

PFC-ABIDJAN : CHOIX MÉTHODOLOGIQUES LIÉS À L'EXTENSION D'UN CORPUS

Béatrice Akissi Boutin

► **To cite this version:**

Béatrice Akissi Boutin. PFC-ABIDJAN : CHOIX MÉTHODOLOGIQUES LIÉS À L'EXTENSION D'UN CORPUS. Revue Texto, Institut Ferdinand de Saussure, 2007, Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation, <http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Boutin.pdf>. hal-02400655

HAL Id: hal-02400655

<https://hal-auf.archives-ouvertes.fr/hal-02400655>

Submitted on 12 Dec 2019

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

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

PFC-ABIDJAN : CHOIX MÉTHODOLOGIQUES LIÉS À L'EXTENSION D'UN CORPUS

Béatrice Akissi BOUTIN
ERSS, Université Toulouse 2

SOMMAIRE

Introduction

1. Le programme PFC : protocole, transcription et codages
2. L'extension de PFC aux pays africains francophones
3. Transcriptions des corpus oraux du français de Côte d'Ivoire
4. Mots et interjections spécifiques
5. Insertion d'énoncés d'une autre langue
6. Contextes particuliers des codages du schwa

Conclusion

***Résumé :** Le corpus PFC (Phonologie du français contemporain : usages, variétés et structure), conçu essentiellement pour examiner la variation géographique, sociale et stylistique de la phonologie du français, comporte de nombreux débouchés puisque les éléments pris en compte dans l'échantillonnage des locuteurs de chaque point d'enquête satisfont aux exigences actuelles de la sociolinguistique et de la linguistique de terrain, dans tout ce qu'elle peut apporter à la phonologie, à la syntaxe, à la sémantique cognitive, à l'analyse de discours, etc.*

Les enquêtes d'Abidjan et Ouagadougou (2004-2005) ont été les premières à étendre le protocole commun de PFC à l'Afrique ; elles seront bientôt suivies de celles du Sénégal et du Mali. Aucun corpus de référence de cette ampleur n'existe encore pour l'Afrique, et aucun corpus africain n'était encore entré dans une telle étude pan-francophone.

Cette extension pose toutefois les questions de l'universalité des outils (ou de l'unité dudit espace francophone), telles que celles du choix de la transcription en français standard et des normes de codage du schwa, domaine privilégié d'observation de PFC.

Nous verrons dans quelle mesure le corpus PFC – Abidjan s'insère dans les problématiques générales de PFC tout en ouvrant de nouvelles perspectives.

Introduction

La communication porte sur l'encodage de phénomènes spécifiquement oraux du français de Côte d'Ivoire dans le cadre du projet PFC « Phonologie du Français Contemporain ». Après une brève présentation du programme PFC et de son extension aux pays africains francophones, je montre la pertinence d'une transcription orthographique, y compris dans plusieurs cas particuliers comme l'alternance d'énoncés d'une autre langue ou les mots et interjections spécifiquement ivoiriens. Je présente aussi la manière dont nous avons adapté le codage des schwas dans des contextes que ne prévoyaient pas les instructions initiales de PFC.

1. Le programme PFC : protocole, transcription et codages

Le projet PFC vise à illustrer la variation phonologique observée dans le français parlé en France et dans les autres zones francophones. La comparabilité des données est garantie par l'adoption d'un protocole et d'outils d'analyses communs dans tous les points d'enquête.

Le protocole comprend deux lectures et deux entretiens ; il est appliqué à une douzaine de locuteurs liés par réseau social dense ou lâche (L. Milroy 1980), par deux enquêteurs, l'un bien connu dans le réseau, l'autre inconnu et présenté comme un ami. Les tâches enregistrées visent à saisir plusieurs registres distincts de chaque locuteur. La durée totale d'enregistrement de chaque locuteur est de 50 minutes environ.

Par la suite, sont entièrement transcrits sous Praat le texte et la liste de mots lus, ainsi que 5 minutes d'entretien guidé et 5 minutes d'entretien libre. La tire de transcription est dupliquée une première fois pour recevoir le codage des schwas et une deuxième pour le codage de la liaison.

2. L'extension de PFC aux pays africains francophones

Le projet PFC s'est voulu, dès sa conception, de portée internationale, mais son extension aux situations africaines le confronte à d'autres dynamiques linguistiques.

En Côte d'Ivoire, le français est en contact avec de nombreuses langues dont la vitalité régresse à peine. Il est cependant le véhiculaire principal et son utilisation s'étend à toutes les situations. Les principaux traits qui montrent le dynamisme de son appropriation communautaire sont sa vernacularisation et sa « nativisation » (R. Chaudenson 2000) en cours. D'autres traits manifestent l'appropriation du français, dans le domaine linguistique comme dans le domaine sociolinguistique : la créativité dans le lexique, la syntaxe et la phonologie, la récurrence de ces spécificités dans tous les registres de la langue, qui vont de la variété académique ivoirienne au nouchi, la prise de conscience de la spécificité des usages, la représentation du caractère homogène et identitaire du français de Côte d'Ivoire. Toutes ces caractéristiques confirment la pertinence de l'extension du programme PFC à la Côte d'Ivoire. On peut ajouter la facilité qu'apporte aux études PFC l'existence d'une norme endogène et de descriptions scientifiques de la variété.

Pour l'ensemble des pays d'Afrique, la réflexion a conduit les chercheurs concernés à valoriser la prise en compte de toutes les zones francophones de façon équivalente et à ne rien ajouter au protocole pour l'instant. En l'occurrence, la situation sociolinguistique de la Côte d'Ivoire permet l'application du protocole sans modification.

3. Transcriptions des corpus oraux du français de Côte d'Ivoire

Les corpus oraux du français de Côte d'Ivoire, particulièrement d'Abidjan, qui existent en Europe ont été transcrits en caractères phonétiques¹. À l'Institut de Linguistique Appliquée de l'Université Cocody – Abidjan existent au contraire des corpus oraux de plusieurs variétés du français de Côte d'Ivoire, y compris du nouchi, transcrits en orthographe standard², ou selon l'usage ivoirien.

L'entrée d'un corpus oral du français de Côte d'Ivoire dans des travaux pan-francophones demande en fait une révision de toutes les questions de transcription. Le problème peut aussi être posé dans l'autre sens (du point de vue africain) : dans quelle mesure les outils et les méthodes de traitement utilisés pour les corpus oraux de variétés euro-américaines de français peuvent-ils être utilisés pour des variétés d'Afrique ?

Mon opinion est que les situations sociolinguistiques sont effectivement différentes ainsi que les domaines de variation, mais qu'un grand nombre de critères peuvent rester les mêmes et acquièrent au contraire toute leur validité en étant confrontés à des situations nouvelles. D'autre part, on peut s'appuyer sur les solutions locales ou régionales, qui ne prétendent pas jusqu'à présent avoir une portée universelle, mais qui pourraient pourtant bien servir pourtant de modèles. L'extension des outils et méthodes à toutes les zones francophones apportera donc, en retour, une réflexion technique et théorique nouvelle sur les corpus oraux.

Dans tous les cas, le choix de l'orthographe standard pour la transcription de l'oral spontané est justifié par un certain nombre de motifs connus, d'ordre éthique, théorique ou purement technique.

Je montrerai comment nous avons choisi de résoudre trois types de problèmes de transcription auxquels nous sommes confrontés : les mots et interjections spécifiques, l'insertion d'énoncés d'une autre langue et les contextes particuliers des codages du schwa

4. Mots et interjections spécifiques

La transcription choisie par PFC est celle de l'orthographe standard du français, que l'on relève dans le matériel lexicographique habituel. Il existe aussi pour notre variété de français des travaux lexicographiques spécifiques, notamment le *Lexique français de Côte d'Ivoire*, de S. Lafage 2003-2004 ou l'*Inventaire des particularités lexicales du français en Afrique noire* (AUELF-UREF 1988), plus général. Ces travaux n'ont pas de portée normative intentionnelle mais constituent une base commune, d'autant qu'ils attestent normalement l'usage ivoirien (ou les usages africains).

Il est important de considérer que la transcription met sur un même pied des variétés de langue de statuts sociolinguistiques très différents. Alors que la norme du français de France est standardisée, fixée, aménagée, suivie pas à pas depuis longtemps, dans les autres pays

¹ M. Jabet 2005 ; K. Ploog 1999 ; Y. Simard 1998 (Voir Simard 2001) ; J.-M. Lescutier 1985 ; J.-L. Hattiger 1981.

² Par exemple : E. Niamien N'Gouan 1997 ; F. Leimdorfer *et al.* 1997 (Voir Leimdorfer *et al.* 2002) ; B.A. Boutin 2002 ; J. Kouadio N'Guessan 2005.

francophones, on en est encore à la question parfois brûlante de la prise en compte des normes endogènes. Les *Etats Généraux de l'Enseignement du Français en Afrique subsaharienne francophone* (Libreville, 17–20 mars 2003) viennent de signer la reconnaissance de leur existence. Une double question éthique se pose si l'on fait se côtoyer dans la même transcription des formes de français standard et des formes écrites de mots n'en faisant pas partie. D'une part tout travail de recherche peut être interprété par le public comme une normalisation des formes. D'autre part, la transcription d'une forme non encore standardisée est une position prise sur sa transcription, inconcevable sans tenir compte des travaux existants déjà.

5. Insertion d'énoncés d'une autre langue

Du fait de la situation plurilingue qui prévaut en Côte d'Ivoire, des insertions d'énoncés plus ou moins longs d'autres langues sont fréquentes et apparaissent effectivement dans les enregistrements. Cela n'empêche pas qu'il est tout à fait naturel de mener une conversation totalement en français (de Côte d'Ivoire), contrairement à ce qui se passe dans les pays voisins. La présence de segments de langues autres que le français pose le problème de la transcription. Les principales langues ouest africaines possèdent actuellement une orthographe officielle sur la base de l'alphabet de Bamako 1966. La seule difficulté est que deux graphèmes n'appartiennent pas à l'alphabet latin mais sont empruntés à l'API : E et . Pour éviter la mixité orthographique (Jacques Durand et Jean-Michel Tarrier, à paraître), j'ai pris le parti de transcrire en caractère SAMPA les énoncés d'une autre langue, par ailleurs rares dans le corpus. La fréquence de ces énoncés dans d'autres points d'enquête nous conduira sans doute à une autre solution.

Un phénomène fréquent est l'insertion du nouchi dans les conversations spontanées. Le nouchi est un argot ivoirien fait d'emprunts lexicaux d'origines diverses qui est usuellement transcrit en caractère latins. La seule divergence avec l'orthographe française est alors l'utilisation de l'accent circonflexe pour le [E] : « ê » et le [] : « ô ». L'écriture de ces mots ne pose donc pas de problème, dans la mesure où peuvent être suivis les critères déjà retenus pour des langues d'Europe comme le basque (Durand et Tarrier, à paraître). Une fois explicité le choix orthographique opéré par le transcripteur, la lecture est d'autant plus facilitée que « l'alignement de la transcription sur le signal sonore permet [...] d'être informé sur la réalité phonético-phonologique du mot transcrit » (Durand et Tarrier, à paraître).

6. Contextes particuliers des codages du schwa

Bien que les critères de codages¹ aient été prévus au départ pour plusieurs zones francophones, l'extension de PFC à l'Afrique, en l'occurrence à la réalité ivoirienne, montre que le codage existant doit faire aussi l'objet d'une extension. Plusieurs cas peuvent ici illustrer les difficultés et les choix méthodologiques opérés :

- **Certains mots sont trop déformés pour être codés**, tout comme *il y a / avait* ([ja] / [javE]) dans les autres variétés de français. Nous avons ainsi : *parce que* prononcé [pas@] / [pa@] ; *comme ça* prononcé [kO~a~] / [kO~O~] ; *quelque chose* prononcé [kESoz] . Dans le premier cas, le schwa est maintenu mais l'identification de la consonne précédente, demandée par le codage, est impossible. Dans les autres cas, le schwa comme son contexte ont disparu.

- **Certaines consonnes finales peuvent être effacées**, en premier lieu les liquides [r] et [l], mais souvent aussi [k] et [t]. Nous avons ainsi : *bord* prononcé [bO], *colère* prononcé [kolE]. On se trouve alors dans une situation exclue du codage, V@, du fait de l'impossibilité d'un schwa dans une telle position quelle que soit la variété de français actuelle. Dans un corpus où ce cas de réduction est si fréquent, on ne peut pas tout simplement se passer de codage : ce serait fausser les résultats et empêcher la comparaison avec d'autres variétés. Une solution a donc dû être trouvée pour coder le phénomène : le chiffre de 3^e position, 5, prévu pour indiquer la réduction d'un groupe consonantique par l'effacement d'une consonne, a été utilisé dans ce cas d'effacement de l'unique consonne. Le codage prend alors place après le dernier son prononcé selon les instructions PFC, c'est-à-dire après la voyelle.

¹ Après chaque occurrence possible d'un schwa apparaissent quatre chiffres. Le premier indique la prononciation ou l'absence de schwa, le second indique la position de la syllabe concernée à l'intérieur du mot, le troisième le contexte de la syllabe précédente et le quatrième le contexte de la syllabe suivante. Je ne discuterai ici que le 3^e chiffre.

- **Un groupe consonantique peut être réduit par l'effacement de la première consonne**, surtout s'il s'agit d'un [R] / [r]. Nous avons ainsi : *Marc* prononcé [mak], *regarde* prononcé [regad]. Ce cas n'est pas prévu par PFC. Seule est envisagée la réduction de *ministre* à [minis], ou de *titre* à [tit], codée par le chiffre 5 en 3^e position. Le codage doit prendre place après la dernière consonne prononcée, donnant par là une première information sur la ou les consonnes effacées. Ce codage a dû être aménagé différemment pour rendre compte de la situation ivoirienne : le 3^e chiffre reste 5 et le codage prend place après les deux consonnes graphiques, et éventuellement après le « e ».

Dans tous ces cas, l'extension du codage existant à la réalité ivoirienne a été réalisé en cohérence avec l'ensemble de la notation. Le même chiffre a pu être conservé pour toutes ces réductions, qui deviennent de ce fait facilement repérables, mais le codage occupe des positions différentes, donnant ainsi des indications sur le type de réduction opérée.

Conclusion

La problématique de PFC réside dans sa visée de comparabilité des résultats. Dans ce cadre, le protocole commun doit être suivi rigoureusement dans tous les points d'enquête et les outils et méthodes doivent permettre de refléter la réalité de toutes les données avec la même limpidité. Chaque extension de l'enquête PFC à une nouvelle zone francophone pose cependant de nouveaux défis au programme. Ceux de la situation ivoirienne sont en partie inconnus dans l'état actuel de l'enquête et, de ce fait, particulièrement intéressants pour la réflexion méthodologique de l'ensemble. Ils font apparaître une démarche empirique – inductive de PFC qui aurait pu rester cachée.

Nous avons déjà montré que les spécificités sociolinguistiques ivoiriennes pouvaient s'insérer dans les variables sociolinguistiques prises en compte par le protocole.

Les premiers résultats du traitement du corpus montrent maintenant que les spécificités linguistiques du français de Côte d'Ivoire peuvent être prises en compte selon les références théoriques minimales de PFC : les sons et les mots spécifiques au français de Côte d'Ivoire peuvent faire l'objet d'une transcription orthographique standard ; les variantes de prononciations et les réductions par rapport au français standard ou à d'autres variétés de français peuvent être codées avec le même système.

L'entrée des variétés africaines de français dans des études d'une telle envergure que celle de PFC représente une avancée considérable pour la recherche en Afrique, mais aussi pour la normalisation et la standardisation de ces variétés.

BIBLIOGRAPHIE

- BILGER, M. (éd.). 2000. *Corpus. Méthodologie et applications linguistiques*, Paris, Honoré Champion et Presses Universitaires de Perpignan, 380 p.
- BLANCHET, Ph. 2000. *La linguistique de terrain. Méthode et théorie. Une approche ethno-sociolinguistique*, Presses Universitaires de Rennes, 145 p.
- BOUTIN, B. A. 2002. *Description de la variation : Etudes transformationnelles des phrases du français de Côte d'Ivoire*, Thèse de doctorat, Université de Grenoble 3, 404 p., Coll. Thèses à la carte, Villeneuve sur Ascq, Presses Universitaires du Septentrion.
- BOUTIN, B. A. 2003. Des attitudes envers le français en Afrique : Enquête au sein de professions dont l'outil est le français en Côte d'Ivoire, *Education et Sociétés Plurilingues*, 14, pp. 69-84, Aosta, Italie.
- BOUTIN, B. A. 2004. PFC en contexte africain : Pré-enquête en Côte d'Ivoire, *Bulletin PFC*, 4, Colloque International *Phonologie et phonétique du français : du segmental au prosodique*, Publication de l'ERSS, Toulouse 2, www.projet-pfc.net/bulletin4/bulletin4.htm.
- DELAIS-ROUSSARIE, E. & DURAND, J. (éds.). 2003. *Corpus et variation en phonologie du français : méthodes et analyses*, Presses Universitaires du Mirail.
- DURAND, J. & TARRIER, J.-M. à paraître, *PFC, corpus et systèmes de transcription*.
- DUMONT, P. 2001. *L'interculturel dans l'espace francophone*, Paris, L'Harmattan, 218 p.
- HATTIGER, J.-L. 1981. *Morpho-syntaxe du groupe nominal dans un corpus de français populaire d'Abidjan*, Thèse de 3^{ème} cycle, Université de Strasbourg.
- Inventaire des particularités lexicales du français en Afrique noire*, 1988, UREF, 443 p., Paris, EDICEF / AUPELF.

- JABET, M. 2005. *Omission de l'article et du pronom sujet dans le français abidjanais*, Thèse de doctorat, Université de Lund, 205 p.
- KOUADIO N'GUESSAN, J. 2005 (à paraître). Le nouchi et les rapports dioula / français, *Des inventaires lexicaux du français en Afrique à la sociologie urbaine ... Hommage à Suzanne Lafage*, *Le français en Afrique, Revue du Réseau des Observatoires du Français Contemporain en Afrique*, 19, Paris, Didier- Erudition.
- LAFAGE, S. 2003 et 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*, 16 et 17, 865 p., Paris, Didier-Erudition.
- LEIMDORFER, F. 2002. L'espace public à Abidjan : individus, acteurs et situations de parole, 40 p. in F. Leimdorfer & A. Marie (éds.), *L'Afrique des citadins (Abidjan, Dakar) sociétés civiles en chantier*, Karthala, 400 p.
- LESCUTIER, J.-M. 1985. *Recherche sur le processus de réactivation. Cas singulier d'un idiolecte relevant du français populaire d'Abidjan*, Thèse de doctorat, Université de Nice.
- MILROY, L. 1980. *Language and Social Networks*, Oxford, Blackwell.
- Rapport Final*, UNESCO, Réunion d'un groupe d'experts pour l'unification des alphabets des langues nationales, Bamako, 28/02 – 05/03/1966,
<http://www.bisharat.net/Documents/index.html>
- NIAMIEN, N'G. E. 1997. *Le français parlé dans les gares routières d'Abidjan*, Mémoire de maîtrise, Université de Cocody - Abidjan.
- PLOOG, K. 1999. *Le premier actant en abidjanais : contribution à la syntaxe du non-standard*, Thèse de doctorat, Université Bordeaux 3.
- SIMARD, Y. 2001. Français de Côte d'Ivoire : l'actualisation du nom chez des locuteurs non scolarisés, R. Nicolăi (éd.), *Leçons d'Afrique. Filiations ruptures et reconstitution de langues. Un hommage à Gabriel Manessy*, Louvain / Paris, Peeters, pp. 483-496.