

Travaux de Linguistique Romane

---

Les variations diasystématiques  
et leurs interdépendances dans les langues romanes

ELIPHI

*Sociolinguistique, dialectologie, variation*

Collection dirigée par Emili Casanova (València), Jean-Paul  
Chauveau (Nancy, ATILF), Hans Goebel (Salzbourg)

Kirsten Jeppesen Kragh  
Jan Lindschouw (éds.)

---

Les variations diasystématiques  
et leurs interdépendances dans les langues romanes

Actes du Colloque DIA II à Copenhague (19-21 nov. 2012)

**ELIPHII**

EDITIONS DE LINGUISTIQUE ET DE PHILOGIE

Ouvrage publié avec l'appui de l'Académie Royale des Sciences et  
Belles-lettres de Danemark et du Lektor Knud Henders Legatfond

La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'article 41, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective », et d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants-droit ou ayants-cause, est illicite » (alinéa 1<sup>er</sup> de l'article 40).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles 425 et suivants du Code Pénal.

ISBN 978-2-37276-002-7

EAN 9782372760027

© Éditions de linguistique et de philologie, Strasbourg 2015.

## Table des matières

Kirsten Jeppesen Kragh, Jan Lindschouw: Introduction: les types de variation diasystématique et leurs interdépendances .....	IX
<i>Variation diachronique</i>	
Paolo Greco: Sull' <i>Accusativus cum Participio</i> in latino: funzioni evidenziali e sviluppi diacronici .....	3
Céline Guillot, Serge Heiden, Alexei Lavrentiev, Bénédicte Pincemin: L'oral représenté dans un corpus de français médiéval (9 <sup>e</sup> -15 <sup>e</sup> ): approche contrastive et outillée de la variation diasystémique .....	15
Pascale Hadermann: <i>D'autant (plus) que</i> et la corrélation en français .....	29
Sylviane Lazard: Premiers indices d'une évolution de la 'scripta' milanaise (XIII <sup>e</sup> -XIV <sup>e</sup> ): la langue de Bonvesin et de Barsegapè .....	45
Sabine Lehmann: La violence verbale dans l'histoire de la langue française. Une perspective diachronique (de l'ancien français au français du XVI <sup>e</sup> s.)	61
Vanessa Meireles: Le sandhi vocalique externe en portugais archaïque et en portugais moderne .....	73
<i>Variation diatopique</i>	
Vanderci de Andrade Aguilera, Dircel Aparecida Kailer: Apagamento do /R/ em coda silábica no Sul do Brasil: um estudo preliminar .....	89
Myriam Bergeron-Maguire: La distance communicative dans les écrits de scripteurs peu lettrés et la variation de <i>septembre</i> en français .....	103
Emilia Calaresu: L'avverbio già da operatore temporale aspettuale a operatore modale di asserzione. Usi preverbal di Sardegna (sardo e italiano regionale) vs. usi olofrastici in italiano standard .....	113
Sascha Diwersy, Sylvain Loiseau: La différenciation du français dans l'espace francophone: l'apport des statistiques lexicales .....	129
Josefa Dorta Luis, José Antonio Martín Gómez, Chaxiraxi Díaz Cabrera: Continuidad prosódica en habla experimental y espontánea de Canarias y Cuba: variación y rango tonal en las interrogativas no pronominales .....	145

Aparecida Negri Isquerdo: Léxico dialetal no português do Brasil: espacialidade, fixação e disseminação .....	161
Carita Klippi, Anne-Laure Kiviniemi: L'écriture de deux frères d'armes, déshérités du français — Une caricature de la langue nationale ? .....	175
Sandra Maria Oliveira Marques: O gerúndio sob a ótica de um contato dialetal .....	191
Jean-Pierre Montreuil: Allophonie, fusion et opacité: /h/ primaire et /h/ secondaire dans les dialectes de l'Ouest de la France .....	203
Francisco Moreno-Fernández: La percepción global de la similitud entre variedades de la lengua española .....	217
Suzana Alice Cardoso, Jacyra Andrade Mota: Os estudos diatópicos no Brasil ..	239
Marcela Moura Torres Paim: A variação semântico-lexical e a identidade social de faixa etária nas capitais do Brasil .....	253
Jean Sibille: La variation inter-individuelle et intra-individuelle dans le parler occitan de Sénailac-Lauzès (Lot, France) .....	265
 <i>Variation diastratique</i>	
Silvana Araujo: A concordância verbal na fala culta e popular do português brasileiro .....	281
Shana Poplack: Norme prescriptive, norme communautaire et variation diaphasique .....	293
 <i>Variation diaphasique</i>	
Carla Bazzanella: Prospettiva pragmatica e complessità della dimensione diafasica .....	323
Françoise Gadet: Le style et les corpus : réflexions à partir d'un corpus de la région parisienne .....	339
Olga Ivanova: De la narración a la descripción: la variación sintáctica en español L2 .....	353
Ildiko Van Tricht: La variation dans la terminologie de l'anatomie de l'œil en moyen français .....	365
 <i>Variation diamésique</i>	
Eva Havu, Michel Pierrard: Observations sur la variation diasystémique du participe passé adjoint en position polaire à l'oral et à l'écrit .....	379
Elisabeth Stark: 'De l'oral dans l'écrit'? – Le statut variationnel des SMS (textos) et leur valeur pour la recherche linguistique .....	395

*Interdépendances entre plusieurs dimensions diasystématiques*

Luisa Amenta: Il diasistema nelle scritture dei semicolti. Analisi degli inediti di Bordonaro .....	409
Marta Andronache: La variation diasystématique protomane. Réflexions à partir de l'expérience du DÉRom .....	421
Gaetano Berruto: Intrecci delle dimensioni di variazione fra variabilità individuale e architettura della lingua .....	431
Paul Cappeau, Catherine Schnedecker: ( <i>Les/des</i> ) gens vs ( <i>les/des</i> ) personnes: évolution diachronique et comparaison oral / écrit. Des SN en voie de pronominalisation ? .....	449
Massimo Cerruti, Elena Maria Pandolfi: 'Standard' coesistenti nell'italiano contemporaneo: i casi di <i>solo più e non più</i> +infinito .....	465
Christine Cuet: Les variations diasystémiques dans le Dictionnaire de Richelet (1680): Compilations des Remarqueurs et enquêtes de terrain ...	479
Josane Moreira de Oliveira, Odete Pereira da Silva Menon: L'expression du futur verbal en portugais brésilien: un cas de variation diasystématique .....	493
Hans Kronning: El condicional epistémico «de atribución» en francés, italiano y español: aspectos diafásicos, diatópicos y diacrónicos .....	507
Marco Mazzoleni: Il diasistema italo-romanzo dei periodi ipotetici .....	521
Antonio Hidalgo Navarro: Sexo, registro y diasistema: usos prosódicos diferenciados para la expresión de (des)cortesía en la conversación .....	533
Concepción Martínez Pasamar: Tradición discursiva, variación concepcional y variedades diasistémicas en un epistolario popular (España, 1925-27) ...	551
Rika Van Deyck: Des concepts cosériens de savoir, norme et tradition à celui des traditions discursives .....	565

# ‘De l’oral dans l’écrit’? – Le statut variationnel des SMS (textos) et leur valeur pour la recherche linguistique

*A Gerold Hilty, cher ami et grand modèle,  
dans la recherche comme dans la vie*

## *Abstract*

*This contribution aims at identifying the variationist profile of text messages, based on a corpus of Swiss text messages in French, together with their general relevance for linguistic research. Empirical analyses show the strikingly conservative character of the data, with very high percentages of standard orthographic marking of morphosyntactic information (clitic subjects, agreement) on the one hand, and interesting deviations from the standard norm, which are partially conform to phonic French (marking of sentential negation without the preverbal particle ne) and partially indicate the presence of universal principles also in the grammar of text messages. Overall, the results do not indicate a new variety of French in text messages, but highlight their special character as a graphic interactional electronic form of communication of particular relevance both for grammar oriented and variationist approaches to language.*

## 1. Introduction

Commençons notre discussion par un exemple tiré du nouveau corpus de référence des SMS suisses (cf. <[www.sms4science.ch](http://www.sms4science.ch)> et partie 2):

(1) *Hey Basto!! On prend un monstre apéro au rosé! Fais chière que tu sois pas là!!!! Pompe!*

Cet exemple comprend bien des variantes linguistiques qui ont une certaine valeur indicative pour différentes dimensions de variation linguistique du français contemporain (selon le modèle de Coseriu repris et élargi par Koch/Oesterreicher <sup>2</sup>2011):

Plusieurs lexèmes indiquent un niveau de langue (variation diaphasique) qu’on pourrait identifier comme ‘français populaire’ ou ‘français familier’, tels que *monstre*, (*faire*) *chier*, *pompe*. La graphie déviante <fais chière> (orthographe standard: <il/ça fait chier>) semble comporter un marquage diastratique bas - l’auteur du SMS soit ne maîtrise pas la graphie standard pour des problèmes culturels/scolaires, soit a fait exprès de choisir ces variantes graphiques pour signaler ‘le langage SMS’ en tant que ‘langage spécial’ d’un groupe (voir plus bas), soit n’a tout simplement pas corrigé ses



fautes de frappe. Des formules d'adresse comme *Hey* ou *tu* suggèrent finalement une certaine connivence entre l'auteur et le destinataire de ce SMS, et nous nous trouvons donc très probablement face à un produit de communication privée, assez informelle, qui relève de 'l'immédiat communicatif' selon Koch/Oesterreicher (<sup>2</sup>2011).

C'est sans doute ce dernier trait qui caractérise la majorité de nos SMS (voir partie 2), ce qui a induit beaucoup de chercheurs à considérer la communication par SMS ou toute la «communication médiée par ordinateur» (CMO, en anglais *computer-mediated communication*, CMC, Walther 1996) comme une communication 'hybride' entre le parlé et l'écrit. Ainsi, Baron (1998) parle, à propos du courrier électronique, de «speech by other means», chez Anis (2007), il est question de «neography», Gadet (2008, 527) classe ces formes de communication comme «de l'écrit avec des caractéristiques de proximité», Cougnon (2008) comme «l'écrit spontané», et Dejong (2002, 30) postule même qu'«[...] on écrit comme on parle!» dans les SMS. On peut apercevoir derrière ces classifications aussi un aspect souvent négligé dans les approches traditionnelles à la variation linguistique, c'est-à-dire le paramètre de la 'synchronicité' dans laquelle se déroulent ou non les différentes formes d'interactions linguistiques:

[...] some researchers have suggested a continuum along which asynchronous CMD [= *computer-mediated discourse*, ES] occupies a position closer to writing, and synchronous CMD occupies a position closer to speaking [...]. (Herring 2007)

Puisque la communication médiée par ordinateur est toujours graphique de par sa nature technologique, mais se passe rapidement et de manière quasi-synchrone, elle englobe aussi des aspects de l'interaction dialogique, traditionnellement phonique – ce qui crée l'impression d'hybridité. Ainsi on pourrait se poser la question de savoir si nous nous voyons confrontés à une *nouvelle variété* voire une *nouvelle langue* avec le 'langage SMS' ou l'écrit électronique, une question qui a été soulevée déjà à plusieurs reprises (cf. Crystal (2001), qui parle de «Netspeak», Dejong (2002), qui assume même une «cyberl@ngue»; cf. Gadet (2008, 523) pour l'hypothèse d'un nouveau genre; à voir aussi Herring (2007) pour une excellente problématisation).

Une explication alternative des particularités que nous avons pu observer dans l'exemple (1) a recours à la seule altération graphique du standard dans l'écrit électronique, qui ne montrerait dans cette perspective que de nouvelles *stratégies d'écriture*:

Or, une des propriétés essentielles de la CéMO est son altération de la graphie. Pour répondre à des besoins d'économie d'espace et de temps [...], les usagers de la CéMO montrent une tendance manifeste à l'abréviation (graphique). (Cougnon 2010, 397)

Dans cette perspective, les nouvelles formes électroniques de communication, dont les SMS, n'offriraient que la possibilité de 'mise à l'écrit' de variantes non-standard d'habitude confinées au code phonique et à l'informalité, plus une tendance générale à l'abréviation, et ne constitueraient pas de variétés, genres, styles ou dimensions de variation propres en elles-mêmes. Par conséquent, nous formulons la question suivante, qui guidera notre discussion dans ce qui suit: Les SMS relèvent-ils

(2) ‘de l’écrit avec des traits du parlé’

ou

(3) seulement de nouvelles stratégies d’écriture?

La première hypothèse s’intéresse à la conception d’un message linguistique selon la relation entre les deux interlocuteurs, les circonstances spatiales et sociales de leur échange etc. ; il s’agit de la dimension de variation entre ‘l’immédiat communicatif’ et ‘la distance communicative’ selon Koch/Oesterreicher (2011). Si cela s’avérait vrai, le ‘langage SMS’ constituerait une nouvelle variété du français contemporain située sur ce continuum. La deuxième hypothèse concerne le médium, le caractère physique des éléments qui transportent le message linguistique (des sons, ondes sonores, ou des lettres, caractères graphiques) et n’aurait pas d’implications pour une éventuelle modification du français et de son espace variationnel.

Après cette brève introduction, nous allons présenter notre base de données ainsi que quelques résultats empiriques de leur analyse dans la partie 2, avant de les discuter d’un point de vue variationnel dans la partie 3. La partie 4 thématise la pertinence de données graphiques non-standard pour la recherche linguistique générale, et la contribution se terminera sur une brève conclusion, qui s’abstient de postuler une nouvelle variété ‘SMS’ du français tout en revendiquant un rôle clé de ce genre de données pour la recherche.

## 2. Notre corpus et quelques résultats d’analyse

Dans ce qui suit, nous présenterons d’abord brièvement le projet *sms4science.ch* dont sont issues nos données ainsi que quelques résultats de leur analyse quantitative qui pourront aider à trancher entre les deux alternatives formulées ci-dessus.

### 2.1. Le corpus

En automne/hiver 2009/2010, les universités de Zurich et de Neuchâtel ont recueilli des SMS authentiques de toute la Suisse, et en mai/juin 2011, un appel ciblé a été lancé en collaboration avec l’université de Berne pour recueillir plus de données italiennes et romanches-grisonnes. Cette action a été soutenue par les médias et la Swisscom, la compagnie téléphonique la plus importante de la Suisse. Les utilisateurs de téléphones mobiles ont été invités à transmettre, à un numéro gratuit, une copie de chaque SMS qu’ils avaient envoyé auparavant (aux différents destinataires originaux) et à remplir un questionnaire sociolinguistique anonyme sur internet. Les données ainsi rassemblées ont ensuite été intégrées au corpus *sms4science.ch*. En remplissant le questionnaire, les participants ont donné leur accord pour l’utilisation de leurs réponses dans le cadre de différents projets scientifiques lancés par les deux universités. Les SMS et les réponses au questionnaire ont été enregistrés dans une banque de données, dans le but d’étudier le phénomène de la communication par SMS en Suisse (cf. <[www.sms4science.ch](http://www.sms4science.ch)>). 25947 SMS (environ 500000 mots) ont

été collectionnés ainsi, dont 18% en français ou une variété du français. Ces SMS ont été envoyés par 2784 personnes différentes. Parmi les participants, 1316 (donc la moitié environ, mais ce sont les auteurs des trois quarts de nos messages !) ont rempli le questionnaire anonyme concernant les facteurs sociodémographiques (domicile, âge, sexe, langue maternelle, éducation etc.). 909 personnes ont déclaré avoir un dialecte suisse allemand comme langue maternelle, 165 ont indiqué l'allemand non-dialectal, 257 le français, 107 l'italien et 64 le rhéto-roman. 129 personnes ont nommé une autre langue comme langue maternelle (plusieurs réponses étaient possibles). 46% des participants étaient âgés de 20 à 29 ans, mais il faut mentionner que 40 personnes de plus de 60 ans ont envoyé leurs SMS et rempli le questionnaire.

Le corpus *sms4science.ch* montre une forte tendance à l'alternance codique. En outre, il montre une préférence pour les dialectes (dans les SMS en allemand, romanche et italien). Au total, le corpus englobe plus de 20 langues et variétés diverses (entre autres l'allemand, le français, l'italien, l'anglais, l'espagnol, le néerlandais, le suédois et le portugais). 10718 SMS (environ 41% ou 218000 mots) sont rédigés en un dialecte suisse alémanique, 7224 SMS (environ 140000 mots) sont en allemand non-dialectal et 4627 SMS (environ 94000 mots) en français (et 26 en francoprovençal). Ces langues et variétés sont les plus fréquemment représentées dans le corpus.

## 2.2. *La réalisation graphique de l'accord sujet-verbe conjugué et sujet/complément d'objet direct-participe passé (cf. Stark 2011, Stark/Riedel 2013)*

L'analyse quantitative des 400 premiers SMS français selon cinq types d'accord sujet-verbe graphiques, publiée dans Stark 2011, a montré que l'accord graphique explicite ou canonique (cf. Corbett 2006) est le type d'accord par défaut dans nos données. Ceci correspond à l'orthographe standard ou un marquage déviant du standard qui marque néanmoins le sujet et les traits d'accord dans une graphie alternative, nommée ici 'graphie SMS' (cf. l'exemple (3) et la partie 2.3). Avec un sujet lexical, on trouve l'accord graphique explicite dans 95,45% des cas (90,45% en orthographe standard), avec un sujet clitique, il est réalisé dans 87,96% des cas (67,23% en orthographe standard, cf. l'exemple (2) en bas).

Pour le marquage de l'accord sujet/cod-participe passé, qui a été analysé dans toutes les données françaises (cf. Stark/Riedel 2013), nous avons constaté un résultat similaire : l'accord suivant l'orthographe standard se présente dans 89,4% des cas, et ceci indépendamment de 'l'audibilité' de l'accord (qui ne se réalise que rarement dans le code phonique, cf. Hunnius 1980). De plus, les 'fautes' trouvées suivent dans leur grande majorité les régularités universelles observées pour d'autres langues et variétés (cf. Greenberg 1963) : on trouve plus de fautes quand le contrôleur suit la cible de l'accord, et plus de fautes avec le genre qu'avec le nombre :

- (2) *Coucou maman, comment vas<sub>2.SG</sub> -tu<sub>2.SG</sub>? Et comment s'est passé<sub>SG.M</sub> cette<sub>SG.F</sub> journée? Est-ce que tu<sub>2.SG.(F.)</sub> es<sub>2.SG</sub> finalement quand même allée<sub>SG.F</sub> chez Sheila?? J'espère... Bisous, bonne soirée*

### 2.3. La réalisation (graphique) du sujet (400 premiers SMS; cf. Stark 2013, Dürscheid/Stark 2013)

Quant à la réalisation graphique du sujet, c'est l'orthographe standard qui l'emporte et qui semble être la graphie par défaut (755), à côté de certains écarts graphiques qui restent tout de même assez proches du standard (71). Dans 80% des cas, les sujets sont marqués comme unités graphiques séparées du verbe et de manière identique ou très proche de l'orthographe standard. Une réalisation graphique 'agglutinante' ou 'fusionnante' du sujet dans les SMS pourrait servir d'indicateur à des tendances de réanalyse en tant que simples marqueurs d'accord (cf. Culbertson 2010, Stark 2013). Or, nous avons constaté en général que notre corpus comprend très peu de sujets lexicaux (73 sur 1059 combinaisons sujet-verbe<sub>conj</sub>) et surtout très peu de réalisations 'agglutinantes' des clitiques préverbaux (cf. l'exemple (4)). Le plus souvent, nous assistons à une simple omission de lettres vocaliques, qui imite l'omission non-standard surtout de la voyelle [y] dans [ty] devant attaque vocalique du verbe conjugué (204 *tokens*, cf. exemple (3)).

Nous trouvons seulement trois variantes graphiques 'agglutinantes' conventionnalisées telles que <g> pour <j'ai>, <c> pour <c'est> et <c t> pour <c'était>, mais ces graphies sont plutôt rares (30 *tokens*):

(3) *T'en as d'autres qui sont bien au moins?*

(4) *Alor Ça c trobien!*

L'exemple (4) contient aussi une 'combinaison XP-cl' (nous n'allons pas discuter ici s'il s'agit d'un redoublement du sujet non marqué ou d'une dislocation à gauche à valeur topicalisante, cf. Culbertson 2010), une structure fréquemment attestée dans des corpus oraux, mais rare dans nos données. Dans les 400 SMS français analysés, nous ne trouvons que 1 cas de syntagme nominal-cl, 4 cas de *moi je* (sur 366 *je*), 2 cas de *toi tu* (sur 196 *tu*), 2 cas de *nous on* (sur 9 *nous*; 115 *on*) et 1 cas de *ça c'est* (sur 80 *ça*).

### 2.4. Le marquage de la négation

Dans Stark (2012a), nous avons pu démontrer, sur une analyse exhaustive de tous nos messages courts en français, que le facteur le plus important pour déclencher l'omission du *ne* de négation, variable sociolinguistique fameuse en français contemporain, est la catégorie grammaticale du sujet, comme il a été relevé à plusieurs reprises depuis du moins les travaux de Ashby (1981) pour différents corpus oraux (cf. p.ex. les chiffres comparables dans Hansen/Malderez 2004, Dufter/Stark 2007, et ceux de Van Compernelle 2008, 2010 pour les chats):

	-ne	+ne	Σ
Sujet lexical	27 (34,2%)	52 (65,8%)	79 (100%)
Sujet pronominal	850 (80,2%)	210 (19,8%)	1060 (100%)

Tableau 1: Omission du *ne* de négation après sujet lexical vs. sujet pronominal (clitique)

Comme en témoignent les chiffres présentés dans le tableau 1 ainsi que les exemples (5) et (6) ci-dessous, le *ne* est très souvent omis après un sujet pronominal, surtout après les sujets clitiques, et il est présent après un sujet lexical :

(5) *Alors ma tenue ne t'a pas séduit hier?*

(6) *Bah, j'ai pas encore trouvé un congrès où je me sentais [...] à ma place*

### 3. Le statut variationnel des SMS

#### 3.1. Résumé des résultats d'analyse morphosyntaxique et leur pertinence pour la linguistique variationnelle

Comme nous avons vu dans l'introduction, la fréquente occurrence de certains indicateurs lexicaux de l'immédiat communicatif nous permet de conclure au marquage 'privé', 'informel' de la grande majorité de nos messages (p.ex. des formules et pronoms d'adresse comme *maman, coucou, tu, bisou* etc.). A part ce constat, nous avons démontré de manière empirique-quantitative, dans la partie 2, une rareté générale de sujets lexicaux, marqueur variationnel morphosyntaxique typique des communications d'immédiat communicatif' (cf. la notion de *Preferred Argument Structure* de Du Bois (1987) et la discussion dans Dufter/Stark 2007). Ces deux phénomènes sont de nature universelle et valent pour toutes les langues et communications.

En ce qui concerne le français contemporain en tant qu'idiome spécifique, notre analyse, empirique à nouveau, du marquage de la négation de phrase avec *ne...pas/plus/point* etc. ou seulement le forclusif négatif postverbal, autre marqueur variationnel par excellence, confirme une régularité interne du français (parlé), partiellement indépendante du médium et souvent observée dans des études de corpus précédentes (cf. Stark 2012a), à savoir l'interdépendance entre le *ne* de négation et la nature de l'élément le précédant (sujet lexical vs. sujet pronominal / clitique). Contrairement à cela et contrairement à ce qu'on trouve dans des corpus oraux, nous avons dû constater une quasi-absence de structures de redoublement du sujet (et d'objet, cf. Stark 2012b). Et finalement, la stabilité de l'orthographe standard comme marquage par défaut – une variante diastratiquement plutôt haute, peut-être due à la surreprésentation de participants avec une formation secondaire, voire académique – s'est avérée une spécificité surprenante de notre corpus, surtout en ce qui concerne le marquage de l'accord entre le sujet et le verbe conjugué d'une part, et entre le sujet ou le complément d'objet et le participe passé, de l'autre.

### 3.2. Conclusion basée sur le résumé des résultats d'analyse morphosyntaxique

Comment interpréter tous ces résultats, partiellement contradictoires entre eux ? Il est hors de doute que nos SMS relèvent dans leur majorité de 'l'immédiat communicatif' (universel). En outre, ils montrent les mêmes régularités que le français tout court (négation), et même du français standard (écrit ; rareté de redoublements, orthographe standard, marquage de l'accord). Les rares graphies déviantes ('fautes') indiquent partiellement (accord) des régularités universelles et l'indépendance du code phonique (cf. surtout les résultats pour l'accord du participe passé, Stark/Riedel 2013). De plus, nous n'avons pu relever dans nos données que très peu d'occurrences des stratégies 'd'écriture SMS' et peu de nouvelles formes graphiques conventionnalisées, c'est-à-dire suffisamment fréquentes et stables pour constituer de vrais indicateurs variationnels de ce 'langage SMS' (cf. aussi Zimmermann 2009, sur le corpus belge cf. Fairon et al. 2006). Ces nouvelles 'stratégies d'écriture' témoignent d'un certain phonocentrisme partiel (cf. les syllabogrammes <c> ou <c t> ou le rébus <t> pour <t'es>). Le bilan global de nos études empiriques est donc mitigé : d'une part, les SMS dévient du standard et possèdent certaines caractéristiques de 'l'immédiat communicatif', quoique dans une moindre mesure que nous n'aurions pensé, de l'autre, ils démontrent des propriétés spécifiques qui les distinguent aussi de corpus oraux de l'immédiat. Tout cela indique la pertinence du médium, du moins pour cette forme de communication électronique, aspect souvent sous-estimé en linguistique variationnelle romane traditionnelle et vers lequel nous allons nous tourner dans la prochaine partie.

## 4. Les données graphiques non-standard dans la recherche linguistique

Depuis le 'phonocentrisme' radical de l'école structuraliste du début du vingtième siècle, la validité générale des sources graphiques pour la recherche linguistique a été constamment mise en question. Ceci en dépit du fait que toutes nos connaissances sur les périodes antérieures à l'époque contemporaine proviennent de sources écrites, graphiques, très souvent artistiques et par là assez éloignées quant à leurs propriétés linguistiques du langage naturellement acquis dans l'enfance. Ce fait rend nécessaire une brève réflexion fondamentale sur le rapport entre 'écrit' et 'parlé' (médium) et le statut de la graphie dans la recherche linguistique. On peut identifier *grosso modo* trois positions en linguistique actuelle : la première part d'une dépendance totale de la graphie de la phonie et la conçoit comme 'miroir fidèle' du code phonique (cf. p.ex. Martinet 1960) ; la deuxième postule une autonomie de la graphie en tant que système sémiotique propre (cf. p.ex. Anis 1988 ; cf. sur la position «phonocentrique» aussi Béguelin 2012). Ni l'une, ni l'autre ne permettraient de prendre des données graphiques comme base de la recherche linguistique, les considérant soit comme superflues et moins fidèles à la réalité que le code phonique soit comme fondamentalement différentes du phonique. Troisièmement, un faisceau de positions qu'on pourrait

dénommer avec Béguelin (2012) comme des « approches pondérées », reconnaît une autonomie partielle de la graphie, mais aussi une forte interdépendance entre graphie et phonie, en étudiant, p. ex., la différence typologique dans le marquage des informations morphosyntaxiques entre le code phonique et le code graphique en français contemporain (cf. Blanche-Benveniste (2010, 152-159)).

S'il est vrai que des données d'une orthographe standard particulièrement morphocentriste et conservatrice comme celle du français n'indiquent que très indirectement la réalité phonique actuelle (contrairement à des systèmes orthographiques 'plats', phonocentristes, comme celui du roumain p.ex., cf. Meisenburg 1996 et Dürscheid/Stark 2013), un corpus de SMS rassemble aussi des 'graphies informelles', libérées en principe de la norme orthographique. Un tel corpus peut, comme il a été remarqué à maintes reprises dans la littérature spécialisée, peut-être dans une mesure exagérée (cf. les citations dans la partie 1) refléter *partiellement* ce qui se passe dans le code phonique en communication informelle ('français de l'immédiat'), sans oublier qu'on y trouvera sans doute aussi des traces des *conventions orthographiques* apprises à l'école. L'écriture (normative) et ses régularités, ne l'oublions pas, n'est jamais acquise de manière naturelle, elle est apprise pendant plusieurs années dans un processus conscient – ce qui n'empêche pas que l'on applique de manière automatisée aussi et surtout dans des contextes privés où le 'correct' ou l'incorrect ne jouent pas de rôle décisif. Cette observation se retrouve aussi partiellement dans la citation suivante à propos de 'l'écriture SMS', qui constate à la fois la présence de certaines procédures 'quasi-phonétiques'<sup>1</sup> et des écritures idéographiques :

[...] [l'] application du principe phono- ou syllabographique, sous-jacent à certains des procédés, est contrebalancée par celle du principe idéographique [...]. (Béguelin 2012, 59)

De là notre hypothèse générale qui revendique même un *rôle clé* des données graphiques non-standard pour la recherche linguistique. Même si leur graphie suit aussi les régularités propres au système orthographique (ce que l'on peut voir, p.ex., dans le codage assez conservateur des informations morphosyntaxiques dans l'orthographe française bien documenté aussi dans nos SMS), qui sont peu pertinentes pour la linguistique/recherche grammaticale, elles reflètent toujours aussi des régularités grammaticales universelles dans les formes déviantes de la norme plus la réalisation quasi-phonétique de certains morphèmes, ce qui est d'une grande pertinence pour la linguistique ou la recherche grammaticale. Tout cela revient à dire que la graphie n'est pas un système, mais un « plurisystème » (cf. Catach 1980), une combinaison du normé (c'est le côté autonome de la graphie, peu pertinent pour la grammaire) et du systématique (à savoir les régularités universelles reflétées dans les 'fautes' et l'imitation de la phonie, toutes les deux centrales pour la grammaire).

<sup>1</sup> A voir aussi les formulations suivantes: « la mimésis de l'oralité » (Béguelin 1998, 240), « mimésis de prononciation » (Gadet 2008, 518), « l'orthographe phonétique » (Anis 2007), « réappropriation spontanée' de la graphie par les sujets » (cf. Béguelin 2012, 53).

## 5. Conclusion

Les analyses empiriques de notre corpus de SMS suisses francophones ont démontré que leur morphosyntaxe est presque inaltérée et marquée même selon l'orthographe standard dans la majorité des cas (peut-être dû partiellement à notre échantillon de participants avec une éducation assez élevée, facteur sociologique non négligeable). Les écarts de l'orthographe que nous avons identifiés relèvent de l'universel (accord) ou de l'idiomatique (français: négation), mais dans la plupart des cas de régularités grammaticales, comme dans tout autre type de corpus/données. Seuls les indicateurs lexicaux comme les formules de salutation et d'adresse et le taux de sujets lexicaux indiquent l'immédiat communicatif universel, et le caractère *graphique* des SMS est sans doute responsable de leur écart partiel des autres variétés (phoniques) de l'immédiat (français), p.ex. en ce qui concerne les structures à redoublement.

Somme toute, notre réponse à la question générale posée dans l'introduction est la suivante : les SMS ne sont en rien une nouvelle variété du français (ils ne possèdent pas de traits linguistiques propres ni de corrélation stable avec des facteurs externes), mais tout au plus une nouvelle *forme de communication* (cf. Dürscheid 2002, 113) avec une certaine liberté vis-à-vis des normes de l'écrit et des stratégies d'écriture propres, pourtant rarement utilisées. Cette 'libération' leur confère un statut central dans la recherche grammaticale, où les écarts de l'orthographe standard nous permettent de jeter un coup d'œil sur ce qui se passe en français phonique contemporain, aussi dans le contexte de potentiels changements linguistiques à l'avenir (cf. Stark 2013).

Université de Zurich

Elisabeth STARK

## Bibliographie

### *Sites web :*

<www.sms4science.ch>

<www.sms4science.org>

### *Études :*

Anis, Jacques, 1988. *L'écriture. Théories et descriptions*, Bruxelles, De Boeck.

Anis, Jacques, 2007. «Neography – Unconventional Spelling in French SMS Text Messages», in: Danet, Brenda / Herring, Susan C. (ed.), *The Multilingual Internet – Language, Culture and Communication Online*, New York, Oxford University Press, 87-115.

Ashby, William, 1981. «The loss of the negative particle *ne* in French. A syntactic change in progress», *Language* 75, 674-687.



- Baron, Naomi, 1998. «Letters by phone or speech by other means: the linguistics of email», *Language and Communication* 18, 133-170.
- Béguelin, Marie-José, 1998. «Le rapport écrit-oral. Tendances dissimilatrices, tendances assimilatrices», *Cahiers de linguistique française* 20, 229-253.
- Béguelin, Marie-José, 2012. «La variation graphique dans le corpus suisse de SMS en français», in: Caddéo, S. / Roubaud, M.-N. / Rouquier, M. / Sabio, F. (ed.), *Penser les langues avec Claire Blanche-Benveniste*, Presse de l'Université de Provence, 47-63.
- Blanche-Benveniste, Claire, [1997] 2010. *Approches de la langue parlée en français*, Gap, Ophrys.
- Catach, Nina, 1980. *L'orthographe française. Traité théorique et pratique*. Avec la collaboration de C. Gruaz & D. Duprez, Paris, Nathan.
- Corbett, Greville G., 2006. *Agreement*, Cambridge, Cambridge University Press.
- Cougnon, Louise-Amélie, 2008. «Le français de Belgique dans 'l'écrit spontané'. Approche d'un corpus de 30.000 SMS», *Travaux du Cercle Belge de Linguistique*.
- Cougnon, Louise-Amélie, 2010. «Orthographe et langue dans les SMS», *Ela. Études de linguistique appliquée* 160-4, 397-410.
- Crystal, David, 2001. *Language and the Internet*, Cambridge, Cambridge University Press.
- Culbertson, Jennifer, 2010. «Convergent evidence for categorial change in French: From subject clitic to agreement marker», *Language* 86-1, 85-132.
- Dejond, Aurélie, 2002. *La cyberl@ngue française*, Tournai, La Renaissance du Livre.
- Du Bois, John W., 1987. «The Discourse Basis of Ergativity», *Language* 63-4, 805-855.
- Dürscheid, Christa, 2002. «E-Mail und SMS – ein Vergleich», in: Ziegler, Arne / Dürscheid, Christa (ed.), *Kommunikationsform E-Mail*, Tübingen, Stauffenburg, 93-114.
- Dürscheid, Christa / Stark, Elisabeth, 2013. «Anything goes? SMS, phonographisches Schreiben und Morphemkonstanz», in: Neef, Martin / Scherer, Carmen (ed.), *Die Schnittstelle von Morphologie und geschriebener Sprache*, Berlin, de Gruyter, 189-209.
- Dufter, Andreas / Stark, Elisabeth, 2007. «La linguistique variationnelle et les changements linguistiques 'mal compris'. Le cas de la 'disparition' du *ne* de négation», in: Marchello-Nizia, Christiane / Combettes, Bernard (ed.), *Etudes sur le changement linguistique en français*, Nancy, Presses Universitaires de Nancy, 115-128.
- Fairon, Cédric / Klein, Jean René / Paumier, Sébastien, 2006. *Le Corpus SMS pour la science. Base de données de 30.000 SMS et logiciels de consultation* (CD-Rom), Louvain-la-Neuve, Presses universitaires.
- Gadet, Françoise, 2008. «Ubi scripta et volat et manent», in: Stark, Elisabeth / Schmidt-Riese, Roland / Stoll, Eva (ed.), *Romanische Syntax im Wandel*, Tübingen, Narr, 513-529.
- Greenberg, Joseph H., 1963. «Some universals of grammar with particular reference to the order of meaningful elements», in: Greenberg, Joseph H. (ed.), *Universals of Language*, Cambridge/Massachusetts, MIT Press, 73-113.
- Hansen, Anita Berit / Malderez, Isabelle, 2004. «Le *ne* de négation en région parisienne. Une étude en temps réel», *Langage et société* 107, 5-30.
- Herring, Susan C., 2007. «A Faceted Classification Scheme for Computer-Mediated Discourse», *Language@Internet* 4, article 1.
- Hunnius, Klaus, 1980. «Sprachnorm und Sprachgebrauch: Zur Geschichte der Kongruenzregeln des *participe passé*», in: Bork, Hans-Dieter / Greive, Artur / Woll, Detlev (ed.), *Romanica Europaea et Americana. Festschrift für Harri Meier zum 75. Geburtstag*, Bonn, Bouvier, 131-169.

- Koch, Peter / Oesterreicher, Wulf, [1990] 2011. *Gesprochene Sprache in der Romania. Französisch – Italienisch – Spanisch*, Berlin, Mouton de Gruyter.
- Martinet, André, 1960. *Eléments de linguistique générale*, Paris, Armand Colin.
- Meisenburg, Trudel, 1996. *Romanische Schriftsysteme im Vergleich. Eine diachrone Studie*, Tübingen, Narr.
- Stark, Elisabeth, 2011. «La morphosyntaxe dans les SMS suisses francophones: Le marquage de l'accord sujet – verbe conjugué», *Linguistik online* 48. <[http://www.linguistik-online.de/48\\_11/stark.htm](http://www.linguistik-online.de/48_11/stark.htm)>
- Stark, Elisabeth, 2012a. «Negation marking in French text messages», *Linguisticae Investigationes* 35-2, 341-366.
- Stark, Elisabeth, 2012b. «Réanalyses dans la graphie – 'l'écrit spontané' des SMS et le statut des pronoms clitiques du français contemporain». Conférence invitée au colloque "Réanalyses" à Neuchâtel, 5-7 septembre 2012.
- Stark, Elisabeth, 2013. «Clitic subjects in French text messages: Does technical change provoke and/or reveal linguistic change?», in: Jeppesen Kragh, Kirsten / Lindschouw, Jan (ed.), *Deixis and Pronouns in Romance Languages*, Amsterdam/Philadelphia, Benjamins, 147-169.
- Stark, Elisabeth / Riedel, Isabelle, 2013. «L'accord du participe passé dans les SMS francophones du corpus SMS suisse», *Romanistisches Jahrbuch* 63(1), 116-138.
- Van Compernelle, Rémi A., 2008. «Morphosyntactic and phonological constraints on negative particle variation in French-language chat discourse», *Language Variation and Change* 20-2, 317-339.
- Van Compernelle, Rémi A., 2010. «The (slightly more) productive use of *ne* in Montreal French chat», *Language Sciences* 32, 447-463.
- Walther, Joseph B., 1996. «Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction», *Communication Research* 23, 3-43.
- Zimmermann, Theo, 2009. *Le 'langage SMS' – une nouvelle variété écrite de la langue française? Une analyse empirique basée sur un corpus de 30'000 SMS sous considération particulière de la relation phonie-graphie*. Mémoire de licence, sous la direction de Elisabeth Stark, Université de Zurich.