *acceptabilité* en termes de représentativité, son *exploitabilité* en termes d' *homogénéité* et de *volume* (Pincemin, 1999).

Tous les champs des SDL ne sont pas concernés au même titre par les corpus et en particulier la sociolinguistique qui va faire l'objet du troisième temps de cette analyse.

III Peut-on parler de corpus en sociolinguistique ?

Ce troisième temps sera construit autour des relations entre corpus et terrain et sur la spécificité de la notion de terrain en SL dans une perspective réflexive et altéritaire.

III. 1 Corpus et terrain (Cappeau, Gadet, 2007). Le corpus comme « construction interprétative du monde social » (Blanchet, 2007).

III. 1. 1 Corpus et terrain

Paul Cappeau et Françoise Gadet dans un article intitulé *L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale* écrivent les lignes suivantes qui interrogent la légitimité des corpus par rapport au terrain en SL :

> 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. Comme il n'est ni souhaitable ni possible de parler de l'ensemble de la sociolinguistique, nous nous limiterons ici à la part de cette discipline qui s'appuie sur le rapport à la langue. Car dans des domaines comme l'ethnographie de la communication, les contacts de langues, le bilinguisme, ou la macro-sociolinguistique, chaque situation et chaque terrain étant spécifiques, ce que l'idée même de « grand corpus » implique d'objectif de représentativité serait dépourvu de pertinence pour les sociolinguistes qui sont, depuis les fondements de la discipline, dans la constitution de collections de données réunies sur le terrain (Cappeau, Gadet, 2007 : 100).

Le corpus, selon ces chercheurs peut acquérir de l'importance, si on considère qu'il permet de regrouper des « données contextualisées et rapportées à des locuteurs bien identifiés » (Cappeau, Gadet, 2007 : 100) comme l'attestent les propos qui suivent :

> La sociolinguistique sera ici prise selon un sens très large, caractérisé davantage par le processus de regroupement de données contextualisées et rapportées à des locuteurs bien identifiés que par la référence à un contexte spécifiquement social ou écologique (Cappeau, Gadet, 2007 : 100).

L'oral apparaît comme le domaine d'étude privilégié de la SL :

> 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).

Cette perspective permet à la SL de ne plus être exclusivement cantonnée au terrain :

> Jusqu'au récent développement des grands corpus, il paraissait pertinent d'isoler, parmi les sous-disciplines des sciences du langage, les disciplines de terrain, qui s'appuient crucialement sur le recueil de données en contexte naturel, dont la sociolinguistique fait partie (Cappeau, Gadet, 2007 : 100).

Enfin on fera l'hypothèse qu'il existe peut-être une complémentarité des approches qui viennent d'être abordées, si on accepte de se situer sur un plan plus « global » (Mayaffre, 2005 : 6) et en particulier celui d'une « sémiotique des cultures » (Rastier, 2001, cité par Mayaffre, 2005 : 6).

Ce détour par l'analyse de F. Gadet et P. Cappeau, qui confère un statut au corpus dans le champ de la SL, nous permet de poser la question de savoir dans un deuxième temps si le corpus ne peut pas être considéré comme une « construction interprétative du monde social par certains de ses acteurs » (Blanchet 2007 : 347) comme on peut le voir à travers l'opposition corpus de type constitutif/ corpus de type restitutif construite par P. Blanchet (Ibid.).

III.2 L'opposition corpus de type constitutif/ corpus de type restitutif en sociolinguistique (Blanchet, 2007)

Sandra Nossik dans un article intitulé « Les récits de vie comme corpus sociolinguistique : une approche discursive et interactionnelle » montre que « Philippe Blanchet oppose […] les corpus de type « constitutif », sur lesquels se fonde exclusivement le travail du linguiste, aux corpus « restitutifs », qui n'offrent qu'un exemplier des données observables sur le terrain

(2007 : 343). Dans le cadre d'une sociolinguistique ethnographique, il pose le statut « secondaire » du corpus vis-à-vis du terrain, dont « la complexité subjective des situations » ne saurait être reflétée par « un "corpus" même méthodiquement construit » (*ibid.*). En cela, l'approche sociolinguistique se distingue de celle de la *linguistique de corpus*, à l'appui de laquelle, P. Blanchet cite un ouvrage de B. Habert, A. Nazarenko et A. Salem (1997). Dans cette dernière perspective, le statut du corpus est *primordial* (Blanchet, 2007 : 343), et la recherche se fonde sur « un ensemble de « données » prélevé dans le « réel » et analysé en lui-même (voire pour lui-même) » (Ibid.) par le chercheur. La méthodologie est axée prioritairement sur la mise en forme du corpus, son *élaboration* et sa *présentabilité* (Blanchet, 2007 : 343). Les aspects techniques, notamment dans le cas des corpus numériques, sont susceptibles de prendre le pas sur les questions de pertinence et de *significativité* du corpus (Blanchet, 2007 : 347). Pour l'approche sociolinguistique, le terrain est prépondérant ainsi que le montre D. de Robillard :

> La construction d'un corpus ne s'arrête bien évidemment pas à son enregistrement matériel […] et […] se poursuit dans d'autres opérations […] par lesquelles on lui donne sens et valeur démonstrative et que je ne peux qu'appeler *« interprétation »*[9] (Robillard, 2007 : 358).

Cette conception conduit à redéfinir la notion de données et à préférer à donnée le terme d'*observable*. P. Blanchet définit les données comme « une « réalité objective extérieure à l'intervention du chercheur » » (2007 : 343). Selon P. Blanchet (2007 : 350) :

> tout est bon à prendre, ça dépend de ce qu'on en fait […] les observables réunis en corpus et présentés dans un discours scientifique ne fondent pas prioritairement la connaissance mais en témoignent en alimentant le processus interprétatif (Ibid.).

Parler d'observables permet d'insister sur le rôle que joue l'analyse du chercheur dans la construction de ces éléments et leur problématisation dans la réflexion scientifique.

Pour Martineau, l'observation est :

> Un outil de cueillette de données où le chercheur devient le témoin des comportements des individus et des pratiques au sein des groupes en séjournant sur les lieux même où ils se déroulent » (Martineau, 2005 : 6).

Le chercheur ajoute que :

> L'observation en situation [peut avoir comme objectif] d'expliquer le plus objectivement possible les faits. En ce cas, l'accent sera mis essentiellement sur la description des événements observés. Par contre, si le chercheur se réfère à une position plutôt interprétative (inspirée de la phénoménologie), l'objectif consistera avant tout à comprendre la signification que les acteurs attribuent à leurs pratiques. Par conséquent, le regard se portera surtout sur la construction du sens. Enfin, si l'on se veut constructiviste […], on poursuivra l'objectif de comprendre les règles de construction du social. C'est pourquoi, le chercheur accordera une attention particulière aux interactions entre les acteurs lorsqu'ils co-construisent leur monde (Martineau, 2005 : ibid).

L'observation permet au chercheur d'acquérir ce que Gumperz appelle un « savoir d'arrière-plan ou background knowledge » (1892), qu'on peut également nommer marqueurs de contextualisation.

[9] Les italiques et les guillemets figurent dans l'article de D. de Robillard (2007 : 358).

La sociolinguistique est une discipline de terrain car d'une part elle « tend à englober toutes les recherches sur le langage dès que celles-ci posent des problèmes sociaux » (Bulot, 2013 : 35 citant Marcellesi, 1986), d'autre part elle « procède par enquête » (Ibid. : 37). Une « méthode d'enquête est d'abord un ensemble de stratégies mises en œuvre par le chercheur pour construire l'objet scientifique » (Bulot, 2013 : 37). En cela, on peut considérer qu'elle procède d'une co-construction.

L'enquête comme co-construction[10]

Éclairante pour notre propos est l'analyse que fait P. Bourdieu de la manière dont il a conduit ses enquêtes sociologiques pendant la guerre d'Algérie, en prenant en compte la façon dont la spécificité du contexte pesait sur les modalités de l'enquête :

> Mener une enquête sociologique en temps de guerre oblige à tout réfléchir, tout contrôler en en particulier ce qui va de soi dans la relation ordinaire entre l'enquêteur et l'enquêté. L'identité des enquêteurs. […]. Le sens même de l'enquête fait question plus que jamais pour les enquêtés eux-mêmes (ne s'agit-il pas de policiers et d'espions ?). Soupçon généralisé : à plusieurs reprises des agents de renseignements viennent, à la suite des enquêteurs, enquêter à leur tour sur la nature de l'interrogation qu'ils avaient menée. […].
>
> On ne peut survivre, au sens propre dans une telle situation qu'au prix d'une réflexivité permanente et pratique qui est indispensable dans les conditions d'urgence et de risques extrêmes, pour interpréter et apprécier instantanément la situation et mobiliser plus ou moins consciemment, les savoirs et les savoir-faire acquis dans la prime expérience sociale (2004 : 68-69).

On abordera à ce stade la question du terrain

III.2.1 La notion de terrain

M. Agier, propose du terrain une définition suivante selon laquelle le terrain constitue un « ensemble de relations personnelles » qui se construisent entre le chercheur et ses informateurs :

> Le terrain n'est pas une chose, ce n'est pas un lieu, ni une catégorie sociale, un groupe ethnique ou une institution (...) c'est d'abord un ensemble de relations personnelles où « on apprend des choses ». « Faire du terrain », c'est établir des relations personnelles avec des gens (Agier, 2004 : 35).

Le terrain peut être défini comme : « l'objet d'étude et le lieu de séjour » (Combessie, 1996 : 15) où se déroule l'enquête sociolinguistique fondée sur l'observation. Il ne doit pas être confondu avec le territoire. Le territoire est défini comme « l'espace social construit dans et par les limites de l'espace physique » (Paugam, 2015 [1ère éd. 2010] : 116).

La question du terrain traverse le champ des sciences humaines et sociales, et notamment celui de la sociolinguistique et de la didactique dans les cas où elle suppose un terrain. Or cette notion n'est pas simple et conduit à interroger la relation du chercheur et particulièrement sa relation intime et subjective au terrain. On postulera qu'une *incertitude irréductible* caractérise cette relation (I. Stengers, 1993 : 162) et qu'on ne peut rigidifier le terrain en le solidifiant et en en faisant un objet de recherche, comparable à l'objet des

[10] À partir de ce titre et jusqu'à la fin du chapitre, le texte est issu de l'article suivant : M.-M. Bertucci, 2007, « Le chercheur et son terrain : peut-on parler d'un « objet de recherche en sciences humaines et sociales », in *Le Français aujourd'hui* 159, pp. 113-118.

sciences dures par exemple. Le terrain préexiste à celui qui l'analyse et échappe aux procédures expérimentales du laboratoire et de ce fait à la production de *preuves stables* (Ibid.). Le terrain est soumis aux aléas de la temporalité et à leur dynamique et échappe au temps scientifique du laboratoire (Ibid.). Il se soustrait aussi à la dichotomie sujet-objet, qui suppose un objet prêt à être analysé par un *observateur neutre* et l'illusion de la séparation du personnel et du professionnel (C. Ghasarian, 2004 : 10). Bien au contraire. Ainsi, certains anthropologues contemporains considèrent que loin d'éviter ou de chercher à contrôler les émotions en jeu sur le terrain, il faut les analyser et les intégrer à la recherche (S. Kleinman et M. Cop, 1993 ; C. Ghasarian, 2004 : 11). C'est ce que P. Bourdieu a décrit comme *l'objectivation participante* (2003), autrement dit l'élucidation du rapport subjectif du chercheur à sa recherche à travers une démarche réflexive.

Ainsi comme J.-C. Kaufmann (2001) et son approche fondée sur l'empathie par exemple, on admettra que loin d'être extérieur à son terrain, le chercheur contribue à la construction des faits. Il les co-construit et ne se borne pas à faire émerger des données préexistantes, qui lui seraient radicalement extérieures (Robillard, 2007 : 360). Comme on l'a vu précédemment, le terrain devient un lieu d'investissement affectif et cognitif auquel se confronte le chercheur (Bertucci, 2007). Cette relation permet la production d'hypothèses sur un mode empirique, l'analyse part d'éléments fragmentaires qui ne visent pas d'entrée de jeu l'universel, la généralisation (Robillard, 2003 : 2-3). C'est la raison pour laquelle elle a été analysée comme réflexive.

III. 2.2 La réflexivité

L'analyse de la réflexivité a largement été initiée par les anthropologues, qui compte tenu de la place du terrain dans leur discipline, ont été amenés à réfléchir à l'incidence de leurs états d'âme et de leurs émotions sur leur sujet d'étude. On peut penser notamment à B. Malinowski dès 1967 ou plus récemment à C Ghasarian (2004) et d'une manière générale aux anthropologues américains, P. Rabinow en 1977, H. F. Wolcott en 1995.

Quel est l'intérêt de la réflexivité pour la recherche en sciences humaines et plus largement pour tout chercheur en relation avec un terrain ? Autrement dit, comment concilier l'objectivation inhérente à la recherche avec la part de subjectivité propre à chaque chercheur (P. Cohen, 2004 : 86) et que dire de soi chercheur au bout du compte en évitant l'anecdotique et le particulier ? Ceci suppose de considérer que la recherche est une expérience humaine qui se constitue en tant que recherche par le processus réflexif.

On insistera d'emblée sur le fait que la réflexivité n'est pas « une introspection psychologisante et autocentrée du chercheur » (C. Ghasarian, 2004 : 14) mais qu'elle est constitutive de la posture de recherche car elle suppose un travail constant du chercheur sur ses positionnements, ses angles d'attaque et une réactivité permanente, qu'illustrent très clairement les propos de P. Bourdieu sur une situation extrême, dans des conditions où la recherche risque d'être mise en cause à tout moment par le contexte de la guerre d'Algérie.

Ainsi comme J.-C. Kaufmann (2001) et son approche fondée sur l'empathie par exemple, on admettra que loin d'être extérieur à son terrain, le chercheur contribue à la construction des faits. Il les co-construit et ne se borne pas à faire émerger des données préexistantes, qui lui seraient radicalement extérieures. En fait, le terrain devient un lieu d'investissement affectif et cognitif auquel se confronte le chercheur. Cette relation permet la production d'hypothèses sur un mode empirique, l'analyse part d'éléments fragmentaires qui ne visent pas d'entrée de jeu l'universel, la généralisation (D. de Robillard, 2003 : 2-3).

C'est la raison pour laquelle on peut postuler que la neutralité du chercheur est « un rôle construit » et que l'histoire et la culture du chercheur forment et déterminent largement sa perspective (C. Ghasarian, 2004 : 15). On formera l'hypothèse que le terrain n'est connu au bout du compte, de manière consciente, qu'à travers le discours qui est produit sur lui, voire à travers le croisement des discours en interrogeant et en éclaircissant les principes épistémologiques, théoriques et politiques des productions, qui ne peuvent éviter de ressortir d'une idéologie, laquelle doit être élucidée par le chercheur. Il est donc peut-être préférable d'assumer la part de subjectivité de la recherche plutôt que de la contenir sans rien en faire, au risque de la voir déborder subrepticement du discours. Dans ces conditions, la recherche en sciences humaines, doit plus être un exercice de construction problématique d'une situation complexe qu'une description ou qu'une série de monographies et ce d'autant plus que l'ampleur, l'instabilité, l'hétérogénéité de fait du terrain rendent difficile une description univoque, stable et systématique. Á ce stade, se pose inévitablement le problème de l'écriture et du statut du texte en sciences humaines dont il faut souligner le caractère problématique. Les théories du « tournant pratique » ou *practice theorist* l'ont bien montré. Celles-ci se retrouvent autour de l'idée qu'il faut comprendre l'activité scientifique comme une globalité dont tous les aspects contribuent à la production du savoir (T.R. Schatzki, K. Knorr Cetina & E. Von Savigny, 2001). Rien n'est donné, tout est à négocier dans un processus réactif, qui prend ses distances avec la fiction du sujet connaissant et de l'objet à connaître et l'illusion de la neutralité[11] (I. Stengers, 2006 : 68). L'écriture des articles scientifiques en est une dimension.

Rien n'est donné, tout est à négocier dans un processus réactif, qui prend ses distances avec la fiction du sujet connaissant et de l'objet à connaître et l'illusion de la neutralité[12] (I. Stengers, 2006 : 68). L'écriture des articles en est une dimension.

III. 2. 3 L'écriture scientifique

L'écriture scientifique repose sur la convention communément admise d'une écriture blanche, comme la voix blanche est sans timbre et les vers blancs sans rime, c'est-à-dire dépourvue d'affects et se défiant des effets de style. Or les voies qu'emprunte le chercheur pour traduire, transcrire ce qu'il poursuit (récits de vie, entretiens, dialogues) ne sont pas en rupture totale avec les formes littéraires. Quelle que soit la discipline, le chercheur est bien aussi d'une certaine manière un auteur, dont on entend la voix en contrepoint dans les récits de vie ou les entretiens par exemple.

En ce sens, les processus discursifs de l'écriture de l'article mériteraient d'être considérés avec plus d'attention, phénomène déjà souligné d'ailleurs par J.-M. Adam (1990) ou F. Affergan (1999) si l'on admet que dans l'écriture scientifique est à l'œuvre un processus de traduction, qui comporte une part de récit. Cette diversité des discours produit un effet de perspective dans le texte scientifique et construit une relation dialogique avec le thème de la recherche, qui permet de construire une distance du dedans. En ce sens, il y a de la créativité dans la recherche par la co-construction déjà évoquée (L. Porcher, 1997 ; I. Stengers, J. Schlanger, 1991 ; J.-C. Kaufman, 2001). On fera l'hypothèse qu'il y a aussi une forme de

[11] On rejoint là les orientations du post-modernisme où la distinction sujet/objet est critiquée au profit de celles de subjectivité et d'hétérogénéité des perspectives, voir notamment sur ces questions, J.-F. Lyotard, *La condition postmoderne : rapport sur le savoir*, Paris, éditions de Minuit, 1979, D. Harvey, *The condition of postmodernity : an enquiry into the origins of cultural change*, Oxford, Blackwell, 1989.

[12] On rejoint là les orientations du post-modernisme où la distinction sujet/objet est critiquée au profit de celles de subjectivité et d'hétérogénéité des perspectives, voir notamment sur ces questions, J.-F. Lyotard, *La condition postmoderne : rapport sur le savoir*, Paris, éditions de Minuit, 1979, D. Harvey, *The condition of postmodernity : an enquiry into the origins of cultural change*, Oxford, Blackwell, 1989.

littérarité, qui ne se reconnaît pas comme telle dans l'écriture scientifique, notamment dans le recours au récit, dans les descriptions ou dans les entretiens. Cette littérarité serait analogue à la littérarité de l'histoire clinique dont parlent les psychanalystes (J Hillman, 2005 : 52) ou de l'histoire, qui s'ordonne comme un récit à partir d'une intrigue, voire à « une tranche de vie », à la manière d'un drame ou d'un roman (P. Veyne, 1979 : 36).

Évoquer l'écriture ne conduit pas cependant à abandonner la finalité première de l'activité scientifique, qui est la construction et la production d'un savoir. Elle engage cependant à prendre en compte la question des relations en jeu dès qu'on travaille non seulement sur les productions des locuteurs mais aussi sur leur contexte en didactique ou sociolinguistique par exemple. Le terrain est fait de la singularité d'expériences successives, de rencontres avec des individus ou de univers sociaux autres dont les caractéristiques conditionnent autant les modalités de la recherche que le savoir qui en est issu et dont l 'hétérogénéité se stabilise dans la durée de l'écriture.

Si le texte scientifique est le lieu d'une « intertextualité savante » pour reprendre l'expression de F. Laplantine (2004), il n'existe qu'irrigué par la parole ou les textes des autres chercheurs et donc par une altérité constitutive dans un processus historique.

L'article n'est pas de fait, et malgré les apparences, clos sur lui-même mais il est nécessairement contextualisé, inscrit dans un processus historique et soumis à de nécessaires réactualisations. Il en est de même de la recherche en sciences humaines si on admet que son objet est instable, évolutif et mouvant. C. Ghasarian va jusqu'à dire que « chaque texte écrit par des chercheurs en sciences humaines n'est pas le reflet d'une réalité mais plutôt celui d'une sensibilité » (Op.cit. : 13). Soutenir qu'il y a une neutralité de l'écriture scientifique revient à méconnaître la part de créativité, qui revient au chercheur, et tout particulièrement lorsqu'il s'agit de décrire un terrain. La part du littéraire dans les mécanismes discursifs des monographies ou des comptes rendus d'entretiens a d'ailleurs été soulignée à plusieurs reprises (J.-M. Adam et *al.* : 1990 ; F. Affergan : 1999). Le texte doit être construit de manière à ce que celui dont on rapporte les propos dans un entretien par exemple ne soit pas effacé, ce qui conduit à poser la nécessité d'une écriture dialogique, polyphonique qui fait apparaître l'intersubjectivité et l'altérité, la polysémie et l'hétérogénéité présentes dans la vie sociale, le terrain étant d'une certaine façon une métonymie de l'autre (D. de Robillard, 2007). L'anthropologie est allée loin dans cette voie puisque certains chercheurs vont jusqu'à prôner la mise en place de différents types d'écriture à mettre en place en fonction de la diversité des contextes et de la rencontre avec le terrain, en l'occurrence avec l'altérité (J. Van Maanen, 1988 : 140).

Conclusion. L'altérité au cœur de la recherche en sciences du langage ?

Or appréhender cet « hors de soi » pour reprendre la formule de F. Affergan (2006 : article altérité) pose un sérieux problème méthodologique. Comment saisir ce qui se présente à première vue comme un autre que soi et qui en fait pourrait bien être un autre soi ? Ce dilemme conduit à attribuer à l'autre des visages successifs : le sauvage, l'étranger, le migrant, jusqu'à la construction de classifications et d'inventaires de traits distinctifs voire discriminatoires. Or sortir de ce dilemme pour construire une méthodologie de recherche prenant en compte l'altérité[13] suppose des modèles de recherche inédits, si on reste dans une posture de recherche classique, sans quoi l'altérité court le risque de demeurer un horizon, une question relevant plus de l'éthique que du champ des sciences humaines et sociales. Cela suppose de renoncer à l'homogénéisation et d'assumer l'éclatement qui en résulte. En sciences du langage, assumer l'altérité oblige à vivre avec les langues, non d'un point de vue

[13] Voir sur ce point ce que B. Latour appelle *l'anthropologie symétrique* (1996 : 54-55), *Petite Réflexion sur le culte moderne des dieux Faitiches,* Paris, Les Empêcheurs de penser en rond.

essentialiste et hiérarchique, mais à admettre qu'elles peuvent être instables, hétérogènes[14], contextualisées (D. de Robillard, 2007) et qu'elles sont en relation étroite avec une temporalité qui les historicise, ce qui métaphorise en quelque sorte la posture de la recherche.

Dans ces conditions, il devient difficile d'envisager un rapport autre que symétrique avec l'autre, ce qui conduit à renoncer à une posture de recherche en surplomb pour adopter une démarche réflexive et autocritique, en décentrant les perspectives.

Bibliographie

Adam J.-M., Borel M.-J., Calame C. , Kilani, M., (1990), *Le discours anthropologique, description, narration, savoir*, Lausanne, Payot.

Affergan F., (1999), *Construire le savoir anthropologique*, Paris, PUF.

Affergan F., (2006), Article *altérité*, dans Mesure S., Savidan P. *Le dictionnaire des sciences humaines*, Paris, PUF.

Agier, M., 2004, *La sagesse de l'ethnologue*, Paris, L'œil neuf.

Bertucci M.-M. 2007 « Le chercheur et son terrain, peut-on parler d'un objet de recherche en sciences humaines et sociales ? ». In *Le français aujourd'hui*. 159. Paris : Armand Colin / AFEF. pp. 113-117.

Blanchet, P., 2007, « Sur le statut épistémologique de la notion de "corpus" dans un cadre ethno-sociolinguistique », *in* Auzanneau M. & Verdoïa S. (éd.) *La mise en œuvre des langues dans l'interaction*. Paris : L'Harmattan, 341-352.

Blanchet, P., 2013, *Une introduction à la sociolinguistique. Pour l'étude des dynamiques de la langue française dans le monde*, Paris : EAC, pp. 27-41.

Bourdieu P., 2003, « L'objectivation participante » dans *Actes de la Recherche en Sciences Sociales*, n° 150, pp.43-58.

Bourdieu P., 2004, *Esquisse pour une auto-analyse*, Paris, Raisons d'agir (éd.).

Bulot, T., 2013 « Objet, terrains et méthodes de la sociolinguistique », in Thierry Bulot,

Cappeau, P., Gadet, F., 2007, « L'exploitation sociolinguistique des grands corpus. Maître-mot et pierre philosophale », in *Revue française de linguistique appliquée*, pp. 99-110.

Cohen P., 2004, « Le chercheur et son double. Á propos d'une recherche sur le vécu des jeunes de La Réunion face au sida » dans Ghasarian C., *op. cit.*, pp. 73-89.

Combessie, J.-C., 1996, *La méthode en sociologie*, Paris : La Découverte.

Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J.-B., Mével, J.-P., 1973, *Dictionnaire de linguistique*, Paris : Larousse.

Gadet, F., 2007 [1ère éd. 2003], *La variation sociale en français*, Paris : Ophrys.

Ghasarian C., 2004, *De l'ethnographie à l'anthropologie réflexive. Nouveaux terrains, nouvelles pratiques, nouveaux enjeux*, Paris, Armand Colin.

Ghasarian C., 2004, « Sur les chemins de l'ethnographie réflexive », dans Ghasarian C., *op. cit.*, pp. 5-33.

Gumperz J. J.,1982. *Discourse strategies*. New York : Cambridge University Press. Trad. fr. (1989). *Engager la conversation, Introduction à la sociolinguistique interactionnelle*. Paris : Minuit.

Habert, B., Nazarenko, A., Salem, A., 1997, *Les linguistiques de corpus*, Paris : Armand

[14] Á cet égard, le témoignage de l'écrivain algérien Slimane Benaïssa est éclairant. Il se définit comme un « métis culturel », modelé par les langues. Parlant aussi bien le français que le kabyle (Taqbaylit), ou l'arabe, il déclarait en 1998, à propos de sa pièce *Les fils de l'amertume* (1999, Paris, Plon) « J'entends mon peuple dans toutes ses langues » (Voir *Périphéries. Rencontre avec Slimane Benaïssa.* www. peripheries.net). Cet exemple illustre particulièrement bien la question de la contextualisation et de l'historicisation des langues, dans un macro-contexte

Colin.
Hillman J., 2005, *La fiction qui soigne*, Paris, Payot.
Kaufmann J.-C., 2001, *L'entretien compréhensif*, Paris, Nathan.
Kleinman S., Copp M., 1993, "Emotions and Fieldwork" dans *Qualitative Research Methods Series* n° 28, Newbury Park CA, Sage Publications.
Laplantine F., 2004, « L'anthropologie genre métis », dans Ghasarian C. *op. cit.*, pp.143-152.
Malinowski B., (1985, éd. fr.), *Journal d'ethnographe*, Paris, Seuil.
Marcellesi, J.-B., 1986, « Éléments et lectures de sociolinguistique pour le DLE », in *Une introduction à la recherche scientifique en didactique des langues,* Paris : CREDIF/Didier, pp. 65-86.
Martineau, 2005, « L'instrumentation dans la collecte des données. L'observation en situation : enjeux, possibilités et limites »
Mayaffre, D., 2002, « Les corpus réflexifs : entre architextualité et intertextualité », in *Corpus*, I, 1, pp. 51-70.
Mayaffre, D., 2011, « François RASTIER — *La mesure et le grain. Sémantique de corpus.* Paris : Champion, Collection *Lettres numériques*, 2011 », compte rendu de lecture, in Varia *Corpus*, 10, pp. 320-325.
OCDE, 2007, *Principes et lignes directrices de l'OCDE pour l'accès aux données de la recherche financée sur fonds publics*, Paris : Les Éditions de l'OCDE.
Paugam, S., 2015 [1ère éd. 2010], *Les 100 mots de la sociologie*, Paris : PUF
Pincemin, B., 1999, « Construire et utiliser un corpus : le point de vue d'une sémantique textuelle interprétative » in A. Condamines, M.-P. Péry-Woodley & C. Fabre (éds), *Atelier corpus et TAL : pour une réflexion méthodologique* (TALN 99), Cargèse, pp. 26-36.
Porcher L., 1997, « Lever de rideau » dans Zarate G., Candelier M. (éds.). *Les représentations en didactique des langues et des cultures*, dans *Notions en questions* n° 2. Paris, CREDIF, pp. 11-27.
Rabinow P., 1985, (trad.fr.), *Un ethnologue au Maroc. Réflexions sur une enquête de terrain*, Paris, Hachette.
Rastier, F., 2001, *Arts et sciences du texte*, Paris : PUF.
Rastier, F., 2002, *Enjeux épistémologiques de la linguistique de corpus*, Texte de la conférence prononcée aux deuxièmes journées de linguistique de corpus à Lorient. *Texto*, [en ligne]. Rubrique Dits et Inédits.
Robillard D. de, 2003, « Français, variation, représentations : quelques éléments de réflexion », dans Billiez J., Robillard D. de (éds.), *Français : variations, pratiques et représentations, Cahiers du français contemporain* n° 8, pp. 35-61.
Robillard D. de, 2007, « La linguistique autrement : altérité, expérienciation, réflexivité, constructivisme, mutiversalité : en attendant que le Titanic ne coule pas », dans *Carnets d'Atelier de Sociolinguistique* n° 1, www.u-picardie.fr/ LESCLAP.
Robillard D. de, 2007, « Lyophiliser ou interpréter l'ornithorynque ? Corpus, métonymie / métaphore », *in* Auzanneau M. (éd.) [avec la collaboration de S. Verdoïa], *La mise en œuvre des langues dans l'interaction*. Paris : L'Harmattan, pp. 352-370.
Schatzki T.R., Knorr Cetina K. & Von Savigny E., 2001, *The Practice turn in contemporary Philosophy*, Londres, Routlege.
Sinclair, J., 1991, *Corpus Concordance Collocation*, Oxford : Oxford University Press.
Sinclair, J., 2005, "Corpus and Text - Basic Principles" in *Developing Linguistic Corpora : a Guide to Good Practice,* ed. M. Wynne. Oxford : Oxbow Books : 1-16.
Stengers I.,1993, *L'invention des sciences modernes*, Paris, La Découverte.
Stengers I., 2006, *La vierge et le neutrino. Les scientifiques dans la tourmente*, Paris, Les Empêcheurs de penser en rond.
Stengers I., Schlanger J., 1991, (éd. fr.), *Les concepts scientifiques*, Paris, Gallimard.

Van Maanen J., (1988), *Tales from the Field. On writing Ethnography*, Chicago, The University of Chicago Press.
Veyne P., 1979, *Comment on écrit l'histoire*, Paris, Seuil.
Wolcott H. F., 1995, *The art of Fieldwork*, Maryland, Alta Mira Press.

Sitographie

CNRTL : http://www.cnrtl.fr/definition/corpus
Guyon, C. 2016. « Le concept d'archives : d'une définition à l'autre. » 2016. <hal-01401138>. https://hal.archives-ouvertes.fr/hal-01401138/document
Mayaffre, D., 2005, « Les corpus politiques : objet, méthode et contenu. Introduction » in *Corpus*, [En ligne], 4. *Les corpus politiques : objet, méthode et contenu*. Mis en ligne le 5 septembre 2006. URL : http://corpus.revues.org/292
Nossik, S., 2011, « Les récits de vie comme corpus sociolinguistique : une approche discursive et interactionnelle », in *Corpus* [En ligne], 10 | 2011, URL : http://corpus.revues.org/2045
Ortolang: URL : Ortolang : URL : https://www.ortolang.fr
Loi n° 2016-925 du 7 juillet 2016 relative à la liberté de la création, à l'architecture et au patrimoine : URL :
www.legifrance.gouv.fr