

# Les représentations graphiques du lexique dans les pratiques atypiques écrites des jeunes

Nicolas Michot

Laboratoire Lexiques Dictionnaires Informatique  
Université de Cergy-Pontoise  
nicolasmichot1@yahoo.fr

## 1 Introduction

Les supports modernes de communication que sont les *chats*, les *blogs*, les *forums*, les *e-mails* ou les *SMS* ont engendré des pratiques linguistiques écrites particulières, notamment chez les jeunes. En effet, s'il ne faut pas confondre écrits sur support moderne et langage des jeunes, il n'en reste pas moins que les discours écrits des jeunes entre dix et vingt-cinq ans sur les supports modernes de communication sont ceux qui présentent le plus grand nombre de phénomènes de variation et d'atypicité<sup>1</sup> sur le plan des représentations graphiques. Ces écrits peuvent être analysés à différents niveaux linguistiques, mais le lexique et ses représentations graphiques constituent une caractéristique centrale et emblématique. D'ailleurs, ces pratiques suscitent beaucoup de réactions et de débats concernant l'influence et les conséquences qu'elles pourraient avoir sur le niveau de maîtrise linguistique chez les jeunes, et plus spécifiquement sur leur niveau de maîtrise orthographique. À l'intérieur d'une réflexion générale consistant à caractériser et à comprendre les écrits atypiques produits par les jeunes tant au niveau des mécanismes linguistiques qu'au niveau de leur place dans la société moderne et mouvante, nous chercherons ici à décrire les différentes logiques de représentations graphiques et orthographiques du lexique dans les discours écrits des jeunes consultables dans les blogs. Il s'agit de rendre compte des divers mécanismes d'encodage du lexique afin de comprendre les manières dont la langue française peut être représentée autrement et parallèlement à la norme dite « standard ».

Quelles sont les différentes tendances d'usage divergentes de la norme et les différentes logiques de représentations graphiques et orthographiques du lexique ? De quelles façons les scripteurs représentent-ils leurs lexiques spécifiques ? Et ces écrits présentent-ils des règles orthographiques particulières et nouvelles ?

Nous commencerons notre étude par la présentation de notre corpus avant de décrire différentes représentations atypiques affectant la graphie des lexèmes et d'analyser les représentations des lexiques spécifiques aux jeunes qui sont normalement utilisés à l'oral. Puis, nous examinerons les problèmes liés aux limites normatives entre les codes dits « standard » et les codes atypiques utilisés.

## 2 Le corpus et les différentes logiques de représentations graphiques

Le recueil de données est une étape importante de notre réflexion. Nous présenterons de ce fait notre corpus avec les enjeux qu'il implique, avant de nous attacher à décrire différents phénomènes de représentations graphiques essentiels pour situer les logiques orthographiques par opposition et parallèlement à la norme académique.

### 2.1 Le corpus

Notre corpus est composé de discours écrits par des jeunes entre dix et vingt-cinq ans dans les blogs. Il s'agit plus précisément de blogs provenant de la plateforme *skyrock.com* dans la mesure où c'est la plateforme de référence des jeunes Français – environ seize millions de blogs à ce jour. Précisons que le moyen de communication que constitue le blog n'est pas spécifique aux jeunes, mais que ces derniers présentent une importante hétérogénéité dans leurs usages linguistiques. La plateforme *skyrock.com* est la

plateforme présentant le plus de variations dans les usages linguistiques et graphiques des jeunes Français.

Mais qu'est-ce qu'une plateforme de blogs ? Et pourquoi étudier les discours écrits sur ce support moderne ? Une plateforme de blogs est un espace Internet proposant un fonctionnement commun et une présentation commune à tous les internautes désirant se créer une page personnelle en ligne. Aujourd'hui, beaucoup de personnes, jeunes et moins jeunes, ont un blog pour se présenter ou pour exprimer leurs idées. Les blogs étant de plus en plus consultés et lus dans une société moderne où tout s'accélère et où tout se renouvelle sans cesse, il nous semble important de chercher à comprendre les différents usages observables. Or, si beaucoup de jeunes écrivent en respectant la norme académique, d'autres écrivent de façon atypique.

Pour sélectionner notre corpus, nous avons donc retenu les critères suivants : l'âge des scripteurs, la plateforme de blogs, et le caractère atypique du discours. Précisons que notre corpus n'a pas été construit à des fins statistiques car, étant donné le nombre de blogs, des résultats chiffrés seraient inévitablement approximatifs. L'enjeu est donc de rendre compte de la diversité et des mécanismes des phénomènes que nous pouvons rencontrer dans les discours atypiques produits par les jeunes à différents niveaux d'analyse linguistique. Pour cela, nous avons consulté plusieurs milliers de blogs de jeunes aux influences et aux pratiques linguistiques diverses afin de rendre compte d'un maximum de phénomènes différents. Cependant, lors de la sélection de notre échantillon, nous devons prendre garde à ne pas orienter notre corpus. En effet, plusieurs méthodes de recherche sont possibles et doivent être manipulées avec précaution. Nous pouvons tout d'abord utiliser le moteur de recherche de la plateforme en tapant des mots-clés de diverses natures. Ainsi, nous n'aurons pas les mêmes résultats si nous lançons une recherche à partir d'un terme désignant un lieu géographique, d'un terme faisant référence à une activité ou à une culture commune à certains jeunes, ou encore d'un terme ou d'une séquence graphique particulière beaucoup ou peu utilisés dans les discours des jeunes. Nous pouvons également effectuer nos recherches à partir des profils des blogueurs pour naviguer de lien en lien et ainsi sélectionner les jeunes en fonction de leur âge précis, de leur ville, ou de leurs centres d'intérêts affichés dans la rubrique intitulée « *ce que j'aime* ». Mais nous pouvons aussi partir des blogs vedettes sur la page d'accueil de *skyrock.com* pour naviguer ensuite de lien en lien. Notons que ces deux dernières méthodes contribuent à mettre en évidence des réseaux de connaissances Internet. Ainsi, nous pouvons parfois remarquer que certains scripteurs adoptent un même style graphique, notamment en ce qui concerne les graphies atypiques particulières à fonction stylistique<sup>2</sup>.

La méthode consistant à écrire une séquence graphique atypique fréquente mais non connotée comme étant « jeune » d'un point de vue lexiculturel (Pruvost et al. 2003) a souvent été utilisée. Ces méthodes de recherche mettent au jour l'importance du lexique et des influences communautaires, et par conséquent, le risque d'orientation d'un corpus est essentiellement dû aux choix effectués au niveau lexical. Elles mettent également au jour l'importance des particularités graphiques formelles de telle ou telle séquence en fonction de la graphie ou du style du scripteur. Ainsi, nous n'aurons pas les mêmes résultats si nous tapons *voudrai*, *voudray*, *vdrai*, ou *vsdré* pour la séquence [vudrɛ] par exemple. Notre corpus a été construit dans le but de présenter un panel des différents usages et des différents styles graphiques possibles dans les blogs écrits par des jeunes. Nous nous inscrivons donc dans une optique descriptive des usages atypiques de la langue écrite par les jeunes afin de les comprendre dans leur fonctionnement. Dans le cadre de notre étude générale, nous avons rassemblé notre corpus à partir des méthodes de recherche décrites précédemment, ainsi qu'à partir de l'observation d'écritures atypiques présentes sur d'autres supports de communication ou de termes et expressions utilisés à l'oral par des jeunes afin de rendre compte des convergences et des divergences d'usage entre les différents supports de communication, et entre la langue parlée et ces types d'écrits souvent qualifiés de « parlécrits ». Nous avons également recherché des informations lors d'entretiens avec des jeunes ou lors de consultations de forums sur le langage SMS et autres langages perçus comme étant « jeunes ».

Le recueil d'un corpus tel que celui-ci met en évidence deux questions essentielles. La première concerne le caractère éphémère d'une page web car un blog peut disparaître du jour au lendemain. Cependant, les phénomènes linguistiques observés ne sont pas remis en cause dans la mesure où ils ont bien existé à un moment donné. Le corpus présente seulement la caractéristique d'être mouvant ou « jetable ». La seconde

concerne la taille du corpus. Un corpus de blogs à échelle humaine ne constituera toujours qu'un infime échantillon de toutes les productions écrites librement consultables sur Internet. Il est donc difficile de rassembler un corpus exhaustif présentant toutes les pratiques atypiques possibles et existantes, d'autant qu'il s'agit d'écrits en perpétuelle évolution caractérisés par la variation graphique.

Enfin, il est important de souligner que tous les exemples seront cités en copie conforme à leur blog source. En effet, toutes les irrégularités au niveau typographique ou à n'importe quel autre niveau graphique sont importantes pour l'analyse variationnelle de ces types de discours dans la mesure où cela peut influencer sur le nombre de caractères utilisés pour écrire une séquence graphique, sur le style graphique adopté par tel ou tel scripteur, ou encore sur les différences d'usage d'un scripteur à l'autre en termes de constance ou d'économie graphique par exemple.

## 2.2 Pour une description de la morphologie graphique du lexique

Dans les discours atypiques produits par les jeunes dans les blogs, le lexique présente des caractéristiques graphiques particulières. En effet, les scripteurs utilisent des procédés fonctionnant avec différents mécanismes, séquences graphiques après séquences graphiques, ou fonctionnant de façon systématique au niveau de tel ou tel graphème à l'intérieur d'un même discours. La morphologie graphique du lexique présente donc des particularités que nous pouvons observer à travers certains types d'usages que nous proposons de décrire.

### 2.2.1 Les représentations à fonctions ornementale et expressive du lexique

Certains discours de blogs présentent une morphologie graphique globale particulière. Différents styles de représentations graphiques sont identifiables sans que le sens soit forcément modifié. Par exemple, nous pouvons relever des usages particuliers des majuscules (1), ou l'utilisation de caractères particuliers liés au choix de la police de caractère (2).

(1) Tjr Fo Ke jPaRl Du 78 Mol Pe Pa M'EmPeCheR LoL BoN BaH aVOuEz Ce De-  
PaRTeMeN i Tue TouT Koi Fo aVouEr

(2) Øκ , Sí τν λάσνεS τØη vlØg ícì , τν eS δάνS lá líSte , et Sí τν λάσνεS  
quάτρεS άυτρεS cØùmS δάνS ηη άυτρε άγτίcλε , je te μετ δάνS μεS ΚυvλØgS  
ρυέfερεS , ά τØí δε vØíα

Dans ces deux cas, l'ornementation du discours consiste à introduire un effet visuel. Le sens n'est pas modifié par l'utilisation atypique de la typographie. D'un point de vue orthographique, nous pouvons remarquer deux logiques différentes. Dans le premier cas, la représentation est atypique et dans le second, la représentation est normée malgré les erreurs qui ont été commises. Par conséquent, les mots affectés d'une typographie particulière ne sont pas forcément représentés de façon atypique d'un point de vue orthographique.

La variation ornementale peut être présente de façon plus ciblée et plus systématique dans la graphie des mots utilisés. En (3) par exemple, nous sommes confronté à l'utilisation de signes graphiques particuliers à la place des signes standard. Dans ce cas, il s'agit de « rendre le texte joli » pour reprendre les termes des jeunes, et donc d'orner le texte dans la mesure où les remplacements sont simples à décrypter. Cette pratique vient du système cryptique appelé *leet speak* par la communauté d'internautes initiés à ce type d'écriture. Ainsi, nous pouvons remarquer les remplacements systématiques du « s » par \$, du « e » par £, du « a » par @, et du « o » par 0 sans modification du sens des séquences graphiques.

(3) v0il@ mn n0uv0 bl0g c0ntre le\$ r@g£ux !!! d0nc \$ì tu v£ p@ssé un c0up d£ gu£l  
Ou k0i k£ \$£ \$0i né\$it£ p@ @ l@ché d c0m pi surt0u @\$Sum£ \$k£ tu di é m£t t0n  
n0m !!!

Le *leet speak* est un langage cryptique consistant principalement à remplacer un signe alphabétique par un ou plusieurs signes graphiques de façon à reproduire avec les possibilités du clavier d'ordinateur les formes approximatives des lettres composant les mots. Le plus souvent, il s'agit de chiffres, mais nous

observons également divers signes réalisables à partir du clavier d'ordinateur tels que \, /, °, =, etc. Ainsi, « leet speak » pourra s'écrire *1337 5p34k* par exemple, mais ce système présente différents degrés de cryptage, et donc, beaucoup de variations. Dans les blogs, les jeunes utilisent ces caractères principalement pour introduire un effet visuel dans leur texte, même s'il s'agit parfois de crypter leur discours (4)<sup>3</sup>. Le degré d'usage est donc différent d'un discours à l'autre comme le montrent les exemples (4) à (6).

(4) CayR13 Si v00 a Riz V4y 4 T00 L1r3 100 AN CHIER zet3s dé Pgm [...] o  
pAç4ge J'3n Pr0fit3 PR p4Ssé 1 M3ç4ge : Si tU 4PpRaysS!3 Cay G3ns I4 Ou Ke tU  
T3 100 Pr0che 2 I4 P3t4sse Ci D3cou, Ou 3nKor Ci 7 !M4ge T3 T0uches, é B!1 Nez  
Z!t3 P4s !!! Lèç3 La DR4içe 2 Ton Bl0g DAN lé Kom loooooooooooooooooo  
(5) J'M£ T4P£ D'SK£ P4RL£NT L£HY RÛM£UR SUR MÔÛ4 TÔÛ SKii M'I  
PÔRTE 4 N'ÔÛ4 B4 C£HY K'MÔÛ4 J'SÔÛ4 FI£R D'MÔÛ4  
(6) j£ P£n££ Qu'il v@%t Mi£µX r£p@rT!R Dµ B@\$ £t T£rM!N£r £n H@ut...  
Qu£ L£ ConTr@!r£!!! J£ r£\$p£ct' tt c£µX k! L@ supPort£ C@r £ll£ £n @ gr@v'  
B£zO!n £t £n + £l l£ m£r!T£... c£ Qu£ j£ n£ cOmpren pO...c'£st c£\$ P@r!\$i£n\$  
pOµr L'OM...

En (4), nous pouvons remarquer qu'en plus de crypter son message avec l'utilisation de signes particuliers relevant du leet speak, le scripteur utilise le digramme « ay », écrit alternativement *ay*, et *4y* pour représenter les sons [e] et [é]. Cependant, l'utilisation de ce digramme n'est pas systématique pour représenter ces mêmes sons puisqu'il utilise aussi les signes *è* et *3*. De plus, le scripteur cumule ces particularités à la fois cryptiques et ornementales avec des procédés de réduction graphique comme l'utilisation de la lettre « k » (*3nKor* pour « encore ») ou la simplification de lettres doubles (*D3cou* pour « dessous »). Nous pouvons aussi remarquer une tendance à orthographier les mots en miroir avec les graphies standard, ce qui engendre une grande variation d'usage. Par exemple, le son [s] est transcrit avec la consonne double « ss » dans *4PpRaysS!3* pour « apprécie » et avec le graphème « ç » dans *Lèç3 La DR4içe* pour « laisse l'adresse ». Enfin, dans cet exemple, tous les chiffres n'ont pas la même valeur : alors que certains sont utilisés en remplacement de lettres comme *4* pour « a » ou *3* pour « e », d'autres sont utilisés par économie (cf. § 2.2.2 *infra*). Concernant les exemples (5) et (6), nous observons des usages très réguliers relevant uniquement du style graphique, et non pas du cryptage. Nous relèverons une particularité d'usage en (5) : l'utilisation de la séquence « ehy » pour représenter les sons [e] et [é]. Si cette séquence produit un effet cryptique, elle est surtout utilisée pour donner un style au discours écrit à l'intérieur même de l'orthographe des mots et non pas seulement en apparence typographique, de la même façon que les scripteurs des exemples (4) ou (7) utilisent des digrammes tels que « ey » ou « ay » pour représenter certains sons.

(7)BN BREF MALGREY TT MOI Y TADOUREUH XD ZEUNE  
DEMOUAZELLE AC LE COEUR A PRENDRE ALOUR NEZIITEUH PAAA  
BISOUW CH'T'ADOWRE

Nous pouvons remarquer que dans ce type de style graphique, la représentation phonétique des termes affectés peut être modifiée par défaut. Ceci étant, il faut différencier les séquences n'introduisant qu'un effet stylistique des séquences phonétiquement modifiées introduisant un effet de proximité dans les interactions entre les scripteurs où les termes les plus affectés sont ceux qui ont un lien avec ce qui se rapporte à l'affectivité. Ainsi, nous pouvons établir une différence entre *zeune*, *tadoureuh* (7), ou *voualou*, *oudray*, *jim* (j'aime), *ousi* (8), où les substitutions de sons contribuent à renforcer la proximité entre le scripteur et la personne concernée dans le message, et *arreytey* (arrêtez), *deuw* (de), *meuw* (me) en (9), ou *lehy* (les), *cehy* (c'est) en (5), qui relèvent d'un style graphique particulier à fonction ornementale. Dans ces derniers cas, il s'agit en fait de graphies se greffant régulièrement sur les mêmes séquences graphiques standard, ou à la même place dans une séquence graphique atypique donnée, ayant un fonctionnement analogue à certains argots à clef, mais n'ayant pas l'objectif premier de crypter les séquences affectées. Ces séquences graphiques ajoutées ou modifiées sont « -euh », « -w » précédé des séquences « o » (*trou* pour « trop »), « ou » (*bisouw* pour « bisous »), « eu » (*meuw* pour « me »), « -ey », « -ay », ou « -ehy ».

- (8) VouAIOU oUdRAy Ou ENCoUR ApElé ma doudoo ! ba cet Meuf JIM TrOP c  
MA tiTE beSTOUllLe Ousl !  
(9) ARREYTEY DEUW PARLEY DEUW MWA.. VOUH ALLEY MEUW  
RENDRE CEYLEYBRE.. xD

De la même façon, en (10) et en (11), nous pouvons remarquer l'accumulation de procédés d'ornementation comme l'alternance entre les majuscules et les minuscules, les étirements graphiques quasi systématiques et courts n'influant pas sur le sens des mots, ou l'ajout de signes « parasites » visant à introduire un style graphique particulier. Notons que le « 'S » ajouté à la fin de certains mots de l'exemple (11) ne modifie ni la représentation phonétique de la séquence graphique, ni le sens des mots qui en sont affectés. Là encore, il s'agit de procédés d'ornementation appliqués régulièrement à la manière d'un argot à clef.

- (10) TkeYtt JeEsPeree moi en tOu Kaa Tu BaiiSseRaa PasSs IOol...  
...Ahhh TrO cOol Sii l'ANnée ProUchaiiNe On sreTrOUVve Ds la Mm CLaSs ...  
(11) Eh vOala G'te presenste ma + rande siista'S bhé uii ma siista'S d'amOur'S kOa  
) bhé iil'S à r1 à diir'S sur Elle !!! guste ke elle et zOliie'S !!! zentille'S'

Enfin, certaines représentations graphiques sont utilisées pour accentuer le sens de ce qui est écrit. C'est le cas notamment en ce qui concerne certaines utilisations des majuscules (12), ou en ce qui concerne les étirements graphiques (exemples 13 et 14). Précisons qu'en (7), le « -euh » de *tadoureh* contribue également à accentuer le sens du terme. Par conséquent, cette séquence joue à la fois un rôle ornemental par l'utilisation régulière du suffixe « -euh » (*neziiteuh*), et un rôle expressif renforçant la proximité entre le scripteur et la personne concernée dans le message.

- (12) ché patu ve kOi etan doné ke « LHIPOCRISIE EST DANS L ETR HUMAIn »  
(13) Alor elle c la plu follllllllllllleeeeeeeeeee ds fille ke je connaisse IL NYA PA PLU  
DéVERGONDER KELLE JE VS ASSURE c ma meilleure amie je l'aime grave JE  
SAI JE NE TE LE DI PA ASSEZ SOUVEN MAI JE TAIIIIIIMMEEEEEEE  
SEURETTE ADORé KISS  
(14) sa c et meuf komen jlm a donf jtador ma kayna adore kssss!!!!!!!!!!!!

Nous pouvons remarquer qu'il existe deux types d'étirements graphiques : il y a d'une part, les étirements de signes divers comme les signes de ponctuation, et d'autre part, les étirements de lettres. Précisons que les étirements graphiques à fonction expressive ne doivent pas être confondus avec les étirements à fonction stylistique. Le plus souvent les étirements ornementaux sont courts et systématiques par opposition aux étirements expressifs qui sont plus longs et qui n'affectent pas nécessairement les mêmes séquences graphiques dans le discours. Enfin, les étirements à fonction expressive portent sur l'intensité du sens des mots alors que les étirements à fonction ornementale portent sur la représentation de tel ou tel graphème qui compose l'orthographe des mots.

## 2.2.2 Les représentations à fonction économique du lexique

Le caractère atypique des écrits des jeunes est souvent symbolisé par les aspects phonétisant et abrégatif. Le côté novateur de ces pratiques linguistiques est l'accumulation et la concentration de ces procédés d'écriture dans un espace graphique restreint. Nous pouvons distinguer plusieurs types d'abréviations affectant la graphie des lexèmes. Il y a tout d'abord les procédés de simplification orthographique comme l'utilisation de la lettre « k » pour représenter le son [k], la généralisation de la lettre « é » pour représenter les sons [e] et [ɛ], les simplifications des consonnes doubles, la chute des lettres muettes, ou encore l'utilisation de la lettre « o » pour représenter le son [o]. Ces phénomènes sont observables dans des exemples tels que (15) ou (16).

- (15) Dc voila ztm vrémen for ma pitchoune, fo surtt pa ke tu change t génaila kom  
ca !! jte remerci pr tt car ss toi ché pa ske je feré, t tjs la pr moi, pr me reconforter é tjs  
la pr se taper dé pure trip!!!  
(16) bref voila jvs léss yeuté trankil.... g kune choz a vs dir c laché vos COMZ!! lol

Cependant, la lettre « è » est également employée pour représenter les sons [e] et [ɛ], et souvent pour introduire un effet stylistique explicable par la représentation du son [ɛ] par défaut. En effet, si certains sons [e] sont fréquemment et spontanément prononcés [ɛ] dans des séquences comme [mɛ] pour « mais » ou [vudrɛ] pour « voudrais », nous n'aurons pas, en situation d'oral spontané, [parle] pour « parler » par exemple, alors que nous pourrions rencontrer *parlè* pour cette même séquence dans les blogs. Ceci étant, pour marquer ce type de transcription graphique à fonction stylistique, les jeunes utiliseront le plus souvent des digrammes tels que « ay » ou « ey » pour transcrire le son [e] (cf. *supra*).

Nous pouvons relever deux grandes catégories d'abréviations. La première concerne les abréviations ne laissant qu'une représentation symbolique de la séquence graphique affectée, et la seconde concerne celles qui laissent des séquences graphiques représentant une séquence phonique correspondante dans son intégralité. Autrement dit, la première catégorie présente des abréviations telles que les *squelettes consonantiques* des mots, pour reprendre la terminologie de Anis (2003 et 2004), comme par exemple *tj* ou *tjs* pour « toujours », *vs* pour « vous », *tt* pour « tout » (repris dans *surtt*, exemple 15), *jms* pour « jamais », *pr* pour « pour », etc. Il s'agit là d'abréviations classiques concernant ces mots, mais ce procédé est également appliqué à des séquences déjà modifiées orthographiquement comme *pk* pour « parce que » (*paske*) ou « pourquoi » (*pourkoi*) ou *tmtc* pour « toi-même tu sais » où le squelette consonantique est lui-même affecté d'un procédé de réduction. Cette catégorie présente également ce que Anis (1999 et 2004) appelle *abréviation unilittère* ou *paralogogramme* comme *v* pour « vais » ou « veux », *t* pour « te » ou « tu », *j* pour « je », *d* pour « de » ou « du », *f* pour « fais », ou *k* pour « que » (*ke*) ou « qui » (*ki*). Nous pouvons remarquer que seuls des termes très fréquents ou des termes grammaticaux sont affectés par ce procédé dans la mesure où l'abréviation consiste à ne laisser que la première lettre d'un mot.

La seconde catégorie concerne des abréviations comme les *syllabogrammes*. Différents types de syllabogrammes sont identifiables. Il y a les syllabogrammes remplaçant une séquence complexe comme *c* pour « c'est » ou « s'est », les syllabogrammes remplaçant un mot comme *d* pour « des », et les syllabogrammes remplaçant une syllabe comme *v* dans *sav* pour « savait » ou *p* dans *zaP* pour « zapper ». Précisons que les syllabogrammes peuvent être agglutinés à un mot, isolés, ou agglutinés entre eux, et qu'ils peuvent parfois être mis en relief par l'utilisation de la majuscule. Ainsi, une séquence telle que *tt* est soit une agglutination de deux syllabogrammes pour représenter la séquence « t'étais », soit le squelette consonantique représentant « tout ». Nous pouvons également relever des *logogrammes* au sens où l'emploie Anis (2004). Différents types de logogrammes sont identifiables. Il y a les logogrammes qui sont utilisés pour représenter leur sens d'origine comme *1* pour « un » ou *7* pour « sept », ceux qui sont utilisés pour représenter une séquence homophonique comme *7* pour « cet(te) » ou *2* pour « de », et ceux qui sont utilisés pour représenter une syllabe ou un son comme *2* et *l* dans *2ml*.

Par ailleurs, nous pouvons relever différentes formes d'*abrégement*, pour reprendre la terminologie de Fairon et al. (2006), comme les troncations ou les différents types de sigles très utilisés dans ces écrits. Ces usages se situent dans une logique abrégative, mais nous ne développerons pas ici leurs caractéristiques morphologiques dans la mesure où ils sortent du cadre de l'analyse orthographique. Nous mettrons simplement au jour quelques variations en ce qui concerne la représentation graphique de certains mots spécifiques aux jeunes (cf. § 2.3 *infra*).

En définitive, nous pouvons observer trois grandes tendances d'usages qui peuvent se combiner entre elles : la première est une dominante d'usage à fonction ornementale ou stylistique, la deuxième présente une logique abrégative, et la troisième concerne l'usage de procédés à fonction expressive. La logique abrégative est celle qui est la plus fréquemment utilisée dans les blogs, et d'une certaine manière, c'est celle qui représente le style atypique des blogs. Par conséquent, en plus d'être motivée par des fonctions d'écriture, elle joue un rôle stylistique par son aspect phonétisant qui semble transcrire directement la langue parlée. De surcroît, elle est connotée « jeune » car ce sont les jeunes qui pratiquent le plus les écrits atypiques sur les supports modernes de communication<sup>4</sup>. Précisons qu'un procédé de représentation graphique très utilisé dans ce type d'écrits se situe à la limite de plusieurs fonctions. Il s'agit de la *variante semi-vocalique* (ou *déconstruction graphique*) du digramme « oi » qui devient *oua*, *oa*, *ôa*, *wa*, ou *woua*. Ces graphies correspondent à la fois à une logique phonétisante sans abréviation et à une logique

stylistique de représentation graphique. Enfin, rappelons que certains usages stylistiques peuvent contenir une dimension cryptique lorsqu'ils sont utilisés de façon extrême. C'est le cas de l'utilisation des caractères relevant du leet speak comme les chiffres, ou encore de la représentation de certains sons par des séquences graphiques spécifiques comme « ehy » pour représenter les sons [e] et [ɛ].

### 2.3 Les représentations graphiques des lexiques habituellement utilisés à l'oral

Les jeunes blogueurs utilisent souvent ce que nous avons l'habitude d'appeler « lexique des jeunes » et/ou « lexique des cités ». En fait, il s'agit des termes massivement utilisés par les jeunes, qu'ils viennent des cités ou d'autres milieux, connotés et perçus comme relevant des usages linguistiques des jeunes. Nous ne proposerons pas ici de critères de distinction pour caractériser les termes des cités par opposition aux termes utilisés par les jeunes de tous milieux et de toutes influences culturelles, puisque l'objet de notre étude est la représentation graphique des lexiques présents dans les écrits des blogs. En effet, ces lexiques présentent beaucoup de variations graphiques malgré les dictionnaires spécialisés existants<sup>5</sup>. Nous pouvons par exemple observer différentes graphies du mot [meskin] dans un même article de blog (17).

(17) [...] ué c pck c + drole (miskine va)... [...] MOI mM je c cke tu vo é meskine tu  
vo pa grd chozzZZZ .....

Précisons que ce terme est écrit *mesquine* dans un répertoire de mots consultable sur le blog *mitar-street-rap*, et *mesquin* dans Goudaillier (2001). D'autre part, nous pouvons observer des divergences de point de vue concernant le terme *seum* (rage, colère). En effet, alors que dans Azor et al. (2007) *seum* est bien segmenté conformément à la transcription présente dans la plupart des blogs comme celle de l'exemple (18), dans certains blogs, *seum* est transcrit de façon agglutinée dans l'expression *talseum(e)* qui signifie « t'as le seum » (19) et qui est devenue une marque de vêtements. Ceci étant, il s'agit d'une expression créée par certains jeunes provenant de *la cité du nord* dans le département de Seine Saint-Denis à partir du mot *seum*, puisqu'en (20), nous pouvons remarquer que le scripteur segmente le terme dans le même énoncé. Dans le blog *mitar-street-rap*, ce terme est transcrit *sum*.

(18) Les Gens Parlent Trop Sans Savoiir ... Sa M'fou Trop Le Seum !!!

(19) En MoDe TaLsEuM!!!!

(20) voilà t'as compris le concept, déguaine ton talseume et fous lui encore plus le seume!!!

Certains autres termes comme [sos] présentent des différences de transcription entre les blogs et les dictionnaires spécialisés. Dans Azor et al. (2007) ce terme est transcrit *sauce*, alors que dans les blogs il est transcrit *sooss* (21) ou *soce* (*mitar-street-rap*) le plus souvent.

(21) SiiSii 7 MeUuUf C Ma SoSs WouLLaH Jii Lm De tRo NN Mé LeSs tRo De  
DeLiR aC cTe Goo !!

Par ailleurs, nous pouvons observer plusieurs degrés de figement dans la représentation graphique de certaines formes abrégées telles que *jtd* signifiant « je t'adore » et déclinée en *jvd* et en *jld* pour « je vous adore » et « je l'adore » ou « je les adore », ou telles que *tmtc*, parfois écrite *tmts*, signifiant « toi-même tu sais », et déclinée en *vmvs* et occasionnellement en *mmjc* pour « vous-même vous savez » et « moi-même je sais ». Les autres formes déclinées sont représentées de manière moins abrégée comme *lui mm il c* par exemple. Nous n'aurons pas dans ce cas *lmic*.

La graphie du lexique n'est donc pas figée et nous pouvons distinguer des logiques de représentation variables en fonction de la perception que le scripteur se fait de la langue parlée des jeunes. Évidemment, certains termes très répandus comme *ouf* (fou), *téma* (mate, regarde), *zarma* (genre, style), *sisi* (terme désignant la connivence dans le rapport entre plusieurs personnes), ou *crari* (faire genre, faire style) sont toujours, ou presque, représentés de la même manière graphiquement, d'autant qu'ils ne sont accordables ni en genre, ni en nombre (y compris *ouf* qui n'est pas suffixé en *oufette* pour désigner le féminin, à la manière de *tigente* pour *tigen* signifiant « gentil »), et que leur représentation graphique présente une relation biunivoque entre chaque graphème et chaque phonème. En définitive, le lexique spécifique qui

n'est pas utilisé par le plus grand nombre ou qui présente généralement les points d'articulation graphique en [k], en [ʒ], en [s], en [z], et parfois en [f], ou au niveau de n'importe quel son vocalique ou semi-vocalique, sauf les sons [y] et [ɔ̃], varie orthographiquement à travers les usages dans les blogs selon une logique phonétisante ou selon une logique de représentation normée, par l'utilisation de l'*archigraphème* (Catach, 2004) ou d'un graphème fréquent dans tel ou tel environnement graphique.

### 3 Les frontières entre l'orthographe « standard » et les représentations graphiques et orthographiques atypiques

Au vu des procédés graphiques affectant la graphie des lexèmes que nous venons de décrire de façon non exhaustive, mais représentative des logiques orthographiques adoptées dans ces types d'écrits, nous pouvons remarquer une sorte de confusion entre les limites normatives et ces logiques d'écriture qui semblent ne contenir aucune contrainte et aucune limite pour représenter la langue. Nous examinerons donc ici certains mécanismes incontournables de ces types d'écrits afin de comprendre comment les scripteurs perçoivent et utilisent les normes orthographiques du français.

#### 3.1 Pouvons-nous parler d'une orthographe spécifique ?

La morphologie graphique globale des écrits atypiques affiche un style graphique souvent phonétisant. Mais s'agit-il d'une orthographe spécifique ? Contrairement aux idées reçues souvent véhiculées par les médias à propos du « langage SMS » ou des langages des jeunes, nous pouvons affirmer qu'il ne s'agit pas d'une orthographe spécifique, compte tenu des nombreuses variations graphiques observables. Si nous reprenons une définition classique et générale de la notion d'*orthographe* (Arrivé et al., 1986 : 443), ces écrits ne correspondraient qu'à des représentations graphiques dysorthographiées de la langue. Nous avons vu que les logiques de représentations graphiques semblent se situer, pour nombre de représentations, en dehors de la logique normée, c'est-à-dire en dehors de toute contrainte et de tout code orthographique. Nous devons cependant nuancer ces affirmations. En effet, même si certaines personnes considèrent ces écrits comme étant une langue étrangère et incompréhensible, il n'en reste pas moins qu'il s'agit de représentations du français. De plus, ces écrits sur supports modernes ne sont pas envisagés comme étant des actes individuels, mais comme des actes à faire partager, et par conséquent, le discours doit rester compréhensible et concevable pour tous. De ce fait, ces pratiques écrites présentent des constantes figurant comme étant des critères incontournables, quelle que soit la logique de représentation graphique adoptée par le scripteur, à savoir la *concevabilité graphique* des séquences représentées, c'est-à-dire la représentation graphique, affectée d'un procédé ou non, décodable et interprétable par tous en une séquence de langue française, l'*intercompréhension*, c'est-à-dire le souci du scripteur d'envisager ses représentations graphiques comme étant concevables pour un groupe restreint et/ou un groupe large de lecteurs, et l'utilisation des *structures syntaxiques* du français écrit et/ou oral. En effet, ce type d'écrit peut aussi bien présenter des caractéristiques propres à l'oral dans les représentations graphiques que des caractéristiques propres à l'écrit. Ainsi, dans l'exemple (22), nous pouvons remarquer la présence dans le même discours d'une caractéristique de prononciation (*chui* pour « je suis »), et de caractéristiques propres à l'écrit comme les accords.

(22) oui je sais, il faut que je fasse un article, un vrai . Mais les gens... chui pas une machine à article...

Ces écrits se situent donc toujours par rapport à la norme et à ses codes, et non pas simplement par rapport à la forme phonique de la langue. D'ailleurs, il ne s'agit en aucun cas d'écrits phonographiques à relation biunivoque comme l'attestent certains procédés d'abréviation tels que les squelettes consonantiques, ou encore certains procédés d'expressivité comme l'étirement graphique. Cependant, les signes graphiques fondamentaux que sont les signes alphabétiques ou les signes mathématiques et numériques sont parfois utilisés différemment de la norme. En réalité, leur fonctionnement est plus complexe dans la mesure où ces signes graphiques peuvent aussi bien être utilisés selon les normes que selon des règles recrées. Par conséquent, ces écrits présentent la particularité de mélanger les codes graphiques normés et



des codes de représentations graphiques atypiques. Ce mélange de codes engendre nécessairement des ambiguïtés graphiques entre des séquences graphiques normées et des séquences graphiques affectées de codes spécifiques. Mais quels sont ces codes spécifiques et quelles sont les ambiguïtés possibles par défaut ?

### 3.2 Analyse des mécanismes de représentation graphique

Le mélange des codes orthographiques met parfois en évidence des contradictions de logique de représentation. Par exemple, nous pouvons rencontrer des termes ou des formes-mots représentant des segments plus longs tels que *eureusement* (heureusement), *aparemmen* (apparemment), *chuis* (je suis), *chais* (je sais), ou *2main* (demain), présentant des contradictions internes de transcription à cause de la prégnance de la norme chez les différents scripteurs. En effet, pourquoi conserver la lettre muette finale dans *eureusement* alors que le « h » a été supprimé, pourquoi transcrire le son [a] avec la graphie normée dans *aparemmen* alors que le « pp » a été simplifié et que la lettre muette finale a été supprimée, pourquoi conserver les désinences verbales dans *chais* et *chuis*, et enfin, pourquoi ne pas remplacer le trigramme « ain » de « demain » par le chiffre 1 ou le graphème « in » qui est plus simple ? Autant de contradictions montrant que ces écrits transcrivent bien la langue française en décalage normatif et avec des codes recréés se greffant massivement sur la norme « standard ». Les scripteurs semblent donc être plus fortement ancrés dans la norme académique qu'il n'y paraît ; ils ne recréent pas une langue ou un langage, ils se fondent seulement sur des connaissances préalables en utilisant parfois des signes graphiques de façon atypique.

Une analyse graphophonique des signes graphiques alphabétiques nous permet de rendre compte des différentes possibilités de représentations et de la complexité de fonctionnement dans les représentations graphiques atypiques. Concernant les consonnes, les mécanismes sont réguliers sauf pour trois cas complexes, *c*, *g*, *s*, et trois cas à part, *h*, *w*, et *x*. En effet, chaque consonne a deux possibilités de représentation dans un mot sauf lorsqu'elle fait partie de procédés symboliques et illisibles comme les squelettes consonantiques ou paralogogrammes. La première est la possibilité conforme à la norme par défaut, et la seconde est la représentation de la séquence phonique correspondant à la prononciation de la lettre en tant que telle. Par exemple, « n » représentera soit le son [n], soit la séquence [en], « p » représentera [p] ou [pe], etc. Précisons que nous donnons ici la représentation d'un son par défaut puisque « n », par exemple, peut également servir à composer des digrammes tels que « en », « an », ou « on ». Les trois consonnes simples à fonctionnement complexe présentent trois possibilités de représentation de son, par défaut, quel que soit leur environnement graphique, c'est-à-dire ce qui précède et ce qui suit graphiquement. Ainsi, « c » peut représenter les sons [k], [s], ou la séquence [se], « g » peut représenter les sons [ʒ], [g], ou la séquence [ʒe], et « s » peut représenter les sons [s], [z], ou la séquence [es]. Nous pouvons citer quelques exemples d'usages atypiques de ces lettres dans des termes ou des segments tels que *guste* (juste), *chang* (change), *biss* (*biz* pour « bisou »), *gacté* (jacter), *osi* (aussi), *cte* (cet), *sengeuler* (s'engueuler), *ca* (ça), *pasé* (passé), ou encore dans des segments tels que *gsper* (j'espère), *lS* (laisse), ou *la FR* (l'affaire). Précisons que si ces lettres peuvent être utilisées en dehors des règles classiques de représentation graphique, elles sont le plus souvent utilisées de façon standard. Concernant les cas à part, le « h » peut être muet ou représenter la séquence [aʃ] lorsqu'il ne compose pas le digramme « ch », le « w » peut représenter les sons [v] et [w], mais il ne représentera jamais la séquence \*[dubleve], et le « x », qui est un cas à la fois singulier et complexe, peut représenter les séquences [iks], [eks], et [gz]. Ainsi, nous pouvons rencontrer *HT* (acheter), *gwo* (« gros » représenté selon une prononciation antillaise stéréotypée), ou *Xtrem* (extrême). Par conséquent, malgré une apparente simplification graphique, les signes graphiques peuvent fonctionner de façon assez complexe par ce que nous appellerons la *valeur multiphonique* des consonnes. Ces règles constituent une constante dans les usages graphiques des jeunes et dans tous les cas où la logique orthographique des discours atypiques se situe dans la reproduction de la séquence phonique complète. Le fonctionnement de ces lettres est valable à l'intérieur d'une séquence graphique, lors de l'apparition de lettres-mots<sup>6</sup> en discours ou de certaines séquences consonantiques courtes comme *gt* ou *ct* pour « j'étais » et « c'était », mais il est important de prendre en considération les

phénomènes de variante libre entre les sons [e] et [ɛ] dans la mesure où les représentations de séquences graphiques symbolisées par une lettre engendrent souvent l'utilisation d'un son pour l'autre par défaut. Enfin, un autre type d'usage plus singulier de consonne a retenu notre attention dans les séquences *Rput* ou *rput* pour « réputation ». En effet, nous pouvons constater que la lettre « r » est utilisée pour représenter la séquence [rɛ], et non pas [ɛr]. Dans ce cas, il ne s'agit pas d'une chute de [ə] caduc comme dans *jsrai* pour « je serai » par exemple, et de ce fait, ce phénomène se situe à l'inverse des règles décrites précédemment.

Les représentations des voyelles sont moins systématiques. En effet, les usages présentent davantage de possibilités pour transcrire les sons. Par exemple, le son [ɛ] peut être transcrit par les graphies « ehy », « ey », « ay », « é », « è », « e », « ai », « ei », « 3 », etc. selon les scripteurs et le type de style graphique propre à chacun. Un analyse phonographique, c'est-à-dire une description des différents graphèmes utilisés pour transcrire tel ou tel son, aurait donc été pertinente, mais comme nous focalisons notre attention sur le fonctionnement des graphèmes utilisés pour transcrire les différents sons dans les discours atypiques, nous poursuivrons notre analyse graphophonique. Nous nous attacherons pour cela à ne décrire que les principaux phénomènes récurrents et représentatifs relatifs à la représentation de tel ou tel son vocalique. Certains graphèmes n'apparaîtront donc pas dans notre description s'ils sont quasi inexistantes et/ou s'ils ne présentent pas véritablement d'atypicité dans les usages. Contrairement aux consonnes, les voyelles présentent différents graphèmes accentués et davantage de graphèmes complexes que sont les digrammes et les trigrammes. En ce qui concerne un grand nombre de voyelles accentuées avec le tréma, l'accent circonflexe, ou, pour certaines voyelles, avec les accents aigu ou grave comme le « u », le « o », ou le « i », l'accent n'a pas de fonction diacritique, c'est-à-dire qu'il ne modifie pas la représentation phonétique de la lettre. Ce type d'accentuation est une caractéristique ornementale des écrits atypiques des blogs.

Nous pouvons différencier les graphèmes représentant plusieurs voyelles selon l'environnement graphique des graphèmes à fonctionnement simple ne représentant qu'un seul son vocalique. Le graphème le plus complexe est le « e » dans la mesure où il peut représenter les sons [ə], [e], [ɛ], [œ], [ø], selon son emploi et selon que le scripteur omette l'accent ou non lorsqu'il est présent dans la graphie standard. Ainsi, nous pouvons observer des séquences comme *ref* (*veuf* pour « frère »), *mef* (*meuf* pour « femme »), *fer* (faire), *pe* (peut), *ke* (que), ou *ekipé* (équipée, dans *eLLe e eKiPé*). Un phénomène très fréquent mérite d'être mis en évidence : il s'agit de la réduction de la graphie « eu » par « e » dans les séquences *pe* (peux), *ve* (veux) et *ce* (ceux). Les graphèmes « é » et « è » ont également un fonctionnement complexe dans la mesure où ils peuvent tous les deux représenter les sons [e] et [ɛ]. Par exemple, nous pouvons rencontrer des séquences comme *mé* (mais), *vréman* (vraiment), *loré* (j'aurais), *parlé* (parler), ou *yallè* (y aller). Certains monogrammes présentent un fonctionnement simple. C'est le cas de « a », « o », « u » représentant respectivement les sons [a], [o] et sa variante [ɔ], et [y]. De la même façon, nous avons des digrammes et des trigrammes qui présentent un fonctionnement simple tels que « ou » pour [u], « ai » et « ei » pour [ɛ], « in », « ain », « ein », « un » pour [ɛ̃] ou [œ̃], « on » pour [ɔ̃], « au », « eau » pour [o]. Cependant, compte tenu de l'analyse graphophonique des consonnes, il est important de préciser que l'assemblage de voyelle(s) et d'une consonne dans les graphèmes composés peut ne pas représenter nécessairement un son vocalique ; ces graphèmes doivent effectivement être considérés comme pouvant occasionnellement représenter une séquence décomposée où chaque signe correspond à un son ou à la valeur phonétique de la lettre. Précisons qu'aucun phénomène de ce type n'a pu être relevé dans notre corpus, à l'exception des graphèmes « em » et « om », ce qui montre que les scripteurs sont souvent attachés à la norme lorsque plusieurs graphies existent déjà. Par ailleurs, nous pouvons remarquer que le son [ɛ̃] pourrait présenter des usages systématiques avec le chiffre 1, mais c'est loin d'être le cas dans les productions consultables dans les blogs. Notons que la séquence « oin » est le plus souvent transcrite *ouin* ou de façon standard, mais rarement avec le chiffre 1. Enfin, un cas singulier présente un fonctionnement complexe. Il s'agit de la séquence graphique « oo » pouvant représenter le son [u] dans *bisoox*, le son [œ] dans *moof*<sup>7</sup>, ou le son [o] dans *noom* (nomme).

Concernant les semi-voyelles, nous pouvons relever des usages particuliers des graphèmes « y », « i », « il », « ou + voyelle », et « w ». Les trois premiers cas représentent le son [j] dans des séquences telles que *fiy* (fille), *ia* ou *ya* (il y a), ou *fil* (fille). Précisons que le graphème « i » peut également représenter le son [i], et que la séquence « il » présente ici un emploi atypique dans la mesure où l'accord n'est pas réalisé et que la séquence homographique se prononçant [fil] crée une ambiguïté au premier abord. Les deux derniers cas représentent le son [w]. Nous pouvons observer l'usage de ces graphies dans des séquences telles que *wé* (ouais), *mwa* (moi), ou *moua* (moi).

D'autre part, nous pouvons relever des variations d'usage ainsi que des ambiguïtés en ce qui concerne les graphèmes « an », « en », et « em » représentant le son [ã]. En effet, nous pouvons rencontrer *komen* ou *koman* pour « comment », par opposition à *comem* signifiant « quand même » et non pas « comment ». Dans ce dernier cas, nous sommes face au phénomène explicité précédemment, à savoir un fonctionnement biunivoque entre chaque signe et chaque son. Cependant, *longtem* est également identifiable, et là encore, nous pouvons observer une certaine prégnance du standard dans les productions atypiques des jeunes. Il est logique que le « em » de *comem* ne représente pas le son [ã] dans la mesure où le « m » n'est présent que devant certaines lettres dans l'orthographe du français. Par opposition, *longtem* a simplement été tronqué de son morphogramme final, ce qui explique le graphème « em » pour représenter le son [ã]. De la même manière, « om » peut soit représenter le son [õ] dans ses usages standard, soit la séquence [õm] comme dans *com* (« commentaire » ou « comme »).

Enfin, le chevauchement des codes orthographiques peut soulever différentes ambiguïtés entre le code « standard » et les codes atypiques. Nous pouvons relever deux procédés importants de désambiguïstation orthographique. Le premier est la matérialisation par un signe graphique du son [e] dans les séquences *cé*, *gé* (ou *jé*), *dé*, ou *té*, le plus souvent représenté par les syllabogrammes correspondants *c*, *g*, *d*, et *t*. En effet, la matérialisation graphique du son [e] permet au lecteur de ne pas hésiter lors du décodage des séquences graphiques auxquelles il est confronté. Notons que cette matérialisation graphique pourrait expliquer la graphie *pencé* pour « pensé » fréquemment employée dans les blogs. Le second procédé est le redoublement de lettres pour marquer le féminin ou l'utilisation de la majuscule pour marquer la prononciation de la lettre muette finale. Par exemple, « gros » par défaut est prononcé [gro], et de ce fait, lorsque le scripteur désire marquer le féminin, nous pouvons rencontrer des graphies telles que *gross* ou *groS*. De la même façon, nous pouvons rencontrer *fill* pour le distinguer de [fil], et *personell* pour représenter le féminin. Cependant, il reste des cas comme *fer* ou *mer*, pour représenter « faire » et « mère » en discours, alors que ces graphies existent également avec un autre sens en français standard.

### 3.3 Les jeunes ont-ils conscience des limites normatives ou y a-t-il des inter-pénétrations entre les différentes logiques orthographiques dans leurs productions écrites ?

Les analyses précédentes nous conduisent naturellement à nous demander comment nous pouvons différencier les procédés graphiques des fautes d'orthographe, et si la notion de faute d'orthographe peut réellement exister dans des écrits perçus comme étant libres et sans contraintes. Deux critères sont à prendre en considération. Il s'agit des critères de *logique* et de *cohérence orthographique*. La logique orthographique définit le type de représentation graphique pour une séquence donnée, alors que la cohérence orthographique concerne davantage la constance d'usage dans un énoncé ou dans un discours donné. Autrement dit, entre *fê* et *fait*, ou *demain* et *2m1*, nous sommes face à deux logiques orthographiques différentes, et si dans un même discours, le scripteur accorde certains mots et pas d'autres, nous serons face à des incohérences graphiques dues à l'oscillation entre différentes logiques orthographiques. Ces critères nous permettent d'affirmer que dans la plupart des cas, c'est le surplus orthographique ou l'oubli de consonne double qui fait l'objet d'erreurs orthographiques clairement identifiables. Ainsi, dans l'exemple (2, *supra*), *quatres* et *met* dans *je te met* sont des erreurs orthographiques nettement marquées dans la mesure où la logique orthographique est standard et où les usages sont constants. En (23), la sup-

pression du « e » à « t'arrives » relève davantage du procédé de réduction que de l'erreur orthographique, étant donné l'oscillation entre la logique standard et les logiques atypiques, et la tendance au raccourcissement dans la deuxième partie de l'énoncé. En (24) la seule erreur orthographique que nous pouvons dégager clairement est la confusion homophonique entre *le est* et « ai » car c'est la seule séquence graphique écrite avec une logique marquée comme étant standard avec *raconte*. Tous les autres mots sont affectés d'un procédé de simplification ou de réduction. Les erreurs orthographiques présentes dans les écrits atypiques sont donc généralement identifiables dans des séquences graphiques écrites avec une logique normée, c'est-à-dire avec les morphogrammes muets.

(23) je sais que dans le crew on a toutes un gros pb psychologik mais alors toi t'arriv  
dans le top five krkrkrkr

(24) j18 est montré la tete ke tu fê ken je raconte des coneri o bahu !!

Par ailleurs, la question de la maîtrise de l'orthographe chez les jeunes est une question délicate qui obsède les Français et qui est au cœur des débats actuels concernant l'apprentissage de la langue à l'écrit. Certains considèrent que le niveau de maîtrise orthographique des jeunes s'est considérablement dégradé et que le développement des moyens modernes de communication a contribué à accélérer cette baisse de niveau. De telles considérations normatives à l'égard de telle ou telle pratique linguistique nous ont interpellé et par conséquent, nous avons entrepris d'examiner attentivement les pratiques orthographiques des jeunes à partir d'une enquête linguistique réalisée au collège, consistant à faire écrire un même texte aux élèves de façon normée et de façon atypique en fournissant différentes versions atypiques s'ils établissent une différence entre les styles graphiques des blogs, des chats, et des SMS. Cette expérience a été réalisée à différents niveaux de classe, mais nous ne retiendrons ici que les productions des élèves de 3<sup>e</sup>. En effet, étant donné que ce sont les élèves les plus âgés des classes de collège, ce sont ceux qui sont censés maîtriser le mieux l'orthographe du français. D'emblée, nous pouvons affirmer que la plupart des élèves établissent une différence parfaitement marquée dans leurs productions entre la logique de représentation standard et les logiques de représentation atypique. Citons un exemple-type de productions ne présentant aucune ambiguïté (exemples 25 et 25') afin d'illustrer notre propos :

(25, français normé) Salut ! Comment vas tu ? Moi bien. Aussi non, quoi de beau ?  
Moi toujours rien. Et as tu vue Mathieu ? parce que j'ai appris qu'il sortait avec Sandy.  
Mort de rire !!! C'est trop marrant !! Je n'aurais pas pensé que Sandy serait sortie  
avec lui ! Il est trop moche ! je rigole ! Bon je te laisse. A plus tard ! Bisous.

(25', version atypique correspondante) Slt ! sa va ? mwa b1. Ossi nan koi 2 bo ? mwa  
tj r1. é ta vu matie ? Pck g apri ki sorté avc Sandy MDR !!! C tro maran !! joré pa  
pensé ke Sandy seré sorti avc ! il e tro moshe ! lol ! Bon jte laisse a+ Bisous

Les logiques orthographiques, telles que nous les avons définies précédemment, sont très clairement tranchées entre (25) et (25') même si le scripteur commet quelques erreurs orthographiques en (25). En (25'), la logique de représentation est phonétisante et abrégative, et l'énoncé présente une forte cohérence dans les usages. Par opposition, nous mettrons en évidence deux exemples singuliers, (26) et (26') d'une part, et (27) et (27') d'autre part, où, outre le faible niveau de compétence orthographique des scripteurs, certaines séquences graphiques mettent au jour des ambiguïtés concernant les confusions entre les logiques de représentations atypiques et la logique standard.

(26, français normé) salut sava ? alors quoi de neuf ? ta fais quoi aujourd'hui ? moi je  
suis parti à Video gag. Mort de rire!! Comme je fais trop rire les gens alors c'est pour  
sa ta vu.

(26', version atypique correspondante) Slt sava ? alr quoi de 9 ? ta fais quoi jourduis ?  
oi je ss parti à Video gag. MDR, lol. Comme je fais trop lol les gens alors c pr sa ta  
vu.

(27, français normé) Salut comment vas tu ? tu vas bien alors la semaine c'est bien  
passé tu es amusé Moi c'est sait bien passé. Alors les filles vous allez bien oui on va  
bien sauf on s'ennuie. Goussainville c'est une ville merveilleuse

(27', version atypique correspondante) wch B1 ou koi sas pass ou koi alr la smaine  
B1ou B1 tu T tapé dé bar ou koi Moi trankil ct B1 normale. alr lé meuf vou alé B1 wé  
on va b1 mé ta vu on galère. Gouss Nike tou représente 95190 GØU\$\$1VILLE

En (26), nous pouvons relever deux erreurs paraissant mélanger les logiques orthographiques : *sava* et *rir* qui sont écrits de façon phonétisante dans une production théoriquement normée. Concernant le cas de *rir*, nous pouvons constater que le mot est orthographié selon la norme académique lors d'une première occurrence dans le même discours. Par conséquent, l'erreur orthographique ne peut s'expliquer par une confusion des limites normatives de telle ou telle logique de représentation. Concernant le cas de *sava*, il semble qu'il y ait confusion entre les différentes logiques de représentations orthographiques. La confusion du phonogramme « s » pour « ç » est très fréquente dans les productions normées des jeunes et réapparaît à la fin du même exemple de façon isolée, mais l'agglutination introduit un doute. Nous considérerons qu'il y a peut-être ici, par défaut, interpénétration entre les différentes logiques normatives sans pouvoir le prouver par le critère de constance d'usage. En effet, si nous comparons (26) et (26'), nous nous apercevons que les logiques d'écriture sont totalement différentes. Certes, (26) comporte des erreurs orthographiques, mais de manière générale, les mots sont orthographiés de façon normée, sans utilisation de procédés graphiques. Par opposition, (26') montre un grand nombre de transcriptions affectées d'un procédé et marquées comme étant atypiques avec l'utilisation de sigles à la place des mots ou encore d'un chiffre pour représenter un mot. En (27), nous pouvons relever plusieurs erreurs orthographiques. Celles qui paraissent les plus importantes sont les erreurs homophoniques *c'est sait bien passé* pour « ça s'est bien passé ». Ceci étant, nous ne sommes confronté qu'à des graphies existantes dans la même logique de représentation utilisée, à savoir la logique standard. En comparant (27) et (27'), nous pouvons d'ailleurs confirmer le fait que le scripteur a bien conscience des limites normatives entre les différentes logiques de représentations graphiques dans la mesure où, orthographiquement, les logiques sont différentes, et où les registres lexicaux s'opposent également radicalement. En (27), le registre de langue est courant, alors qu'en (27'), il s'agit de lexiques spécifiques aux jeunes. En conséquence, même dans des cas de faible maîtrise linguistique, les jeunes ont conscience des limites normatives entre différentes logiques possibles pour transcrire la langue, et savent reconnaître la norme académique même s'ils ne la maîtrisent pas pleinement.

Enfin, une enquête linguistique avec la collaboration de plusieurs équipes d'enseignants nous a permis d'observer les usages des jeunes dans les copies, c'est-à-dire dans des productions considérées comme étant sérieuses. Très peu d'erreurs de logique normative ont été repérées. Une seule erreur relevant d'un mélange de codes a été relevée : il s'agit de *orait* au lieu de « aurait » dans une copie présentant très peu de fautes d'orthographe. Les graphies phonétisantes en « k » n'ont pas été identifiées à ce jour, ce qui n'exclut pas la possibilité qu'elles apparaissent de façon très singulière dans certaines copies de jeunes Français. D'autre part, quelques abréviations courantes ont été identifiées de façon occasionnelle en fin de copie à cause d'un manque de temps pour terminer l'examen, ce qui montre, là encore, qu'il n'y a pas de réelles confusions entre les différentes logiques de représentations graphiques.

#### 4 Conclusion

Les écrits atypiques produits par les jeunes dans les blogs ne présentent pas véritablement de règles orthographiques, mais des tendances d'usages marquées par des styles graphiques aux différentes fonctions. Le but est de se démarquer par l'utilisation d'un style particulier et de se faire comprendre. Autrement dit, les seules règles en vigueur sont celles qui concernent la concevabilité graphique et l'intercompréhension. Ceci étant, le lexique est affecté de différentes manières par les procédés et par la norme académique qui sont utilisés alternativement en fonction des cas. Les lexiques les plus touchés par les procédés sont les termes fréquemment utilisés ainsi que les lexiques spécifiques aux jeunes. Les lexiques spécifiques présentent beaucoup d'instabilité graphique lorsqu'il s'agit de séquences complexes, nouvelles, ou lorsqu'ils ne sont pas conventionnellement attestés dans les dictionnaires de langue. Par opposition aux lexiques usuels, les lexiques peu utilisés ou orthographiquement complexes sont généralement transcrits de façon normée même lorsque la logique globale de représentation est atypique. Enfin, même si ces écrits sont caractérisés par la variation graphique, certaines constantes dans les mécanismes en usage sont observables. Néanmoins, il ne s'agit pas d'une orthographe nouvelle dans la mesure où les normes standard sont utilisées, mais plutôt de représentations graphiques présentant des décalages normatifs dans l'utilisation divergente des codes et des signes graphiques élémentaires. Par conséquent, nos analyses

orthographiques nous permettent de comprendre les différents types de représentations graphiques du lexique afin de pouvoir le décoder, bien que celui-ci renferme d'autres particularités au niveau de ses usages syntaxiques et de sa diversité.

### Références bibliographiques

- Anis, J. (1999). Chats et usages graphiques. In Anis, J. (éd.), *Internet communication et langue française*. Paris : Hermès. 71-90.
- Anis, J. (2003). Communication électronique scripturale et formes langagières : chat et SMS. In *Actes des Quatrièmes Rencontres Réseaux Humains / Réseaux Technologiques*. Université de Poitiers. <http://edel.univ-poitiers.fr/rhrt/>.
- Anis, J. (2004). Les abréviations dans la communication électronique (en français et en anglais). In Andrieux-Reix, N., Branca-Rosoff, S., et Puech, C. (éd.), *Écritures abrégées (notes, notules, messages, codes...) : L'abréviation entre pratiques spontanées, codifications, modernité et histoire*. Paris/Gap : Ophrys. 97-112.
- Arrivé, M., Gadet, F., Galmiche, M. (1986). *La grammaire d'aujourd'hui : guide alphabétique de linguistique française*. Paris : Flammarion.
- Azor, C., Azor, M. M., Longepied, F., Nagau, C., Pérez, M., Rajef, I., Sarré, A., Sarré, A., Sylla, B., Sylla, K., Touré, D. (2007). *Lexik des cités*. Paris : Fleuve Noir.
- Catach, N. (2004). *L'orthographe*. 9<sup>e</sup> édition. Paris : P.U.F. collection « Que sais-je ? ».
- Fairon, C., Klein, J. R., Paumier, S. (2006). *Le langage SMS : étude d'un corpus informatisé à partir de l'enquête « faites don de vos SMS à la science »*. Louvain : Presse Universitaire de Louvain.
- Gadet, F. (1992). *Le français populaire*. Paris : P.U.F. collection « Que sais-je ? ».
- Gadet, F. (2007). *La variation sociale en français*. 2<sup>e</sup> édition. Paris/Gap : Ophrys.
- Goudaillier, J.-P. (2001). *Comment tu tchatches ! Dictionnaire du français contemporain des cités*. 3<sup>e</sup> édition. Paris : Maisonneuve et Larose.
- Pierre-Adolphe, P., Mamoud, M., Tzanos, G.-O. (1995). *Le Dico de la banlieue*. Boulogne : La Sirène.
- Pruvost, J., Lino, M. T. (2003). *Mots et lexiculture : hommage à Robert Gallisson*. Paris : Honoré Champion.
- Schifres, A. (1982). Le jeune tel qu'on le parle. *Le Nouvel Observateur*, 943, 60-62.

### Blogs

- <http://blackcat04.skyblog.com/>
- <http://caroline191301.skyblog.com/>
- <http://cycy-jtm-grave.skyblog.com/>
- <http://f4shii0n-viip.skyblog.com/>
- <http://fraicheur-kali.skyrock.com/>
- <http://fuck-les-rageux.skyblog.com/>
- <http://jimbo77220.skyrock.com/>
- <http://kos-tos1213.skyrock.com/>
- <http://lamiss-lapetasse.skyblog.com/>
- <http://leoufdu17.skyblog.com/>
- <http://lestatxxx.skyrock.com/>
- <http://lolo937sang.skyblog.com/>

<http://missespagne03.skyrock.com/>

<http://misskirri.skyblog.com/>

<http://misst3ak-91.skyblog.com/>

<http://nat190.skyblog.com/>

<http://reilly.skyblog.com/>

<http://snoop937sang.skyblog.com/>

<http://vryne-78-c.skyblog.com/>

<http://x-just-iin-l0v3-x.skyrock.com/>

<http://xx-la-rageuse-xx.skyblog.com/>

<http://xx-miss-tos-xx.skyrock.com/>

---

<sup>1</sup> Précisons que tous les jeunes n'écrivent pas de façon atypique sur les supports modernes de communication.

<sup>2</sup> Les blogs <http://kos-tos1213.skyrock.com/> et <http://missespagne03.skyrock.com/> qui sont en contact par réseau amical présentent les mêmes caractéristiques graphiques singulières. Ainsi, lorsqu'on effectue une recherche à partir du moteur de recherche de *skyblog.com* en tapant une séquence graphique singulière présente dans l'un des deux blogs pour savoir s'il s'agit d'un style graphique répandu, nous sommes face à ces deux blogs ainsi qu'à quelques autres.

<sup>3</sup> Traduction de cet exemple : « Sérieux si vous arrivez à tout lire sans en chier, vous êtes des progamers (= très forts, dans ce contexte). [...] Au passage j'en profite pour passer un message : si tu apprécies ces gens-là ou que tu te sens proche de la pétasse ci-dessous, ou encore si cette image te touche, eh bien n'hésite pas !!! Laisse l'adresse de ton blog dans les commentaires (rire). »

<sup>4</sup> De manière générale, les langages des jeunes écrits ou oraux sont représentés de façon stéréotypée à l'écrit. En effet, Schifres (1982), journaliste au *Nouvel Observateur*, en témoigne dans la mesure où il représentait déjà, bien avant la démocratisation du téléphone portable et d'Internet, le lexique des jeunes avec une orthographe phonétisante pour renvoyer un aspect « branché ».

<sup>5</sup> Les dictionnaires spécialisés et les listes de mots consultables dans des études spécifiques des langages des jeunes présentent beaucoup de variations graphiques. Par exemple, Pierre-Adolphe et al. (1995) écriront *cheulou* alors que Goudaillier (2001) et Azor et al. (2007) écriront plutôt *chelou*.

<sup>6</sup> Les écrits atypiques des blogs présentent un grand nombre de formes-mots ou de lettres-mots, c'est-à-dire de formes autonomes paraissant être des mots composant la phrase. En réalité, il s'agit le plus souvent d'agglutinations ou de formes résultant d'un procédé graphique représentant une séquence complexe, et non pas un lexème.

<sup>7</sup> Cette séquence est attestée dans le blog <http://xx-miss-tos-xx.skyrock.com/>, dans l'énoncé suivant : *Ste mOof Elle S'nOom d'Jenn[ii]fer*.