

Le lexique de Joseph Zobel, auteur antillais: extraction semi-automatique des particularismes lexicaux

1. Introduction

La lexicographie différentielle francophone a beaucoup progressé ces dernières années (DSR, DHFQ, DRF, DRFA), entre autres grâce à des banques textuelles telles que Frantext, Suistext ou Québétext; or, le français régional des Antilles est resté jusqu'à maintenant un parent pauvre de cette discipline (malgré l'existence de Telchid 1997, très perfectible). Pourtant, l'importance de cette variété pour l'histoire du français colonial, du français nord-américain, du français parlé, et en définitive du français tout court, est primordiale. C'est pour cette raison que nous avons entrepris l'élaboration d'une banque de données textuelles de littérature antillaise francophone, dont nous présentons ici le premier jet.

2. Le corpus

Une subvention de la Mission pour la Science et la Technologie de l'ambassade de France au Canada nous a permis de numériser la totalité de l'œuvre en prose de l'écrivain martiniquais Joseph Zobel (Rivière-Salée, 1915 – Alès, 2006). Il s'agit, en ordre chronologique, des ouvrages suivants:

- Les jours immobiles*. Fort-de-France: Imprimerie officielle, 1946. 217 pages.
- Diab'-la*. Paris: Nouvelles éditions latines, 1947. 175 pages.
- La rue Cases-Nègres*. Paris: Froissart, 1950. 314 pages.
- La Fête à Paris*. Paris: La Table ronde, 1953. 257 pages.
- Le Soleil partagé*. Paris: Présence africaine, 1964. 208 pages.
- Laghia de la mort*. Paris: Présence africaine, 1978. 111 pages.
- Les Mains pleines d'oiseaux*. Paris: Nouvelles éditions latines, 1978. 157 pages.
- Quand la neige aura fondu*. Paris: Éditions caribéennes, 1979. 145 pages.
- Et si la mer n'était pas bleue: nouvelles*. Paris: Éditions caribéennes, 1982. 88 pages.
- Mas Badara*. Paris: Nouvelles éditions latines, 1983. 150 pages.
- Gertal: et autres nouvelles*. Matoury (Guyane): Ibis rouge éditions, 2002. 222 pages.

Nous allons très brièvement présenter la méthodologie mise en œuvre lors de la numérisation des ouvrages, puis nous jetterons un coup d'œil aux résultats préliminaires. Le corpus total, qui réunit un peu plus de 2000 pages, a aussi été dépouillé manuellement; il s'agissait de voir si les algorithmes et autres stratégies d'interrogation semi-automatique permettraient d'identifier les diatopismes que permet de récolter un dépouillement manuel, afin de pouvoir les appliquer éventuellement à l'extraction de diatopismes sur un corpus beaucoup plus vaste, qu'un lecteur humain ne saurait dominer.

3. La méthodologie

Les textes numérisés (à l'aide d'Omnipage SE de Scansoft) ont été annotés en XML, ce qui comporte de nombreux avantages: compatibilité avec la Text Encoding Initiative; indépendance face aux outils utilisés; passage facilité à de nombreux concordanciers; possibilité d'annotation en couches. Il a fallu construire une DTD, ce qui fut fait à l'aide du logiciel <oXygen>. Les balises structurantes retenues ont été les suivantes: document; roman, recueil_nouvelles; nouvelle; préface, introduction, chapitre, postface; page; titre, sous-titre; saut_de_section. D'autres balises, métalinguistiques et métadiscursives, sont venues s'ajouter aux premières: langue (pour les passages en créole, en anglais, etc.); typo (pour l'italique et le gras); césure (pour que les mots coupés à la ligne soient identifiables lors d'une recherche, ou de l'établissement d'un concordancier); sic (pour les coquilles corrigées d'autorité); note (pour les notes de bas de page); appel (pour les appels de notes de bas de page); discours (discours direct, discours rapporté, passages narratifs, etc.).

4. Les résultats des stratégies d'extraction semi-automatique

4.1 L'exploitation de l'annotation

L'exploitation de l'annotation permet déjà de faire une belle récolte: les mots ayant fait l'objet d'une mise en relief typographique à l'aide de l'italique, par exemple, sont en général très intéressants pour le linguiste, et le fait de pouvoir en obtenir automatiquement la liste permet de récupérer rapidement des matériaux diatopiquement marqués (*gadez* «regardez», *chabine*, *mussieu*, *mazouk*, *madanm*, *mamzé*, etc.); il en va de même pour les guillemets (*coco-de-l'œil* «conjonctive», *fifine* «bruine», *parer* «préparer», etc.). Les notes de bas de page, chez Zobel, sont presque toujours consacrées à des commentaires métalinguistiques, d'où la possibilité de dresser automatiquement un mini-glossaire (très rudimentaire, cf. Thibault 2006, mais tout de même précieux) de régionalismes antillais: *zombi* «esprit malin», *volant* «poisson volant», *bois mapou* «bois excessivement tendre», etc. Des stratégies d'interrogation de nature plus discursive permettent également de recueillir des données intéressantes, comme une simple recherche portant sur la séquence

textuelle «comme on dit»: «et comme on dit chez nous «Balai neuf balaie bien»»; «je leur conseille de ne pas perdre leur temps avec moi qui suis un poteau pourri, comme on dit chez nous»; «et il est obligé de faire un grand geste des bras et des jambes – un «siac», comme on dit de ceux que les vapeurs de rhum font tituber»; «Alex parlait un tiers français, un tiers créole, un tiers anglais de cuisine, anglais banane, comme on dit là-bas».

4.2 Listes de spécificités

L'un des avantages indéniables des banques de données textuelles est qu'elles permettent de faire ressortir objectivement des phénomènes de fréquence¹ qui, bien souvent, nous mettent sur la piste de lexèmes diasystémiquement marqués. En confrontant le corpus à un corpus de référence (en l'occurrence, le texte intégral de plusieurs années du journal *Le Monde*) et en demandant au programme de comparer systématiquement les fréquences relatives des lexèmes dans chaque corpus, on peut obtenir une liste des lexèmes du corpus Zobel par ordre décroissant de fréquence marquée (c'est-à-dire, des lexèmes dont la fréquence chez Zobel est anormalement élevée par rapport à celle qu'elle affiche dans *Le Monde*). La comparaison est effectuée à l'aide du calcul des spécificités tel que proposé par Lafon 1980. Il s'avère que la grande majorité des éléments lexicaux ainsi mis en évidence appartient bel et bien au stock de mots identifiés par le dépouillement manuel comme diatopiquement marqués, ce qui signifie que ce procédé d'extraction automatique est valable et pourrait être étendu à des corpus de taille beaucoup plus considérable. Voici la liste des 20 premiers mots relevés; seul trois ou quatre d'entre eux n'y figurent que pour des raisons strictement thématiques, tous les autres présentant un intérêt diatopique certain pour la lexicographie différentielle:

1. *case* n. f. «maison (quelle qu'elle soit)». Le mot en français des îles s'applique à toute habitation et ne se limite pas à un logis rudimentaire fait de matériaux légers.

2. *canot* n. m. «barque de pêcheur». Le mot ne souffre guère de la concurrence de l'anglicisme *canoë*, qui n'apparaît pas une seule fois chez Zobel (alors que les attestations de *canot* abondent).

3. *nègre* n. m. «homme». Le mot n'a pas en français des Antilles de connotation péjorative, mais il équivaut souvent, du point de vue du registre, à «mec», «type».

4. *coui* n. m. «moitié dealebasse évidée et séchée, utilisée comme récipient». Diatopisme emblématique du français des Antilles. – Telchid 1997. – Sources créoles: Jourdain 1956: 298 (*coui*); Ludwig *et al.* 2002: 190 s.v. *kwi*.

5. *bourg* n. m. «petite agglomération urbaine». Le français des Antilles n'utilise guère *village*, mot extrêmement rare chez Zobel (à moins qu'il ne s'agisse de contextes européens ou africains). Cette préférence pour *bourg* constitue un héritage du français de l'Ouest (où le mot s'oppose tout de même à *village*, lequel a pris la place de *hameau*; cf. l'article de J.-P. Chauveau dans le DRF). – DRF 2001 s.v. *bourg* (où il conviendrait de rappeler l'emploi antillais). – Sources créoles: Jourdain 1956: 231 (*bouc*); Ludwig *et al.* 2002: 81 s.v. *bou/bouk*.

¹ Sur les régionalismes de fréquence et leur repérage, cf. Thibault 2007.

6. *poitrine* n. f.: ce mot ne présente aucun intérêt diasystémique et n'apparaît dans la liste que pour des raisons thématiques. C'est un cas de «bruit» dans les résultats.

7. *rhum* n. m.: le mot n'est pas seulement un régionalisme de fréquence; il est aussi intéressant linguistiquement car il entre dans la formation de nombreuses locutions substantives précisant sa nature (*rhum fort*, *rhum blanc*, *rhum clair*, *rhum sec*, *rhum pur*, *rhum de mélasse*, *rhum non réduit*) et qu'il donne lieu à des dérivés (*rhumerie* n. f.); dans la mesure où il sert aussi dans les romans de Zobel d'antiseptique et d'alcool à friction, sa définition mériterait d'être précisée en fonction du contexte.

8. *robe* n. f.: tout comme *poitrine*, le mot ne présente aucun intérêt diasystémique.

9. *manioc* n. m.: régionalisme de fréquence, entrant lui aussi dans la formation de locutions substantives et verbales caractéristiques de la prose de Zobel (*pièce de manioc*, *pied de manioc*, *farine de manioc*, *amidon de manioc*, *moulin à manioc*; *gratter le manioc*); on relève également son emploi fréquent au pluriel (*maniocs* n. m. pl. «plants de manioc»). Ce mot mériterait, pour rendre justice à tous ces emplois, un bon article dans un dictionnaire différentiel.

10. *punch* n. m.: autre antillanisme emblématique (cf. Telchid: «rhum pris avec du sucre brun et du citron ou avec un sirop de fruit et du citron»), qui entre dans la formation de nombreuses locutions dont plusieurs mériteraient d'être définies (*petit punch*, *ti punch*; *punch au lait*; *punch au citron*; *punch glacé*; *punch créole*; *punch martiniquais*; *verre à punch*).

12. *manger* v.tr.; n. m.: la fréquence relativement élevée de cette forme chez Zobel par rapport au corpus de référence s'explique probablement par son emploi substantivé, qu'on peut considérer ici comme un créolisme, même si on le rencontre également en français de référence (mais il est «vx.» au sens de «fait, acte de manger» et «pop.» au sens de «nourriture, repas», cf. Nouveau Petit Robert 2008; l'article du TLF est décevant du point de vue du marquage). Telchid 1997 y consacre un article. Le mot apparaît également dans des composés (*manger-cochon*, aussi dans Telchid; *manger-coulies*).

13. *canne* n. f.: chez Zobel, *canne à sucre* devient la plupart du temps par ellipse *canne* tout court; le mot est très fréquent pour des raisons thématiques, et entre dans la formation de nombreuses locutions, parfois au pluriel: *champs de canne(s)*, *pièce de canne(s)*, *plantation de canne(s)*; *pieds de canne*, *tiges de canne*, *feuilles de canne*, *paille de canne*, *touffes de cannes*, *coupeur de canne*; *jus de canne*, *sucre de canne*, *sirop de sucre de canne*; *travailler dans les cannes*. L'emploi du mot au pluriel n'est pas répertorié dans l'article du TLF.

14. *maman* n. f.: outre les éventuelles raisons thématiques, le mot apparaît chez Zobel avec la valeur, empruntée au créole, d'élément de formation à valeur augmentative: cf. *maman-violon* n. m. «violoncelle» (*La Rue Cases-Nègres*, p. 59). On le trouve aussi très souvent antéposé à un nom de femme, avec la valeur de «maman» mais aussi bien souvent de «Madame»; dans cette fonction, toutefois, il est fortement concurrencé par son équivalent créolisant *Man* (cf. Telchid 1997). On le relève aussi chez Zobel dans le composé *grand-maman*, vieilli en français de référence (cf. DSR s.v., et Thibault 1999).

15. *soleil* n. m.: tout comme *poitrine* et *robe*, ce mot ne présente aucun intérêt diatopique et sa fréquence exceptionnelle chez Zobel est due à des raisons de contenu narratif.

16. *morne* n. m. «colline»: mot caractéristique des français d’outre-mer; antillanisme à la base, le mot s’est répandu en français canadien (cf. l’ILQ) et dans l’Océan Indien (cf. Beniamino 1996: 206; Chaudenson 1974: 619).

17. *béké* n. m. «blanc créole»: autre antillanisme emblématique, d’origine inconnue; on trouve aussi chez Zobel le féminin *békée*, ainsi que des emplois adjectivaux ou en apposition (*femme béké*). – Telchid 1997. – Sources créoles: Jourdain 1956: 171 et n. 1;² Tourneux / Barbotin 1990: 47; Ludwig *et al.* 2002: 71.

18. *tam-tam* n. m. «instrument à percussion»: le mot n’est pas d’origine antillaise, mais le référent auquel il renvoie est fréquemment évoqué dans les romans de Zobel; sur ce mot d’origine indienne, cf. Chaudenson 1974: 1073-1074 et TLF.

19. *madras* n. m.: plutôt que l’étoffe, le mot désigne le plus souvent chez Zobel, par métonymie, la coiffure «formée d’un fichu de cette étoffe noué sur la tête» (TLF). – Telchid 1997. – Sources créoles: Jourdain 1956: 106 et n. 1; Ludwig *et al.* 2002: 216.

20. *canari* n. m. «marmite, cocotte»: ce mot originaire d’une langue amérindienne de l’aire caraïbe n’a rien à voir avec son homonyme désignant un volatile. Sa fréquence est très élevée chez Zobel, dans des descriptions de la vie domestique. – TLF s.v. *canari*²; Telchid. – Sources créoles: Jourdain 1956: 83 (*kānari*), 298 (*canari*); Ludwig *et al.* 2002: 164 s.v. *kannari*.

4.3 Mots à fréquence zéro dans le corpus de référence

Si la stratégie exposée au paragraphe précédent permet de faire ressortir des diatopismes sémantiques (*case*³, *canari*⁴), grammaticaux (*manger* substantivé) ou simplement de fréquence (*bourg*, *punch*, *rhum*), les diatopismes lexématiques quant à eux sont encore plus faciles à repérer: la confrontation avec un corpus de référence les fait tout de suite apparaître. Le procédé est simple et très rentable. Voici une douzaine de mots ainsi relevés:

1. *bacoua*, *bakoua* n. m. «chapeau fait de feuilles séchées de l’arbuste du même nom» (Telchid 1997). – Ø TLF (sauf sous la graphie *vacoua*, et en référence à l’arbre; att. de 1832 de G. Sand s.v. *ancrer*). – Sources créoles: Jourdain 1956: 268 (*bacoua* ou *vacoua* < Océanie *pandanus utilis* (BORY)); Chaudenson 1974: 54, 61, 184-185, 307 et surtout 1074-1076 (*vacoua*); Tourneux / Barbotin 1990: 39 s.v. *bakwa*.

2. *calebassier* n. m. «arbre d’Amérique appartenant à la famille des Bignoniacées» (TLF). Le mot figure à la nomenclature de la plupart des dictionnaires français; sa présence dans cette liste est un artefact dû au fait que le mot était absent du corpus de référence (*Le Monde*). Les résultats doivent donc évidemment être confrontés à un corpus d’exclusion de nature lexicographique.

² «On explique ainsi à la Martinique l’origine de ce mot: *eh bé qué?* (eh ben quoi?) question revenant fréquemment sur les lèvres des premiers colons». Il s’agirait alors d’un délocutif. Pour d’autres cas de noms ethniques délocutifs, cf. Buchi 1995.

³ Cas de polysémie.

⁴ Cas d’homonymie.

3. *chabin*, *-ine* n. m., f. «personne de race noire à la peau claire parfois parsemée de taches de rousseur et aux cheveux blonds ou roux crépus ou bouclés» (Telchid 1997). – Sources créoles: Jourdain 1956: 174 et n. 2; Tourneux / Barbotin 1990: 73 s.v. *chaben*, *chabin*; Ludwig *et al.* 2002: 89 s.v. *chaben*, *chabin*. – Selon Jourdain (loc. cit.), «Cette appellation vient du mouton <chabin> qu'on a cru à tort issu du croisement du bouc et de la brebis». Cf. encore FEW 2, II, 309b, CAPRINUS I 1: centr. *chabin* «[...] sobriquet d'une personne aux cheveux frisés».

4. *cici* n. m. «petit oiseau». – Ø Telchid 1997. – Sources créoles: *cici*, *sissi* n. «petit oiseau (*Tiaris bicolor omissa*, JARD)» Jourdain 1956: 31, 300 (mentionne deux espèces: *sissi zhèbe* et *sissi mǎnioc*; postule une origine caraïbe); *sisi-zèb* n. «oiseau (*Tiaris bicolor*, Fringilidae)» Tourneux / Barbotin 1990: 374; *sisi*, *sisi-zèb* n. «très petit oiseau» Ludwig *et al.* 2002: 293.

5. *corossol* n. m. «gros fruit tropical, variété d'anone, dont la peau est hérissée de pointes» (Nouveau Petit Robert 2008⁵). – Une seule att. dans TLF (qui cite Lar 19^c), s.v. *cachiment* dont il serait un synonyme. – Telchid 1997. – Sources créoles: Jourdain 1956: 275; Chaudenson 1974: 609 (1^{re} att.: 1654); Tourneux / Barbotin 1990: 207-208; Ludwig *et al.* 2002: 182.

6. *débèloire* n. f. «cafetière» (*Les jours immobiles*, p. 136). – Cf. FEW 23, 36a: Ancecy *dubelloir* m. f. «cafetière en fer-blanc ou en terre», Lant. *dèbèlwar* f. «nom populaire de la cafetière; filtre de la cafetière»; cf. encore TLF s.v. *débèloire*, qui ne le relève que chez Giono et cite Rheims 1969 pour sa proposition étymologique non étayée. – On relève aussi *dubelloir* n. m., de même sens, dans Martin / Pellet 1987: 80.

7. *gèreur* n. m. «contremaître d'une plantation de canne à sucre». Dérivé original sur *gérer*, suffixe d'agent *-eur*. – Ø TLF; pourtant, le mot est bien attesté aujourd'hui dans Frantext (27 att., de Zobel, Leiris, Chamoiseau, Chandernagor; toutes en contexte antillais). – Telchid 1997; Nouveau Petit Robert 2008. – Sources créoles: Jourdain 1956: 197 (*gèrèu* «régisseur»);⁶ Tourneux / Barbotin 1990: 167 s.v. *jèrè*; Ludwig *et al.* 2002: 153 s.v. *jèrè*.

8. *iche* n. m. «fils». – Ø TLF, Telchid 1997. – Première att. dans Frantext: 1921, René Maran (né à Fort-de-France), *Batouala, véritable roman nègre*, p. 48. – Sources créoles: Jourdain 1956: 44 (*fai yiche* «donner naissance, mettre au monde»), 115 (Guadeloupe *yiches zenfants* «la postérité, la descendance, la lignée, les enfants»)⁷, 116 (*yiche* «l'enfant»); Ø Tourneux / Barbotin 1990; Ø Ludwig *et al.* 2002. Selon Hazaël-Massieux 1989: 291, l'emploi du mot en interpellatif avec l'adjectif possessif dans *mon iche* chez Zobel serait une «forme tout à fait artificielle» (il faut supposer, mais l'auteure ne le dit pas explicitement, que les locuteurs antillais diront soit *iche-à-moin* en créole, soit *mon enfant* en français, mais jamais *mon iche*).

⁵ Cette source donne «fin XIV^e» comme première attestation (non référencée), ce qui semble très précoce pour un mot «du créole des Antilles» (ibid.).

⁶ «Le contremaître agricole s'appelle le «commandeur», au-dessus de lui vient «l'économiste» et au-dessus le «gèreur» c'est-à-dire le régisseur ou le propriétaire si celui-ci fait valoir lui-même son bien» (197, n. 2).

⁷ Selon l'auteure, «< de *hijo* (espagnol) = fils» (p. 115; cf. encore p. 296). Cela n'est pas entièrement impossible, la jota espagnole s'étant prononcée comme une constrictive pré-palatale sourde à l'époque coloniale.

9. *laghia* n. m. «danse populaire d'origine africaine» (*Jours immobiles*, p. 75, n. 1). – Ø TLF. – «lutte dans les veillées mortuaires» Telchid 1997 (définition maladroite, cf. note 8 ci-dessous). – Sources créoles: Jourdain 1956: 185⁸, 296; Ludwig *et al.* 2002: 195 s.v. *lagya*.

10. *moussache* n. f. «farine obtenue à partir du manioc râpé, puis séché» Telchid 1997. – Ø TLF; Frantext: 2 att. de Zobel. – Sources créoles: Jourdain 1956: 19 (*farine moussache* «féculé de manioc à très gros grains»), 196 (*fai moussache*), 196, n. 1 «Avec le manioc on fait deux sortes de farines, avec le gluten: la farine de manioc avec l'amidon: la farine de moussache, avec celle-ci on confectionne des galettes appelées cassaves < mot africain influencé par l'espagnol» (*sic*, ponctuation); Tourneux / Barbotin 1990: 280 (s.v. *mouchach*), 282 (s.v. *mousach*²); Ludwig *et al.* 2002: 233 s.v. *mouchach*. – Le mot figure aux Inconnus du FEW (21, 208ab); il est attesté dp. Enc 10 (1765), d'abord sous la forme *mouchache*. Cette source voudrait le rattacher à l'esp. *muchacho* n. m. «enfant», à travers la métaphore «le petit du manioc», mais le FEW signale l'absence d'un tel sens en espagnol d'Amérique, tout en admettant qu'il n'existe pas de types correspondants dans les langues indigènes des Caraïbes.

11. *petites-bandes* n. f. pl. «groupes de jeunes enfants recrutés pour travailler dans les plantations de canne à sucre». Composé de l'adj. antéposé *petit*, qui doit évoquer ici non seulement la taille du groupe mais aussi le jeune âge des enfants, et du substantif *bande* n. f. «groupe de plusieurs personnes rassemblées pour un but précis». Très fréquent dans *La rue Cases-Nègres* (cf. Frantext); toujours au pluriel. – Ø TLF. – Mot sans tradition lexicographique.

12. *quimbois* n. m. «sortilège, maléfice» (Telchid 1997). – Ø TLF. – Sources créoles: Jourdain 1956: 254 et n. 1 (*quimbois*);⁹ Tourneux / Barbotin 1990: 188 s.v. *kenbwa*; Ludwig *et al.* 2002: 170 s.v. *kenbwa/kyenbwa*.

5. Conclusion

Comme le montre ce trop bref exposé, la numérisation et le balisage de l'œuvre en prose de Joseph Zobel se sont avérés très rentables pour la recherche lexicologique. Les mises en relief typographique de l'auteur, ainsi que les notes de bas de page, livrent au chercheur des

⁸ «[...] *l'aguia*, divertissement dansé par les hommes et qui simulait une lutte, un combat guerrier: nom et chose venant d'Afrique. La race *aguia*, très belliqueuse, occupe la région d'Athihémé au Dahomey, son nom est encore vivant à la Martinique non seulement dans cette danse (peut-être disparue actuellement), mais dans l'épithète attribuée à un serpent vigoureux et combatif: *sèpent aguia*».

⁹ Selon l'auteure, «< tiens, bois! phrase dite par le sorcier qui administre un philtre; c'est du moins l'explication la plus plausible qu'on ait trouvée jusqu'ici pour ce mot». Nous aurions donc affaire à un délocutif. Une telle formation n'est pas entièrement farfelue; la palatalisation *tiens* > *quim*-est normale en créole martiniquais. En outre, des délocutifs issus de formules de libation sont attestés en français, cf. Buchi 1995: 147.

matériaux dont l'intérêt diasystémique se confirme dans la grande majorité des cas. La liste de spécificités, tout comme celle des mots à fréquence nulle dans le corpus de référence, ont également fourni des matériaux qui dans la plupart des cas se sont révélés pertinents, comme les vérifications manuelles l'ont démontré. Il serait donc rentable d'appliquer les mêmes stratégies de repérage à des corpus beaucoup plus étendus, non gérables par les dépouillements manuels classiques. Des phénomènes intéressants passeraient évidemment à travers les mailles du filet, mais la pêche n'en serait pas moins abondante et de bonne qualité. Quant au choix des œuvres de Zobel pour cette première ébauche d'une banque de littérature antillaise, on ne peut que s'en féliciter: les mots intéressants y foisonnent, et les contextes définitoires et variés vont nous permettre d'élaborer un glossaire qui devrait faire ressortir tout l'intérêt du français des Antilles pour l'histoire de la langue française.

Bibliographie

- Beniamino, Michel (1996): *Le français de la Réunion: Inventaire des particularités lexicales*. Vanves: EDICEF.
- Buchi, Eva (1995): *Typologie des délocutifs galloromans*. In: *Estudis de lingüística i filologia oferts a Antoni M. Badia Margarit*. Vol. 1. Barcelone: Publicacions de l'Abadia de Montserrat, 141-163.
- Chaudenson, Robert (1974): *Le lexique du parler créole de la Réunion* (2 voll.). Paris: Champion.
- DHFQ 1998: Poirier, Claude (ed.) (1998): *Dictionnaire historique du français québécois: Monographies lexicographiques de québécoismes*. Québec: PUL.
- DRF 2001: Rézeau, Pierre (ed.) (2001): *Dictionnaire des régionalismes de France: Géographie et histoire d'un patrimoine linguistique*. Bruxelles: De Boeck / Duculot.
- DRFA 2007: Rézeau, Pierre (2007): *Dictionnaire des régionalismes du français en Alsace*. Strasbourg: PUS.
- DSR 1997: Thibault, André (1997): *Dictionnaire suisse romand: Particularités lexicales du français contemporain*. Genève: Zoé.
- DSR 2004: Thibault, André (2004): *Dictionnaire suisse romand: Particularités lexicales du français contemporain*. Nouvelle édition revue et augmentée par Pierre Knecht. Genève: Zoé.
- FEW: Wartburg, Walther von (1922-2002): *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes* (25 voll.). Bonn / Leipzig / Bâle: Teubner / Klopp / Zbinden.
- Hazaël-Massieux, Marie-Christine (1989): *La littérature créole: entre l'oral et l'écrit?* In: Ludwig (1989), 277-305.
- ILQ: *Index lexicologique québécois* (www.tlfq.ulaval.ca/ilq/).
- Jourdain, Élodie (1956): *Le vocabulaire du parler créole de la Martinique*. Paris: Klincksieck.
- Lafon, Pierre (1980): *Sur la variabilité de la fréquence des formes dans un corpus*. In: *Mots* 1, 128-165.
- Ludwig, Ralph (ed.) (1989): *Les créoles français entre l'oral et l'écrit*. Tübingen: Gunter Narr.
- Ludwig, Ralph et al. (2002): *Dictionnaire créole français*. Vol. 1. Servedit / Éditions Jasor.
- Martin, Jean-Baptiste / Pellet, Jean (1987): *Les richesses du français régional: Mots du Nord-Dauphiné recueillis à Meyrieu-les-Étangs*. Paris: Éd. du CNRS.
- Nouveau Petit Robert 2008: *Le Nouveau Petit Robert: Dictionnaire alphabétique et analogique de la langue française*. Nouvelle édition du Petit Robert de Paul Robert. Texte remanié et amplifié sous la direction de Josette Rey-Debove et Alain Rey. Paris: Dictionnaires Le Robert.

- Rheims, Maurice (1969): *Dictionnaire des mots sauvages: écrivains des XIX^e et XX^e siècles*. Paris: Larousse.
- Telchid, Sylviane (1997): *Dictionnaire du français régional des Antilles*. Paris: Bonneton.
- Thibault, André (1999): *Grand-maman et grand-papa en costume de bain au petit-déjeuner! Contribution à l'histoire de quelques lexies complexes*. In: *Cahiers de lexicologie* 75, 2, 35-54.
- (2006): *Glossairistique et littérature francophone*. In: *RLiR* 70, 143-179.
- (2007): *Banques de données textuelles, régionalismes de fréquence et régionalismes négatifs*. In: Trotter, David (ed.): *Actes du XXIV^e Congrès International de Linguistique et de Philologie Romanes, Aberystwyth 2004*. Vol. 1. Tübingen: Niemeyer, 467-480.
- TLF: *Trésor de la langue française. Dictionnaire de la langue du XIX^e et du XX^e siècle* (1971-1994). sous la dir. de P. Imbs (vol. 1-7) puis de B. Quemada (vol. 8-16). Paris: Gallimard.
- Tourneux, Henri / Barbotin, Maurice (1990): *Dictionnaire pratique du créole de Guadeloupe (Marie-Galante) suivi d'un index français-créole*. Paris: Karthala / A.C.C.T.

