

# RÉFLEXION MÉTHODOLOGIQUE AUTOUR DE LA CONSTRUCTION D'UN OBJET DE RECHERCHE : LA DYNAMIQUE IDENTITAIRE CHEZ LES MIGRANTS AFRICAINS FRANCOPHONES AU CAP (AFRIQUE DU SUD)

Résumé de thèse soutenue à l'Université de Nanterre-Paris-X

Cécile B. Vigouroux

L'objet central de ce travail est une réflexion méthodologique construite à partir d'une recherche de terrain conduite au Cap, à l'extrémité ouest de l'Afrique du Sud, sur la dynamique identitaire chez des migrants africains francophones. L'une des particularités importante de cette étude tient au fait que ces francophones, qui sont aussi des locuteurs de plusieurs vernaculaires africains, se retrouvent dans une écologie sociolinguistique où l'anglais – la langue coloniale de prestige – et l'afrikaans – une langue coloniale stigmatisée – fonctionnent comme les vraies langues officielles à côté de plusieurs vernaculaires bantus du sud, également reconnus par la Constitution comme officiels (voir ci-dessous). Plus qu'un déplacement géographique, la migration en Afrique du Sud constitue, pour ces francophones, une expérience singulière de leur "africanité" dans un pays qui pendant plus d'un siècle s'est construit sur la séparation raciale et l'exclusion – politique, sociale et culturelle – de la population noire autochtone. Ces migrants arrivent en Afrique du Sud riches d'une expérience que les Sud-Africains noirs, compte tenu de leur histoire, ne peuvent avoir. Au sein de cette rencontre entre Africains c'est la rencontre avec le *même autre* mais c'est aussi la rencontre de l'expérience de l'ouverture face à celle de la fermeture. Comment dans cette écologie nouvelle le migrant africain francophone va-t-il construire son altérité ? Si la question identitaire est centrale pour le terrain analysé, elle l'est également pour le phénomène migratoire lui-même, qui peut être envisagé comme une mise à l'épreuve de l'identité d'un sujet et de l'identité de sa langue.

Le terme de *réflexion* contenu dans le titre de cette thèse est à envisager au sens de cheminement et de regard réflexif sur notre propre travail de chercheur. À travers la description et les résultats du travail de recherche, il s'agit d'examiner la démarche méthodologique qui les a produits. En rendant transparentes et intelligibles les étapes de la recherche, depuis l'élaboration du projet de recherche jusqu'à l'analyse des données langagières, nous voulons montrer comment s'opère, au fil du travail, la reconfiguration de l'objet d'étude. Si les réorientations sont courantes au cours d'un travail de recherche – notamment lorsqu'il dure plusieurs années – elles sont rarement exhibées. Dans notre cas, partant d'un travail initial sur les "représentations" linguistiques nous sommes passée à une réflexion plus restreinte et plus ambitieuse sur la notion "makwerekwere", hétéro-désignation stigmatisante utilisée par certains Sud-Africains pour nommer les migrants africains qu'ils soient francophones ou non. Retraçant ainsi le cheminement au fil duquel s'est opérée la transformation de notre objet de recherche, nous nous engageons – à

partir d'une interrogation sur des questions identitaires dans le contexte sud-africain – dans une réflexion sur la construction des données et des objets de savoir du sociolinguiste que nous sommes.

### **La dynamique identitaire en Afrique du Sud**

Tout au long de son histoire, depuis l'arrivée des premiers colons portugais jusqu'à la mise en place et l'institutionnalisation d'un système politique fondé sur la division raciale et raciste (l'apartheid), l'Afrique du Sud n'a cessé de (re)mettre en question la notion d'identité. En érigeant des frontières au sein même du pays, en classant et en divisant, l'Afrique du Sud peut être envisagée comme le lieu où se sont jouées toutes les variations, où se sont confondues toutes les ombres du dedans et du dehors, de l'"autochtonie" et de l'"étrangéité". De ces divisions et de ces classifications n'a jamais émergé un éloge de l'altérité : en marginalisant plus de 80% de sa population sur la base de critères raciaux et culturels, la société sud-africaine s'est bâtie sur l'idée qu'il ne pouvait y avoir de rencontre possible entre gens de langues et de cultures différentes. C'est ainsi qu'à l'intérieur du pays, entre les Noirs, les Blancs, les Colorés ("coloureds") et les Indiens chacun a été l'étranger d'un autre étranger. Si durant l'apartheid les êtres ont été réduits à des espèces classées selon des groupes raciaux distincts et hiérarchisés (*Population Registration Act*, 1950), les espaces ont quant à eux été réduits à des zones (*Group Areas Act*, 1950) : le centre-ville était réservé aux Blancs, les Colorés et les Noirs étant refoulés à la périphérie de la ville, dans les townships.

Il en va de la géographie comme de la politique linguistique menée pendant plus de deux siècles. Celle-ci a joué un rôle majeur dans la mise en place du système ségrégationniste élaboré par les théoriciens de l'apartheid. La singularité de la situation sociolinguistique de l'Afrique du Sud ne réside pas dans son multilinguisme (depuis 1994, onze langues sont reconnues comme officielles) mais est à chercher dans l'histoire de sa politique linguistique qui a conduit à la protection des structures du pouvoir blanc et à la redistribution des ressources économiques. La politique de la langue a créé des stigmates sociaux et culturels encore visibles dans la société sud-africaine d'aujourd'hui. Elle a influencé le rapport des locuteurs à leurs propres langues, avec une nette dévalorisation des langues indigènes parlées majoritairement par la population noire africaine, et construit leurs représentations de l'Autre et de sa langue.

C'est dans ce contexte social, linguistique, culturel et politique extrêmement complexe qu'arrive le migrant africain francophone. L'Afrique du Sud, notamment urbaine, lui rappelle en bien peu de points son expérience de l'Afrique tant dans son paysage extrêmement divers que dans la composition de sa population où les Blancs sont devenus depuis des générations des autochtones. La rencontre des migrants avec cette "autre" Afrique va s'accompagner pour eux d'une expérience singulière et parfois douloureuse de leur condition d'étranger et d'Africain.

### **La migration africaine francophone dans le contexte post-apartheid**

L'Afrique du Sud connaît depuis le début des années 90 une vague importante de migrations venues des pays francophones de l'Afrique. Plusieurs

raisons peuvent être avancées à cela. D'une part les changements politiques intervenus en Afrique du Sud avec l'élection de Nelson Mandela, ainsi que le potentiel économique que représente le pays, constituent un pôle d'attraction pour ces Africains en quête d'une vie meilleure. D'autre part les conflits qui ont secoué ces dernières années l'Afrique Centrale (les deux Congos) et l'Afrique des Grands Lacs (Rwanda, Burundi) ont entraîné, surtout à partir de 1994, l'arrivée massive de ces populations francophones dans cette partie du continent. Ces migrations ont aussi été encouragées par des facilités administratives offertes, jusqu'en 1995, aux ressortissants de pays noirs africains (Mali, les deux Congos, Côte d'Ivoire, etc.).

Il est difficile de donner une estimation du nombre de migrants "continentaux" en Afrique du Sud. S'il existe des chiffres, ils sont rares, peu précis et pas fiables. On peut néanmoins estimer que les migrants francophones représentent une minorité comparativement aux autres ressortissants étrangers africains présents en Afrique du Sud, notamment les Mozambicains et les Zimbabwéens. On estime leur nombre au minimum de l'ordre de 50 000 personnes en 1995-1996, la communauté Congolaise (ex zaïroise) étant la plus nombreuse.

Si l'on s'en réfère à nos enquêtes menées au Cap – relativement représentatives de la tendance générale – on peut dire qu'il s'agit d'une immigration plutôt masculine (81% d'hommes contre 19% de femmes), composée en majorité de célibataires. Relativement jeunes (58 % âgés entre 26 et 35 ans), les migrants ont un niveau d'instruction élevé avec 40% des enquêtés ayant atteint une scolarité de niveau lycée et 34% ayant été inscrits dans un cycle supérieur.

### **La pertinence de la notion de " communauté francophone " en Afrique du Sud**

L'une des premières questions qui s'est posée, dès l'élaboration du projet de recherche, est la pertinence de constituer un échantillon d'enquêtés sur la base d'un critère linguistique, le français, et sur celle de l'appartenance à un pays francophone africain (ex-colonies françaises et belges) quand l'on sait la diversité des situations socioculturelles et linguistiques que recouvrent les notions de *francophonie* et *francophone*. Postuler l'existence d'une communauté africaine francophone au Cap n'est-ce pas postuler une catégorie *a priori* et homogénéisante qui laisserait de côté la diversité socioéconomique, culturelle, religieuse (musulmans, chrétiens, animistes) et linguistique de ces populations, catégories construites par notre regard de chercheur sans considération des réalités du terrain ? Les premières enquêtes de terrain menées au début de la recherche ont révélé, chez les migrants interrogés, l'existence d'un fort sentiment d'appartenance à un " monde " francophone où l'expérience de la migration conduit à une recomposition identitaire qui passe – entre autres – par la revendication d'une " francophonité " dans une stratégie de valorisation et de différenciation par rapport à la population locale, notamment noire.

### **Un travail réflexif**

En prenant le parti de montrer un objet de recherche en train de se construire, nous voulons insister sur l'importance d'une démarche qui pense son objet en terme dynamique, opérant un va-et-vient entre objet de recherche et regard réflexif sur la fabrication de cet objet. L'exposition en détail des différentes étapes

du travail (l'enquête de terrain, la constitution du corpus, la transcription, l'analyse des données, le travail d'écriture), nous permet de montrer comment le chercheur que nous sommes construit au fur et à mesure les données.

Si nous parlons de *construction des données* (vs *recueil des données*) c'est que nous pensons que celles-ci ne préexistent pas au travail du chercheur mais qu'elles sont, à chaque stade de la recherche, configurées par ce dernier. Cette perspective invite à réfléchir sur les positions occupées par le chercheur d'un bout à l'autre de la *chaîne de fabrication des données* et à repenser la coupure traditionnelle faite entre travail de terrain, travail de transcription et travail d'analyse. Si les diverses positions occupées par le chercheur correspondent à des étapes du travail chronologiquement distinctes, elles sont à envisager en terme de continuum. Il n'y a en effet jamais de coupure nette entre toutes ces positions, chacune est, en partie, le prolongement des positions antérieures. Un transcripteur, par exemple, qui a été aussi l'enquêteur puisera insensiblement dans sa mémoire de sujet-participant de l'entretien pour " façonner " à l'écrit la parole de l'enquêté qu'il retranscrit.

La réflexion méthodologique conduite ici n'a jamais constitué un objet de recherche de départ mais s'est imposée au cours du travail, notamment en raison de la difficulté à traiter les données construites sur le terrain. Cette difficulté relevait pour une part de la diversité des dispositifs d'enquête utilisés (l'entretien dirigé, semi-dirigé et l'observation) et de la complexité à les articuler pour le traitement des données langagières. De retour du terrain, l'une des questions récurrentes était celle des choix à faire pour procéder à l'analyse de données construites à partir de dispositifs aussi différents : peut-on traiter les données de l'entretien dirigé au même titre que celles de l'entretien semi-dirigé ? L'autre raison de la difficulté de l'analyse des données tenait à l'hétérogénéité du corpus d'enquête (cf. *infra*) composé de locuteurs qui n'avaient que peu de points communs à part le fait d'être Noirs africains et d'être originaires d'une zone francophone de l'Afrique.

### **Une démarche empirique**

Notre démarche se veut résolument empirique dans le sens où elle ne projette pas de cadre d'analyse *a priori* mais part de l'étude des données pour proposer des outils qui sont pertinents à leur analyse. Les choix théoriques que nous avons faits et les outils d'analyse utilisés ont répondu à notre volonté de faire surgir les données dans toute leur complexité – en les " dénaturant " le moins possible – et d'atteindre un certain degré de généralité qui ne soit pas un appauvrissement de ces dernières. Cette démarche se traduit, d'un point de vue théorique, par l'hétérogénéité des disciplines convoquées : la sociolinguistique, l'ethnographie de la communication, la linguistique interactionnelle, l'analyse conversationnelle, et l'analyse de discours. Les données sont ainsi travaillées, tout au long de la thèse, à partir de procédures d'analyse diverses qui marquent autant d'étapes dans notre réflexion.

Le travail de mise à nu de toutes les strates successives de la recherche nous permet d'aborder des questions importantes pour la linguistique. Notre thèse s'apparente à une sorte d'hypertexte où sont soulevées, à la lumière des préoccupations de recherche qui sont les nôtres, des notions comme celle de corpus,

de transcription, de catégories linguistiques et extralinguistiques, d'écriture des données. Nous exposons quelques-uns des points développés dans notre travail.

### Une écriture polyphonique

Notre réflexion sur les différentes postures occupées par le chercheur au cours de la recherche s'articule aussi à partir d'un travail sur les voix (*voicing*), celles qui parlent et celles dont on parle dans la thèse. Comment coexistent toutes ces voix et comment participent-elles à la construction des objets de savoir ? Notre texte est à bien des égards polyphonique : il y a la voix des enquêtés à travers les discours reproduits, il y a celle du chercheur de terrain qui décrit et décrypte le travail d'enquête, il y a celle du transcripateur qui pourrait métaphoriquement être vu comme un " passeur de voix ", il y a celle de l'enquêtrice qui interroge, et enfin celle de l'analyste. La polyphonie de toutes ces voix est rendue lisible à travers " la mise en texte " que nous en faisons : utilisation de polices différentes quand il s'agit du discours des enquêtés ou de celui du narrateur. À la polyphonie de ces voix s'ajoute la pluralité des écritures à travers l'utilisation de matériaux textuels hétérogènes : cartes géographiques, textes juridiques, citations, chronologies, articles de presse, notes de terrain, extraits de correspondance Email etc. Ces différentes instances énonciatives visent à apporter un éclairage différent sur le savoir en construction où le discours intime (ex : notes de terrain) se mêle au discours institutionnel pour produire du sens. Notre écriture oscille ainsi perpétuellement entre un " je " et un " nous ", entre récit personnel et discours scientifique descriptif et analytique.

L'architecture donnée à la thèse participe également de ce travail d'écriture. L'organisation textuelle invite à une lecture à deux niveaux : une première, linéaire, où le lecteur se déplace au fil des pages recto qui constituent le corps du texte et une seconde, en miroir, où la page verso vient éclairer le discours présenté en page de droite. Les matériaux figurant sur la page gauche du texte (cartes, photographies, fiche aide-mémoire, correspondance Email) constituent un " hors texte " qui, s'ils échappent à l'analyse, participent à la réflexion sur la construction des données.

À travers cette attention portée à l'écriture – sous différentes formes – nous voulons montrer comment celle-ci structure le savoir et participe à l'émergence de la connaissance.

### Une approche dynamique du corpus

La présentation de notre cheminement de recherche s'articule à partir d'une réflexion sur la notion de corpus. Nous avons, dans ce travail, une conception dynamique du corpus dans le sens où nous l'envisageons comme sujet à recomposition. C'est pourquoi, plutôt que de parler de corpus au singulier – comme c'est généralement le cas – nous en parlons ici au pluriel en montrant comment chacun des corpus que nous distinguons – trois au total – se structure par rapport à l'autre, et comment tous se complètent en questionnant de manière différente les données.

Le premier corpus à avoir été chronologiquement constitué est le **corpus d'enquête**. Il représente une sorte de " supra-corpus " à partir duquel ont été élaborés les deux autres. Construit durant la période d'enquête, il a été élaboré à

partir des objectifs initiaux de la recherche (travail sur les pratiques langagières et les “représentations” linguistiques des migrants) et des contraintes du terrain (difficulté de rencontrer les enquêtés, dangerosité de certaines zones, etc.). Il a été constitué à partir de trois dispositifs d’enquête : l’entretien dirigé, l’entretien semi-dirigé, et l’observation (participante et non participante). L’échantillon des enquêtés a été constitué à partir de différents critères (origine nationale, âge, niveau scolaire, durée de séjour dans le pays d’accueil, emploi exercé au Cap, etc.) afin de mesurer l’incidence des facteurs extralinguistiques sur les pratiques langagières et les “représentations” linguistiques des migrants. C’est au total 148 enquêtés qui ont été interviewés, originaires de 12 communautés nationales différentes (Bénin, Burundi, Cameroun, Congo Brazzaville, Congo RDC, Côte d’Ivoire, Gabon, Guinée Conakry, Mali, Sénégal, Rwanda et Togo). Ce corpus représente environ 100 heures d’entretiens enregistrés auxquels s’ajoutent 21 heures de productions langagières en situation (ex : enregistrement d’offices religieux, conversations spontanées) et des notes ethnographiques. En intégrant le travail d’enquête dans une réflexion plus vaste sur la notion de corpus, (le terrain est souvent présenté dans les travaux de recherche comme une partie autonome, détachée du reste) nous voulons insister sur le terrain comme lieu de configuration des données.

Après un premier travail de sélection des matériaux, opéré à partir d’une redéfinition des objectifs de recherche (abandon du volet sur les pratiques langagières en contexte) et d’un tri des matériaux (les notes d’observation ne sont pas retenues, les entretiens à plus d’un interlocuteur sont écartés, etc.) nous avons constitué le **corpus ressource**. Nous avons choisi de l’appeler “ressource” parce qu’il constitue une “banque de données” dans laquelle nous allons puiser pour mettre en valeur certains points de notre réflexion sur les données. L’échantillon de ce corpus est plus restreint que le précédent : 124 locuteurs, équivalent à 90 heures d’enregistrement. Ce second corpus correspond à une première étape d’archivage des matériaux sonores (ex : constitution de fiches aide-mémoire pour chaque entretien) qui ont été structurées et systématisés dans une base de données informatique (*Visual dBase*). L’outil informatique donne au chercheur une vue synoptique de son corpus (d’autant plus quand il s’agit d’un corpus vaste et hétérogène comme le nôtre difficilement saisissable dans sa globalité) et lui permet d’élaborer des premières hypothèses rapidement vérifiables (corrélations entre plusieurs facteurs extralinguistiques par exemple). En cela, on peut dire que la transformation du matériau oral en données informatiques constitue une première étape de l’analyse des données.

Dernier corpus à avoir été constitué, le **corpus d’analyse** est celui à partir duquel vont être analysées les données langagières (3<sup>e</sup> partie de la thèse). Il procède à nouveau d’un travail de sélection opéré cette fois à partir d’un cheminement d’analyse qui a conduit à un repositionnement théorique et à une réorientation des perspectives de recherche. Ce corpus, de nature tout à fait différente des deux autres, est le fruit d’un long travail de maturation sur les données et d’une remise en question de notre démarche fondée, jusqu’ici, sur la constitution d’un échantillon à partir de facteurs extralinguistiques. L’hétérogénéité des données, la diversité des enquêtés et la singularité de chaque entretien rendaient difficile la constitution d’un corpus à partir de facteurs extralinguistiques : quel(s) critère(s) de sélection

privilégier plutôt qu'un autre ? La difficulté d'élaborer un corpus cohérent et homogène, suffisamment restreint pour mener des analyses qualitatives, nous a conduit à changer notre approche du corpus : plutôt que de l'appréhender à partir des locuteurs, c'est-à-dire sélectionner des enquêtés selon des critères extralinguistiques, nous sommes partie des données linguistiques elles-mêmes. Ce changement de perspective qui marque un tournant majeur dans notre appréhension des données, s'est accompagné d'une (re)mise en question de la pertinence des catégories élaborées lors du travail d'enquête. En construisant *a priori*, c'est-à-dire avant même l'analyse des données linguistiques des catégories extralinguistiques (" Congolais ", " agent de sécurité ", " migrant en Afrique du Sud "), le chercheur de terrain que nous avons été postulait leur pertinence sans en interroger leur validité. De plus, en pré-catégorisant nous présupposons une parfaite adéquation entre les catégories construites par l'enquêteur-analyste et celles élaborées par l'interviewé. À partir de l'analyse discursive de plusieurs séquences d'entretien nous montrons qu'il n'y a pas toujours coïncidence entre les catégorisations élaborées par l'enquêté et celles construites par l'interviewer, notamment dans un contexte interculturel comme le nôtre. Un des exemples cité est celui de la conception différentielle de l'appartenance : pour ToMoh né et grandi au Burkina Faso, de langue première mosi et de parents togolais, elle renvoie à une " origine du sang " alors que pour l'interviewer il s'agit d'une " origine du sol ". Le danger d'une démarche reposant sur la catégorisation *a priori* est d'entraîner une description circulaire, l'analyste cherchant à faire coïncider ce qu'il enregistre et observe avec les cadres catégoriels qu'il aura lui-même élaborés.

Si la pertinence des catégories extralinguistiques est à interroger, il en va de même pour les catégories linguistiques utilisées dans notre guide d'entretien pour le travail sur les " représentations " linguistiques. L'analyse de trois séquences révèle des définitions très contrastées, chez chacun des enquêtés et chez l'enquêteur, de la notion de " langue maternelle ". On peut alors se demander comment à partir de catégories floues, qui n'émergent pas du discours de l'enquêté mais qui lui sont imposées par les questions de l'enquêteur, on peut produire une analyse pertinente qui rendrait compte des " représentations " linguistiques des interviewés ? L'entretien ou le questionnaire, généralement utilisés dans les travaux sur les représentations, est-il le plus approprié pour accéder aux discours sur la langue, même si les occasions sont rares de recueillir des discours spontanées – c'est-à-dire non provoqués par la situation d'interview – sur la langue ? Toutes ces questions invitent à se demander si les dispositifs d'enquête utilisés pour accéder aux discours sur la langue ne révèlent pas davantage les mythes de la langue chez linguiste-enquêteur que les " représentations " linguistiques des interviewés.

C'est dans le but de ne pas présupposer un cadre interprétatif sous-jacent à l'analyse des données linguistiques que nous décidons de constituer un corpus qui abandonne les catégories construites *a priori*. Cette nouvelle démarche s'accompagne d'une réorientation de notre objet d'étude où la dynamique identitaire des migrants n'est plus appréhendée à partir des " représentations " linguistiques mais à partir de l'étude de l'hétéro-dénomination " makwerekwere ". Terme à connotation fortement péjorative – équivalent du français *barbare* – il apparaît, d'un point de vue linguistique, comme le symbole le plus fort de la stigmatisation des

étrangers africains en Afrique du Sud. Le corpus d'analyse se compose de 30 séquences<sup>1</sup> d'entretien où l'enquêté, sur sollicitation de l'enquêteur, donne sa définition du terme "makwerekwere". Les séquences ont été retenues en fonction de leur degré d'explicitation (les réponses laconiques de type oui/non ont été écartées), de leur représentativité ou, pour trois d'entre elles, de leur singularité par rapport à l'ensemble du corpus ressource.

### **La transcription : enjeux théoriques d'une pratique**

Le travail de transcription est envisagé ici comme le second lieu, après le terrain, de la configuration des données. Notre réflexion est marquée par nos préoccupations sur le corpus recueilli, du champ dans lequel nous nous inscrivons : l'étude du discours et par les positions que nous assumons dans la chaîne de construction des données : enquêteur, transcripteur, analyste, scripteur. En montrant comment nos transcriptions relèvent de choix théoriques mais sont également influencées par notre position d'enquêteur, nous voulons montrer que la séparation des rôles (enquêteur-transcripteur-analyste) n'est pas aussi clairement tranchée qu'on le laisse souvent entendre. À travers cette réflexion sur la transcription, nous souhaitons problématiser le rôle du transcripteur dans la *fabrication des données*, celui-ci étant trop souvent relégué à un rôle subalterne (cela est d'ailleurs particulièrement visible dans la répartition des tâches où le travail de transcription est souvent confié à un "apprenti"/étudiant, l'analyse étant réservée au "maître"/linguiste). Cette réflexion sur le travail de transcription s'articule à partir de l'examen critique et comparatif des trois étapes de transcription qui ont jalonné le travail. Ces étapes sont envisagées comme une mémoire textuelle d'un objet d'analyse en construction. En procédant ainsi, nous voulons insister sur le fait que la transcription ne relève pas de choix *a priori* mais s'inscrit directement dans la dynamique de l'analyse. En cela, une transcription témoigne davantage d'un **état d'analyse** qu'elle ne constitue une version définitive et intouchable.

L'examen comparatif des trois versions de transcription permet de mettre en valeur un certain nombre de problèmes que nous avons eu à résoudre au cours du travail et témoigne également en filigrane d'une certaine idéologie de la langue qui sous-tend certains choix opérés. Nous citerons, à titre d'exemple, la prégnance de l'écrit dans notre première transcription (ex : conventions de transcription utilisées : virgule pour marquer une pause, point d'interrogation pour un questionnement, etc.) qui marque une représentation de la langue comme forme avant tout scripturale. En transcrivant, nous faisons plus qu'emprunter à l'écrit son système de classification nous construisons, sur le modèle de l'écrit, des catégories linguistiques et discursives. L'organisation physique du discours sur la page en est un exemple révélateur : le retour à la ligne organise le discours selon des modalités qui appartiennent à l'écrit et qui de ce fait construisent des catégories qui lui sont propres, indépendamment de la parole proférée.

Un des autres points soulevé, ayant plus directement trait à la spécificité de notre corpus, est les choix à opérer pour rendre compte de variétés de langue non

---

<sup>1</sup> " Séquence " est entendu comme suite d'éléments discursifs et envisagé comme découpage de l'interaction selon un critère thématique.

standard. Faut-il privilégier l'aménagement orthographique (ex : aphérèse : *dicap* pour *handicap* ; liaison impropre : *elle est plus* : *elle est zutile*) pour rendre au plus près l'oralité des locuteurs ? Le danger de tels " trucages orthographiques " n'est-il pas de conduire à une stéréotypie péjorative des locuteurs, notamment quand ceux-ci appartiennent à une couche sociale défavorisée ? Il est à noter, dans l'histoire littéraire française par exemple, que le recours à l'aménagement orthographique a souvent été utilisé pour représenter le parler des couches populaires et de ce fait a toujours eu une connotation péjorative dans la description donnée, par les auteurs, des locuteurs pratiquant cette variété de langue. C'est pourquoi il nous semble important que le transcripateur soit sensible aux représentations que ses choix de transcription risquent de construire des locuteurs dont il représente le discours. On voit donc qu'à travers le passage d'un état sonore à un état graphique la transcription interroge la langue dans sa dimension sociale, du côté de l'émission de la parole (l'enquêté) et de la réception par un tiers absent (le lecteur). Le travail de transcription, loin de se limiter à des problèmes techniques, engage d'abord le transcripateur et plus généralement le linguiste dans un questionnement théorique sur les données et leur traitement.

### **Le travail sur les données : le traitement de " makwerekwere "**

L'analyse de " makwerekwere " se présente chronologiquement comme la dernière étape de notre réflexion sur la construction des données. Élaboré au fur et à mesure du travail, l'objet de recherche " makwerekwere " constitue l'aboutissement de notre réflexion méthodologique et marque le début d'une nouvelle orientation de notre réflexion sur la construction identitaire des migrants africains francophones dans le contexte post-apartheid. L'identité n'est pas envisagée ici d'un point de vue substantialiste mais comme le produit des diverses situations d'interaction sociale et linguistique dans lesquelles sont engagés les locuteurs, d'un point de vue macro – le contexte sud-africain – et d'un point de vue micro – la situation d'entretien. Notre cheminement de recherche a en effet montré la nécessité d'envisager les discours dans une perspective interactionnelle, c'est-à-dire analysés à la lumière de leur contexte d'émergence (l'entretien) et comme le produit d'une co-construction entre enquêteur et enquêté. Le dernier chapitre de notre thèse est à envisager en terme programmatique où sont explorées des pistes possibles d'analyse qui tiennent compte de toute la réflexion menée au cours du travail sur les données. Elles sont autant de points de vue différents qui visent à saisir les discours dans leur hétérogénéité.

La première piste explorée est celle de l'analyse contrastive des séquences qui permet de réfléchir sur les conditions d'émergence du terme " makwerekwere " et d'élaborer une série d'hypothèses sur les modalités de son interprétation par les enquêtés. Il apparaît, à travers cette analyse, une grande hétérogénéité d'une séquence à une autre à la fois dans les réponses des interviewés (ex : large disparité dans le volume langagier), dans l'organisation des séquences (variation du canevas des questions) et dans les questions de l'enquêteur (différentes formes interrogatives, hétéro-catégorisation plurielle des étrangers : *étrangers africains, étrangers, étrangers noirs*, références multiples aux Sud-Africains : *certaines Sud-Africains, les Sud-Africains, une partie des Sud-Africains*, etc.). Les variantes

relevées sont examinées à partir de facteurs externes à l'entretien : 1. la modalité relationnelle entre enquêteur et enquêté (ex : relation privilégiée vs relation distanciée) ; 2. le temps de l'enquête (différence significative entre les premières interviews et les dernières) ; 3. la compétence linguistique de l'enquêté (cette hypothèse s'est révélée peu fructueuse) et de facteurs relevant plus directement du processus interactionnel : forte dépendance (syntaxique, pragmatique et lexicale) du couple question/réponse. L'analyse contrastée des séquences met en évidence la nécessité d'analyser les discours dans leur dynamique interactionnelle afin de comprendre comment s'élabore et se déploie la catégorie " makwerekwerek ". Cette perspective nous semble d'autant plus importante que la définition du terme n'est jamais initiée par l'enquêté mais sollicitée par l'enquêteur dans un contexte discursif structurant la réponse.

La seconde piste explorée est l'étude de séquences où " makwerekwerek " émerge " spontanément " au cours de l'entretien, c'est-à-dire sans être initié par l'interviewer. Il s'agit de voir si la catégorie est construite différemment lorsqu'elle n'est pas le produit d'une sollicitation. Le premier constat que l'on peut faire est que le terme émerge très rarement hors questionnement (1 fois au cours des 30 séquences retenues). Cela peut relever d'un refus, de la part de l'enquêté, d'être catégorisé par un terme à forte connotation péjorative ou plutôt du processus interactionnel de l'entretien : le caractère directif des questions ne favorise pas l'émergence de " makwerekwerek ", les thèmes abordés donnent peu d'occasion à l'enquêté d'utiliser le terme ou l'enquêté méconnaît les procédures de sollicitation (*elicitation*) utilisées par l'enquêteur. L'analyse de " makwerekwerek " hors entretien met en évidence les procédures catégorielles (auto-désignations, hétéronominations) mises en œuvre par les enquêtés. Il apparaît ainsi que le terme est perçu négativement et associé exclusivement à la population noire sud-africaine qui l'utilise. L'analyse des contextes d'émergence de " makwerekwerek " a l'avantage de montrer, contrairement aux séquences en entretien, l'accomplissement discursif des catégorisations, c'est-à-dire le remodelage des désignations au fil du discours.

Enfin, la troisième piste d'analyse suggérée est celle d'une étude transversale thématique. Il s'agit ici de répertorier les différentes significations attribuées à " makwerekwerek " et de les analyser à la lumière de facteurs extralinguistiques (origine de l'enquêté, temps de séjour en Afrique du Sud, etc.). Celle-ci montre que le terme recouvre des sens différents : il renvoie à une absence de compétence linguistique (celui qui ne parle pas la langue locale), il est synonyme d'*étranger* (avec ou sans connotation péjorative), il fait référence à l'espace territorial (celui qui n'est pas de l'Afrique du Sud), etc. L'étude transversale thématique présente l'avantage de donner une photographie d'ensemble du corpus et nous permet de repérer les tendances qui s'en dégagent. La prise en compte de facteurs extralinguistiques nous permet d'échafauder des hypothèses sur la perception des migrants de leur propre " statut " d'étranger et de Noir Africain en Afrique du Sud. Elle permet également de voir les stratégies qu'ils déploient pour réagir au rejet dont ils sont l'objet de la part d'une partie de la population locale. Néanmoins, la recherche de corrélations entre facteurs extralinguistiques et discours produits sur " makwerekwerek " s'est révélée, à ce stade de l'analyse, peu fructueuse. La singularité des discours, l'hétérogénéité des enquêtés (du point de vue de leur

appartenance, de leur situation dans le pays d'accueil, etc.) rendent difficile l'établissement de telles corrélations.

Les lectures plurielles que nous proposons d'un corpus comme " makwerekwere " visent à tisser des réseaux de sens pour faire surgir les discours dans leur complexité. Celle-ci nous semble pouvoir être saisie en **prenant compte** et en **rendant compte** des différentes strates contextuelles qui participent à leur construction.

Notre étude a voulu montrer la nécessité, pour tout travail de recherche, d'établir un dialogue permanent entre l'objet d'étude, la situation d'enquête, les acteurs comme individus réagissant à la situation, et le chercheur lui-même comme observateur et acteur. C'est une invitation à engager une réflexion épistémologique sur les objets de savoir produits en sociolinguistique.