

ÉTUDIER LE CAMFRANGLAIS : RECUEIL DES DONNÉES ET TRANSCRIPTION¹

Carole de Féral

Université de Nice-Sophia Antipolis
UMR 6039 *Bases, Corpus et Langage*

1. Recueil des données et objet de description

1.1. Le problème du « paradoxe de l'observateur »

Dans le projet que j'ai proposé pour constituer une équipe virtuelle sur les français identitaires, j'ai souligné l'importance qu'il y avait à recueillir des interactions verbales spontanées, c'est-à-dire l'importance de régler, autant que faire se peut, le « paradoxe de l'observateur » (« observer la façon dont les gens se servent du langage quand on ne les observe pas », Labov, 1976 : 116). En effet, si pour l'élaboration d'un inventaire lexical, on peut se contenter d'énoncés recueillis après avoir demandé à des locuteurs de « parler camfranglais », on ne peut, pour essayer de cerner l'objet « camfranglais » dans sa totalité, s'en tenir à ce type d'énoncés, produits dans le seul but de faire plaisir à l'enquêteur ou de mettre en valeur sa propre compétence en camfranglais. Dans ces conditions, en effet, on doit s'attendre à ce que les productions obtenues comportent une fréquence de termes perçus comme « camfranglais » beaucoup plus importante que ce que l'on va trouver dans des discours spontanés. Car ces termes « camfranglais » ne sont généralement pas utilisés à cause d'un manque d'équivalents en français, comme c'est le cas de termes empruntés qui font référence à des *realia* locales (plats culinaires, danses, musiques, etc.) : ils sont souvent synonymes de termes appartenant au français courant. Et si le recours aux emprunts peut être systématique, dans des textes travaillés comme les chansons de rap, par exemple, l'alternance entre mot français et mot emprunté est relativement fréquente dans des discours spontanés comme dans ces extraits d'une conversation qui a été enregistrée entre des étudiants à Yaoundé en novembre 2004 (cf. aussi Féral : (page 259), ici même, pour l'alternance *do/argent*) :

¹ Ceci est une version très légèrement remaniée d'un texte qui s'adressait à l'origine aux collaborateurs de l'équipe virtuelle *Français identitaires chez les jeunes en Afrique*, consultable sur le site du réseau *Sociolinguistique et Dynamique des Langues* de l'AUF (<http://www.sdl.auf.org/Etudier-le-camfranglais-recueil>).

nga/fille

La **nga** dit que ici là zéro + il a regardé la **fille** là comme ça là + il a regardé la **fille** là comme ça là cinq minutes + ses habits se sont brûlés+ elle-même ne s'est pas brûlée + elle était nue dans le train++ Il a regardé la **fille** cinq minutes + la **nga** était nue dans le train + sans caleçon + comme elle est née ++ Ses habits se brûlaient mais son corps ne chauffait pas

du/faire :

L1 : Il est venu **faire** ça au Cameroun?

L2 : Mais chez lui non + il y a les Camer qui **du** ça non

1.2. Le problème de la non-systématicité du camfranglais

Cette alternance (ici entre les mots français *argent*, *faire*, *fille* et les mots d'origine anglaise/pidgin *do*, *du*, *nga*) est souvent interphrastique et on voit l'importance de ne pas limiter l'analyse au cadre de la phrase. En outre, dans certaines conversations, perçues comme étant en camfranglais par les Camerounais que nous avons interrogés, plusieurs phrases à la suite peuvent être très proches et même identiques au français standard, comme on peut le voir dans les échanges cités pour *do* et *argent*.

Il n'est donc pas obligatoire, pour parler camfranglais, de recourir à chaque fois que c'est possible, à un terme camfranglais plutôt qu'à un terme en français courant, comme ce serait le cas si le camfranglais était une langue étrangère. Par conséquent, s'il est nécessaire, dans un premier temps, d'établir un inventaire lexical des termes qui contribuent à donner une identité camfranglaise au discours, il faut se mettre en garde contre la tentation de regarder ce type d'inventaire comme s'il s'agissait d'un dictionnaire bilingue. Par exemple, quand je consulte le glossaire de camfranglais sur *cameroonline.org*, et que je vois « acheter : buy (exemple : j' ai *buy* un *book*, j'ai acheté un livre) », je dois savoir que je suis en présence d'un énoncé tout à fait possible mais je dois savoir aussi que « j'ai acheté un *buk* » ou « j'ai *buy* un livre » seront également tout à fait possibles sans perdre pour autant leur identité camfranglaise. Et pour aller plus loin, je dirai que « j' ai acheté un livre », conforme au français standard, peut faire partie d'un ensemble, plus grand que la phrase, qui sera perçu comme camfranglais à partir du moment où ailleurs dans l'intervention du locuteur (ou même plus loin dans la séquence des échanges analysés ?) se trouveront un ou plusieurs termes qui suffiront à donner une identité camfranglaise au discours. La question que l'on doit alors se poser c'est quelle unité d'analyse choisir.

Faire appel à un terme camfranglais n'est donc pas une contrainte linguistique mais un choix discursif, dont il nous appartient de découvrir les motivations². Ce n'est que le jour où faire appel à ces termes deviendra une contrainte, le jour où des termes comme *nga*, *do*, *du* remplaceront systématiquement *fille*, *argent*, *faire*, que nous pourrions nous demander si nous ne sommes pas en train d'assister à la naissance d'une nouvelle langue.

² Une tentative d'interprétation est proposée dans Féral, à paraître.

Cette non-systématicité est une des raisons essentielles pour lesquelles on ne peut limiter la définition du camfranglais à un simple inventaire lexical. Un autre problème est celui de la non-spécificité de certains termes (la majorité ?) perçus comme camfranglais. Des mots comme *skul*, *go* (« école », « aller », empruntés à l'anglais/au pidgin *school*, *go*) ou *bachot*, *galère*, ne sont pas spécifiquement « camfranglais » et ne peuvent être identifiés comme tels que s'ils sont actualisés dans un discours en français dans des circonstances bien déterminées (conversation entre jeunes Camerounais...). Quand un petit Français qui habite en France dit « on est *go* » (on est partis), il ne parle pas camfranglais, il parle juste comme un jeune Français. Mais si c'est un jeune Camerounais qui le dit à un autre jeune au Cameroun, il parle camfranglais. De même, *galère* et *bachot*, s'ils sont perçus comme camfranglais au Cameroun, sont considérés par des locuteurs français comme tout à fait hexagonaux (*galère*, « situation difficile » est cité dans Goudaillier 2001 et Seguin et Teillard 1996, mais il est également répertorié par *Le Petit Larousse* 2005 comme familier ; *bachot*, quant à lui, est qualifié par ce même dictionnaire de familier et de vieilli).

2. Le problème de la transcription

Transcrire un corpus, c'est faire des choix, savoir quel est le but de l'étude, tenir compte des problèmes que la méthode du recueil des données ou que les données elles-mêmes peuvent engendrer. Comme l'objet de l'étude n'est ni phonétique ni phonologique et que la langue étudiée est du français, une transcription du type GARS me semble être adéquate. Cependant, elle ne résout pas le problème posé par la transcription des emprunts, notamment ceux au pidgin/à l'anglais. En effet, il est impossible dans de nombreux cas, de décider si les emprunts d'origine anglaise sont des emprunts faits au pidgin ou directement à l'anglais puisque la grande majorité du lexique pidgin est d'origine anglaise.

Quand on regarde des extraits de corpus qui ont été transcrits par quatre auteurs différents³ que remarque-t-on ?

On remarque, tout d'abord, que la graphie utilisée n'est pas homogène non seulement d'un transcripateur à l'autre mais aussi chez un même transcripateur : les mêmes critères ne sont pas appliqués à tous les termes. On peut ainsi relever dans chaque corpus deux principaux types de transcription en concurrence : soit les emprunts au pidgin/à l'anglais sont transcrits selon l'orthographe anglaise (graphie étymologisante), soit on écrit ce qu'on entend (graphie phonologisante), en s'inspirant parfois de l'orthographe du français (graphie francisante). S'ajoutent à cela des transcriptions dont il est difficile de savoir à quels principes elles obéissent et que j'appellerai « indélices ».

Voici quelques exemples qui concernent des termes dont on ne peut décider s'ils ont été empruntés au pidgin ou directement à l'anglais.

³ Le corpus A est cité dans Biloa 1999. Les corpus B, C et D ne sont pas publiés et appartiennent respectivement à A.S.H.N. et E.G. La réflexion proposée ici a fait l'objet d'un exposé donné au séminaire de DEA d'A. Queffelec le 12 novembre 2004 à l'Université d'Aix-Marseille 1.

2.1. Orthographe anglaise

do « faire » est transcrit dans les quatre corpus avec l'orthographe anglaise mais il a pour homographe, mais non homophone, *do* (< angl. *dough* (?)) ou, plus vraisemblablement, *do*(*llars*) : « argent », « fric » :

- (1) ma *remé* m'a *gui* le *work* à **do** (A)
- (2) c'est chacun qui **do** non? (B)
- (3) il a **do** semblant de *laugh* (C)
- (4) elle pouvait tout **do** tout avoir (D)
- (5) il ne veut pas *komot* les **do** de sa poche pour *buy* (B)
- (6) tu as eu le *ton*, le *ton* de ne pas dépenser tes **do** de *taco* (A)
- (7) je n'avais plus les **dos** de *takesh* pour *back* à la *piole* (C)
- (8) j'ai *give* all mes **ndo** (C)
- (9) un *djo* qui a les *dos* qui a les **dos** mon ami (D)

Dans les exemples (1) à (4) il s'agit de *du*⁴, transcrit selon l'orthographe anglaise, mais dans les exemples (6) et (8), il s'agit de *do*, transcrit phonétiquement (*do*~*ndo*) mais auquel on ajoute pour (7) et (9) la marque « s » du pluriel en français écrit (non prononcée).

2.2. Graphie anglaise/graphie francisante

Pour, [gi], (*donner*), on trouve selon les corpus soit l'orthographe anglaise soit une graphie francisante (*gui*)

- (8) j'ai **give** all mes *ndo* (C)
- (10) il *gui* cinquante *kollos* (D1)
- (11) ma *remé* m'a **gui** le *work* à *do* (A1)

2.3. Orthographe anglaise/graphie phonologisante

Dans l'extrait de corpus C est utilisée l'orthographe anglaise *laugh* (*rire*) mais dans B et D on préfère une graphie phonétique *lap* (selon les locuteurs on prononcera [lap] ou [laf]).

- (12) Pourquoi tu **lap**? (D)
- (13) on a éclaté de **lap** (B)
- (14) : il a *do* semblant de **laugh** (C)

Pour *no*, « savoir », alternent aussi, selon les corpus, orthographe anglaise (*know*) et graphie phonologisante (*no*). Mais on remarque dans A non seulement les deux graphies mais une troisième, (*now*), qui fait à tort penser à angl. *now*, « maintenant » :

- (15) tu **know** que ma *remé* m'a *gui* le *work* à *do* (p.173)
- (16) tu **now** que comme il y avait les *mbindi* il a voulu sortir (p. 173)
- (17) toi-même tu **now** (p. 174)

⁴ J'utilise, pour ma part, une graphie fondée sur une transcription phonétique large (cf. 2.6.).

2.4. Graphie « indécise »

En revanche *komot* (angl *come out*) ne bénéficie jamais de l'orthographe anglaise, sans doute parce que ce verbe, très courant en pidgin, a un signifié et une distribution qui se distinguent nettement de l'anglais *come out*. L'exemple (18) nous montre une transcription phonétique tandis que (19) et (20) illustrent ce que j'ai appelé plus haut une transcription indécise :

- (18) il ne veut pas **komot** [« sortir »] les *do* de sa poche pour *buy* (B)
- (19) en **comotant** (C)
- (20) il faut qu'on **commot** il faut qu'on **commot** ensemble (D)

2.5. Morpho-syntaxe

Si on ne prend pas un grand soin dans la transcription des données, celle-ci peut engendrer des erreurs dans l'analyse. Par exemple, choisir d'écrire *came* (ex. 21 à 24) est ennuyeux car si on se réfère au principe de la graphie francisante on va être tenté de prononcer [kam] et si on se réfère au principe de la graphie étymologisante on va être tenté de prononcer [kem] et de voir un temps passé de l'anglais pour l'énoncé 21 :

- (21) je **came** du school (« je sors de l'école » et non « je suis sorti ») (A)
- (22) elle va **came** [...] donc au lieu de **came** il relance le huit là (B)
- (23) *wait* je **came** (C)
- (24) il est **came** (D)

Pour (25), **loss** ne pose pas de problème particulier, mis à part le côté arbitraire de sa transcription mais la graphie étymologisante **lost** (prétérit et participe passé de *lose*, « perdre ») dans les ex. 26 et 27 risque de conduire le transcripateur linguiste à analyser *los* comme le participe passé anglais, induit par la présence de l'auxiliaire *avoir* alors qu'il est bien plus vraisemblable (cf. 2.6.2.) de considérer *los* comme un emprunt au pidgin *los* (« perdre »). La transcription phonétique qui figure dans le corpus C indique bien d'ailleurs que le *t* de l'énoncé 27 n'est pas prononcé).

- (25) mais est-ce qu'on **loss** le bac tchadien (B)
- (26) je n'ai rien **lost** (A)
- (27) donc j'ai **lost** (C)

Le corpus D présente le même problème avec *changed* :

- (28) la vie a **changed** la vie a **changed** de sens (en note : prononciation sans « ed ») (D)

2.6. Quelle transcription choisir ?

Choisir une transcription étymologisante exige que l'on connaisse l'origine de chaque terme, ce qui n'est pas toujours une condition remplie. Pour les emprunts aux langues camerounaises, en outre, il faut savoir que si certaines possèdent un et même plusieurs systèmes graphiques⁵, ils sont d'un usage très restreint. Et on doit

⁵ C'est le cas notamment du basaa et de l'ewondo.

s'attendre à ce que ces systèmes ne soient en général pas connus des transcripteurs et encore moins des lecteurs. En ce qui concerne les emprunts au pidgin/ à l'anglais, choisir systématiquement l'orthographe anglaise pour des mots d'origine anglaise dont certains ont vraisemblablement transité par le pidgin incite à surévaluer le rôle de l'anglais et à occulter celui du pidgin comme langue prêteuse. Adopter une transcription phonologisante permet plus d'objectivité. Elle permet de laisser le doute lorsqu'on ne peut décider, selon des critères strictement linguistiques, s'il s'agit d'emprunts au pidgin ou à l'anglais et de ne pas occulter le rôle du pidgin pour des termes qui, de toute évidence, sont des emprunts au pidgin, comme *hia* et *mimba* :

(28) il me dit qu'il **hea** bad, Monsieur, il commence à trembler (B2)

hia (pidg. *hia* < angl. *hear*, « entendre ») signifie ici, comme en pidgin, « se sentir ».

(29) quelque part je **mimba**-ais qu'il allait recame (C3)

mimba (pidg. *mimba* < angl. *remember*) ne signifie pas ici « se souvenir » mais « penser » conformément à son usage en pidgin. L'orthographe anglaise aurait ici non seulement occulté le rôle très vraisemblable du pidgin comme langue prêteuse mais n'aurait pu rendre compte de la disparition de la première syllabe en pidgin.

Choisir une transcription phonétique large, une transcription phonologisante, pour les mots qui ne sont pas reconnus comme étant du français me paraît donc être la solution la moins dangereuse. Evidemment, elle ne règle pas tout. Si dans un corpus on trouve pour « rire », [lap] en variation avec [laf], il faudra choisir. Mais on pourra toujours, en note, indiquer un écart par rapport à la norme de transcription.

2.6.1. Accord

Comment, cependant, régler le problème posé par l'accord ? Par exemple l'accord en genre et en nombre, présent à l'écrit en français mais souvent absent à l'oral ?

On a vu que pour *do* (ex. 5 à 9) qui semble être systématiquement précédé de l'article pluriel, le « s » était parfois noté (ex. 7 et 9) comme dans :

(30) ses **dos** sont *bolé* (D).

Dans ce dernier exemple, le fait que *bolé*, avec son accent aigu, soit de façon non ambiguë une transcription francisante nous fait nous demander pourquoi l'accord en nombre n'est pas noté pour ce terme (*bolés*) alors qu'il l'est pour *do*. D'un autre côté, la graphie francisante de *bolé* pourrait nous inciter à voir dans *dos* non pas *do* auquel est ajouté le pluralisateur « s » mais le substantif français « dos » !

À partir du moment où les termes non reconnus comme français ont une graphie phonologisante et se distinguent bien du reste de l'énoncé - à l'aide de caractères gras, par exemple-, qui est écrit conformément aux règles de l'orthographe française, je pense qu'il est souhaitable de ne pas marquer le « s » du pluriel du français écrit puisqu'à l'oral cette marque du pluriel est généralement absente.

2.6.2. désinences temporelles

Le camfranglais respecte généralement la structure du français. Comme je l'ai dit ailleurs (Féral 2005), on note cependant une omission de la marque désinentielle pour les mots empruntés lorsqu'elle est ressentie comme redondante. Ainsi la marque du participe passé est généralement absente dans un temps au passé composé puisque l'auxiliaire « avoir » donne déjà l'indication sur le temps mais la désinence de l'imparfait est bien présente. Doit-on alors la noter en phonétique ou conformément à l'orthographe française ?

Soit l'exemple :

(29) quelque part je **mimba-ais** qu'il allait *recame* (C)

Ecrire la désinence de l'imparfait en phonétique *mimbaje* ou *mimba-e* nuirait, à mon avis, à la lisibilité du corpus et ferait perdre de vue que la structure du camfranglais est bien celle du français. Garder le *-ais* du français me paraît donc justifié.

De même dans :

(31) il est *comot* premier en **beginnant** par derrière (B)

Noter conformément à l'orthographe du français la marque du participe présent me paraît être la bonne solution. En revanche, les deux « n » de *beginnant* me semblent être le fait d'une graphie francisante. Je suggère de séparer la marque morpho-syntaxique, écrite conformément à l'orthographe française, du lexème, qui serait écrit phonétiquement : *bigin-ant*.

2.7. Propositions

Je propose donc une transcription phonétique très large, une transcription phonologisante, qui utilise les signes dont on dispose normalement sur notre clavier. Il n'est pas utile à mon avis de faire une distinction entre « o » ouvert et « o » fermé, « e » ouvert et « e » fermé. Les chuintantes [ʃ] et [ʒ] peuvent être transcrites *sh* et *j* (ex. : *ledj*, « village »). Le *y* peut transcrire, si nécessaire la semi-consonne [j]. Il n'est pas essentiel ici, à mon avis, de noter les tons dans le cas d'emprunts aux langues à tons.

Il faut évidemment ne pas omettre, les chevauchements entre les locuteurs, les hésitations, les mots inachevés.... Enfin, tout ce qui est caractéristique d'interactions orales authentiques et spontanées.

Pause : + note une pause courte, ++, une pause plus longue etc.

Chevauchement : le soulignement marque le chevauchement entre deux locuteurs :

L1 : quand le Bassa te lance la maladie c'est pour partir

L2 : tu meurs (corpus Féral)

Segment incompréhensible : xxx (un « x » par syllabe, si possible).

Allongement syllabique :

: pour un allongement relativement court : lorsqu'il et plus long, ::: lorsqu'il est encore plus long etc. :

il est parti depuis :::

Bibliographie

- BILOA, Edmond (1999) : « Structure phrastique du camfranglais » in G. Echu et A.W. Grundstrom (éds), p.147-174. : *Bilinguisme officiel et communication linguistique au Cameroun*, New York, Peter Lang. Publié également dans Biloa (2003) *La langue française au Cameroun*, chapitre 13, p. 247-278.
- FERAL, Carole de (à paraître) : « Décrire un « parler jeune » : le cas du camfranglais (Cameroun) », communication présentée au colloque : *Des inventaires lexicaux du français en Afrique à la sociolinguistique urbaine*, Paris, MSH, 28 février 2005, cf. ici même.
- FERAL, Carole de (à paraître) : « Ce que parler comafanglais n'est pas : de quelques problèmes posés par la description d'un parler jeune », in *Langues, cultures et interaction*, (Actes du IV^e colloque du réseau Français de Sociolinguistique *Faits de langues et de cultures dans la dynamique interactionnelle*, Université Paris V, Paris, 7-8 octobre 2005, Paris, l'Harmattan.
- GOUDAILLIER, Jean-Pierre (2001) : *Comment tu tchatches! Dictionnaire du français contemporain des cités*, Paris, Maisonneuve et Larose.
- LABOV, William (1976) : *Sociolinguistique*, Paris, Editions de Minuit.
- SEGUIN, Boris et TEILLARD, Frédéric (1996) : *Les Céfrans parlent aux Français*, Paris, Calmann-Lévy.