

Le français contemporain en Afrique et dans l'Océan Indien : usage, variétés et structure

Chantal Lyche et Ingse Skattum

Université d'Oslo

Dans ce travail, nous nous proposons de décrire le projet *Le français contemporain en Afrique et dans l'Océan Indien : usage, variétés et structure* (CFA)¹, qui a pour objectif principal la description d'un ensemble de variétés de français pour un ressourcement empirique et un renouvellement des descriptions phonologiques, syntaxiques et sociolinguistiques des usages. Une équipe internationale de 18 chercheurs (11 d'Europe et 7 d'Afrique/Océan indien) envisage sur la base d'un protocole d'enquête commun de décrire 8 variétés de français (Côte d'Ivoire, Burkina Faso, Cameroun, Centrafrique, Sénégal, Mali, Maurice, Réunion) afin de proposer une analyse comparative de ces variétés. Nous partons de l'hypothèse qu'il existe un français panafricain qui transcende les différents contextes sociolinguistiques et les différents substrats, et que cette "variété" panafricaine s'inscrit dans un cadre plus vaste, le français panlectal (Chaudenson *et alii* 1993).

Les huit variétés qui constitueront notre base empirique seront décrites et analysées selon trois axes: phonologie, syntaxe et sociolinguistique, mais sans cloisons disciplinaires nettes, l'étude des interfaces menant à une meilleure compréhension globale. Notre ambition est de créer, à terme, une grande base de données numérisées, offrant une transcription orthographique alignée au signal. Nous nous concentrerons dans cette présentation sur les volets phonologique et sociolinguistique du projet.

La phonologie dans CFA

Le projet CFA s'inscrit au sein du projet PFC (www.projet-pfc.net) pour la partie phonologique, tout en s'en distanciant afin de prendre en compte les particularités locales.

¹ De son nom anglais *Contemporary French in Africa and the Indian Ocean*.

Cette section sera donc consacrée au projet PFC et à sa méthodologie avant d'envisager brièvement l'originalité phonologique de CFA.

Le PFC: le protocole d'enquête et les enregistrements

Le projet PFC se propose sur la base d'un protocole unique de rassembler des données phonologiques comparables à travers la francophonie. L'un des objectifs du projet (Durand, Laks et Lyche 2005; Durand et Lyche 2003)² concerne la constitution d'une base de données de référence sur le français oral à partir d'une méthodologie commune. La méthodologie adoptée au sein du projet s'est construite sur des expériences de français langue première et nous nous interrogerons ici sur les défis qu'elle pose au contexte africain.

PFC s'inscrit dans la tradition des grandes enquêtes phonologiques (Laks, 2003) avec une cinquantaine de points d'enquête en France (y compris les DOM) et hors de France (Afrique, Belgique, Canada, Louisiane, Suisse) choisis pour assurer une grande représentativité aussi bien géographique que sociologique en s'efforçant d'équilibrer les paramètres âge et sexe et de cibler des témoins provenant de différents milieux socio-économiques. Dès sa conception le projet s'est voulu francophone avec la Belgique et la Suisse parmi les premiers points d'enquête. L'incorporation d'une vingtaine de points d'enquête hors de l'Hexagone donne au projet une dimension diatopique d'une grande originalité, mais soulève un ensemble de questions sur sa faisabilité : peut-on réellement envisager de mener une enquête en Louisiane, au Sénégal ou à la Réunion sur exactement les mêmes principes que ceux qui ont été adoptés pour la France hexagonale ? La difficulté provient de ce que PFC se caractérise par l'application dans tous les points d'enquête d'un

²Le site du projet offre une description complète du protocole, des bulletins réguliers et l'annonce des conférences. Voir Durand et Lyche (2003) pour un aperçu plus complet de la méthodologie de PFC.

protocole unique visant à rassembler des données provenant de quatre registres distincts.³ Il en ressort des données robustes et strictement comparables, ce qui fait de PFC un projet pilote et permet la constitution d'une grande base de données inégalée à ce jour.

La définition des quatre registres s'inspire des grands travaux de Labov et chaque témoin est soumis à deux tâches de lecture (une liste de mots et un texte) ainsi qu'à deux entretiens, un entretien semi-dirigé et une conversation libre, chacun d'une durée de 20 à 30 minutes. Le protocole est défini *a minima*, c'est-à-dire que nous demandons à tous les locuteurs de lire la liste de mots ainsi que le texte, mais les particularités phonologiques locales peuvent être testées plus en détails à l'aide de liste ou de texte complémentaires. Les témoins sont recrutés par le biais de réseaux denses (Milroy 1980) et selon les points d'enquête, nous opérons avec un ou deux enquêteurs dont l'un au moins est un familier des témoins. L'entretien semi-dirigé est conduit par une personne étrangère au réseau étudié mais qui est introduite par l'un des membres du réseau (Durand et Lyche 2003). Le choix des témoins s'effectue en vertu de leur degré d'ancrage dans la communauté et nous favorisons les personnes qui ont fait toute leur scolarité dans le lieu de l'enquête. Nous nous efforçons également d'enregistrer plusieurs générations au sein d'une même famille afin de tester les changements phonologiques en temps apparent ainsi que les phénomènes d'effet de l'âge (Durand et Lyche, 2003).

Le travail de lecture exclut cependant tout témoin dont le niveau d'éducation ne lui permet pas d'effectuer cette tâche, ce qui réduit considérablement notre couverture diastratique. Si ce choix méthodologique semble légitime dans les zones francophones européennes puisqu'il garantit une certaine cohésion des enquêtes, il faut s'interroger sur son bien-fondé dans des zones où le français n'est maîtrisé qu'oralement. Tel est le cas par exemple en Louisiane où le protocole a été adapté afin de prendre en compte des populations

³ Gadet (2000: 164) par exemple, dans son étude des corpus sur la négation regrette l'absence de possibilité de comparaison : « Mais aucun auteur n'effectue de comparaison entre zones. Il n'y a pas à ma connaissance, de tentative pour comparer un même type d'enregistrement d'une région à l'autre. »

de français langue maternelle, mais qui ne le lisent que très peu ou pas du tout. En Louisiane, nous partons de l'anglais et demandons aux locuteurs de traduire en français la liste de mots, nous leur demandons également de raconter quelques histoires bien connues en pays cadien (Klingler 2006, Lyche 2006).

Une fois les données enregistrées, elles sont numérisées et transcrites sous PRAAT⁴ sur une tire spécifique avec alignement au signal. Toutes les transcriptions sont effectuées en graphie standard sans aucune modification, sans aucun diacritique qui permettrait d'effectuer un lien avec la prononciation. Seuls les éléments absents ne sont pas transcrits, comme par exemple le *ne* de la négation. Nous avons suivi ici une pratique adoptée par de grands corpus oraux afin d'assurer la mutualisation des données, jugeant que toute autre prise de position impliquerait un biais théorique, un début d'analyse. Dans les régions où l'alternance codique est fréquente, nous transcrivons le texte dans la L2 (par exemple anglais au Canada ou créole à la Réunion), et nous faisons appel aux lexiques locaux lorsque cela s'avère nécessaire. Pour la Côte d'Ivoire par exemple, nous avons eu recours au *Lexique français de Côte d'Ivoire* (Lafage 2003-2004) et repris les mêmes principes de transcription pour des termes qui ne figuraient pas dans les lexiques (Boutin 2006). Le travail de transcription facilite l'accès aux données et prépare à l'analyse mais ne doit en aucun cas constituer une analyse. Le retour systématique au signal s'impose pour toute approche phonologique et est souhaitable dans bien d'autres domaines. Ce retour au signal se fait très simplement sous PRAAT, soit à travers l'écoute d'un morceau transcrit, soit grâce à la visualisation de la courbe de fréquence fondamentale (F0) et des formants vocaliques. Au-delà de la possibilité de retrouver le signal correspondant à n'importe quelle séquence, nous fournissons au chercheur un ensemble de données statistiques qui peuvent être obtenues à partir d'une politique de codage.

⁴ Voir <http://www.fon.hum.uva.nl/praat/>

Les axes phonologiques de PFC

PFC propose trois grands axes de recherche qui doivent fournir une vision globale de la phonologie du français : (1) les systèmes phonémiques de toutes les variétés de français, (2) le schwa et (3) la liaison, (2) et (3) représentant les phénomènes les plus caractéristiques et les plus étudiés de la phonologie française. Le profil phonologique de chaque locuteur est établi dans un premier temps sur la base de la liste de 94 mots lus dont les dix derniers sont des paires minimales. Le système ainsi élaboré est ensuite confronté à la lecture du texte qui reprend certains mots de la liste pour se voir finalisé après étude des deux conversations. Afin de faciliter l'analyse des systèmes phonémiques, la liste de mots et le texte sont également transcrits orthographiquement sous Praat et avec un alignement par mot dans le cas de la liste permettant ainsi une extraction automatique des structures formantiques, travail effectué sous la direction du laboratoire LPL à Aix-en-Provence.

Au-delà de la tire de transcription, le protocole recommande l'usage de deux/trois tires supplémentaires,⁵ servant à des fins de codage. La politique de codages développée au sein du projet s'appuie sur une méthodologie commune : il s'agit de fournir aux chercheurs un premier balayage des données mais en aucun cas une analyse. Afin d'atteindre cet objectif les codages sont aussi athéoriques que faire se peut, reposant sur un consensus observé dans la littérature scientifique. Sont codés sur des tires spécifiques le schwa, la liaison et dans une moindre mesure la prosodie. Le schwa et la liaison peuvent être approchés comme des phénomènes non gradients et nous nous concentrons sur leur présence/absence et sur les contextes droits et gauches, ce qui implique un degré de granularité de codage relativement grossier. Le texte dans son entier est codé pour chaque locuteur de la base et sont aussi codées trois minutes de conversations dirigée et libre pour schwa et cinq minutes des deux conversations pour la liaison. Le codage-schwa comporte quatre chiffres : (1)

⁵ Précisons qu'en aucun cas Praat ne limite le nombre de tires. Certains chercheurs utilisent jusqu'à 15 tires pour des analyses prosodiques (phonèmes, syllabes, pieds, etc.).

présence/absence, (2) position dans le mot, (3) contexte gauche, (4) contexte droit. Le codage-liaison indique la longueur du mot liaisonnant, la nature de la consonne de liaison si cette dernière est réalisée, la nature de la voyelle en cas de liaison avec consonne nasale, la nature de la liaison (enchaînée ou non, épenthétique). L'importance des contextes considérés pour le schwa a largement été mise en valeur par Dell (1973/1985) et pour la liaison, nous avons puisé dans les travaux tout aussi classiques de Delattre (1966). Le codage-schwa s'inscrit à la suite de tout *e* graphique mais également à la suite de toute consonne graphique finale prononcée. La prise en compte des consonnes graphiques finales s'explique par un double objectif : elle permet de relever la fréquence de schwas prépausals comme dans *bonjour-e*, tout comme elle offre des indications précieuses sur les simplifications consonantiques. Soit l'exemple à *table*: codé à *table1423*, il correspond à la prononciation [atablø] avec un schwa final réalisé. Le codage à *tab0453le* indiquera quant à lui que le schwa n'est pas prononcé, que la dernière consonne perçue est le [b] et que l'on observe une simplification consonantique. Nous verrons en 2.3 que ce codage peut s'étendre aisément à l'étude de la disparition du /r/ dans les variétés africaines du français. Le site PFC met à la disposition de tous les chercheurs des outils de recherche par code, il autorise également les requêtes selon un contexte orthographique précis, pour un point d'enquête précis, etc. Dans tous les cas, l'exemple s'affiche en contexte et il est possible non seulement d'avoir accès à la bande sonore mais également au signal par le biais d'une fenêtre Praat.⁶

Dans certains points d'enquête et pour quelques locuteurs par point d'enquête, nous proposons le codage de la prosodie. Praat permet l'affichage de la F0, ce qui donne une première approximation des mouvements mélodiques, mais la tâche de codage prosodique est bien plus ambitieuse que pour le schwa et la liaison et sa faisabilité peu évidente. Le domaine prosodique ne bénéficie pas du même consensus théorique de base qu'en phonologie

⁶ Le moteur de recherche autorise également de placer l'exemple considéré dans un contexte plus large, par ex. 10 secondes avant et 5 secondes après, etc.

purement segmentale (voir Lacheret et Beaujendre 1999 pour une revue d'un ensemble de modèles) et dans un premier temps, il nous a semblé prudent de nous limiter à l'étude des relations entre le schwa et la prosodie (Lacheret et Lyche 2006). Le codage est effectué sur une quatrième tire sous Praat mais les unités codées ne sont plus des mots graphiques mais des syllabes transcrites dans l'alphabet phonétique Sampa.

La base de données PFC comprend à ce jour 220 locuteurs répertoriés, transcrits, **codées** (accord avec données, un peu loin?) selon des principes identiques et le travail laborieux de transcription et de codage commence à porter ses fruits. Le protocole, la méthodologie générale du projet semblent suffisamment bien rodés pour être appliqués à l'étude des variétés africaines à une plus grande échelle, mais un élargissement diastratique ne saurait être mis en place sans réflexion méthodologique approfondie.

PFC et l'Afrique/l'Océan indien

La base PFC ambitionne de recueillir des données comparables pour toute la francophonie. En Europe, sont déjà représentées la Belgique et la Suisse, en Amérique du Nord, le Canada et la Louisiane. Pour la zone Sud, la base PFC inclut deux points d'enquête en Afrique (Côte d'Ivoire et Burkina Faso) et un point d'enquête à l'île de la Réunion. PFC a opéré certains choix méthodologiques afin de garantir la comparabilité des données (Boutin, Guèye, Lyche et Prignitz, à paraître). Les exigences du protocole ont imposé le recrutement de locuteurs scolarisés, capables d'effectuer les tâches de lecture avec une certaine aisance. Nous n'avons donc pas pour ces variétés modifié de quelque façon que ce soit le protocole, nous avons choisi d'éliminer les locuteurs qui ne répondaient pas aux critères recherchés. Ce choix, légitime dans un projet à dessein phonologique, ne l'est plus dès que l'on affiche des ambitions sociolinguistiques et syntaxiques. Au-delà d'une description générale de l'inventaire phonémique des locuteurs, les intérêts phonologiques du projet se déclinent en

trois points : le phonème /r/, la liaison et la prosodie. Bordal (2006) met largement à profit le codage du schwa tel qu'il est prescrit dans le protocole PFC dans son traitement du /r/ dans le français régional de l'île de la Réunion, ce qui lui permet de rendre compte de la présence/absence de la consonne en finale absolue de mot ou comme premier élément d'un groupe consonantique final.⁷ Cette stratégie de codage fournit un ensemble de données statistiques sur le fonctionnement du /r/ en coda de syllabe accentuable, mais elle est impuissante à relever toutes les autres instances d'absence, comme dans *pa'tir* par exemple. Le projet CFA s'engage à développer rapidement un codage alphanumérique du /r/ selon les mêmes principes que les autres codages en vigueur dans PFC.

Au sein des grands débats autour du traitement théorique de la liaison,⁸ revient régulièrement le rôle joué par la connaissance de la graphie sur la réalisation ou non de la consonne de liaison. Aucune étude jusqu'à présent n'a comparé les réalisations de populations lettrées à celles de populations non lettrées. Si l'on choisit d'inclure des témoins non lettrés, les données africaines seront d'une grande valeur non seulement en comparaison au français hexagonal mais entre elles. Dans la mesure où le degré d'appropriation du français varie d'un pays à l'autre, on pourrait s'attendre à observer tout un continuum dans la réalisation de certaines liaisons. La liaison présente un autre avantage, celui de constituer un lien direct avec la syntaxe et avec la prosodie, deux domaines centraux dans le projet.

La liaison met en question l'existence phonologique du mot en français, tout au moins comme unité accentuelle, or nous savons que toutes les langues africaines avec lesquelles le français est en contact sont des langues à tons. Si, en français, l'accent frappe la dernière syllabe d'un groupe à des fins démarcatives, l'accent dans les langues en contact avec le français n'a pas de fonction contrastive ni démarcative, mais plutôt une fonction expressive.

⁷ Nous ne nous attardons pas ici sur la structure syllabique, et il est bien évident que dans le cas de *morte*, le /r/ peut être considéré comme coda suivie d'une syllabe extramétrique.

⁸ Voir Durand et Lyche (à paraître) pour une description de la problématique et pour une revue des différents courants théoriques.

Dans ces conditions, la rencontre de deux systèmes prosodiques opposés donne comme résultat un système hybride et un membre de notre équipe, Béatrice Boutin (2006) a montré, sur la base de données PFC, comment le français ivoirien maintient des différences tonales. On sait par ailleurs qu'au Mali par exemple, les locuteurs disent pouvoir déterminer la langue première (L1) du locuteur à l'écoute de son français. Nous faisons donc l'hypothèse que le locuteur africain s'approprie le français à travers le filtre de sa propre langue et conserve dans cette L2 des distinctions tonales qu'il met à profit à des fins distinctives. Plus encore que les différences au niveau purement segmental (comme par exemple la disparition du /r/ avec allongement compensatoire ou non), ce serait dans le suprasegmental qu'il serait nécessaire de rechercher les caractéristiques phonologiques des variétés du français d'Afrique.

L'approche sociolinguistique

Dans la mesure où le protocole d'enquête sociolinguistique est toujours en voie de finalisation, nous nous contenterons dans cette section de présenter la méthodologie et les principes qui sous-tendent les choix proposés. L'approche sociolinguistique a comme objectif (1) d'assurer la validité des données à l'aide de critères de sélection adaptés aux objectifs et, (2) de mettre les données sociolinguistiques obtenues en rapport avec les traits phonologiques et syntaxiques définis à travers l'analyse des enregistrements.

Nous présenterons d'abord nos critères de sélection des points d'enquête et des témoins. Ensuite nous discuterons l'incidence des registres sur la forme linguistique et la nécessité d'ajouter à ces registres des genres écologiques (Gadet 2007). Pour terminer, nous présenterons le questionnaire sociolinguistique que nous proposons en complément aux autres outils d'enquête.

Sélection des points d'enquête

Les 8 points d'enquête ont été choisis pour illustrer trois types de contexte sociolinguistique. Une première distinction s'opère selon la présence/absence d'une langue locale majoritaire. Le Sénégal, le Mali et la Centrafrique ont chacun une langue majoritaire (wolof, bambara, sango) au niveau national, qui sert de langue de communication interethnique et qui, à divers degrés, cantonne le français dans le domaine formel. Dans ces cas, il y a *discontinuité* (Wald *et alii* 1973) entre le français et les langues locales, ce qui rapproche ce français du français standard. La Côte d'Ivoire, le Burkina Faso et le Cameroun sont caractérisés par l'absence d'une langue locale dominante au niveau national, absence favorisant le français comme *lingua franca*. Dans ces pays, il y a plutôt *continuité* entre les langues locales et le français. Le *corpus* (usage réel) du français, (qui s'oppose au *status*, c'est-à-dire la position, Chaudenson 2000) est plus important dans ces derniers pays, et surtout en Côte d'Ivoire (Knutsen 2007a,b). Le Mali se situe de l'autre côté du spectre, son corpus français étant parmi les plus insignifiants de l'Afrique (Skattum 2008). La Réunion et l'île Maurice constituent un troisième cas: le français, qui a été introduit dans des îles inhabitées, coexiste avec les créoles, qui se sont développés parallèlement au français. Ce type de contexte invite aussi les langues partenaires à entrer dans un continuum. Cependant, les trois types principaux englobent à leur tour des contextes particuliers. C'est pourquoi nous avons tenu à étudier plusieurs terrains pour diversifier nos données.⁹ Les situations décrites ne sont toutefois pas statiques et nous nous intéresserons aussi à la dynamique des rapports entre français et langues locales, et l'incidence que cette dynamique peut avoir sur l'utilisation du français et sur ses formes linguistiques.

Sélection des témoins

⁹ Notre travail se construit bien évidemment sur une vaste littérature que le cadre de cet article ne nous permet pas de citer.

Nous avons choisi pour le volet phonologique du projet de retenir les 4 registres de PFC. La transcription de ces données, l'annotation et la vérification représentent une lourde tâche. Il s'ensuit que, pour des raisons de faisabilité, il convient de limiter dans le protocole le nombre de locuteurs, dans notre cas 10-12 pour chaque point d'enquête,¹⁰ réduisant d'autant l'éventail des paramètres à retenir pour la sélection des témoins. Nous avons résolu de nous concentrer sur quatre paramètres, qui tous s'avèrent pertinents dans tous les points d'enquête concernés et qui sont donnés par ordre d'importance décroissante : (1) le niveau d'instruction; (2) l'âge; (3) la L1 ; (4) le sexe. Les choix opérés demandent justification, ce vers quoi nous nous tournons.

Tout d'abord, il est légitime de s'interroger sur nos préoccupations concernant le choix des locuteurs. Françoise Gadet remarque (communication à Oslo, le 28.08.07) qu'en sociolinguistique, l'échantillonnage des locuteurs a eu tendance à primer sur la typologie des situations ou des genres discursifs qui, pourtant, ont autant sinon plus d'incidence sur la forme de la langue. Nous reviendrons ci-dessous sur le rôle des genres dans notre corpus. Nous faisons cependant l'hypothèse que dans le contexte africain/Océan indien, les états de langues varient considérablement selon les locuteurs et qu'une description linguistique globale ne peut se passer des distinctions que nous proposons.

Le niveau d'instruction est le paramètre le plus déterminant pour la forme de français parlé ainsi que pour la conscience métalinguistique des locuteurs. Nous l'avons défini en fonction des diplômes obtenus et nous avons distingué trois niveaux. (Les catégories ci-dessous sont celles en vigueur au Mali; il faut évidemment les adapter aux systèmes locaux):

- CEP (6 ans = le premier cycle de l'école fondamentale)
- DEF (9 ans = le second cycle de l'école fondamentale/ le premier cycle de l'école secondaire/ le collège)
- Bac/ Bac+ (12 ans = l'école secondaire/le second cycle de l'école secondaire/ le lycée) plus éventuellement des études supérieures.

¹⁰ Chaque équipe garde évidemment la liberté d'accroître ce chiffre en fonction des moyens humains et financiers dont elle dispose.

Pour avoir une image plus complète et plus réelle du français parlé, il serait souhaitable d'inclure des témoins ayant appris le français en dehors de l'école. Nous n'avons cependant pas retenu cette catégorie comme obligatoire pour le protocole commun. Les équipes qui désirent inclure des locuteurs non scolarisés le feront en omettant les tâches de lecture.

L'âge est un paramètre indispensable pour vérifier l'évolution de la langue en temps apparent et pour tester les phénomènes de normalisation lorsque la situation socioéconomique du locuteur évolue favorablement. Nous avons défini trois catégories : (1) 20-39 ans; (2) 40-59 ans; (3) 60 ans ou plus. Pour juger du résultat de la scolarisation avons inclus les étudiants, qui représentent une strate sociale importante dans la première tranche d'âge, les 20-39 ans, constituant probablement la future élite du pays. Pour les 40-59 ans nous verrons ce qui leur a permis de pratiquer ou non le français, et ce qui reste de l'enseignement du français d'il y a 10 à 40 ans. Ils sont aussi parents: quels sont leur comportement et leur attitude envers leurs enfants en ce qui concerne l'apprentissage du français ? Enfin, les 60 ans et plus ont vécu l'ancien système scolaire, encore empreint de l'école coloniale. Pratiquent-ils encore le français, et de quel(s) français s'agit-il? Quelles sont aujourd'hui leurs attitudes envers le français et son rapport avec les langues locales ?

Les conversations libres entre pairs sont considérées comme plus révélatrices des traits linguistiques d'une tranche d'âge que les conversations en famille, mais nous n'imposons pas le choix de témoins pour ce registre. Il suffit de noter les rapports entre les interlocuteurs et de les prendre en compte dans l'analyse.

Il serait intéressant d'initier des enquêtes partielles sur les tranches d'âge, par exemple la langue des jeunes en ville, la langue des locuteurs âgés, etc. et de comparer les données entre les zones géographiques. A notre connaissance, ce terrain reste entièrement vierge en ce qui concerne l'Afrique.

La LI des locuteurs est un paramètre pertinent en Afrique, où les gens affirment qu'il leur est possible à travers l'accent de français de détecter la L1 d'un locuteur. C'est un domaine (avant tout phonologique/prosodique) qui a été peu exploité. Pour tester l'hypothèse de l'influence de la L1 sur la prononciation du français en Afrique, il suffit de confronter deux L1 africaines pour chaque point d'enquête, de préférence des langues de groupes linguistiques distincts. Il faut choisir des locuteurs qui pratiquent encore leur L1 et, pour les langues véhiculaires comme le wolof, le dioula, le bambara, le peul et le sango, choisir des locuteurs natifs et éviter des locuteurs parlant ces langues comme L2. Si possible, le travail de terrain devrait s'effectuer dans la région où la L1 en question est dominante et son influence sur le français pas encore (trop) « contaminée » par les langues véhiculaires. Cette approche exclut bien évidemment les strates les plus scolarisées, qui ont fait leurs études en ville.

Le sexe est un paramètre standard dans les études sociolinguistiques, mais semble moins pertinent dans le contexte africain. Faute d'hypothèse forte, nous avons décidé d'inclure, lorsque faire se peut, le sexe comme paramètre, mais de privilégier les trois autres critères.

Si on devait inclure les quatre paramètres mentionnés de manière systématique et enregistrer deux locuteurs dans chaque « case », on obtiendrait une grille de 48 locuteurs pour chaque point d'enquête, tâche irréaliste dès que l'on prend en compte le coût de la transcription. En guise de conclusion rappelons nos *trois paramètres* pour un protocole *a minima* : le niveau d'instruction, l'âge et la L1.

Registres et genres

Le protocole des 4 registres, appliqué à des contextes diversifiés à travers le monde francophone, permet de tester l'hypothèse d'un français panlectal. L'existence éventuel d'un français panafricain a été discutée par Manessy qui, en référant au concept de la sémantaxe,

parle des "mêmes catégorisations de l'expérience" dans "les parlers de populations appartenant à une même aire de civilisation" (Manessy 1995: 234). Hazaël-Massieux (1993) réfléchit sur les traits syntaxiques communs entre variétés du français et langues africaines parlées, mais cherche l'explication dans le contexte commun, qui est celui de l'oralité. Aucune étude systématique basée sur des données empiriques provenant de différents pays n'a, à notre connaissance, été faite pour tester l'hypothèse d'un français panafricain. Pour décrire les particularités africaines locales et tester cette hypothèse, le CFA doit adapter le protocole au contexte de l'Afrique au sud du Sahara et de l'Océan indien. C'est ce que nous nous proposons de faire en modifiant le contenu de l'entretien semi-dirigé et en ajoutant des genres écologiques et un questionnaire sociolinguistique.

L'entretien semi-dirigé intéresse les trois volets phonologique, syntaxique et sociolinguistique. Il peut porter sur différents sujets selon la fonction qu'on lui assigne. Dans notre optique, il servira à provoquer certaines constructions syntaxiques. On sollicitera ainsi un récit du passé ou de l'avenir pour les temps verbaux et les prépositions temporelles, un récit sur ce que le témoin ferait si ..., pour les constructions hypothétiques et une description d'un itinéraire et un lieu (domicile, travail) pour les prépositions spatiales et les qualificatifs. Les entretiens seront conduits pendant environ 30 minutes. Dans un premier temps, 10 minutes ciblées d'enregistrements seront transcrites, mais les enregistrements seront entièrement numérisés et accessibles pour analyse.

La conversation libre est le registre qui pose le plus de problème en Afrique et dans l'Océan indien, car le parler informel se fait normalement dans les langues locales. Son sort sera déterminé par chaque équipe selon le contexte local (français langue véhiculaire ou non). En complément de ou en substitution à ce registre peu adapté au contexte, nous allons introduire des *genres écologiques*, des discours "authentiques" non sollicités par le chercheur. Les genres écologiques permettent aussi dans une certaine mesure de rétablir l'équilibre des genres dans les corpus oraux, car ceux-ci sont, dans leur grande majorité, constitués d'entretiens semi-dirigés. Comme exemples de genres écologiques, citons : un cours magistral, un prêche, des débats à la télévision, etc.

Le questionnaire sociolinguistique

Nous prévoyons enfin un questionnaire sociolinguistique, qui aura deux fonctions: (1) collecter les informations sociolinguistiques nécessaires à l'interprétation des données linguistiques ; (2) mener une enquête sociolinguistique quantitative. La première fonction suppose que les témoins enregistrés répondent également au questionnaire. La seconde demande un échantillon de 200 à 400 personnes.

Le questionnaire doit poser des questions sur la vie linguistique de l'enquêté: langues parlées par lui-même et son entourage, informations à mettre en rapport avec la scolarisation, la/les profession(s), les langues des parents, les lieux d'habitation (les différentes régions parlant différentes langues). Viendront ensuite des questions sur ses attitudes linguistiques déclarées envers les différentes langues, quel(s) rôle(s) ont-elles ou doivent-elles joué/er et aussi des jugements sur la langue, sur la variation locale, sur la norme (centrale et endogène)... Nous prévoyons une pré-enquête pour tester les questions, et environ 400 témoins, avec environ 50% de retour. Sauf pour les témoins enregistrés, la passation doit être anonyme, pour ne pas infléchir les réponses.

Conclusion

Pour conclure nous rappellerons les éléments principaux du protocole *a minima* mis en place pour le projet CFA.

(1) Phonologie : nous suivrons le protocole PFC, mais développerons des méthodes appropriées pour le codage du /r/ et de la prosodie, deux domaines privilégiés pour les études phonologiques.

(2) Sociolinguistique : notre protocole définit les critères de sélection des témoins, les registres et genres à enregistrer et analyser.

L'échantillon minimal sera constitué de 10-12 locuteurs, choisis de manière systématique selon les deux paramètres du niveau d'instruction et de l'âge. La L1 sera obligatoirement prise en compte, mais de manière non systématique, tout en s'attachant à une certaine diversification: au moins 2 locuteurs natifs de 2 langues du terroir parmi les témoins.

Les langues doivent appartenir à 2 groupes linguistiques différents. Les enregistrements incluront les 4 registres du PFC plus des genres écologiques (discours formels et informels). L'entretien portera sur des thèmes choisis à des fins syntaxiques et durera 30 minutes environ. Dans les deux types de parole continue, la transcription se fera sur 10 minutes choisies.

La base de données ainsi constituée, soumise à des outils d'indexation phonologique et syntaxique, nous permettra d'atteindre notre objectif de renouvellement de l'empirie. Le protocole strictement appliqué dans tous les points d'enquête nous fournira des données non seulement originales mais comparables pour mieux cerner l'espace francophone dans sa totalité. Si la variation est incontournable, la linguistique moderne l'a trop longtemps négligée, se repliant derrière le caractère épars et non systématique des données souvent reliées à des erreurs de performance. Un projet comme CFA, par la rigueur de son protocole, a l'ambition de systématiser cette variation et de mettre à jour les facteurs qui l'animent en Afrique.

Références

Bordal, G. 2006. *Traces de la créolisation dans un français régional. Le cas du /R/ à l'île de la Réunion*, mémoire de DEA, Universitetet i Oslo.

Boutin, B. A. 2006. Les corpus PFC - Afrique : questions de transcriptions et de codages. In *Phonologie du Français Contemporain : données et enjeux théoriques*. Journées PFC. Paris, MSH. 3-4 février 2006. <http://www.projet-pfc.net/?accueil:bibiliopfc:parisfebruary2006>.

Boutin, B. A., Lyche, C. et G. Prignitz. à paraître. Les enquêtes PFC en Afrique. In *Bulletin PFC* 7.

Chaudenson, R., Mougenon, R. et Beniak, E. 1993: *Vers une approche panlectale de la variation du français*. Paris: Didier Erudition.

Chaudenson, R. 2000. *Grille d'analyse des situations linguistiques*. Paris: Didier Erudition.

Delattre, P. 1966. *Studies in French and Comparative Phonetics*. The Hague: Mouton.

Dell, F. 1973/1985. *Les règles et les sons*. Paris : Hermann.

Durand, J., Laks, B. et Lyche, C. 2005. Un corpus numérisé pour la phonologie du français. In: G. Williams (éd.), *La linguistique de corpus*, 205-217. Rennes: Presses Universitaires de Rennes.

Durand, J. et Lyche, C. 2003. Le projet 'Phonologie du Français Contemporain' (PFC) et sa méthodologie. In: E. Delais et J. Durand (éds.), *Corpus et variation en phonologie du français : méthodes et analyses*, 212-276. Toulouse: Presses Universitaires du Mirail.

Gadet, F. 2000. Des corpus pour (ne)...pas. In M. Bilger (ed.), *Corpus. Méthodologie et applications linguistiques*, 156- 167. Paris : Champion.

Gadet, F. 2007. *La variation sociale en français*. Nouv.éd. revue et augmentée. Paris: Ophrys.

Hazaël-Massieux, M.-C. 1993. Oralité et variation du français. In: D. de Robillard, et M. Beniamino, *Le français dans l'espace francophone*. Paris, Champion. Tome 1: 371-384.

Klingler, T. 2006. PFC en terrain louisianais: défis et adaptation du protocole. Colloque international *Phonologie du français contemporain : données et enjeux théoriques*, Paris, MSH, 3-5 février 2006. http://www.projet-pfc.net/?u_s=4&u_a=49&.

Knutsen, A. M. 2007a. *Variation du français à Abidjan (Côte d'Ivoire). Etude d'un continuum linguistique et social*. Thèse. Oslo, Universitetet i Oslo.

Knutsen, A. M. 2007. Le français à Abidjan (Côte d'Ivoire): vers une analyse multidimensionnelle de la variation. *Le français en Afrique, Revue du Réseau des Observatoires du Français Contemporain en Afrique*, n° 22.

Lacheret, A. et Beaujendré, F. 1999. *La prosodie du français*. Paris : Editions du CNRS.

Lacheret, A. et Lyche, C. 2006. Le rôle des facteurs prosodiques dans l'analyse du schwa et de la liaison. In: Simon, A.-C. (éd.), *Bulletin PFC* 6, 27-49. ERSS, Université de Toulouse le Mirail et CNRS.

Lafage, S. 2003-2004. *Le lexique français de Côte d'Ivoire, appropriation et créativité*. Tomes 1 et 2. *Le français en Afrique, Revue du Réseau des Observatoires du Français Contemporain en Afrique*, n° 16 & 17. Paris : Didier Erudition.

Laks, B. 2003. Les grandes enquêtes phonologiques en France. *La Tribune Internationale des Langues Vivantes*, 33, 10-17.

Lyche, C. 2006. Corpus en domaine francophone : la Louisiane. Ecole Thématique CNRS, Linguistique de corpus oraux, Nantes 19-24 June, 2006.

Manessy, G. 1995. Créoles, pidgins, variétés véhiculaires. Procès et genèse. Paris: CNRS Editions.

Milroy, L. 1980. *Language and Social Networks*. Oxford: Blackwell.

Skattum, I. 2008. Mali: In Defence of Cultural and Linguistic Pluralism. In Simpson, A. (éd.) *Language and National Identity in Africa*. Oxford, Oxford University Press: 98-121.

Wald, P., Chesney, Hily, M.-A. et Poutignat, P. 1973. Continuité et discontinuité sociolinguistiques: hypothèses pour une recherche sur le français en Afrique noire. *Bulletin du Centre des études des plurilinguismes*, octobre/décembre.