View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by OpenEdition

Cahiers de praxématique

71 | 2018
Médium et discours

De l'influence du médium sur des écrits hors-normes. Une étude comparatiste de la variation graphique dans des écrits de peu lettrés de la Grande Guerre et des SMS contemporains

On the influence of the medium on non-standard written texts. A comparative study of spelling variation in texts written during the Great War by people with little education and contemporary SMSs

Frédéric André et Beatrice Dal Bo

Édition électronique
URL : http://journals.openedition.org/praxematique/5042
ISSN : 2111-5044

Éditeur
Presses universitaires de la Méditerranée

Référence électronique
Frédéric André et Beatrice Dal Bo, « De l'influence du médium sur des écrits hors-normes. Une étude comparatiste de la variation graphique dans des écrits de peu lettrés de la Grande Guerre et des SMS contemporains », *Cahiers de praxématique* [En ligne], 71 | 2018, mis en ligne le 09 avril 2019, consulté le 01 mai 2019. URL : http://journals.openedition.org/praxematique/5042

Ce document a été généré automatiquement le 1 mai 2019.

Tous droits réservés

De l'influence du médium sur des écrits hors-normes. Une étude comparatiste de la variation graphique dans des écrits de peu lettrés de la Grande Guerre et des SMS contemporains

On the influence of the medium on non-standard written texts. A comparative study of spelling variation in texts written during the Great War by people with little education and contemporary SMSs

Frédéric André et Beatrice Dal Bo

Introduction

1 Cet article se propose d'interroger le rôle du *médium* dans un contexte spécifique de production discursive, celui de la communication privée, sur des écrits que nous considérons comme *hors-normes.* En effet, ces pratiques rendent compte d'un usage particulier du français, car plusieurs formes de variation par rapport à la norme orthographique et syntaxique standard du français peuvent y être relevées[1].

2 Ainsi, nous allons comparer d'une part, des lettres manuscrites rédigées au cours de la Grande Guerre par des soldats et des femmes peu scolarisés, recueillies dans les fonds du projet *Corpus 14* (Praxiling, 2014), et d'autre part, d'authentiques SMS contemporains issus du corpus intitulé *88milSMS* (Panckhurst *et al.*, 2014), collecté en 2011 dans le cadre du projet international *sms4science* (*www.sms4science.org*).

3 Ces deux pratiques de l'écriture, que près d'un siècle sépare, peuvent sembler très éloignées : nous confrontons en effet ici une forme de communication de type épistolaire

à une autre, beaucoup plus récente, prenant place au sein d'un ensemble plus vaste de pratiques renvoyant à la communication électronique. En outre, ces deux types d'écrits sont produits sur des supports fort différents. Cependant, ces deux pratiques partagent des points communs, comme nous le verrons : elles ont un contexte de production discursive similaire, les deux média ont des caractéristiques communes et les deux types d'écrits présentent des variations par rapport à la norme standard du français.

4 Ainsi, nous nous interrogerons tout d'abord sur la nature de ces deux média : quelles caractéristiques les séparent ou les rapprochent ? Nous allons ensuite chercher à déterminer si ces supports d'écriture peuvent potentiellement avoir une influence sur ces écrits *hors-normes* : certaines tendances concernant la variation graphique pourraient-elles être favorisées ou limitées par les caractéristiques spécifiques du médium employé ?

1. Présentation des corpus

5 Pour cette étude, nous avons fait le choix de nous baser uniquement sur des données authentiques, issues de véritables échanges. Ainsi, nous avons eu recours d'une part aux lettres des scripteurs peu scolarisés, datant de la Grande Guerre, recueillies dans *Corpus 14* , et d'autre part, à d'authentiques SMS contemporains issus du corpus intitulé *88milSMS*.

1.1. Corpus 14

6 *Corpus 14* est un projet de recherche développé au sein du laboratoire Praxiling (UMR 5267 CNRS, Université Paul-Valery Montpellier) et dirigé par Agnès Steuckardt. Initié en 2014, ce projet a pour but de recueillir et analyser des correspondances de la Grande Guerre, échangées par des scripteurs peu lettrés[2]. Pour la première fois éloignées les unes des autres – les hommes au front, les femmes à l'arrière –, ces personnes sont contraintes d'avoir recours à l'écriture, pratique non habituelle pour elles, afin de maintenir un contact (Steuckardt, 2015a) : la Grande Guerre joue ainsi le rôle de phénomène déclencheur de l'entrée dans l'écrit d'une grande partie de cette population. Ces sources donnent accès à des écrits s'éloignant de la norme linguistique à différents niveaux – orthographique, morphosyntaxique, lexical, stylistique – et permettent d'interroger l'usage écrit de la langue qui est fait par des scripteurs peu expérimentés, ayant un niveau d'instruction élémentaire, au début du XXe siècle. Les textes sont mis à disposition de la communauté scientifique sur le portail TXM du projet (http://textometrie.univ-montp3.fr).

1.2. 88milSMS

7 Le corpus *88milSMS* est le résultat d'une collecte de 6 mois, qui fut menée en 2011 dans la région de Montpellier. Cette initiative fait partie du projet international *sms4science*, qui vise l'étude de la communication par SMS par le développement d'une méthodologie de collecte de grands corpus authentiques, basée sur la contribution volontaire de participants aux profils très hétérogènes. En effet, après avoir rempli s'il le désirait un questionnaire permettant de définir son profil (âge, sexe, profession, niveau d'étude, téléphone utilisé, etc.), chaque donateur était amené à transmettre directement ses échanges par SMS aux responsables de chaque région participant au projet. Ainsi, entre 2004 et 2011, ce projet a permis de collecter et traiter (en termes d'anonymisation et

d'annotation) de grands corpus comparables de données authentiques dans plusieurs régions francophones à travers le monde : en Belgique, en Suisse, au Québec, sur l'île de la Réunion, et enfin en France.

2. Comparaison des deux média

8 Dans un premier temps, nous comparons les deux média en fonction d'une classification adaptée de Herring (2007). Il s'agira d'un modèle de classification à facettes regroupant 17 caractéristiques. Bien qu'ayant été développé pour mieux appréhender différentes formes de communications électroniques, il permet de définir un médium bien au-delà de son simple aspect physique. Comme l'auteure le rappelle, « le but d'une classification des discours médiés par ordinateur est d'articuler les aspects du contexte – à la fois techniques et sociaux – qui influencent potentiellement l'usage du discours dans les environnements médiés par ordinateur, et de les porter ainsi à l'attention consciente du chercheur »[3] (2007 : 1-2). Nous utilisons ainsi ce modèle pour confronter le médium de la lettre manuscrite à celui des SMS.

9 Les résultats de cette analyse indiquent que les deux média ne divergent pas autant qu'on aurait pu le supposer : si certains critères les opposent, la plupart ont tendance à les rapprocher, quand ils ne sont pas totalement partagés.

2.1. Caractéristiques communes

10 L'analyse de ces deux média nous amène à dégager dix caractéristiques qu'ils partagent :

- le fait que la communication s'établit par un code strictement écrit,
- le caractère privé de la communication, en opposition à la communication publique permise par certains systèmes de CéMO (Communication écrite Médiée par Ordinateur[4] (Cougnon, 2015),
- la persistance des transcriptions, fragiles sur papier, par défaut dans le cas des SMS (les messages restent enregistrés dans le système tant que l'utilisateur ne les efface pas, et sont également conservés dans les bases de données des opérateurs téléphoniques),
- la transmission individuelle de chaque message, qui se fait message par message (ou lettre par lettre), sans que le destinataire ait connaissance de la production d'un message qui lui est adressé tant qu'il ne l'a pas reçu[5],
- le fait que les utilisateurs puissent s'approprier la norme écrite standard, comme notre corpus en témoigne,
- la structure conversationnelle, normalement constituée d'un émetteur unique et, généralement, d'un seul destinataire[6],
- la variabilité des caractéristiques sociales des participants (âge, sexe, profession, compétences linguistiques, maitrise du médium utilisé, etc.),
- la variabilité des buts de l'interaction,
- la pluralité potentielle des sujets traités au sein du même message,
- la diversité des tons employés.

2.2. Caractéristiques partiellement partagées

11 Quatre critères rapprochent les deux média, bien que des divergences apparaissent parfois :

- l'anonymat : dans un mode de fonctionnement canonique, émetteurs et destinataires sont normalement connus ; toutefois, bien qu'il s'agisse de cas très rares, on peut imaginer que ces deux média peuvent également être employés lorsque l'un des deux participants reste anonyme. Ce genre d'utilisation très particulière sort cependant du cadre de nos corpus, aussi nous nous bornerons à traiter des cas où émetteur et destinataire restent connus.
- le délai de réponse, qui est généralement long. Toutefois, il peut devenir rapide en cas d'envoi de SMS courts, simulant une conversation en coprésence.
- les systèmes de filtrage des messages : ils sont absents par défaut dans les deux média, mais dans le cas des SMS, des numéros spécifiques peuvent être bloqués par le service.
- le fait que les échanges s'établissent de façon asynchrone dans les deux cas. Néanmoins, certains échanges de SMS peuvent être très rapides, dans le cadre de la simulation d'une conversation en coprésence. Il ne s'agit donc plus vraiment d'un échange asynchrone, mais nous parlerons, à la suite de Garcia & Jabob (1999), de « quasi-synchronie ».

2.3. Caractéristiques divergentes

12 Les seuls points qui distinguent indéniablement ces deux médias sont au nombre de trois, et tiennent notamment aux « affordances technologiques des systèmes de la communication médiée par ordinateur » (dans le texte source : "technological affordances of CMC systems", Herring, 2007 : 16).

13 Le premier concerne la possibilité d'insérer la citation d'une partie du message précédent dans la réponse. Ceci peut se faire de façon assistée pour les SMS, comme dans n'importe quel logiciel de traitement de texte, exigeant quelques manipulations spécifiques en fonction du service offert par le téléphone. Dans le cas de la lettre, les citations des contenus des lettres précédentes sont également possibles, mais elles sont effectuées au moyen de différentes structures linguistiques, telles que la reprise diaphonique (voir également Gomila, 2015), comme dans cet exemple : « **tu me dis sur ta l'etre** bien chér Pierre **que** je ne perde pas des moments de repos pour técrire » (Marie, 11.10.1914).

14 Deuxièmement, les deux média se distinguent par le format du message, concernant l'ordre d'apparition des messages et leur présentation visuelle. Dans le cas des SMS, ils sont limités à l'interface du téléphone (un écran de taille réduite par le biais duquel l'utilisateur peut naviguer dans le texte, alignant automatiquement chaque ligne, et offrant peu de possibilités quant à l'orientation des différentes parties d'un message) ; en ce qui concerne les lettres, cela dépend de facteurs extérieurs, comme la poste, qui peuvent causer du retard ou des chevauchements dans la réception du courrier, ainsi que du type de papier utilisé pour l'écriture (feuilles, cartes postales, etc.). Ce support permet en outre l'appropriation de toute la surface disponible pour l'écriture, en produisant ainsi parfois des écritures en marge, verticales, à l'envers, etc.

15 Enfin, la taille maximale du message constitue le dernier point distinctif de ces deux média. En effet, les lettres ne sont limitées qu'à l'espace physique offert par le support sur lequel un scripteur rédige, mais la taille des SMS, lors de leur apparition, était limitée par des caractéristiques techniques qui limitent l'utilisateur à 160 caractères par message

(tout dépassement entrainant une augmentation du coût du service). Cette contrainte s'est réduite avec le temps, en parallèle des évolutions du marché de la téléphonie mobile, comme notamment l'apparition des forfaits SMS illimités à partir de 2007. Cependant, selon les résultats d'études précédentes sur les corpus de *sms4sciences*, cette évolution n'a fait que réduire la taille moyenne des messages (Cougnon, 2015 ; André, 2017). En effet, on pourrait croire que la réduction des contraintes économiques conduirait à l'augmentation de la taille des SMS, mais au contraire, les scripteurs, qui semblaient auparavant « rentabiliser » l'espace graphique « payant » en essayant de s'approcher des 160 caractères, rédigent aujourd'hui des messages généralement plus courts, comme « OK ».

16 Ainsi, à la suite de cette première comparaison, nous pouvons mieux définir ces deux types d'écrits *hors-normes* : les résultats indiquent que l'on se situe, dans les deux cas, dans un cadre privé de communication interpersonnelle, s'établissant par l'intermédiaire d'un code écrit qui offre la possibilité de s'approprier les normes graphiques standard, et renvoyant de ce fait à la proximité communicative telle qu'elle est définie par Koch & Oesterreicher (2001).

3. Analyse de la variation graphique

17 Dans un deuxième temps, nous comparons différents phénomènes de variation graphique au sein de chaque médium à partir d'une typologie inspirée des travaux de Panckhurst (2009) et de Lopez *et al.* (2015), afin de déceler certaines tendances d'utilisation et d'appropriation du code graphique selon le type de médium employé. Pour cela, nous avons réalisé des sélections de données au sein des deux corpus *Corpus 14* et *88milSMS*.

3.1. Protocole expérimental

3.1.1. Sélection des données

18 Aussi larges soient-ils, ces deux corpus posent la question de leur représentativité, n'étant qu'un échantillon de la pratique globale dont ils sont respectivement issus. Par exemple, pour ce qui est des SMS, et plus particulièrement du corpus *88milSMS*, nous pouvons calculer que ce corpus regroupe seulement 0,003 % des 2 610 millions de messages envoyés dans la région Languedoc-Roussillon durant les 6 mois de la collecte. Ainsi, ne pouvant pas décrire ces pratiques de l'écriture dans leur globalité, nous avons fait le choix de relever des tendances marquées apparaissant dans l'utilisation qui en est faite par plusieurs profils différents. Nous avons donc cherché à établir des sélections, basées sur différents profils de scripteurs.

19 En ce qui concerne *Corpus 14*, nous avons sélectionné cinq scripteurs (les époux Fabre, Marie et Pierre, les époux Arcis, Victoria et Félicien, ainsi que Laurent Pouchet), dont nous avons étudié les dix premières lettres échangées. Cette première sélection regroupe donc un total de 50 lettres, soit 87 706 caractères (espaces compris).

20 Nous avons par la suite cherché à regrouper un nombre équivalent de SMS, permettant la comparaison de nos sélections. Pour cela, nous nous sommes basés sur le nombre de caractères (espaces compris). En effet, dans le cadre de SMS tout comme dans les lettres de *Corpus 14*, la liberté scripturale dont disposent les scripteurs conduit à des modifications de la norme du français standard, et donc parfois à une modification de la segmentation des mots par l'intermédiaire de divers phénomènes de variation (décris

infra) qui remet en question l'identification de l'unité « mot », dont la définition est aussi problématique (*inter alia*, Béguelin, 2000 ; Haspelmath, 2011). Ainsi, en nous basant sur le nombre de caractères, nous avons sélectionné un total de 1 640 messages produits par 23 profils de scripteurs différents et très hétérogènes, représentant 87 700 caractères.

21 Une fois deux corpus de taille équivalente constitués, nous avons procédé à leur comparaison.

3.1.2. Méthodologie de comparaison : typologie de la variation du code graphique et traitement des données

22 Notre seconde question de recherche ayant pour objectif de comprendre si certaines tendances concernant la variation graphique pourraient être favorisées ou limitées par les caractéristiques spécifiques du médium employé, nous avons cherché à comparer l'utilisation globale de différents types de variation graphique au sein de chaque médium.

23 Afin de distinguer ces différents types, nous avons adapté une typologie de l'*écrit-SMS* proposée par Panckhurst en 2009. En effet, cette typologie recense et organise l'ensemble des phénomènes de variation classifiés dans les typologies précédentes (notamment celles proposées par Anis *et al.*, 2004 ; Fairon *et al.*, 2006 ; Véronis & Guimier de Neef, 2006 ; Liénard, 2008), dans l'optique de « clarifier certains aspects qui ne semblent pas toujours distingués de manière limpide » (Panckhurst, 2009 : 39). Elle regroupe quatre grands types de phénomènes, qui sont les substitutions, les réductions, les suppressions et les ajouts, et elle pose également une distinction entre phénomènes d'ordre graphique et phonétisé.

24 Les substitutions, graphiques tout d'abord, affectent le code sans être oralisées, et correspondent à des cas de remplacements de caractères par d'autres. Elles affectent la typographie et la ponctuation (« *m en* » pour « m'en »[7]), l'utilisation de symboles (« *a+* »[8]), ou correspondent à des cas de remplacements de caractères n'affectant pas l'oralisation (« *bisoux* » pour « bisous »). Les substitutions phonétisées, par opposition aux substitutions graphiques, consistent à remplacer un élément destiné à être oralisé (cognitivement). Elles peuvent être entières, lorsqu'une chaîne de caractères est remplacée par un seul (« *o* » pour « au »), partielles, lorsque le remplacement des caractères n'est que partiel, sans modifier l'oralisation (« *allé* » pour « aller »), ou encore faire varier l'oralisation par rapport à la norme standard, formant des cas définis comme des substitutions phonétisées avec variation (« *nan* » pour « non »).

25 Les réductions consistent à réduire la taille des messages par une réduction volontaire du nombre de caractères utilisés dans la norme du français standard. On y regroupe donc les cas de réductions graphiques, affectant tout d'abord les caractères muets (« *vou* » pour « vous », « *douch* » pour « douche »), les cas d'abréviations (« *dc* » pour « donc », « *ele* » pour « elle »), et les agglutinations (« *jattends* » pour « j'attends »). Les réductions phonétisées correspondent quant à elles aux cas de troncation (« *ordi* » pour « ordinateur »), aux sigles et acronymes (« *mdr* » pour « mort de rire »), ou encore aux réductions créant une modification lors de l'oralisation (« *i* » pour « il », « *ui* » pour « oui »).

26 Les cas de suppression affectent uniquement les diacritiques (« *ca* » pour « ça » ; « *voila* » pour « voilà ») ou les signes de ponctuation. Cette catégorie est un peu particulière, car ces derniers n'ont pas pu être comptabilisés dans le cadre de notre sélection issue de *Corpus 14* : selon la typologie d'origine, ces cas correspondent à des messages ne laissant

apparaitre aucun signe de ponctuation[9]. Il est donc possible de repérer et comptabiliser les SMS dont la ponctuation a été totalement supprimée, ce qui permet de poser les bases d'une comparaison, mais à l'échelle d'une lettre, cette méthode n'est pas applicable. Compte tenu du grand nombre de cas de suppression (Steuckardt, 2015b), les relever selon cette méthode conduirait à calculer une moyenne d'utilisation de 100 %, puisqu'ils apparaissent dans chaque lettre. Nous avons préféré les écarter en relevant cependant qu'il s'agit d'un phénomène très courant, dans les deux cas. En ce qui concerne les diacritiques, Lucci & Millet ont observé des tendances similaires dans des écrits manuscrits quotidiens de gens ordinaires : « près de 2/3 des variations qui affectent l'ensemble des diacritiques relèvent d'une omission et surtout d'une neutralisation- des accents grave et aigu » (1994 : 54). Nos deux corpus appartenant à cette catégorie, ils ne semblent pas faire exception.

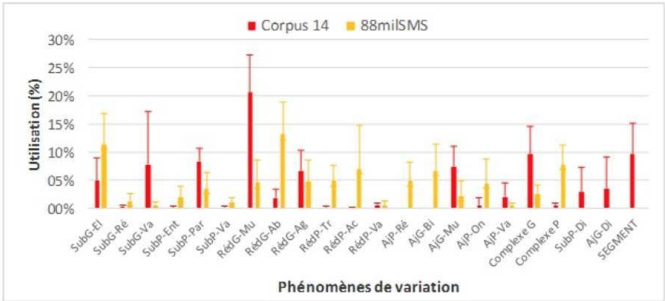
27 Enfin, les ajouts de caractères correspondent à tous types d'ajouts, qu'il s'agisse d'émoticônes («☺»), de répétitions (« *suuuuuuuper* »), de transcriptions d'onomatopées/ d'interjections (« *mouai* », « *bof* ») ou encore de l'ajout d'un seul caractère. Cependant, nous avons fait le choix de dissocier l'ajout d'un caractère muet, constituant une modification d'ordre graphique (« *je penses* » pour « je pense »), et l'ajout d'un caractère oralisé, constituant une modification phonétisée (« *oki* » pour « ok »), afin de maintenir la distinction entre les phénomènes d'ordre graphiques et phonétisés. De plus, afin de nous adapter aux phénomènes trop rares pour apparaitre dans la typologie d'origine, nous y avons également ajouté l'ajout et la substitution de diacritiques[10] (respectivement « *chér époux* » pour « cher époux » et « *chére femme* » pour « chère femme »), ainsi qu'une catégorie très problématique liée à la segmentation des mots[11] (« *cela na rive rapas* » pour « cela n'arrivera pas »). Ces cas de segmentations pourraient être considérés comme des phénomènes complexes, c'est-à-dire une combinaison de phénomènes simples employés sur la même chaîne de caractères, comme les cas d'écrasement de signes (« kestufé » pour « qu'est-ce que tu fais ? », Liénard, 2008). Cependant, nous avons fait le choix de les considérer à part afin d'être en mesure de comparer leur usage par rapport aux autres phénomènes employés dans notre sélection issue de *Corpus 14*.

28 Nous avons analysé le corpus réalisé pour cette étude en fonction de cette typologie « tout-terrain », manuellement, afin d'établir des pourcentages d'utilisation de chaque phénomène par profils, puis de calculer sur cette base des moyennes globales par médium. Comme vu *supra*, ces résultats ne sont pas représentatifs de la graphie d'un médium dans son ensemble, mais ont néanmoins permis de classer les huit types de phénomènes les plus fréquents dans chaque corpus par ordre d'utilisation.

3.2. Résultats

29 La Figure 1 présente les résultats de la comparaison des pourcentages d'utilisation de chaque phénomène de variation de notre typologie, calculés sur la base de nos relevés effectués au sein de chaque corpus. Les écarts types (dorénavant ET) illustrent les différences de proportion d'emploi d'un phénomène entre les différents profils sélectionnés.

Figure 1 : comparaison des pourcentages d'utilisation de chaque phénomène de variation (avec ET)



Légende :

Substitutions graphiques :
SubG-El => élision typographie/majuscules
SubG-Ré => icône / symboles / rébus
SubG-Va => avec variation

Substitutions phonétisées :
SubP-Ent => entière
SubP-Par => partielle
SubP-Va => avec variation

Réductions graphiques :
RédG-Mu => début/fin de mot muette /chute de *e* instables
RédG-Ab => squelette consonantique / abréviation / consonnes doubles / abréviation sémantisée
RédG-Ag => agglutinations

Réductions phonétisées :
RédP-Tr => troncations
RédP-Ac => acronymes
RédP-Va => avec variation

Ajouts graphiques :
AjG-Rép => répétition de caractères
AjG-Bi => binettes
AjG-Mu => ajout de caractère muet

Ajouts phonétisés :
AjP-On => onomatopées
AjP-Va => avec variation

Autres :
SupP-Di => Suppression phonétisée : signes diacritiques
AjG-Di => Ajout graphique de diacritique
SEGMENT => segmentation problématique

30 La figure 1 présente des pourcentages moyens d'utilisation de chaque type de phénomène de variation, calculés à partir des résultats obtenus pour chaque profil de scripteur. On peut tout d'abord observer des ET relativement importants, illustrant le fait que chaque scripteur semble faire usage de l'écriture, et plus particulièrement de la variation graphique, de façon personnelle. Dans certains cas, notamment chez les scripteurs de SMS, on pourrait même parfois parler de « style graphique personnel ».

31 En observant maintenant les moyennes globales obtenues en fonction de chaque médium, la figure 1 laisse apparaitre très clairement deux profils d'utilisation de la variation bien distincts. En effet, très peu de types de phénomènes sont employés dans des proportions similaires, comme les réductions graphiques correspondant à des cas d'agglutination (RedG-Ag). Cependant, ce type de phénomène particulier est loin d'être le plus courant, dans les deux cas. De plus, certains types sont représentés de façon exclusive dans les données issues de l'un ou l'autre médium. Par exemple, les cas de réductions phonétisées par troncation ou emploi d'un sigle/acronyme (RedP-Tr et RedP-Ac), ainsi que les répétitions de caractères et ajouts d'émoticônes (AjP-Ré et AjG-Bi), ne se retrouvent que dans le corpus constitué de SMS, alors que les substitutions et ajouts de diacritiques (SubP-Di et AjG-Di), comme les phénomènes complexes relatifs à la segmentation (SEGMENT) n'apparaissent quant à eux que dans les lettres de *Corpus 14*. Enfin, des divergences marquées apparaissent également concernant les phénomènes les plus employés dans chaque corpus : il s'agit de la réduction graphique affectant les lettres muettes (RedG-Mu) dans le cas des lettres manuscrites, et de la réduction graphique par utilisation d'abréviations (RedG-Ab) dans les SMS.

32 En conclusion, le tableau 1 établit un bilan chiffré de cette comparaison, en classant les phénomènes les plus employés dans chaque corpus en fonction de leur fréquence d'utilisation. Le même tableau, enrichi d'exemples, se trouve en annexe.

Tableau 1 : Bilan de la comparaison des phénomènes de variation les plus employés par corpus, classés par fréquence d'utilisation

CORPUS 14		88milSMS	
Type	Utilisation	Type	Utilisation
Réduction graphique de lettres muettes	20,72 %	Réduction graphique par abréviation	13,19 %
Segmentation	9,8 %	Substitution graphique - typographie et ponctuation	11,49 %
Complexe graphique	9,8 %	Complexe phonétisé	7,95 %
Substitution phonétisée partielle	8,38 %	Réduction phonétisée par sigles/ acronymes	7,17 %
Substitution graphique avec variation	7,79 %	Ajout - émoticônes	6,73 %
Ajout de lettres muettes	7,56 %	Ajout - répétition de caractères	5,02 %
Réduction graphique par agglutination/soudure	6,79 %	Réduction phonétisée par troncation	5,01 %
Substitution graphique - typographie et ponctuation	4,98 %	Réduction graphique par agglutination/ soudure	4,91 %

33 Le tableau 1 montre finalement que les phénomènes les plus employés divergent d'un médium à l'autre, à l'exception, nous l'avons vu *supra*, des suppressions de diacritiques et de la ponctuation. Ainsi, l'emploi de tels phénomènes peut-il avoir été influencé par les caractéristiques spécifiques des média qui véhiculent ces écrits ?

3.3. Discussion

34 L'ensemble des phénomènes de variation majoritairement employés dans les lettres de *Corpus 14* est constitué par des cas d'ajouts et de réductions de lettres muettes (respectivement AjG-Mu et RédG-Mu), ainsi que de substitutions phonétisées partielles (SubP-Par) et de substitutions graphiques avec variation (SubG-Va). Les phénomènes réunis dans ces catégories correspondent à des variations dans les désinences verbales, ainsi que dans les accords en genre et en nombre, et coïncident avec les principales « zones de fragilité » du système orthographique français relevées par Lucci & Millet (1994).

35 Premièrement, les réductions graphiques de lettres muettes sont les plus courantes, et concernent généralement les morphogrammes grammaticaux, comme l'exemple [1] placé ci-après en témoigne.

> [1] « les lettr**e** ne so**n** pas régulièr**e** car nous somm**e** dans les tranché**e** » (Laurent, 28.10.1914)

36 Ici, la marque du pluriel des substantifs *lettres* et *tranchées* et de l'adjectif *régulières* n'est pas réalisée à l'écrit, de même que les dernières lettres des formes verbales *sont* et *sommes* . Cependant, tous ces cas de réduction de lettres muettes n'ont aucune incidence sur l'oralisation du message, qui est tout à fait compréhensible à la lecture. Ainsi, les solutions graphiques adoptées par le scripteur répondent pleinement au but de rendre possible la communication avec son destinataire.

37 Un autre type de variation à dominante morphogrammique est également attesté dans le cas de nombreux ajouts de lettres muettes (un phénomène employé très rarement dans les SMS), comme dans les exemples [2] et [3].

> [2] « je les ai**t** reçu » (Félicien, 19.09.1914)
> [3] « Nous avons fini**e** la vendange » (Victoria, 10.10.1914)

38 Ainsi, on relève l'ajout d'un *t* à la flexion du verbe *avoir* dans la forme composée *ai reçu* (exemple [2]), ou bien l'ajout d'un *e,* marque du féminin, au participe passé *fini* (exemple [3]). Ce dernier cas semble indiquer que la règle grammaticale de l'accord du participe avec un substantif féminin est acquise (cf. *la vendange*) ; toutefois, elle est employée dans un contexte non prévu par les règles prescriptives. S'agit-il d'un cas d'hypercorrection, témoignant d'un effort d'adhésion à une norme orthographique dont on connait l'existence ? Ou est-il question d'une influence du nom féminin *vendange* sur le participe *finie* ? Il est difficile d'en décider.

39 La variation portant sur les morphogrammes grammaticaux concerne aussi les cas de substitutions graphiques avec variation (exemple [4]), et les cas de substitutions phonétisées partielles (exemple [5]).

> [4] « ne te fai**t** pas au moins du mauvais sang » (Marie, 20.09.1914)
> [5] « je t'**est** repondu tout de suite » (Victoria, 12.08.1914)

40 Dans l'exemple [4], la désinence *t* est substituée à la désinence *s* de la 2e personne du singulier de l'impératif présent du verbe *faire* ; dans l'exemple [5], la 1re personne de l'indicatif présent du verbe *avoir*, *ai*, est substituée par la 3e personne de l'indicatif présent du verbe *être*, *est*, pour la construction du passé composé. L'homophonie de ces formes verbales[12] favorise ainsi leur substitution ; cependant, comme dans tous les exemples précédents, ces cas de variation ne font nullement obstacle à la compréhension du message lors de son oralisation.

41 Deuxièmement, l'ensemble des autres phénomènes les plus employés dans les lettres de *Corpus 14* se rapporte au découpage de la chaîne graphique : de nombreux cas de segmentations non normées y sont attestés ; nous allons les décrire en adoptant la typologie de Cappeau & Roubaud (2018). Tout d'abord, ces segmentations non conformes à la norme peuvent aboutir au phénomène de sous-segmentation, qui équivaut à la soudure de plusieurs éléments autonomes, comme dans l'exemple [6].

> [6] « je **nai** pas le temp » (Marie, 03.09.1914)

42 L'adverbe de négation *ne* et le verbe *avoir* ne sont pas séparés par une apostrophe, comme la norme orthographique l'indique en cas d'élision d'une voyelle finale devant un mot commençant par une voyelle, mais sont ici agglutinés, en formant ainsi une seule unité, *nai*. Le même phénomène se produit dans le cas de l'exemple [7].

[7] « quoi qu'il te dise ne **tefreye** pas » (Pierre, 15.10.1914)

43 Dans [7], le pronom personnel objet *te* élidé est soudé à la forme verbale qui le suit, ce qui aboutit à la forme *tefreye* à la place de la réalisation normée *t'effraye*. Ensuite, le phénomène inverse, celui de la sur-segmentation, est également attesté : un même mot est découpé en plusieurs éléments, comme dans l'exemple [8].

[8] « **a laur** on poura le soir parler un peu du temp passer » (Laurent, 09.10.14)

44 Ici, l'adverbe *alors* est sur-segmenté en deux unités, *a* et *laur* ; le découpage semble être calqué sur une division syllabique : sa réalisation écrite comporte un premier élément existant en tant qu'unité standard de la langue (préposition *à*), à la différence du deuxième, qui semble servir à compléter la suite, conformément au schéma décrit par Cappeau & Roubaud (2018). Enfin, un troisième phénomène concerne la marque surnuméraire à l'écrit d'une consonne de liaison (exemple [9]).

[9] « que j'ai reçu avec plaisir **en na prenand** » (Laurent, 28.09.1914)

45 Dans l'exemple [9], la liaison de la consonne *n* est marquée à l'écrit par l'ajout d'un *n* devant la voyelle suivante (*e**n** **n**a)*. Le participe présent *apprenant* est en outre sur-segmenté (*n**a** **prenand***) : comme cet extrait l'illustre, ces trois phénomènes peuvent se cumuler. Ainsi, dans l'exemple [10], nous retrouvons les trois.

[10] « **[nous] sana von** un peut plus aprésent » (Laurent, 01.10.1914)

46 La suite *nous en avons* présente la marque surnuméraire de la consonne de liaison *s (**s**ana)*, la sur-segmentation du verbe *avoir* en deux éléments (***a von***), ainsi que la soudure du premier élément *a* à la préposition qui le précède (***ana** von*). Comme nous l'avons observé pour les exemples de variation morphogrammique précédents (exemples [1]-[5]), ces cas de segmentations non normées de la chaîne graphique (exemples [6]-[10]) n'empêchent pas la compréhension du texte lorsqu'il est oralisé.

47 De fait, les phénomènes de variation les plus employés dans notre sélection de lettres de *Corpus 14* semblent indiquer des solutions graphiques adoptées par ces scripteurs lors de la pratique de l'écrit, dont ils n'avaient pas l'habitude avant le commencement de la guerre. La dimension phonogrammique se retrouve dans tous les types de variation orthographique ici observés : elle semble répondre au besoin des scripteurs de trouver avant tout une façon de réaliser les correspondances entre phonèmes et graphèmes. De là semblent découler les autres dimensions de variation (morphogrammique, de segmentation, etc.), lorsque cette réalisation ne correspond pas à la forme graphique normée : ainsi, dans ces exemples, les variations orthographiques ne produisent pas d'altération phonique par rapport au standard. Il en résulte que la compréhension du message n'est pas entravée (Pellat, 2015 ; Surcouf, 2018a, 2018b), assurant ainsi les finalités de ces écrits : donner des nouvelles, garder le contact avec ses proches et,

notamment pour les soldats, donner un signe de vie (Rousseau, 2014 : 56-63 ; Caffarena, 2005 : 70, 2007 : 649 ; Gibelli, 2014 : 14 ; Gomila, 2015).

48 Du côté des SMS, deux visées semblent être privilégiées par les scripteurs d'après les phénomènes de variation les plus utilisés de notre sélection issue du corpus *88milSMS* : la diminution réfléchie du nombre de caractères, pour notamment réduire le nombre de manipulations à effectuer et donc limiter les contraintes rédactionnelles[13], et une mise en valeur de l'expressivité par divers procédés graphiques.

49 Les réductions graphiques par abréviation (exemple [11]), ainsi que les cas de réduction par sigle/acronyme (exemple [12]), viennent tout d'abord réduire les contraintes rédactionnelles tout en limitant le nombre de caractères.

[11] « Je peux manger **ac** toi a la cantine ? » (SMS n° 3 000)

50 L'exemple [11] présente un cas simple de réduction graphique par abréviation, *ac* venant ici remplacer *avec*. Ce genre d'abréviations permet de réduire le nombre de caractères requis, mais n'est pas destiné à être lu tel qu'il apparait.

[12] « **Lol** pour toi c'est legerement plus long **mdr** » (SMS n° 8 315)

51 L'exemple [12] fait apparaitre deux cas distincts de réductions phonétisées par sigle/acronyme. Le premier, *Lol*, est un acronyme très courant dans les SMS et la CéMO en général, formé à partir de l'expression anglaise *laughing out loud*, et utilisé pour signifier un rire appuyé. Bien qu'issu de la langue anglaise, la généralisation de son emploi l'a conduit à constituer une entrée du Petit Robert depuis 2014 ; dans notre sélection, il apparait en effet dans 3,96 % des messages. La seconde occurrence, *mdr*, est un sigle, signifiant *mort de rire*. Il correspond au pendant francophone de *lol*, mais son utilisation semble moins répandue : dans notre corpus, il apparait dans 2,07 % des messages. Ces types particuliers de réductions phonétisées très populaires participent directement du caractère méta-langagier de l'écriture SMS (Cougnon, 2015), donnant au destinataire un aperçu de l'état émotionnel du scripteur. Ils revêtent généralement une fonction démarcative, remplaçant certains signes de ponctuation, ou se placent « à la fin des « blocs paragraphes » pour accentuer la dimension interactive du texte » (Sperlinga Gerner, 2015 : 126).

52 Par suite, l'emploi de substitutions graphiques en lien avec la typographie et la ponctuation (exemples [13] et [14]), ainsi que de réductions graphiques par agglutination (exemple [15]), semble de prime abord être un choix effectué par les scripteurs pour faciliter la rédaction.

[13] « Qu'es**t** **c**e que tu fais ? » (SMS n° 60 039)

53 Dans l'exemple [13], le tiret « - » marquant l'inversion entre *est* et *ce* de l'expression « est-ce que » a été remplacé par un espace, limitant ainsi les manipulations du médium, parfois complexes en fonction du modèle de téléphone ou de l'interface.

54 Néanmoins, la substitution de minuscules en majuscules (d'ordre typographique) peut également venir modifier le ton d'un message (exemple [14]).

[14] « Ehh la j'en peux plus mdrr tu sais je sais plus si je te l'avais dit mais on a enfin acheter la machine a laver aujourd'hui et donc je dis : elle est belle ? Quel est blanche ? elle marche bien ? Et la mon pere me dit : elle est verte ! ! Nan mais **VERTE** quoi ! **MDRRRR** » (SMS n° 20 475)

55 L'exemple [14] présente deux emphases qui sont créées par le scripteur par l'intermédiaire de substitutions minuscules/majuscules : *VERTE* et *MDRRRR*. *MDRRRR* correspond à un phénomène complexe, puisqu'il s'agit à la fois d'une réduction par sigle, d'une répétition de caractères, par la répétition du « R », et d'une substitution d'ordre typographique. Ce phénomène complexe permet ainsi d'accentuer l'emphase, et de créer une nuance entre un *mdr*, correspondant à un rire simple, et un *MDRRRR*, correspondant à un rire long, fort, et très appuyé. La substitution correspondant à *VERTE* est quant à elle beaucoup plus explicite en matière de modification du ton du message, puisque cette occurrence prend place à la suite de son homologue non accentué : *verte*.

> [15] « **Jp**asse la soirée avec <PRE_4> ce soir.et **jp**ars sur toulouse demain soir.
> **jv**ais voir <SUR_5> samedi soir **da**illeurs. » (SMS n° 6 356)

56 Dans un même ordre d'idée, l'exemple [15] présente différents cas d'agglutinations, mis en place pour éviter d'avoir à rédiger les apostrophes. En outre, l'agglutination des trois pronoms personnels *je* avec le verbe auquel ils sont associés (*passe*, *pars* et *vais*), dans des contextes où il n'y a pas d'élision normée, pourrait relever, à l'instar des modifications de tons, d'une « oralité simulée » (Anis, 1999 : 154-155). En effet, dans ce message particulier, *jpasse* semble par exemple destiné à être oralisé [ʒpas :], et non [ʒəpasə] : l'élision graphique des *e* instables se retrouve lors de l'oralisation du message.

57 L'expressivité du message semble être une visée mise en avant dans certains SMS, dans lesquels figurent de façon récurrente les derniers phénomènes les plus employés de notre sélection : l'ajout d'émoticônes, smileys et autres binettes (exemple [16]), ainsi que la répétition de caractères (exemple [17]). Bien souvent, ces différents types de phénomènes apparaissent au sein d'un même message, et se combinent parfois avec des modifications typographiques, pour former un complexe phonétisé (exemple [18]).

> [16] « C'est bien aussi, et heureusement, sinon maman s'ennuierait ! (et viendrait me voir ! ! ! **:s**) » (SMS n° 19 003)
> [17] « Okiiiii ca maaaaaarche :D » (SMS n° 44 540)
> [18] « VACAAAAANCES:D:D:D:D » (SMS n° 46 534)

58 La binette finale apparaissant dans l'exemple [16], *:s* , a été mise en place par le scripteur dans le but de créer un sous-entendu, transmettant au destinataire la gêne que la situation décrite engendrerait. Par ce biais, il modalise son discours tout en laissant le destinataire imaginer son expression faciale à un instant t. Ainsi, ces phénomènes illustratifs « permettent de rendre plus accessibles les sentiments et les émotions de l'auteur du message » (Marcoccia & Gauducheau, 2007 : 6), et participent directement à la simulation d'une coprésence en apportant une nouvelle dimension d'expressivité à la graphie.

59 Dans l'exemple [17], par l'utilisation de deux répétitions de caractères, *Okiiiii* et *maaaaaarche*, le scripteur transmet à son destinataire la prosodie qu'il souhaiterait le voir mettre en place lors de l'oralisation cognitive du message, en lui signifiant des modifications du rythme par l'allongement des voyelles *i* de *oki* et *a* de *marche* : [oki : : :] et [ma : : :ʁʃə]. Ce caractère très ludique est par ailleurs accentué par le fait que *Okiiiii* est en réalité un phénomène complexe produit par l'ajout d'un caractère oralisé répété (*i* étant ajouté puis répété cinq fois à la suite de *ok*), ainsi que par l'ajout d'un smiley, *:D* , à la fin du message.

60 Enfin, l'exemple [18] fait apparaitre un nouveau phénomène phonétisé complexe, *VACAAAAANCES*, la variante normée *vacances* ayant été modifiée à la fois par une répétition de caractères demandant au destinataire d'allonger la syllabe « an » ([ã : : :]), ainsi que par une substitution minuscule/majuscule, demandant d'apporter une intonation forte.

61 Notre analyse des phénomènes de variation les plus fréquents relevés dans notre sélection de SMS renvoie finalement aux « caractéristiques formelles » de l'*écrit-SMS* définies par Anis (2000) : les effets d'oralité, l'abréviation et l'iconicité. D'abord, on retrouve le besoin de réduire le nombre de caractères et ainsi les contraintes rédactionnelles relatives à l'interface d'un téléphone portable, mais également le besoin de transmettre, par des procédés graphiques dont nous avons pu avoir un aperçu, des « directives » d'expression, à l'instar de didascalies (Mourlhon-Dallies & Colin, 1999).

62 Cette distinction marquée par rapport aux résultats obtenus dans le cadre de notre sélection issue de *Corpus 14*, où les phénomènes utilisés semblent indiquer des solutions graphiques adoptées par des scripteurs peu expérimentés lorsqu'ils sont confrontés à l'écriture, peut avoir été favorisée par une des caractéristiques des média comparés qui s'est révélée souvent divergente : la quasi-synchronie des échanges par SMS. En effet, les évolutions progressives des supports au cours du XXe siècle, rendant possible des échanges écrits quasi-instantanés entre deux locuteurs distants, ont pu influencer la diffusion de procédés graphiques accentuant la « coprésence », et permettant parfois la simulation de véritables conversations en transmettant des informations para-verbales, voire non-verbales, donnant alors à l'écrit-SMS un caractère multicanal[14] (Cosnier & Vaysse, 1997), malgré sa nature exclusivement graphique. Ainsi, une des caractéristiques des média comparés semble bien avoir pu influencer certaines spécificités des graphies qu'ils véhiculent.

Conclusion

63 Cette étude propose de comparer deux média à l'origine de la diffusion d'écrits définis comme *hors-normes*, constitués de formes graphiques s'éloignant de la norme standard du français. Nous nous sommes tout d'abord interrogés sur leur nature : quelles caractéristiques séparent ou rapprochent ces deux média ?

64 À l'aide d'une classification à facettes proposée par Herring (2007), nous avons pu déterminer que la correspondance de guerre partage de nombreux points communs avec la communication par SMS. En effet, plusieurs critères convergent, comme notamment le cadre privé de communication interpersonnelle, et l'emploi régulier de formes graphiques renvoyant à la proximité communicative définie par Koch & Oesterreicher (2001), alors que peu de critères les séparent, ceux-ci étant en lien avec les possibilités qu'offrent aujourd'hui les formes de communications électroniques.

65 Nous avons ainsi exploré les caractéristiques de ces média dans l'optique de déterminer si ces supports d'écriture peuvent potentiellement exercer une influence sur les écrits *hors-normes* dont ils permettent la diffusion. Nous avons donc cherché à répondre à une seconde question : certaines tendances concernant la variation graphique pourraient-elles être favorisées ou limitées par les caractéristiques spécifiques du médium employé ?

66 En comparant les graphies de deux média bien distincts, nous avons pu observer que dans le cadre des échanges épistolaires de guerre, les variations les plus courantes semblent

témoigner des solutions graphiques mises en place par les scripteurs lors de la pratique de l'écrit pour permettre la communication avec le destinataire, tandis que les échanges par SMS font parfois apparaitre des formes graphiques qui accentuent la simulation de la coprésence en conférant à l'*écrit-SMS* un caractère multicanal (Cosnier & Vaysse, 1997), en parallèle de la diminution, voire disparition, des tendances relevées précédemment. En considérant les évolutions séparant les supports de communication analysés, comme notamment le développement du caractère quasi-synchrone de la communication par SMS, « réduisant la distance » qui sépare deux scripteurs, les spécificités des graphies analysées semblent bien avoir pu être favorisées par une des caractéristiques du médium employé. Un angle sociolinguistique pourrait certainement apporter un éclairage nouveau sur ces premières comparaisons, en s'intéressant notamment, à la suite de Lucci & Millet (1994), au « capital scolaire » des auteurs des SMS.

67 Dans cette étude préliminaire, nous nous sommes concentrés sur le code graphique, mais tout au long du travail d'analyse, nous avons également pu observer des différences marquées sur le plan syntaxique. La formulation d'interrogations en est un exemple : particulièrement récurrente dans les SMS (dans près de 25 % des messages), elle est assez rare dans les échanges épistolaires. Notre analyse diachronique initiale pourra donc être considérablement enrichie à court et moyen termes, et pourra également être ouverte à plusieurs autres média, tels que des cartes postales contemporaines ou des mails, qu'il serait intéressant de confronter à *Corpus 14.*

BIBLIOGRAPHIE

ANDRÉ F., 2017, *Pratiques scripturales et écriture SMS. Analyse linguistique d'un corpus de langue française*, Thèse de doctorat en Langue française, Université Paris-Sorbonne.

ANIS J., 2000, « L'écrit des conversations écrites de l'Internet », *Le français aujourd'hui* 129, 59-69.

ANIS J., 1999, *Texte et Ordinateur. L'écriture réinventée ?,* Bruxelles, De Boeck Université.

ANIS J., DE FORNEL M. & FRAENKEL B. (organisateurs), 2004, « La communication électronique : Approches linguistiques et anthropologiques », Colloque international, EHESS, Paris, 5-6 février 2004.

BEGUELIN M.-J., 2000, *De la phrase aux énoncés : grammaire scolaire et descriptions linguistiques*, Bruxelles, De Boeck Duculot.

BRANCA-ROSOFF S. & SCHNEIDER N., 1994, *L'écriture des citoyens : une analyse linguistique de l'écriture des peu-lettrés pendant la période révolutionnaire*, Paris, Klincksieck.

BRUN A., 1931, *Le français de Marseille. Étude de parler régional*, Marseille, Institut historique de Provence.

CAFFARENA F., 2007, "Le scritture dei soldati semplici", in A. Gibellli (éd.), *La prima guerra mondiale*, Torino, Einaudi, 645-660 [édition italienne d'Adoin-Rouzeau S., Becker J.-J. (éds), 2004, *Encyclopédie de la Grande Guerre 1914-1918 : histoire et culture*, Paris, Éditions Bayard].

CAFFARENA F., 2005, *Lettere dalla Grande Guerra. Scritture del quotidiano, monumenti della memoria, fonti per la storia. Il caso italiano*, Milano, Unicopli.

CAPPEAU P. & ROUBAUD M.-N., 2018, *Regards linguistiques sur les textes d'élèves (de 5 à 12 ans)*, Clermont-Ferrand, Presses universitaires Blaise-Pascal.

COSNIER J. & VAYSSE J., 1997, « Sémiotique des gestes communicatifs », *Nouveaux Actes Sémiotiques* 52, 7-28.

COUGNON L.-A., 2015, *Langage et Sms*, Louvain-la-Neuve, Presses Universitaires de Louvain.

CRYSTAL D., 2008, *Txtng: The gr8 db8*, Oxford, Oxford University Press.

DURAND J. & EYCHENNE J., 2016, « Le français méridional », in S. Detey, I. Racine, Y. Kawaguchi & J. Eychenne (éds), *La prononciation du français dans le monde. Du natif à l'apprenant*, Paris, CLE International, 34-38.

FAIRON C., KLEIN J. R. & PAUMIER S., 2006, *SMS pour la science. Corpus de 30.000 SMS et logiciel de consultation*, Louvain-la-Neuve, Presses Universitaires de Louvain.

FREI H., de 1929. *La grammaire des fautes*. Réédition Presses Universitaires de Rennes, Rennes, 2011.

GIBELLI A., 2014, *La guerra grande : storie di gente comune*, Bari, Laterza.

GOMILA C., 2015, « Paroles de Poilus : une caresse sous la mitraille », in A. Steukardt (éd.), *Entre village et tranchées. L'écriture de Poilus ordinaires*, Uzès, Inclinaison, 157-167.

HASPELMATH M., 2011, « The indeterminacy of word segmentation and the nature of morphology and syntax », *Folia Linguistica* 45/1, 31-80. DOI: 10.1515/flin.2011.002

HERRING S.C., 2007, « A faceted classification scheme for computer-mediated discourse », *Language@Internet* 4.

KOCH P. & OESTERREICHER W., 2001, « Gesprochene Sprache und geschriebene Sprache. Langage parlé et langage écrit », in G. Holtus, M. Metzeltin & C. Schmitt (eds), *Lexikon der romanistischen Linguistik* 1-2, 584-627.

LIÉNARD F., 2008, « Analyse linguistique et sociopragmatique de l'écriture électronique. Le cas du SMS tchaté », in J. Gerbault (éd.), *La langue du cyberespace : de la diversité aux normes*, Paris, l'Harmattan, 265-278.

LOPEZ C., ROCHE M. & PANCKHURST R., 2015, « Classification des items inconnus de 88milSMS : aide à l'identification automatique de la créativité scripturale », *Travaux neuchâtelois de linguistique* 63, 71-86.

LUCCI V. & MILLET A. (éds), 1994, *L'orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*, Paris, Honoré Champion.

MARCOCCIA M. & GAUDUCHEAU N., 2007, « L'analyse du rôle des smileys en production et en réception : un retour sur la question de l'oralité des écrits numériques », *Glottopol* [En ligne] 10, 39-55, consulté le 25 novembre 2019. URL : http://glottopol.univ-rouen.fr/telecharger/numero_10/gpl10_03marcoccia.pdf

MOURLHON-DALLIES F. & COLIN J.-Y., 1999, « Des didascalies sur l'Internet ? », in J. Anis (éd.), *Internet, communication et langue française*, Paris, Hermès.

PANCKHURST R., 2009, « Short Message Service (SMS) : typologie et problématiques futures », in T. Arnavielle (éd.), *Polyphonies, pour Michelle Lanvin*, Université Paul-Valéry Montpellier 3, 32-52.

PANCKHURST R., DÉTRIE C., LOPEZ C., MOÏSE C., ROCHE M. & VERINE B., 2014, « 88milSMS. A corpus of authentic text messages in French », produit par l'Université Paul-Valéry Montpellier 3 et le CNRS, en collaboration avec l'Université catholique de Louvain, financé grâce au soutien de la MSH-M et du Ministère de la Culture (Délégation générale à la langue française et aux langues de France) et avec la participation de Praxiling, Lirmm, Lidilem, Tetis, Viseo. ISLRN: 024-713-187-947-8, http://88milsms.huma-num.fr/corpus.html

PELLAT J.-C., 2015, « Les graphies des Poilus, loin des canons orthographiques », in A. Steukardt (éd.), *Entre village et tranchées. L'écriture de Poilus ordinaires*, Uzès, Inclinaison, 67-77.

PRAXILING, 2014, *Corpus 14*. https://hdl.handle.net/11403/corpus14/v1

ROUSSEAU F., [1999] 2014, *La guerre censurée. Une histoire des combattants européens de 14-18*, Paris, Seuil.

SPERLINGA GERNER M.-M., 2015, *Variations graphiques des textes des forums sur Internet,* Thèse de doctorat en Sciences du langage, Université de Strasbourg.

STARK E., 2015, « 'De l'oral dans l'écrit' ? – Le profil variationnel des SMS (textos) et leur valeur pour la recherche linguistique », in K. A. Jeppesen Kragh & J. J. Lindschouw (éds), *Les variations diasystématiques et leurs interdépendances dans les langues romanes. Actes du Colloque DIA II à Copenhague (19-21 nov. 2012)*, Strasbourg, Éditions de Linguistique et de Philologie ELiPhi, 395-405.

STARK E., 2012, « Negation marking in French text messages », *Lingvisticæ Investigationes* 35-2, 341-366.

STEUCKARDT A., à paraître, « Hors-normes et norme dans des écrits peu-lettrés anciens », Actes du colloque *Analyses des discours hors-normes : approches, concepts et méthodes*, Sherbrooke, Éditions de l'Université de Sherbrooke ÉDUS.

STEUCKARDT A., 2015a, « Introduction », in A. Steuckardt (éd.), *Entre village et tranchées. L'écriture de Poilus ordinaires*, Uzès, Inclinaison, 9-17.

STEUCKARDT A., 2015b, « Sans point ni virgule », in A. Steuckardt (éd.), *Entre village et tranchées. L'écriture de Poilus ordinaires*, Uzès, Inclinaison, 91-100.

SURCOUF C., 2018a, « « tu me di que je ne connai pas lecri ture ébien çi chaque foi que tu écri je le connai » - Analyse des graphies des formes verbales dans la correspondance des Poilu du Corpus14 », in F. Neveu, B. Harmegnies, L. Hriba & S. Prévost (éds), *6e Congrès Mondial de Linguistique Française*, SHS Web of Conferences 46, 06001. DOI : 10.1051/shsconf/20184606001

SURCOUF C., 2018b, « L'efficacité orthographique des « peu-lettrés » : une analyse des graphies des Poilus du Corpus 14 ». Communication au colloque *La Grande Guerre des gens « ordinaires ». Correspondances, récits, témoignages*, organisé par A. Steuckardt, C. Gomila & C. Wionet, 14-16 juin 2018, Montpellier.

WARNANT L., 2006, *Orthographe et prononciation en français*, Bruxelles, Éditions Duculot.

ANNEXES

Tableau 1b : Bilan de la comparaison des phénomènes de variation les plus employés par corpus, classés par fréquence d'utilisation

CORPUS 14			88milSMS		

Type	Exemple	Utilisation	Type	Exemple	Utilisation
Réduction graphique de lettres muettes	« mai comme nous étion »	20,72 %	Réduction graphique par abréviation	« tt de suite ds ton sac »	13,19 %
Segmentation	« en voyer la dresse »	9,8 %	Substitution graphique – typographie et ponctuation	« J essaie de trouver »	11,49 %
Complexe graphique	« si pré de la mort »	9,8 %	Complexe phonétisé	« hier jt ecris »	7,95 %
Substitution phonétisée partielle	« donné lui bien le bonjour »	8,38 %	Réduction phonétisée par sigles/ acronymes	« Bon we »	7,17 %
Substitution graphique avec variation	« je te disait »	7,79 %	Ajout - émoticônes	« Encore au taff. : (»	6,73 %
Ajout de lettres muettes	« saint et sauf »	7,56 %	Ajout - répétition de caractères	« Tu verras bien demain..... ... »	5,02 %
Réduction graphique par agglutination/ soudure	« on ne ma pas manqué »	6,79 %	Réduction phonétisée par troncation	« ton ordi port »	5,01 %
Substitution graphique - typographie et ponctuation	« console toi »	4,98 %	Réduction graphique par agglutination/ soudure	« Jai vu ;) »	4,91 %

NOTES

1. Cependant, malgré les variations attestées, la norme reste dominante dans les deux types d'écrits (Steuckardt, *à paraître* ; Crystal, 2008 ; Stark, 2012)

2. Le terme *peu lettré* a été défini par Branca-Rosoff et Schneider en 1994 : il désigne le « groupe de ceux qui emploient une langue non conforme », et qui se situent donc entre les groupes polaires des *lettrés*, qui maitrisent la langue correctement, et des *illettrés*, qui au contraire ne savent pas écrire.

3. Dans le texte source : "the goal of the CMD scheme is to articulate aspects of context – both technical and social – that potentially influence discourse usage in CMC environments, and thereby to bring them to the conscious attention of the researcher".

4. Le terme *ordinateur* prend ici un sens extrêmement large, renvoyant à tout matériel électronique servant de support de communication (ordinateur personnel, tablette, téléphone portable, montre connectée, etc.).

5. Contrairement aux cadres d'utilisation comme le chat ou la messagerie instantanée, dans lesquels le destinataire peut parfois être informé en temps réel qu'un message qui lui est destiné est en train d'être rédigé.

6. Cependant, le nombre de destinataires peut varier dans les deux média, les destinataires d'une lettre pouvant être multipliés, comme dans cet exemple : « Bien cher beau Parents, Epouse et fils » (Laurent 15.09.17), de même que dans les SMS.

7. Dans cet exemple, l'apostrophe n'est pas supprimée, mais est remplacée par un espace. Bien que faisant apparaitre un « blanc » dans le code graphique, l'espace, en matière de codage informatique, reste un caractère. Il s'agit donc bien ici d'un cas de substitution d'un caractère par un autre.

8. Bien que les symboles comme « + » soient oralisés et permettent parfois de former des rébus simples, la typologie d'origine considère que ce type de modification reste principalement d'ordre graphique. Bien que discutable, ce choix permet de distinguer l'utilisation de symboles d'autres cas correspondant à des substitutions phonétisées entières et partielles.

9. Exemple : « je c ke c chian de devoir se connecter mé en fait on é noté ossi sur le nombre d'échange et ce qu'on dira dc si vous avez un ordi à disposition ca seré cool d'envoyer des infos sur d sites par ex pr le dossier ou se genre de truc pr le site je pense ke ca devré allé vite je vou envéré un message ds les jours suivants sur chaque page k on peu faire et chacun poura dire ce ki lui plé le plus de faire moi je m en fou voila. » (Panckhurst, 2009 : 35)

10. Dans les SMS, lorsque les diacritiques sont affectés par des modifications par rapport à la norme, ils sont généralement supprimés, selon un potentiel principe de simplicité réduisant les contraintes rédactionnelles. Bien que des cas de substitution puissent exister, nous n'avons relevé aucune occurrence dans notre sélection issue de *88milSMS*.

11. Si dans les SMS la segmentation des mots peut être modifiée, notamment dans le cadre d'agglutinations, celle-ci demeure relativement stable (Stark, 2015), et aucun cas de segmentation comparable à ceux observés dans *Corpus 14* n'a pu être relevé.

12. Bien que la prononciation du français parisien cultivé, adoptée comme norme par Warnant 2006, indique les prononciations /ɛ/ pour *est* (parfois /e/), /e/ pour *ai*, dans le « français méridional », l'opposition entre voyelles mi-fermées et mi-ouvertes est neutralisée (Durand & Eychenne, 2016 : 34). La fermeture progressive de l'*e* final (/ɛ/ > /e/) est signalée pour le français en général par Frei (1929 : 93) et, pour le français de Provence, par Brun (1931 : 32).

13. La simple réduction du nombre de caractères était une préoccupation des scripteurs au début des années 2000 (Fairon *et al.*, 2006), mais les scripteurs de notre corpus collecté en 2011 y semblent moins attachés, probablement suite à l'apparition des forfaits « sms illimités », à partir de 2007 (Panckhurst *et al.*, 2014).

14. Selon Cosnier & Vaysse (1997), le caractère multicanal se définit par un mélange de verbal, de non verbal et de para-verbal (vocal), en proportions variables, la composante gestuelle étant considérée comme consubstantielle avec le verbal.

RÉSUMÉS

Cet article se propose d'interroger le rôle du *médium* dans un contexte spécifique de production discursive, celui de la communication privée, sur des écrits *hors-normes.* Le corpus de cette étude est constitué de deux types d'écrits : d'une part, des lettres manuscrites rédigées au cours de la Grande Guerre par des soldats et des femmes peu scolarisés, recueillies dans les fonds du projet *Corpus 14*, et d'autre part, d'authentiques SMS contemporains, collectés dans le cadre du projet international *sms4science*. L'étude s'articule en deux parties : dans un premier temps, nous comparons les deux média en fonction d'une classification à facettes ; dans un deuxième temps, nous comparons l'usage qui est fait de différents phénomènes de variation orthographique. Les résultats de ces analyses indiquent que l'on se situe, dans les deux cas, dans un cadre privé de communication interpersonnelle, s'établissant par l'intermédiaire d'un code écrit qui offre la possibilité de s'approprier les normes graphiques standard. En outre, certaines spécificités du code graphique relevées semblent avoir été favorisées par une des caractéristiques des deux média qui s'est révélée souvent divergente : la quasi-synchronie des échanges par SMS.

This paper examines the role of the medium in a specific context of language production, that is private communication in non-standard written texts. The corpus for this study consists of two types of texts: letters written during the Great War by soldiers and women with little formal education, which were taken from the collection of the project *Corpus 14*, and a set of genuine, contemporary SMSs collected within the international project *sms4science*. The study is divided into two parts: we first compare the two media according to a faceted classification, and we then compare how different phenomena of spelling variation are used. The results of these analyses indicate that in both cases we are dealing with interpersonal communications in a private setting that take place via a written code enabling the writers to acquire the standard spelling norms. Furthermore, some of the specificities of the written code that have been noted seem to be favoured by one characteristic that has turned out to be often distinct between the two media, namely the quasi-synchrony of SMS communication.

INDEX

Keywords : medium, letters, SMSs, spelling variation
Mots-clés : médium, lettres, SMS, variation orthographique

AUTEURS

FRÉDÉRIC ANDRÉ

Université Paris-Sorbonne, STIH EA 4509

BEATRICE DAL BO

Université Paul-Valéry Montpellier, Praxiling UMR 5267