

**Réflexions épistémologiques sur la constitution
d'un corpus sociolinguistique.
La relation d'enquête en question**

Souheila HEDID¹

Le présent travail interroge un aspect important de la recherche en sociolinguistique, celui de la méthodologie des enquêtes de terrain. L'objectif de ce travail est de comprendre les multiples interactions entre le chercheur et le terrain d'enquête (M. Houis, 1971). Dans les sciences humaines et sociales, la relation d'investigation a souvent été difficile à saisir. Ce que l'enquêté sait de l'enquêteur c'est ce que ce dernier lui montre. De nombreuses études ont mis l'accent sur les négociations implicites et instantanées que les deux partenaires accomplissent lors de leur premier contact.

La présence de l'enquêteur sur le territoire de l'informateur peut être appréhendée comme une intrusion à multiples facettes à cause de plusieurs éléments : le physique, le sexe, l'âge, la couleur, le style vestimentaire, l'appartenance sociale, l'accent, la façon de parler..., etc. Nous allons apporter plus d'éclairage à cette question en interrogeant des travaux déjà réalisés. La réflexivité épistémologique permet dans ce cas de faire des retours critiques plus méthodiques sur les données récoltées ultérieurement (H. Boyer, 2010). Cette façon de procéder rend plus rentable la lecture des corpus présentés sous un autre regard. Nous serons en mesure de mettre l'accent sur les éléments non dits, qui sont passés implicitement dans les analyses effectuées au départ.

1. Le corpus en sciences humaines et sociales et en linguistique

Pour introduire une publication scientifique portant sur des questions méthodologiques en sciences sociales, Charlier et Moens (2006 : 7) expliquent que : « Les images sont nombreuses pour décrire ou dénoncer la complexité

¹Université Frères Mentouri, Constantine 1, Algérie.

des rapports qui, en sciences sociales et particulièrement en sociologie, unissent le scientifique à son objet d'étude. La question classique – "l'homme peut-il tout à la fois être l'objet et le sujet d'une science ?" – demeure en filigrane des productions de nos disciplines. Elle est souvent confondue avec celle de l'objectivité, qu'on s'accorde aujourd'hui à considérer définitivement hors d'atteinte, ou de l'objectivation, qui s'y est substituée ». Effectivement, le positionnement épistémologique du chercheur dans le circuit de la démarche scientifique a toujours été un sujet à controverse. Les écarts entre les sciences sociales et les sciences dures constituent dans ce cas une contrainte de taille dans le cadrage de toutes problématiques. Si l'objet d'étude est différent, son appréhension par le scientifique va certainement obéir à des restrictions méthodologiques différentes, dictées par la nature des données étudiées. Pour atteindre une réponse à la question-filigrane évoquée dans la citation de Charlier et Moens, il nous semble pertinent de mettre tout d'abord l'accent sur le point de départ de la réflexion scientifique expérimentale, celle de la constitution du corpus.

D'un état brut à une donnée à analyser, le corpus est pour la science un objet d'étude et une source de questionnement permanent. Élément essentiel voire incontournable, il est le point de départ de toute réflexion scientifique. En sciences humaines et sociales, le corpus est souvent considéré dans une perspective de complexité. L'hétérogénéité des éléments qui le composent, la diversité des contextes et des populations étudiées font que le choix des outils d'investigation et du cadrage méthodologique devient problématique. Que la méthode préconisée soit qualitative ou quantitative, la conception de ce produit scientifique doit répondre à des exigences, telles que la fiabilité et la représentativité. L'analyse doit s'effectuer dans une perspective réflexive qui permet un retour sur son élaboration et son parcours d'interprétation.

Les études faites en linguistique attestent de l'omniprésence des données d'appui. Bien qu'ils soient différents, diversifiés et hétérogènes, ces éléments se convergent vers la même option, confirmer ou infirmer les hypothèses du départ et accorder à la problématique annoncée plus de fiabilité. Plusieurs branches de cette discipline sont identifiées grâce à ces données d'appui : ainsi, la linguistique de corpus, par exemple, met en avant des contenus tirés des textes réels pour cerner les différentes problématiques qu'elle traite. La linguistique de terrain, de son côté, s'appuie sur le terrain et tout ce qu'il peut offrir comme pistes représentatives des réalités sociolinguistiques des contextes observés.

La réflexion linguistique trouve son origine dans les corpus étudiés, de l'observation à l'exploitation en passant par la sélection et la collecte. Les pistes épistémologiques développées autour des préoccupations des chercheurs mettent en avant la vision fragmentée de l'univers et la perte de certitude devant l'objet d'étude (Morin, 1977). Or, c'est cette incertitude qui sous-tend l'esprit scientifique et l'oriente vers sa dimension hypothético-déductive ou empirico-inductive.

2. L'architecture de l'argumentation

La réflexion que l'on tente d'appréhender trouve ses origines dans les travaux de terrain que nous effectuons depuis plusieurs années en Algérie. Difficile à interroger, le locuteur algérien semble entretenir une relation conflictuelle avec la question. Ce fait a souvent rendu nos investigations pénibles et a donné lieu à des malentendus qui nous ont, parfois, paralysés. La lecture des travaux réalisés en sciences du langage laisse apparaître une hétérogénéité frappante en matière de cadrages méthodologiques, et ce à plusieurs niveaux : le choix des outils d'investigation, la sélection des terrains d'enquête ou celle des informateurs, ainsi que les protocoles et les démarches analytiques.

Les parcours méthodologiques des enquêtes de terrain diffèrent selon le contexte étudié. Le principe de l'adaptation de l'enquête aux contextes est une des préoccupations majeures des chercheurs. Les protocoles mises en place par les scientifiques reflètent parfaitement cet aspect méthodologique, où l'on relève, pour une seule problématique plusieurs protocoles. L'adaptation, dont il est question ici, est à placer en amont et en aval de toute l'investigation. Elle concerne tous les éléments évoqués ci-dessus mais elle est déterminante dans la relation qu'entretient le chercheur avec son terrain et avec ses informateurs. Or, la problématisation dont elle fait l'objet dans les sciences humaines et sociales la place comme facteur sous-tendant toutes les opérations même réflexives de la recherche scientifique.

La mise en exergue des multiples interactions entre le chercheur et son terrain (Houis, 1971 ; Bourdieu, 1993) définissent le cadrage de la relation d'enquête. Nous estimons que plusieurs facteurs interviennent et font que cette relation soit particulièrement codifiée et indubitablement liée au contexte d'enquête. Pour mieux appréhender ces pistes de recherche, nous orientons notre réflexion autour de deux interrogations : Quels sont les facteurs qui influent sur la relation d'enquête dans les investigations sociolinguistiques en

Algérie ? Comment le jeune chercheur en sociolinguistique peut-il gérer cet aspect au moment de la constitution de son corpus ? On ne s'étonnera donc pas de voir, dans cette recherche, un échafaudage plus théorique qu'expérimental. Où les données exploitées seront tirées directement des travaux déjà réalisés.

3. À l'origine du questionnement ...

Les chercheurs sont aujourd'hui unanimes sur le fait que la constitution et l'analyse d'un corpus sont en passe de modifier les pratiques voire les théories en lettres et en sciences sociales. Parmi toutes les opérations qui sous-tendent l'élaboration d'un corpus, celle qui consiste à interpellier les informateurs est certainement la plus difficile à cadrer. Ce qui semble important à signaler dans ce cas, c'est l'effet qu'elle exerce sur la suite et l'ensemble même de l'investigation. Dans cette perspective conceptuelle, Bourdieu (1993 : 4) explique : « Si la relation d'enquête se distingue de la plupart des échanges de l'existence ordinaire en ce qu'elle se donne des fins de pure connaissance, elle reste, quoi qu'on fasse, une relation sociale qui exerce des effets (variables selon les différents paramètres qui peuvent l'affecter) sur les résultats obtenus ». Effectivement, le premier biais, le plus innocent sans doute, est celui de l'importance de la relation d'enquête dans le tissu interactionnel de toutes les communautés. Le second atteste de la part sociale de la relation d'enquête et l'influence que peut avoir la société sur son accomplissement.

En effet, la relation entre le chercheur et l'informateur a connu des mutations facilement perceptibles dans les recherches publiées à travers les témoignages livrés par les scientifiques. Le facteur social y est pour beaucoup, les coutumes, les traditions, la stratification sociale, et bien d'autres facteurs, sont autant d'éléments susceptibles d'influencer les rapports établis entre l'enquêteur et son enquêté.

De plus, le changement, des paradigmes sociologiques, des systèmes économiques, des régimes politiques, des courants de linguistique et des mouvements littéraires, vient s'ajouter comme un facteur déclencheur des modifications ou des perturbations de l'ordre établi par les deux partenaires dans l'interaction d'enquête. En évoquant la particularité de la relation d'enquête et la difficulté épistémologique de la cerner, Payet et Giuliani (2010 : 7) expliquent : « Se rapprocher, aller voir plus près, entrer en relation, écouter, partager des activités, inscrire ces démarches dans une durée : qu'y a-t-il de nouveau à cela qui ne soit pas le credo de l'ethnologie – et les ethnologues,

anciens et néo, n'ont-ils pas déjà tout dit de l'art de l'enquête impliquée ? Notre affaire ici n'est pas disciplinaire. Elle est heuristique et éthique. Heuristique : nous voulons prendre au sérieux les contraintes de l'enquête auprès d'individus disqualifiés, comprendre les modes de résolution de ces contraintes par les enquêteurs, identifier les conséquences sur le travail d'interprétation des données. Éthique : il ne suffit pas de bonnes intentions pour dépasser les risques d'indécence de l'enquête de proximité. Heuristique et éthique : deux niveaux interagissent, car les conditions de connaissance et les conditions de décence sont imbriquées ». Les deux volets traités ici (Heuristique et éthique) concernent la gestion de cette relation dans les sciences humaines et sociales, où l'interprétation des données de l'enquête est déterminée par la nature du lien que l'enquêteur conçoit avec son informateur.

4. La relation d'enquête en sociolinguistique

En sociolinguistique, une lecture attentive des différentes recherches montre que la relation d'enquête a beaucoup changé. Pour mieux étudier cet aspect, nous mettons l'accent sur les différents facteurs qui interviennent habituellement dans la conception de ce lien. Les données que nous allons exploiter sont toutes tirées des travaux de terrain réalisés par des chercheurs algériens. Ils ne sont pas évoqués de façon directe mais utilisés pour renforcer l'argumentation et évoquer les difficultés rencontrées lors de la collecte des corpus. La liste présentée n'est pas exhaustive, mais elle comporte les données les plus récurrentes. Nous traitons dans ce qui suit, quatre éléments :

- a) L'épanouissement de la linguistique et notamment des sciences du langage ;
- b) L'intégration des technologies (Internet, logiciel, plateformes, applications,...) ;
- c) La durée de l'enquête et le paradoxe de l'observateur ;
- d) L'épanouissement de la linguistique et notamment des sciences du langage.

Souvent considéré comme un postulat de départ de toute réflexion scientifique, un changement des paradigmes épistémologiques implique un changement dans les relations d'enquête. En linguistique, le corpus n'avait pas engendré ce genre de problématiques auparavant. La linguistique structurale travaillait à partir de données préconçues, où la langue est avant tout un système codifié. L'émergence des corpus sociolinguistiques a donné naissance à des pistes variationnistes et interactionnelles.

En abordant la question, Boyer (2002) explique que : « Le statut des corpus a profondément évolué du structuralisme jusqu'à l'émergence de la sociolinguistique comme champ disciplinaire autonome. Cependant, à l'intérieur même de ce champ en pleine expansion, qui avait opté tout d'abord pour l'hétérogénéité comme critère de construction des corpus à traiter, il est possible d'observer une orientation interactionniste qui s'appuie sur des corpus très limités dont on cherche à mettre en évidence la densité représentationnelle en même temps qu'on souligne leur exemplarité ».

Effectivement, la conception épistémologique du corpus en linguistique a connu un changement radical. Son statut est passé d'un ensemble de données figées, d'une langue considérée en elle-même et pour elle-même comme un système d'éléments et de connexions, à un ensemble de données issues de situations de communications authentiques, dont l'objectif est d'étudier le fonctionnement de la langue en contexte. Pour mieux cadrer cette conception, Boyer (2002) ajoute qu' : « 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 ». Les orientations méthodologiques qui s'opèrent au sein de la sociolinguistique font que le corpus change d'apparence. Les deux niveaux analytiques, macrosociolinguistique et microsociolinguistique donnent lieu à des choix méthodologiques et à des approches critiques différents. Si le premier met l'accent sur la structure sociale, les pratiques langagières des groupes et des communautés linguistiques, et sur les types de variation... le second vise les interactions entre les locuteurs et les enjeux que ces pratiques impliquent. Que l'on se place dans la première ou la deuxième approche, le corpus semble conditionné par les caractéristiques de chacune. C'est à partir de ces considérations que le corpus sociolinguistique a pu se construire comme un ensemble de données à partir desquelles se décrivent les différentes situations étudiées.

a) *Bouleversements sociologiques, politiques, économiques, ...*

Sur ce plan, plusieurs facteurs peuvent être évoqués. Les sociologues parlent des changements profonds des sociétés et des communautés qui y vivent. Des événements historiques, des bouleversements sociopolitiques, des

changements des paradigmes de la pensée collective, interviennent souvent et font que les communautés changent de profil.

Les régimes politiques et les systèmes économiques incarnent les balises de la stabilité d'un peuple et définissent ainsi, leurs points de repères dans le circuit des interactions sociales et interpersonnelles. Ces éléments ne peuvent échapper à l'imaginaire collectif des individus et ils marquent indéniablement leurs représentations.

1. La stratification sociale s'est accentuée : l'ouverture des sociétés sur la mondialisation (et sur le marché mondial) a rendu plus fluide leurs structures et leurs organisations. Ainsi, une stratification sociale très intense s'est installée et elle a pesé très lourd sur les rapports interpersonnels entre les locuteurs. Elle a généré des jugements, des stéréotypes et des discours épilinguistiques qui ont à leur tour influencé les travaux scientifiques et les enquêtes de terrain. La stratification sociale a donné lieu à des questions telles que : Qui parle à qui ? Quelles images portent les uns sur les autres ? Quelles conceptions a l'enquêteur de son informateur et du terrain qu'il exploite ? Quels regards porte l'informateur sur l'enquêteur et sur sa présence sur son territoire ? Quelle est l'origine de l'un et de l'autre ? À quelle catégorie sociale appartient l'un et l'autre ?

2. Le rôle de la femme dans la société a changé : elle a un accès plus facile au domaine professionnel, elle occupe des postes, avant, réservés uniquement aux hommes. Ce qui est ici en question concerne l'image que la société a de la femme et du rôle qu'elle doit jouer. Sur les lieux d'enquête, les informateurs ne sont pas tous favorables à une présence féminine. Cette présence est considérée comme indésirable, une intrusion dans un univers qui lui est étranger. L'hostilité et la ségrégation sexuelle marquent la relation d'enquête. L'image que l'informateur a de la femme et la force de caractère de celle-ci sont autant d'éléments qui définissent la relation qu'elle va établir une fois sur le terrain.

3. Une ouverture sur le monde : qui a engendré, selon les cas, soit une fluidité dans les interactions sociales, soit un fort sentiment de haine, d'altérité et d'enfermement. Les situations d'enquête sont le lieu où les interactions sociales avec les étrangers se multiplient. Les images que les uns portent sur les autres sont parfaitement capables de modifier le parcours de la rencontre.

4. Les répercussions des événements politiques et historiques (guerres et conflits) : les situations conflictuelles (guerres, conflits, insécurité,...) provoquent le plus souvent un sentiment de peur et d'enfermement. Les

informateurs ne se laissent pas interroger facilement. En Algérie, la peur de se donner à un enquêteur se fait ressentir dans plusieurs travaux. Les événements historiques que les Algériens ont traversés (guerres, décennie noires,..) leurs rappellent des périodes de peur durant lesquelles les personnes interrogées étaient souvent coupables ou accusées de quelque chose ou risquaient d'être sévèrement punies. Ces données pèsent très lourd sur la relation d'enquête et rendent les interactions entre enquêteur et informateur plus difficiles à établir.

b) La technologie pour une relation d'enquête à distance

L'usage des logiciels, des magnétophones, des plateformes, des applications et bien d'autres moyens technologiques ont rendu les investigations sur le terrain plus abordables. La relation enquêteurs/enquêtés est fortement imprégnée par cet aspect. Les moyens technologiques ont instauré de nouvelles pratiques et ont inauguré ainsi une nouvelle ère où la distance entre ces partenaires est presque légitime. Beaucoup d'enquêtes s'effectuent actuellement sans qu'une rencontre de face à face entre les deux partenaires ne soit organisée. Les questionnaires en ligne, par exemple, très utilisés de nos jours, impliquent une relation d'enquête complètement différente : l'enquêteur envoie son questionnaire sur Internet, et les réactions des informateurs sont complètement ignorées et jamais signalées sur les plateformes. L'enquêteur reste loin du terrain et de ses informateurs.

c) La durée de l'enquête et le paradoxe de l'observateur

Depuis l'instauration du système LMD en Algérie, la durée, accordée, avant, aux jeunes chercheurs pour effectuer leurs travaux de recherche, a nettement diminué. L'architecture du système LMD stipule une réduction des années d'apprentissage et ce à tous les niveaux. Ce fait rend plus brève la période d'investigation et oblige les jeunes chercheurs à réétudier leurs protocoles d'enquête. Ainsi, la relation enquêteur/enquêtés doit être obligatoirement revue et réduite. Avoir le maximum d'informations en si peu de temps implique une gestion minutieuse des échanges et des interactions établies entre les partenaires.

En sciences du langage, l'enquêteur tente par les moyens les plus efficaces d'être très proche de son terrain et de ses informateurs. Il estime que sa proximité discrète garantit une observation plus efficace et des résultats plus précis. Il se propose d'« observer la façon dont les gens se servent du langage quand on ne les observe pas » (Labov, 1976 : 116). Cet objectif implique par

ailleurs que l'observateur s'engage à laisser à ces informateurs un espace assez large pour qu'ils se sentent à l'aise dans leurs interactions sociales et verbales. L'on parle dans ce cas du paradoxe de l'observateur, qui implique qu'on observe « le langage que parlent les gens quand on ne les observe pas » (Traverso, 1999 : 22). Le regard que porte le chercheur sur son objet est souvent conditionné par sa démarche et par ses choix méthodologiques. Les réalités sociolinguistiques étudiées doivent être soumises à des observations minutieuses. De plus, les terrains choisis, s'ils ne sont pas soigneusement exploités, ne dévoileront jamais leurs caractéristiques. La relation d'enquête devient dès lors un contrat tacite impliquant des acteurs dont le statut et le rôle sont définis en fonction de leur distance par rapport au terrain et à l'objet étudié.

Conclusion

Établir une bonne relation d'enquête n'est pas une chose aisée. Plusieurs facteurs interviennent et font qu'elle devient problématique. Les éléments évoqués ici ne doivent pas être perçus comme paralysants mais plutôt comme des paramètres à prendre en considération pour mieux cadrer le corpus et les données récoltées. Une lecture attentive des travaux de terrain réalisés en Algérie permet de confirmer que l'Algérien est difficile à interroger. Ceci n'est pas un défaut ni un problème en soi, c'est une donnée qui peut bien orienter le parcours d'une investigation, car elle renseigne l'enquêteur sur les éventuelles conceptions (idéologiques surtout) de ses informateurs.

La constitution d'un corpus en sciences humaines et sociales, et, particulièrement dans les sciences du langage, ouvre souvent la voie vers de nouvelles pistes épistémologiques. Rastier et Ballabriga (2006) expliquent que : « La constitution et l'analyse de corpus est en passe de modifier les pratiques voire les théories en lettres et sciences sociales », ce qui implique que le fondement même de la pensée scientifique, où l'on s'interroge sur la nature de la donnée et sur la démarche à suivre pour la traiter, soit souvent mis en question à chaque fois que l'on procède à la conception d'un corpus ou à son analyse.

Bibliographie

- Bourdieu, P., « À propos de la famille comme catégorie réalisée », in *Actes de la recherche en sciences sociales*, Vol. 100., 1993, p. 32-36.
- Houis, M., *Anthropologie linguistique de l'Afrique Noire*, Édité par Presses Universitaires De France, 1971.
- Boyer, H., « Sociolinguistique : faire *corpus* de toute (s) voix ? », *Mots. Les langages du politique* [En ligne], 69 | 2002, mis en ligne le 14 mai 2008, consulté le 25 avril 2018. URL : <http://journals.openedition.org/mots/10553> ; DOI : 10.4000/mots.10553
- Charlier J.-E., Moens F., *Observer, décrire, interpréter : Un état des méthodes en sciences sociales*, Éducation Politiques Sociétés, 2007.
- Gérard, Ch., 2004, *Contribution à une sémantique interprétative des styles : Étude de deux œuvres de la modernité poétique : Jacques Dupin et Gérard Macé. Sciences de l'Homme et Société*, Université de Toulouse 2 Le Mirail, 2004.
- Labov, W., *Sociolinguistique*. Traduit de l'anglais par Alain Kihm. Présentation de Pierre Encrevé, Paris, Le Minuit, 1976.
- Morin, E., *La méthode 1. La Nature de la Nature*, Paris, Seuil, 1977.
- Payet, J.-P., Rostaing, C., Giuliani F. (dir.), 2010, *La relation d'enquête*, Presses universitaires de Rennes, 2010.
- Rastier F., « Rhétorique et interprétation- Ou le miroir et les larmes, Texto ! juin 2002 [en ligne], disponible sur http://www.revue-texto.net/Lettre/Rastier_Rhetorique.html, consulté le 2 mars 2019.
- Rastier F., « Sémiotique et linguistique de corpus », *Signata* [En ligne], 1 | 2010, mis en ligne le 26 avril 2016, consulté le 25 avril 2018. URL : <http://journals.openedition.org/signata/278> ; DOI : 10.4000/signata.278