

Néologie et sms

Louise-Amélie Cougnon et Richard Beaufort
Université catholique de Louvain – CENTAL

Abstract

Cet article livre un aperçu des potentialités de recherche en linguistique que peut révéler un corpus de sms, en exploitant la voie de la néologie. Après avoir circonscrit et défini la notion d'*écrit sms*, les auteurs présentent les corpus francophones étudiés ainsi que la méthodologie d'extraction précise qui a mené à plus d'une centaine de candidats, propres au contexte de la communication médiée par ordinateur. Ces candidats sont ensuite détaillés et analysés selon leur provenance géographique, leur procédé de création et leur contexte d'apparition.

Introduction

La communication médiée par ordinateur (CMO, Panckhurst : 1997) est aujourd'hui une pratique du quotidien (courrier électronique, messagerie instantanée, sms¹, sites de réseautage social, etc.) qui s'illustre dans toutes les sphères socioprofessionnelles (scolaire, professionnelle, privée, etc.) et à pratiquement tous les âges. Cette omniprésence s'est réalisée rapidement et a donné lieu à un vif intérêt scientifique multidisciplinaire (linguistique, psychologie, anthropologie, etc.) autant qu'à un enthousiasme de la population, qui tantôt voit dans cette pratique un moyen de renouer avec l'écrit, tantôt crie au scandale devant ce qu'elle considère comme un nouvel outil de destruction de la langue. Une démarche descriptive permet toutefois de sortir de ce débat et de proposer une série de phénomènes propres à la CMO (Crystal : 2006 ; Fairon et al. : 2006a) et de comportements envers la CMO (Guitteny, 2007).

Cette optique descriptive s'applique également à l'*écrit sms*² (Panckhurst : 2008), un écrit qui interpelle un certain nombre d'auteurs (Baron, 2008 ; Anis, 2007 ; Liénard, 2005). Nous avons précédemment démontré la diversité et la complexité linguistiques qui résident dans l'écrit sms, notamment sa richesse lexicale : régionalismes, néologismes et argot s'y retrouvent de façon très récurrente (Cougnon, 2008 ; Cougnon & Ledegen, 2010). Pour compléter l'engouement scientifique intense pour la recherche de néologismes dans des corpus écrits (ou retranscrits de l'oral) de type journalistique, littéraire ou politique (Sablayrolles, 2000 ; Tournier, 2002 ; Cabré et Yzaguirre, 1995), nous avons précédemment décidé d'exploiter la voie de la néologie dans les sms en nous basant sur un corpus de sms belge collecté en 2004 (Cougnon, 2010). En effet, il nous semblait important d'analyser les pratiques réelles de la population, la néologie dans « l'usage quotidien ». Ce cheminement a montré des résultats que nous sommes à même de compléter grâce à deux nouveaux corpus de sms francophones : un corpus réunionnais et un corpus suisse, collectés respectivement en 2008 et 2009. Ces nouveaux corpus ouvrent deux voies : d'une part, ils permettent de tester la vitalité des candidats à la néologie de 2004 (certains se sont répandus dans la communication

¹ Ou *short message service*.

² Nous définirons cette notion plus précisément au point 1.

par sms, d'autres ont disparu, certains enfin viennent de voir le jour) ; d'autre part, ils offrent une analyse diatopique des néologismes, une comparaison de nos données néologiques à travers ces trois aires de la francophonie.

Dans la première partie de cet article, nous nous pencherons sur la notion d'*écrit sms* et présenterons nos trois corpus, en détaillant la méthode de collecte. Nous présenterons également la méthodologie d'extraction des néologismes. Dans la troisième partie, nous détaillerons les candidats à la néologie provenant des trois corpus. Nous aborderons ensuite les possibles variations diatopiques et diachroniques illustrées dans ces corpus. Enfin, nous proposerons de futures améliorations de notre méthodologie et de nos résultats.

1. Écrit sms et constitution de corpus

La notion d'*écrit sms* rend compte de plusieurs propriétés importantes pour la néologie : il s'agit médialement parlant d'une forme d'écrit (et non d'oral ou d'oral transcrit) et d'un écrit nouveau (depuis 1985), le sms, son lieu de création. Il s'agit d'un écrit électronique qui permet donc des recherches automatiques sur corpus. Mais cet écrit montre aussi des caractéristiques de l'oral³, notamment la spontanéité de communication qui s'illustre plus souvent à l'oral qu'à l'écrit, et qui a valu à l'écrit sms l'étiquette de « parlécrit » (Dejond, 2006). Du point de vue du registre, les messages ont un caractère peu officiel et montrent un langage plus familier (tout en restant dans le registre général⁴) que l'on assimilerait volontiers à la 'Nähesprache'⁵ (Koch & Oesterreicher, 1985)⁶ (Cougnon & Ledegen, 2010).

Certaines caractéristiques originales de la pratique du sms entraînent une très grande créativité linguistique : en effet, cette pratique a tendance à inhiber les peurs traditionnelles face à la communication. Cette pratique, qui s'illustre le plus souvent en situations informelles, fait disparaître le poids de la norme (aucune autorité normative⁷) et amoindrit la peur d'être jugé par l'interlocuteur (absence physique de ce dernier). En outre, la communication par sms montre un caractère ludique, autant dans le contenu que dans la forme des messages (Moise, 2007) : nous l'observons dans les jeux sur les mots, les rébus, les phonétisations, etc. Par ces diverses propriétés, l'écrit sms favorise la création. Les néographismes de la CMO (*slt koi de 9 ?*⁸) en sont la manifestation la plus évidente, à tel point qu'on oublie souvent de prêter attention à d'autres phénomènes qui s'illustrent également dans le domaine lexical (comme cités précédemment : usage de régionalismes, recours à l'argot, aux emprunts, etc.). Les néologismes n'y font pas exception.

L'écrit sms sera donc caractérisé comme suit : un écrit souvent spontané et familier, presque toujours créatif et ludique, autant de critères qui encouragent à la formation de mots et de sens nouveaux.

³ Ce que relèvent de nombreux auteurs tels que Dejond (2006), Marcoccia (2000), Moise (2007), etc.

⁴ Des études spécifiques se sont penchées sur la néologie du français populaire ou de l'argot (voir par exemple François-Geiger, 1991).

⁵ « Langue de proximité », « communication relâchée ».

⁶ Il s'agit évidemment d'une tendance générale, à relativiser : en effet, des messages formels et brefs, adressés au milieu professionnel ou à un service (par exemple, prise de rendez-vous auprès d'un professionnel de la santé), apparaissent également dans le corpus et ont tendance à pénétrer de plus en plus le marché du sms.

⁷ Mis à part le dictionnaire intégré au téléphone portable, qui est facilement désactivé (et très souvent, selon les données sociologiques de la collecte en Belgique).

⁸ Pour *Salut quoi de neuf ?*

Abordons à présent les trois corpus qui ont servi de matériau de base à notre étude néologique. Pour ce faire, il faut brièvement détailler la méthode de collecte qui a été suivie dans les trois zones géographiques : la Belgique francophone, la Réunion et la Suisse. Le corpus de sms belge a été collecté en 2004 par deux centres de recherche de l'Université catholique de Louvain (UCL)⁹. Il a été constitué dans le cadre d'un projet nommé *Faites don de vos sms à la science* et a permis de récolter 75 000 sms en deux mois. Un sous-ensemble de 30 000 sms, comprenant les textes originaux (ou « bruts ») et leur transcription manuelle en graphie standard est disponible pour la recherche et diffusé sur CD-Rom (Fairon et al., 2006b). C'est cette dernière ressource qui a servi de base à notre étude. Le corpus réunionnais, pour sa part, s'inscrit dans l'élargissement international du projet belge, qui se nomme *sms4science*¹⁰ (sms for science) ; il a été assemblé en 2008 par l'Université de la Réunion¹¹ et a offert un ensemble de 20 000 sms. La Suisse, troisième région investie le projet, a collecté en 2009 plus de 20 000 sms dans les trois aires linguistiques, dont plus de 4 000 pour la partie francophone. Les universités de Zurich et Neuchâtel ont mis en place ce projet¹². Pour les corpus réunionnais et suisse et contrairement au corpus belge, nous avons eu recours à une transcription automatique¹³, qui a servi de base à nos recherches.

La réalisation des collectes de sms requiert, dans chaque zone géographique, l'obtention de sms provenant d'un échantillon le plus représentatif possible¹⁴ des usagers réels du sms : nous avons indiqué vouloir obtenir, dans chaque région, des messages envoyés lors de véritables échanges conversationnels, et non des messages qui nous seraient adressés directement. La procédure de transmission et de recopiage automatique a permis d'éviter les biais obtenus lors du recopiage manuel pratiqué habituellement pour ce type d'enquête. Une fois la collecte terminée, nous avons traité ces corpus afin qu'ils répondent aux exigences juridiques et scientifiques : nettoyage du corpus¹⁵, anonymisation de toutes les données personnelles. Enfin, nous avons réalisé une transcription (manuelle ou automatique), nécessaire pour une étude néologique, en graphie standard permettant l'exploration automatique sur des formes lexicales reconnues.

Par ailleurs, les participants à la collecte étaient également invités à remplir un questionnaire sociolinguistique sur l'Internet. Pour une grande partie des messages, ce questionnaire nous a permis de récolter de nombreuses informations concernant les auteurs (l'âge, le sexe, la profession, la langue maternelle, la langue seconde, le domicile, etc.), mais également leurs pratiques sms. Au total, 2 700 participants belges, 120 réunionnais et 1 250 suisses ont

⁹ Le Centre de traitement automatique du langage (Cental) et le Centre d'étude des lexiques romans (Celexrom).

¹⁰ Pour plus d'information, voir www.sms4science.org.

¹¹ Pour plus d'information au sujet du projet à la Réunion : www.lareunion4science.org

¹² Pour plus d'information au sujet du projet en Suisse : <http://www.sms4science.ch/>.

¹³ Cette transcription automatique constitue la première partie du projet général *Vocalise*, un projet de synthèse vocale des SMS (<http://cental.fltr.ucl.ac.be/projects/vocalise/>) et est décrite en détail dans Cougnon et Beaufort (2010). Elle permet des transcriptions de type *Cc ca va? G tjs pa dcred c pr ca kjécrici pa. É dmin soir jpe pa pr lcinè jss ac mé pote dlecol sur gnev vers Coucou ça va? J'ai toujours pas de crédit c'est pour ça que je t'écris pas. Et demain soir je peux pas pour le ciné je suis avec mes potes de l'école sur Genève*. Notons que ce transcripateur a été conçu à partir du corpus belge et qu'il n'est donc pas encore tout à fait adapté à la transcription de corpus d'autres parties de la francophonie.

¹⁴ Pour atteindre cette représentativité, nous avons mis en place des appels à participation diversifiés (articles dans la presse, actions sur les sites de réseautage social, etc.). Pourtant, le problème de la représentativité est complexe lorsqu'il s'agit de corpus dépendant du bon vouloir de la population. Lors de précédentes recherches, nous avons montré l'impossibilité d'assurer cette représentativité de façon satisfaisante (Cougnon et François, 2010).

¹⁵ Elimination de messages qui étaient adressés à l'équipe, de messages publicitaires ou circulaires (ex. : chaîne de l'amour, etc.).

répondu à cette enquête¹⁶. Nous ne détaillerons pas ces profils, qui ne nous intéressent pas dans l'étude menée ici.

2. Méthodologie d'extraction de néologismes

Notre recherche a porté sur les créations lexicales formelles¹⁷ apparues de 2004 à 2010 (période qui couvre les collectes de sms). Pour la mise en œuvre de notre méthodologie, nous avons tiré parti de la retranscription des sms en graphie standard. Nous avons choisi d'appliquer sur ces transcriptions un dictionnaire électronique à large couverture (le DELAF¹⁸), qui a servi de corpus d'exclusion. Les processus d'identification et d'application d'un corpus d'exclusion ont été réalisés à l'aide de l'outil *Unitex*¹⁹. Cette procédure informatique permet de récupérer un fichier contenant les mots connus du dictionnaire et un fichier contenant les mots inconnus qui seront traités comme autant de candidats à la néologie.

Les fichiers de mots inconnus comptaient 6 609 lexies pour la Belgique, 2 311 pour La Réunion et 1 177 pour la Suisse. Chaque liste de lexies inconnues a été triée manuellement : nous avons éliminé les appellations non ambiguës (prénoms, diminutifs et autres hypocoristiques²⁰), les doublons, les toponymes, les mots issus de fautes de frappe de la part du transcripateur (pour le corpus belge), les erreurs du transcripateur automatique (pour les deux autres corpus), les séquences rédigées en une langue autre que le français²¹, les marques, les interjections ainsi que les mots connus du français *non conventionnel*²².

Chaque lexie de cette liste épurée a ensuite été analysée en cotexte²³, en étudiant l'ensemble des environnements linguistiques de la lexie dans le corpus. Pour ce faire, nous avons recherché chaque candidat de chacune des listes et ses formes fléchies dans les corpus : une liste de cotextes pour chaque candidat a permis d'affiner la sélection de néologismes. Nous avons uniquement gardé les formes qui apparaissaient clairement comme des candidats à la néologie, et non comme des néographismes, des jeux sur la forme ou autres phénomènes qui mériteraient chacun, une analyse à part entière, et qui ont par ailleurs déjà fait l'objet de diverses études (Fairon, 2006a, pp. 31-48), mais que nous ne traitons pas ici.

¹⁶ Ces chiffres ont été arrondis et, pour la Suisse qui n'a pas encore achevé le traitement de ses données, il s'agit d'un calcul préliminaire, et concernent les Suisses toutes aires linguistiques confondues.

¹⁷ Définies comme des « mots qui n'existaient pas dans un état antérieur de la langue et qui sont obtenus, pour la plupart, par dérivation ou composition » (Sablayrolles, 2000 : 43).

¹⁸ Le dictionnaire DELA des formes fléchies du français est un dictionnaire électronique constitué par l'Université de Marne-la-Vallée et composé de 683 824 entrées simples pour 102 073 lemmes différents et 108 436 entrées composées pour 83 604 lemmes différents (Courtois, 2004). Ce dictionnaire n'a pas été mis à jour depuis 2001, d'où notre choix d'étudier les néologismes des dix dernières années.

¹⁹ *Unitex* est une plate-forme libre de traitement de corpus : <http://www-igm.univ-mlv.fr/~unitex/>.

²⁰ Ces petits noms affectifs sont facilement identifiables : ils se situent souvent en tête de message, sont précédés d'un terme de salutation et sont constitués de diminutifs de mots, d'agglutinations et/ou de répétitions de syllabes. Ex. : *bibou*, *boutchou*, *chouchou*.

²¹ À ne pas confondre avec les *emprunts*, lexies simples ou composées provenant d'une langue étrangère et intégrées au lexique français, utilisées en contexte de communication francophone.

²² Notion inventée par Jacques Cellard (1980) et qui regroupe les notions de français familier, populaire et argotique.

²³ Le cotexte étant défini comme « l'environnement textuel immédiat d'une forme linguistique à interpréter (Cislaru et Stiri, 2009).

Un certain nombre de problèmes ont entravé notre recherche. Le plus important concernait le corpus réunionnais : en effet, ce corpus a montré de nombreux messages où le créole était particulièrement présent, en alternance avec le français ; puisque le transcritteur automatique en graphie standard reconnaît uniquement les lexies françaises, la liste de lexies inconnues du dictionnaire était très étendue et l'absence de locuteur créolophone n'a pas permis, dans certains cas, de distinguer les mots créoles des néologismes français de la Réunion ou de lexies restées inconnues après le tri manuel. Les variétés régionales du français ont également rendu difficile notre analyse. Ainsi une lexie telle que *bec*²⁴ (291 occurrences dans le corpus suisse) aurait pu être considérée comme un néologisme sémantique, alors qu'elle apparaît clairement comme une variété régionale de *bisou*.

3. Candidats à la néologie

Nous commenterons dans cette partie les listes épurées et analysées en cotexte, qui sont détaillées dans les tableaux en annexe²⁵. Ces tableaux reprennent le candidat à la néologie, le type de néologisme (le procédé de création dont il est issu), le nombre d'occurrences dans le corpus pour ce candidat et un exemple de cotexte du candidat dans le corpus²⁶. En ce qui concerne les procédés néologiques, nous avons opté pour une typologie générale simple reprenant les grands types suivants :

- Dérivation (que nous abrègerons en *dériv*) ;
- Composition (*compo*) ;
- Agglutination (*agglut*) ;
- Emprunt (*emp*) ;
- Conversion (*conv*) ;
- Troncation (*tronc*) ;
- Mot-valise (*m-v*) ;
- Siglaison (*sigl*), qui comprend le sigle lui-même et l'acronyme ;
- Antonomase (*antono*).

Une analyse transversale des candidats à la néologie issus des trois tableaux fait ressortir plusieurs tendances.

On remarquera tout d'abord le nombre élevé d'hapax (57 hapax pour 142 lexies au total). Ce phénomène est la conséquence de plusieurs facteurs, tels que la taille limitée du corpus²⁷ ou encore le contexte ludique dans lequel ces formes sont apparues. Malgré la prépondérance des hapax, notre corpus total a montré la récurrence d'un certain nombre de créations lexicales : ainsi *lol* (emprunt à l'anglais, acronyme à partir de *laughing out loud*, littéralement « rire à

²⁴ Exemple de cotexte : *Quand je ne suis pas en vacances, je suis au taquet ! Au vu de ta méchanceté, tu n'as droit qu'à un petit bec* :*. Notons que nous n'avons pas trouvé *bec*, néologisme sémantique, par le biais de notre méthodologie d'extraction de néologismes formels, mais par le hasard de nos recherches sur corpus.

²⁵ Chaque tableau représente la liste des néologismes pour une aire de la francophonie donnée : belge, réunionnaise et suisse.

²⁶ Pour ces exemples, nous avons le choix de proposer des textes bruts ou des textes transcrits en français standard. La transcription en français standard s'est faite essentiellement sur la graphie, en prenant garde de bien respecter la syntaxe spontanée des usagers du sms. Ainsi *quel site utilise tu pour télécharger* est retranscrit en *quel site utilises-tu pour télécharger*, et *Jaim po le crumble moi* est devenu *J'aime pas le crumble moi*. Afin de faciliter la lecture des cotextes, nous avons choisi de proposer les sms transcrits.

²⁷ 54 000 messages constituent le corpus total d'analyse, ce qui représente, somme toute, un corpus restreint de situations de communication. Dans cette perspective, on considérera que certains hapax n'ont pas eu l'occasion, au travers de ce corpus restreint, de montrer l'ampleur de leur apparition en situation de communication écrite et spontanée.

haute voix ») qui apparaît 2 750 fois dans les trois corpus, *sms* (emprunt à l'anglais, sigle à partir de *short message service*, littéralement « service de message court »), avec 1 002 occurrences, *aprèm* ou *aprem* (troncation et agglutination de *après-midi*, très répandue dans ce que l'on a appelé précédemment le *français non conventionnel*, et récemment lexicalisée), avec 724 occurrences, ou encore *mdr* (siglaison à partir de *mort de rire*) qui apparaît à 571 reprises.

On notera ensuite la quantité importante d'emprunts (deuxième procédé de création dans la communication par sms), notamment parmi les lexies les plus récurrentes, principalement à la langue anglaise. Ainsi *asap* (acronyme de *as soon as possible*, « aussi vite que possible »), *bug* (néologie sémantique à partir de *bug*, « insecte », pour signifier « anomalie de fonctionnement »), *email* (néographisme de l'agglutination de *electronic mail*, « courrier électronique ») ou encore *fitness* (dérivation de l'adjectif *fit*). Nous avons également trouvé un certain nombre de lexies étrangères qui suivent le contexte linguistique de chaque aire linguistique donnée (le néerlandais pour la Belgique, le créole pour la Réunion, l'allemand pour la Suisse) ; mais, comme nous l'avons exprimé précédemment, les séquences de plus d'une lexie ont été écartées de l'analyse néologique parce qu'elles ne font pas partie du lexique français et sont utilisées en cotexte allogène. A travers nos trois corpus, nous remarquons également la capacité des emprunts à engendrer un nombre important de dérivés dans la communication par sms ; *looker*, *loveuse*, *débriefing*, *chatter*, *speecher*, *smser* et *e-mailiser* sont des exemples pertinents de ces dérivations.

Concernant toujours la typologie des néologismes, c'est bien sûr la dérivation qui est le premier procédé de création dans nos corpus. Etonnamment, la composition n'est pas particulièrement propice à la création dans les sms (moins d'une dizaine de lexies). Nous avons encore été surpris de rencontrer, dans une liste de 142 néologismes, pas moins de 4 mots-valises, procédé d'ordinaire peu répandu ; il s'agit de *bichique* (variante orthographique de l'agglutination de *biche* et de *chic*), *bisounours*²⁸ (issu de *bisou* et de *nounours*), *moisversaire* (jour de chaque mois durant lequel on fête un événement, créé sur le modèle de *anniversaire*) et *cybernamoreutisme* (issu de *cybernétique* et *amoureuxisme*, lui-même néologisme dérivé de l'adjectif *amoureux*).

Enfin, si l'on analyse les domaines ou champs notionnels concernés par la création lexicale dans les sms, on note des néologismes répondant à des technologies nouvelles : les usagers du sms *textent*, s'offrent des *divix* et des *dvd*, usent du *wap* et des *emails*. Certaines nouvelles réalités arrivent en francophonie en même temps que leur dénomination : c'est le cas de l'emprunt à l'allemand *Intervision*, utilisé pour désigner des groupes d'échange de pratiques et d'expériences, voire des séances de questions-réponses dans le but de résoudre certains problèmes dans le milieu socioprofessionnel. Lors de la communication par sms, il est souvent question de personnes que l'on s'amuse à décrire, et on voit donc apparaître des créations lexicales qui se réfèrent à des traits humains : on parle de *footeux* (amateurs de football), de *grunges* (conversion, personnes qui s'habillent et vivent comme les musiciens et chanteurs du style pop rock *grunge*), de *nymphos* (de *nymphomanes*) et de *loveuses* (« amoureuses »). Le domaine des sentiments est très convoité dans la communication par sms, notamment au travers des hypocoristiques communs aux trois aires de la francophonie *bisette* et *bisouille* (pour « petit bisou » ou « petite bise »), mais aussi les dérivés plus familiers *dodoter* (« faire un petit dodo ») et *craquous* (issu de *crack*, signifiant « petits cracks »). De manière générale, on conclura que les champs notionnels abordés dans les sms

²⁸ Sans référence au dessin animé (voir annexes 1 et 3), notons que ce mot-valise, n'est pas un hapax, et qu'il apparaît 11 fois dans nos corpus.

touchent majoritairement au domaine comprenant la nature humaine et les loisirs, contrairement au cadre des corpus journalistiques et politiques, par exemple, pour lequel ce domaine n'arrive qu'en troisième position (Sablayrolles, 2000 : 327).

4. Variations diatopique et diachronique

Les néologismes extraits de notre corpus peuvent également être analysés sous un angle diatopique ou diachronique. En effet, tout d'abord, nous remarquons des différences de résultats en fonction de chaque aire géographique (belge, réunionnaise et suisse). S'il existe bien des créations lexicales répertoriées dans les trois corpus²⁹, telles que les célèbres *lol*, *mdr*, *sms* ou *aprèm*, certaines lexies, que l'on pensait rencontrer partout, sont absentes d'un, voire de deux corpus ; ainsi nous remarquons que *blog*, *djembé* et *smser* ne se trouvent pas dans le corpus suisse. Cette absence peut s'expliquer par la taille réduite du corpus suisse : un corpus de 4 000 sms ne permet pas d'aborder le même nombre de thématiques et ne contient pas un lexique aussi varié que des corpus de 20 000 ou 30 000 sms. Toutefois, cette explication ne vaut pas pour certaines lexies, telles que *gsm* et *texto*, deux lexies communément employées dans la communication par sms et lexicalisées. Ces absences particulières sont révélatrices d'une différence dans les variétés du français étudiées : les nouvelles réalités sont nommées différemment dans les diverses aires de la francophonie. Le cas de *texto* est lui-même très pertinent : son dérivé, le verbe transitif direct *texter*, répertorié 137 fois dans le corpus réunionnais, n'apparaît pas une seule fois, ni dans le corpus belge, ni dans le corpus suisse³⁰. De la même façon, nous trouvons *préloc* en Suisse, apocope évidente de *prélocation* qui ne pouvait se trouver dans le corpus belge où l'on emploie plus volontiers sa variante *prévente*. Tel est également le cas du verbe *soupercommuter*, hapax, composition à partir des lexies belges *souper* (« dîner ») et *commu* (« cuisine et living communautaires dans les dortoirs étudiants »).

D'un point de vue grammatical, tous corpus confondus, nous constatons que plus de la moitié des néologismes sont des substantifs, la deuxième catégorie la plus fréquente étant l'adjectif, suivi du verbe³¹. Or, ici aussi, les aires géographiques se distinguent les unes des autres. En effet, dans le corpus réunionnais, le verbe passe en deuxième position, tout comme dans le corpus Suisse, où l'on ne rencontre d'ailleurs aucun néologisme adjectival.

Nous remarquons également des phénomènes néologiques évoluant dans le temps. Il serait dangereux de conclure à propos de l'absence dans les corpus de 2008 et 2009 de certains néologismes apparus dans le corpus de 2004, puisque ces absences peuvent résulter, comme nous l'avons déjà dit, de la différence de taille de corpus ou, comme nous venons de le voir, de la variation diatopique. Toutefois, nous pouvons mettre en évidence plusieurs phénomènes significatifs. Premièrement, il est intéressant de trouver par exemple *bug*, non présent dans les dictionnaires en 2001, qui a depuis lors engendré des dérivés, tels que le verbe intransitif *buger* (encore orthographié *bugger* ou *buguer*) que l'on retrouve dans les deux corpus tardifs suisse et réunionnais.

Il est également pertinent de remarquer que certains néologismes de 2004 sont entrés dans le dictionnaire³² par la suite. Ainsi, nous trouvons sans surprise les plus récurrents *sms*, *gsm*, *aprèm*, *dvd*, *appart*, *fitness*, *blog*, *tiramisu* et *bug*, mais nous y découvrons également des hapax de 2004 : *débriefing*, *nympho* et *squatteur*. Enfin, soulignons que certaines lexies très

²⁹ Pour une liste exhaustive, voir annexe 4.

³⁰ Ceci s'explique notamment par l'usage quasi unique de l'appellation *texto* à la place de *sms* en France, la Réunion suivant la tendance métropolitaine.

³¹ Notons que cet ordre de fréquence N – adj – V suit les tendances générales énoncées notamment dans Sablayrolles (2000 : 318).

³² Pour définir cette liste, nous avons eu recours à un dictionnaire récent, le Nouveau Petit Robert 2009.

récurrentes et présentes dans les trois aires de la francophonie, avec une fréquence identique, n'ont jamais été lexicalisées ; c'est le cas de l'acronyme *lol*, du sigle *mdr* et de la variante orthographique *email*. Il s'agit peut-être de créations lexicales trop spécifiques à la communication médiée par ordinateur pour être considérées par la lexicographie générale.

Conclusions et perspectives

L'objectif de cette étude était de livrer un aperçu du potentiel que peut représenter un corpus de sms pour la recherche en linguistique, surtout lorsque l'on étudie des phénomènes propres au discours spontané. Nous avons également la volonté d'attirer l'attention sur la notion d'*écrit spontané*. Les résultats obtenus montrent que les usagers du sms emploient des néologismes qui sont entrés dans le dictionnaire et dans le discours plus formel, autant qu'ils continuent à innover dans la création lexicale.

L'étude que nous venons de mener est toutefois limitée par les données que nous possédons et gagnerait sans aucun doute à s'enrichir de nouvelles données. Nous prévoyons deux nouveaux apports importants : un corpus conséquent est actuellement en constitution au Québec et permettra de compléter l'étude des singularités et points communs de la francophonie. Ensuite, nous avons en projet de réinitialiser une collecte dans une aire déjà étudiée, par exemple en Belgique francophone, et de comparer les résultats pour une même aire à une dizaine d'années d'intervalle (quelles lexies ont disparu ? d'autres se sont-elles répandues ? etc.).

Notre méthode d'identification des néologismes pourrait elle aussi être améliorée. En effet, le transcripteur automatique va recevoir des mises à jour importantes qui permettront d'obtenir des résultats plus précis (notamment reconnaître un plus grand nombre d'unités lexicales). Ensuite, cette méthode qui travaille par reconnaissance formelle, ne permet pas de détecter les néologismes issus de conversions et les néologismes sémantiques. Nous travaillons sur une probable amélioration à ce niveau également (notamment en ayant recours à un analyseur morphosyntaxique pour reconnaître les conversions).

Bibliographie

- Le nouveau Petit Robert de la langue française* (2009). Paris, Dictionnaires Le Robert.
- Le Petit Robert. Dictionnaire alphabétique et analogique de la langue française* (2004), Paris, Dictionnaires Le Robert.
- Jacques Anis (2007). « Neography – Unconventional Spelling in French SMS Text Messages ». Dans Brenda Danet et Susan C. Herring (Hrsg.). *The Multilingual Internet – Language, Culture and Communication Online*. New York: Oxford University Press, pp. 87-115.
- Naomi Baron (2008). *Always On: Language in an Online and Mobile World*. Oxford University Press.
- Maria-Teresa Cabré et Lluís de Yzaguirre (1995). « Stratégie pour la détection semi-automatique des néologismes de presse ». Dans *TTR : Traduction, Terminologie, Rédaction : Études Sur le Texte et Ses Transformations*, Montréal, PQ, Canada, 8 : 2, pp. 89-100.
- Jacques Cellard et Alain Rey (1980). *Dictionnaire du français non conventionnel*. Paris : Hachette.
- Georgeta Cislaru et Frédérique Stiri (2009). « Texte et discours. Corpus, co-texte et analyse automatique du point de vue de l'analyse de discours ». Dans *Corpus*, n° 8, *Corpus de textes, textes en corpus* [en ligne] <http://corpus.revues.org/index1678.html> (consulté le 20 mars 2011).
- Louise-Amélie Cougnon (2010). « La néologie dans l'écrit spontané'. Etude d'un corpus de SMS en Belgique francophone ». In *Actes du Congrès International de la néologie dans les langues romanes*. Barcelone. Série Activitats, 22, pp. 1139-1154.
- Louise-Amélie Cougnon (2008). « Le français de Belgique dans l'écrit spontané'. Approche d'un corpus de 30.000 SMS ». Dans *Travaux du Cercle Belge de Linguistique* [En ligne] <http://webh01.ua.ac.be/linguist/SBKL/sbkl2008/cou2008.pdf> (consulté le 14 mars 2011).
- Louise-Amélie Cougnon et Richard Beaufort (2010). SSLD: a French SMS to Standard Language Dictionary. Dans S. Granger & M. Paquot (dirs) *eLexicography in the 21st century: New applications, new challenges*. Proceedings of eLEX2009. Cahiers du Cental 7. Louvain-la-Neuve: Presses universitaires de Louvain.
- Louise-Amélie Cougnon et Thomas François (2010). « Quelques contributions des statistiques à l'analyse sociolinguistique d'un corpus de SMS ». In Sergio Bolasco, Isabella Chiari et Luca Giuliano (éds). *Statistical Analysis of Textual Data. Proceedings of 10th International Conference JADT*, 9-11 juin 2010, volume 1. Sapienza University of Rome, pp. 619-630.
- Louise-Amélie Cougnon et Gudrun Ledegen (2010). « *c'est écrire comme je parle*. Une étude comparatiste de variétés de français dans l'écrit sms' ». In Michaël Abecassis et Gudrun Ledegen (éds). *Les voix des Français*, vol. 2, Modern French Identities, 94. Peter Lang, pp. 39-57.
- Blandine Courtois (2004). « Dictionnaires électroniques DELAF anglais et français ». Dans Ch. Leclère et al. (dirs). *Lexique, syntaxe et lexique-grammaire; syntax, lexis & lexicon-grammar*, pp. 113-123.
- David Crystal (2006). *Language and the Internet*. 2nd éd. Cambridge, Cambridge University Press.
- Aurélia Dejong (2006), *Cyberlangage*, Bruxelles, Éditions Racine.
- Cédric Fairon, Jean René Klein et Sébastien Paumier (2006a). *Le langage SMS*. Louvain-la-Neuve : Presses universitaires de Louvain, Cahiers du Cental, 3.1.

- Cédric Fairon, Jean René Klein et Sébastien Paumier (2006b). *Le Corpus SMS pour la science. Base de données de 30.000 SMS et logiciels de consultation*, CD-Rom, Louvain-la-Neuve : Presses universitaires de Louvain, Cahiers du Cental, 3.2.
- Denise François-Geiger (1991). « Panorama des argots contemporains ». Dans *Langue Française*, 90, pp.- 5-9.
- Pierre Guitteny (2007). « Langue des signes, communauté sourde et CMT ». Dans Jeannine Gerbault (éd.) (2007). *La langue du cyberspace : de la diversité aux normes*, Paris, Editions L'Harmattan, pp. 67-74.
- John Humbley, Christine Chodkiewicz et Didier Bourigault (2002). « Making a workable glossary out of a specialised corpus: term extraction and expert knowledge ». Dans Bengt Altenberg et Sylviane Granger (dirs) (2002). *Lexis in contrast, corpus-based approaches*, John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 249-267.
- Peter Koch et Wulf Oesterreicher (1985). « Sprache der Nähe – Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte ». Dans *Romanistisches Jahrbuch*, 36, pp. 15-43.
- Michèle Lenoble-Pinson et Christian Delcourt (dirs) (2006). *Le point sur la langue française. Hommage à André Goosse*, Bruxelles, Le Livre Timperman et la Revue belge de philologie et d'histoire.
- Fabien Liénard (2005). « Langage texto et langage contrôlé, description et problèmes ». Dans *Linguisticae Investigationes*, 28(1), pp. 49–60.
- Michel Marcoccia (2000). « La représentation du non-verbal dans la communication écrite médiatisée par ordinateur ». Dans *Communication & Organisation*, 18, pp. 265-274.
- Raluca Moise (2007). « Le SMS chez les jeunes. Premiers éléments de réflexion, à partir d'un point de vue ethnolinguistique ». Dans *Glottopol*, 10.
- Rachel Panckhurst (2008). « Short Message Service (SMS) : typologie et problématiques futures ». Dans Teddy Arnavielle (coord.) (2008). *Polyphonies*, pour Michelle Lanvin, Montpellier, Éditions Lu.
- Rachel Panckhurst (1997). « La communication 'médiatisée' par ordinateur ou la communication 'médiée' par ordinateur ? ». Dans *Terminologies nouvelles*, 17, pp. 56-58.
- Jean-François Sablayrolles (2000). *La néologie en français contemporain. Examen du concept et analyse de productions néologiques récentes*, Paris, Champion.
- Maurice Tournier (2002), « Des mots en politique: Préfixes branchés de la communication ». Dans *Mots*, 68, mars 2002, pp. 131-138.

Annexes

Lexie	Type	Nb	Cotexte
lol	emp	1386	<i>Pas t'éclipser en douce hein ! Pas bien ça lol ;_) je déconne si t'en as marre t'as raison casse-toi mais pas trop souvent !</i>
sms*	emp + conv	742	<i>tu réponds pas aux SMS donc OK, pas très sympa!</i>
gsm*	emp	390	<i>elle t'a dit emprunte le GSM de ta mère pour l'appeler 2 minutes elle veut te parler</i>
aprèm* ou aprem	tronc + agglut	378	<i>je pars au volley toute l'aprèm donc je risque de pas répondre avant 16h30/17h.</i>
mdr	sigl	219	<i>Bisou petite poulette signé "ta" petite soeur mdr souvenir souvenir!!</i>
dvd*	emp	106	<i>viens chez moi on va regarder une K7 j'ai pas de DVD</i>
email	emp	48	<i>Michel a reçu mon email. Il me dira quoi avant midi si c'est possible ce jour vers 14h.</i>
appart*	tronc	40	<i>Voilà c'est bon on a un appart de cinq... Faut juste qu'on aille rendre ses papiers le plus vite possible.</i>
fitness*	emp	15	<i>J'ai fitness jusque 20h30 et toi t'as cours jusque quelle heure?</i>
kiffer	dériv	15	<i>OK, donc il paraît que tu kiffes le prof de maths!</i>
bisouille	dériv	12	<i>Je te bisouille tout plein.</i>
bisounours	m-v	11	<i>Bisounours à toi mon babouin des steppes.</i>
adsl	emp	10	<i>J'ai eu Koko au téléphone j'ai l'ADSL maintenant.</i>
smser	dériv	8	<i>Et tu me SMSeras si c'est Kerry qui a gagné parce que j'aurai pas la TV à Leuven..</i>
dodoter	dériv	7	<i>J'arrive pas à dodoter : je me réveille chaque heure</i>
blog*	emp	6	<i>Tu peux me transmettre l'adresse de ton blog par SMS?</i>
tiramisu*	emp	6	<i>Vous avez mangé le tiramisu? C'était bon?</i>
asap	emp	5	<i>OK mais j'ai besoin d'une confirmation asap</i>
bug* ³³	emp	5	<i>Je viens de me rendre compte qu'y a un bug avec les cars</i>
texto*	dériv	5	<i>Je t'envoie 1 petit texto car on se parlera pas ce soir. Vais aller au dodo</i>
divx	emp	4	<i>Dis, si t'as copié d'autres DivX depuis l'autre fois, tu pourrais me les déposer dans ma boîte aux lettres, stp?</i>
néonatalité	compo	4	<i>Alors la néonatalité, ça répond bien à tes attentes?</i>
déj	tronc	3	<i>Petit déj chez nous avant départ, café croissants</i>
overbooké	dériv	3	<i>ne me demande pas à quelle heure, il est toujours overbooké</i>
rebisous	dériv	3	<i>petit SMS pour te faire de gros bisous. Ici tout va bien. Rebisous</i>
wap*	emp	3	<i>C'est normal j'ai arnaqué le WAP et il m'a désactivé mes MMS..</i>
chatter*	dériv	2	<i>Non, j'ai pas encore eu l'occasion de chatter avec lui!</i>
circocabaret	compo	2	<i>Soirée super! Circocabaret : clown plus bonne musique!</i>

³³ Les lexies marquées d'un astérisque montrent une attestation dans *Le nouveau Petit Robert de la langue française* après 2001 ; il s'agit donc de termes lexicalisés. Or, on pourrait considérer que « la lexie néologique est néologique dans l'intervalle compris entre le moment de sa création [...] et celui de son insertion dans le dictionnaire » ; dès lors, seules les lexies non marquées d'un astérisque devraient être retenues comme néologiques.

footeux	dériv	2	<i>Contente que tu t'amuses mon petit footeux :)</i>
improshow	compo	2	<i>les 3 dernières représentations de l'improshow, les 8, 22 et 29</i>
rebooter	dériv	2	<i>Si je dois graver et rebooter mon PC</i>
aquaplouf	compo	1	<i>Waouw t'as même un GSM aquatique ? (aquaplouf) GÉNIAL !</i>
autogonflant	compo	1	<i>Oublie pas le matelas autogonflant de Sof. Kiss</i>
bisette	dériv	1	<i>Pas de soucis, je digère l'affaire petit à petit. Passe une belle soirée. Bisette</i>
bisounuit	compo	1	<i>Je regarde une émission sur le sida, c'est cool, ça fout le moral. Bisounuit</i>
bouratcha	dériv	1	<i>Je suis désolée pour hier, c'est con, j'étais bouratcha</i>
cafarisé	dériv	1	<i>je suis à Léon Lepage, c'est bien mais c'est un peu cafarisé</i>
cambresque	dériv	1	<i>Lundi cambresque mais week-end dégagé ou dégageable!</i>
casage	dériv	1	<i>on se prend des délires tout le temps mais aucune trace de casage</i>
comatage	dériv	1	<i>Oui, l'endroit même de nos ébats amoureux, nos comatages</i>
comater	dériv	1	<i>Tu comates? Ou je peux venir?</i>
constructivement	dériv	1	<i>on se voit samedi constructivement</i>
cybernamoureuxisme	m-v	1	<i>On va lancer une mode de cybernamoureuxisme</i>
débriefing*	dériv	1	<i>Yep, petite fondue au Bauloy...fini votre débriefing?</i>
décacher	dériv	1	<i>Peux-tu me dire comment on décasse le numéro stp...</i>
dédramatiseur	dériv	1	<i>Calamity est là, fidèle à son poste de dédramatiseuse..</i>
dégageable	dériv	1	<i>Lundi cambresque mais week-end dégagé ou dégageable!</i>
dégoûtage	dériv	1	<i>Gros dégoûtage!!! Y a plus Eternal Sunshine au Carollywood</i>
désaoûlé	dériv	1	<i>Sinon, t'as bien dormi? T'as désaoûlé?</i>
dink	emp	1	<i>moi j'étais content d'être un couple DINK</i>
dispensatoire	dériv	1	<i>j'ai test dispensatoire en religion lundi</i>
djembé*	emp	1	<i>J'ai vraiment pas envie de me retrouver toute seule avec mon djembé, mdr!</i>
donnage	dériv	1	<i>Et toi ça marche ton donnage de cours?</i>
dormage	dériv	1	<i>Bonjour Nuno! Ça été le dormage? Moi oui</i>
éclatade	dériv	1	<i>Bonne éclatade chez Fabienne...-</i>
e-mailiser	dériv	1	<i>aurais-tu la gentillesse de me e-mailiser ton texte</i>
embiéré	dériv	1	<i>des nouvelles concernant les péripéties de ton portable "embiéré"?</i>
émollience	dériv	1	<i>Fabrication Artificielle de Malheurs Interminables et de Longue Lymphatique Émollience</i>
ensapinée	dériv	1	<i>Fais de beaux rêves, miss Ensapinée</i>
étudieusement	dériv	1	<i>Bonne merde pour ton étudieusement</i>
feignasse	dériv	1	<i>éternueuse, feignasse, simplette et casse-couilles</i>
gatopouilles	?	1	<i>j'ai envie de me blottir contre toi et de faire des gatopouilles</i>
géreux	dériv	1	<i>à bientôt Seby.. Oui la semi-géreux</i>
gnoler	dériv	1	<i>Vous vous êtes gnolés la gueule</i>

gravage	dériv	1	<i>Anne n'a rien sur son CD! Le gravage a foiré!</i>
grunge	convers	1	<i>j'ai compris avec les grunges et Damien</i>
guili	convers	1	<i>te faire des guilis dans le bas du dos</i>
hallucinamment	dériv	1	<i>Hallucinamment Aimable et Rayonnante Personne?</i>
indéplaçable	dériv	1	<i>ton "congé indéplaçable sous peine de divorce"</i>
kitchissime	dériv	1	<i>plein de couleurs, qui clignotent, sur la maison, dans le jardin! Kitchissime</i>
labelloté	antono + dériv	1	<i>Bisous tout labelloté</i>
looker	dériv	1	<i>N'oublie pas de regarder après les extraits.</i>
mamme	tronc	1	<i>à tes reins la mer a perlé rousse à tes mammes vermeilles</i>
markopsien	dériv	1	<i>Passe une bonne journée markopsienne</i>
memerde	dériv	1	<i>C'est la merde</i>
miterrandienne	dériv	1	<i>Votre respectueux et dévoué serviteur. Bise miterrandienne</i>
moisversaire	m-v	1	<i>je te souhaite 1 bon moisversaire</i>
mojo	?	1	<i>alors t'as le mojo cette semaine toi :D</i>
nympho*	tronc	1	<i>Ma mère nous invite à manger ce soir avec la cousine nympho</i>
phoner	tronc ou emp	1	<i>Bon Loïc, t'as phoné le chanteur?</i>
pichant	?	1	<i>entrée: 7€ mais vestiaires/djoks gratuits et sorteurs pichants</i>
reposage	dériv	1	<i>Bon reposage ce soir et bon cours demain</i>
ressourçage	dériv	1	<i>Ton ressourçage se passe bien? Bonne murge ce soir</i>
réunionite	dériv	1	<i>Ça s'appelle de la réunionite aiguë !</i>
solutionnaire	dériv	1	<i>si tu désires le solutionnaire pour le petit jeu d'association tu peux appeler</i>
soupercommuter	agglut + compo	1	<i>Ce soir on soupercommute puis on sort à la Casa</i>
souplette	dériv	1	<i>je viendrai- y aura un peu de souplette pour moi?</i>
speecher	dériv	1	<i>Ça va, t'as bien "speeché" sur Canal C</i>
squatteur*	dériv	1	<i>Ta grande amie et squatteuse de cours de l'année dernière</i>
stressage	dériv	1	<i>Tu vois le "stressage" de ton frère devient contagieux</i>
studier	emp	1	<i>mon père m'oblige de studier. Désolé</i>

Annexe 1. Liste de candidats à la néologie présents dans le corpus sms pour la science (Belgique), classée par nombre d'occurrences

Lexie	Type	Nb	Cotexte
lol	emp	1286	<i>T'inquiète, je me suis vengé, lol.</i>
mdr	sigl	331	<i>Mouais mais bon. L'essence c'est de l'argent mdr</i>
aprèm* ou aprem	tronc + agglut	291	<i>Coucou my love, cet aprèm pourrais-tu m'amener à mon cours 2 piano vers dix-sept heures trente stp?</i>
sms*	emp + conv	211	<i>Excuse, c'était pas pour toi le sms</i>
texter	dériv	137	<i>je te texte dans la journée si il y a possibilité qu'on se voit</i>
gsm*	emp	89	<i>Excuse moi pour le retard, mon gsm était sur mon lit!</i>
appart*	tronc	12	<i>Le temps est pourri ! Là je rentre à l'appart. Je vais préparer mon sac.</i>
blog*	emp	11	<i>Ah ok ! Té c'est triste ! J'ai pas de mot ! Clarisse a un</i>

			<i>blog?</i>
bug*	emp	7	<i>Mon imprimante bug!</i>
dvd*	emp	6	<i>C'est un film en dvd ! Lol ! Et on avait regarde le 1 et là on regarde le 2!</i>
djembé*	emp	4	<i>Ce soir j'ai mon premier cours de djembé ! Youpi.</i>
kiffer	dériv	4	<i>Dis-moi un truc y a ma copine qui kiffe un gars mais le gars c'est pas trop ça et elle veut l'oublier.</i>
email	emp	3	<i>je crois que tu n'as pas reçu mes derniers emails à ton boulot aurais-tu un email</i>
looker	dériv	3	<i>Euh non pourquoi ? Je look la télé et toi ?</i>
bisette	dériv	2	<i>Bonne soirée à toi. Bisette bisette</i>
déj	tronc	2	<i>On est à la plage, y a des petits oiseaux qui attendent nos miettes de petit déj, c'est mignon !</i>
wap*	emp	2	<i>Désolée ma connexion wap plane bonne journée bisous bye</i>
smser	dériv	1	<i>Bon, je commence à te smser quand je suis dans la voiture pour aller dans un resto à plateau caillou.</i>
adsl	emp	1	<i>Tu viens tout juste de mettre internet ? Adsl?</i>
bichique	m-v	1	<i>Je t'aime fort ma bichique de France !!! Bisou</i>
fitness*	emp	1	<i>Lol du fitness. Tiens j'ai fait un an de judo.</i>
megatop	compo	1	<i>Ben oui je viens ce midi j'ai espagnol en dernière heure du matin ! T'as pas cours là ou quoi ? Pas MEGATOP en tout cas!</i>
bisouille	dériv	1	<i>C'est ok ! Plein de bisouille cocotte</i>
divx	emp	1	<i>Fais aller. Là je mate un divx.</i>
feignasse	dériv	1	<i>Coucou c'est émi, lève toi feignasse mdr!</i>
félicitation*	tronc	1	<i>Félicitation. Tu as eu ton semestre avec 10,765. Bisou. J'ai un problème avec mon gsm</i>
loveuse	dériv	1	<i>Stp aidez et donnez-moi une réponse soutenue et digne d'une jeune fille tendre et loveuse.</i>
nympho*	tronc	1	<i>My god, une poupette frigo-nympho en crise de tétanie qu'il faut rassurer! Je suis choqué!</i>
texto*	dériv	1	<i>j'arrive dès que j'ai le bus et je t'envoie un texto dès que je suis à la fac. Désolée du retard</i>
textoter	dériv	1	<i>Chou je vais dodo moi! Suis crevé! Textote moi quand même!</i>

Annexe 2. Liste de candidats à la néologie présents dans le corpus réunionnais, , classée par nombre d'occurrences

Lexie	Type	Nb	Cotexte
lol	emp	78	<i>Alors miss, encore une grasse mat? T'es vraiment incroyable, lol. Becs à plus.</i>
aprèm* ou aprem	tronc + agglut	55	<i>Comme tu l'a dit ce sera une longue journée. Bonne aprèm, a domani.</i>
sms*	emp + conv	49	<i>Je te sms quand on est dans le bus, bisous! À toute!</i>
mdr	sigl	21	<i>J'ai raté mon train pour une minute mdr je te dis ces temps tout va mal.</i>
dvd*	emp	12	<i>tu pourrais me donner le dvd casino de james bond please ?</i>
appart*	tronc	11	<i>Ouf! On a trouvé notre appart. La rue n'est effectivement pas très engageante, un clochard y dort, et on a pas encore vu la chambre, mais on a déposé nos bagages pas loin.</i>
fitness*	emp	7	<i>Salut miss! Comment va ? Tu t'es motivée pour le fitness ? moi me réjouis trop de reprendre !</i>
bisouille	dériv	4	<i>Oui, moi, je suis déjà à la fac pour notre exposé sur les</i>

			<i>sectes donc, je me motive pour les stats, je suis trop sérieuse! Bisouilles</i>
email		3	<i>Je viens de vous envoyer les citations par email.</i>
bisette	dériv	2	<i>Moi, j'arrive à la t chaud à dix-sept heures. Bonne nuit! Bisette :_)</i>
bug*	emp	2	<i>Essaie d'aller regarder si t'es inscrite au cours en arrivant-sinon on est dans la merde. Mais si on est les 2 dans le même cas, c'est peut-être parce que y a eu un bug.</i>
intervision	emp	2	<i>À demain pour notre intervention, je me réjouis!</i>
asap	emp	1	<i>J'attends des news pour une visite d'appartement. Si ça marche pas, je prendrai le train vers seize heures vingt et je viens direct chez toi! Je te redis asap.</i>
bisounours	m-v	1	<i>Excuse-moi, j'ai été retenue par le boulot. Du coup ni eu le temps de venir ni celui de t'avertir... :_(Bisounours!</i>
bubulle	dériv	1	<i>Ha la la, c'était assez comique. Je pense que je vais me taper une grosse bubulle!! Enfin bref. J'ai tout mais au pif, donc il ne me reste plus qu'à prier. Amène!</i>
craquous	dériv	1	<i>Coucou les craquous, vos proposition son très sympas mais ne tombent pas du tout au bon moment pour moi</i>
débriefing*	dériv	1	<i>Me réjouis tellement d'y être! On se fera un débriefing à notre retour! À bientôt, bise</i>
pétage	dériv	1	<i>Ça va toujours ou c'est le pétage de plomb total? Bon aprèm! Bisous!</i>
préloc	tronc	1	<i>Salut ! Désolée de t'écrire si tard. Mais lauri m'a dit que t'allais peut-être chercher des prélocs demain ! Tu pourrais aussi nous en prendre deux et on te donne les sous samedi ? Becs</i>
psyoter	dériv	1	<i>Alors ça a été cette lecture jeudi passé ? Tu psyotes à fond ? Courage, plus qu'un et c'est bon ! Bisous bisous !</i>
tiramisu*	emp	1	<i>Je fais un tiramisu aux framboises pour le dessert de dimanche! Bisous</i>

Annexe 3. Liste de candidats à la néologie présents dans le corpus suisse, classée par nombre d'occurrences

Lexie
appart
aprèm ou aprem
bisette
bisouille
bug
dvd
email
fitness
lol
mdr
sms

Annexe 4. Liste des candidats à la néologie présents dans les 3 corpus, classée par ordre alphabétique