

La variabilité formelle des questions dans les écrits SMS

Alexander GURYEV^{1 2} & François DELAFONTAINE¹

¹Institut des sciences du langage et de la communication, Université de Neuchâtel

²ED 268 Langage et langues, Université de Paris 3 Sorbonne Nouvelle

In classical sociolinguistic studies, the variety in French interrogative structures is accounted for in terms of the different socio-stylistic values that they express. However the study of SMS data provides us with evidence that this approach needs to be challenged. This paper argues in favor of an alternative approach, which views various interrogative structures as "multiple forms in concurrence" (Quillard 2000) with their own structural and functional particularities, instead of viewing various interrogative structures as alternative ways of expressing the same thing. In line with this position we will discuss some methodological issues and the design of this approach, leading us through the article from the annotation of linguistic data to some preliminary results of the study.

1. Introduction

La variété formelle de l'interrogative en français a été étudiée dans le cadre de plusieurs approches, notamment les trois suivantes: (i) les analyses syntaxiques, qui ciblent les contraintes grammaticales ou structurelles dans la production des interrogatives (Kayne 1972; Danjou-Flaux & Dessaux 1976; Obenauer 1976; Jones 1999; Druetta 2009), (ii) les analyses sémantico-pragmatiques, qui s'intéressent aux relations entre une forme interrogative et sa fonction vue comme usage potentiel (Weinrich 1989; Mosegaard-Hansen 2001; Berrendonner 2005; Abeillé & al. 2012), ou encore (iii) les approches variationnistes, qui tentent de rendre compte de la variété des interrogatives en considérant une multiplicité de facteurs, tant linguistiques (grammaticaux) que non linguistiques (communicatifs ou sociolinguistiques), susceptibles d'avoir une incidence sur le choix des interrogatives (Coveney 1996¹=2002²; Quillard 2000; Elsig 2009; Adli 2006, 2013). Plusieurs de ces études se sont intéressées dans une perspective quantitative aux corpus de français parlé. Il a ainsi été démontré que l'oral ordinaire est caractérisé, d'une part, par l'extrême fréquence de formes interrogatives à maintien de l'ordre canonique Sujet-Verbe et, d'autre part, par la rareté des autres formes, notamment structures inversées.

Quant à cet article, il s'intéresse à la variabilité formelle des questions dans les écrits SMS (angl. *Short Message Service*) et vise particulièrement à répondre aux questionnements suivants:

- Les tendances observées dans les corpus oraux informels sont-elles aussi valides pour d'autres types de discours informels, à savoir les nouvelles formes de communication écrite véhiculées par les SMS?
- Ce type des données a-t-il quelque chose à nous apprendre à propos de la variation dans l'usage des interrogatives?
- De quelle manière faut-il aborder, de manière plus générale, les faits de variation syntaxique?

Sans prétendre à l'exhaustivité, cet article se focalisera sur les questions d'ordre méthodologique et procédera à une analyse exploratoire de la variabilité formelle des questions dans les messages du Corpus suisse de SMS en français¹. Dans un premier temps, nous présenterons les données utilisées, avant de comparer les faits de langue de ce type de corpus avec les données de l'oral spontané. Dans un deuxième temps, nous discuterons de l'approche à adopter pour analyser la variabilité des questions en général. Pour terminer, nous présenterons une procédure d'annotation qui nous permettra de mettre en évidence le rôle des paramètres morphosyntaxiques dans la sélection des formes interrogatives à disposition du locuteur.

2. Le Corpus suisse de SMS

Le projet *sms4science* a été lancé en 2004 par le CENTAL de l'Université de Louvain en Belgique (prof. C. Fairon), via l'opération 'Faites don de vos SMS à la science'. L'objectif des initiateurs était de collecter un très vaste corpus de messages afin de "contribuer à l'étude de la communication par SMS ('SMS', 'txt', 'texto', etc.) et à l'étude du langage qu'elle véhicule"². Quelques années plus tard, ce projet a pris une dimension internationale avec la participation de la Réunion (2008), de la Suisse (2009), du Québec au Canada, des Hautes Alpes et de l'Isère en France (2010). En Suisse, le projet s'est développé autour des Universités de Zurich, de Neuchâtel et de Berne avec la collaboration de l'Université de Leipzig (Allemagne). Il a notamment donné lieu à sept thèses de doctorat dont les questions de recherche sont explicitées dans ce numéro de Tranel (voir les articles de Robert-Tissot, de Cathomas, Ferretti, Bucher et Morel, et de Frick en collaboration avec Gazin et Meisner).

Le corpus collecté en Suisse sous l'égide des Universités de Zurich et de Neuchâtel se monte à 25'947 textos rédigés essentiellement, mais non exclusivement, dans les quatre langues nationales de la Suisse (allemand et

¹ Cf. <https://sms.linguistik.uzh.ch> (Stark, Ueberwasser & Ruef 2009-2014).

² Cf. <http://www.sms4science.org/>, le 29 mai 2015

dialectes alémaniques, français, italien, romanche). Il comprend 4'619 messages en français qui nous ont servi de données dans cette étude (fig. 1)³:

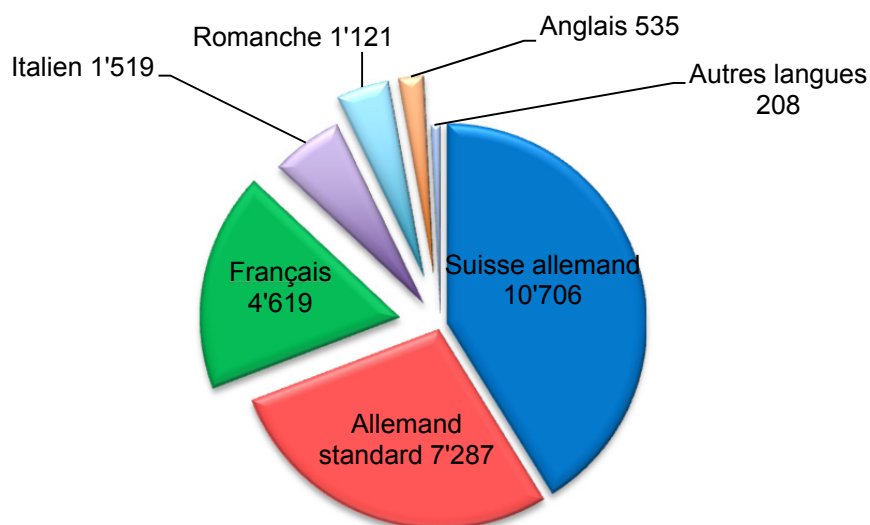


Fig.1: Répartition des langues dans le Corpus suisse de SMS

De manière globale, les données ont été "légues à la science" par 2'784 donateurs âgés de 12 à 77 ans, dont 46 % ont entre 20 et 29 ans (pour plus de détails, voir le site www.sms4science.ch).

3. Les écrits SMS en tant que données linguistiques

Si les écrits SMS sont souvent associés dans le large public à des productions langagières malformées et à une écriture peu compréhensible, voire agrammaticale, nous verrons ci-dessous que la réalité est loin de ce stéréotype.

3.1 Néographie⁴ vs orthographe standard

Au premier abord, la pratique du SMS fait penser à une rupture totale avec la norme de l'orthographe standard, surtout face à des messages qui abondent en procédés de néographie et d'abréviations:

- (1) Hello, comen va? Jespèr ke ton operation c bi1 pacé é ke ta pa du alé en pédiatri come moi!lol,ya dé joli pti desin o mur la ba! Peutu mdir kel tail de chemis il te fau pr lé scout? Merci biz a+ (18395)⁵

³ Lors des analyses qualitatives et afin de renforcer nos observations, nous avons aussi utilisé quelques exemples en provenance du Corpus "88milSMS" recueilli à Montpellier en France par Panckhurst et ses collègues (Panckhurst & al. 2014).

⁴ Nous nous référons ici au terme anglais "neography" employé par Anis (2007).

Toutefois, diverses études portant sur le français (Cougnon & François 2010), l'allemand (Bieswanger 2007), l'anglais (Thurlow & Brown 2003), ou encore le norvégien (Ling 2005) montrent qu'en dépit des préjugés dont fait l'objet la pratique du texto, les scripteurs, toutes catégories sociales confondues, se passent en majorité d'abréviations (cf. Krummes & al. 2014)

Par ailleurs, si l'écriture non-conventionnelle est favorisée par l'interaction par SMS en tant que genre de discours, plusieurs études montrent que les procédés associés à l'usage de graphie non-standard sont néanmoins ordonnés, relèvent d'une logique (Béguelin 2012a) et assurent diverses fonctions (Cougnon 2010). Par exemple, comme le remarque Cougnon (2010: 401), qui a étudié divers corpus de SMS issus de différents pays francophones, les auteurs des textos peuvent "marquer leur appartenance géographique et renforcer une complicité avec leur interlocuteur de même appartenance". Elle montre ainsi qu'en Suisse certains scripteurs peuvent être amenés, par imitation d'une prononciation locale, à procéder par "allongement de la syllabe finale [u] en [uə]", ou par "réalisation approximative de [v] en [w]" (2) (2010: 404):

- (2) Holà! Dis-woir... Qu'est-ce que tu lui as fait, samedi, à ma sœur?... Elle me parle bien souvent de toi... il me semble... (< Cougnon)

3.2 *Quelle grammaire pour le texto?*

Un autre préjugé concerne la structuration discursive de la communication par SMS. On pourrait ainsi croire que la façon d'écrire le texto se rapproche des productions orales "relâchées". Ce deuxième préjugé semble encore découler du premier, discuté ci-dessus, du fait que chez beaucoup de locuteurs, sous la pression des institutions sociales et scolaires, la maîtrise de la grammaire va de pair avec la maîtrise de l'orthographe (Béguelin 2012b):

Du côté du grand public scolarisé, notamment dans le monde francophone, il existe ainsi une tendance très nette au "graphocentrisme", caractéristique d'une société à forte tradition écrite; pour bon nombre de ses usagers, le français tend ainsi à être identifié, d'abord et avant tout, à sa forme écrite (2012b: 40)⁶

Si le texto semble approprié à l'usage de formes linguistiques familières ou expressives⁷, cela n'empêche pas que les conditions dans lesquelles sont

⁵ Par la suite, tous les exemples en provenance du corpus suisse seront accompagnés du numéro de SMS dans la base de données sms4science pour la Suisse. Par ailleurs, nous avons conservé dans tous les messages les particularités graphiques d'origine.

⁶ Sur ce point, Béguelin renvoie notamment le lecteur aux réflexions de Blanche-Benveniste (2003: 317; 1997: 11).

⁷ Cela s'explique entre autres par le fait que très souvent les interlocuteurs sont proches psycho-socialement: ils viennent du même milieu social et ont accès aux mêmes références partagées suite aux activités mutuelles (cf. Anis 2007).

produits les SMS diffèrent d'une interaction orale ordinaire. Le SMS relève de la graphie et partant correspond à un "produit fini" (Béguelin & al. 2000: 233); de ce fait, on n'y trouve que rarement les phénomènes de bribes et autres traces d'élaboration du discours en temps réel qui caractérisent l'oral spontané, cf.:

- (3) je suis pas très euh patriote ou nationaliste je suis | je suis je suis je suis genevois
oui parce qu'en fait euh | _ | Ge- Genève c'est un peu | _ | on fait partie de la
Suisse c'est sûr on fait partie euh on est suisse | _ | mais on est quand même un
peu international (OFROM⁸)

Dans l'exemple (3), les scories en question se manifestent sous la forme d'hésitations (*euh*), de répétitions (*je suis je suis je suis genevois*), de tronctions (*Ge- Genève*), ou encore de reformulations (*Genève c'est un peu [...] on est quand même un peu international*) (cf. Blanche-Benveniste & Martin 2011).

Cette divergence entre l'oral spontané et le texto est plus nette encore dans la façon d'exploiter les différentes ressources grammaticales dont certaines sont réputées rares en français "relâché". Ainsi, dans le domaine des temps verbaux, Labeau (2014) a démontré, en se basant sur une étude effectuée à partir de 500 SMS belges, que dans le texto on aurait tendance à employer davantage le futur simple (*je ferai*) que le futur périphrastique (*je vais faire*): "Dans le domaine du futur, la forme synthétique apparaît majoritaire (64,86%), contrairement à ce qui est attesté dans l'oral spontané" (2014: 136).

D'autre part, contrairement à ce qu'on pouvait penser, le texto atteste de l'emploi du passé simple, même si les formes concernées sont plutôt des tournures stabilisées du type "ce fut". Ces emplois ne sont pas uniquement caractéristiques des groupes âgés, mais se rencontrent aussi chez les locuteurs plus jeunes, comme l'atteste le corpus suisse (4)⁹:

- (4) a. F, 19 ans: Bonne nuit na yo!merci pr ce week end,on s'est tro marré et *ce fut* très creu et tt et tt et TOUTOU!mdr! [...] dors bien,je t'aime renaud!ps: osa kitoko na coiffure na yo:) (20682)
b. F, 25: Mais de rien! *Ce fut* un plaisir pour nous aussi. On se réjouit déjà de vous revoir. Bise et bon we. (22308)

Enfin, dans le domaine de l'interrogation, objet de cette étude, la réalisation des formes interrogatives dans le texto n'est pas exactement la même qu'à l'oral spontané. À titre indicatif, le tableau ci-dessous illustre que la réalisation des interrogatives totales est différente dans les deux types de corpus. Le premier constat surprenant, dans le corpus suisse de SMS – si c'est toujours

⁸ Corpus du français parlé en Suisse Romande (Avanzi & al. 2012-2015).

⁹ Nous rencontrons aussi les emplois du passé simple dans le Corpus français "88milSMS" (Panckhurst & al. 2014): "Lo! Je lai pas rendu et *ce fut* HORRIBLE! Jsais pas qui ya demain. Afait le cours daujourdhui a ete annulé je sais pas si je te lai dis x3 bonne nuit et a demaaaaaaain!! :3 (ID 394)".

la variante par maintien de l'ordre assertif qui domine, tout comme dans les corpus oraux – est que la variante par inversion, qui est réputée soutenue, dépasse clairement la variante en "est-ce que". Cette dernière est en revanche plus fréquente à l'oral d'après le corpus de Coveney:

	[SV] <i>Tu vas au cours?</i>	[V-ScI] <i>Vas-tu au cours?</i>	[ESV] <i>Est-ce que tu vas au cours?</i>
Coveney (1996 ¹ =2002 ² : 118): Français parlé	79,4% (143)	-	20,6% (37)
Mosegaard Hansen (2001: 520): Français parlé	86% (86)	-	14% (14)
Corpus suisse de SMS (2009- 2014)	87,2% (1445)	8,3% (138) ¹⁰	4,5% (75)

Fig.2: Réalisation des interrogatives totales dans deux corpus oraux vs un corpus de SMS

L'écart entre le texte et l'oral ordinaire devient encore plus important quand – à côté des structures par inversion simple (5) – on s'intéresse aux interrogatives par inversion complexe (6):

- (5) a. Salut les ami(e)s. *Seriez vous là demain soir pour une bouffe chez nous. J'ai envie de feter avec vous ma récente obtention de mon titre de psychothérapeute.* [...] A+ (9872)
 b. *Emploies tu encore mon vieux natel? Sinon apporte le. Merci* (15382)
 c. *Suis-je un cadeau de Noel suffisant pour toi? :-P* (21925)
- (6) a. *L un de vous aura t il son ordi avec lui? Sinon j prends le mien, pr vous montrer les fantastiques statistiques de notre site :)* (10807)
 b. *Sam vs a-t-il laissés dormir un peu ce matin? Belle journée et gros bisous. Votre mam's qui vs aime* (20411)
 c. *Alors Lyon? Le paradis? Vos postes vous comblent-ils? Avez-vous trouvé vos repères? Des bisous de la Cote d'Az' à tous !* (88milSMS, ID 148)
 d. *Coucou [...], je suis désolée de t'avertir comme ça, au jour le jour, mais j'ai un peu de mal à gérer mon emploi du temps...! Est-ce que t'as pu monter qalif vendredi dernier? Et est-ce que ça t'es possible de le monter aujourd'hui? La carrière est-elle praticable ou pas? Merci... Bisous* (88milSMS, ID 452)

À titre de rappel, l'interrogation par inversion complexe est réputée très formelle ou littéraire, et pratiquement non-existante en français parlé européen (Elsig & Poplack 2006). Pourtant, on la trouve dans le contexte réputé familier du SMS.

Ces observations nous amènent à conclure que, d'une part, la répartition des faits grammaticaux réputés rares dans les discours oral informel¹¹ ne peut

¹⁰ Dont huit sont des structures dites par inversion complexe: *Cela te va-t-il?*

s'expliquer exclusivement via le paradigme "formel/informel". Nous pensons ainsi qu'on aurait tort de ramener les procédés grammaticaux à deux grandes catégories opposées, en leur collant les étiquettes soit de français "standard", soit de français "non-standard" ou "avancé" (Rowlett 2007). Par ailleurs, comme attestent certains exemples, la frontière entre les catégories "standard"/"non-standard" semble être assez floue:

- (7) Yo Baboun, la forme? Jmapprête a te poster ma feuille APG et comme tu me connais g d question. 1) *dois je* a nouveau la filer au SPO ou *puis - je* me contenter de cocher la case "aucune modif par rapport..." 2) si jte l'envoi aujourd'hui *je serai riche quand?* (12578)

Nous voyons ainsi dans (7) que les variantes "standard" (*dois-je...puis-je...?*) vont de pair avec les variantes dites "avancées" (*la forme? je serai riche quand?*).

D'autre part, les différences que nous avons constatées montrent que le recours aux divers procédés grammaticaux n'est pas aléatoire, mais suit certaines tendances, de sorte que des pratiques langagières différentes affichent des préférences différentes dans leur façon d'exploiter la grammaire. Suite aux réflexions de Berrendonner (1987), nous pensons ainsi que les phénomènes variationnels, en tant que diversité de ressources grammaticales mises à disposition du locuteur, peuvent être envisagés en termes de "stratégies d'exploitation" (1987: 49).

4. Quelle approche pour la variation formelle des questions?

Les observations faites jusqu'ici nous amènent à adopter une approche moins traditionnelle selon laquelle le choix des interrogatives ne se fait pas forcément en fonction de la valeur socio-stylistique qui leur correspond, mais est motivé par l'ensemble des facteurs caractéristiques d'une situation énonciative. Ci-dessous, nous exposerons en plusieurs points à quoi pourrait ressembler cette approche.

4.1 Les interrogatives en français sont-elles ou non des variantes?

S'il y a lieu de parler de variation syntaxique dans le système des interrogatives en français, c'est parce que dans chaque type de situation communicationnelle, la grammaire française met à disposition du locuteur une diversité de moyens en vue de produire des énoncés interrogatifs. Ainsi, dans le cas de l'interrogation totale, qui porte sur l'ensemble du contenu propositionnel, nous avons relevé trois variantes dans le corpus suisse de

¹¹ Voir aussi Tagliamonte & Denis (2008) pour l'étude de l'anglais dans le cadre de la messagerie instantanée, où les auteurs constatent qu'à côté des emplois de formes caractéristiques du langage parlé, les locuteurs usent de procédés grammaticaux qui passent pour formels.

SMS (fig. 2). Traditionnellement, ces variantes sont reconnues dans la plupart des études comme trois possibilités de poser la question totale ¹²:

- (8) a. Maintien de l'ordre assertif [SV]:
Yo, *ca roule?* alors 7raclette? [...] *Vs êtes tjs motiv pr sortir, si oui où?* (7367)
b. Inversion du sujet clitique [V-ScI]:
Hello Angèle, *est-tu libre demain midi?* Bonne journée. Sonthaya (12522)
c. "Est-ce que" [ESV]:
Coucou ma belle! Oui on sréjouit tp, c déjà ce we!:-) [...] *Sinon est-ce que vs avez besoin de qqch de Suisse?* Redis moi au cas où! Sinon on se retrouve donc samedi, vers 15h15! Je t'enverrai 1 sms avant! On sréjouit, bisous!:-) (16342)

Concernant l'interrogation partielle, qui porte sur une partie du contenu propositionnel et qui s'appuie sur l'usage de la variable en *Qu- {qui/quoi/quand/où/comment/combien/pourquoi/quel}*, le choix devient encore plus important. Nous avons ainsi relevé dans nos données jusqu'à 7 variantes différentes (voir la fig. 3 pour les pourcentages):

- (9) a. In situ [SVQ]:
alors on sretrouve à quelle h ds les vestiaires? (12662)
b. Antéposition [QSV]:
Comment il s'appelle le groupe autrichien avec la rousse et l' autre type? (20130)
c. Inversion du sujet clitique [QV-ScI]:
Dans quelle direction dois-je me diriger ce soir pr aller boire du thé au beurre? (13230)
d. Postposition du sujet nominal [QV SN]:
Comment s'est passé cette reprise nordique :-)? (23419)
e. Tour interrogatif "est-ce que" [QESV]:
Alor qd S kon fête sa? (11215)
f. Le mot interrogatif est "Sujet" [Q=S V]:
Qui dit mieux? (17428)
g. Clivage [seQkSV]:
sinon c'est qui qui m'amènera de la joie et de la bonne humeur? (20890)

1. [SVQ]	<i>Tu parles de quoi?</i>	53,1% (227)
2. [QSV]	<i>De quoi tu parles?</i>	16,6% (71)
3. [QV-ScI]	<i>De quoi parles-tu?</i>	14,8% (63)
4. [QV SN]	<i>De quoi parle ton frère?</i>	8,2% (35)
5. [QESV]	<i>De quoi est-ce que tu parles?</i>	5,9% (25)
6. [Q=S V]	<i>Qui parle?</i>	0,9% (4)
7. [seQkSV]	<i>C'est qui qui parle?</i>	0,5% (2)

Fig.3: Forme syntaxique des interrogatives partielles dans le corpus suisse de SMS

¹² Par commodité, nous désignerons nos variantes par des abréviations, comme cela se fait chez les autres chercheurs qui ont étudié les structures interrogatives en français (Coveney 1996¹=2002²; Quillard 2000; Mosegaard Hansen 2001; Elsig 2009). Ici, nous reprenons le système d'abréviations utilisé par Coveney (2011).

En ce qui concerne le terme de "variante", il peut être utilisé avec sens différent d'une approche linguistique à l'autre, notamment en fonction des objectifs de programmes de recherche (cf. Jakobson 1989). Ainsi, dans la méthodologie sociolinguistique labovienne les variantes représentent "des façons alternatives de dire la même chose" (notre traduction de l'anglais, cf. Winford 1996: 177). Selon cette conception, les variantes doivent être interchangeables sémantico-pragmatiquement; si elles ne le sont pas, les cas considérés ne sont plus des variantes, mais relèvent des contextes catégoriques¹³.

De notre côté, nous pensons qu'au lieu d'imposer la condition d'interchangeabilité sémantico-pragmatique, il serait plus judicieux d'accorder à la définition de "variantes" plus de souplesse et de traiter les structures en cause, à l'instar de l'approche de Quillard (2000), comme "plusieurs formes en concurrence" (2000: 297), et ce pour plusieurs raisons.

4.2 Approche en termes de concurrence

Il nous semble en effet que l'approche en termes de concurrence présente plusieurs avantages. Premièrement, elle nous permet d'envisager l'étude des variantes dans leur ensemble sans exclure les cas où elles ne sont pas interchangeables en raison des contraintes grammaticales ou sémantico-pragmatiques. Ainsi, si l'on prend l'exemple des interrogatives totales (10-12), nous verrons que leur réalisation fait l'objet de plusieurs types de contraintes, qu'elles soient d'ordre structurel (10), syntactico-sémantique (11), ou pragmatique (12):

- (10) Ça va?/ Est-ce que ça va?/*Va ça?
- (11) a. Tu as faim, je crois?/*As-tu faim, je crois?/*Est-ce que tu as faim, je crois?
(cf. Borillo 1982, Abeillé et al. 2012)
b. Une bêtise il a fait?/*Une bêtise a-t-il fait?/*Une bêtise est-ce qu'il a fait?
(< Rossari & Gachet 2013)¹⁴
- (12) a. Tu coupes un peu le magnétophone?/# Coupes-tu un peu le magnétophone?/#
Est-ce que tu coupes le magnétophone? (< Coveney 1996¹=2002²)
b. A. Je vais à Paris ce mardi.
B. Ah bon? Tu vas à Paris?
B`. Ah bon? # Vas-tu à Paris?/# Est-ce que tu vas à Paris?

¹³ Voir par exemple Guy (2007: 3): "The prevailing consensus is that, while certain structures may have different meanings in some of the contexts they occur in, there are often other contexts in which they function as alternants. Therefore, productive variationist analyses can be conducted, given careful attention to contexts and meaning."

¹⁴ Nous remercions Frédéric Gachet pour les commentaires de (11 b), où l'intonation interrogative porte sur le SN *une bêtise*, alors que *il a fait* est un segment postfocus à contour intonatif réduit. D'où vient cette impossibilité d'employer les structures par inversion ou par "est-ce que" en position post focale.

Deuxièmement, cette approche nous permet de considérer la valeur socio-stylistique directement en lien avec la fonction de la variante (Lavandera 1978: 179; Gadet 1997: 14). Nous pouvons ainsi constater que la notion de valeur socio-stylistique n'est pas en soi absolue, la même variante étant susceptible de recevoir deux interprétations différentes selon telle ou telle situation (cf. Dewaele 1999: 5). En effet, si certains contextes favorisent l'emploi de variantes dites "soutenues" ou "familières", cela n'exclut guère leur apparition dans d'autres situations, comme c'est le cas des emplois des structures inversées dans le SMS. Par exemple, la variante par inversion peut être employée dans un contexte formel de la communication, où elle permet de souligner une distance interpersonnelle entre les locuteurs (dans les écrits SMS ce type d'emploi est plutôt rare):

- (13) a. Bonsoir, des cours sont annulés pour raison d'un conseil de lycée extraordinaire. Je finis donc les cours à 15h15 au lieu de 17h. *Est-il possible d'avancer notre rdv?* Bonne soirée rick (8310)
 b. Bonjour Clerc, *viens-tu au cours de francais aujourd'hui?* Nous n'avons PAS de cours le mercredi 11 novembre (la semaine prochaine). Cordialement. (8155)

Mais employée dans le contexte familial du SMS, elle n'est plus perçue comme "formelle", la graphie non-standard attestant aussi du caractère "familier" de la communication:

- (14) a. Hello il y aura finalement pas de séance au lycée vieux-castel samedi proch. On a réussi à tout planifier samedi. je vous envoie le programme mardi. *par contre pour la séance du matin au lycée éclair as-tu déjà idé d'activité?* A+ (10455)
 b. Hello, comment va? J'espère que ton opération ça va bien parce que ça va pas du tout en pédiatrie comme moi! lol, ça va être joli petit dessin ou mur là-bas! *Peut-être m'indiquer quel tailleur de chemise il te faut pour le scout?* Merci bis a+ (18395)

Cette diversité d'interprétations nous conduit à conclure que la valeur socio-stylistique n'est pas stable en soi. En revanche, c'est le fonctionnement syntactico-sémantique des variantes qui s'avère plus stable (supra 10-12). Nous pensons ainsi qu'il sera plus instructif d'étudier la valeur socio-stylistique en lien avec les propriétés syntactico-sémantiques des variantes (cf. Berrendonner 1988; Gadet 2007; Langacker 2010).

En somme, l'approche que nous préconisons ici insiste sur la nécessité de tenir compte de plusieurs facteurs, de caractère aussi bien grammatical que pragmatique ou encore socio-personnel (cf. Coveney 1996¹=2002²; Quillard 2000; Druetta 2009). Ainsi, sous l'influence de ces facteurs, les emplois des formes en concurrence, ou des variantes, "ne se font pas de manière entièrement aléatoire", au sens de Berrendonner (1987), mais "donnent lieu au développement de stratégies rationnelles et explicables" (1987: 49). À noter qu'il peut y avoir une interaction entre deux ou plusieurs de ces facteurs dans le choix de telle ou telle variante (cf. Coveney 1996¹=2002²). Plus généralement, nous adopterons une perspective fonctionnaliste d'après laquelle le choix entre plusieurs variantes permet au locuteur-scripteur de mieux s'adapter à ses besoins communicatifs: sous la pression de l'ensemble des facteurs constitutifs d'un contexte de communication particulier, il recourt

à une variante plus efficace que les autres afin d'arriver au mieux à ses fins communicatives. De ce point de vue, nous nous rallions aux propos de Berrendonner (1987: 49), Catach (1995: 25), Haspelmath (1999: 186), ou encore Gadet (2007: 163), selon lesquels les phénomènes variationnels peuvent être envisagés respectivement en termes de "stratégies", "palette de possibilités", "adaptation" ou "ajustement dynamique".

4.3 *Étude multifactorielle*

Il est essentiel, en traitant la variation syntaxique dans les structures interrogatives, de tenir compte de la pluralité des contraintes, d'où le choix d'une étude multifactorielle. La tâche consiste alors à dégager les grandes tendances observées dans la distribution des variantes en lien avec les paramètres (i) linguistiques, (ii) communicatifs et (iii) socio-personnels.

Tout d'abord, en lien avec les paramètres linguistiques, il s'agira d'effectuer une étude distributionnelle et de voir comment différents *paramètres morphosyntaxiques* conditionnent l'emploi des structures interrogatives.

En deuxième lieu, les tendances linguistiques établies devront être expliquées qualitativement en lien avec les contraintes communicatives propres à l'interaction par texto, car différents genres de discours n'exploitent pas de la même façon toutes les possibilités de la grammaire (Fillmore 1973: 394; Newmeyer 2010: 32).

Enfin, du point de vue sociolinguistique, il s'agira d'examiner dans quelle mesure il peut y avoir une corrélation entre le choix de telle ou telle variante et d'éventuelles réalités d'ordre socio-personnel. Ainsi, nos premières observations montrent que l'usage de l'inversion semble progresser avec l'âge. À ce propos, nous renvoyons aussi aux observations de Quillard qui a étudié les emplois des interrogatives dans les corpus oraux. Elle montre ainsi que le modèle par inversion aurait plus de chances d'être employé par des locuteurs de plus de 35 ans et de la catégorie socio-professionnelle "haute" (2000: 295-296). De ce point de vue, il serait aussi intéressant de voir si ces observations sont valides dans le contexte de communication suisse. D'autre part, dans les cas impliquant l'emploi des verbes modaux, il arrive qu'un scripteur choisisse une variante (par exemple, *peux-tu...?* vs *tu peux...?*) et l'applique à de nombreuses situations prototypiques. Cela peut s'expliquer par des ritualisations formulaires, des habitudes ou des préférences personnelles (cf. Quillard 2001, Guryev 2013).

Le besoin de recourir à un modèle multifactoriel dans le traitement des structures interrogatives se justifie entre autres par le fait que la sélection d'un paramètre unique risquerait de négliger des observations importantes faites par d'autres linguistes et de rencontrer par là ses limites.

Dans la suite de l'article nous nous centrerons exclusivement sur l'étude des paramètres linguistiques, ou morphosyntaxiques. Nous observerons quelques tendances propres à la réalisation des variantes dans le cas de l'interrogation totale dans le corpus suisse de SMS.

5. L'annotation des paramètres morphosyntaxiques et ses enseignements

L'approche que nous proposons nécessite une annotation méthodique des données, ce qui nous permettra par la suite de dégager quelques tendances linguistiques dans la réalisation des interrogatives totales dans les données SMS. En outre, nous mettrons en évidence la pertinence du concept de *configuration syntaxique*, car nos données attestent que certains contextes linguistiques s'avèrent plus propices que d'autres à l'alternance entre les variantes.

5.1 Annotation des paramètres morphosyntaxiques et traitement des données

Le codage a été effectué via le logiciel MMAX2, conçu spécialement en vue d'annoter les données linguistiques à des niveaux multiples (Müller & Strube 2006)¹⁵. Nous avons annoté deux types de variables:

- (i) Au premier niveau ont été annotées les variantes de structures interrogatives, qui constituent la variable dépendante, c'est-à-dire l'objet de recherche;
- (ii) Au deuxième niveau ont été annotés tous les facteurs que nous avons estimés sur la base des travaux antérieurs, susceptibles d'influencer la production des variantes interrogatives et qui correspondent à la variable indépendante.

S'agissant du premier niveau, il a d'abord fallu définir tous les types d'interrogatives attestés dans nos données puis les intégrer dans MMAX2 sous forme de schémas d'annotation. Ces schémas ont été établis suite à des observations initiales sur les données, afin de recouvrir les différents types d'interrogatives relevés dans le corpus SMS. Ils ont été testés à l'aide d'un travail d'annotation préalable portant sur cent, trois cents puis cinq cents

¹⁵ Dans le cadre du projet *sms4science* en Suisse, qui regroupe plusieurs universités partenaires, le programme d'annotation MMAX2 a été choisi en vue d'un partage ultérieur des données, au travers du programme ANNIS (Zeldes & al. 2009), lequel permet des requêtes sur des données annotées à des niveaux multiples.

messages SMS. Une case "autres questions"¹⁶ a été utilisée quand aucune des catégories établies ne donnait satisfaction (fig. 4):

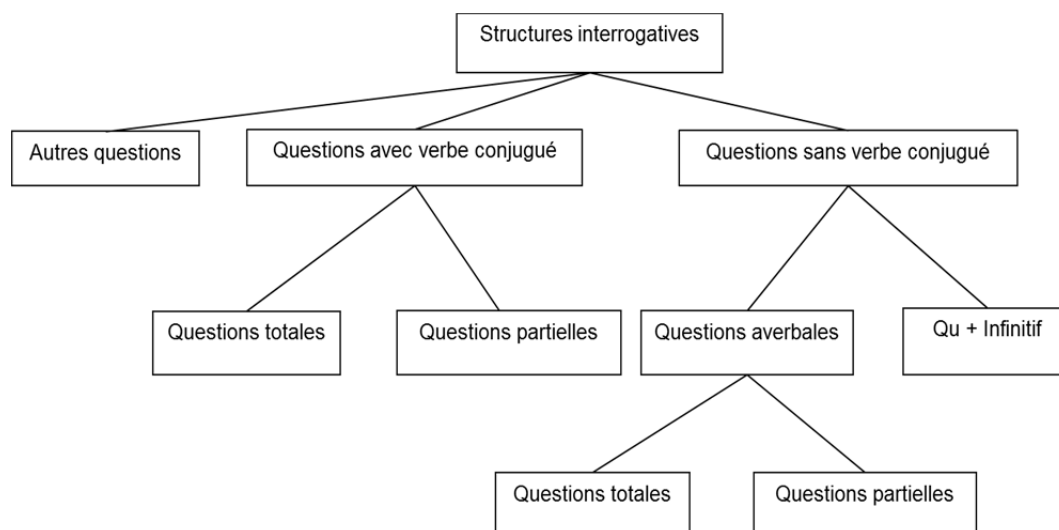


Fig.4: Niveau 1 - Types de structures interrogatives

Bien que les schémas cherchent à documenter les différents types d'interrogatives attestés dans le corpus étudié, le but ultime du travail, en accord avec l'état de l'art dans le domaine, était de procéder à l'analyse quantitative et qualitative des questions à verbe conjugué, qui étaient réparties en deux grandes classes: (i) les questions totales (supra 8), qui comptent trois variantes dans notre corpus, et (ii) les questions partielles (supra 9), qui comptent dans notre cas jusqu'à 7 variantes. Enfin, ces questions ont aussi été annotées selon qu'elles étaient marquées ou non par un point d'interrogation et selon qu'elles étaient coordonnées ou non à d'autres questions.

Quant au deuxième niveau, le travail a consisté à annoter les différents paramètres morphosyntaxiques caractérisant l'occurrence des énoncés interrogatifs à verbe conjugué, ceux-ci constituant l'objet principal de l'étude. Le cadre syntaxique adopté se rattache, *mutatis mutandis*, aux conceptions de Blanche-Benveniste, pour laquelle il est primordial de commencer l'analyse linguistique des énoncés par le verbe, qui est "le principe organisateur, avec son sujet et ses compléments, régis par ce verbe" (1990: 19, cf. 2013: 129-138). Dans cette perspective, nous avons été amenés à documenter les relations entre le verbe de l'interrogative et d'autres constituants à caractère

¹⁶ Pour donner quelques exemples de ce qui a été rangé dans cette catégorie, nous citerons les questions en "si" *Et si on allait au cinéma?* (16936), différents cas d'actes interrogatifs indirects: *j'espère que ça va bien en Allemagne !?* (11542), *alors redis moi si t'as besoin d'un coup de main* (20893), *Alors il paraît que ça va mieux?* (17316), etc., mais aussi des cas d'alternances codiques: *Toi ça gehts ou bien?* (21841). À noter que 174 questions de notre corpus ont été attribuées à la catégorie "autres questions".

obligatoire ou accessoire, de sorte que l'énoncé interrogatif *Et le texte de Labov, on l'aura quand environ? !* (20354) a été codé comme suit:

- (15) [Et]: Renforceur
 [le texte de Labov]: Constituant disloqué
 [on]: Sujet
 [le]: Complément clitique
 [aura]: Verbe
 [quand]: Mot interrogatif
 [environ]: Complément circonstanciel

Nous avons ainsi distingué neuf éléments différents, verbe compris, dont certains ont été identifiés à un degré ou à un autre dans les travaux de référence (Coveney 1996¹=2002²; Quillard 2000; Druetta 2009; Elsig 2009) (fig. 5):

[Sujet] [Complément clitique] [VERBE] [Complément régime] [Complément circonstanciel]¹⁷ [Mot interrogatif] [P subordonnée] [Constituant disloqué et/ou extra-prédicatif] [Renforceur]¹⁸

Fig.5: Niveau 2 - Paramètres morphosyntaxiques pertinents pour les questions avec verbe conjugué

Ensuite, chacun de ces constituants a été détaillé selon la complexité paramétrique attestée. Par exemple, dans le cas du verbe, nous prenons en compte (i) son type de diathèse (Berrendonner 2011): construction active, passive, impersonnelle, pronominale, causative en [*faire* + V^{Inf}]; (ii) son temps: présent, passé composé, futur périphrastique, futur simple, imparfait, plus-que-parfait, conditionnel, autre temps; (iii) le nombre de syllabes de la forme finie: 1-2 syllabes, 3 syllabes et plus (Elsig 2009: 63); et (iv) sa modalité: modal (*pouvoir, vouloir, devoir*, autre verbe) et non modal, cette dernière catégorie étant encore répartie en constructions à un seul verbe fini (*Tu vas au cours?*), et en celles à plusieurs verbes, où la forme verbale finie est suivie d'une forme infinitive (*Tu vas lire?*).

L'annotation des interrogatives dans le corpus suisse de SMS a en tout dégagé 2'677 occurrences de structures interrogatives au premier niveau, et 7'488 unités morphosyntaxiques au deuxième. Pour être exploitées, les données annotées ont été exportées depuis MMAX2 sur Excel (fig. 6)¹⁹:

¹⁷ À noter que dans le cas des catégories [Complément régime] et [Complément circonstanciel], il s'agissait des compléments autres que clitiques.

¹⁸ Concernant le terme de "renforceur", nous nous référons à Druetta (2009).

¹⁹ Certaines colonnes ont été masquées à des fins d'illustration.

AO11		fx		.redouble							
	A	B	C	D	E	F	K	AI	AJ		
1	sms_ID	phrase_texte	st	â	car	l	p_1_verbale	e_element	e_1_su		
2	.8314	Tu veux venir avec moi	631	16	2000	M	.totale	.sujet.verbe.compL_circ	.clit		
3	.9210	Alors tu viens faire ce massage ?	727	47	2400	F	.totale	.compL_pro.verbe.sujet.renforc	.clit		
4	.8310	Est - il possible d' avancer notre rdv ?	631	16	2000	M	.totale	.compL_pro.sujet.verbe	.clit		
5	.9281	Ça joue pr toi ?	736	21	1747	F	.totale	.sujet.compL_circ.verbe	.clit		
6	.11102	Ca va ?	810	24	1005	M	.totale	.sujet.verbe	.clit		
7	.11092	alors et bien ce wknd à ATL ?	810	24	1005	M	.totale	.compL_pro.sujet.renforc.verbe	.clit		
8	.11108	ca va ?	810	24	1005	M	.totale	.sujet.verbe	.clit		
9	.11102	Le médecin a dit koi ?	810	24	1005	M	.partielle	.sujet.verbe.mot_int	.sn		
10	.7982	si tu viens a l'entraînement ce soir , pourrais je venir avec toi ?	643	28	1700	M	.totale	.sujet.verbe.disloc.compL_circ	.clit		
11	.11126	ca va ta jambe ?	735	22	2034	F	.totale	.verbe.sujet	.clit		
12	.9237	Ça tient toujours ?	698	18	2000	F	.totale	.adverbe.verbe.sujet	.clit		
13	.11108	lao ou piscine ca te dit ?	810	24	1005	M	.totale	.sujet.compL_clit.verbe	.clit		
14	.7983	comment vas - tu ?	643	28	1700	M	.partielle	.mot_int.sujet.verbe	.clit		
15	.11129	Viens tu à la nuit des contes ?	719	46	1965	F	.totale	.verbe.sujet.compL_pro	.clit		
16	.11202	Tu pourrais me tel ou tel à fabrice rapidos ?	749	27	2000	F	.totale	.sujet.verbe.adverbe.compL_clit.compL_pro	.clit		
17	.11208	Tu seras au chauffage dans 1 heure ?	833	38	2000	F	.totale	.sujet.verbe.compL_circ.compL_circ	.clit		
18	.11215	Alor qd S kon fête sa ?	836	19	2000	F	.partielle	.compL_pro.renforc.mot_int.verbe.sujet	.clit		
19	.11215	di on se voi jeudi ?	836	19	2000	F	.totale	.compL_circ.renforc.sujet.verbe	.clit		
20	.8305	T' es où ?	605	24	2000	M	.partielle	.sujet.verbe	.clit		
21	.9237	Tu pourras venir me chercher après la pièce de théâtre :)	698	18	2000	F	.totale	.compL_circ.compL_clit.sujet.verbe	.clit		
22	.8305	Y a moyen de se joindre ce soir ?	605	24	2000	M	.totale	.sujet.verbe.compL_pro	.clit		
24	.11221	sava ?	836	19	2000	F	.totale	.verbe.sujet	.clit		
25	.7666	Peut - tu me confirmer que tu l' a bien reçu ?	527	22	8600	F	.totale	.que_sip.sujet.verbe.compL_clit	.clit		
26	.11217	omt ca va ?	836	19	2000	F	.partielle	.verbe.mot_int.sujet	.clit		

Fig.6: Annotations réorganisées sur Excel

La réorganisation des données annotées sous forme de tableau Excel nous a permis d'atteindre deux objectifs: non seulement nous disposons du nombre et du type précis des éléments qui composent les structures interrogatives, ce qui permet de répondre aux besoins d'une analyse quantitative; mais nous pouvons en même temps accéder au texte complet de chaque structure, ce qui facilite considérablement l'analyse qualitative.

Ce tableau contient notamment une colonne (colonne B) qui fournit l'énoncé interrogatif, et une seconde colonne (colonne AI) qui fournit la liste exhaustive des éléments qui constituent cet énoncé. Le détail des niveaux d'annotation apparaît dans les autres colonnes: ainsi, pour le premier niveau, les différentes variantes de la variable dépendante, ou encore les types de structure interrogative (colonnes K et suivantes) et, pour le second niveau, les différents facteurs de la variable indépendante, ou paramètres morphosyntaxiques (colonnes AJ et suivantes). Enfin, un type d'information nous a été fourni directement par sms4science: au travers de leur numéro d'identification, les messages (colonne A) donnaient accès à leur locuteur (colonne C) et aux informations sociodémographiques qui le concernent (colonnes D et suivantes).

5.2 Des paramètres morphosyntaxiques à un type de configuration syntaxique

Comme nous l'avons vu plus haut, la procédure d'annotation adoptée nous permet, d'une part, de préciser la forme sous laquelle a été réalisée la question, et, d'autre part, de documenter les paramètres morphosyntaxiques caractéristiques de l'énoncé en question. À titre de rappel, dans le cadre du traitement de l'énoncé interrogatif suivant: *Dis-moi...Qu'est-ce que tu lui as fait, samedi, à ma soeur?* (11787), nous retrouvons deux sortes d'informations

désignant la variable dépendante (niveau 1) et la variable indépendante (niveau 2):

- (16) Niveau 1: Question verbale > Question partielle > Variante par 'est-ce que' > ponctuée
- Niveau 2: Sujet [tu]: clitique, 2e personne, singulier
 Verbe [as fait]: voix active, passé composé, forme finie à une syllabe, non-modal, construction à un seul verbe fini
 Complément [lui]: clitique, 3e personne, singulier, objet indirect, datif
 Mot interrogatif [que]: objet direct
 Complément circonstanciel [samedi]: référence temporelle, adverbe
 Constituant disloqué [à ma sœur]: dislocation à droite, objet retardé
 Renforçateur [dis-voir]: position initiale

Le but ultime de notre analyse étant d'étudier l'impact des facteurs linguistiques (niveau 2) sur la sélection des variantes (niveau 1), la prochaine étape consiste à modéliser les contraintes linguistiques, avec une préférence pour l'usage du concept de *configuration syntaxique* qui désigne une combinaison quelconque des paramètres morphosyntaxiques constitutifs d'un énoncé interrogatif. Par exemple, dans le cas de la question *Tu y es?*, la configuration syntaxique sera schématisée comme suit:

- (17) [Sujet clitique] + [Complément clitique] + [Verbe]
 Tu *y* *es*

La pertinence du concept de configuration syntaxique pour notre analyse s'explique avant tout par notre préoccupation d'aborder l'étude des contraintes linguistiques sous un angle large intégrant dans l'analyse plusieurs facteurs linguistiques en même temps, au lieu de traiter séparément de l'incidence de chaque facteur linguistique sur la réalisation des variantes (voir Guryev in prep. pour plus de détails). Plus précisément, cette analyse, qui est au cœur de l'approche proposée, demande d'inventorier plusieurs types de configurations et de procéder à leur tri selon deux types *d'environnement morphosyntaxique*:

- (i) Environnement morphosyntaxique à variabilité faible, regroupant toutes les configurations syntaxiques qui favorisent l'emploi d'une variante au détriment des autres;
- (ii) Environnement morphosyntaxique à variabilité remarquable, regroupant toutes les configurations syntaxiques qui favorisent une alternance entre les variantes.

Suite à cette distinction, nous verrons ci-dessous quelques tendances dans la sélection des interrogatives totales en fonction des types de configurations syntaxiques où elles figurent.

5.3 Quelques tendances dans la sélection des interrogatives totales dans le corpus suisse de SMS

L'analyse de l'incidence des paramètres morphosyntaxiques sur la réalisation des variantes dans le cas de l'interrogation totale (1'658 occurrences) montre clairement que la sélection des variantes, que nous désignerons désormais par [SV] *Tu viens au cours?* (1'445 occurrences), [V-Scl] *Viens-tu au cours?* (138 occurrences) et [ESV] *Est-ce que tu viens au cours?* (75 occurrences), ne se fait pas aléatoirement mais qu'elle est soumise à une certaine logique.

En premier lieu, nous avons constaté que certains contextes grammaticaux n'admettent pas toutes les variantes, l'inversion n'étant pas possible dans les configurations à sujet pronominal *ça* (supra 10). De plus, les configurations de ce type privilégient largement l'emploi de [SV]: dans 387 cas sur 399 ces constructions ont été réalisées avec cette variante, et seulement 12 avec [ESV]. En raison de ce comportement, nous n'avons pas tenu compte des constructions en *ça* dans l'analyse ultérieure.

Dans un deuxième temps, nous pouvons constater que la réalisation des variantes est sensible au type de construction verbale:

- Constructions à un seul verbe fini [V1];
- Constructions à un verbe tensé suivi d'un infinitif [V+Inf].

Le test de chi-carré (χ^2) a ainsi montré que la différence dans le comportement des constructions [V1] vs [V+Inf] est significative ($\chi^2=53,802$ $df=2$ $p < 0.00001$) (fig. 7):

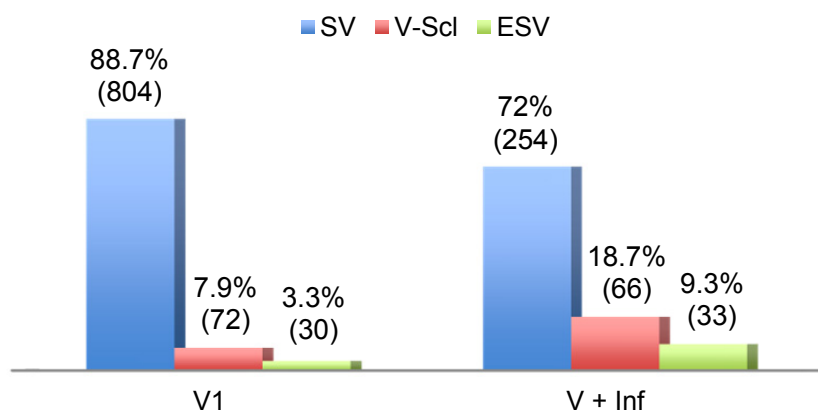


Fig.7: Construction verbale simple vs suivie d'un infinitif

Il s'avère ainsi que les constructions [V+Inf] favorisent davantage l'alternance entre les trois variantes:

- (18) a. tu veux pas prendre le suivant?? Qu' on puisse parler ... avant d' arriver au manège ... (20426)
 b. Ta soeur viens mangé avec nous? (17740)
 c. dois-je a nouveau la filer au SPO ou puis-je me contenter de cocher la case aucune modif par rapport ... (10152)
 d. Viens tu t'entraîner ce soir? (11270)
 e. penses - tu venir au Réseau de ton patient cet a-m? (17027)

- f. est-ce que je pourrais passer vite chez toi pour te poser 2 petites questions?
(13397)
g. est-ce que tu vas skier aujourd ' hui? (12530)

À noter aussi que dans ces constructions, nous ne retrouvons pas seulement les verbes typiquement reconnus comme modaux (18 a, c, f), quand bien même ils restent majoritaires, mais aussi d'autres verbes (18 b, d, e, g).

En ce qui concerne les constructions [V1], elles sont de loin réalisées avec [SV]:

- (19) a. t' es en t-shirt dehors !? ! (8531)
b. alors t es toute bien installée? (13092)
c. Les cours de renforcement en infocom ils se compensent , comme au sein d' un module? (16330)

Il faut pourtant admettre que si à la base les constructions [V1] sont moins propices à la variabilité dans l'usage des interrogatives totales, elles peuvent compter des configurations qui facilitent l'emploi des variantes autres que [SV]:

- (20) a. As - tu déjà 1 tache? (9193)
b. Est-c ' que par l' plus grand des hasard tu prends l' train d' 20h?? (22701)
c. est-ce que tu aurais le num de véra? pcq le matel de Benj est provisoirement bloqué (8316)

En effet, nous verrons par la suite que les constructions [V1] peuvent à leur tour être subdivisées en deux types de configurations syntaxiques, selon la variabilité (faible vs remarquable) qu'elles attestent dans la sélection des variantes. Nous avons encore procédé à la sélection, parmi tous les paramètres morphosyntaxiques, de ceux qui se sont révélés importants dans le cas de la réalisation des interrogatives totales.

Parmi les configurations syntaxiques qui concernent les constructions [V1] et qui ont un faible potentiel de déclencher la variabilité dans l'emploi des interrogatives, nous citerons les deux configurations suivantes:

- (21) [Sujet clitique] + [Verbe]:
Tu viens/Viens-tu/Est-ce que tu viens?
- (22) [Sujet clitique] + [Complément clitique] + [Verbe]:
Tu les prends/Les prends-tu/Est-ce que tu les prends?

Ces configurations sélectionnent le plus souvent la variante à maintien de l'ordre assertif [SV] (259 cas sur 265), alors que les emplois de [V-ScI] et [ESV] sont marginaux. Voici quelques exemples de ces emplois:

- (23) a. Yo madame!qu'est-ce qui se passe?tu es malade?tu déprime? (10763)
b. Haha!elle à fait exprès?elle à aimé?miam miam,ca me donne l'eau à la bouche!lol!elle à pris à quoi?becs,bon aprem,jvm fort (11566)
c. Trop cool :) vous avez gagner hier? Le rdv est à quel heure ce soir? J'ai déjà oublierxD bisous à ce soir alors. (19177)
d. Coucou! Tu as fini? Tu prends le train de quelle heure? A toute! (13309)
e. Il n y a pas de message - tu es sorti? (18188)

- (24) a. Chai pas...*t'y va toi?* (11579)
 b. *Tu m'abandonnes?* :-((21839)
 c. Oh non mince!!C vmt pas sympa d'avoir dit ca devant lui!! *t'en a parlé du coup??*T parents ont surement peur pr toi... [...] (21720)
 d. Je v à l ancienne gare. *T y es?* (19483)
 e. *Tu t' arretes à Lsne après tes cours mercredi?* (10687)

Enfin, parmi les configurations syntaxiques qui relèvent des constructions [V1] à fort potentiel de variabilité, nous citerons les suivantes:

- (25) [Sujet clitique] + [Verbe] + [Complément régime SN/Pro SN]:
Tu as des nouvelles/As-tu des nouvelles/Est-ce que tu as des nouvelles?
- (26) [Sujet SN/Pro SN] + [Verbe]:
Ton frère va mieux/Ton frère va-t-il mieux/Est-ce que ton frère va mieux?

En effet, les configurations syntaxiques (25-26) font preuve d'une certaine neutralité quant à la sélection des variantes. Quand bien même [SV] reste hors concurrence, les variantes [V-ScI] et [ESV] voient leurs chances significativement augmenter: autour de 94% de leurs emplois dans les constructions à un seul verbe fini, soit 96 occurrences sur 102, se font avec les configurations syntaxiques de ce type:

- (27) a. As tu du temps après midi pour le visionner chez nous? (12274)
 b. Emploies tu encore mon vieux natel? (15382)
 c. marches - tu à Berne mardi prochain (10875)
 d. Suis - je un cadeau de Noel suffisant pour toi? (21925)
 e. Yo man,yep ca me dirait bien.Rlini frogeu:est ce un souper entre hommes?javais prévu de sortir avec LA... (17732)
 f. est-ce que tu aurais un moment pour boire un verre histoire que je me fasse pardonner la dernière fois? (21131)
 g. et est ce que y a un code d entrée? (9433)
 h. Est-ce que tu es finalement quand meme allée chez Sibylle?? (13764)
- (28) a. Mon ... chauffeur va-t-il mieux? (15045)
 b. Sam vs a-t-il laissés dormir un peu ce matin? (20411)
 c. L un de vous aura t il son ordi avec lui? (10807)
 d. Est-ce que demain aprem 13h - 18h t' irait pr kekea? (10686)
 e. et eske le matelat est au sol? (14962)
 f. Est-ce que clau a des envies pour noël? (20700)
 g. Est ce que alexis et marcus ont rompu??? (23049)

Ce qui est encore intéressant dans ces observations, c'est que l'écart entre les variantes [V-ScI]/[ESV] et leur concurrente [SV] devient moins important dans la configuration avec sujet SN/Pro SN (*Ton frère travaille?*) (28), comparé à leurs emplois dans les configurations avec sujet clitique (*Il travaille?*) (fig. 8):

	[SV]	[V-ScI]/[ESV]
Configuration avec sujet clitique	89.8% (769)	10.2% (87)
Configuration avec sujet SN/Pro SN	70% (35)	30% (15)

Fig.8: Emploi des variantes dans les constructions [V1] avec sujet clitique vs sujet SN/Pro SN

Le test de chi-carré a aussi montré que la différence entre les emplois des variantes avec différent type de sujet (clitique vs non clitique) est significative: $\chi^2=18,606$ $df=1$ $p < 0.00001$ ²⁰.

Ces résultats pourraient certes être plus amplement discutés. Néanmoins, ils permettent de conclure qu'il est nécessaire de tenir compte de l'influence des facteurs linguistiques dans l'étude des interrogatives. Nous avons vu en effet que la sélection des variantes ne se fait pas d'une façon aléatoire, mais suit régulièrement certaines tendances²¹. Ces tendances peuvent être saisies sous forme de liens connexes entre l'ensemble des paramètres morphosyntaxiques de l'énoncé interrogatif et le choix d'une forme interrogative, ou variante, sous laquelle est réalisée la question.

6. Conclusions

À travers cet article, nous avons montré que l'étude des interrogatives en tant que phénomène variationnel nécessitait l'analyse approfondie de plusieurs composantes. Nous préconisons donc un modèle multifactoriel prenant en compte la pluralité des facteurs caractéristiques de la situation d'énonciation, en particulier (i) les paramètres linguistiques, (ii) les paramètres communicatifs, (iii) et les paramètres socio-personnels. Sous cet angle, nous ne considérons pas les variantes syntaxiques comme des formes sémantico-pragmatiquement interchangeable, opposées entre elles par leurs valeurs socio-stylistiques. En effet, pour nous, chaque variante est dotée de ses propres spécificités fonctionnelles. Nous pensons notamment que le choix des variantes, qui sont en concurrence entre elles (Quillard 2000), ne se fait pas aléatoirement mais est motivé par l'ensemble des facteurs constitutifs d'un contexte de communication particulier. Nous avons ainsi vu au cours de cet article qu'il y a plusieurs arguments en faveur de cette approche.

Tout d'abord, l'étude des données SMS montre que l'hypothèse classique, selon laquelle le choix des variantes se fait en accord avec la valeur socio-stylistique qui leur est propre, ne permet pas toujours de rendre compte de la réalité. D'une part, dans nos données, les variantes par inversion, dites formelles ou soutenues, s'emploient dans le cadre informel du texto et peuvent

²⁰ Notons aussi que dans le cas des configurations à deux arguments nominaux, dont le deuxième est un objet direct: "*L un de vous aura t il son ordi avec lui?*" (10807), la variante [SV] n'est plus dominante dans notre corpus: 3 occurrences de [V-ScI]/[ESV] vs 2 occurrences de [SV]. Ces observations donnent à penser que de manière générale, les configurations syntaxiques avec arguments autres que clitiques constituent un milieu favorable à l'alternance entre les variantes. Gardons-nous toutefois de généraliser ces observations vu la faible représentation de ces cas.

²¹ Là encore, en fonction des genres de discours et des types d'interaction, les tendances dans la sélection des variantes seront sujettes à variation.

être coprésentes avec les variantes à maintien de l'ordre canonique SV, dites familières. D'autre part, plusieurs indices montrent que la valeur socio-stylistique des variantes, même s'il y a lieu dans certains cas de parler d'une opposition entre emplois "familiers" vs "soutenus" vs "neutres", n'est pas arbitraire, mais qu'elle est motivée par les propriétés sémantico-pragmatiques des variantes qui résultent entre autre des spécificités structurelles de leurs formes (Mosegaard Hansen 2001).

Enfin, l'autre argument majeur en faveur d'une approche multifactorielle est que, comme le montre l'analyse du corpus suisse de SMS, la sélection des variantes se fait selon le type de configurations syntaxiques. Ces configurations, en tant que combinaisons différentes de paramètres morphosyntaxiques, relèvent de deux types d'environnement: (i) environnement à variabilité faible, sélectionnant d'une façon discriminatoire une seule variante parmi plusieurs, et (ii) environnement à variabilité remarquable, propice à l'alternance entre plusieurs variantes. Soulignons toutefois que les tendances linguistiques observées demanderont à être expliquées qualitativement, en lien avec les contraintes de la communication par SMS, différents genres de discours exploitant différemment les ressources de la grammaire.

À l'avenir, il conviendra de comparer les tendances relevées dans le corpus suisse de SMS avec ce qui se passe dans d'autres corpus, afin de voir dans quelle mesure elles sont extrapolables. Ceci nous aidera à mieux comprendre la nature des phénomènes variationnels, qui ne sont autre chose, selon nous, que des "stratégies d'exploitation" de structures appartenant à une même grammaire (cf. Berrendonner 1987).

Bibliographie

- Abeillé, A., Crabbé, B., Godard, D. & Marandin, J.-M. (2012): French questioning declaratives: a corpus study. In: *Proceedings of SemDial 2012 (SeineDial): The 16th Workshop on the Semantics and Pragmatics of Dialogue*. Université Paris-Diderot (Paris 7), Paris Sorbonne-Cité, September 2012.
- Adli, A. (2006): French wh-in-situ Questions and Syntactic Optionality: Evidence from Three Data Types. In: *Zeitschrift für Sprachwissenschaft*, 25, 163-203.
- (2013): Syntactic variation in French Wh-questions: A quantitative study from the angle of Bourdieu's sociocultural theory. In: *Linguistics*, 51(3), 473-515.
- Anis, J. (2007): Neography – Unconventional Spelling in French SMS Text Messages. In: Danet, B. & Herring, S. C. (éds.), *The Multilingual Internet – Language, Culture and Communication Online*. New York (Oxford University Press), 87-115.
- Avanzi, M., Béguelin, M.-J. & Diémoz, F. (2012-2015): Présentation du corpus OFROM – corpus oral de français de Suisse romande. Université de Neuchâtel, <http://www.unine.ch/ofrom>.
- Béguelin, M.-J. (dir.), avec Matthey, M., Bronckart, J.P. & Canelas, S. (2000): *De la phrase aux énoncés. Grammaire scolaire et descriptions linguistiques*. Bruxelles (De Boeck-Duculot).

- Béguelin, M.-J. (2012 a): La variation graphique dans le corpus suisse de SMS en français. In: Caddéo, S., Roubaud, M.-N., Rouquier, M. & Sabio, F. (éds.): *Penser les langues avec Claire Blanche-Benveniste*. Presses de l'Université de Provence, 47-63.
- (2012 b): Le statut de l'écriture. In: Ruggero Druetta (éd.). *Claire Blanche-Benveniste (1935-2010). A l'école de l'oral*. Sylvain-les-Moulins: Editions GERFLINT, (Collection Essais Francophones, no1), 39-54.
- Bernicot, J., Volckaert-Legrier, O., Goumi, A. & Bert-Erboul, A. (2012): Forms and functions of SMS messages: A study of variations in a corpus written by adolescents. In: *Journal of Pragmatics*, 44.
- Berrendonner, A. (1987). Stratégies morpho-syntaxiques et argumentatives. In: *Protée* 15/3, 48-58.
- (1988): Normes et variations. In: Schoeni, G., Bronckart, J.-P. & Perrenoud, P. (éds.), *La langue française est-elle gouvernable?* Neuchâtel/Paris (Delachaux & Niestlé), 43-62.
- (2004): Grammaire de l'écrit vs grammaire de l'oral: le jeu des composantes micro-et macro-syntaxiques. In: *Interactions orales en contexte didactique*, 249-262.
- (2005): Questions et mémoire discursive. In: Rossari, C., Beaulieu-Masson, A., Cojocariu, C. & Razgoulieva, A. (éds), *Les Etats de la question*. Editions Nota Bene, 147-173.
- (2011): *Rudiment de Grammaire Française* (support de cours, non publié), Université de Fribourg.
- Bieswanger, M. (2007): 2 abbrvi8 or not 2 abbreui8: A contrastive analysis of different shortening strategies in English and German text messages. In: *SALSA XIV*.
- Blanche-Benveniste, C. (1990): Un modèle d'analyse syntaxique 'en grilles' pour les productions orales. In: *Anuario de psicología*, 47, 11-28.
- (1997): *Approches de la langue parlée en français*. Ophrys, Paris.
- (2003): La langue parlée. In: *Le grand livre de la langue française*. Paris (Seuil), 317-344.
- (2013): De la nécessité de commencer l'étude de la syntaxe par le verbe et non par la phrase: de la nécessité d'étudier les constructions verbales avec des classifieurs. Manuscrit du 25 avril 1977, K.U. Leuven, Département Linguistick. In: *Tranel*, 58, 129-138.
- Blanche-Benveniste, C. & Martin, P. (2011): Structuration prosodique, dernière réorganisation avant énonciation. In: *Langue française*, 170 (2), 127-142.
- Borillo, A. (1982): Deux aspects de la modalité assertive: croire et savoir. In: *Langages*, 67, 33-53.
- Catach, N. (1990): Le problème des variantes graphiques: variantes du passé, du présent et de l'avenir. In: *Langue française*, 108/1, 25-32.
- Cedergen, H. J. & Sankoff, D. (1974): Variable rules: Performance as a statistical reflection of competence. In: *Language*, 50.2 (86), 333-355.
- Cougnon, L.-A. (2010): Orthographe et langue dans les SMS. In: *Études de linguistique appliquée*, 4, 397-410.
- Cougnon, L.-A. & François, T. (2010): Quelques contributions des statistiques à l'analyse sociolinguistique d'un corpus de SMS. In: *Actes du colloque JADT 2010*, Vol. 1, 619-630.
- Coveney, A. (2002): *Variability in Spoken French: interrogation and negation*. Bristol (Intellect Books).
- (2011): L'interrogation directe. In: *Travaux de linguistique*, 63, 112-145.
- Danjou-Flaux, N. & Dessaux, A.-M. (1976): L'interrogation en français: Données linguistiques et traitements transformationnels. In: *Grammaire transformationnelle: syntaxe et lexique*. Études réunies par Jean-Claude Chevalier. Publications de l'Université de Lille, 139-231.
- Dewaele, J.-M. (1999): Word order variation in interrogative structures of native and non-native French. In: *International Journal of Applied Linguistics*, 123, 161-180.
- Druetta, R. (2009): *La question en français parlé: étude distributionnelle*. Trauben (Edizioni Torino).
- Elsig, M. (2009): *Grammatical variation across space and time: the French interrogative system*. Amsterdam (John Benjamins).
- Elsig, M. & Poplack, S. (2006): Transplanted dialects and language change: Question formation in Québec. In: *University of Pennsylvania Working Papers in Linguistics*, 12(2), 8.

- Fairon, C. & Paumier, S. (2007): Un corpus SMS est-il un corpus comme les autres. In: Camugli Gal-Lardo, C., Constant, M. & Dister, A. (éds.), 209-216.
- Fillmore, C. J. (1973): Pragmatics and the description of discourse. In: Kasher, A. (éd.), (1998): Pragmatics: Critical concepts. London (Routledge).
- Frei, H. (2003): La grammaire des fautes. Rennes. La première édition en 1929.
- Gadet, F. (1989): Le Français ordinaire. Paris (Colin).
- (1997): La variation, plus qu'une écume. La variation en syntaxe. In: *Langue française*, 115, 5-17.
- (2007): La Variation Sociale en Français. Collection l'Essentiel Français. Paris (Ophrys).
- Guryev, A. (2013): Comment traiter la variation dans la communication par SMS? Le cas de l'interrogation totale. In: *JéTou 2013. Variation et variabilité dans les sciences du langage: analyser, mesurer, contextualiser. Actes de la 4ème édition des JéTou 16 et 17 mai 2013*, Toulouse, France. Toulouse: Université de Toulouse II - Le Mirail, 76-87.
- (in prep.): "Suis-je un cadeau de Noël suffisant pour toi?": quelques enseignements du SMS à propos de la variation dans les interrogatives totales. In: Sélection d'articles rédigés dans la foulée du Colloque "La syntaxe des interrogatives en français", Université de Neuchâtel, 13-14.06.2013.
- Guy, G. R. (2007): Grammar and usage: The discussion continues (Letters to Language). In: *Language*, 83 (1), 2-4.
- Haspelmath, M. (1999): Optimality and Diachronic Adaptation. In: *Zeitschrift für Sprachwissenschaft*, 18, 180-205.
- Jacobson, S. (1989): Some approaches to syntactic variation. In: Fasold, R. W. & Schiffrin, D. (éds.), *Language change and variation*. Amsterdam (John Benjamins), 381-394.
- Jones, M. A. (1999): Subject-clitic inversion and inflectional hierarchies. In: *French Language Studies*, 9, 181-209.
- Kayne, R. S. (1972): Subject Inversion in French Interrogatives. In: Casagrande, J. & Saciuk, B. (éds.), *Generative Studies in Romance Languages*. Rowley, Mass. (Newbury House), 70-126.
- Koch, P. & Oesterreicher, W. (2001): *Gesprochene Sprache und geschriebene Sprache / Langage parlé et langage écrit. Lexikon der Romanistischen Linguistik 1/2*. Tübingen (Max Niemeyer Verlag), 584-627.
- Krummes, C., Guryev, A. & Morel, E. (2014). Les fautes, l'anglais, la langue en danger: Quelques mythes sur les textos. Communication faite en vue de la Semaine de la langue française et de la francophonie, 19: La langue en folie: des textos aux jeux littéraires, Neuchâtel. Disponible: <http://cedrickrummes.org/ckfinder/userfiles/files/SLFF-mythes%20final.pdf> (02.07.2015)
- Labeau, E. (2014): Quand l'analytique se fait synthétique: les formes verbales périphrastiques dans le texto. In: *Studii de Lingvistică*, 4.
- Langacker, R. W. (2010): How not to disagree: The emergence of structure from usage. In: *Language usage and language structure*, 213, 107.
- Lavandera, B. (1978): Where does the sociolinguistic variable stop? In: *Language in Society*, 7/2, 171-182. (Revisited version of Lavandera 1977).
- Ling, R. (2005): The sociolinguistics of SMS: An analysis of SMS use by a random sample of Norwegians. In: *Mobile Communications*. London (Springer), 335-349.
- Mosegaard Hansen, M.-B. (2001): Syntax in interaction. Form and function of yes/no interrogatives in spoken standard French. In: *Studies in language*, 463-520.
- Müller, C. & Strube, M. (2006): Multi-Level Annotation of Linguistic Data with MMAX2. In: Braun, S., Kohn, K. & Mukherjee, J. (éds.): *Corpus Technology and Language Pedagogy. New Resources, New Tools, New Methods*. Frankfurt (Peter Lang), 197-214.
- Newmeyer, F. J. (2010): What conversational English tells us about the nature of grammar: A critique of Thompson's analysis of object complements. In: *Language usage and language structure*, 3-44.
- Obenauer, H.-G. (1976): *Etudes de syntaxe interrogative du français*. Tübingen (Max Niemeyer Verlag).

- 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 III 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, Lirimm, Lidilem, Tetis, Viseo. ISLRN : 024-713-187-947-8
- Preacher, K. J. (2001): Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Ohio State University. Disponible: <http://quantpsy.org>
- Quillard, V. (2000): Interroger en français parlé: études syntaxique, pragmatique et sociolinguistique. Thèse de doctorat. Université de Tours, France.
- (2001): La diversité des formes interrogatives : comment l'interpréter? In: Langage et société, 95, 57-72.
- Rettie, R. (2009): SMS: exploiting the interactional characteristics of near-synchrony. Information. In: Communication & Society, 12/8, 1131-1148.
- Rossari, C. & Gachet, F. (2014): Parenthetical verbs as a challenge for discourse units. In: Discourse Segmentation in Romance Languages, 250, 95.
- Rowlett, P. (2007): The Syntax of French. Cambridge (CUP).
- Sankoff, D. & Labov, W. (1979): On the uses of variable rules. In: Language in society, 8(2-3), 189-222.
- Stark, E., Ueberwasser, S. & Ruef, B. (2009-2014): Swiss SMS Corpus. University of Zurich. Disponible: <https://sms.linguistik.uzh.ch> (1.7.2015)
- Tagliamonte, S. A. & Denis, D. (2008). Linguistic ruin? LOL! Instant messaging and teen language. American speech, 83 (1), 3-34.
- Thurlow, C. & Brown, A. (2003): Generation Txt? The sociolinguistics of young people's text-messaging. In: Discourse analysis online, 1(1), 30.
- Weinrich, H. (1989): Grammaire textuelle du français. Paris (Didier/Hatier).
- Winford, D. (1996): The problem of syntactic variation. In: Sociolinguistic Variation: Data, Theory, and Analysis: Selected Papers from Nwv23 At Stanford, 177-192.
- Zeldes, A., Lüdeling, A., Ritz, J. & Chiarcos, C. (2009): ANNIS: A search tool for multi-layer annotated corpora.