Dialectologie grapho-phonématique de deux communautés virtuelles : pour une approche discursive des communaulectes.

Laura Renée Gabrielle Goudet

▶ **To cite this version:**

Laura Renée Gabrielle Goudet. Dialectologie grapho-phonématique de deux communautés virtuelles : pour une approche discursive des communaulectes.. Linguistique. Université Paris-Nord - Paris XIII, 2014. Français. ⟨NNT : 2014PA131018⟩. ⟨tel-01230557⟩

HAL Id: tel-01230557

https://theses.hal.science/tel-01230557v1

Submitted on 18 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HAL Authorization

Université Paris XIII, Sorbonne Paris Cité

UFR de Lettres, Langues Sciences Humaines et des Sociétés

Département d'études anglophones

École doctorale ERASME (ED 493)

Laboratoire : CRIDAF- Pléiade (EA 453)

Thèse de linguistique anglaise

Présentée pour l'obtention du titre de

Docteur de l'Université Paris XIII – Sorbonne Paris Cité

par Laura Gabrielle Goudet

DIALECTOLOGIE GRAPHO-PHONÉMATIQUE DE DEUX COMMUNAUTÉS VIRTUELLES

POUR UNE APPROCHE DISCURSIVE DES COMMUNAULECTES

Soutenue le 1er Décembre 2014

Président :	Professeur Jean-Louis Duchet (Université de Poitiers)
Rapporteurs :	Professeur Philip Carr (Université Paul-Valéry-Montpellier 3)
	Professeur Jean-Michel Fournier (Université de Tours)
Examinateurs :	Professeur Nicolas Ballier (Directeur de thèse, Université Paris-Diderot)
	Professeure Marie-Anne Paveau (Université Paris 13-Sorbonne Paris Cité)

Remerciements

J'adresse mes premiers remerciements à Olivier Simonin, pour m'avoir présentée à Nicolas Ballier, autour d'une coupe de champagne à sa soutenance de thèse. Sans lui, rien n'aurait été possible, et je suis très heureuse qu'il ait pensé à cette collaboration, qui s'est avérée fructueuse.

Mes plus sincères remerciements vont à Nicolas Ballier, pour avoir supporté et déplié les Lauraïsmes sibyllins comme j'ai dû déplier les concepts complexes qu'il m'a présentés tout au long de ce travail. Le chemin était droit mais la pente était raide, et je suis très reconnaissante de l'avoir gravi avec un directeur de thèse d'une telle qualité. Je remercie particulièrement les membres du jury, Madame la Professeure Marie-Anne Paveau et Messieurs les Professeurs Philip Carr, Alain Deschamps, Jean-Louis Duchet et Jean-Michel Fournier, qui ont tous contribué humainement ou théoriquement à ce travail de thèse.

J'ai eu la grande chance d'avoir rencontré des personnes exceptionnelles pendant parcours universitaire, d'abord à Paris IV. Je remercie Richard Lilly pour avoir cru en moi dès le début et dont le soutien a été sans faille. Découvrir la constitution d'une théorie linguistique ou phonologique avec lui fut inoubliable, et je lui en serai éternellement redevable. Je suis également très reconnaissante à Leo Carruthers, Professeur hors-pair, dont la rigueur scientifique alliée à une intelligence perçante et humaine m'a confirmée dans ma volonté de poursuivre mes études jusqu'au doctorat. C'est émue que j'adresse des remerciements au regretté Michel Viel (1946-2011) dont l'esprit traversait si aisément la structure des mots ; et à feu André Crépin (1928-2013), si bienveillant malgré son érudition encyclopédique. Vous avez tous laissé une empreinte profonde dans la constitution de mon esprit scientifique.

Je pense à tous mes collègues doctorants (ou docteurs !), compagnons de fatigue ou de discussions passionnées. Il est possible que j'oublie certains noms dans la liste, mais j'espère que vous aurez la mansuétude de pardonner à une doctorante en fin de thèse. Tout d'abord, je remercie les membres du séminaire de sciences du langage de Marie-Anne Paveau : Noémie Marignier, pour nos séances de travail et ces discussions stimulantes, Catherine Ruchon, pour sa joie de vivre si contagieuse dans laquelle je vois une camarade de corpus alternatif, et Gaëlle Labartha, si patiente et qui joue aussi à l'équilibriste entre deux disciplines. Grâce à vous, je me suis sentie entourée et les perspectives ouvertes par nos parcours différents ont beaucoup contribué aux problématiques de cette thèse. Je suis reconnaissante aux doctorants de Nicolas Ballier, surtout Charles Mignot, pour son soutien lors de mes années à l'Université du Texas, Issa Kanté l'indéfectible ami discret, et Véronique Pouillon pour nos discussions *on-* et *off-topic*. Je remercie également Thomas Gaillat et Adrien Méli pour les aides technico-conceptuelles qu'ils m'ont apportées si gentiment.

Les collègues de mes disciplines m'ont accueillie avec amitié et indulgence, et je tiens à les remercier également, Sophie Herment et Gabor Tucsan en tête. Ils ont été des alliés formidables, autant humainement qu'intellectuellement, et ils ont rassuré mes débuts de communications en conférence.

Je retrouve toujours Paris-Diderot avec plaisir, grâce notamment à Anne Talbot, Philippe Martin, Alain Diana et Emmanuel Ferragne, dont l'aide en phonologie ou en programmation a été plus que précieuse–presque autant que nos conversations plus détendues ! Merci à Jérémy Castanier, compagnon de corpus en tableaux colorés, et de conférences. Je remercie également James Costa pour les conversations stimulantes sur « l'écosscots » et Cécile Rey pour son aide au début de la rédaction.

Je remercie l'ensemble des collègues de Rouen portés par Alain Lauzanne, qui m'ont fait confiance et m'ont entourée–surtout Anne Besnault-Levita, Cécile Vallée, Marc Martinez, Carine Lounissi et Laurence Costa (la liste est longue !). Merci pour ces moments de sérieux et de détente, pour les conseils toujours bienvenus. Grâce à vous, j'ai pu prendre des responsabilités et être écoutée, et je vous en suis reconnaissante.

Ce projet de thèse n'aurait pas abouti sans le forum PoC, dont je suis membre depuis 2002 et modératrice de section, pour m'avoir fait comprendre les mécaniques des forums. Ses membres et mes amis, surtout Engels Karl, *l'orgueil du parti* grâce à qui j'ai pu décompresser, wam [sic], Domina et Requyem, et Grumph, qui m'a donné tellement envie de faire une thèse ; je remercie les membres qui ont fréquenté le chan et surtout le CdC, à qui je dois le gros de mon *postcount*. Je me souviens encore de mes tâtonnements pour comprendre le vocabulaire des forumeux, entre les MàJ et autres Con(ventions) organisées. Mon univers numérique s'est encore accru quand, grâce à @mapav8, j'ai découvert la communauté de Twitter. Je remercie surtout Stéphanie Messal (@misanthropolog), Mélody Faury (@infusoir) pour les moments de détente et surtout le soutien pendant la #Redaction.

Merci à mes amis proches, qui ont suivi le déroulement de cette petite chose, surtout Florence Cabaret, si calme et si drôle. Je te remercie tout particulièrement pour les parenthèses pendant la rédaction, qui ont été essentielles à la conservation de mon pep légendaire. Merci à Annick Moisan et Adrien Lemaître pour l'amitié transatlantique, et dont j'ai suivi les péripéties de thèse avant d'en suivre les pas.

Merci à nonscience devenu Stephane Haas, pour sa patience infinie lorsque je lui dépliais mes théories linguistiques ; et pour avoir été repêcher des morceaux de brouillons égarés (à cause de divers complots félins). Merci d'être mon ami et mon acolyte depuis 11 ans.

Enfin, je clos cette longue liste de remerciements avec Anny Goudet, qui a supporté mes manifestations pour apprendre à écrire. Depuis le premier jour de la maternelle où je l'ai congédiée de cette phrase, elle sait que « j'ai beaucoup de travail. ». Je m'enivrais du bruit des touches de ton clavier, j'ai lu ta bibliothèque et tes articles, et je sais que tu es fière de moi autant que je suis heureuse de te montrer l'aboutissement de ce travail de longue haleine.

Table des matières

Table des Illustrations v
Liste des tableaux vii
Acronymes vii
Conventions Typographiques xvii

Introduction **1**

1 Analyse des graphies fictionnelles **13**
1.1 Caractérisation du corpus secondaire 14
1.2 *Push* et *Precious* 15
1.2.1 *Push* 16
1.2.1.1 Structures narratives de *Push* 16
1.2.1.2 Analyse onomastique 18
1.2.2 Typologie des graphies 20
1.2.2.1 Graphie gabaritique 21
1.2.2.2 Graphies alternatives 23
1.2.2.3 Modalité de la correction 24
1.2.2.4 Graphies transcriptionnelles 25
1.2.2.4.1 Consonnes 26
1.2.2.4.2 Voyelles 31
1.2.2.5 Étude épistolaire 35
1.2.2.5.1 Étude longitudinale des réélaborations graphiques 37

1.2.2.5.2 Discours métalinguistique et épilinguistique 42
1.2.3 *Precious* 44
1.2.3.1 Caractéristiques du média 45
1.2.3.2 Homogénéité des parlers 46
1.2.4 Traits communs des sous-titres et du texte 50
1.3 *Trainspotting* 53
1.3.1 Analyse multimodale 54
1.3.1.1 Structure narrative du livre 54
1.3.1.1.1 Variation interscripteurs 58
1.3.1.1.2 Scottismes 62
1.3.1.1.3 Argot d'Édimbourg et de Leith 64
1.3.1.2 Film 66
1.3.2 Graphies communes et scottismes dans les sous-titres et le livre 71

2 Caractérisation de l'Afro-Américain Vernaculaire (AAV) et du Scots 77
2.1 Sociolinguistique du statut du scots 80
2.1.1 Le scots 80
2.1.1.1 Recensement et place du scots 83
2.1.1.1.1 Explications sur la section langage du questionnaire et comparaisons 84
2.1.1.1.2 Diffusion du Scots 86
2.1.2 Utilisation du gaélique 89
2.2 Spécificités du scots 91
2.2.1 Lexique 91
2.2.2 Graphies 93
2.2.2.1 Graphie anciennes du Scots 94
2.2.2.2 Standardisation contemporaine des graphies 95
2.2.2.3 Typologie des écarts à la norme 99
2.2.3 Syntaxe 101

2.2.3.1 Particularités verbales . . . 101
2.2.3.2 Particularités nominales . . . 103
2.2.3.3 Prépositions et adverbes . . . 106
2.2.4 Phonologie . . . 108
2.2.4.1 Voyelles du Scots . . . 108
2.2.4.2 Comparaison entre les différentes variétés de scots . . . 110
2.2.4.3 Consonnes du Scots . . . 112
2.3 Présentation de l'Afro-Américain Vernaculaire . . . 114
2.3.1 Analyse critique des formes désignationnelles . . . 114
2.3.1.1 Désignation des locuteurs . . . 117
2.3.1.2 Représentations de l'AAV et précautions méthodologiques . . . 117
2.3.2 Interactions avec l'Anglais Blanc du Sud . . . 119
2.4 Examen des spécificités de l'AAV . . . 123
2.4.1 Lexique . . . 123
2.4.2 Graphie . . . 124
2.4.2.1 Authenticité des graphies scriptées . . . 128
2.4.3 Syntaxe . . . 130
2.4.3.1 Caractéristiques verbales . . . 131
2.4.3.2 Caractéristiques nominales . . . 133
2.4.3.3 Caractéristiques adverbiales . . . 134
2.4.4 Phonologie . . . 137
2.4.4.1 Réalisation des voyelles . . . 137
2.4.4.2 Réalisation des consonnes . . . 139
2.5 Statut sociolinguistique des systèmes considérés . . . 142
2.5.1 Dictionnaires et ouvrages de référence . . . 142
2.5.2 Prescriptivisme ou descriptivisme . . . 145
2.6 Analyse des représentations . . . 148
2.6.1 Statut du Scots et de l'AAV . . . 148

2.6.2 Portrait des locuteurs . . . 149
2.6.2.1 Dialectologie perceptive . . . 149
2.6.2.2 Différences sociales et ethniques . . . 150
2.6.3 Évaluation par les locuteurs d'autres dialectes . . . 152
2.6.3.1 *Ebonics* dans l'*Urban Dictionary* (UD) . . . 153
2.6.3.1.1 *Fo Shizzle my nizzle* . . . 155
2.6.3.1.2 Description des Afro-Américains . . . 157
2.6.3.2 *Scottish* dans l'*Urban Dictionary* . . . 159
2.6.3.2.1 Stéréotypie écossaise . . . 161

3 Essai de caractérisation des pratiques discursives sur les forums internet 167
3.1 Approche sémiologique de la matérialité des échanges dans les communautés virtuelles 173
3.1.1 Topographie des sujets . . . 176
3.1.2 Positionnement communicationnel . . . 180
3.1.3 Caractérisation sémiologique des échanges communicationnels . . . 184
3.1.4 Propriétés discursives de la communication asynchrone et des effets de citationnalité . . . 189
3.2 Analyse des systèmes écologiques . . . 194
3.2.1 Présentation de *Black Planet* . . . 194
3.2.1.1 Architecture de *Black Planet* (BP) . . . 198
3.2.2 Présentation de *Scotster* . . . 201
3.2.2.1 Architecture de *Scotster* . . . 203
3.2.3 Justification du choix et des limites du corpus . . . 208
3.3 Positionnement épistémologique et collecte des données . . . 209
3.3.1 Méthodologie de recherche . . . 210
3.3.1.1 Méthodologie de la constitution du corpus . . . 210
3.3.1.2 Problèmes lors de la constitution du corpus . . . 211
3.3.1.2.1 Problèmes techniques . . . 211
3.3.1.2.2 Problèmes humains . . . 211

3.3.1.3 Corpus . . . 212
3.3.2 Éthique de la recherche . . . 215
3.3.2.1 Questionnements concernant les ethnies . . . 215
3.3.2.2 Anonymat des données . . . 215
3.3.3 Choix des fils de discussion étudiés . . . 217
3.4 Essai de circonscription de l'identité en ligne . . . 218
3.4.1 Fragilité des données socio-démographiques . . . 220
3.4.2 Description fonctionnelle de l'avatar . . . 226
3.4.3 Caractérisation des pseudonymes . . . 228
3.4.3.1 Importance du pseudonyme . . . 228
3.4.3.2 Pseudonymes sur BP et *Scotster* (SC) . . . 230
3.4.3.3 'Etude de cas de deux matrices de pseudonymes sur BP . . . 231
3.4.4 Présentation personnelle et monologues . . . 236
3.5 Identités causées . . . 238
3.5.0.1 Néologismes des communautés virtuelles . . . 239
3.5.0.2 Paradoxe de l'observateur . . . 242
3.5.0.3 Positionnement du chercheur . . . 244

4 Analyse de communaulectes 249
4.1 Définition de communaulecte . . . 250
4.1.1 Communaulectes et langues minoritaires . . . 253
4.2 Équivalence graphophonémique en anglais standard . . . 257
4.2.1 Équivalences consonantiques . . . 258
4.2.2 Équivalences vocaliques . . . 260
4.3 Fautes d'orthographe . . . 262
4.3.1 Verbalisation . . . 264
4.3.2 Graphie . . . 265
4.3.3 Saisie sur clavier . . . 266
4.3.4 Néologismes et mots inconnus . . . 268

4.3.5 Dimension normative . . . 269
4.4 Graphies alternatives . . . 274
4.4.1 Utilisations propres à un forum . . . 275
4.4.1.1 Etude de cas : la parole noire et blanche sur internet . . . 275
4.4.2 Netspeak . . . 282
4.4.3 Mèmes . . . 284
4.4.3.1 Étude de cas : les LOLcats . . . 290
4.4.3.2 *Engrish : All your base are belong to us* . . . 294
4.4.4 Intersection entre mème et communaulecte . . . 300
4.4.4.1 Concept exprimé en un seul mot, une seule expression . . . 302
4.4.4.2 Créations Lexicales : *Derp* . . . 304
4.4.5 Censure et détournement . . . 306
4.4.5.1 Études de cas : la cryptographie de gangs . . . 306
4.4.5.2 Censure dans les messages . . . 309
4.5 Comparaison d'utilisation . . . 312
4.5.1 Types d'échanges . . . 312
4.5.1.1 Insultes et disputes . . . 312
4.5.1.2 Discours valorisant . . . 318
4.5.1.3 Discours empathique . . . 320
4.5.2 Fonction complice du terme communaulectal . . . 322
4.5.2.1 Réaction aux utilisations . . . 325
4.6 Le communaulecte, entre graphie et discours . . . 327
4.6.1 Originalité du discours . . . 327
4.6.1.1 Communaulecte graphique . . . 329

5 Graphophonématique des communaulectes 335
5.0.2 Remarques méthodologiques . . . 339
5.1 Production en langues minoritaires . . . 340
5.1.1 Étude des limites perçues du scots . . . 343

5.2 Variation allographique de la représentation des consonnes 348
5.2.1 Traitement des occlusives 353
5.2.1.1 Variantes graphiques des plosives sur SC 353
5.2.1.2 Variantes graphiques des plosives sur BP 354
5.2.1.3 Récapitulation des variantes graphiques des plosives dans le corpus 357
5.2.2 Traitement des fricatives 358
5.2.2.1 Variantes graphiques des fricatives sur SC 358
5.2.2.2 Variantes graphiques des fricatives sur BP 360
5.2.2.3 Récapitulation des variantes graphiques des fricatives dans le corpus 364
5.2.3 Traitement des affriquées 365
5.2.3.1 Variantes graphiques des affriquées sur SC 365
5.2.3.2 Variantes graphiques des affriquées sur BP 365
5.2.3.3 Récapitulation des variantes graphiques des affriquées dans le corpus 367
5.2.4 Traitement des nasales 367
5.2.4.1 Variantes graphiques des nasales sur SC 367
5.2.4.2 Variantes graphiques des nasales sur BP 368
5.2.4.3 Récapitulation des variantes graphiques des nasales dans le corpus 370
5.2.5 Rhoticité et /r/ 371
5.2.5.1 Variantes graphiques des <r> sur SC 371
5.2.5.2 Variantes graphiques des <r> sur BP 372
5.2.5.3 Récapitulation des variantes graphiques de <r> dans le corpus . . . 373
5.2.6 Traitement de /l/ et ses allophones 374
5.2.6.1 Variantes graphiques des <l> sur SC 374
5.2.6.2 Variantes graphiques des <l> sur BP 375
5.2.6.3 Récapitulation des variantes graphiques de <l> 376
5.2.7 Traitement des glissées 377
5.2.7.1 Variantes graphiques des glissées sur SC 377
5.2.7.2 Variantes graphiques des glissées sur BP 378

5.2.7.3 Récapitulation des variantes graphiques des glissées dans le corpus 379
5.2.8 Réduction de groupes consonantiques . . . 380
5.3 Variation allographique des représentations des voyelles . . . 382
5.3.1 Variations graphiques des voyelles brèves en AAV . . . 386
5.3.2 Variations graphiques des voyelles brèves sur SC . . . 388
5.3.3 Multigraphies pour voyelles inaccentuées . . . 392
5.3.4 Diphtongues . . . 395
5.4 Tests de la spécificité des mots sélectionnés . . . 400
5.4.1 Méthode . . . 400
5.4.2 Résultats . . . 401
5.4.3 Discussion . . . 402
5.5 Études de cas lexicaux . . . 403
5.5.1 Variations graphiques de *nigger* . . . 405
5.5.2 Variations graphiques de *bitch* . . . 411
5.5.3 Variations graphiques de *fuck* . . . 416
5.5.4 Autres phénomènes de variances graphiques . . . 423

Conclusion **429**

Bibliographie **435**

A Annexe **461**
A.1 Arbre de parenté des langues . . . 461
A.2 Precious . . . 463
A.2.1 Inventaire des équivalences graphophonématiques des voyelles simples . . . 463
A.2.2 Entrées du journal de Precious dans *Push* . . . 463
A.2.3 Casting du film . . . 464
A.3 Trainspotting . . . 464
A.3.1 Casting du film . . . 464
A.4 Classes d'isophones - Récapitulation . . . 466

A.5 Fréquence des lettres en anglais . . . 466
A.6 *Scotster* et *Black Planet* . . . 467
A.6.1 *Scotster* . . . 467
A.6.1.1 Welcome to Scotster . . . 467
A.6.1.2 Graphies alternatives - SC . . . 468
A.6.1.3 Architecture de SC . . . 469
A.6.2 *Black Planet* . . . 469
A.6.2.1 Architecture de BP . . . 469
A.6.3 Graphies alternatives - BP . . . 469
A.7 Liste de lieux virtuels évoqués dans cette thèse . . . 470
A.7.1 Données avec distances de Levenshtein . . . 472
A.8 Résultats du test de Wilcoxon . . . 473
A.9 Codes Informatiques . . . 474
A.9.1 Macros en Visual Basic Advanced [Excel] . . . 474
A.9.1.1 Macro pour le calcul pondéré de la distance de Levenshtein-Damerau sous Excel . . . 474
A.9.1.2 Macro pour le calcul de la proportion de similarité orthographique 477
A.10 Travaux universitaires réalisés au cours de cette thèse . . . 478
A.10.1 Communications . . . 478
A.10.2 Publications . . . 482
A.10.3 Autres travaux . . . 482
A.11 Utiliser la clef USB d'annexes . . . 483

Index **485**

Index Nominum **485**

Index Rerum **489**

Liste des Illustrations

1 Organisation du plan de thèse 10

1.1 Entrée du journal de Precious avec corrections 22
1.2 Répartition des graphies des sons vocaliques 33
1.3 Schéma récapitulatif des graphies de Precious 34
1.4 Exemple de correction par Precious (p.98) 36
1.5 Occupation spatiale d'une graphie squelettique (p. 70) 36
1.6 Exemple de remplissage dans Precious (p.98) 37
1.7 Pourcentage d'erreurs dans les entrées 38
1.8 Résultats de Flesch-Kincaid et courbe de tendance 40
1.9 Poème « Rootine Break », entrée 24 du journal de Precious (partie V) 41
1.10 Extrait de *Precious* où la protagoniste utilise sa voix de star 47
1.11 Titre du film 51
1.12 Extrait de Precious avec sous-titres (00 :41 :30) 52
1.13 Extrait de Precious avec sous-titres (00 :49 :00) 52
1.14 Proportion de discours suivant les scripteurs 57
1.15 Localisation de Leith à Édimbourg (d'après Skilling, 2012 et Google Maps) 64
1.16 Captation sonore de Spud (00 :14 :47-00 :14 :48) 68
1.17 Un exemple de scottisme dans les sous-titres de *Trainspotting* 72
1.18 Panorama des usages des graphies alternatives et insertion écologique 75

2.1 Différents dialectes du scots proposés à l'écoute (*Aye Can*) 87
2.2 Distribution du Scots (**shepherd_historical_1956** ; Ross, 2001) 88

2.3 Répartition des locuteurs du gaélique écossais en 2001 (General Register Office (Scotland), 2005) . . . 90
2.4 Correspondances phonographématiques principales anglais-scots . . . 96
2.5 Voyelles en Scots (Scoobie, Gordeeva & Matthews, 2006)Nominum]Scoobie, J. . . 108
2.6 Aire où l'Anglais Blanc du Sud (US) (ABS) est pratiqué (Labov, 2006)Nominum]Labov, W. . . . 120
2.7 Voyelles réparties selon leurs traits articulatoires (Kortmann, 2004)Nominum]Kortmann, B. . . . 137
2.8 Monophtongues de *General American* (GA) (Ladefoged, 1993, 1999)Nominum]Ladefoged, P. . . . 138
2.9 Entrée la plus populaire d'*Ebonics* avec nombre de votes, au 17 février 2013. . . . 153
2.10 Proportion de définitions et d'exemples caricaturaux et neutres . . . 158
2.11 Représentation stéréotypique individuelle associée à « Scottish » sur l'*Urban Dictionary* (UD) . . . 162

3.1 Trois états du clavier de téléphone (iPhone 4S) . . . 170
3.2 Captures d'écran de l'application BP (source : iTunes Store) . . . 171
3.3 Tweet du 9 février 2013 . . . 172
3.4 Photographies du profil d'un membre de SC (SCPH4) . . . 176
3.5 Capture d'écran du début de SCPH3, photographie et commentaires . . . 177
3.6 Extrait de Corpus En Contexte (CEC) (SCP10) . . . 179
3.7 Page d'accueil de SC au 29 juillet 2014 . . . 181
3.8 Page d'accueil de SC pour les inscrits, au 29 juillet 2014 . . . 182
3.9 Page d'accueil de BP au 29 juillet 2014 . . . 183
3.10 Extrait du fil « Random Thoughts MMXIV » (BP12E5582) . . . 185
3.11 Extrait du fil « Hi Fae Dundee » SC1A82 . . . 186
3.12 Facteurs communicationnels et fonctions du langage (Jakobson ; source : Gibbons, 1997) . . . 188
3.13 Schéma de la communication appliqué aux forums . . . 188
3.14 Extrait de « Random Thoughts 2013 » (BP12E17) . . . 191
3.15 Schéma explicatif . . . 192

3.16 Extrait de « Loch Ness monster cited as “proof” of creationism » (SC11F16) 193
3.17 Architecture des sections et sous-sections de BP 199
3.18 Organisation des sections du forum 200
3.19 Extrait de la première page de la sous-section « Customizing » de la section Cars . . 201
3.20 Extrait de la page « Browse Forums » de SC 204
3.21 Nuage de mots-clef du menu « Explore Tag » de SC (2 août 2014) 205
3.22 Carte des sections et sous-sections de SC 206
3.23 La page de choix de « race » et « ethnie » sur BP 216
3.24 Panorama de l’identité numérique (Fred Cavazza, 2006b) 219
3.25 Répartition de 55 membres de gangs autoproclamés (BP) 222
3.26 Répartition géographique des participants d’un fil sur la généalogie écossaise (SC2V13) 224
3.27 Répartition géographique des participants d’un fil sur les langues écossaises (SC2N59) 225
3.28 Comparaison de profil avec et sans photo (SC) 227
3.29 Avatar neutre de membre sans photo (BP) 228
3.30 Commentaires enchâssés sur *Slashdot* 229
3.31 Message d’un compte fermé, d’un compte ouvert sans avatar et d’un compte ouvert avec avatar (BP) 230
3.32 Trois exemples de renseignements personnels préliminaires à un message (SC) 231
3.33 Distribution par genre des pseudonymes utilisant *bitch* 233
3.34 Distribution par genre des pseudonymes utilisant *nigga* 234
3.35 Mème « Sudden Clarity Clarence » illustrant le couple *dawg*/*bitch* 235
3.36 Extrait de profil sur BP 237
3.37 Profil d’un membre de SC (SCP12) 238
3.38 La page d’accueil « What’s New ? » de BP au 28 juin 2013 241
3.39 le « Pulse » de SC au 28 juin 2013 242

4.1 Équivalences graphophonémiques pour les voyelles de l’anglais. 260
4.2 Modèle de Ciura & Deorowicz (Deorowicz & Ciura, 2005) 263
4.3 Le logo non officiel des « Grammar Nazi » 270

4.4 Utilisation 1 de « Grammar Nazi » (BP12E3357) . . . 271
4.5 Utilisation 2 de « Grammar Nazi » (BP12E699) . . . 272
4.6 Image accompagnant « Nigger Language »—originellement légendée « Black Honor Student » . . . 279
4.7 Extrait de la page de *Know Your Meme* qui recense les *Rage Comics* . . . 286
4.8 Un exemple du mème « Philosoraptor », auteur inconnu (Source : www.diylol.com . 287
4.9 Les 6 mèmes du « niveau divin »de *Meme Generator*. Le nom et le nombre de votes positifs figurent sous chaque image. . . . 289
4.10 Photographie de 1905 (Harry Whittier Frees) . . . 291
4.11 Le premier lolcat, créé par Cheezburger (Eric Nakagawa) . . . 292
4.12 Image postée sur engrish.com en 2008, tirée d'une bannière d'élection de Konomiya (Japon) . . . 295
4.13 Intérêt et nombre de recherches pour « All your base » (en bleu) et « All your base are belong to us » (en rouge) entre 2004 et 2011. . . . 297
4.14 Une des histoires en greentext publiée sur 4Chan . . . 301
4.15 Extrait d'un profil personnel . . . 309
4.16 Page d'inscription au forum *Scotster* . . . 325
4.17 Variante de « Ceci est un forum, pas un portable » . . . 330

5.1 Extrait de profil personnel (BPPP17) . . . 355
5.2 Position et zone des doigts typiques sur un clavier QWERTY (source : Wales et Sanger, 2001) . . . 376
5.3 Occurrences de <massa> dans la littérature (Google Ngrams, résultats du 23 août 2014)382
5.4 Exemples de graphies alternatives sur voyelles brèves en AAV . . . 386
5.5 Changement vocalique afro-américain (Thomas, 2007)Nominum]Thomas, E. . . . 387
5.6 Exemples de regraphies sur voyelles brèves en scots . . . 389
5.7 Évolutions vocaliques suggérées par le corpus SC . . . 391
5.8 Écran de fin du jeu « Coo Game » . . . 396
5.9 Rubrique du site SC . . . 396
5.10 Extrait de profil comportant la rubrique « Heilan Coo » . . . 397
5.11 Classement des formes de <nigger> en fonction de leur fréquence sur BP . . . 406

5.12 Dendrogramme des distances entre les graphies de <nigger> 409
5.13 Classement des formes de <bitch> en fonction de leur fréquence sur BP 413
5.14 Dendrogramme des distances entre les graphies de <bitch> 415
5.15 Classement de toutes les formes de <fuck> en fonction de leur fréquence sur BP . . . 416
5.16 Classement des formes de <fuck> étudiées, en fonction de leur fréquence sur BP . . . 417
5.17 Extrait de livre d'or d'un profil contenant <fucc> et <fuccn> (BPGB5) 419
5.18 Extrait de profil contenant <fuccn> (BPPP28) 420
5.19 Publicité de la marque FCUK (1997) 420
5.20 Dendrogramme des graphies de <fuck> 422
5.21 Extraits de réponses à un statut contenant <noe>, '*now*' (SCP9) 423

A.1 Carte de l'origine géographique des acteurs de *Trainspotting* 465
A.2 Fréquence des lettres en langue anglaise (Lewand, 2000) 467
A.3 Panorama des réseaux sociaux 470

Liste des tableaux

1.1 Contenu de Push . . . 18
1.2 Groupement des graphies gabaritiques de la seconde entrée du journal de Precious . . 22
1.3 Réélaborations graphiques des monophtongues . . . 27
1.4 Retranscriptions erronées utilisant <u> . . . 32
1.5 Récapitulation des phonostyles de Precious . . . 47
1.6 Répartition des parlers des personnages de *Precious* . . . 49
1.7 Composition de *Trainspotting* . . . 56
1.8 Proportion de mots imputée à chaque personnage . . . 58
1.9 Les 25 lexèmes les plus courants dans les narrations de Renton et de Spud. . . . 60
1.10 Amplitude allographique de <glaikit> . . . 63

2.1 Questions figurant sur les formulaires familiaux et individuels de recensement 2001 et 2011 . . . 84
2.2 Résultats de Macafee et McGarrity, 1999 portant sur l'utilisation du scots . . . 92
2.3 Réalisation des voyelles d'après Stuart-Smith, 1999 . . . 109
2.4 Voyelles et réalisations vocaliques du scots, comparé à l'anglais écossais standard . . 110
2.5 Différences dialectales dans la prononciation des classes d'isophones(Stuart-Smith, 2003)Nominum]Stuart-Smith, J. . . . 112
2.6 Répartition des consonnes (Kortmann, Upton, Schneider, Burridge & Mesthrie, 2008)Nominum]Kortmann, B. . . . 112
2.7 Traits phonologiques saillants de l'AAV d'après Mufwene, 1998 . . . 122
2.8 Comparaison entre les scripts automatiques d'afro-américanisation . . . 129

2.9 Spécificités de l'AAV suivant la zone géographique d'après Kortmann et Lunkenheimer, 2008 . . . 136
2.10 Réalisations vocaliques en AAV(Kortmann, 2004)Nominum]Kortmann, B. . . . 139
2.11 Consonnes de l'AAV classées par mode et lieu d'articulation . . . 139
2.12 Variations consonantiques en AAV . . . 141
2.13 Les dix mots les plus fréquents dans les exemples de l'UD . . . 155

3.1 Éléments d'environnement influant sur le discours en ligne . . . 168
3.2 Conditions de communication par type de plateforme virtuelle . . . 174
3.3 Commentaires 1 - 4 de la photographie 3.5 (SCPH3) . . . 177
3.4 Message 612 de *Random Thoughts 2011* . . . 178
3.5 Comparaison des théories de Rourke et Jakobson . . . 187
3.6 Caractérisation des interactions sur BP . . . 196
3.7 Architecture de Scotster . . . 207
3.8 Récapitulatif de la nomenclature des indices extra-forums de BP et SC . . . 213
3.9 Répartition des pseudonymes en *bitch* et *nigga* par graphie . . . 232

4.1 Correspondances phonie-graphie des consonnes en anglais standard . . . 259
4.2 Messages 1,3 de « Anyone get tricked by a chick with a gurdle ? » (BP12E5532) . . . 265
4.3 Messages 62 - 78 de *Stupid Black Women on Black Planet.....Ironic* (BP12E3630) . . . 273
4.4 Réponses au sondage sur « Whitemanitus » . . . 278
4.5 Extrait de l'introduction de Zero Wing traduite en anglais. . . . 296
4.6 Messages 8 - 12 de *Coincidence ? Or...lightskinned-racism ???* (BP12E932) . . . 305
4.7 Expressions et graphèmes tabous pour les membres des Bloods et des Crips . . . 307
4.8 Extraits de *You porch monkeys are just subhuman, degenerate rapists and murderers* (BP3D22) . . . 315
4.9 Réactions au message initial de *Peter D. is single and lookin* (SC20E5) . . . 317
4.10 Messages 43 - 53 de *Loss of our Language* (SC2X50) . . . 319
4.11 Messages 1 - 20 de *With Sadness* (SC25K1) . . . 322
4.12 Messages 84 - 87 de *Loss of our language* (SC2N59) . . . 325

4.13 Messages 3 et 10 de *Scottish Jokes* (SC2Y29) . . . 326
4.14 Messages 51 - 52 de *Loss of our language* (SC2N59) . . . 328

5.1 Nombre d'occurrences d'[EFTER] et [AHINT] sur Scotster . . . 336
5.2 Exemples de substitutions alphanumériques dans le corpus . . . 340
5.3 Messages 71 - 78 de *Weekly Scots Vocab List* (SC2N27) . . . 345
5.4 Messages 135 - 136 de *Loss of our language* (SC2N59) . . . 347
5.5 Messages 28 - 30 de *Loss of our language* (SC2N59) . . . 347
5.6 Hiérarchie des sons (ZWICKY, 1972) . . . 349
5.7 Archigraphèmes des consonnes de l'anglais . . . 351
5.8 Variation des graphies des occlusives en scots . . . 353
5.9 Variation des graphies des plosives en AAV . . . 354
5.10 Récapitulation des changements graphiques impliquant des occlusives dans les deux corpus . . . 357
5.11 Variation des graphies des fricatives en scots . . . 358
5.12 Variation des graphies des fricatives en AAV . . . 360
5.13 Distribution des changements graphiques de <th> . . . 363
5.14 Récapitulation des changements graphiques impliquant des fricatives dans les deux corpus . . . 364
5.15 Variation des graphies des affriquées en scots . . . 365
5.16 Variation des graphies des affriquées en AAV . . . 366
5.17 Récapitulation des changements graphiques impliquant des affriquées dans les deux corpus . . . 367
5.18 Variation des graphies des nasales en scots . . . 368
5.19 Variation des graphies des nasales en AAV . . . 368
5.20 Récapitulation des changements graphiques impliquant des nasales dans les deux corpus 370
5.21 Variation des graphies des /r/ en scots . . . 371
5.22 Variation des graphies des /r/ en AAV . . . 372
5.23 Récapitulation des changements graphiques impliquant des /r/ dans les deux corpus . 374
5.24 Variation des graphies des latérales en scots . . . 374

5.25 Variation des graphies des liquides en AAV . . . 375
5.26 Récapitulation des changements graphiques impliquant des fricatives dans les deux corpus . . . 376
5.27 Variation des graphies des glissées en scots . . . 377
5.28 Variation des graphies des glissées en AAV . . . 378
5.29 Récapitulation des changements impliquant des glissées dans les deux corpus . . . 379
5.30 Variation des groupes consonantiques . . . 380
5.31 Tentative d'inventaire des archigraphèmes des voyelles de l'anglais . . . 384
5.32 Voyelles et classes d'isophones à partir de DESCHAMPS, 1994 . . . 385
5.33 Distribution graphique des représentations de COMMA LETTER et HAPPY dans les deux corpus . . . 392
5.34 Changements graphiques qui suggèrent des monophtongaisons dans le corpus SC . . . 395
5.35 Changements graphiques qui suggèrent des diphtongaisons dans le corpus BP . . . 398
5.36 Changements impliquant des diphtongues dans les deux corpus . . . 399
5.37 Décomposition phonologique et graphique de nigger . . . 405
5.38 Variation allographique de <nigger> . . . 406
5.39 Décomposition phonologique et graphique de *bitch* . . . 411
5.40 Variation allographique de <bitch> . . . 414
5.41 Décomposition phonologique et graphique de *fuck* en AAV . . . 416
5.42 Variation allographique de <fuck> . . . 418
5.43 Messages 106 - 113 de *Loss of our Language* (SC2N59) . . . 425

A.1 Équivalences phonographématiques des voyelles simples <a, e, i, o> dans les erreurs graphiques de Precious . . . 463
A.2 Casting de *Precious* et provenance des acteurs . . . 464
A.3 Casting de *Trainspotting*et provenance des acteurs . . . 464
A.4 Voyelles de l'AAV et du scots, comparées aux classes d'isophones du GA et du *Received Pronunciation* (anglais britannique) (RP) . . . 466

Acronymes

AAV Afro-Américain Vernaculaire 77
ABS Anglais Blanc du Sud (US) 114
ADSL *Asymmetric Digital Subscriber Line* 172
AES Anglais Écossais Standard 91
ASS Anglais Standard du Sud (GB) 99
AYB *All your base [are belong to us]* 294
BET *Black Entertainment Television* 197
BNC *British National Corpus* 98
BP *Black Planet* 79
CEC Corpus En Contexte 176
CMO Communication Médiée par Ordinateur 168
CNDP Centre National de Documentation Pédagogique 89
CNIL Commission Nationale de l'Informatique et des Libertés 217
COCA *Corpus of Contemporary American English* 214
DOST *Dictionary of the Older Scottish Tongue* 82
DSL *Dictionary of the Scots Language* 148
ED *Encyclopaedia Dramatica* 276
FAQ Foire aux Questions 84
FB Facebook 180
GA *General American* 138
GLOWBE *Corpus of Global Web-Based English* 214
GTA *Grand Theft Auto* 116
IRC *Internet Chat Relay* 195
IRL *In Real Life* 172
LPD *Longman Pronunciation Dictionary* 386
MSN *Microsoft (Messenger Service) Network* 174
NTIC Nouvelles Technologies de l'Information et de la Communication 168

OSD *Online Scots Dictionary* 147
PWI *Predominantly White Institution* 119
RP *Received Pronunciation* (anglais britannique) 260
SA *Something Awful* 188
SC *Scotster* 91
SCOTS *Scottish Corpus of Text and Speech* 101
SDD *Scots Dialect Dictionary* 82
TIC Technologies de l'Information et de la Communication 168
TLF Trésor de la Langue Française 79
TR-F *Trainspotting* (Film) 359
TR-L *Trainspotting* (Livre) 359
UD *Urban Dictionary* 78

Conventions Typographiques

Les conventions graphiques suivantes seront utilisées :

Les lexèmes isolés en langues étrangères seront écrits en italiques. Les traductions en anglais seront encadrées de guillemets simples (‘ ’), et ne seront traduits en français que lorsque ces mots seront peu fréquents en anglais, ou que cela sera jugé nécessaire. Les citations longues seront mises en forme suivant les règles de la typographie utilisées dans la feuille de style (police 10…) et les traductions en anglais ou en français seront données en vis-à-vis, lorsque cela sera jugé pertinent. Les citations de forums sont copiées *verbatim* : les fautes d'orthographe ne seront pas toujours marquées par un [sic].

Les transcriptions phonétiques seront encadrées de crochets carrés [], les transcriptions phonémiques, de barres obliques // et les séquences graphiques de chevrons <>. Les exemples tirées d'autres corpus (littéraires et autres) seront encadrés par des guillemets français (« »). J'utilise les abréviations courantes de représentations phonétiques (<C> pour consonne, <V> pour voyelle…), les petites majuscules pour les classes d'isophones (exemple : KIT) et les lemmes sont en petites majuscules entourées de crochets carrés (exemple : le lemme $[\text{CHILD}]_N$). Les gloses et traductions sont précédées de guillemets anglais simples : <‘*exemple*’>.

L'ancienne notation des affriquées /ʤ/ et /ʧ/ est préférée à /$\widehat{\text{dʒ}}$/et /$\widehat{\text{tʃ}}$/.

La figuration de lignes de code ou de script en HTML sera faite en police Courier, ainsi que les adresses de sites internet dans la bibliographie. Cette thèse a été composée sous LaTeX, et ce choix oriente ainsi la typographie.

L'orthographe rectifiée définie par l'Académie française en 1990 est utilisée dans cette thèse, notamment l'utilisation des mots composés soudés (*plateforme*), et le changement de certaines accentuations (accents circonflexes, accent grave). Les mots étrangers « scots » et « netspeak » ne sont pas italicisés. Ce sont des représentants d'une terminologie linguistique qui n'existe pas en français, et leur intégration à part entière au vocabulaire utilisé dans ce travail me semble importante, sans avoir recours aux italiques qui les singulariseraient.

Malgré tout le soin apporté aux tableaux, ils sont considérés comme des objets flottants, et n'apparaissent pas forcément à la page où ils sont évoqués. Une recherche du lien hypertexte (dans la version .pdf de ce mémoire) sera parfois nécessaire pour pallier ce problème que je n'ai pas toujours pu résoudre.

Pour la version .pdf de la thèse, les hyperliens actifs sont indiqués par des couleurs différentes :

— La table des matières et les références internes sont indiquées en rouge. De la même façon, un clic sur les abréviations renvoie à leur liste dans le préambule.

— pour les éléments bibliographiques, les dates dans le corps du texte (en vert) renvoient aux références complètes, dans la bibliographie.
— les liens hors-thèse (en ligne) sont écrits en rose. Un clic ouvre un nouvel onglet dans votre navigateur.

Le .pdf de thèse a également une table des matières cliquable (onglet « signets » dans votre logiciel de lecture si la table des matières n'apparaît pas automatiquement à gauche de votre écran). Pour utiliser les logiciels et plateformes anti-plagiat (type Compilatio.net), préférez les versions .tex des fichiers pour éviter des problèmes de repérages liés à la compilation de LaTeX. Pour ouvrir les fichiers .tex, vous pouvez utiliser Notepad ++ si vous ne disposez pas d'éditeur (Miktex + Texmaker pour Windows, par exemple). Un premier passage d'une version antérieure montrait 6% de copie exacte d'une autre source, à cause des citations longues de *Push*, déplacées depuis dans les annexes sur clef USB.

Introduction

Présentation

Le projet de cette thèse vient de la nécessité de vouloir décrire la communication écrite sur internet, une forme d'écriture qui n'est ni orale, ni tout à fait écrite, et qui s'éloigne des prescriptions associées à l'écrit. Cette écriture sur internet n'est pas la langue des dominants, et la notion d'orthographe est loin des pratiques, qui s'en éloignent volontairement. Cette thèse est portée par la conviction que les graphies alternatives volontaires sur internet sont natives d'un discours que je me propose d'exposer ici.

La distinction entre langue et langage, la faculté humaine à communiquer et les divers systèmes linguistiques est ancienne et il a été établi que la structure des langues pouvait être résumée à un réagencement d'éléments de base : « une langue ne comporte jamais qu'un nombre réduit d'éléments de base mais ces éléments […] se prêtent à un grand nombre de combinaisons. »(Benveniste, 1966). Cette perspective ne prévoyait certainement pas le réagencement des graphèmes, et les changements sémiotiques créés par de simples substitutions par rapport aux mots originaux, comme un <y> à la place d'un <i>. La graphie <bytch> est souvent plus valorisante que son équivalent orthographique ; et <Amerikkka> n'a pas le même sens qu'« America ». Le nombre réduit d'éléments de base se prêtent donc à des combinaisons non prévues par la norme de référence qui a cours actuellement. Le sens ne se donne pas de la même manière selon la graphie. Pour le dire en termes frégéens, le sens (*Bedeutung*) est tout à fait dépendant du référent (*Sinn*) (Frege, 1892). Rien n'est moins vrai en ce qui concerne

les langues minoritaires proches de l'anglais, qui sont très proches d'une forme standard.

Cette thèse s'intéresse à la figuration graphique du scots et de l'AAV sur des forums internet de niche. Un site de niche s'inspire des niches de marché, où un sous-ensemble de consommateurs sont concernés par l'offre. Par exemple, le marché du *retrogaming* qui ne s'adresse qu'à une sous-catégorie des joueurs de jeux vidéo qui aiment les jeux anciens. Les deux forums sélectionnés sont dits « de niche » parce qu'ils n'adressent leurs services qu'aux utilisateurs qui ont un intérêt pour la culture afro-américaine, ou écossaise (par extension, qui sont Afro-Américains ou Écossais), à l'encontre de site communautaires comme *4Chan* ; ou de réseaux sociaux comme *Facebook*, qui s'adressent à tous.

L'AAV et le scots ont été sélectionnés parce qu'elles partagent des caractéristiques communes. Ce sont deux langues minoritaires de deux pays anglophones, parlées par deux groupes aux pratiques langagières différentes de la norme. Les locuteurs du scots et de l'AAV parlent un sociolecte qui s'étend sur tout le pays, avec des sous-dialectes distincts. L'AAV et le scots sont tous les deux des langues stigmatisées, non enseignées à l'école, mais elles sont toutes les deux documentées grâce à des corpus. Certains, comme l'eWAVE (*electronic World Atlas of Varieties of English*) sont généraux et traitent de toutes les variantes de l'anglais, créoles et langues minoritaires incluses (Kortmann & Lunkenheimer, 2008) ; d'autres sont textuels et oraux, comme le *Scottish Corpus of Text and Speech* (SCOTS) (Corbett et al., 2007). Certains représentent des états de langue antérieurs, comme le corpus BLUR (Blues Lyrics Collected at the University of Regensburg) (Miethaner & E. Schneider, 1997) qui regroupe des paroles de chansons de blues (Miethaner, 2012). Ce ne sont que quelques exemples de corpus importants et disponibles à tous, sur internet ou CD-ROM. Ces deux langues sont parlées par des locuteurs qui maîtrisent la variété de l'anglais standard de leur pays, et parfois, oublient ou ignorent qu'ils parlent également une langue à part entière. Ces deux langues doivent cohabiter dans la même zone linguistique avec d'autres variétés de l'anglais. Pour le scots, les langues et dialectes voisins sont l'anglais britannique standard, ainsi que l'écossais standard, et, dans une moindre mesure, le gaélique écossais au nord de la ligne Oban-Perth (Durand, 2004). L'AAV est en contact constant avec les divers parlers de l'anglais américain, et partage certaines de ses caractéristiques phonologiques avec l'anglais blanc du sud, notamment. Cette proximité géographique ne doit pas faire oublier

qu'il existe des similarités entre l'AAV et d'autres parlers vernaculaires, et autres dialectes de l'anglais américain, étant donné que l'AAV est utilisé dans l'ensemble des États-Unis. Je ferai une comparaison entre l'AAV et la variété d'anglais parlé par les Blancs au sud des États-Unis.

Les recherches sur internet, et sur les propriétés discursives des forums de discussion soulignent le caractère novateur de ces discussions, qui complexifient les analyses conversationnelles, notamment (Marcoccia, 2004). Ces corpus sont archivés automatiquement, comme le fait remarquer D. Maingueneau :« Ces corpus sont homogènes, définis par leur mise en mémoire et par le dispositif ou l'institution qui a assuré cette mise en mémoire » (Maingueneau, 1991), ce dont les participants ont conscience. Cette conservation, alliée au fait qu'il n'y ait pas besoin d'observateur pour que les données soient préservées, rend ce terrain idéal pour les analyses conversationnelles. Si l'on reprend une définition fonctionnelle des forums de discussion comme celle-ci : « Les forums de discussion peuvent être définis comme des dispositifs hybrides de communication interpersonnelle de masse (Lievrouw & Livingstone, 2002) dans la mesure où ils permettent à la fois l'échange interpersonnel (A répond à B) et la communication de masse (A poste un message lisible par un nombre potentiellement illimité d'internautes). » (Marcoccia, 2004). La communication est tout à fait définie du point de vue conversationnel, mais pas au niveau écologique, dans toutes les potentialités que le forum apporte. La mise en forme du message, de la présentation du scripteur, les liens hypertextes disponibles, la modélisation des données participent tous de l'environnement écologique de la plateforme de discussion. Par exemple, les membres de SC sont présentés avec leur pseudonyme et leur vrai nom, au format <(Prénom Initiale du patronyme)>, et ils sont nombreux à avoir renseigné un prénom et donné un pseudonyme, alors que les membres de BP n'utilisent que des pseudonymes, parce que le forum ne leur donne pas cette possibilité.

Le forum est donc aussi l'endroit où montrer son identité numérique, définie ainsi : « L'identité numérique d'un individu est composée de données formelles (coordonnées, certificats…) et informelles (commentaires, notes, billets, photos…). Toutes ces bribes d'information composent une identité numérique plus globale qui caractérise un individu, sa personnalité, son entourage et ses habitudes. » (Fred Cavazza, 2006a). Cette thèse met l'identité numérique et l'identification aux valeurs d'une

communauté internet au centre des problématiques abordées. L'hypothèse d'un communaulecte, qui serait un ensemble de configurations discursives propres aux communautés de scripteurs des forums de niche conduit à problématiser sur l'identité. L'identité numérique telle qu'elle est montrée dans les profils personnels et les pseudonymes sera étudiée, parce qu'elle est un vecteur important de graphies communaulectales (notamment chez les membres de gangs autoproclamés, par exemple).

L'AAV et le scots ne sont pas au même niveau de développement. Si l'AAV influence la culture populaire aux États-Unis grâce à la musique (rap, R'N'B, hip hop) interprétée par des Afro-Américains (Kermit E Campbell, 2007), le scots est une langue en érosion, qui ne s'exporte pas, mais qui est soutenue par ses locuteurs en tant que marqueur d'identité patriotique (Costa, 2009). Un mouvement identique (rejet de certaines valeurs de l'Amérique blanche) se retrouve dans l'AAV, qui incorpore également des resémantisations militantes (« Amerikkka » pour *America*). Dans les deux cas, ces langues servent en priorité à la communication orale, et si les échanges écrits sont secondaires, la production littéraire est beaucoup plus faible que d'autres formes de communication à l'écrit, comme les discussions sur internet, ou les SMS. La communication virtuelle sert à rassembler des locuteurs de ces deux variétés de l'anglais, autour de leur amour de la langue ou de la culture dont elles font partie à part entière. Ainsi cette recherche portera-t-elle sur les discussions des forums d'internet, pour plusieurs raisons : tout d'abord, ces forums sont des espaces de discussion libres, mais néanmoins codifiés par des règles différentes d'un forum à un autre. Le caractère asynchrone, dépourvu de la spontanéité d'expression des salons de clavardage, aide cette recherche, car il offre un corpus de messages réfléchis par leurs auteurs avant d'être ajoutés à la discussion. Ces messages sont également archivés, et les premières occurrences de graphies sont traçables. Ceci permet d'éviter le problème du « forgeur inconnu » (Lecercle, 2004), ou le premier utilisateur d'une expression ensuite grammaticalisée, et dont on ne peut avoir de trace, à cause de sa nature habituellement orale.

Problématique contextualisée

Le scots et l'AAV sont considérés comme des langues à part entière par certains de leurs locuteurs, et sont documentés dans des ouvrages de référence qui s'intéressent principalement aux dimensions diachroniques et essentiellement orales (Corbett, McClure & Stuart-Smith, 2003 ; Mufwene, 1998). Ils sont transmis plus volontiers par oral que par écrit. Il existe néanmoins un corpus littéraire afro-américain et écossais fourni, sur plusieurs siècles. La présente étude s'attache à analyser le discours produit par des utilisateurs de réseaux sociaux, où toute latitude leur est laissée afin de composer des productions écrites pouvant mêler variété standard de l'anglais, langue minoritaire et éléments discursifs propres à internet.

Avec l'apparition des réseaux sociaux, l'accès à une plateforme pour discuter et se rencontrer en ligne est facilité. Les communications sur internet présentent des caractéristiques inédites, grâce à l'utilisation de l'écrit comme biais quasi-oralisé, dans le sens où les productions linéaires ne sont pas régulées par une instance normative, et où l'accent est plus souvent mis sur la création d'un esprit de groupe. Ces propriétés expliquent en partie l'émergence d'une culture internet à part entière. Les forums de discussion sont des plateformes où les échanges sont contrôlés par des instances physiques, des modérateurs, connus des utilisateurs (comme ceux des plateformes de chat instantané de groupe de type *Internet Chat Relay* (IRC)). Malgré cela, les utilisateurs ont une grande liberté discursive, et échangent sur divers sujets. Les productions écrites ne sont pas fixées, et, contrairement aux discussions instantanées, peuvent être réfléchies et amendées ultérieurement par leurs auteurs. Les formes observées dans ce travail sont typiques des pratiques internet sur les forums (Keegan, 2006).

Le sujet principal de cette thèse est l'émergence d'un lecte propre à une communauté virtuelle, le communaulecte, fortement inspiré de langues minoritaires dans ce cas précis. L'importance de la représentation graphique en tant qu'indice de la création d'une identité de communauté virtuelle est une problématique primordiale pour cette étude. Les messages sur les forums sont ordonnés de façon thématique, ce qui permet la création de petits agglomérats d'utilisateurs autour d'intérêts intracommunautaires. Ces agglomérats ont tendance à faire naître des graphies alternatives stables, utilisées

par plusieurs scripteurs. La variation inter-scripteurs et la création, les causes et les manifestations de regraphies (graphies alternatives stables) sont des éléments centraux à la problématique de cette thèse. Certaines graphies n'ont pour but que de contourner la censure, mais une seule regraphie n'est pas le seul candidat pour certains mots répétés. Certaines regraphies sont préférées, mais la création est libre et une dizaine de candidats regraphiques se dispute pour représenter le mot « fuck » sur BP, par exemple.

Ces regraphies sont également socioculturelles, comme les graphie spécifiques aux gangs. Elles sont principalement utilisées hors-dialogue, dans les présentations personnelles, et ont souvent une forme figée : elles peuvent être des jeux sur des acronymes, des variations sur des poèmes connus, ou des particularités linguistiques, comme le *roll call* ou *shout-out*, ou liste de noms ou surnoms des amis de l'utilisateur.

Les pratiques discursives se situent sur un continuum linguistique : à l'évidente maîtrise de l'anglais standard s'ajoutent la maîtrise d'une langue minoritaire parlée et transcrite à l'écrit (scots et AAV), ainsi qu'un sociolecte internet (Netspeak) et d'un lecte commun, développé sur chaque plateforme virtuelle, un communaulecte commun à tous les utilisateurs d'un forum. Le choix de forums internet dédiés aux intérêts afro-américains et écossais permet de pouvoir évaluer l'importance des langues minoritaires utilisées en synchronie, autant dans les discussions que dans les productions isolées, où les utilisateurs s'expriment sans attendre de réponse de la part des autres membres de la communauté. Pour certains, l'expression d'appartenance à un groupe social, virtuel ou non, sera traduite par l'alternance codique.

Cette thèse s'intéresse à la fois à la forme du signe qui, s'emplissant d'un sens différent, marque des réalités différentes. Le « but idéologique », pour reprendre Bakhtine [1] est alors d'exprimer la multiplicité en tant que locuteur de multiples lectes. Les codes de l'anglais standard sont connus et exploités, comme ceux du scots et de l'AAV. Ils peuvent être discrètement distillés (utilisation fréquente de « you know what I mean ? », dans certains textes afro-américains, ou restriction au scottisme « wee » ou « tae » dans certaines productions de SC). Ces signes peuvent aller jusqu'au lexème non recon-

1. Bakhtine, 1977.

naissable pour un non-locuteur du dialecte [2]. Ces locuteurs sont également utilisateurs d'internet, et sont donc exposés à des phénomènes numériques indigènes, et à la maîtrise de réalités du forum qu'ils fréquentent. Par exemple, l'équivalent le plus proche du statut de Facebook (FB) s'appelle « What's New ? » sur BP, alors qu'il s'appelle « The Pulse » sur SC. L'exposition aux forums est aussi celle des productions des autres participants, qui viennent parfois à être imitées.

Le clivage entre le monde réel et le monde supposé virtuel est tombé, à mesure que les outils technologiques ont évolué, permettant ainsi cette porosité. L'AAV ou le scots pratiqués à l'oral vont également être incorporés en ligne. Le lexème *GIF* est devenu mot de l'année 2012 selon l'Oxford English Dictionary. Il désigne pourtant un format d'image qui peut être animé et montré en boucle sur n'importe quelle page web, et cette popularité montre que même les lexicographes les plus reconnus prennent compte des changements intervenus grâce à l'informatique. Ce travail de thèse s'inscrit dans une perspective postdualiste, qui considère que le clivage corps/esprit n'existe pas, et où l'esprit humain imprime sa marque sur les objets et technologies qui l'entoure. Le discours est un assemblage d'éléments langagiers et non-langagiers (Paveau, 2013b). Le corollaire de ce postdualisme est de considérer que l'humain est au moins « augmenté » : la position du cyborg est tout à fait celle dans laquelle les diverses alternances codiques s'inscrivent : « cyborg politics is the struggle for language and the struggle against perfect communication, against the one code that translates all meaning perfectly » (Haraway, 2009). Il y a une diversité de codes employés sur les forums, et leur interaction est l'un des enjeux de cette thèse.

Justification du choix du titre

Les pratiques discursives des Afro-Américains et des Écossais est donc le point principal de ce travail. Elles relèvent de plusieurs problématiques, identitaires et communautaires principalement. L'appartenance à une communauté (virtuelle ou physique) est un des points principaux sur lesquels repose la création et l'exploitation d'un communaulecte par et pour les membres de chaque communauté.

2. Si *jotter* est compréhensible grâce à sa parenté avec le verbe *jot*, *dricht* est moins transparent à moins d'avoir une culture de l'Écosse.

Les langues parlées sont également un point important sur lequel les graphies alternatives sont créées. Les reflets graphiques de langues parlées sur un média communautaire permettent aux membres de se prouver qu'ils partagent et approuvent des valeurs communes.

La grapho-phonématique sera donc primordiale pour analyser les pratiques graphiques alternatives. Par souci de simplification, le terme général de « dialectologie » est employé pour parler autant des langues minoritaires (AAV et scots) que du netspeak, voire des deux formes d'anglais standard britannique et américain. Toutes ces strates linguistiques se confrontent sur les forums, dans des alternances codiques fréquentes. L'aspect discursif est essentiel dans ces jeux graphiques, et l'utilisation de quelques outils de l'analyse de discours semble incontournable pour prendre en compte toutes les spécificités du communaulecte.

Plan

La première partie de cette thèse s'appuie sur la figuration fictionnelle des marques de l'AAV et du scots dans les graphies. Les graphies alternatives stylistiquement incorporées à ces écrits littéraires sont un premier pas vers une dichotomie entre standard anglais et standard en scots et en AAV. Ces deux dispositifs romanesques donnent accès à des variantes allographiques, auxquels donnent lieu les variations de point de vue symbolisées par des variations selon le « scripteur ». Ces œuvres ont été ensuite transposées au cinéma, et les productions orales des acteurs permettent de comparer les éléments graphophonématiques des graphies alternatives.

Le second chapitre quadrille ce que les deux langues étudiés représentent, à la fois pour leurs utilisateurs et pour les universitaires qui les ont théorisées. Les caractéristiques lexicales, graphiques et surtout phonologiques sont abordées. Le chapitre poursuit avec une étude du statut sociolinguistique de ces deux langues à travers la représentation offerte par les dictionnaires, ainsi que les représentations de ces deux langues sur internet. Je montrerai qu'elles sont stigmatisées, au travers de leurs locuteurs, et les représentations trouvées sur internet sont stéréotypées.

Ces deux premiers chapitres servent à délimiter les langues minoritaires étudiées telles qu'elles

sont utilisées, décrites et perçues, hors ligne et sur internet. Le troisième chapitre permet d'établir les spécificités de la Communication Médiée par Ordinateur (CMO) et son rapport à la norme. Le corpus doit être analysé en interaction, et les caractéristiques discursives des deux forums doivent être délimitées. La méthodologie de la collecte de données sera expliquée dans cette partie, à la place qui semble légitime, c'est-à-dire après l'exposé de ce qui la justifie. La spécificité d'un corpus internet a soulevé des problèmes méthodologiques, que je décrirai ensuite. Les membres du forum, et les chercheurs qui recueillent des corpus en ligne doivent trouver leur place dans cet écosystème très différent des enquêtes de terrain traditionnelles. L'identité personnelle et l'identification à et sur une communauté virtuelle est la dernière partie de ce chapitre, pour les utilisateurs, mais également pour l'observateur, qui voit se déplacer le paradoxe de l'observateur [3].

Le quatrième chapitre conceptualise la notion de communaulecte à travers l'analyse de fils de discussion, en tant qu'interface entre des configurations discursives et les langues minoritaires. Je montrerai d'abord la différence entre fautes d'orthographe et graphies alternatives, ces dernières composent une grande partie des pratiques linguistiques sur internet. Les conditions d'utilisation des graphies alternatives, dans des fils de discussion polémiques, ou au contraire, dans des contextes laudatifs, moqueurs, y sont analysées, grâce à des études de cas.

Le cinquième et dernier chapitre traite de la figuration graphique de la prononciation dans les graphies alternatives. Les variations de graphie des consonnes et des voyelles éclaireront les constatations du second chapitre sur la théorie phonologique. Les graphies alternatives étudiées sont aussi exploitées automatiquement pour voir si elles sont spécifiques aux deux communautés. Le chapitre s'achève sur une étude des variations graphiques de certains mots iconiques sur BP et SC.

Le schéma reprenant l'organisation des différents chapitres les uns par rapport aux autres est présenté dans l'illustration 1. Le premier chapitre est à part, parce qu'il sert de partie préliminaire à l'élaboration du corpus internet : les pratiques qui y sont décrites appartiennent à un corpus fictionnel multimodal (romans et films) qui sert de test pour vérifier la cohérence des relations graphophonématiques utilisées par des locuteurs d'AAV et de scots. Ces œuvres de fiction sont cependant compréhensibles par le grand public, et constituent donc une introduction toute trouvée pour cette thèse. L'analyse

3. Le recueil de données est toujours influencé par la présence du chercheur (Labov, 1972).

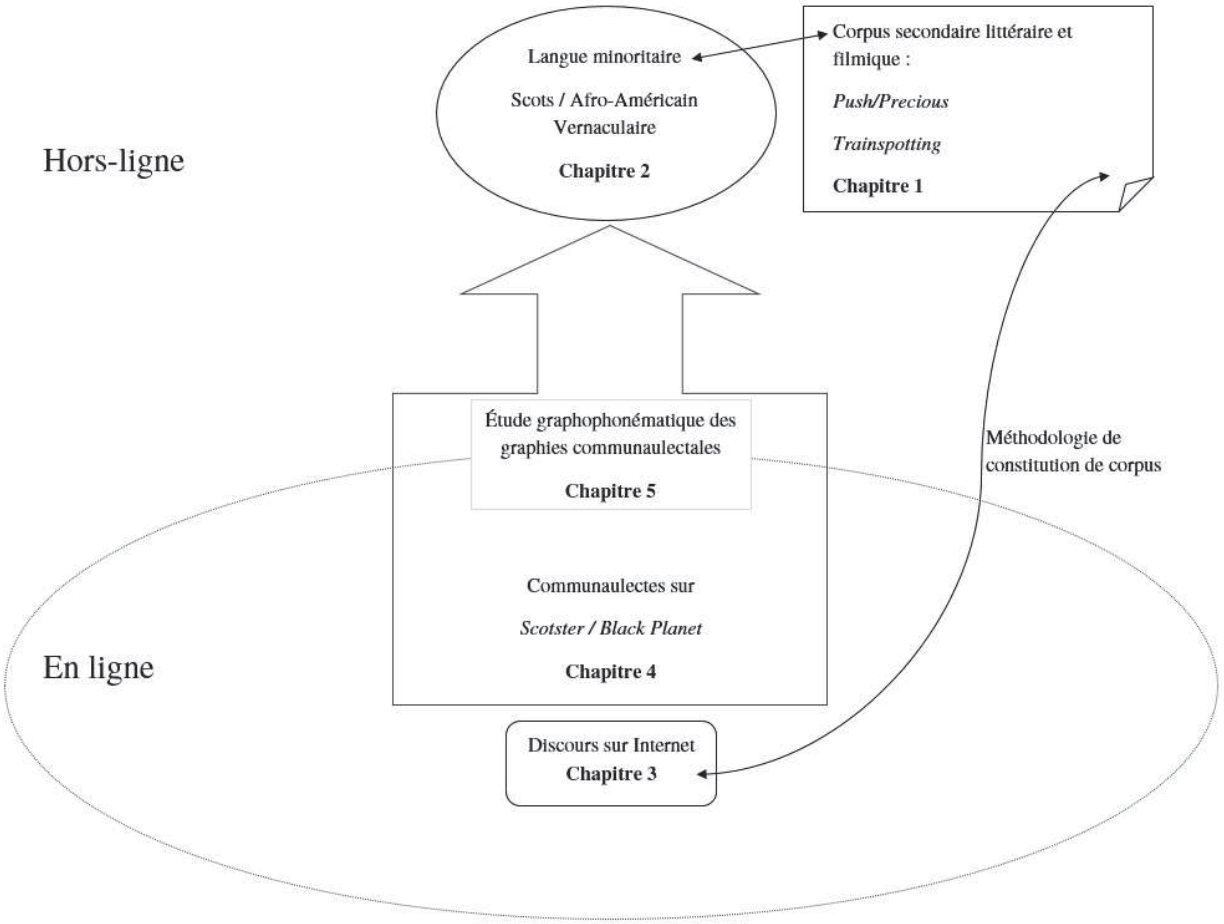


Illustration 1 – Organisation du plan de thèse

phonologique des corpus littéraires est une pratique déjà établie (D. R. Preston, 1982 ; N. F. Blake, 1981). La description formelle et théorique des langues minoritaires qui y sont utilisées, ainsi que leurs relations avec leurs voisins linguistiques (l'ABS pour l'AAV, le gaélique et l'Anglais Écossais Standard (AES) pour le scots) est faite dans le second chapitre, qui introduit les études sur des corpus et des ressources sur internet. Le troisième chapitre décrit l'environnement dans lequel les productions des utilisateurs d'AAV et de scots évoluent en ligne, et cette partie est indispensable pour comprendre le communaulecte et les pratiques discursives propres à internet étudiées dans le quatrième chapitre, qui se concentre sur les productions des membres de BP et de SC. Le développement, qui se veut cyclique, montre les rapports entre communication hors-ligne et en ligne. Le corpus constitué dans le chapitre un sert de base pour la méthodologie utilisée pour recueillir les données, décrite dans le chapitre trois.

Les trois derniers chapitres s'intéressent à BP et SC, et le corpus est primordial dans les deux derniers chapitres. Les formes plutôt parallélépipédiques du schéma évoquent des ensembles semi-fermés, les romans, films et les forums, qui ont servi à constituer le corpus. Les formes plutôt arrondies évoquent les ensembles linguistiques eux-mêmes (les langues minoritaires, le discours sur internet). L'ellipse « en ligne » est représentée en pointillés pour montrer sa porosité avec le discours hors-ligne.

Chapitre 1

Analyse des graphies fictionnelles

Thas the alphabet. Twenty-six letters in all.
Them letters make up words.
Them words everything.

Sapphire - *Push*

Le premier chapitre a pour but de déterminer les limites de la graphophonématique, appliquée à l'afro-américain et au scots, en se servant d'un corpus secondaire qui servira d'étalon afin de baliser les corpus naturels plus facilement, et de théoriser une typologie des graphies alternatives à partir d'un petit corpus fermé, rédigé par deux auteurs seulement, dans des œuvres de fiction. Les deux corpus secondaires possèdent des caractéristiques communes : ce sont des romans écrits par des locuteurs natifs des dialectes, contenant une proportion de marqueurs dialectaux importante, adaptés ensuite en film.

1.1 Caractérisation du corpus secondaire

J'étudie donc *Push* de Sapphire, et son adaptation *Precious*, puis *Trainspotting* d'Irvine Welsh, au format roman et film. L'avantage de ce corpus préliminaire est l'avantage offert par les dialogues des films qui sont à leur tour écrits dans les sous-titres. La comparaison entre les mêmes éléments peut être faite plusieurs fois (comparaison écrit-film, comparaison film-sous-titres et écrit-sous-titre) et apportent ainsi une première balise sur l'utilisation de l'afro-américain vernaculaire et du scots.

Ces deux œuvres ont été écrites par le même type de scripteur : deux auteurs, parlant couramment les deux dialectes, qu'ils font figurer à l'écrit en utilisant des graphies alternatives ou démarcatives, ainsi que des formes syntaxiques propres à l'AAV et au scots. Les graphies alternatives seront examinées plus en détail dans la section 4.4. C'est un concept pilier pour ce travail, et les graphies alternatives sont toutes les réélaborations graphiques qui s'éloignent du standard orthographique, et qui sont désirées par leurs scripteurs. Les graphies démarcatives sont une sous-catégorie des graphies alternatives, dont le but est de s'éloigner du standard graphique anglais en s'en démarquant. Cependant, ces deux œuvres et leurs adaptations ne sauraient pas remplacer l'étude des deux langues minoritaires, le scots et l'AAV. Ni I. Welsh, ni Sapphire ne sont linguistes, et leurs écrits ont plus à voir avec une stylisation graphique basée sur des parlers retranscrits qu'avec une quelconque authenticité. Leurs écrits s'insèrent dans la tradition littéraire qui a recours à une graphie dialectale. De nombreux textes littéraires ont recours à ces graphies pour retranscrire divers parlers ou accents (texan, canadien, yiddish…) (D. R. Preston, 1982). Dans les deux œuvres, les représentations sont cohérentes, et les formes alternatives sont stables. Elles ne sont pas éloignées de celles des utilisateurs des forums, et leur volonté de s'approcher le plus possible d'une « parlécriture » dans le sens de Freire (Trovato, 2010), qui utilise une technique pédagogique pour l'alphabétisation, sans stigmate normatif. Je suis évidemment convaincue qu'il s'agit dans ce corpus d'une parlécriture fictionnelle, créée par des auteurs maîtrisant parfaitement le code standard—et la subversion de l'écrit pour offrir au lecteur une expérience quasi-authentique.

Ce corpus secondaire est constitué de retranscriptions de parlers vernaculaires, sinon de versions

oralisées et fortement dialectalisées de l'anglais. En outre, les deux romans se passent dans des périodes relativement proches (1987 pour *Push*, et la fin des années 80, sans plus de précision, pour *Trainspotting*) Pour ce faire, les deux auteurs ont à cœur d'utiliser des graphies alternatives, sur lesquelles s'assoit la méthodologie et le protocole de typologisation du corpus principal, qui sera traité aux chapitres suivants.

Ce corpus sert à sérier les thèmes et les étiquettes nécessaires pour la graphométrie du corpus principal, et permet d'apporter un corpus étalon, adapté aux outils de linguistique quantitative dont je me servirai au chapitre 5. L'analyse n'est pas littéraire, dans le sens où j'évoque l'auteur et les narrateurs homodiégétiques de façon interchangeable, tout en le précisant dans le développement. C'est une caractéristique de *Trainspotting* puisque les narrateurs successifs sont autant de codages graphiques, syntaxiques ou lexicaux différents, tous liés à I. Welsh. On retrouve cette question dans le corpus *Push*, où les écrits de la narratrice, de ses camarades, ainsi que de son professeur sont incorporés « tels quels » dans le corps du texte.

1.2 *Push* et *Precious*

Les deux œuvres, présentées plus en détail dans leur partie respective, sont connues du grand public. Le livre *Push* s'était déjà vendu à 280 000 exemplaires lorsque Sapphire a été approchée par Lee Daniels pour l'adapter au cinéma, avec l'aide d'Oprah Winfrey, qui devint l'une des productrices du film. Malgré le choix d'une actrice inconnue pour tenir le rôle principal et un budget relativement modeste de 10 millions de dollars, le film a généré un succès certain (près de 45 millions de dollars en 2010, selon l'IMDb [1]).

1. L'*Internet Movie Database* (Needham, 1990) est l'un des 50 sites les plus visités selon Alexa (« Alexa - The Web Information Company », 2000) et l'une des sources les plus fiables concernant les films et les séries télévisées.

1.2.1 *Push*

Push est un roman fait d'embriquements discursifs et situationnels : écrit par Sapphire, une locutrice de l'AAV, il est également autobiographique, car il s'appuie sur son expérience en tant que pédagogue dans une école pour les jeunes défavorisés. Le texte est tout de même une fiction, puisque les détails sont romancés, et le personnage de Precious est une invention basée sur plusieurs jeunes filles qui ont suivi les cours de Sapphire. Je laisse la question de la fiction dans la narration aux littéraires, et je limite les explorations à moins de se tourner vers l'analyse de discours : le roman est pris comme un ensemble discursif représentatif des graphies potentiellement réalisables par les locuteurs de l'AAV qui voudraient retranscrire leur façon de parler.

Je considère donc ce livre comme un ensemble de productions graphiques, et leur contenu comme un ensemble de parlers et de graphies de plusieurs locuteurs, imités ou émulés [2]. Le changement de registre, et de parole, est donc attendu. Il n'y est presque pas question de polyphonie de scripteurs différents, puisque seuls les écrits de Precious et de Blu Rain, sa professeure, seront étudiés. J'écarte donc les transcriptions de l'anglais jamaïcain du personnage de Rhonda, par exemple. Ce phénomène est également caractéristique du corpus principal, où toutes les productions forment un ensemble homogène (à l'échelle de la communauté, et des communaulectes (étudiés au chapitre 4), qui est en fait un faisceau de productions différentes. L'imitation et l'émulation d'un changement de registre ou de « parole » a une proportion importante [3].

Dans la partie suivante, je décrirai les structures narratives du livre avant de passer à une courte analyse onomastique. Ces considérations sont indispensables pour analyser la totalité des alternances graphiques en fonction de l'évolution de la scriptrice fictive Precious.

1.2.1.1 Structures narratives de *Push*

Push est un roman publié par Sapphire (nom de plume de Romona Lofton) en 1996. Il retrace l'histoire de Claireence Precious Jones dite Precious, une jeune femme afro-américaine illettrée qui

2. On pourrait fantasmer un « recopiés de sources primaires » , sans jamais en avoir la preuve.
3. et ce phénomène est encore plus visible dans *Trainspotting*.

apprend progressivement à écrire et à lire dans une école alternative. Cette narratrice homodiégétique tient un carnet de bord, dont les graphies évoluent à mesure que sa maîtrise de l'écrit s'affirme.

Le livre est structuré en cinq mouvements : la première partie introduit l'histoire jusqu'au moment où Precious s'inscrit à l'école alternative (jusqu'à la p. 38 [4]), la seconde couvre le début de son apprentissage (p.76), la troisième s'ouvre sur l'accouchement du second enfant de Precious. La quatrième partie (p. 109) s'amorce alors que Precious décide de ne pas abandonner ses études après avoir appris qu'elle avait le SIDA. La dernière partie n'est pas à proprement parler intégrée dans la narration, étant donné qu'elle est composée par les différentes autobiographies rédigées par les étudiantes de l'école alternative (p. 154). En tant que telle, cette partie ne sera pas exploitée avec autant d'attention que les autres dans cette étude. Les graphies présentées répondent plus au standard que celles du livre, et elles ne sont pas en soi des productions de la protagoniste (à l'exception du premier et du dernier texte). De plus, elles sont présentées à part dans le livre, dans une police de caractère différente, et les pages ne sont pas numérotées.

Les différentes parties sont relativement équilibrées et découpent effectivement le récit en quatre parties, mais certaines contiennent les échanges épistolaires entre Precious et Blu Rain, sa professeur de l'école alternative (parties II, III et IV). Ces textes sont insérés dans le cœur du texte, mais leur graphie contraste souvent avec celle du récit en tant que tel. Blu Rain [5] écrit dans un anglais tout à fait standard, et ses lettres ne seront donc pas étudiées. La dernière partie est la seule qui contienne des textes d'autres scripteurs : les quatre autres étudiantes de l'école alternative écrivent leur histoire, et des textes de Precious ouvrent et ferment cette partie. Les parties épistolaires ne contiennent quasiment que des graphies alternatives, c'est pourquoi leur proportion est importante dans cette étude. Le cœur de la narration est une extension de ces carnets de bord, et subit des changements graphiques et grammaticaux plus subtils et lents que ceux des carnets de bord, mais suivent peu ou prou le même mode d'évolution. Pour l'étude des graphies présentées dans *Push*, l'ensemble des productions de Precious (la narration pure, et ses cahiers de bord) seront prises en compte.

Le livre est divisé en cinq parties, et les genres littéraires s'y mêlent, ce qui rend la circonscription

4. Les indications de pages correspondent à l'édition de travail (Sapphire, 1996)
5. Le nom de ce personnage est graphié soit « Blue », soit « Blu ». La dernière graphie a été choisie, parce qu'elle est confirmée par le film.

Partie du roman	Nombre de mots	Parties épistolaires rédigées par Precious
I	7989	0
II	8385	109 (1.2%)
III	7191	739 (10.3%)
IV	9673	1718 (17.8%)
"LIFE STORIES Our Class Book" (V)	5155	343 (6.6%)
Total	38398	2909 (7.5%)

Tab. 1.1 – Contenu de Push

du corpus plus difficile. Les parties épistolaires et diaristes se mêlent à la narration homodiégétique, et les poèmes de la dernière partie du livre sont autant de fictions de l'AAV et de lectes différents (comme l'anglais jamaïcain parlé, puis écrit par Rhonda, la camarade de Precious). Les dialogues sont autant de façons de confronter ces accents et mots différents.

L'analyse linguistique de ce sous-corpus secondaire pose donc le problème du mélange des genres. Les parties et sous-parties du roman, ainsi que leurs implications discursives (utilisation de formules phatiques pour ouvrir et fermer les entrées du journal, mise en forme…) se mêlent, de la même façon que les différents registres, comme les lettres ouvertes avec phatiques, diatribes, pages personnelles…se complètent sur SC et BP.

1.2.1.2 Analyse onomastique

Tout comme les pseudonymes, les prénoms en disent long sur leur porteur, au moins de façon stéréotypique, surtout dans une oeuvre de fiction, où les noms sont sciemment choisis. C'est la raison pour laquelle je fais cette courte analyse du nom de certains personnages de *Push*. Dans cette sous-partie, j'étudie le cas de trois personnages, dont les noms sont à double entrée.

La fille de Precious s'appelle Mongo, une référence double. Selon Precious, il s'agit de l'apocope du mot « Mongoloid » , étant donné qu'elle est atteinte de trisomie 21, et d'un mot qu'elle perçoit comme ayant des consonances espagnoles « Mongo sound[s] Spanish do[es]n't it ? Yeah, tha[t']s why

I chose it, but what it is is short for Mongoloid Down Sinder [Syndrom], which is what she is »[6]. Ce mot est utilisé pour décrire les gens atteints de trisomie de façon très péjorative (ce lexème a 29 entrées sur l'*Urban Dictionary*, et le sème le plus fréquent est '*handicapé mental*'ou '*stupide*'est l'un des deux sens les plus fréquemment décrits[7]).

Mongo peut également être perçu comme un jeu de mots de Sapphire, sur le personnage de Mungo, l'esclave représentant le stéréotype négatif de l'esclave africain dans la pièce de Bickerstaff, *The Padlock* (présentée à Londres en 1768[8]). Ce personnage est caractérisé par ses défauts (ivrogne, voleur et hypocrite), et s'exprime dans une représentation du dialecte des esclaves issus des Caraïbes. Le nom est devenu, par extension, une façon péjorative de parler de tous les gens d'origine africaine.

Le prénom de Precious, Claireence, est très afro-américain dans sa forme. La graphie mélangeant *Claire* et *Clarence* est caractéristique des prénoms afro-américains, souvent caractérisés par des concaténations de prénoms, ou l'utilisation de préfixes comme La- ou Le-, Da- ou De-. Les graphies personnelles ou les mots-valises à partir de plusieurs prénoms (« Condoleezza » de l'italien 'con dolcezza', '*avec douceur*') sont légion, ainsi que les prénoms d'origine grecque ou arabe (Malik, Darius, Jazmin[9]). La graphie <Claireence>, sur Google, ne renvoie que 269 résultats[10], et seulement 9 si on force une recherche sans <Precious>[11]. Cette fréquence extrêmement faible est cohérente avec le désir d'arriver à un hapax, une des caractéristiques des prénoms afro-américains. Il est à noter que « Precious » lui-même fait partie de la liste des prénoms féminins perçus comme étant les plus « afro-américains » . Cette constatation peut aussi être appliquée au prénom de la professeure, Blu (Rain), sur lequel une camarade de Precious s'étonne :

> « "OK, I'll start. My name is Blue Rain—"
> "Thas your real name !" This from girl with boy suit on.
> "Urn hmmm, that's my for real hope to die if I'm lying name."
> "Your first name Blue ?" same girl say.
> "Urn hmm," Ms Rain say this like she tired of mannish girl.

6. partie II, p. 34
7. L'infirmier qui est venu la secourir lorsqu'elle accouchait étant d'origine hispanique, Precious associe des caractéristiques positives à cette origine ethnique.
8. Bickerstaffe, 1775.
9. La liste complète des noms les plus « caucasiens » et les plus « afro-américains » est tirée de *Freakonomics* (Levitt & Dubner, 2005)
10. Recherche stricte « Claireence » en novembre 2012
11. Recherche stricte « "Claireence" - Precious »

> “Splain that !”
> “Well,” say Ms Rain real proper. “I don’t feel I have to explain my name.” »

Les noms sémantiquement différents, comme les prénoms afro-américains qui ajoutent le préfixe <La-> aux prénoms féminins (et masculins), ou graphiquement différents, comme <Blu>, qui s’éloigne du standard orthographique, interrogent les autres sur leur perception. Ce qui importe n’est pas la cause sous-jacente, c’est le fait qu’ils existent sous cette forme, et les réactions extérieures sont uniquement des indices de l’écart, face à une norme intégrée. Aucun nom n’étant laissé au hasard, les explications de noms et de pseudonymes dans le corpus principal sont importantes parce que les utilisateurs de forums inventent eux aussi leur dénomination virtuelle. On retrouvera ces questions onomastiques au chapitre 3, et plus particulièrement en 3.4.3. Expliquer son nom y est parfois un problème (comme l’exemple de Blu Rain le montre), et ces questions de pseudonymat vont puiser dans la sphère intime de l’élaboration de son identité.

Le choix des pseudonymes et des noms de personnages est une étude en circuit très fermé des graphies alternatives choisies par Sapphire. Dans la partie suivante, je vais m’attacher à dresser une typologie des graphies alternatives dans les écrits de Precious.

1.2.2 Typologie des graphies

Le corpus secondaire est un galop d’essai pour l’explicitation des types de graphies différentes, rencontrées dans le corpus principal. Certaines stratégies de création ou d’adaptation lexicales sont prototypiques et seront considérées comme des schémas que l’on retrouve sur les forums comme dans le corpus secondaire. Ces types de graphies et d’élaboration à partir de *l’image sonore des signifiants* (Saussure, 1917) pratiquées par Precious sont transversales.

Les graphies alternatives adoptées par Sapphire reflètent différentes réalités. Certaines sont le miroir de celles utilisées par la population générale (« gonna »), et touchent surtout les quasimodaux : « There is also a plausible candidate for a product-oriented schema in English based on the phonological structure of a family of construction schemas. These are the so-called quasimodals, or recently grammaticalized constructions […] All of the quasimodal forms end in an alveolar consonant follo-

wed by schwa, although this ending represents the reduction of different source forms [12] » (Croft & Cruse, 2004). Sur ce modèle, les Afro-Américains ont adopté certaines graphies constatées dans le corpus, comme « usta ».

> I usta blame all my mishaps on everyone else, but no one is to blame for my unhappiness but me. (BPPP11)
> It never usta be on my mind. (p. 109)

Il y a 11 occurrences de « usta » dans *Push*, contre 1 100 dans BP, et cette concaténation à partir d'un verbe + de $[\text{to}]_{prep}$ est assez proche en termes de construction que le lexème « tryna », dont l'on trouve environ 122 000 occurrences dans le corpus BP et une occurrence dans SC. Certaines graphies émulent la prononciation imaginée de la protagoniste, ou ses erreurs de graphie (comme les parties épistolaires, qui marquent les progrès de Precious) et ces erreurs de graphies, ces gabarits, constituent tous des éléments que l'on peut appliquer au corpus principal.

1.2.2.1 Graphie gabaritique

Les graphies gabaritiques sont présentées comme des graphies à part des autres graphies alternatives. Elles constituent le premier stade d'apprentissage de la graphie dans le roman, et seuls les sons les plus saillants sont représentés. Ce sont notamment les sons les plus proches de la prononciation des lettres de l'alphabet et les moins durs à reconnaître phoniquement, le plus souvent des consonnes. Le récepteur-lecteur doit donc faire des efforts pour les comprendre, et elles sont généralement accompagnées d'annotations dans le livre. L'une des premières entrées du journal de Precious contient un alphabet personnalisé (sur le modèle « A is for… » et ainsi de suite), où l'on peut donc supposer une connaissance des lettres, de leur ordre (une aptitude fraîchement acquise), et de leur valeur phonique. Pour les étudier, je vais les regrouper dans différentes catégories qui correspondent à leur type de gabarit utilisé dans le tableau 1.2. Les retranscriptions normatives sont ajoutées à la suite des entrées, comme le montre l'illustration 1.1.

Les représentations consonantiques sont bien plus fréquentes que les autres types de gabarits. Elles assoient la position dominante de la consonne, au niveau du mot, dans un système qui rappelle les

12. Les formes dont Croft parle sont *to*, *have*, *better* dans *useta* [sic], *coulda* et *betta*.

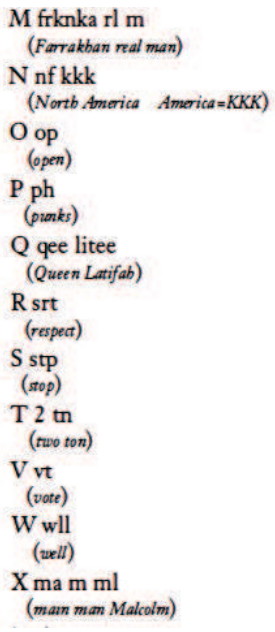

M frknka rl m
(*Farrakhan real man*)
N nf kkk
(*North America America=KKK*)
O op
(*open*)
P ph
(*punks*)
Q qee litee
(*Queen Latifah*)
R srt
(*respect*)
S stp
(*stop*)
T 2 tn
(*two ton*)
V vt
(*vote*)
W wll
(*well*)
X ma m ml
(*main man Malcolm*)

Illustration 1.1 – Entrée du journal de Precious avec corrections

Type de gabarit	Exemples
Position initiale	m (‘*man*’), ma (‘*main*’), Jerm (‘*Jermaine*’), litee (‘*Latifah*’), op (‘*open*’), ph (‘*punks*’), qee (‘*queen*’), somb (‘*somebody*’)
Consonantique C[V]C	Afrc (‘*Africa*’), bk (‘*black*’), bt (‘*but*’), cl (‘*colored*’), el (‘*evil*’), Frknka (‘*Farrakhan*’), hm (‘*home*’), kl (‘*kill*’), lv (‘*love*’), ml (‘*Malcolm*’), nf (‘*north*’), rl (‘*real*’), srt (‘*respect*’), tn (‘*ton*’), vt (‘*vote*’), wll (‘*well*’), zk (‘*zonked*’)
Vocalique CV[C]V	ba e (‘*baby*’)

Tab. 1.2 – Groupement des graphies gabaritiques de la seconde entrée du journal de Precious

langues chamito-sémitiques. Les groupes consonantiques sont plus prompts à se réduire. Le seul restant dans ces exemples est /fr/ dans <Afrc>. Dans la plupart (« bt, hm, kl, lv, nf, rl, tn, vt, wll, zk ») des graphies de cet ordre, le cœur (vocalique) du mot en CVC a disparu, et ces graphies évoquent les écritures natives en ligne : si <luv> est une graphie alternative « historique » de « love », où le <e> final a disparu, on peut imaginer une graphie <lv> qui retrouverait ce gabarit C[V]C.

L’écriture de *main* <ma> semble peu logique de prime abord. Elle s’explique par le possible amuïssement de la consonne finale, et donc d’une prononciation suggérée [mæ̃],qui correspond à l’omission des nasales en coda syllabique [13], qui constitue l’un des traits de la phonologie de l’AAV (voir le tableau

13. En revanche, le transfert de la nasalité sur la voyelle précédente n’est évidemment pas vérifiable.

2.12). Les gabarits initiaux ne sont pas tant des gabarits que des refus (ou plutôt, des impossibilités) d'écrire les mots correctement : l'apocope est involontaire. Pour *punks*, la prononciation de Precious pourrait être le standard /ˈpʰʌŋks/,où seul le [pʰ] et son aspiration seraient faciles à retracer pour qui ne sait pas écrire, contrairement aux autres lettres. Il pourrait s'agir d'une lénition du son /p/ en /f/, mais il n'y a pas d'autres exemples d'un tel phénomène dans le corpus *Push*.

Quant à *queen* retranscrit « qee », le gabarit initial est, en fait, le seul disponible pour la locutrice, semble-t-il : la disparition du /n/ final et une prononciation [kʷi], voire [kʷɪ] (qui correspondraient au système phonologique de l'AAV) n'aurait pas à s'embarrasser d'un /w/ peut-être trop dur à retranscrire, ou d'un /n/ en coda certainement non prononcé. Le problème réside ici dans les conjectures, trop nombreuses et impossibles à confirmer ou à infirmer, mais cette piste semble être la plus satisfaisante pour expliquer la disparition du <n> final, amuï dans une prononciation relâchée de « queen ».

Ces graphies semblent suivre ce que j'appellerais un « principe consonantique » qui rappelle le système des langues chamito-sémitiques, où seules les consonnes et les voyelles longues sont marquées : les consonnes sont les éléments les plus saillants, et les graphies disponibles pour chaque phonème sont moins ambiguës que pour les voyelles (se référer au tableau 4.1 et à l'étude des équivalences graphophonémiques pour les consonnes en 4.2.1). Les consonnes sont les éléments les plus garants de la graphie, ce que l'on trouve également en ligne, dans les tachygraphies, comme « thx », '*thanks*' où les sons consonantiques sont conservés : <thx> représente /θks/où les sons /æŋ/, sont amuïs (mais retraçables) grâce à la couronne consonantique reliquaire qui ne laisse pas d'autre candidat que « thanks » (Goudet, 2011).

1.2.2.2 Graphies alternatives

Il est difficile d'observer les graphies alternatives pures et de les différencier des fautes d'orthographe de Precious, surtout dans les dernières parties du livre, où la distinction entre le contenu de son carnet d'écriture et la narration sont très ténues (voire inexistantes). Le problème principal est le dispositif scriptural, entre les graphies de l'auteure qui retranscrit graphiquement des usages comme ceux de la protagoniste, qui essaye d'écrire le plus correctement possible. La subversion s'appuie sur

la connaissance d'un code standard, que Sapphire maîtrise. Elle crée l'œuvre à partir de ses connaissances en apprentissage de l'anglais écrit, et on peut considérer l'ensemble des graphies erronées de Precious comme étant des graphies alternatives créées par Sapphire [14].

Du point de vue de Precious, il n'est guère question de s'éloigner d'un code sur lequel elle s'appuie (comme l'alphabet, récité, utilisé comme mantra et écrit plusieurs fois dans le livre le montrent). Apprendre à lire et à écrire est primordial, et la graphie standard est le but à atteindre. Les graphies ne sont alternatives que dans le sens où elles sont des transcriptions de la prononciation de Precious : non alternatives pour elle, mais fautives du point de vue du standard graphique, pas de la prononciation, surtout pour les productions de la dernière partie du livre. En fait de graphies alternatives, je vais traiter les erreurs graphiques dans la sous-section suivante. La distinction entre une graphie alternative désirée et une erreur est primordiale, et le corpus *Push* explore le rapport à la normalité de l'écrit. Les corrections apportées sur quelques éléments du journal épistolarisé de Precious sont des éléments préliminaires à la compréhension des *grammar nazis*, et autres acteurs de la correction.

1.2.2.3 Modalité de la correction

Les erreurs graphiques sont les plus fréquemment faites dans le corpus *Push* [15]. Lorsqu'elles sont insérées dans le texte en tant qu'entrées de son journal, elles sont toujours suivies par une rectification d'enseignante, effectuée par Blu Rain.

Le corps du texte est parsemé de corrections et de visions normalisées de l'anglais américain standard, données par Blu Rain, la professeure d'anglais de Precious. Les dimensions normatives sont attendues, puisque le but en est de faire apprendre à écrire en anglais standard. La double entrée (entre l'erreur graphique et la correction) est tirée vers l'orthographe effectuée par les correcteurs automatiques sur ordinateur. Sur les logiciels de traitement de texte, ou dans les navigateurs internet, voire sur les courriels, les lexèmes fautifs sont soulignés. Ce soulignement en rouge est un processus d'emphase visuelle, qui est identique au soulignement de Blu Rain (ou aux tirets bas tracés par Precious

14. Cette remarque est également valable pour *Trainspotting* et le travail de Welsh sur la graphie du scots et de l'idiolecte parlé par chacun des personnages.
15. Si on prend en compte les écrits de Precious comme émanant d'elle-même et non de Sapphire.

lorsqu'elle sait que la graphie d'un mot va lui échapper). Les conventionnelles vaguelettes rouges attirent l'attention du lecteur tout en offrant une correction (s'il prend la peine de cliquer sur le mot). Dans le cas de *Push*, les corrections aident tous les lecteurs : le personnage du livre et sa compétence orthographique, et le lecteur, pour qu'il comprenne ce qu'il lit. Ceci me permet d'introduire le principe d'intelligibilité : pour qu'un texte soit lu, et donc pour ne pas rebuter le destinataire, il doit au moins être assez lisible pour lui. Ceci constitue le frein principal à la créativité des inventeurs de graphies alternatives, malgré l'adaptativité de l'humain par rapport à la lecture.

La standardisation graphique de Blu Rain, loin d'être castratrice pour la créativité, sert Precious. Contrairement à celle effectuée par les « grammar nazis » d'internet (dont je parlerai en 4.3.5), elle représente les étapes d'apprentissage de la langue écrite. Pour que les utilisateurs de forums subvertissent les codes, il faut qu'ils en aient une certaine maîtrise. Apprendre à écrire passe par la mémorisation de la forme de certains mots (« étape logographique ») ; puis par la prise de conscience d'équivalences grapho-phonématiques, où l'apprenant se rend compte de la complexité syllabique (« étape phonologique »). En dernier vient l'étape orthographique, où les morphèmes et les éléments lexicaux sont pris en compte pour composer les mots, conséquence de l'accroissement de la compétence orthographique dans la langue (« étape orthographique ») (Dehaene, 2007a).

1.2.2.4 Graphies transcriptionnelles

Les écrits de Precious, dans leur mauvaise retranscription de la graphie, posent la question de l'asymétrie graphophonémique. Il n'y a pas d'équivalence correcte entre les sons et les séquences graphiques. L'anglais orthographique est le fruit de la stabilisation progressive de la langue écrite, et Quirk remarque la stagnation de l'écrit, par rapport à l'oral : « our present-day system of spelling represents roughly the speech of four or five hundred years ago » (Quirk, 1963). Il souligne l'absence d'équivalent pour la voyelle ə, et surtout, le changement surprenant de prononciation pour se rapprocher de la forme écrite, et non l'inverse (*forehead* prononcé /ˈfɔːhed/au lieu de son ancienne prononciation /ˈfɒrɪd/, par exemple). On pourrait néanmoins opposer les graphies qui tendent à se rapprocher de la prononciation, mais elles sont plus mouvantes pour les prénoms, où l'écriture est libre, que pour les

noms communs [16].

Je n'évoquerai pas la puissance normalisatrice du dictionnaire concernant l'orthographe, pour me concentrer sur le caractère systémique de la graphie de l'anglais, comme celle de Precious, qui adapte ses graphies à sa prononciation. Ces erreurs graphiques sont donc leur reflet chez elle, surtout dans les deux dernières parties du livre, où elle maîtrise assez la graphie pour pouvoir écrire les mots en entier (et ne pas utiliser de graphies gabaritiques ou initiales). Les graphies tardives ne sont pas corrigées, puisqu'elles ne sont pas inclues dans les carnets montrés au professeur, et elles sont la preuve de l'évolution de la maîtrise de la langue par Precious. Elles permettent de faire des premières analyses sur les équivalences mentales entre prononciation et graphie. Il est possible de retirer des observations sur la répartition des consonnes et des voyelles (et donc des sons consonantiques et vocaliques), objet des deux sous-parties suivantes.

1.2.2.4.1 Consonnes

Les hypothèses de départ sont simples : les consonnes non prononcées ne sont tout simplement pas écrites, comme le <k> de la séquence <kn->, ou le <w> de <wr->, les consonnes doubles sont simplifiées, et les digraphes consonantiques courants sont réduits (ainsi, <ph> deviendrait <f>, et <k> remplacerait <ck> dans le système grapho-phonématique de l'anglais. Ce que l'on conserve, c'est ce que l'on dit, et ceci vaut également pour le corpus internet, où certains retranscrivent leur prononciation (pour certains mots à la graphie inconnue) avec plus ou moins de succès. En regardant le corpus, et la liste de mots à la graphie alternative, la plupart de ces hypothèses se confirme–avec plus ou moins de régularité, comme le montre le tableau 1.3.

Dans le tableau 1.3, les exemples et contre-exemples viennent tous de la liste des graphies alternatives, afin de montrer que même certaines graphies se voulant erronées peuvent être justes sur d'autres points. Cet inventaire est, en outre, très parcellaire [17] et ne prend pas en compte la dimension chronologique du développement graphique, pour se concentrer plus particulièrement sur les formes

16. Ainsi, Stephen est remplacé par Steven, et Catherine par Katherine : en 1999, Steven était le 47ème prénom le plus donné, contre Stephen en 86ème position (pour la même année, Katherine était le 36ème prénom féminin le plus populaire, contre Catherine en 98ème position (of America, 2014).

17. Pour une recension exacte de toutes les formes graphiques et grammaticales alternatives, se référer aux annexes très complètes de (Hémart, 2001).

Phénomène	Sous-règle	Valeur phonétique	Exemples	Contre-exemples
<th> (θ, ð)	<th> → <f(f)>, <fe>	/θ/ → /f/	bafroom (‘*bathroom*’), bofe (‘*both*’), maff (‘*math*’), def (‘*death*’), nuffin (‘*nothing*’)	nothin’, filthee (‘*filthy*’)
	<th> → <t>	/θ/ → /t/	tink (‘*think*’)	somethin’, thankfill (‘*thankful*’)
	<th> → <d>	/ð/ → /d/	dis (‘*this*’), dese (‘*these*’), das (‘*that’s*’)	thas (‘*that’s*’)
	<th> → <v>	/ð/ → /v/	fahver (‘*father*’), muver (‘*mother*’), breave (‘*breathe*’)	mutherfuckin, breathin’
Simplification des groupes consonantiques	<st> → <s(s)>	/st/ → /s/	firs (‘*first*’), tess (‘*test*’), bress (‘*breast*’), breakfus (‘*breakfast*’)	mostlee (‘*mostly*’), musta (‘*must have*’)
	<ng> → <n(’)>	/ŋ/ → /n/	nothin’, walkin	amung (‘*among*’), borning (‘*born*’)
	<nd> → <n(n)>	/nd/ → /n/	fren (‘*friend*’), n (‘*and*’), unnerstand (‘*understand*’)	nd (‘*and*’), underneaf (‘*underneath*’)
	<gh> → <ff>	/f/ → /f/	enuff (‘*enough*’), laff (‘*laugh*’)	–
	<sk> → <x>	/sk/ → /ks/	ax (‘*ask*’)	–
	<wh> → <w>	/w/ → /w/	wat, wen	wher
	<wr> → <r>	/r/ → /r/	rite (‘*write*’)	–
	<ld> → <l(e)>	/ld/ → /l/	ole (‘*old*’), hole (‘*hold*’)	oldern (‘*older*’), holdin’
→ séquence hétérosyllabique	<mb> → <mm>	/mb/ → /m/	crummel (‘*crumble*’)	–
Monographie des digraphes	<ff> → <f>	/f/ → /f/	difrent (‘*different*’)	diffrent
	<dd> → <d>	/d/ → /d/	dady (‘*daddy*’)	–
	<cc> → <c>	/k(s)/ → /k(s)/	cept (‘*accept*’)	–
	<ch> → <f>	/ʧ/ → /f/	birf (‘*birch*’)	culchure (‘*culture*’)

Tab. 1.3 – Réélaborations graphiques des monophtongues

que prennent les mots, surtout vers la fin de l’ouvrage.

Les différents cas de changements concernant /θ, ð/, soit le TH-fronting des consonnes en /f, v/ respectivement, ou l’alvéolarisation en /t, d/ sont attendus. Le digraphe <th> est souvent retranscrit par une seule consonne, dans le cas de <t, d, v> mais pour /θ/, le choix graphique entre <f>, <fe> et <ff> est complexe, et semble tenir de plusieurs facteurs sur la même abscisse : de la simple trans-

cription, <f>, jusqu'à la conservation du nombre de lettres dans le mot, pour la lisibilité du lecteur, surtout dans ces monosyllabiques. On peut alors imaginer une tension entre la transcription graphique d'un personnage d'une part, et celle d'une auteure qui applique les lois grapho-phonématiques qu'elle maîtrise de l'autre. C'est la raison pour laquelle un mot courant, « bathroom » devient <bafroom> : en position coda, avant un autre mot et dans un contexte où les phrases sont elles-mêmes réduites à leur plus simple expression [18] (« Afterward I go bafroom » [sic], p.111), la graphie la plus simple est adoptée. C'est également le choix stylistique adopté pour transcrire /θ/ dans <def> : la phrase est tirée d'un contexte violent (« if she ever come [sic] at me like that again I will stab her to def », p.75), et la compréhension ne pâtit pas de la graphie alternative, même d'un mot si court. En effet, « death » fait partie de l'expression « stab [pronom] to death » et n'a pas besoin d'être écrit avec des précautions particulières pour être compris.

En revanche, pour des mots plus courts et moins courants, les graphies se rapprochent du standard, au moins grâce au nombre de lettres. C'est la raison pour laquelle <fe> et <ff> sont des substitutions toutes trouvées pour le digraphe <th>. La conservation du nombre de lettres est importante pour la lecture (Larson, 2004) et la maîtrise de l'écriture guide au moins inconsciemment la conservation d'un digraphe dans ce cas. Le digraphe pour un phonème est un cas différent de la simplification des digraphes (polyphonémiques) du deuxième cas du tableau 1.3.

Lorsqu'un groupe consonantique symbolise deux phonèmes, ces graphies vont également vers la simplification. Les sons plosifs sont les plus prompts à disparaître, d'où la figuration de <st> en <s(s)> dans <pas('), tess> ou <breakfus>, ou celle de <ld> en <l(e)>. Le groupe n'est sauvegardé que lorsque la consonne finale est suivie par une voyelle, comme <musta> (où le <a> figure *have*).

Les consonnes muettes dans le système phonologique de l'anglais standard, comme le <b> de *crumble* [19] <crummel> disparaissent logiquement, parce que ne sont écrites que les consonnes prononcées dans ces groupes–mais encore une fois, la consonne <m> est doublée pour ne pas perdre de lettres dans le mot. On aurait pu imaginer que conserver la longueur du mot aide à sa lisibilité pour le lecteur lambda. La bouma, ou forme générale du mot considérée comme lue en un bloc par les ty-

18. La phrase est tirée d'une scène de viol incestueux.
19. Le /b/ est prononcé dans *crumble* en anglais standard, mais pas en AAV.

pographes et les spécialistes de la lecture, n'est en fait pas si déterminante pour la lisibilité. J'ai déjà évoqué la conservation de la longueur d'un mot, mais elle me semble plus importante que la bouma en soi. Le contour global des mots ne joue qu'un rôle auxiliaire dans la lecture : « La reconnaissance visuelle des mots ne repose pas sur une appréhension globale de son contour, mais sur sa décomposition en éléments simples, les lettres et les graphèmes » (Dehaene, 2007b). Mais l'élaboration consciente de graphies alternatives, le principe de conservation du nombre de lettres du mot joue un rôle pour que <crummel> n'ait certes pas la même bouma que <crumble>, mais que le mot en conserve le même nombre de lettres. Il était beaucoup plus difficile de lire des mots réduits à une ou deux lettres, comme <m> pour *man*. En revanche, il semble être plus aisé de lire des mots de même longueur, et les transcriptions plus longues comme <mostlee> (*mostly*) sont rares dans le corpus. Les scripteurs n'y pensent pas, parce qu'allonger un mot est contre-intuitif. Remplacer une lettre d'un mot par une autre, en conservant le nombre de caractères dans le mot, semble être un type de changement courant.

C'est également la raison pour laquelle *understand* est retranscrit <unnerstand>, alors qu'un hypothétique *<unerstand> aurait représenté les mêmes sons avec autant d'efficacité [20]. Cette graphie a d'ailleurs plus de poids que le groupe consonantique final <nd> du même mot, écrit <unnerstan> (occurrence hapaxique). La trace au moins symbolique des lettres omises, qui correspondent à des sons omis est primordiale, sauf pour les formes validées dans l'anglais vernaculaire.

<Nd> est également altéré pour montrer la réduction de /nd/ avec d'autres stratégies, comme le montre l'exemple « I mine my bizness » (p. 37). L'agrégat consonantique final /nd/ de *mine* '*mind*' est réduit, et la dentale finale est amuïe. L'homographie avec « mine » /ˈmaɪn/permet de garder la tension de la voyelle dans un environnement <_C <e># > tout en ne perdant pas de lettres dans le mot. La simplification de groupes consonantiques (en particulier les groupes terminés par une plosive) est une caractéristique des *connected speech processes* de l'anglais parlé.

La dévélarisation de /ŋ/est complète dans le corpus de *Push*, mais les apostrophes d'excuse qui suivent le <n> final sont légion (79% des occurrences de <ng> final transcrit <n'>) : il n'y a que 9 occurrences de participes présents finissant en <-in> ce qui montre le souci que représente la disparition totale d'une marque graphique inculquée. L'apostrophe-symbole de la lettre manquante est la marque

20. Si l'on imagine une prononciation avec un /d/ flappé, /ˌʌɾ̃ˈstænd/.

de l'auteure qui connaît la graphie standard. Cette pratique est répandue dans le corpus internet, où les utilisateurs ont le même parcours entre la standardisation de la graphie et le détachement ultérieur, qui laisse des traces graphiques [21].

Quant aux graphies complexes et irrégulières, comme <gh> pour /f/, où le phonème ne correspond pas à la combinaison des phonèmes, ils sont régularisés en <ff>. Ce digraphe est particulièrement problématique, à cause de sa valeur silencieuse (dans <ough>, comme « through », écrit <threw, throo, thru>, ou « daughter » transcrit <dater> dans le livre). Le flottement est résolu avec le digraphe <ff>, qui souligne encore une fois le respect de la conservation de deux lettres pour figurer <gh>.

La disparition du <h> de la séquence <#wh> est due à la nature grammaticale de ces mots (*what*, *when*). Leur fréquence rend leur érosion plus facile, et l'amuïssement complet du /h/ ne joue pas en sa faveur. C'est la raison pour laquelle il disparaît de presque toutes les occurrences de <#wh>.

En revanche, la monographie de digraphes orthographiques est récurrente, et, si elle semble aller à l'encontre de la démonstration précédente, elle est soutenue par la règle de simplification. Les consonnes répétées, comme <dd> dans *daddy* délivrent le même message alors que les groupes consonantiques, même ceux dont les graphèmes sont amuïs, gardent un poids inconscient.

L'appréhension de la phonologie et le poids de la prononciation jouent un rôle important dans les graphies alternatives, qui contreviennent parfois à la conservation du nombre de lettres dans un mot. Au contraire, elles peuvent parfois rallonger les mots : la graphie <culchure> (transcription de /ˈkʌlʧər/) est plus longue que le standard *culture*, mais elle transcrit la perception de la fricative /ʧ/ et sa transcription en <ch>, au détriment du <t> abandonné en syllabe inaccentuée. La fricativisation de /t/ en /ʧ/ ou /ʃ/ avant certaines voyelles (comme dans *nation* ou *nature*), et la graphie de *church* en <chuc(h)> dans le corpus indique la connaissance de l'équivalence graphophonémique entre /ʧ/ et <ch>. C'est la raison pour laquelle cette graphie, plus logique, a été adoptée pour transcrire ce phonème.

Certaines règles graphiques ne sont pas enfreintes, comme l'absence (sauf certaines langues [22]) de consonne répétée en début de mot (*accept* est écrit <cept> et non *<ccept>) ; phonotactiquement

21. La comparaison avec un corpus d'apprenants de l'anglais serait ici très intéressant, pour connaitre la proportion de <ng>, <n> voire <n'> pour transcrire /ŋ/

22. Les <ll> initiaux de <llama> ou <Lloyd> étant des emprunts.

parlant, et il y a peu de séquences jugées improbables en anglais comme *<hv>, au moins à l'initiale (cet exemple apparaît cependant dans <fahver> (‘*father*’). Le <h> est en fait accolé à la voyelle : c'est une construction en <fah> + <ver>. Precious a déjà élaboré la correspondance /æ/ ; /aː/ = <ah> dans la graphie <all-tur-nah-tive>, où les tirets montrent qu'elle décompose le mot, et où aucune des syllabes n'est réduite, puisqu'elle tente de s'approprier le mot en le décomposant. La consonne graphique a une valeur vocalique, sinon postvocalique dans ces deux graphies. Ce passage à une utilisation vocalique permet d'enchaîner avec la partie suivante, qui traite des voyelles et de leur représentation graphique.

1.2.2.4.2 Voyelles

Precious s'appuie sur sa prononciation pour graphier les mots, comme en témoignent les graphies de mots inconnus. Ainsi, les tensions et longueurs des voyelles, ainsi que les voyelles inaccentuées sont rajoutés par Sapphire, pour coller au plus près à la prononciation de la protagoniste. Les voyelles de l'anglais sont un sujet délicat à traiter, à cause du nombre de prononciations différentes pour chaque caractère graphique. Pour Precious, le système graphique, notamment des voyelles non-accentuées, est problématique. Elle utilise un système peu régulier, et ses graphies changent. Dans cette partie, je pars de la phonologie pour expliquer les choix graphiques. Les schwas sont les voyelles les plus intéressantes, et leur transcription suit deux trajectoires différentes. Soit ils sont représentés par un <u>, soit par un <e>, à cause de la proximité phonique entre /ʌ/ et /ə/, qui semblent fusionner dans son parler, et donc dans ses écrits. Les graphies alternatives marquant une voyelle épenthétique (devant /m, n, l/) contiennent un <u> préconsonantique, comme dans les graphies <appul, Cambull, emoshunal, sumebuddy>. En revanche, dans <crummel> (/ˈkrʌmbəl/), Sapphire a décidé de ne pas doubler le <u> graphique, et n'a pas écrit <crummul>. D'ailleurs, les consonnes syllabiques introduisent également des suppressions totales de voyelle graphique, comme dans <Harlm, journl, coln> (‘*Harlem, journal, colon*’). Certaines lettres muettes sont d'ailleurs également supprimées, comme dans *Catholic*, qui devient <Kathlic>. Cette graphie élimine entièrement l'hypothèse d'une prononciation du <o>, comme la variante attestée /ˈkæθəlɪk/et assoit la prononciation /ˈkæθlɪk/. Les graphies montrent un choix de compressions syllabiques.

Le <u> est une sorte de joker graphique, qui recouvre le plus de sons par rapport aux autres voyelles dans la graphie de Precious, comme le montre le tableau 1.4. Son statut est un peu ambigu en anglais standard, car c'est la seule lettre qui puisse représenter deux sons en syllabe accentuées et fermées, /ʌ/ ou /ʊ/, avec palatalisation optionnelle de /ʊ/. Ici, elle sert à figurer presque toutes les voyelles postérieures et centrales moyennes. Les sons représentés sont, pour la plupart, peu éloignés des

Son représenté	Exemple	Transcription phonétique (LPD)	Orthographe
	Syllabe accentuée		
/uː/	impr**u**v	/ɪmˈpruːv/	improve
/ʊ/	g**u**d	/ˈgʊd/	good
/ɜː/	all-t**u**r-nah-tive	/ɔːlˈtɜːnətɪv/	alternative
/ʌ/	s**u**mebuddy	/ˈsʌmbədi/	somebody
	Syllabe inaccentuée		
/ə/	app**u**l	/ˈæpəl/	apple
/ɑː/	anyb**u**ddy	/ˈenibɑːdi/	anybody
/ɒ/	wh**u**tevr	/wɒtˈevər/	whatever

Tab. 1.4 – Retranscriptions erronées utilisant <u>

valeurs attendues phonographémiquement. Il existe une certaine confusion au niveau des voyelles d'arrière, qui sont presque toutes représentées par < u > dans Push (sauf les deux voyelles moyennes postérieures /o/ et /ɔ/). Ce choix est systématisé sauf dans le cas ambigu de <anybuddy>. Elle est certainement due à une confusion à cause de la graphie de *buddy* moins qu'une transcription de la prononciation, qui peut se doubler de l'indication de la réduction vocalique en /ˈenibədi/, logique avec le choix presque constant de < u > pour symboliser un schwa. Quant à la graphie < impruv >, elle est presque logique : Precious a conscience de l'équivalence entre le son /uː/et <u> en syllabe ouverte (comme les graphies <blu, tru> '*blue, true*' le prouvent, plus loin dans le livre), mais elle applique la même règle en syllabe fermée. Les principes graphophonémiques d'anglais standard sont donc presque intégrés, mais parfois, la systématisation est approximative [23].

Le résumé des sons vocaliques recouverts par les graphies de Precious (diphtongues et digraphes exclus) est présenté dans l'illustration 1.2 qui accompagne le tableau A.1 (placé dans les annexes). Cette répartition est logique, la plupart du temps et illustre notamment la fusion de phonèmes entre /ɪ/

23. La séquence graphique <ov> aurait pu faire penser aux autres mots de ce type, comme *government* ou *love*, dont les prononciations en /ʌ/rappelle la graphie médiévale en <uv>. Pour *improve*, apparu en anglais au 15ème siècle et graphié <improuwe>, la question ne se pose pas.

et /e/ . Encore une fois, seuls des mots aux graphies alternatives ont été choisis, pour montrer les choix conscients de la scripteuse dans ses réélaborations graphiques.

En revanche, comme il n'y a aucune volonté de la part de Precious de produire des graphies alter-

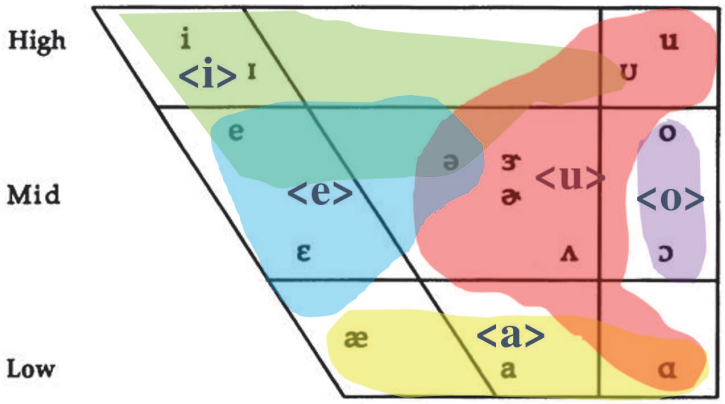


Illustration 1.2 – Répartition des graphies des sons vocaliques

natives, la voyelle <y> est absente des réélaborations graphiques pour les sons vocaliques simples (à l'inverse de la graphie alternative la plus connue de *bitch*, <bytch>). Les <y> n'apparaissent qu'à leur place (<dyin'>) ou dans des diphtongues, qui sont étudiées dans les paragraphes suivants.

La graphie <vaykent> retranscrit la diphtongue /ˈveɪkənt/ et prend l'allongement de /eɪ/ en compte. Le choix graphique d'un <ay> pour représenter /eɪ/ s'explique par la fréquence de mots comme *[Mon]day*, que Precious doit connaître pour écrire dans son journal, bien que la séquence soit dans une syllabe inaccentuée. La représentation par un digraphe d'une diphtongue est plutôt logique, d'autant plus que Precious a tendance à rétablir les équivalences phonographématiques dans ses écrits (comme la graphie <pritty> pour *pretty*, ou <bild> pour *build*). D'autres diphtongues comme /oʊ/ sont rétablies, comme le montre la graphie <goe> pour go [24]. Dans l'étude du corpus secondaire, je ne développerai pas le traitement des voyelles suivant leur valeur réduite, accentuée ni des monophtongues et des diphtongues, que je garde pour l'analyse du corpus principal.

Quant aux voyelles tendues, elles sont traitées instinctivement comme des lettres doublées. Ce système permet d'indiquer la longueur de la voyelle par la longueur du mot. <Concreet> ('*concrete*'en

24. Qui représente un « goes » dans le texte—mais la troisième personne du singulier n'est souvent pas marquée dans *Push* (Hémart, 2001).

est une bonne représentation. La voyelle longue /i:/ est représentée par un doublement de la lettre graphique, proche des exagérations pour retranscrire la parole comme <pleeezzze> (*please*). Les deux phénomènes montrent la maîtrise de l'équivalence phonographématique entre la séquence <ee> et le son /i:/ . Le <e> final n'est pas conservé, puisqu'il est muet.

Dans toute l'oeuvre, Precious a conscience de la façon dont elle prononce les mots, et cette prononciation influence la graphie. Les graphies fluctuantes, comme celles des utilisateurs de forums, n'en ont cependant pas les mêmes causes. L'intérêt du corpus de *Push* réside dans le fait que ces graphies alternatives interrogent la pseudo-régularité du système phonographématique de l'anglais. Loin de vouloir se distinguer, Precious essaie d'être la plus rigoureuse possible dans son apprentissage, et ces graphies représentent le cheminement vers l'acquisition de la langue, avant la subversion du système. Le placement relatif de ces deux éléments est représenté dans le schéma 1.3.

Les productions choisies pour Precious (en vert) sont la première phrase de son journal, puis trois

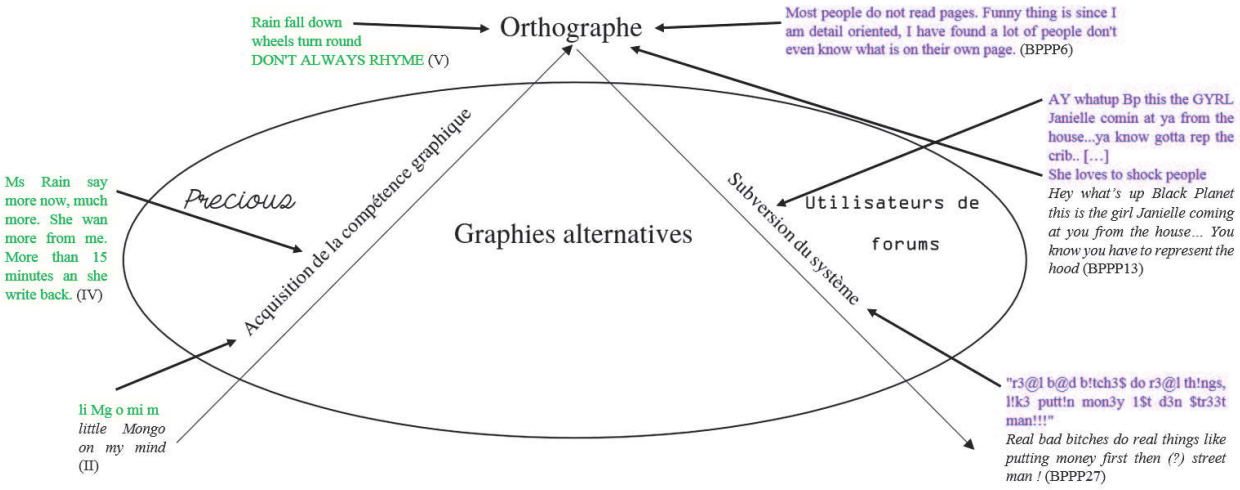


Illustration 1.3 – Schéma récapitulatif des graphies de Precious

phrases tirées de la IVème partie, et enfin un extrait des poèmes insérés dans la Vème partie. Les trois citations illustrent la progression de ses compétences graphiques : le premier est un état des lieux de sa littéracie orthographique au début de son apprentissage à l'école alternative, le second montre ses progrès (entérinés par la requête de sa professeure, qui lui demande d'écrire plus dans son journal). Les trois vers de sa composition qui servent à illustrer sa compétence orthographique ne sont pas exempts

d'une faute de grammaire (accord sujet-verbe entre « rain » et « fall »), mais la graphie de l'anglais standard est maîtrisée.

J'ai choisi de faire un parallèle avec des profils personnels sur BP, qui présentent le même type de dimension diariste épistolaire [25]. À l'acquisition des compétences graphiques tardives chez Precious répondent les subversions du système orthographique que l'on peut supposer appris à l'école par les membres de BP. L'extrait de BPPP6 (le plus proche de l'étiquette « Orthographe ») est orthographiquement irréprochable. L'utilisateur s'y vante d'ailleurs de faire attention aux détails, et tout son profil est écrit dans ce style. Le second exemple choisi est un profil d'une jeune utilisatrice de BP (BPPP13). Elle ouvre son profil avec plusieurs graphies alternatives, comme <gyrl>, <comin> et <ya>, mais dans la seconde partie de son profil (où elle se décrit à la troisième personne du singulier), elle repasse à une graphie normée, sans même utiliser de mots argotiques jusqu'à la fin de son profil. Dans ce cas, le code orthographique est maîtrisé, ce que l'on peut voir dans cette seconde partie, mais l'utilisatrice a désiré ouvrir son profil en injectant quelques graphies alternatives. Le troisième exemple, où le degré de subversion du système orthographique me semble le plus important, l'utilisatrice remplace systématiquement les voyelles <a>, <e>, <i> et la consonne <s> par <@>, <3>, < !> et <$>. Le deuxième et le troisième extrait de profils personnels contiennent la graphie alternative <n> pour <ng> en fin de participe présent, ce qui veut dire que les mêmes stratégies sont utilisées par divers scripteurs. Néanmoins, ce remplacement n'est pas systématique, puisque la terminaison de <ng> de <th !ngs> n'est pas affectée. Cela se rapproche également des graphies de Precious.

Le dernier pan de l'analyse des graphies de Precious concerne les entrées de son journal, et la façon dont elles évoluent à mesure que le roman progresse.

1.2.2.5 Étude épistolaire

Les parties épistolaires sont souvent introduites par la mention de la date, et contiennent des marques phatiques de la communication écrite par lettres (date, adresse, et signature. Rain reproche d'ailleurs à Precious de ne pas écrire la date, ce que lui retourne ironiquement Precious quelques lettres plus tard

25. La réponse du coénonciateur potentiel est contenue dans le « Guest Book » sur ce site.

(« Don't forget to put the date, 1/18/88, on your journal entries » (p. 70) ; « Dan frget rite day Ms R », '*Don't forget to write the date Ms Rain*' (p. 72)).

Dans les parties II et III, Rain corrige les écrits directement sous les productions de Precious, et cela est signifié au lecteur par la présence de parenthèses qui encadrent un texte écrit en police inférieure à celui du corps du texte. Le volume de ces corrections est beaucoup plus faible quantitativement que les entrées rédigées par Precious. Ces corrections n'interviennent qu'aux parties épistolaires, mais le corps du roman est stylistiquement attribué à Precious, avec de nombreuses utilisations de grammaire et de graphies alternatives afro-américaines et substandard. Les corrections constituent une sorte de retour à la norme, et donc d'arrachement au corps du texte, c'est pourquoi les graphies standard non idiosyncrasiques sont dévaluées et perdent donc leur taille sur la page. En revanche, quand Precious se corrige elle-même, elle barre et remplace le mot, à la suite de la graphie erronée, juste avant de continuer à écrire dans son journal. Ce phénomène n'apparaît qu'à la fin du livre (dernier tiers, partie IV), lorsqu'elle maîtrise assez l'orthographe et a un recul suffisant sur ce qu'elle produit.

D'autres mécanismes typographiques entrent en jeu pour montrer au lecteur les difficultés rencon-

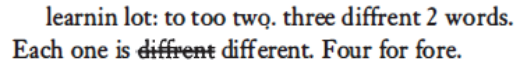

learnin lot: to too twǫ. three diffrent 2 words.
Each one is ~~diffrent~~ different. Four for fore.

Illustration 1.4 – Exemple de correction par Precious (p.98)

trées par Precious : quand elle ne sait pas écrire un mot, elle en écrit la première lettre et tire un trait : « Gr___ » (p. 71), pour '*grandmother*'. C'est une technique que lui a apprise sa professeure : « When I ka can not spell a word Ms Rain tell me sound out firs[t] lett[e]r c___ and draw a lin[e].[…] Latrer [later] she will fill in rite [right] spelling for me. » (p.112). Ces deux techniques de correction sont présentées dans l'illustration 1.5.

Aux représentations gabaritiques succèdent donc des graphies squelettiques, signalées comme fau-

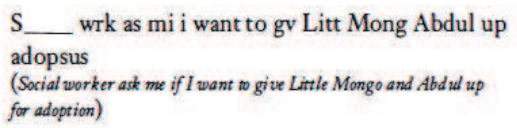

S____ wrk as mi i want to gv Litt Mong Abdul up
adopsus
(*Social worker ask me if I want to give Little Mongo and Abdul up for adoption*)

Illustration 1.5 – Occupation spatiale d'une graphie squelettique (p. 70)

tives. Precious fait attention à ce qu'elle écrit, et signale par ses soulignements qu'elle estime se trom-

per. Le même phénomène se retrouve dans les forums, où les utilisateurs signalent leurs hésitations par un point d'interrogation entre parenthèses <(?)> après un mot, ou même une petite incise, comme « Jazmine Harrington (Probably spelt wrong) » (BPPP13) ou « Onassis (probably spelt wrong) » (SC11J42). Dans les lettres de Precious, le soulignement est le symbole le plus simplifié de cet avertissement. Le soulignement de caractères inexistants (ou non-écrits) montre aussi l'importance de la longueur physique du mot, et de sa spatialisation dans la page [26]. La longueur ou l'espace sont donc primordiaux. Ces soulignements physiques peuvent aussi signaler une évolution ultérieure : le remplissage, effectué par Blu Rain ou Precious (illustration 1.6).

À partir de la quatrième partie, Blu Rain ne corrige plus ce que Precious écrit avec un système

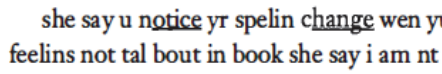


Illustration 1.6 – Exemple de remplissage dans Precious (p.98)

de parenthésage. Le carnet de Precious devient une sorte de journal moins qu'un exercice d'écriture corrigé par Blu Rain. Precious ne le lui montre plus systématiquement et les extraits présentés dans le roman sont de plus en plus longs et rédigés.

1.2.2.5.1 Étude longitudinale des réélaborations graphiques

Dans cette section, je ferai une courte étude quantitative des graphies alternatives. Contrairement aux forums, le roman est au service d'une stratégie narrative. Je vais donc montrer la convergence entre l'évolution des formes alternatives et le développement narratif. Cette tâche est aisée parce que ce corpus clos se prête à une analyse quantitative précise. Les données ne sont pas labiles comme celles des forums étudiés.

De simple syntagme nominal « li Mg o mi m » '*Little Mongo on my mind*' (p.61), Precious passe progressivement à des phrases décousues mais complètes, puis à des descriptions de parties de Harlem (p. 102-104). Bien qu'elle ait commencé à partir de syntagmes très simples ou de phrases assertives

26. Laisser un espace blanc est une pratique courante d'étudiants en traduction, lorsqu'ils omettent de traduire un mot, par exemple.

pour finir par de la poésie, déjà évoquée par Precious p. 90 (elle signe « Precious Jones – The Poet »).

Il y a 24 entrées distinctes au journal de Precious (23 et une distincte, dans la dernière partie du livre). Mon hypothèse de travail de départ consistait à dire que l'évolution chronologique, qui correspond à l'incrément des entrées (qui sont présentées dans l'ordre dans le livre) correspondait à une diminution du pourcentage de fautes d'orthographe. Par conséquent, mon hypothèse zéro est la suivante : l'évolution chronologique, qui correspond à l'incrément des entrées n'avait aucun rapport avec le pourcentage d'erreurs. J'ai donc créé un graphique avec le pourcentage d'erreur en abscisse, le facteur « MisPercent » et les entrées de son carnet (facteur « Entry ») dans leur ordre d'appartition, en ordonnée. Les données traitées sont présentées sur la clef USB d'annexe. Ce graphique 1.7 se présente ainsi :

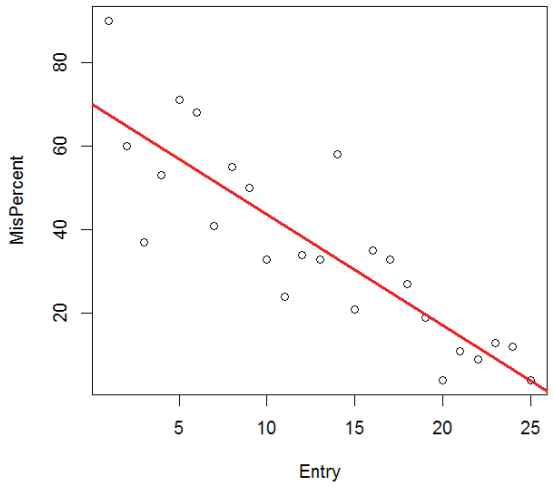


Illustration 1.7 – Pourcentage d'erreurs dans les entrées

La marge de progression de Precious est constante et visible. La ligne rouge représente la moyenne des erreurs par entrée, et les points (qui correspondent aux entrées par rapport à leur pourcentage d'erreurs) restent effectivement proches de celle-ci. J'ai effectué un test de corrélation sous R [27], et j'ai opté pour la méthode de corrélation de Spearman parce qu'elle est efficace pour traiter des relations à

27. Formule : `cor.test(Entry,Mispercent, alternative = c(``less''), method = c(``spearman''))` sous R.

un seul prédicteur (en l'occurrence, le pourcentage d'erreurs graphiques) de moins de 50 échantillons. La valeur p obtenue [28] était de 3.244^{-09}. L'hypothèse nulle est donc rejetée : la littéracie de Precious augmente donc significativement avec le temps, et elle fait de moins en moins de fautes d'orthographe.

D'autres études quantitatives peuvent être tirées de ces écrits, car le pourcentage de fautes d'orthographe ne montre pas tout : une graphie alternative peut être tout à fait compréhensible. La facilité à lire un texte n'est pas un corollaire de la proportion de fautes d'orthographe, c'est pourquoi, pour peindre un tableau plus précis du style graphique, j'ai choisi d'utiliser le score de lisibilité Flesch-Kincaid. Ce test est basé sur le nombre de mots, de syllabes et de phrases, et donne un résultat sur 100. Les textes les plus difficiles à lire ayant un score compris entre 0 et 30 ; un score entre 90 et 100 correspond à ce qu'un enfant de 11 ans peut lire sans problème. La formule pour calculer la lisibilité de Flesch-Kincaid est la suivante (Kincaid, Fishburne, Rogers & Chissom, 1975) :

$$206.835 - 1.015\left(\frac{Nombre\ total\ de\ mots}{Nombre\ total\ de\ phrases}\right) - 84.6\left(\frac{Nombre\ total\ de\ syllabes}{Nombre\ total\ de\ mots}\right)$$

Mon hypothèse de départ est la suivante : plus elle écrit, et plus ses écrits sont lisibles, ou intelligibles. En revanche, l'absence de ponctuation dans le texte (par ignorance) peut éventuellement influencer les résultats. Je postule donc une grande facilité à lire les deux ou trois premières entrées, à cause de leur caractère proto-gabaritique, où les mots sont tellement courts et où il y a si peu de phrases qu'elles affecteront les paramètres du nombre total de mots, de syllabes et de phrases. À quoi succède une diminution de la facilité de lecture, à cause de l'accroissement de la proportion de mots polysyllabiques, suivie d'une simplification à mesure que Precious maîtrise les signes de ponctuation, et crée des phrases plus courtes. Les données brutes [29] ont permis de tirer le graphique et tableau récapitulatif 1.8. Visiblement, les données suivent une progression chaotique, et cette irrégularité est confirmée dans le graphique tiré de ces données et présenté en 1.8.

28. Qui détermine si l'hypothèse zéro est entièrement rejetée, et donc si « le résultat du test est statistiquement significatif » (Wasserman, 2004). Pour que l'hypothèse zéro soit rejetée, il faut que la valeur p soit inférieure à 1.

29. Les tests de Flesch-Kincaid ont été réalisés sur le site <https ://readability-score.com/> (dernier accès le 23 avril 2014) et les résultats ont été exploités sous Excel.

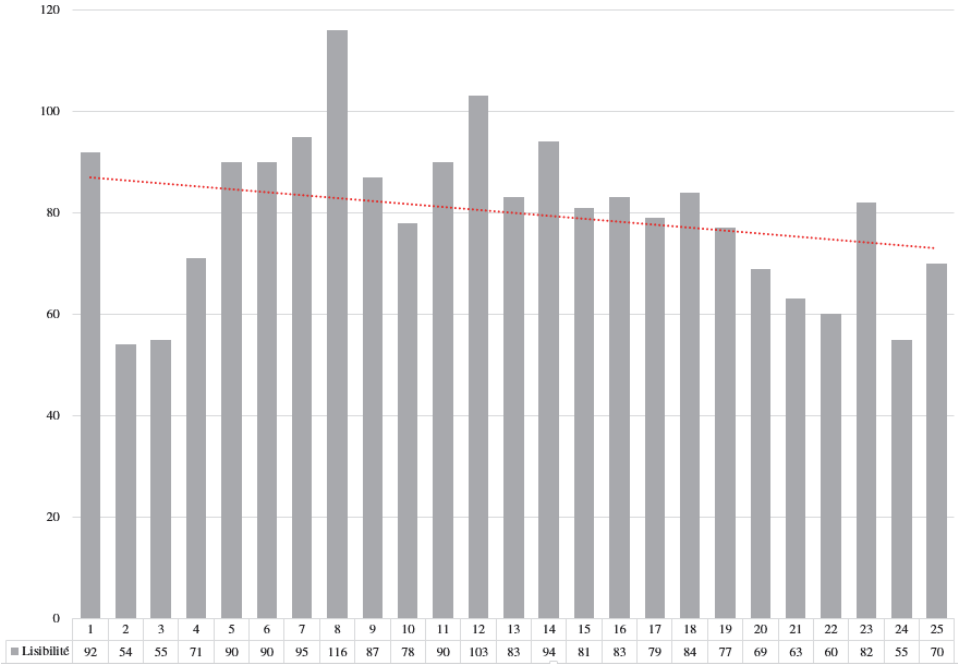


Illustration 1.8 – Résultats de Flesch-Kincaid et courbe de tendance

Les études quantitatives peuvent ne pas avoir de sens, avec un type de corpus aussi inconventionnel, et il n'y a pas d'indication de période supposée d'apparition (ou du délai entre les entrées du carnet). La lisibilité trop grande (au dessus de 100) est un indice de l'inadéquation de ce corpus avec les outils traditionnels d'exploration textométrique. Les syllabes n'existent parfois pas en tant que tel (comme dans un texte dans lequel elles pourraient être calculées) puisqu'elles sont parfois, surtout dans les premières entrées, symbolisées par des lettres isolées, comme « li Mg o mi m » '*Little Mongo on my mind*'. Les résultats sont donc exagérés, et si la courbe de tendance (en rouge sur le graphique) indique une plus grande difficulté de lisibilité, ce résultat est mis en échec. La baisse de la lisibilité s'explique par la complexification de ses entrées et sa plus grande maîtrise de la langue écrite. Pour un lecteur humain, les graphies proto-gabaritiques sont difficiles à comprendre, mais les entrées sont effectivement courtes et moins élaborées, et les phrases sont souvent plus simples. On peut ainsi comparer « th wfr hlp mma it help mi. Precious Mi » '*The welfare help mama. It help me ?*' (entrée 8), où les phrases sont minimales et où la syntaxe conventionnelle pour poser une question n'est pas respectée, et « Whut is my erliest ~~memry~~ memory of my mother ? a room that's small fillt up wif my parents, it

smell, can of mackerel left open in the kitchen on hot day that's what make me remember […] » (entrée 21). Ici, la syntaxe interrogative est rétablie, la phrase suivante contient une proposition relative et des appositions, même si toutes les règles de grammaire ne sont pas respectées. Les mots sont presque tous écrits correctement, et présentent beaucoup moins de graphies gabaritiques que dans la septième entrée. L'évolution scripturale s'accompagne d'une complexification, et l'on part d'une simplicité plus qu'enfantine à des phrases complexes. Quant aux dernières entrées du journal de Precious étiquetées en tant que telles, elles sont constituées de parties poétiques. Le texte est mis en forme et les lettres comptent, ainsi que les indices (dessinés) par Precious. La sémiotique en est tout à fait différente, et c'est la raison pour laquelle l'entrée 24 (illustration 1.9) a un score bien plus bas que les autres (à 67.4, le second score le plus bas le dépasse de près de cinq points). Le jeu visuel de placement des mots n'est pas calculable par un algorithme et montre clairement la fonction poétique assumée par le texte. L'absence de ponctuation est compensée par la présence de sauts de ligne et de mise en forme du texte.

Sur internet, on pourrait penser à un score de style inverse : plus on écrit, et plus on se forme un

```
rootine b
         r
          e
           a k
i run get dress
     fix tea (don't like coffee)
     grab books
               walk
     mornin wet the streets
     amung the vakent trees
          is secrit plots
          of green diamonds
          call grass.
```

Illustration 1.9 – Poème « Rootine Break », entrée 24 du journal de Precious (partie V)

habitus fait de ses particularités graphiques personnelles, prises dans un communaulecte propre à la communauté. À force de tests et de tâtonnements linguistiques, l'on réussit à se forger une identité qui passe par ses écrits. De la même façon, le discours de gang est forgé par les prescriptions spécifiques au milieu culturel. La poétique de Precious trouve un écho dans les efforts de mise en forme sur certains

profils personnels sur BP, où changer de police, de taille de caractère…montre la maîtrise des outils de mise en forme.

Le discours épistolaire de Precious et sa sophistication croissante répond à ses propres questions en matière de compétence linguistique. À mesure qu'elle acquiert des connaissances graphiques, elle remet en question le système linguistique, et ces explorations à la fois métalinguistique et épilinguistiques sont le sujet de la partie suivante.

1.2.2.5.2 Discours métalinguistique et épilinguistique

Precious fait quelques mentions ou commentaires sur la façon d'utiliser l'anglais écrit, ou sur la façon dont la langue se construit. Le premier est « capital letter is how you start off sentence or say something with deep shit meaning like Fuck with capital F you mad or some shit like that ! » (Partie IV, p. 114). Cette citation contient deux remarques linguistiques pertinentes pour cette étude : la lettre majuscule en début de phrase, et l'emphase apportée par les capitales sont deux phénomèmes différents.

Les conventions graphiques et l'emphase (que l'on peut voir sur les forums, de la même façon) les mots en majuscules n'y sont pas pour uniquement marquer *something with deep shit meaning*, pour reprendre les mots de Precious, mais aussi les expressions de l'émotion (colère, humour pour certains, emphase, excitation). La mise en majuscule est un indice d'importance du mot– ou une convention– visuelle forte. Il est surprenant de voir ce genre de remarque, alors que l'on pourrait s'attendre à une habitude de « lire » ou d'être exposé aux panneaux et autres textes écrits, et ce même sans les comprendre.

Les lettres majuscules ont parfois des formes différentes de leurs équivalent minuscule [30], et c'est la raison pour laquelle certains documents pédagogiques conseillent aux professeurs des écoles d'apprendre l'alphabet en lettres capitales après l'alphabet en lettres minuscules, et ce, dans de nombreux pays (Ontario, 2006 ; Gregorio, 2006 ; de Clermont-Ferrand, 2011 ; University of Texas System, 2009). La compréhension des conventions de la langue écrite est centrale, et la première remarque

30. Même en écriture en script, qui est utilisée aux États-Unis pour apprendre aux enfants à écrire. Grâce à ce style d'écriture, chaque caractère est indépendant, contrairement à l'écriture cursive où les lettres sont liées.

porte sur la construction graphique de la phrase, telle qu'elle est enseignée traditionnellement à l'école. Les majuscules sont à part, parce qu'elles sont cantonnées aux initiales ou au discours d'avertissement, comme le panneau STOP, blanc sur rouge, à la forme exceptionnelle, par exemple. L'apprentissage de la convention des majuscules pour remplacer un cri, par opposition aux lettres minuscules qui représentent un discours normal(isé). Et effectivement, l'apparition d'un paradigme « [mot] with a capital [lettre initiale] » n'existe qu'à l'écrit, où l'on appuie l'importance du mot en lui mettant une majuscule. Il est écrit dans *Push*, mais aussi dans les corpus internet, où le recours à l'oral est impossible et où l'insistance peut se faire en modulant la taille des caractères. Les lettres majuscules ont donc un statut de variation du discours à valeur quasiment intonatoire, et montrent dans les deux corpus l'importance du mot écrit ainsi.

Le discours de Precious se fait parfois l'écho de celui de certains utilisateurs de *Black Planet* ou *Scotster* à propos de la stigmatisation d' utilisateurs de langues minoritaires en pays anglophones : « Secon' thas when the "I'mma joke" start. » (‘*Second [grade], that's when the 'I'mma'jokes started*’) (partie II, p. 36). Les Afro-Américains sont tournés en dérision par l'utilisation de leur langue par les locuteurs d'autres lectes. *I'mma* est une expression afro-américaine courante, qui signifie ‘*I'm going to*’ et elle était associée (surtout à l'époque où *Push* a été écrit) à un parler vernaculaire et socialement réprouvé par ceux qui ne le pratiquent pas. Le commentaire social sous-jacent est sans appel, et nous sommes confrontés à un bel exemple de linguistique populaire qui se révèle être perceptuelle (Paveau, 2007), où chacun juge selon ses propres critères internalisés ce qui constitue la norme ou non. Par extension, faire partie de la norme internalisée signifie ne pas être confronté à la raillerie sur sa façon de parler, ou, au moins, ne pas être celui qui reçoit les railleries.

L'écriture est donc pensée par Precious, et elle s'empare de tous les outils linguistiques à sa disposition. Les discours épilinguistiques sont rares, mais ils montrent parfaitement la démarche cognitive que nous effectuons tous. D'une part, l'adaptation du discours aux émotions ressenties, et d'autre part, la difficulté à s'exprimer dans une variété d'anglais ou dans une langue minoritaire jugée comme étant inférieure, lorsqu'on est jugé par les autres. C'est la raison pour laquelle certains sont fiers des fautes qu'ils font à l'écrit, ou les incorporent comme faisant partie de leur personnalité :

i know how 2 spell i aint dumb dats just da ghetto talk (BPPP14)

And by the way, I know how to spell, its just the ghetto in me talkin‘ to ya‘ll right now ! (BPPP15)

Les deux utilisateurs justifient leurs graphies comme étant « ghetto », mais ne parlent de cela qu'à la fin de leur description. Le choix du non-recours à la normativité est intéressant : les deux utilisateurs disent savoir épeler correctement les mots, mais choisissent tout de même d'écrire comme il leur chante. Ces exemples montrent l'importance extrême de la norme sur laquelle on s'appuie pour prouver que les graphies alternatives ont été réfléchies et représentent soit une émotion, soit sa personnalité (« the ghetto in me »).

Dans le dispositif fictionnel, les graphies de Precious et leur évolution est le point le plus important de cette étude. Ces graphies présentent des caractéristiques communes avec celles trouvées dans le corpus principal. Le carnet de bord de Precious est une caractéristique unique de ce roman, et il permet de commencer à appréhender la différence de discours des utilisateurs de forums entre leurs interventions dans des conversations, et leur profil personnel. Le corps de la narration n'est pas exempt de graphies alternatives, et contient le discours épilinguistique de Precious. Ceci permet d'appréhender les diverses opinions autour du standard graphique. Celui-ci est volontairement altéré par les utilisateurs de BP, tout comme il l'est par Sapphire, qui a décidé de donner plus de pouvoir stylistique à Precious. C'est pourquoi elle a écrit tout le roman en ayant recours à des graphies alternatives.

Je vais maintenant orienter l'étude sur le film tiré de cet ouvrage, afin d'en explorer les liens linguistiques, et surtout, la dimension orale du parlécrit qui charpente le livre.

1.2.3 *Precious*

Precious est un film américain, réalisé en collaboration avec Sapphire par Lee Daniel. L'adaptation a forcément édulcoré certaines parties du livre, mais aussi le style de parole de Precious et des autres personnages, afin que les dialogues soient compréhensibles mais sonnent tout de même authentiques, grâce à des acteurs qui parlent AAV ou l'anglais de New York. Le public américain, qui est le premier concerné par le film, n'a pas connu les mêmes déboires que celui de *Trainspotting*(cet aspect sera traité

en 1.3.1.2).

1.2.3.1 Caractéristiques du média

La distribution oscille entre acteurs inconnus, comme Gabourey Sidibe (Precious), dont c'était le premier rôle, ainsi que des personnalités du monde de la musique, ou des humoristes comme Mo'Nique (Mary) ou Mariah Carey (Mrs Weiss), ainsi que Lenny Kravitz (Nurse John)[31]. Le film suit peu ou prou la même progression que le livre, de l'exclusion de Precious de son collège à l'entretien entre Precious, sa mère et l'assistante sociale, où sa mère décrit les abus dont Precious a été victime. La dernière scène correspond à la page 136 du livre, quelques pages avant la cinquième partie qui recueille les histoires des étudiantes de Blu Rain et quitte la narration. Le seul personnage qui subit une transformation radicale est Nurse John (L. Kravitz), qui apparaît dans *Push* dans le souvenir de l'accouchement de Mongo en tant qu'un infirmier hispanique. Dans *Precious*, il se lie d'amitié avec elle et les étudiantes de l'école alternative *Each One Teach One* quand Precious accouche d'Abdul, son second enfant.

La forme du film est souvent dépouillée : il y a peu de musique de fond, ce qui permet une étude phonétique de certaines phrases tirées du film. Je présenterai donc des transcriptions phonétiques et des captations de son lorsque la qualité le permet. En dépit de l'absence de musique, peu de passages sont exploitables à cause des bruits de fond et des voix souvent étouffées (voir le spectrogramme en 1.10).

Le réalisateur Lee Daniels a travaillé tout particulièrement sur la voix de Gabourey Sidibe. Il lui a demandé de prendre un accent fort et de changer la tessiture de sa voix : « 'I don't have what we call the blackccent,' she says. 'But I knew that Precious would. And Lee thought my voice should be deeper. It was his idea of what a girl my size would sound like, so he would yell at me, "As deep as you can go, Gabby ! Deeper !" I'd be hoarse at the end of the day.' » (Garratt, 2010). Le travail sur la voix et l'accent va donc vers la caractérisation sociétale de Precious, ce qui est logique compte tenu de ses

31. La liste des acteurs et de leur rôle, ainsi que leur lieu de naissance est indiqué en annexes.

écrits. L'accent et les marqueurs sociolinguistiques des protagonistes est le sujet de la section suivante.

1.2.3.2 Homogénéité des parlers

L'exploration des graphies alternatives et de leur portage à l'écran montre plusieurs tendances contradictoires : ces graphies et les paroles qui doivent les représenter doivent être à la fois intelligibles et distincts, tout en étant spécifiques. Tout en s'éloignant de la norme, ils doivent représenter plusieurs groupes (accent jamaïcain, diverses strates d'AAV, accent hispanique…). Les graphies alternatives devaient donc refléter une phonologie sous-jacente, modulée. Il est impossible d'être assez exhaustive dans cette étude pour traiter la question de l'ensemble allographique cohérent qui représenterait des variantes sociolinguistiques, mais le problème mérite d'être présenté. Dans cette partie, je vais surtout poser la question de l'homogénéité des parlers.

L'hypothèse pour étudier l'éventuelle homogénéité des parlers est la suivante : Precious s'exprime de la même façon que sa mère, dans un parler fortement teinté d'AAV. Ses camarades de classe s'expriment également en AAV ou avec des accents forts. Blu Rain et les autres professeurs s'expriment mieux, dans un anglais américain proche du GA.

C'est effectivement le cas, avec quelques nuances. La plus importante, et celle qui déroge le plus au livre, à cause de la forme est le travail vocal sur le personnage de Precious. Elle a trois voix distinctes dans le film, qui correspondent à ses trois différents registres d'apparence dans le film. La voix de narration externe du film (Precious-Narratrice) correspond au style général de l'oeuvre littéraire, son accent est beaucoup plus marqué dans les dialogues du film (Precious-Personnage), et enfin, lors de ses épisodes d'absence (lors de la scène du viol par son père, notamment), elle s'imagine être une star du cinéma, ou une chanteuse (Precious-Star) et adapte son discours à ce statut fantasmé. La prosodie de Precious-Star est beaucoup plus lente, et correspond à une élocution exagérée. Ces trois styles sont récapitulés dans le tableau 1.5.

Type de voix	Caractéristiques	Exemple dans le livre	Exemple dans le film
Personnage	Rapide, accent marqué	« I jus' say, "My mother is busy" » p.9	/ma ˈmʌvə ˈbɪzɪ/
Narratrice	Lent, accent léger	« There has always been something wrong wif the tesses. » p.30	/ðæz ˈɔːlweɪz ˈsʌmn̩ ˈwrɒŋ widɪz ˈtesɪz/
Star	Lent et exagéré. Accent standard	—-	/aɪ ˈθɒt ɪ wəz ˈfəæbjʊləs/

Tab. 1.5 – Récapitulation des phonostyles de Precious

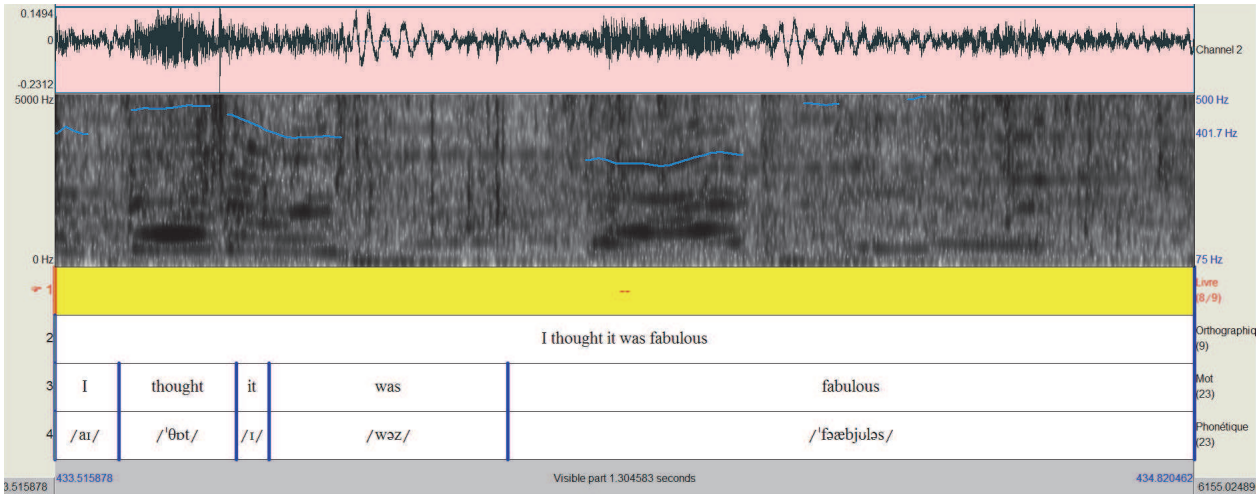


Illustration 1.10 – Extrait de Precious où la protagoniste utilise sa voix de star

Dans l'extrait présenté dans l'illustration 1.10, le mot *fabulous* est exagérément prononcé, et la première syllabe, qui porte l'accent primaire, est allongée dans une imitation prosodique indexicale d'une classe sociale supérieure. L'extraction sous Praat n'est pas très heureuse, parce que les rares exemples de Precious-Star ont toujours beaucoup de bruit de fond (musique très forte, brouhaha et crépitement des flashs d'appareils photo). Cette prosodie est marginale, et aucun personnage du film ne parle de cette façon. Lorsque Precious-Personnage s'exprime, les marques phonétiques présentes dans les graphies alternatives du livre sont transposées dans son parler : elle prononce « tests » [ˈtesɪz], les fricatives dentales sont souvent antériorisées (/ˈmʌvə/ pour /ˈmʌðə/), ce qui correspond à la graphie <muver> dans le livre. Les métathèses trouvées dans le livre, notamment le verbe *ask* graphié <ax> est confirmée dans le film : la prononciation d'*asking* est /ˈæksɪn/, et, au passage, l'on peut constater la dévélarisation de /ŋ/ . Comme l'on pouvait s'y attendre, le groupe consonantique /sk/ est particulièrement sensible à

la métathèse : Precious prononce également escape /eksˈkeɪp/, avec un rétablissement de /k/ en début de syllabe portant l'accent primaire. Le mot est absent de *Push*, mais cette prononciation est peu étonnante lorsque l'on considère la graphie alternative d'*escape*, <xscape>, qui la rappelle et la justifie. Sur internet, la graphie <excape> apparaît 17 300 000 fois, ce qui représente 13% des 134 000 000 de résultats pour la graphie standard [32]. Toutes ces observations tendent vers la constatation que les observations de Sapphire traduites dans son texte dans ces graphies alternatives sont proches de la réalité prononcée par les personnages de *Precious*. Même les hapax qui semblent plus s'éloigner des observations sur la phonologie de l'AAV, comme la graphie <Down Sinder> pour *Down syndrome*, sont attestés dans le film. Precious prononce ces deux mots /ˌdaəʊn ˈsɪndər/, avec un allongement de la diphtongue de « down », et une scission du groupe consonnantique /dr/ avec monophtongaison de la dernière diphtongue du mot, et chute du /m/ final. Internet ne semble pas avoir d'occurrence de <syndrome> écrit <sinder>, certainement à cause de l'homophonie avec « cinder » [33] qui force la regraphie, en ajoutant aux recherches pour <sinder> celles pour <cinder>. Cette prononciation exceptionnelle de Precious, dans le livre et le film, semble due aux légers problèmes de surdité de l'héroïne, desquels on se rend compte lorsqu'elle prononce *incest* [ˈɪnsekt] (p.123-124).

Les particularités grammaticales de l'AAV comme la double négation ou l'absence de la copule be sont fréquents dans le discours de Precious et de Mary : « School ain't gonna help none » ; « don't nobody want you » ; « something gonna happen » illustrent ces phénomènes, surtout dans la scène de dispute entre Precious et sa mère dans le premier tiers du film. D'autres personnages, comme Blu Rain, ne montrent que des signes grammaticaux d'AAV :« You in ABE ? », dans une de ses premières prises de parole à l'écran. Ce trait est exceptionnel dans le parler de Blu Rain, et l'on pourrait l'interpréter comme la prise de contact avec Precious, et l'adaptation au registre de ses élèves, dans une demande d'information ponctuelle, non reliée à ce qu'elle fait en cours.

D'autres n'ont pas d'accent perceptible et il s'agit surtout de personnages secondaires, comme les travailleurs sociaux, dont Miss Weiss, qui représente, comme Blu Rain, la société qui ne cautionne pas

32. La graphie <excape>, qui aurait pu être un candidat graphique pour cette prononciation ne représente que 0.6% des résultats, avec 930 000 occurrences. Recherches effectuées le 4 juin 2014 sur *Google*.

33. « S/Cinder » apparaît également dans la caractérisation de la robe de boas, sur des forums spécialisés, ce qui trompe les résultats de la recherche « have/s a s/cinder ».

les abus. En revanche, les jeunes personnages comme Precious et ses camarades de cours, et la mère de Precious, ont des accents marqués. Joann, l'une des camarades de Precious, est une exception : son accent AAV peu marqué coïncide avec son ambition de « devenir quelqu'un », qu'elle précise dès sa première apparition à l'écran. Ce découpage accentuel est donc une façon extrêmement simple et rapide d'identifier une classe sociale et/ou un niveau d'éducation ou de misère sociale pour les spectateurs. J'ai sélectionné certains traits phonologiques s'éloignant le plus de l'anglais américain standard (Switzer, Brasseau & Youngblood, 2010) :

— Smoothing de voyelles successives (hairrun pour « heroin »),

— Changement de la prononciation des fricatives dentales (muver (/ˈmʌvə/)pour « mother »),

— Réduction des groupes consonantiques en position coda (*chile* pour « child »),

— Métathèses (*ax* pour « ask »),

— Dévélarisation de /ŋ/ en /n/ en position coda (*feelins* pour « feelings »),

— Chute des /t/ après une contraction (*don* pour « don't »)

Si l'on prend ces critères comme base de gradation d'accents, avec certains indices grammaticaux, comme l'absence de be-copule, les personnages se répartissent donc en trois catégories (cf tableau 1.6).

Les parlers sont donc relativement homogènes : l'utilisation d'une langue minoritaire est plus visible lorsque Rhonda utilise du jamaïcain pour s'exprimer (« Shut the bumba ras clot », '*shut your motherfucking mouth*') et combine deux expressions au sens similaire, *bumba clot* et *ras clot* qui si-

Type	Caractéristiques	Personnages
Accent afro-américain	Tous les traits listés	Precious-Personnage, Precious-Narratrice, Mary
Accent « étranger »	Diphtongaison des V accentuées	jamaïcain : Rhonda ; hispanique : Rita
Accent GA	—	Precious-Star, Miss Weiss, Cornrows, Blu Rain

Tab. 1.6 – Répartition des parlers des personnages de Precious

gnifient « fuck » [34], écrites au tableau de la classe de Blu Rain, comme pour renforcer l'indication d'un accent et d'une langue minoritaire étrangère.

Les dialogues sont marqués sociolinguistiquement, et les parlers se regroupent autour de trois pôles phonostylistiques. Les dialogues oraux sont complétés par une piste de sous-titrage pour certains spectateurs. La préservation éventuelle de la dimension sociophonétique dans ces sous-titres, qui sont le seul moyen d'accéder au film pour certains spectateurs, est le sujet de la partie suivante.

1.2.4 Traits communs des sous-titres et du texte

Les sous-titres doivent correspondre à plusieurs critères : lisibilité et discrétion (Golding & A. Bailey, 1999 ; Gambier, 1996), pour que la compréhension déserve le moins possible la visibilité de l'image. Le sous-titre est cantonné à la marge visuelle (en bas de l'écran, où il se passe le moins de choses possible). Les normes régissant l'apparition des sous-titres (délai minimum d'apparition, laps de temps entre deux sous-titres, longueur…) n'est pas soumise à une convention générale. Je n'examinerai que les sous-titres officiels du DVD dans cette sous-partie (Daniels et al., 2011).

Le film adapte le discours à un public plus large, qui doit s'approprier le contenu en moins de temps que le lecteur de roman : la lecture immersive, où l'on apprend progressivement à comprendre le code ne s'adapte pas à la lecture des sous-titres. Ils ne sont utilisés que dans des cas bien précis. Le sous-titrage anglais d'un film en anglais est surtout fait pour les sourds et malentendants et non pour les spectateurs anglophones lambdas [35]. L'accent afro-américain vernaculaire, ainsi que les autres accents (comme l'accent jamaïcain de Rhonda) sont bien connus des spectateurs américains, tout comme les lexèmes argotiques, comme « eighty-six » (‘*se débarrasser de quelque chose*’) (Gove & Merriam-Webster, 1993).

Les sous-titres du DVD sont rédigés pour les sourds et malentendants. En conséquence, ils sont épurés et ne contiennent pas de graphies alternatives. En revanche, les caractéristiques grammaticales propres à l'AAV, comme l'utilisation de « done » résultatif, sont conservées.

34. De « blood cloth » ou « rag cloth » (l'interprétation « blood clot » est possible), ces deux expressions désignent originellement le tissu avec lequel les saignements menstruels étaient essuyés (Sullivan, 2009).
35. Un autre type de spectateur ayant besoin de sous-titres en langue originale pourrait être l'apprenant dont l'anglais est la L2.

La valeur iconique de la graphie de Precious n'est conservée que dans le titre du film, (illustration 1.11), avec la double lecture d'un manuscrit maladroit en haut, et d'une orthographie dans une police plus conventionnelle pour la correction. Le titre du film, ou, au moins, le mot le plus important, Precious, n'est pas affecté par cette graphie alternative. Paradoxalement, il n'y a pas d'image du film qui montre Precious en train d'écrire, ou ses productions écrites dans le film, qui constituent pourtant une proportion importante du livre (7.5% est rédigé par Precious dans son carnet, sans compter le jeu de transcription en anglais standard de Blu Rain, comme exprimé dans le tableau 1.1. Les illustrations 1.12 et 1.13 montrent la conservation d'éléments grammaticaux afro-américains. Ils démontrent tous les deux qu'il est plus facile de comprendre un texte bien écrit, mais à la grammaire non standard, que l'inverse. Toutes les pistes de sous-titres trouvées sur internet ou les DVD n'incluent pas les terminaisons verbales des participes présents en <-in'> au lieu de <-ing>. Ce n'est pas le parti-pris des sous-titres officiels non plus, qui se contentent de quelques « gonna » ou « wanna ». Y figurent également des formes grammaticales comme le « done » résultatif, ou des utilisations de *be* alternatives (forme à la troisième personne du singulier pour un pronom à la deuxième personne, comme

Illustration 1.11 – Titre du film

dans l'illustration 1.12). Les graphies sont standardisées, et l'on peut imaginer des locuteurs qui

Illustration 1.12 – Extrait de Precious avec sous-titres (00 :41 :30)

Illustration 1.13 – Extrait de Precious avec sous-titres (00 :49 :00)

s'appuient sur les sous-titres pour comprendre l'accent afro-américain, surtout pour comprendre la grammaire alternative et s'assurer ainsi qu'ils ont bien compris ce qui avait été prononcé par les personnages. Les sous-titres sont écrits dans un anglais standard, et la vraie co-incidence son et graphie est en fait entre le livre et la transposition orale. Elle laisse le sous-titrage, normé, de côté.

Cette première approche du corpus secondaire a permis de définir des principes graphiques et des bases pour l'étude de la comparaison graphophonématique. *Push*, avec son narrateur interne, est ce-

pendant bien différent de *Trainspotting* avec ses multiples personnages, tous « scripteurs » de leur histoire. Le second volet de cette partie servira donc à explorer la variation graphique inter-scripteurs d'une même langue. minoritaire.

1.3 *Trainspotting*

La seconde moitié du corpus secondaire est plus complexe, à cause de la longueur du livre, et de la nature des lexèmes utilisés, qui tiennent plus du scots standard que de graphies alternatives reflétant la phonologie ou l'acquisition de la langue écrite. Dans *Trainspotting*, il s'agit plus souvent d'un jeu stylistique complexe mêlant lexèmes en scots, argot, et même imitations de marques phonologiques. L'œuvre est plus longue que *Push*, et fait appel à plus de narrateurs intradiégétiques, ce qui complique la tâche. De plus, je suis confrontée au même problème de graphies fluctuantes ou évolutives : dans Push, il s'agissait d'acquérir des compétences graphiques et Precious écrivait plusieurs variations intrascriptoriales. Par exemple, une graphie alternative peut succéder à une orthographie, comme « show » qui apparaît p. 38, écrit <sho> plus loin, p. 101. C'est une graphie régressive, qui ne correspond pas au mouvement progressif du livre. Dans *Trainspotting*, les allographies sont dues au changement de narrateur, et il n'est y est pas question de progression, mais de style personnel traduit par la graphie utilisée.

Les études de cette partie du corpus secondaire suivront peu ou prou le même plan que celles de *Push* et *Precious*. En revanche, le prisme d'acquisition de formes graphiques est tout à fait écarté, puisqu'il ne s'agit nullement d'un récit censé être écrit par des personnages, mais plus d'une prise de parole stylistiquement chargée en régionalismes et en langue minoritaire. Grâce aux jalons posés dans la partie précédente, cette partie abordera les graphies en scots, puis les usages langagiers conservés dans le film, pour passer à une courte étude des sous-titres et des scottismes qui y sont conservés.

1.3.1 Analyse multimodale

Pour différencier le livre du film, éponyme, le film (Boyle et al., 2004) sera appelé *Trainspotting* (Film) (TR-F) et le livre, (Welsh, 1993/2011) *Trainspotting* (Livre) (TR-L) en cas d'ambiguïté.

Les deux œuvres n'ont que trois ans d'écart : le livre est sorti en 1993, le film, en 1996, et tous deux se situent principalement à Édimbourg. Ce film n'est pas la première adaptation de *Trainspotting*, qui a d'abord été porté au théâtre en 1994. Ewen Bremner (Spud dans le film) a d'ailleurs joué le rôle de Mark Renton, et les trois formats (livre, pièce et film) ont rencontré un succès certain, voire gagné un statut d'œuvre culte. En Grande-Bretagne, plus d'un million d'exemplaires ont été vendus (chiffre datant de 2012), et le film a généré 24 millions de dollars au box office à l'échelle mondiale (dont 16 au Royaume-Uni). Tous les livres d'Irvine Welsh contiennent un anglais fortement écossisé, voire scottisé. L'auteur s'exprime d'ailleurs avec un accent écossais marqué (Macdonald, 2010), bien qu'il ne réside plus à Édimbourg (Welsh, 2007).

1.3.1.1 Structure narrative du livre

Trainspotting est un livre écrit par Irvine Welsh. Écrit en discours direct par la plupart des protagonistes de l'histoire, il présente des formes caractéristiques écossaises fortes (notamment par les personnages Begbie et Spud). Dans le film tiré de ce livre, les personnages principaux utilisent également des scottismes et parlent un anglais très « écossisé ».

La proportion de dialogues n'est pas pertinente dans le livre, étant donné que la plupart des nouvelles composant l'œuvre sont des narrations homodiégétiques. Welsh y imprime des caractéristiques oralisantes dans son écrit, pour créer des pratiques qui servent de signature linguistique aux protagonistes. Par exemple, Spud réfère aux humains comme étant des « cats » (avec des dérivés, comme « catboy »). Pour Begbie, les graphies sont scottisées, et il s'exprime avec beaucoup plus de grossièretés que les autres personnages : « Obviously, ah wis cairryin ma fuckin chib n aw. Too fuckin right. » (‘*Obviously, I was carrying my fucking knife and all. Too fucking right.*’ (p. 85). Les phénomènes d'imitation dans *Trainspotting*sont également présents : certains personnages s'imitent les uns

les autres, ou passent à des codes qu'ils ne maîtrisent pas forcément, pour se faire passer pour un autre. Ainsi, Renton adopte un accent Cockney pour ne pas être importuné « You're orlroight Jock, ah tell um, in ma best Cockney. » (V, 1). Simon-Sick Boy imite souvent Sean Connery, avec un chuintement systématique des fricatives, traduite par la graphie <Shimon> pour son prénom, par exemple.

L'organisation globale de l'œuvre est résumée dans le tableau 1.7. Quand je citerai le livre, j'appellerai d'abord la partie dont l'extrait est tiré en chiffre romain, et le numéro du chapitre en chiffre arabe (par exemple, Junk Dilemmas No. 63 est I, 2).

Partie	Nom du chapitre	Narrateur
Kicking	The Skag Boys, Jean-Claude Van Damme and Mother Superior	Renton
Kicking	Junk Dilemmas No. 63	Renton
Kicking	The First Day of the Edinburgh Festival	Renton
Kicking	In Overdrive	Sick Boy
Kicking	Growing Up in Public	Diégétique
Kicking	Victory on New Year's Day	Diégétique
Kicking	It Goes Without Saying	Renton
Kicking	Junk Dilemmas No. 64	Renton
Kicking	Her Man	Second Prize
Kicking	Speedy Recruitment	Diégétique ; Renton ; Spud
Relapsing	Scotlad Takes Drugs in Psychic Defence	Tommy
Relapsing	The Glass	Renton
Relapsing	A Disappointment	Begbie
Relapsing	Cock Problems	Renton
Relapsing	Traditional Sunday Breakfast	Davie
Relapsing	Junk Dilemmas No. 65	Renton
Relapsing	Grieving and Mourning in Port Sunshine	Diégétique
Kicking Again	Inner Shitty	Begbie
Kicking Again	Na Na and Other Nazis	Spud
Kicking Again	The First Shag in Ages	Diégétique
Kicking Again	Strolling Through the Meadows	Spud
Blowing it	Courting Disaster	Renton
Blowing it	Junk Dilemmas No. 66	Renton
Blowing it	Deid Drugs	Sick Boy
Blowing it	Searching for the Inner Man	Renton
Blowing it	House Arrest	Renton
Blowing it	Bang to Rites	Renton
Blowing it	Junk Dilemmas No. 67	Renton
Exile	London Crawling	Renton
Exile	Bad Blood	Davie
Exile	There is a Light that Never Goes Out	Diégétique
Exile	Feeling Free	Kelly
Exile	The Elusive Mr Hunt	Diégétique
Home	Easy Money for the Professionals	Spud
Home	A Present	Renton
Home	Memories of Matty	Diégétique
Home	Straight Dilemmas No. 1	Renton
Home	Eating Out	Kelly
Home	Trainspotting at Leith Central Station	Renton
Home	A Leg-Over Situation	Renton
Home	Winter in West Granton	Renton
Home	A Scottish Soldier	Diégétique
Exit	Station to Station	Diégétique

Tab. 1.7 – Composition de Trainspotting

À première vue, aucun des titres des chapitres (à part '*Deid Drugs*') ne contient de mots écossais, et les utilisations d'argot sont très rares comme *skag* ('*drogue*'). Les titres sont donc trompeurs, bien qu'ils puissent être ironiques, surtout après en avoir lu le contenu [36] mais surtout, ces titres ne sont pas révélateurs de ce qu'ils recèlent.

Comme indiqué dans le tableau 1.7, les personnages n'ont pas la même fréquence d'apparition dans la narration, et ne sont pas les scripteurs présumés de la même proportion d'éléments. Le tableau suivant complète le premier, et montre le volume discursif des personnages de *Trainspotting*.

L'étiquette « Autres » n'est attribuée qu'aux productions uniques (personnages n'ayant « contri-

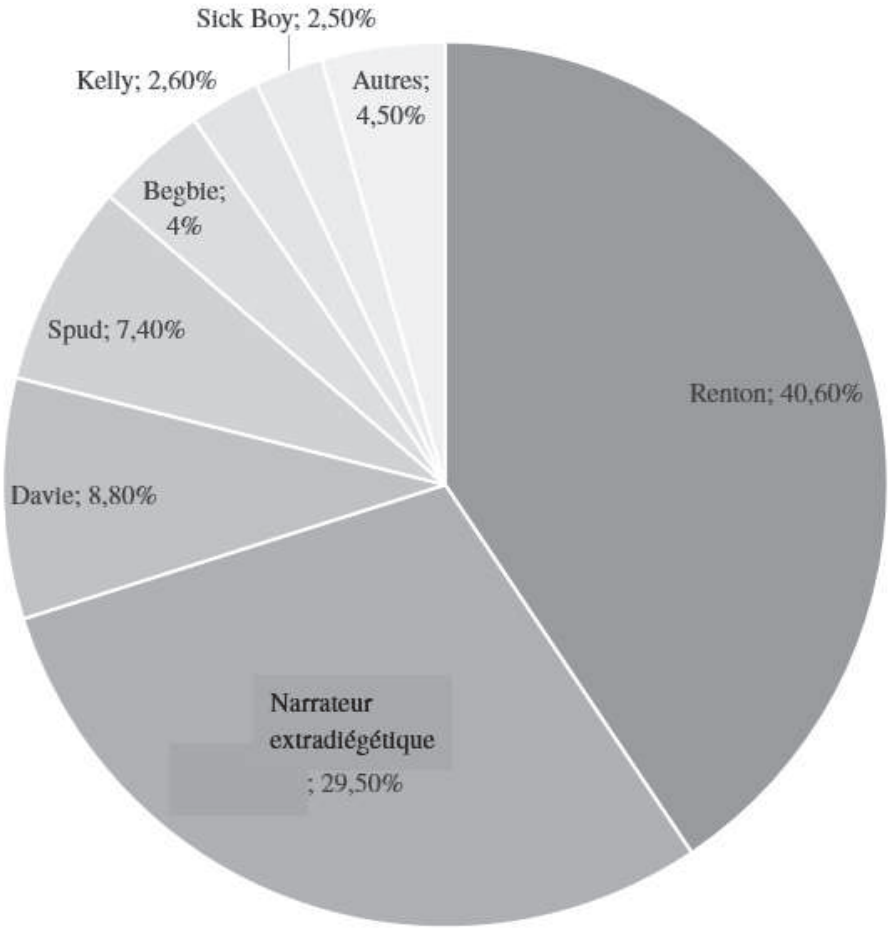


Illustration 1.14 – Proportion de discours suivant les scripteurs

36. Le meilleur exemple étant certainement '*The Elusive Mr Hunt*', qui n'est qu'une description d'une blague téléphonique et sexiste faite dans un bar, où la prononciation de la protagoniste de « Mark Hunt » était homophone avec « my cunt ».

bué » qu'à un seul chapitre, comme Tommy) ou aux chapitres écrits par plusieurs personnages. Les données brutes correspondant au graphique 1.14 sont les suivantes :

Personnage	Nombre de mots	Pourcentage
Renton	42 113	40.6%
Davie	9 214	8.8%
Spud	7 759	7.4%
Autres	4 747	4.5%
Begbie	4 196	4%
Kelly	2 666	2.6%
Sick Boy	2 577	2.6%
Narrateur extradiégétique	30 980	29.5%
Total		100%

Tab. 1.8 – Proportion de mots imputée à chaque personnage

L'impact de certains personnages est donc bien mineur, et les productions de Renton sont les plus importantes. Les parties rédigées par un narrateur externe contiennent tout de même des graphies alternatives, grâce aux dialogues échangés qui réutilisent du scots. En outre, il n'est nulle part mention de garder des traces écrites par les personnages, et ces parties du roman ne sont que des narrations autodiégétiques. La partie concernant le roman *Trainspotting*est divisée entre une étude qualitative des écrits de plusieurs « scripteurs », ainsi que des scottismes et de l'argot de Leith.

1.3.1.1.1 Variation interscripteurs

Les narrations homodiégétiques provoquent des écarts dans la fréquence des mots employés, ce qui est un bon exemple de questions soulevées par un corpus. Renton s'exprime en scots d'Édimbourg, Spud utilise très fréquemment le mot « cat » ou « catboy » pour parler de ses amis. Dans cette partie, j'ai recours à AntConc, un concordancier qui compte et classe les mots par fréquence. Les effets de style sont graphiques ou lexicaux, suivant les personnages, et j'ai choisi de comparer les productions de Renton à celles de Spud [37]. Renton est en quelque sorte l'étalon graphique du scots, puisque c'est le personnage qui apparaît le plus fréquemment, alors que la représentation de Spud est tout à fait unique dans le livre, car ce personnage est dépeint comme ayant de nombreux tics de langage (« It's too hot

37. Tout en gardant à l'esprit que ce livre est une fiction linguistique, et que les variations observées sont en très grande partie à imputer à I. Welsh.

for a fair-skinned punter, likesay, ken ? » (p. 119)), et il s'exprime de manière très personnelle.

J'ai d'abord utilisé le classement par fréquence des graphies alternatives sur toutes les parties attribuées à Spud (III, 2 et 4 et VI, 1). Ces deux jeux de données sont sensiblement identiques dans leur structure narrative (présence de discours direct et récit au passé…). Je les ai ensuite comparées aux données extraites d'un corpus de taille sensiblement identique à la sous-partie du livre écrite du point de vue de Renton (pour correspondre aux 7 759 mots du corpus Spud, j'ai pris les parties I, 1 et I,3, qui totalisent 8 118 mots).

Les 25 graphies alternatives les plus fréquentes dans les deux échantillons sont présentées dans le tableau 1.9. Je n'ai conservé que les scottismes purs, les formes élidées passant mal le concordancier d'AntConc et sont moins pertinentes à l'analyse stylistique et textométrique que je mène. Par exemple, le mot « fah-kin » n'a tout simplement pas été traité par AntConc, et c'est la même chose pour les mots qui contiennent une apostrophe.

Les mots grammaticaux sont communs aux deux personnages, et dans des proportions similaires. On peut mettre à part les marqueurs grammaticaux, prépositions et pronoms (les deux personnages utilisent à peu près également *ah*, *tae*, *ay*, *ye*, *wi*, *ma*, *aw*, *oan*, *aboot*, *oot*, *thit* et *fir*), et les verbes les plus courants, comme *git* et *goat* (get), *wis* (be), *hud* (have) et *dinnae* (do[n't])…. Une fois les mots grammaticaux écartés, les mots les plus fréquents sont en fait des marqueurs identitaires, non seulement du scots (amplement représenté dans l'œuvre grâce au remplacement de mots grammaticaux fréquents) mais de son milieu social. Les insultes comme *fuckin*, très fréquent (4ème et 6ème place) et *shite* (seulement Renton, 19ème position) marquent l'appartenance à une classe où elles peuvent être fréquemment prononcées sans que cela ne soit problématique, dans des phrases comme « This swan though, is fuckin beautiful, likes. » (Spud, III, 2) ou « Custom-fuckin-designed fir your needs. » (Renton, I, 3). *Fuckin* sert souvent d'intensifieur, et perd sa valeur tabou à force d'être employé (ce qui ne veut pas dire qu'il soit tout à fait déparé de sa valeur insulte, comme le montre la citation « You shut yir fuckin mouth ! Fuckin radge. Ah'll fuckin tell you whin tae speak. Just shut yir fuckin erse. You dinnae like the company, you kin git tae fuck. End ay fuckin story. » (I, 3), où l'accumulation ne l'affaiblit pas, mais montre simplement l'énervement du personnage. Il nous reste donc peu de marqueurs

Ordre	Spud		Renton	
	Lexème	Tokens	Lexème	Tokens
1	Likesay	175	Ah	290
2	Ah (I)	172	Tae	202
3	Tae	125	Ay	133
4	Fuckin	93	Ma	125
5	Ken	88	Oan	79
6	Ay (of)	72	Fuckin	74
7	Ye (you)	60	Ye	56
8	Wi	54	Wis	54
9	Wis	52	Wi	44
10	Ma	44	Intae	30
11	Jist	43	Aboot	29
12	Sortay	40	Oot	28
13	Aw	39	Aw	26
14	an	36	Goat	26
15	Wee	36	Fae	24
16	Aboot	31	Fir	24
17	Oot	31	Git	23
18	Sais	30	Sais	21
19	Cat	27	Shite	18
20	Goat (got)	27	Nae	17
21	Thit	23	Bit (but)	16
22	Whae	23	Hud (had)	16
23	Eh	20	Yir	16
24	Git	19	Dinnae	14
25	Fir	18	Oaf	14

Tab. 1.9 – Les 25 lexèmes les plus courants dans les narrations de Renton et de Spud.

indexicaux du discours de Renton et de Spud dans les mots fréquents, mais ils sont assez saillants pour pouvoir en tirer une analyse.

Les hésitations de Spud, en récit ou en dialogue, se repèrent parfaitement dans la fréquence des mots qu'il utilise. Ainsi, « likesay », « sortay » et « eh » (‘*euh*’). Ces mots sont typiques de l'expression orale, et participent à la création du style de Spud. Ils sont complètement absents du corpus Scotster, et de ceux des autres personnages (mis à part les lignes de dialogue de Spud).

Les mots grammaticaux sont des éléments porteurs du style donné à chaque personnage. Renton raconte plus volontiers des histoires qui lui sont arrivées : I, 1 est une ouverture où il rencontre ses amis pour se droguer, et il relate sa première tentative de désintoxication dans I, 3 [38]. La focalisation est plus volontiers interne, ce qui explique la plus grande proportion de pronoms personnels à la pre-

38. Cette scène, et notamment « The Worst public toilet in Scotland » est recréée sans complaisance dans le film.

mière personne comme « ay » ou « ma », que celles du corpus de Spud. Ses récits (au passé), par rapport à ceux de Spud, se traduisent par le plus grand recours à des verbes au prétérit comme « hud » et « dinnae », qui sont absents de la liste de Spud, et une plus grande proportion de « goat » et même « wis ».

Certaines de ces graphies sont inexistantes sur SC, comme « goat » mais la plupart des graphies de Welsh sont attestées dans le corpus SC. La différence majeure entre ces deux corpus est la polygraphie de certains mots : si « aboot » est la seule graphie alternative attestée sur SC, les autres mots sont graphiés différemment par les utilisateurs divers du forum. Ainsi, *all* est graphié <aw> dans le corpus TR-L, mais sur SC, on le trouve sous les formes <aa> et <au>, suivant l'utilisateur. Le cas de la graphie <wi> est encore plus unique, puisqu'il signifie *with* dans TR-L, mais qu'il est utilisé indifféremment pour *we* et *with*, le dernier étant parfois graphié <wie>. Les limites de l'invention graphie de Welsh sont certainement dictées par des questions de confort d'écriture et de lecture grâce à une graphie alternative unique, ce qui n'empêche pas d'exceptionnels problèmes d'homographies.

Ce corpus est un peu semblable à celui des forums internet, où les locuteurs sont tous différents et utilisent des expressions personnelles, mélange de style et de graphies particulières. Ces expressions peuvent très bien ne pas être très courantes à l'échelle du corpus, et servent de signature à certains même s'ils ne les emploient pas à outrance. Au fil de la lecture, il devient de plus en plus facile de lire les entrées et de distinguer entre les personnages, tout comme c'est le cas sur les forums. Aucune étude de réception ne pouvant être menée, cette constatation se borne à être subjective. Ces constatations sur les productions de Renton et de Spud sont évidemment très microscopiques parce qu'il n'y a que deux personnages comparés ; sur les forums, les échanges et le nombre de participants rend ce type d'étude difficile. Elles sont alors conduites avec d'autres outils, comme la thèse de S. Nowson, qui analyse les blogs de plusieurs internautes pour en comparer les traits psychologiques différents. Il dispose d'un corpus stable de productions d'un seul internaute par blog, et donc d'une série de sous-corpus non-conversationnels où il peut s'assurer que les internautes ne vont pas s'influencer mutuellement (Nowson, 2006).

Les variations inter-scripteurs fictionnels se rejoignent autour des scottismes, qui sont la marque

de fabrique d'I. Welsh. C'est le sujet des deux parties suivantes, qui décrivent les utilisations des scottismes généraux et l'argot de Leith, mobilisé dans l'oeuvre.

1.3.1.1.2 Scottismes

Les scottismes peuvent poser problème parce que certains mots correspondent à des graphies existantes en anglais standard. Ce problème d'homographie est aussi présent dans le corpus *Push*, avec la graphie <four> pour *for*, par exemple. Même en ne sélectionnant que les entrées de Renton, le narrateur le plus prolifique, il subsiste un problème de stabilité des usages de Welsh, qui attribue plusieurs graphies d'un mot au même personnage : « Deek » et « deek » désignent respectivement *Derek* et *deek*, '*regarder*' en scots.

Les scottismes reconnus sont listés dans certains glossaires en ligne, comme (« A dictionary for the dialects in "Trainspotting" », 2007) qui contient 133 entrées, ou d'autres articles de presse qui citent des paragraphes entiers pour faire comprendre à leurs lecteurs la difficulté de rentrer dans *Trainspotting*. Le style de chaque personnage, qui a été exploré dans la section précédente permet de les reconnaître, mais le livre dans son ensemble est un recueil de scottisme.

Les graphies qui sont les plus stables sont souvent les graphies attestées, rentrées dans la langue commune dans une certaine mesure, et très indexicales. « Ken » et surtout « wee » sont des graphies courantes en ligne. Elles ont en outre l'avantage d'une certaine stabilité graphique. L'indexicalité de ces deux mots est très grande, et ils sont très bien représentés en ligne (84 millions de tokens pour « wee »), et signalent immédiatement le registre et l'Écosse pour les lecteurs. The Scotsman a intitulé un de ses tests en ligne « Scots language, if you ken whit ah mean ? » (Scotsman, 2006). La graphie du scots attestée n'est pas unique, et les variantes graphiques abondent dans le *Dictionary of the Scots Language* (DSL). Ainsi, le mot *glaikit* (' *idiot*'), qui est la forme de l'entrée dans le dictionnaire, est attesté sous les formes <glaiket>, <glaiked>, <glaykit>, <glaikyt> et <glakyt>. Dans *Trainspotting*, il est utilisé à deux reprises, sous les formes <glaikit> et <glakit>. Cette dernière est logique, bien que non attestée dans le DSL.

L'amplitude allographique de ce mot est récapitulée dans le tableau suivant (1.10). Il ne présente

que les candidats graphiques et leur position, moins que les contraintes de ces graphies (par exemple <d> n'apparaît qu'après <e>). :

GL	A	Y I –	K	I Y E	T D

Tab. 1.10 – Amplitude allographique de <glaikit>

Le choix d'I. Welsh est logique, compte tenu des graphies variées du dictionnaire. Les consonnes sont encore une fois un peu plus stables que les voyelles, et leur concurrent est phonologiquement très proche, /t/ et /d/ n'étant distingués que par leur trait [± sonore]. Les scottismes se concentrent autour des transcriptions de prononciation, non seulement écossaises, mais aussi profondément orales : dans l'extrait « how ah love ya, how ah love yah, my dear old Swanney… » (I, 1), la graphie de *you* fluctue. Il n'y a que deux occurrences de <yah> dans tout le corpus TR-L, qui semble fortement oral et influencé par la répétition de la phrase (prononcée lors d'une séance de prise de drogues). La graphie la plus courante pour *you* est <ye> (546 occurrences), et correspond au pronom écossais prononcé /jɪ/ (source : DSL). L'auteur essaye donc de coller au plus près aux graphies oralisantes et aux scottismes connus.

Certaines graphies sont donc évidemment communes à SC et *Trainspotting*. Leur exploitation est un peu délicate, parce que quelques graphies de *Trainspotting* correspondent à des homographes en ligne : <alane> (‘*alone*’dans *Trainspotting*) est le pseudonyme d'une membre de SC, ce qui fausse les résultats. Une autre variante proposée pour le même mot dans *Trainspotting*est <alain>, mais les membres canadiens de SC parlent de Val Alain au Québec, et aucun résultat ne renvoie vers le mot *alone*. La très faible fréquence des occurrences des deux graphies (40 et 41, respectivement) prouve que les mots ne sont pas considérés comme les membres de SC comme des candidats graphiques possibles. D'ailleurs, les graphies de *Trainspotting*attestées dans SC sont souvent les plus saillantes en scots, comme <aw> (‘*all*’), <ken> (‘*know*’) ou la graphie des mots grammaticaux en général (<oot>, ‘*out*’, ou <jist>, ‘*just*’), sauf lorsqu'ils sont homographes à des mots existants. La graphie <is> pour *as* n'est pas répertoriée dans SC, certainement parce que les membres veulent éviter des candidats

allographiques déjà existants.

Les Écossais fictionnels ou réels utilisent les mêmes graphies parce qu'elles sont fréquentes, connues, et appartiennent à ce qu'ils connaissent (les poèmes de R. Burns utilisent la graphie <ken>, par exemple (Burns, 1867)). Elles représentent un standard du scots actualisé dans les deux corpus écrits TR-L et SC. Il existe cependant un sous-ensemble de scottismes dans l'œuvre : les utilisations de l'argot de Leith et d'Édimbourg. C'est le sujet de la partie suivante, qui clôt les études graphiques sur TR-L.

1.3.1.1.3 Argot d'Édimbourg et de Leith

La plupart des personnages utilisent non seulement le scots, mais l'argot d'Édimbourg, et plus précisément, de la région de Leith (au nord de la ville, vers le port). L'illustration 1.15 montre où elle se situe précisément.

Illustration 1.15 – Localisation de Leith à Édimbourg (d'après Skilling, 2012 et Google Maps)

Le livre a permis en partie la compilation d'un dictionnaire d'argot de Leith en ligne (Mountford, 2005) en tant que seule source publiée sur lequel il s'appuie, et ce dictionnaire cite le livre et le script du film en exemples d'argot de Leith. Malheureusement, les entrées de ce dictionnaire ne sont

pas consultables comme dans un dictionnaire traditionnel, et l'on ne peut y accéder qu'en recherchant des mots spécifiques. Il ne reste que la possibilité de chercher les mots, et de voir s'ils retournent des résultats, ou leur contrepartie en anglais dans la case « Definition ». Une recherche des citations de *Trainspotting*ne renvoie que 65 résultats dans ce dictionnaire, soit 7.5% du corpus *Trainspotting*. C'est une ébauche de dictionnaire élaboré pour un projet de l'Université de Lancaster, et il n'est plus mis à jour depuis 2005, ce qui explique les limitations décrites.

Les éléments du dictionnaire de Leith ne concernent ni les mots grammaticaux, ni les formes argotiques d'autres variétés. Le vocabulaire relevé dans le dictionnaire d'argot se concentre particulièrement autour de mots ayant trait à la violence (*swedge*, *pagger*, '*combat*' *chore*, '*voler*') ou aux insultes (*cunt*, *hoor*, apparenté à « whore », *schemie* '*personne bénéficiant des aides sociales, RMiste*')…Le dictionnaire offre des descriptions sociologiques des utilisations de certains lexèmes, comme pour *cunt* :

> In Leith 'that cunt' can be used for 'that guy,' […] if the emphasis on both words is the same, no offense is intended, and is indeed more likely to be used amongst friends. The usage of 'cunt' in Leith can be likened to the usage of 'cat' in jazz culture (?). If the word 'cunt' is emphasised, however, the meaning is derogatory.

Ces observations sont relativement logiques, étant donné que la récupération d'insultes pour l'utiliser entre amis est connue (Allan & Burridge, 2006). La phrase « Happy New Year, ya wee cunt ! » (I, 6) est presque l'un des seuls exemples d'utilisation positive du mot, souvent modifié par l'intensifieur « doss ». Le dictionnaire ne se veut pas glossaire de la langue de *Trainspotting*, étant donné, entre autres, que l'argot de Leith a beaucoup évolué entre la date de l'écriture du livre et du dictionnaire (environ 10 ans d'écart). Certains mots sont tombés en désuétude, ou n'ont plus le même sens, comme « tidy », qui signifie '*fort (au combat)*' dans le livre, mais '*séduisant*' pour le groupe étudié par Mountford.

Enfin, l'auteure du dictionnaire reconnaît la prononciation alternative des lexèmes, mais les impute non au lieu, mais à l'accent écossais en général : « Many words in Scots like 'aheid' for 'ahead,' 'aff' for 'off' which represent words in the English language pronounced with a Scottish accent and not everyone's accent is the same, even in Leith. ». Cette phrase peu engagée prouve l'argument avancé dans la section précédente : I. Welsh a adapté ses amplitudes allographiques aux prononciations des

personnages du roman. En revanche, l'accent local et les choix lexicaux dessinent la carte d'un certain scots de Leith, teinté d'activités illégales, comme tout bon argot urbain. Sans surprise, cet argot ne se retrouve pas, ou très marginalement, dans le corpus SC, avec des mots comme « bevvy » … Mais l'utilisateur (SCP7) qui l'emploie vient de Paisley (vers Glasgow) [39]. Un lexème simple ne permet pas de donner une indication géographique précise, mais l'ensemble des mots spécifiques à *Trainspotting*le situe donc dans un contexte socioculturel précis.

La transposition de l'œuvre en film oblige de nombreux arrangements avec le parlécrit du roman. C'est le propos de la prochaine partie, qui délimitera les caractéristiques du scots qui ont pu être préservées. La prononciation réelle de ces graphies alternatives, transcriptions du parler, est aussi abordée dans cette partie.

1.3.1.2 Film

L'extraction des dialogues à partir d'un film est difficile, puisque ils sont toujours sur des fonds sonores (bruits d'ambiance, ou musique) dans TR-F. La transcription doit être réalisée « à la main », sans l'aide d'outils comme *Winpitch* ou *Praat* pour affiner les analyses. Dans le film, l'accent écossais des personnages est peu marqué, par volonté de D. Boyle, qui a voulu rendre son film accessible au public américain, et a donc réenregistré les vingt premières minutes de dialogue audio (Jenkins, 1996), pour laisser aux spectateurs le temps de s'adapter, avant que l'accent ne devienne plus fort. Une piste envisagée, puis abandonnée consistait à sous-titrer les dialogues, mais D. Boyle a estimé que le film perdrait de sa force, et que la *linguistique-fiction* (Sorlin, 2010) reposait sur le scots. Il évoque le problème de la langue dans plusieurs interviews : « The biggest problem was language because the vernacular acrobatics in the book can't be done on film, it would have to be dubbed » (Russell Powell, 1996). La langue parlée pose donc un réel problème d'intelligibilité, tout comme la langue écrite du livre : l'une des éditions sorties pour le public américain contient d'ailleurs un glossaire, ce qui rappelle certaines éditions d'*Orange mécanique*, qui a posé des problèmes similaires à cause du

39. Un autre membre, de Balloch (également au sud-ouest) utilise aussi ce mot.

nadsat inventé par Burgess [40]. Ces œuvres mobilisent un vocabulaire en partie cryptique, pour que le lecteur lambda ou non-membre de la catégorie sociologique des personnages fournisse un effort pour en comprendre le contenu. Le temps alloué à la lecture ne ressemble pas au temps filmique, où le spectateur n'a pas le loisir de revenir en arrière. Les informations doivent être données rapidement, et l'exploration linguistique est incompatible avec ce média. L'hétéroglossie des personnages est donc très édulcorée, et seuls certains mots demeurent, comme « bird », '*girl*', « shite », « skag », '*heroin*' .

Les mots les plus susceptibles d'être compris rapidement, suivant le contexte, sont les candidats parfaits pour être conservés. En outre, les scottismes donnent des détails discrets et « couleur locale » pour les spectateurs non-Écossais. C'est la raison pour laquelle des mots comme « aye » ou les négations en <-nae> sont utilisés dans l'ensemble du film.

Les seules traces d'un accent écossais sont les /r/ d'E. McGregor (Renton) qui sont flappés, comme ceux de J. Lee Miller (Sick Boy) [41]. P. Mullan (Swanney) a l'accent le plus marqué de tous, mais c'est un personnage mineur, qui a peu de lignes de dialogue, et cet accent est plus caractéristique du nord de l'Écosse que celui des autres personnages, P. Mullan étant originaire de l'Aberdeenshire. Phonologiquement parlant, il est difficile de ne pas limiter les observations à une simple description, la captation des sons étant rendue difficile, à cause de musiques ou d'effets d'échos. Je me bornerai donc à travailler sur des extraits très courts, au son le plus propre possible.

La captation 1.16 est de mauvaise qualité (à cause d'un écho marqué), mais elle illustre plusieurs problèmes de cette étude. J'ai découpé chaque extrait avec le passage du livre le cas échéant, le sous-titre à l'écran, le découpage mot à mot, et une transcription phonologique. Le livre (premier *tier*), est beaucoup plus fourni que le film, sur les maniérismes linguistiques de Spud. Les sous-titres respectent ici ce qui est dit, avec des changements mineurs : l'auxiliaire est rétabli, ce qui n'est pas étonnant. La différence entre le livre et les sous-titres est minimale, si l'on fait exception des reprises de Spud.

No Ah went tae […] Craigy.
No I've been to Craigy.

40. L'adaptation par S. Kubrick l'a déçu, en partie à cause de l'abandon de l'essentiel du vocabulaire.
41. Une liste complète du casting (tableau A.3), ainsi qu'une carte de la provenance des acteurs du film (carte A.1) est disponible dans les annexes. Après la première référence, seul le nom de l'acteur sera donné dans le corps du texte.

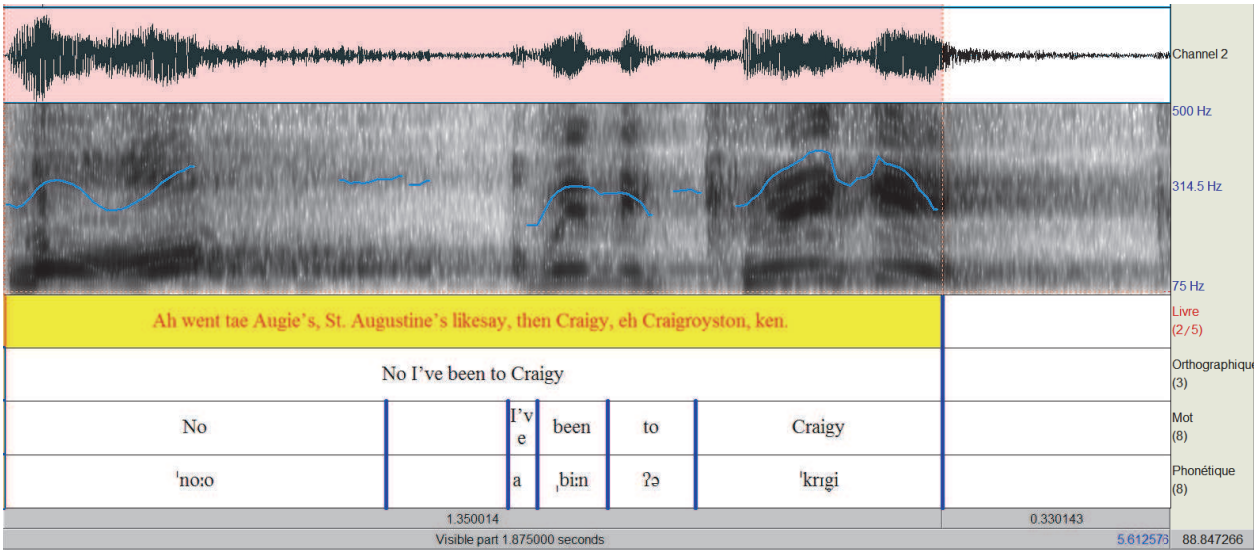


Illustration 1.16 – Captation sonore de Spud (00 :14 :47-00 :14 :48)

En revanche, l'intonation de Spud a été relativement bien captée par Praat. Les bornes de mots semblent mal placées, mais les réverbérations n'aident pas la lecture sans fichier audio (la scène est disponible sur la clef USB d'annexes, sous le nom « Spud job interview »).

L'intonation descendante-montante sur le premier mot est marquée dans la transcription, parce que ce contour intonatoire allonge la voyelle. Comme l'on pouvait s'y attendre, le passage contient des phonèmes scottisants : la glottalisation du seul /t/ et la prononciation trillée du /r/ de <Craigy>. La prononciation de Spud est prototypique de l'accent écossais, même dans un extrait aussi court. Les graphies indicatives de la prononciation sont également respectées : la monophtongaison de /aɪ/ et son ouverture en /a/ est retranscrite <ah>. Le <h> n'a pas de valeur phonétique et représente peut-être l'allongement du son /a/. <Tae>, ou plutôt le passage de <o> à <ae> figure l'aperture plus importante en scots qu'en anglais, qui n'est pas exploitée dans ce passage. La forme réduite [tø] est attestée en scots (Eagle, 2002), et cette aperture est symbolisée par la graphie <tae>, descendante de <taa> (DSL).

Pour compléter cette étude, j'ai retranscrit l'introduction du film et monologue connu d'Ewan McGregor, « Choose Life ». Cet extrait provient de « Searching for the Inner Man » (IV, 4), et conclut ce chapitre du livre. J'ai utilisé les graphies du livre lorsqu'elles figuraient effectivement

dans le passage (en gras, TR-L) et adapté le texte original suivant pour pouvoir faire figurer la transcription phonologique :

> Choose us. Choose life. Choose mortgage payments ; choose washing machines ; choose cars ; choose sitting oan a couch watching mind-numbing and spirit-crushing game shows, stuffing fuckin junk food intae yir mooth. Choose rotting away, pishing and shiteing yersel in a home, a total fuckin embarrassment tae the selfish, fucked-up brats ye've produced. Choose life.

La partie choisie du monologue dure environ une minute, et la transcription est rendue plus difficile à cause du fond musical (la chanson « Lust for life » d'Iggy Pop)[42].

Choose life. Choose a job. Choose a career. Choose a family.
[ʧʉz ləif /] [ʧʉz ə ʤoḇ/ ʧʉz ə kəˈɾɪə/ ʧʉz ə ˈfamli]
Choose a fucking big television. **Choose washing machines**, **cars**, compact-disc players,

[ʧʉz ə ˈfakɪn bɪg ˌtɛləvɪʃn̩] [ʧʉzˈwoʒɪn məˌʃinz/ kạz / ˌkəmpak ˈdɪsk ˈpleəz]

and electrical tin openers.

[ən əˈlɛktrɪkl̩ˌtin ˈopɛnəz]

Choose good health, low cholesterol, and dental insurance.

[ʧʉz gëd hɛlθ / ləʊ kəˈlɛstərol /ən dɛntl̩ɪnˈʃʌɾənts]

Choose fixed-interest mortgage repayments.

[ʧʉz ˌfɪkst ˈɪntrɛst ˈmɔɾʔgeɪʤ rəˈpemənts]

Choose a starter home. Choose your friends. Choose leisurewear and matching luggage.

[ʧʉz ə ˈstaːtə hom] [ʧʉz jə frɛndz] [ʧʉz ˈleʒəwe ən ˈmatʃɪn ˈlɤ̈gɪʤ]

Choose a three-piece suite [sic] on hire purchase in a range of fucking fabrics.

[ʧʉz ə ˌθri pis ˈswit ən ˈhəiə ˈpɤ̈ʧəs ɪn ə renʤ əv ˈfɤ̈kɪn ˈfabrɪks]

Choose D.I.Y. and wondering who you are on a Sunday morning.

[ʧʉz ˌdi əiˈwəi ən ˈwɤ̈ndrɪn hu ðə fak juạr ɒn ə ˌsɤ̈nde ˈmɔɾnɪn]

Choose sitting oan a couch, **watching mind-numbing**,

[ʧʉz ˈsɪʔɪn ən ə kʌutʃ / ˈwoʧɪn ˈməin ˌnɤ̈mɪŋ]

spirit-crushing game shows, **stuffing fuckin junk food intae yir mooth**.

42. Cet extrait figure sur la clef USB d'annexes, et porte le titre « Choose Life ».

[ˈspɪrɪt ˌkrÄʃɪn gem ʃɒz/ ˈstÄfɪn ˈfÄkɪn ˈʤÄŋk fud ɪntə jə mʌuθ]

Choose rotting away at the end of it, **pishing** your last **in a** miserable **home**,

[ʧʉz ˈroʔɪn əˈwəi aʔ ðɪ ən əv ɪʔ] ˈpɪsɪŋ jə last ɪn ə mɪsˈɾɛbl̩ hɔm]

Nothing more than **an embarrassment tae the selfish**, **fucked-up brats**

[ˈnÄðɪn mɔ ðən ən əmˈbarəsmənt tʊ ðə ˈsɛlfɪʃ fɐkt ɐp brats]

that **ye've** spawned to replace yourself.

[ðat juv spond tə rɪˈples jəˈsɛlvz]

Choose your future. **Choose life**.

[ʧʉz jə ˈfjuʧə] [ʧʉz ləif]

Dans le film, l'extrait du livre a été étoffé, et certaines transcriptions ont été changées : <pishing>, qui figure certainement une prononciation /ˈpɪʃɪŋ/ est prononcé [ˈpɪsɪŋ]. La post-alvéolarisation est attendue, au vu du corpus *Trainspotting*, car les deux formes du lemme $[\text{PISS}]_N$ sont graphiées <pish>. Une autre alvéolarisation de la consonne finale trouvée dans cet extrait est la prononciation de « yourself » [jəˈsɛlvz], qui est en fait la transcription de la forme au pluriel, fréquemment utilisée dans le livre pour désigner un référent singulier. Par exemple « us » pour I, par exemple « Ah remember you sitting oan toap ay us. » (IV, 5).

Ewan McGregor flappe la plupart de ses /r/, et produit des occlusives glottales (dans des mots tels que [ˈrɒʔɪn], 'rotting' ; [ˌkrÄʃɪn], 'crushing') et peu de voyelles sont altérées, certainement pour que le spectateur lambda puisse suivre le film. L'une des seules ambiguïtés de la transcription phonétique provient de la prononciation /ən/ pour *on*, qui n'est pas retranscrit dans les sous-titres. Comme vu précédemment avec l'exemple de Spud, les voyelles ont tendance à être antériorisées et abaissées, comme le mot « fuck » et ses dérivés le prouvent, avec les transcriptions /ˈfakɪn/ voire /fɐkt/ . L'aperture plus importante, et l'antériorisation sont visibles dans les transcriptions de Welsh, qui utilise les graphies <faahk>, <fahk>, <fah-k> voire <fark> pour figurer la prononciation.

Dans le film, j'ai recensé une dizaine de mots en scots ou d'utilisations de scottismes : « shite » ;

« bird », '*girl*' ; « skag », '*drugs*' ; « giro », '*RMI*' ; « ken » ; « aye » ; « doss » ; « punter », '*patron (of a bar, for instance)*' ; « wee ». Ce sont les seuls mots qui illustrent réellement le vocabulaire propre à l'œuvre, et ils se limitent, encore une fois, à des thèmes bien précis : les femmes, la drogue, les insultes ; les scottismes connus par tous, comme « aye » ou « wee ». Leur utilisation dans le film est extrêmement marginale, et ce sont tous des hapax legomena. Il convient maintenant de voir comment ces scottisme et cette langue standardisée dans TR-F est représentée dans les sous-titrages. Il est possible que les graphies du livre rejoignent celles des sous-titres, ou qu'ils soient écrits en anglais standard, comme c'était le cas pour *Precious*. Si les scottismes sont conservés, leur proportion et les mots sur lesquels ils portent sont également de bons indicateurs du lien entre graphie initiale et adaptations successives (en film, puis en report de dialogues du film).

1.3.2 Graphies communes et scottismes dans les sous-titres et le livre

Pour créer un lien avec la sous-partie précédente, je vais encore avoir recours à l'extrait « Choose life » afin de montrer les coupes effectuées dans les sous-titres, en les entourant de crochets quand ils n'y figurent pas.

> Choose life. Choose a job. Choose a career, [Choose] a family.
> Choose a fucking big television. Choose washing machines, cars, compact-disc players [and electrical tin openers].
> Choose good health, low cholesterol and dental insurance. Choose [fixed-interest] mortgage repayments.
> Choose a [starter] home. Choose your friends ! Choose leisurewear and matching luggage.
> Choose a three-piece suite on hire purchase [in a range of fucking fabrics. Choose D.I.Y. and] wondering who the fuck you are on a Sunday morning.
> Choose [sitting on a couch,] watching mind-numbing, spirit-crushing game shows, stuffing fucking junk food into your mouth.
> Choose rotting away at the end [of it, pissing your last] in a miserable home, nothing more than an embarrassment to the [selfish, fucked up] brats [that] you've spawned [to replace yourself].
> Choose your future. Choose life.

Ces sous-titres sont évidemment épurés pour pouvoir être lus dans le temps imparti (entre une seconde et demie et deux secondes, pour deux lignes comportant environ six ou sept mots au total). Le débit de parole est standard, avec un total de 130 mots en 70 secondes avec des pauses assez importantes. À titre d'information, le débit moyen de parole est de 160 mots par minute (Williams,

1998). Les doublons sont effacés, les adjectifs aussi. Le plus surprenant est l'épuration des insultes : sur les trois occurrences de « fuck », deux sont éliminées des sous-titrages. La démarche prescriptiviste attendue de sous-titres en anglais courant, sans presque aucun scottisme, se double d'une censure du discours tenu.

Les graphies vernaculaires connues, comme la figuration de la dévélarisation de /ŋ/ (<ing> ⟶ <in(')>). Il n'y a qu'un exemple de scottisme dans le sous-titrage du film, l'illustration 1.17. Cette graphie écossaise, <wouldnae> est hapaxique, et semble donc être une exception à la perspective normative du sous-titrage du film.

Les sous-titres ne sont pas représentatifs des dialogues, pour des questions de brièveté du nombre

Illustration 1.17 – Un exemple de scottisme dans les sous-titres de Trainspotting

de caractères affichables à l'écran et une tendance à éviter les insultes. Les scottismes ne sont pas épargnés, et presque tous sont écrits en anglais standard. Le sous-titrage est donc une retranscription à part, hors de la volonté de se rapprocher le plus possible du scots, et ce malgré les promesses du film (profondément écossais, et évoquant des thèmes subversifs). Les sous-titres sont donc écartés de la constitution du corpus secondaire pour toutes ces raisons.

Ce chapitre et les différentes forces qui tirent les graphies alternatives (argot local, appropriation

des graphies par stylisation, expressivité qui passe par l'utilisation de lettres doublées…) a permis de voir quels types de graphies deux locuteurs de l'AAV et du scots mettent en place dans leur dispositif fictionnel, ainsi que les prononciations simplifiées adoptées par les acteurs qui portent leurs personnages à l'écran.

Conclusion du chapitre

Les lexiques compilés à partir des œuvres de fiction sont utilisés comme base pour la compréhension des graphies tirées du corpus internet, mais certains lemmes sont presque inexistants dans les œuvres de fiction : ainsi, *fuck*, *bitch* ou *nigga* nécessitent de constituer un sous-corpus à part, et leurs utilisations ne reflètent pas le jeu du chat et de la souris que nous allons étudier plus loin (au chapitre 4). En revanche, les démarches des auteurs, très différentes de celles des scripteurs d'internet, servent les intrigues des ouvrages de fiction–par l'élaboration d'un portrait identifiant immédiatement le personnage, (dans le cas de *Trainspotting*). Ce portrait peut éventuellement évoluer, comme l'on peut le voir dans *Push*. Les graphies sont donc éminemment identitaires, autant au niveau d'un scripteur que d'une communauté, ce que corrobore la suite de cette thèse.

Ce chapitre était l'occasion d'exposer les disparités sous-tendant la création de graphies alternatives, ainsi que d'esquisser la méthode d'exploration du corpus principal. Les graphies de *Push* et *Trainspotting* se ressemblent par la volonté de leur auteur de les éloigner de l'anglais standard, pour des raisons stylistiques différentes (ou en suivant des buts stylistiques différents). Leur différence principale avec le corpus des forums est leur manque de spontanéité : travaillées et réfléchies par les auteurs, elles ont été relues à maintes reprises avant l'édition finale. Malgré cela, elles sont identiques ou presque aux graphies des scripteurs de forums. Ceci démontre leur légitimité en tant qu'exemples de graphies représentatives du scots et de l'AAV.

Afin de représenter les types de graphies, ainsi que les pressions que leurs scripteurs subissent, je les ai résumées dans l'illustration 1.18. La norme se nourrit de symboles qui la préservent, comme les dictionnaires, et elle est l'outil de certains actants/militants qui la défendent, comme les professeurs

(Blu Rain, dans *Push*, par exemple) ou les *Grammar Nazis*. Les langues minoritaires occupent un espace marginal, entre la norme standard de la langue dominante, et les pratiques graphiques alternatives. Celles-ci peuvent partager les mêmes caractéristiques que les fautes d'orthographe, à la différence que ces dernières sont involontaires. De la même façon que les scripteurs de forums qui savent parfois qu'ils ne graphient pas les mots correctement (comme <logerithum>, utilisé par une membre de SC qui dit ne pas savoir comment orthographier le mot), les graphies gabaritiques de Precious au début de son carnet de bord ne sont pas volontairement graphiées ainsi.

Par rapport à l'Orthographe, les graphies alternatives sont moins légitimes, mais elles ont une logique interne, systémique, qui les rend cohérentes. La plupart des graphies que j'étudie, dans ce chapitre et les suivants, appartient au grand cercle des graphies alternatives, dont le socle est la consonne—qui est également le socle des graphies gabaritiques, qui seront peu utilisées dans ce travail de thèse. La stabilité de l'amplitude allographique repose souvent sur les consonnes à la manière du système des langues chamito-sémitiques. Les voyelles sont souvent plus fluctuantes. L'initialisme et l'utilisation d'abréviations (exemple dans *Trainspotting* : « k.b'd ») pose plus de problème à cause de l'homonymie : K.B., veut autant dire « Kitchen Bitch », « Kick Back », '*petite fête*', ou « Knock Back », son sens au nord de l'Angleterre et en Écosse (source : UD). Ces initialismes sont donc placés dans le pôle moins intelligible des graphies alternatives. L'intelligibilité est l'un des éléments déterminants : l'aisance de lecture est optimale pour les textes orthographiés, mais les premières entrées du journal de Precious ne sont pas lisibles : plus les graphies sont basses dans le cercle que je propose pour synthétiser les phénomènes étudiés (illustration 1.18), plus elles sont cryptiques.

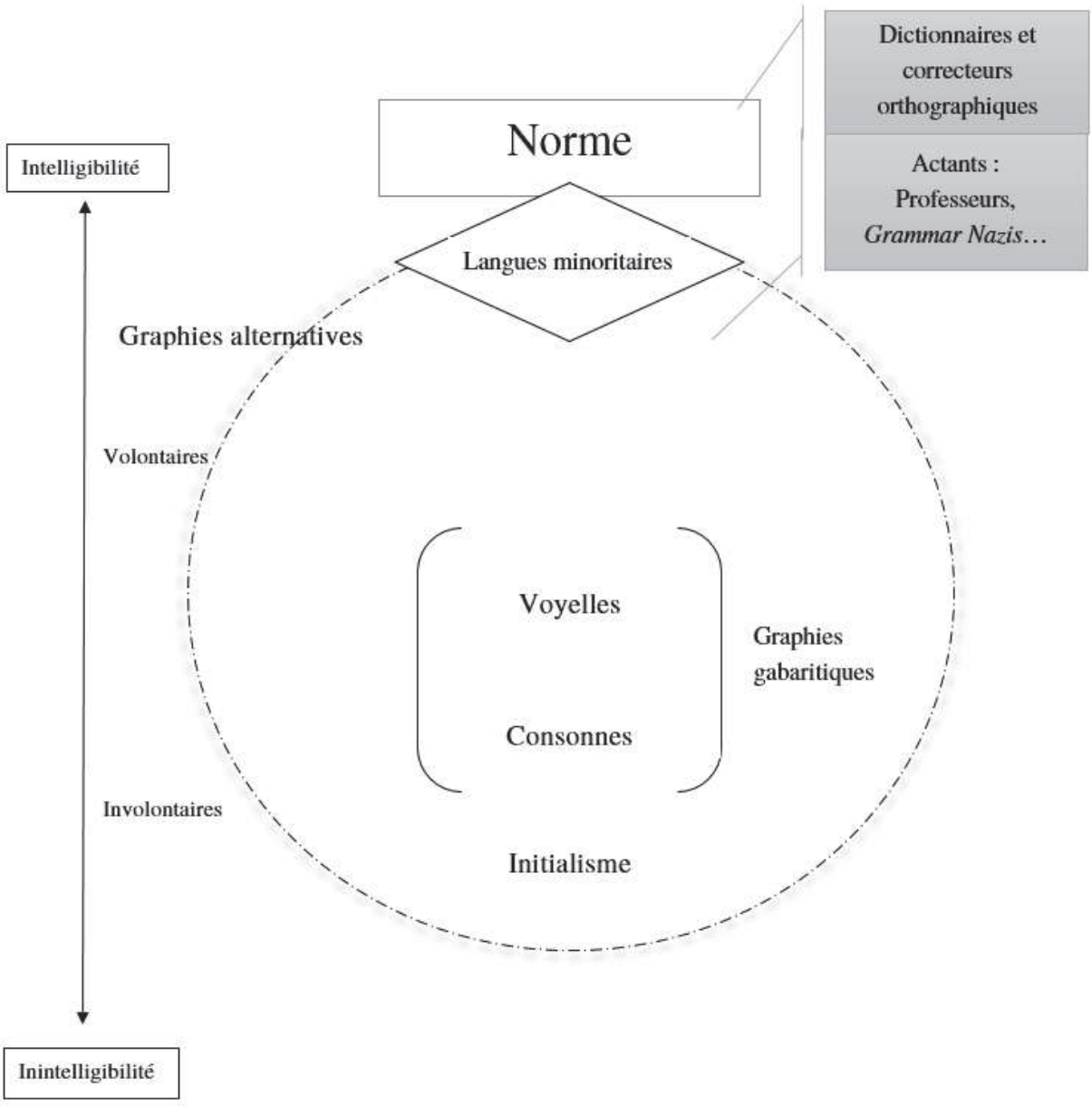


Illustration 1.18 – Panorama des usages des graphies alternatives et insertion écologique

Pour conclure ce chapitre, je citerai l'une des références en matière d'orthographe du français : « La forme des mots est trompeuse : *ce qui compte sur le plan graphique, c'est le rapport entre l'écrit et l'oral à une époque donnée, autrement dit la traduction du graphème en phonème et inversement.* » (Catach, 1978, emphase apportée par l'auteure). Malgré les vues fondamentalement prescriptivistes

de N. Catach, je souscris tout à fait à cette façon d'appréhender le rapport entre la graphie et la phonie. Les graphies alternatives appartiennent à des espaces-époques en dehors de l'orthographe et de la norme, mais elles sont représentatives, en synchronie, du rapport phono-graphématique. Elles ont donc un statut de norme alternative, représentant des langues minoritaires, ou des usages qui ne se placent pas sous la juridiction des organismes normatifs.

Les particularités discursives des ensembles linguistiques, ou langues minorées en présence, sont décrites dans des travaux scientifiques. C'est l'objet du chapitre suivant, qui rendra compte non plus de représentations graphiques de leur parler par leurs locuteurs, imaginaires ou non, mais de la description formelle et théorisée de ces langues par la communauté universitaire.

Chapitre 2

Caractérisation de l'AAV et du Scots

Ah hope ye'll like the efforts
In this wee book o rhyme
The verse is often rotten
Bit the language is sublime

C. Eliott - *Robert Henryson Poems*

Dans le chapitre précédent, nous avons exploré les bases de la graphophonématique des deux langues (dont le statut a déjà été discuté dans l'introduction) lors de la constitution d'un corpus théorique et contrastif. Dans ce chapitre, il est question des langues minoritaires et de la dimension de stigmatisation. Ce chapitre vise à faire un relevé de quelques caractérisations du scots et de l'Afro-Américain Vernaculaire (AAV) tels qu'ils sont décrits par les universitaires. Ces deux langues minoritaires seront décrites tant au point de vue du lexique, de la graphie (ou des particularités graphiques), de la syntaxe et de la phonologie.

Cette étude est théorique et inspirée de travaux de linguistes et d'organismes officiels (textes gouvernementaux, recommandations dans un cadre scolaire, feuilles de style en ligne…). Les travaux de L. Green (Green, 2002), sur le statut phonologique et sociétal de l'AAV sont le point introductif

de l'étude contemporaine de la langue, d'un point de vue endocentrique. Les polémiques soulevées par la décision du bureau d'éducation d'Oakland de déclarer l'AAV comme langue des enfants afro-américains a déclenché des débats sur la *Linguist List* (Labov, Wolfram, Baugh, Torbert & Rickford, 1996), à l'époque où l'AAV était encore appelé *Ebonics* (je préciserai ce point dans la partie 2.4).

Les travaux d'E. Schneider (Edgar W Schneider, 1989) sont importants pour comprendre l'histoire de l'AAV, notamment les caractéristiques phonologiques et lexicales antérieures. Les études de Dillard (Dillard, 1972) et de Wolfram (Wolfram, 1998) dans la partie 2.4.4 sont importantes pour comprendre la façon dont la phonologie a évolué, dans les cent dernières années. Les sources concernant l'AAV et qui n'ont pas été écrites par des linguistes sont multiples : elles proviennent soit d'utilisateurs d'internet (comme les scripts de « traduction » afro-américaine), soit de sites reconnus comme l'*Urban Dictionary* (UD) (« Urban Dictionary », 2003).

Quant au scots, le travail de J. Stuart-Smith (Stuart-Smith, 2003 ; Stuart-Smith, Timmins & Tweedie, 2007) a grandement influencé cette thèse, surtout dans l'analyse de la phonologie, et ceux d'Aitken et de Corbett (Corbett et al., 2003 ; McClure, Aitken & Low, 1980) pour l'étude linguistique. Les problèmes de l'évolution du scots et son intégration dans les programmes scolaires, dans une période relativement récente où il n'avait pas sa place, sont couverts par le livre de Kay (Kay, 1986). Les travaux de la Saltire Society [1], destinés aux auteurs désireux d'écrire en scots, et les initiatives gouvernementales, ou autrement légitimes comme *Aye Can*, le site de l'éducation nationale écossaise, la BBC…illustrent l'intérêt pour le scots dans la sphère numérique.

Le scots (et son interaction avec le gaélique) seront abordés dans cette partie, puis l'AAV (en comparaison avec l'anglais blanc du sud). Une fois l'étape descriptive passée, les perceptions des deux langues minorées seront étudiés. Bien que l'AAV et le scots soient reconnus dans cette thèse comme des langues légitimes (c'est-à-dire en tant qu'ethnolectes complexes, parlés non par appartenance génétique à un groupe donné, mais par appartenance sociale, voire goût personnel), ils souffrent de la stigmation des locuteurs d'autres variétés de l'anglais. Les entrées de la définition d'*Ebonics* sur l'UD seront étudiées en 2.6, ainsi que les perceptions de l'écossais (le peuple) par rapport à leur(s) langue(s)

1. Purves et Saltire Society, 2002.

par des locuteurs anglais et américains, et toujours dans l'UD : les mots *scots*, *scottish*, *inglis* et *lallans* seront analysés. Les différentes désignations des langues (dont le jeu entre *Ebonics* et *e-bonics*, point traité en 2.4.2) sont indicatrices d'une certaine forme de stigmatisation, par les locuteurs d'autres variétés de l'anglais. L'introduction de la couleur de la peau dans *ebonics* ou *black English*, par exemple, sont indicatrices d'une certaine mise à distance (l'AAV est également pratiqué par les blancs—et autres ethnies), alors que la tendance actuelle va vers l'appellation *African American language* (sans mention de *vernacular*).

L'anglais américain et britannique standard dont je parle dans ce chapitre sont des constructions à partir de l'anglais tel qu'il est pratiqué dans ces pays, et à partir des grammaires, forcément normatives (Sorlin, 2012). C'est une vue de l'esprit, qui est représentée par l'anglais tel qu'il est utilisé dans les médias et les livres de ces deux pays. La prononciation correcte, « normale », ou normée, est une invention des locuteurs qui ressemble à un condensé du *génie de la langue* dans son incarnation la plus extrême. L'absence d'organisme régulant la langue britannique et américaine, à la manière de l'Académie française—qui n'est d'ailleurs pas à jour sur les derniers développements lexicaux— provoque également cette dureté envers les langues minorées ou les dialectes des autres. Si le but de l'Académie est de conserver la langue (ou de la défendre), le Trésor de la Langue Française (TLF) n'est pas utilisé quotidiennement par les locuteurs du français. L'absence d'autorité américaine et britannique n'empêche néanmoins pas les locuteurs d'émettre des jugements et de tenter de deviner l'appartenance socioculturelle à partir de l'accent (le fil de discussion « Linguistic Profiling » de *Black Planet* (BP) est un exemple de ces débats). La norme construite existe dans la tête de chacun, née de leur expérience linguistique, de leur milieu socioculturel, entre autres, et elle est une prescriptrice intransigeante.

2.1 Sociolinguistique du statut du scots

2.1.1 Le scots

Outre l'anglais britannique standard, les Écossais ont à leur disposition l'anglais standard écossais, le scots, ainsi que le gaélique. L'anglais écossais standard est défini comme « the characteristic speech of the professional class [in Scotland] and the accepted norm in schools » McClure (Hogg et al., 1992)).

Quant à savoir ce qu'est la norme acceptée à l'école à une période où l'exposition et la production en scots sont vivement encouragées, il existe des différences de perception et d'enseignement : le scots est censé être cantonné à des cours bien particuliers (notamment la littérature, et l'étude des oeuvres de R. Burns), mais les élèves interviennent à l'oral en scots [2] et n'arrivent parfois pas à faire la différence entre l'anglais écossais standard et le scots.

La répartition de la connaissance de ces langues est essentiellement géographique, le gaélique et le scots étant des langues développées diachroniquement par deux types de populations, influencées différemment. Le gaélique est pratiqué par l'ancienne population celtique, alors que le scots a subi une influence anglienne, voire angli-française (Murison, 1977). Tous les Écossais utilisent l'anglais écossais standard, et comprennent, ou pratiquent l'anglais britannique standard, qui est une forme de l'anglais courante à la télévision, dans les journaux, et est la langue maternelle de la plupart des Britanniques.

Les statistiques à propos de l'utilisation du scots sont très rares, étant donné qu'il existe un problème de caractérisation, voire de discrimination, contre la langue (voir 2.6.3.2 ci-dessous). Comme précisé en introduction, le scots est considéré dans ce travail comme étant une langue à part entière, ce qui n'est pas toujours le cas dans la recherche, surtout dans le monde anglo-saxon (Hickey, 2004). Le scots est tout de même reconnu comme langue dans la seconde édition de l'*Encyclopedia of Language and Linguistics* (Brown & A. Anderson, 2006), ce qui montre à quel point les perspectives changent en fonction du degré d'éloignement avec la langue. Plus elle est considérée de manière distanciée, et

2. Données personnelles ; étayées par (Costa, 2010).

plus elle regagne ses caractéristiques de langue légitime.

Cette attitude, anglophone, ressemble à celle des Écossais envers la graphie du scots (voir 2.2.2 ci-dessous). Le scots est souvent défini comme de l'anglais auquel on aurait retiré ou ajouté des graphèmes, ou qui serait prononcé différemment. Ces deux exemples relèvent du centrisme autour de la langue que l'on parle et de celle que l'on considère—il est toujours plus simple de comparer avec une autre langue et d'émettre des jugements axiologiques que de tenter de décrire un système en soi. Non que l'anglais et le scots n'aient rien en commun, mais une fois la comparaison initiale passée, il serait peut-être préférable de caractériser les particularités de la langue au lieu de revenir à la langue majoritaire. Cette attitude est également visible dans les documents produits par les sociétés de défense et de promotion du scots tels le Makar's Club, qui a édité des recommandations d'usage graphique, et les cercles écossais de promotion de la langue, qui éditent des livres comme le *Orkney Language and Culture Group* et son dictionnaire dialectal des Orcades (Flaws & Lamb, 2001). Les perspectives y varient entre une considération du scots en tant que langue à part entière, tout en incluant un volet didactique où il est présenté aux lecteurs comme n'étant composé que de mots en anglais écrits différemment.

Cette approche renie le lexique du scots et montre une ignorance profonde envers cette langue et sa culture (le premier texte écrit en scots date tout de même du 14ème siècle), ce qui n'est pas étonnant lorsque l'on se place dans un contexte historique britannique. Pendant longtemps, les nations minoritaires (Pays de Galles, Écosse, Irlande) furent reléguées au second rang par une Angleterre prompte à considérer qu'en tant que pays dominant, seule sa culture (et sa langue) est importante. L'Angleterre n'a pas été le seul pays à être tombé dans ces écueils ; la France s'est également illustrée par des politiques linguistiques niant l'importance des langues régionales. La relative stabilité de la langue majorée (langue de prestige) en Angleterre et en France explique certainement une partie de ces attitudes : les gouvernements ont tenté d'écraser les particularités régionales pour fédérer ces pays, et les langues ont fait les frais de ces politiques d'arasement culturel.

Cette analyse est à nuancer, étant donné les efforts faits dans les dernières décennies pour raviver le gallois et les langues régionales. Ces acquis sont aussi nés de la dévolution des pouvoirs au Pays de

Galles et à l'Écosse, entre les années 1970 et le début du 21ème siècle. Le vif regain d'intérêt envers le gallois a provoqué sa renaissance, et le gallois est devenu une langue activement pratiquée depuis le 18ème siècle (Aitchison, Carter & Aitchison, 2000). Le scots a récemment fait son entrée à l'école, depuis qu'il a été reconnu comme une langue régionale à part entière. Des sites pilotés indirectement par le gouvernement écossais comme *Scuilwab - For Teachers an Young People fae Nursery tae University* (Chris, 2010) offrent des activités scolaires en scots de la maternelle au lycée. Ceci participe à l'évolution de la perception du scots en Écosse. Le scots a longtemps été négligé en tant que langue et cette attitude est visible dans la hiérarchisation des efforts linguistiques. Ceux-ci ont d'abord été dirigés vers le gaélique (qui n'a pas connu de réforme orthographique, contrairement à son pendant irlandais), puis enfin vers le scots.

La polysémie du mot « langue » se retrouve dans le mot *tongue*, utilisé pour décrire ou parler de la langue écossaise. L'expression *the Scottish tongue* est présente dans de nombreuses ressources écossaises pour parler de la langue, comme le *Dictionary of the Older Scottish Tongue* (DOST) *Dictionary of the Older Scottish Tongue*, ou *Scots : the Mither tongue* (Kay, 1986). Selon le *Scots Dialect Dictionary* (SDD) *Scots Dialect Dictionary* , le signifiant *tongue* peut avoir de multiples sens en scots : « 1 dialect. 2 manner of speaking. 3 abuse, violent language ». Les sens du verbe *tongue* sont également surprenants : « 1 to talk immoderately. 2 to scold, abuse » (Warrack, 2006). Le fait qu'un mot ait des signifiés aussi différents semble illustrer la position sociale traditionnellement basse des locuteurs du scots, ainsi que le traitement péjoratif du scots en tant que langue à part entière.

La confusion entre fierté nationale d'une part, et traitement péjoratif d'une langue aux traits si similaires à l'anglais de l'autre, explique ce déficit en données afférant directement au scots. La démarcation entre anglais écossais, scots des Basses Terres, et variétés d'anglais parlées au nord de l'Angleterre est floue, et le scots comprend plusieurs dialectes, que l'on peut regrouper en scots du sud, du centre et du nord du pays, sans compter les variétés insulaires parlées dans les archipels des Orcades et aux Shetland (Kay, 1986).

La stigmatisation de la langue peut être datée au 17ème siècle, où le déplacement de la cour royale écossaise à Londres dû à l'accession au trône anglais de Jacques VI d'Écosse (Jacques I d'Angleterre)

marque le déclin progressif de l'utilisation du scots en faveur de l'anglais. À partir du Traité de l'Union (1707), l'*Inglis* (ou scots) fut relégué au rang de dialecte vernaculaire, employé par les classes pauvres, ou dans des contextes domestiques. Ainsi, les Écossais ont souffert de la dévalorisation de leur langue, et la confusion entre simple accent (Trudgill & Hannah, 2007), continuum dialectal entre le nord de l'Angleterre et les dialectes des Basses Terres (*Lallans*, ou *Border Scots*). L'apprentissage du scots formel est relégué à des bases historiques et littéraires à l'école, notamment avec les poèmes de R. Burns, qui éloignent la langue pour les Écossais, et leur donne l'impression qu'ils ne parlent pas une variété de scots aussi prestigieuse, ou éloignée d'eux. L'étude de Burns et de son oeuvre sont abordés dès la première année de l'école primaire, et sont intégrés au programme d'excellence de l'Écosse (Scotland. Scottish Government, 2010). À ce scots formel, quelque peu archaïsant, s'ajoute l'étude du scots, depuis le début du 21ème siècle.

2.1.1.1 Recensement et place du scots

Il n'existe pas de recensement fiable sur le nombre exact de locuteurs de scots à l'heure actuelle. Les statistiques fluctuent entre 1,5 et 2,7 millions de locuteurs à la fin du 20ème siècle (Kay et al., 1995). La reconnaissance du scots pose problème : moins de locuteurs supposés répondent favorablement à des questions telles que « Parlez-vous le scots ? » qu'à une tournure différente : « Parlez-vous le scots ou une de ses formes dialectales ? » dans le recensement de 2001, ce qui n'a pas été le cas en 2011 (Scotland, 2011 ; Great Britain. Registrar General Scotland & Scotland. General Register Office, 2001).

La place de la langue et de l'identité nationale en Écosse est illustrée par les contrastes entre les questions posées dans deux derniers formulaires de recensement de 2001 et 2011. Le questionnaire de 2001 reprenait en substance celui de 1991, mais le changement dans l'attitude britannique et écossaise envers ces questions apparaît dans le format du questionnaire 2011, où le nom de cette langue régionale apparaît enfin en toutes lettres.

Les questions 16 à 18 sont particulièrement intéressantes. Les réponses à cocher sont indiquées dans le tableau en italique (les questions admettant des réponses multiples sont marquées d'un asté-

risque) ; la question 16-2011 est un tableau à 16 cases, dont l'abscisse et l'ordonnée sont les réponses. Les questions non pertinentes à notre sujet d'étude sont grisées (dans le tableau 2.1).

Questions	Recensement 2001	Recensement 2011
14	What religion, religious denomination or body were you brought up in ?	What do you feel is your national identity ?* Scottish/English/Welsh/Northern Irish/British/Other
15	What is your ethnic group ?	
16	Can you understand, speak, read or write Scottish Gaelic ?* Understand spoken Gaelic/Speak Gaelic/ Read Gaelic/Write Gaelic/None of these	Which of these can you do ? Understand/Speak/Write/Read English/Scottish Gaelic/Scots
		How well can you speak English ? *Very well/well/not well/not at all.*
17	Do you look after, or give any help or support to family members, friends, neighbours or others […]	Do you use a language other than English at home ?* *No, English only/British Sign Language/Yes (specify)*
18		

Tab. 2.1 – Questions figurant sur les formulaires familiaux et individuels de recensement 2001 et 2011

Le recensement a ajouté une question à propos du scots après que le Royaume-Uni a ratifié la Charte Européenne des Langues Régionales ou Minoritaires en 2001, reconnaissant ainsi le statut de langue régionale au scots en vertu de la partie II de la charte (Europe, 1992). Durant les séances de consultation à propos du contenu du questionnaire de recensement de 2011, l'inclusion de questions sur l'utilisation des langues, notamment le scots a été demandée fréquemment, selon la Foire aux Questions (FAQ) du site de consultation du recensement 2011 [3].

2.1.1.1.1 Explications sur la section langage du questionnaire et comparaisons

Les informations demandées diffèrent, et le recensement de 2011 voit enfin l'apparition de la notion de scots dans les questionnaires. Les recensements antérieurs ne font pas grand cas de questions de langues, et il est étonnant de voir que la connaissance de l'anglais semblait aller de soi. Aucune

3. http ://www.scotlandscensus.gov.uk/en/faqs/detailed/censusquestions.html# q1 consulté le 06/01/2013.

auto-évaluation du niveau de maîtrise des répondants n'était demandée. À titre d'exemple, le gallois est traité de façon relativement similaire dans le recensement du Pays de Galles (Government, 2012). Le recensement gallois de 2011 contient une question proche de la Q16-2001 (« Can you understand, speak, read or write Welsh ? » qui est formulée différemment de sa version en 1991. Elle commençait par *does the person* […]. Le changement de style semble avoir eu une influence légère sur les résultats, avec un passage de 18.7% de réponses positives en 1991 à 20.8% en 2001 [4].

En revanche, la question 18 est « What is your main language ? », avec un choix entre le gallois, l'anglais ou une autre langue : il existe une ouverture relative vers le fait que le gallois pourrait être la langue principalement utilisée par les répondants, ce qui est corrélé par le fait que l'anglais et le gallois sont traités ensemble dans les résultats du recensement. La conséquence (négative, pour cette étude) est le fait que les statistiques du gallois en tant que langue primaire sont inexistantes en soi. Dans l'ensemble Angleterre-Pays de Galles [5], seuls 58 répondants disent utiliser le gaélique écossais comme langue principale, contre 562 000 galloisants dans la même zone (Poulter & Chorley, 2013). Le recensement est néanmoins proposé en anglais et en gallois, ainsi qu'en gaélique écossais depuis 2001–mais pas en scots.

En 2011, la question de la maîtrise de l'anglais sert de base et de critère à l'intégration de populations étrangères (par exemple, en offrant des documents traduits dans les langues parlées par les populations qui ne maîtrisent pas l'anglais), ainsi qu'au repérage potentiel des communautés illettrées [6].

L'aspect linguistique du recensement n'était que partiel avant 2011, la maîtrise des quatre aptitudes demandées traditionnellement dans l'étude d'une langue (compréhension, conversation, lecture, écriture) n'ayant été ajoutée qu'en 2001. Le recensement de 1991 contenait une question similaire à 16-2001, H13, mais celle-ci ne comprenait pas le volet de compréhension du gaélique écossais (« *Can the person speak, read or write Scottish Gaelic ? Can speak Gaelic/Can read Gaelic/Can write Gaelic/None of the above* »), éliminant ainsi des statistiques une partie de la population qui comprenait le

4. Le recensement écossais de 1991 (*Household*) était déjà formulé avec « can you […] »
5. Les résultats pour l'Écosse et l'Irlande du Nord ne sont pas encore publiés
6. Informations tirées de <http ://www.gro-scotland.gov.uk/files2/the-census/policy/2011_ census_ recommendation_ paper.pdf > consulté le 06/01/2013.

gaélique sans le parler. Le problème, aussi visible dans le recensement gallois, est la non-cohérence de certains résultats : certains peuvent ne choisir que de répondre parler ou écrire le gaélique (ou le gallois), en omettant de cocher la case « compréhension de la langue » [7]. Des résultats similaires sont inaccessibles pour l'Écosse parce que la connaissance minimale de la langue correspond au fait d'avoir coché au moins une des cases. Près de 92 000 répondants avaient dit maîtriser quelques éléments linguistiques du gaélique (General Register Office (Scotland), 2005).

La place plus grande accordée à l'identité infra-nationale, ou régionale (Q14-2011 (Scotland, 2011)) au sein de la Grande-Bretagne fait écho au nouveau regain d'intérêt envers toutes les langues et dialectes parlés sur le territoire, et non plus seulement envers le gaélique écossais. Le scots figure pour la première fois dans le recensement, signe qu'il est enfin reconnu en tant que langue. Cela est renforcé par la Q18-2011, où les situations de bilinguisme (utilisation d'une langue autre que la langue nationale au domicile) sont interrogées. Les résultats du recensement 2011 ne sont pas encore disponibles pour l'Écosse.

2.1.1.1.2 Diffusion du Scots

La définition du scots peut être problématique pour certains répondants, qui n'ont jamais envisagé le scots comme une langue à part entière, mais plutôt comme une particularité régionale. C'est surtout le cas des locuteurs plus âgés, qui n'ont pas été sensibilisés au scots à l'école comme les plus jeunes, et voient cette langue comme une langue conservée dans les poèmes de R. Burns, ou, plus récemment, dans ceux d'H. MacDiarmid. C'est la raison pour laquelle des sites comme *Aye Can* [8], orchestré par les autorités écossaises, clarifient cette entité linguistique en vue du recensement 2011. Il a été conçu de façon didactique, et comprend une FAQ rédigée en des termes simples : « Can I understand Scots ? » ; « I would say I speak Scots, but I don't recognise some of the words used by others. Why is that ? ». Les réponses apportées sont très concises, et invitent le visiteur à écouter les extraits audio présentés sur le site. Ce sont des enregistrements de discours relativement spontané en scots, qui servent à reconnaître la particularité régionale du scots parlé : les locuteurs du scots d'Inverness pourraient ne pas

7. Il y a environ 15% de répondants dans ce cas au Pays de Galles en 2011
8. Accessible à l'adresse <http ://www.ayecan.com>

se reconnaître dans un scots pratiqué à Glasgow, et *vice versa*.

La distribution des locuteurs de scots est le négatif de la distribution de locuteurs de gaélique écossais dans le pays (comme le montre la carte ci-dessous, en 2.1.2). Le gaélique et le scots se sont développés parallèlement, et ne concernent pas les mêmes populations, comme expliqué sur *Aye Can* : le scots s'est développé dans les régions sud et est du pays. Quant au gaélique, il est maintenant en recul et n'est plus parlé que dans les régions nord et ouest de l'Écosse.

Illustration 2.1 – Différents dialectes du scots proposés à l'écoute (Aye Can)

Seuls dix échantillons sonores sont disponibles sur *Aye Can*, qui correspondent peu ou prou aux différences entre les topolectes du scots : du nord-ouest au nord-est, les échantillons traitent le scots de Galloway, celui des Marches Écossaises, du centre ouest et du centre-est ; de Fife, d'Angus, le dorique (parlé au nord-est), et le scots de Caithness. Enfin, il y a également des illustrations sonores de

documents des Orcades de l'archipel des Shetlands. Des documents comme *Mapping Languages in Edinburgh* (McPake, 2002) arrivent à des conclusions relativement similaires : le shetlandais et le dorique sont agglomérés au scots, et il n'y a que 15 enfants qui déclarent connaître le gaélique, soit parce qu'ils viennent de famille gaélophones, soit parce qu'ils prennent des cours du soir. En revanche, le scots est relativement peu utilisé. Seuls 2.6% des répondants disent le pratiquer, chiffre qui peut s'expliquer par l'ambiguïté de la définition de scots.

L'illustration 2.2 montre les zones de développement successif du scots en Ulster et en Écosse : la zone rouge correspond au développement du substrat northumbrien et vieil-anglais au 9ème siècle, auquel s'ajoute ensuite la zone orange, où le vieil-écossais ou *inglis* a commencé à émerger (au 15ème siècle). Ces deux zones ainsi que la jaune symbolisent les lieux où le scots est encore pratiqué aujourd'hui– le territoire bleu correspond, quant à lui, à l'aire de développement du gaélique.

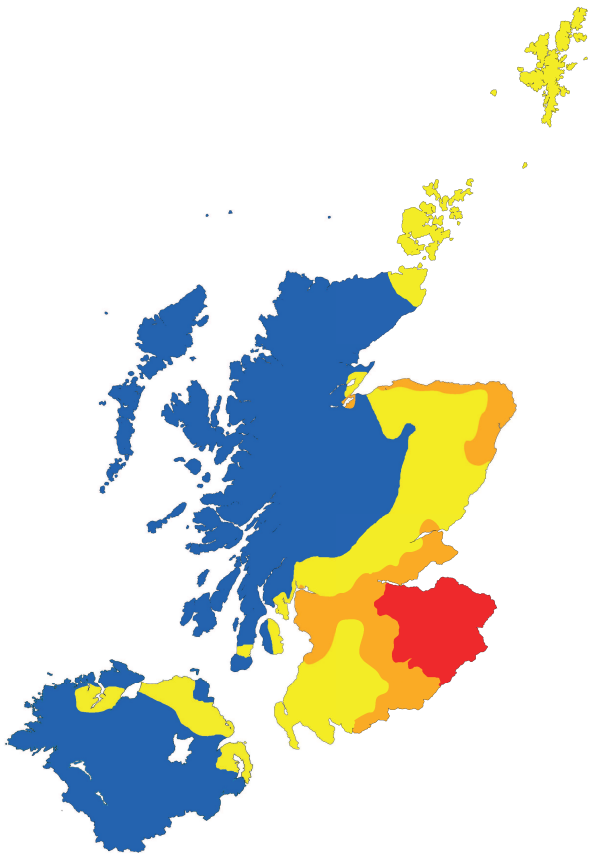


*Illustration 2.2 – Distribution du Scots (***shepherd_historical_1956** *; Ross, 2001)*

Le site est extrêmement didactique, et les contenus enregistrés sont des extraits de conversation spontanée. *Aye Can* est emblématique des efforts déployés pour la compréhension du volet dialectal du scots, pour le reconnaître comme une langue minoritaire à part entière.

À titre d'exemple, le gallo en France n'est pas illustré en ligne par le même type de site, bien que des initiatives soient prises pour en démocratiser l'usage en Bretagne (épreuve optionnelle au baccalauréat depuis 1984, version gallo et française dans les transports en commun à Rennes…). Le site qui pourrait être le plus proche d'*Aye Can* est le dossier thématique du Centre National de Documentation Pédagogique (CNDP) [9], qui est beaucoup plus axé sur l'apprentissage et présume une ignorance totale de la part du lecteur. Quelques détails comme les emprunts des mots au breton, dans les zones de contact sont évoqués, mais le gallo y est vu comme un tout linguistique, et pas comme une langue réellement parlée.

2.1.2 Utilisation du gaélique

À l'inverse, le gaélique écossais ou *erse* (nom également utilisé pour la variété de gaélique parlée en Irlande) est pratiqué dans les Hautes-Terres, ainsi que sur l'île de Skye, les Hébrides et Glasgow. Le dernier recensement, effectué en 2001, estime le nombre de locuteurs à 58 700. Les locuteurs natifs, de moins en moins nombreux, se concentrent surtout dans le nord du pays et dans les îles (principalement, les Hébrides). Dès 1986, Kay estime qu'il ne reste plus d'adulte qui ne parle que gaélique écossais (Kay, 1986). La dernière locutrice de la variété des basses-terres du gaélique (gaélique gallovidien, parlé dans les îles de Galloway !) est morte en 1760 [10].

Le gaélique est présent sous la forme de journaux, programmes à la télévision nationale ainsi qu'à la radio. Une station de radio, *BBC nan Gàidheal*, diffuse depuis 1985 des programmes en gaélique écossais exclusivement (mais n'est disponible sur la bande FM qu'au centre du pays, les régions au sud et les îles au nord n'y ayant pas accès). Depuis 2008, une chaîne de télévision, BBC Alba, ne diffuse que des programmes en gaélique écossais, bien qu'elle n'émette que sept heures par jour maximum.

9. Accessible à la page <http ://www.cndp.fr/crdp-rennes/crdp/crdp_dossiers/dossiers/gallo/langue/accueil.htm>
10. Un siècle plus tard (en 1850) mourrait le dernier locuteur de norne, langue scandinave parlée dans les Orcades et les Shetland.

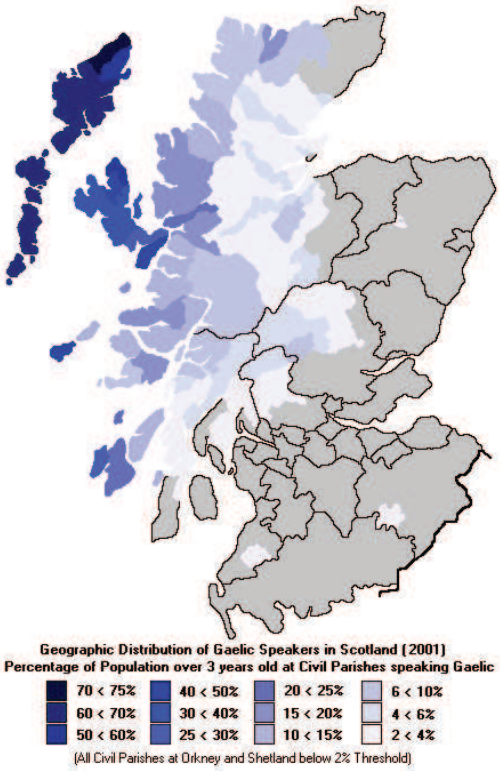


Illustration 2.3 – Répartition des locuteurs du gaélique écossais en 2001 (General Register Office (Scotland), 2005)

La plupart des programmes est cependant sous-titrée en anglais (sauf les programmes pour enfants), et cette chaîne est accessible par tous les Écossais.

Des podcasts sont disponibles sur Internet, ainsi que du contenu multimédia, afin d'apprendre le gaélique, ou de maintenir son niveau. Beaucoup d'initiatives pour l'apprentissage du gaélique naissent du désir d'un internaute de faire découvrir sa langue ou de l'enseigner (Alexander, 2012), alors que d'autres sites proviennent de sources officielles, à l'échelle de la région ou du pays comme *Beag Air Bheag* (BBC, 2008). Il existe également des communautés virtuelles en gaélique écossais, dont le forum sobrement intitulé « Forum en Gaélique » (« Fòram na Gàidhlig », 2007) où les internautes s'échangent des liens pour apprendre le gaélique, parlent d'autres sujets dans une section « off-topic »… La vie virtuelle du gaélique écossais est assez prospère, et sa présence sur des réseaux sociaux très fréquentés (Facebook, Twitter) le rend accessible à tous.

En outre, le gaélique jouit d'une image positive, et fait partie intégrante des traditions écossaises. Cette langue est fortement caractérisée par sa différence avec l'anglais, et est donc bien plus simple à

assimiler à la frange celtique européenne (le gaélique écossais et le gaélique irlandais sont mutuellement intelligibles) ainsi qu'au folklore associé. Quatre festivals annuels célèbrent la culture écossaise, et plus précisément gaélique, à travers tout le pays (Celtic Connections est le plus connu, à Glasgow et réunit plus de 100 000 participants ; l'édition 2012 accueillit 2 100 artistes pour 18 jours de festival entre la mi-janvier et le début février). Ces événements culturels participent à la cohésion de cette communauté gaélophile, et lui donne un dynamisme sur lequel se fondent ces initiatives d'apprentissages du gaélique. Sur *Scotster* (SC), les utilisateurs font de fréquentes références au gaélique, et échangent des messages dans cette langue. De nombreux membres demandent aux plus experts de traduire des phrases en gaélique, et les autres forums sont parfois cités comme de bonnes sources pour apprendre ou pratiquer le gaélique.

2.2 Spécificités du scots

Comment le Scots est-il décrit par les universitaires, et qu'est-ce qui en fait une langue ? C'est un ensemble de pratiques discursives, lexicales, graphiques, syntaxiques et phonologiques qui ne sont pas toujours utilisées dans leur ensemble par tous les locuteurs du scots, mais ces spécificités le mettent à part de l'Anglais Écossais Standard (AES), déjà différent de l'anglais britannique standard.

2.2.1 Lexique

Le scots a commencé à se développer indépendamment de l'anglais parlé au sud à partir du 14ème siècle. Le premier texte contenant des mots en vieux scots (*Older Scots*) est *The Brus* et date de 1375. Il est probable que les dialectes qui ont composé le vieil anglais (le kentien, le saxon de l'ouest, le mercien et le northumbrien [11] se soient tous entre-influencés—sans compter les emprunts au norrois, et aux langues proto-germaniques— mais le vocabulaire du scots a conservé des mots qui n'existent plus en anglais comme *dicht* '*essuyer*', *reik* '*fumée*' [12] qui sont encore attestés dans des dictionnaires

11. Les deux derniers dialectes étant parfois regroupés sous le nom d' « anglien »)
12. La forme courante de *reik* est *reek*, ce qui pose des problèmes d'homographie en anglais.

récents comme le SDD. Le vieux norrois, parlé par les populations viking en Scandinavie et dans les régions conquises par les Vikings ainsi que le gaélique ont laissé leur empreinte dans le vocabulaire du scots, tels *kirk* (‘*église*’) du norrois *circe*, ou le très connu *whisky*, *uisge beatha*, ou ‘*eau de vie*’ en gaélique irlandais.

Les langues romanes (le français, le latin, et l'anglo-normand) ont également influencé le vocabulaire du scots pendant le Moyen Âge. Le vocabulaire législatif et religieux particulièrement, ont été influencés par ces langues, tandis que la Vieille Alliance (1295) a permis l'introduction de mots appartenant au français parisien.

Le vocabulaire du scots est en constante érosion, par rapport à l'anglais britannique. Certaines études (Macafee & McGarrity, 1999) ont prouvé que plus les locuteurs étaient jeunes, et moins ils maîtrisaient le vocabulaire du scots. En moyenne, les utilisateurs plus âgés maîtrisent 1% de vocabulaire en plus que les plus jeunes, par tranche de six années.

Les différences entre l'utilisation, et la connaissance du vocabulaire montrent que le scots est surtout une langue orale, et dont la connaissance lexicale est majoritairement passive. Tous les groupes d'âge ont montré qu'ils connaissaient au moins deux fois plus de vocabulaire en scots qu'ils n'en utilisaient. Les statistiques semblent séparer les locuteurs en deux catégories : les 20–45 ans et les 46 ans et plus, c'est-à-dire les gens nés avant ou après la seconde Guerre Mondiale.

Groupe d'âges	Ensemble des sondés	Connaissances	Utilisation	Utilisation occasionnelle à fréquente	Utilisation fréquente
20-33	47	24	10	8	5
34-45	60	30	13	11	6
45-59	82	38	17	15	11
60+	93	41	23	18	11

Tab. 2.2 – Résultats de Macafee et McGarrity, 1999 portant sur l'utilisation du scots

Les locuteurs du scots utilisent certains mots-signatures de la langue. Bien que le vocabulaire scots « pur » puisse être assez pauvre, ces locuteurs utilisent très fréquemment des lexèmes comme *wee*

(‘*petit*’) ; *bairn* (‘*enfant*’), *kirk* (‘*église*’), qui servent de repères pour les non écossais pour les caractériser facilement. Ainsi, le *Groundskeeper Willie* [13], personnage récurrent des *Simpson* qui illustre la vision américaine de l’Écossais (à tel point que la ville de Glasgow a listé Willie comme « figure locale célèbre ») utilise fréquemment des expressions écossaises comme *aye* et *nay*, ainsi que des scottismes acceptés et reconnus en tant que tels comme « the wee one » (« ”Treehouse Of Horror V” The Simpsons », 1994). Ce portrait caricatural est mal perçu en Écosse (« Groundskeeper Willie Defines U.S. Image of Scots », 2007) mais il est linguistiquement pertinent dans la mesure où il est caractéristique des représentations extérieures sur cette langue. Les autres Écossais célèbres sont caractérisés par leur accent (le chuintement des fricatives de Sean Connery, par exemple [14]), et peu par des utilisations lexicales en scots. L’altenance codique lexicale est rare, mis à part les quelques mots reconnus comme étant typiquement écossais.

En revanche, les écrits écossais ne suivent pas les mêmes règles. L’attachement à la graphie du scots a repris de l’ampleur dernièrement, bien qu’elle n’ait pas été enseignée à l’école avant le début du 21ème siècle.

2.2.2 Graphies

Authors cannot be expected to be phoneticians,
and consequently their spelling are often imperfect and inconsistent

C. Wise - *Applied Phonetics*

Le scots est écrit de façon relativement constante, grâce à des conventions graphiques apparues lorsque il est également devenue une langue utilisée dans des discours littéraires. Il a connu deux vagues de représentation graphique, avant l’annexation de l’Écosse à la Grande-Bretagne, que je vais traiter ici. Les graphies anciennes sont encore intelligibles par les locuteurs de scots actuels ; mais les écrits actuels tendent vers un certain prescriptivisme, bien qu’il n’émane pas forcément des sources

13. Un des personnages secondaires de la populaire série « The Simpsons ». Sa réplique à propos des Français dans l’épisode 22 de la 6ème saison « Cheese-eating surrender monkeys » est entrée dans le vocabulaire courant (elle est présente dans un dictionnaire de citations d’Oxford (Shorto, 2007) notamment pendant la guerre en Irak.
14. Comme nous l’avons vu au chapitre précédent, avec les imitations de Sick Boy dans *Trainspotting*

officielles (ministère de l'Éducation écossais…).

2.2.2.1 Graphie anciennes du Scots

L'exemple de l'*Historia Gentis Scotorum* (Boece, 1527) écrite d'abord en latin puis traduite en scots par J. Bellenden en 1536 (Boece & Bellenden, 1536/1941) illustre les différences graphiques entre le scots et le moyen anglais [15]. L'extrait suivant, tiré de l'histoire de Macbeth illustre assez bien les utilisations graphiques écossaises (comparées à leur traduction anglaise postérieure [16]) :

> The samyne **tyme** happynnit ane wounderfull thing. **Quhen** Makbeth and **Banquho** war **passand** to Fores, quhair King Duncan wes for the **tyme**, thai mett be the gaitt thre weird sisteris or wiches, quhilk cam to **thame** with elrege clething. The first o **thame sayid** to Makbeth ; "Hayill, Thayne of Glammys !" The saicund sayid : "Hayill Thayn of Cawdor !" The thrid sayid : "Haill Makbeth, that sallbe sum tyme King of Scotland.
>
> At the same **time**, a wonderful thing happened. **When** Macbeth and **Banquo** were **travelling** to Forres, where King Duncan was at the **time**, they met by the road three paranormally-gifted sisters or witches, who came to **them** wearing otherworldly clothes. The first of **them said** to Macbeth, 'Hail, Thane of Glamis !'. The second said, 'Hail, Thane of Cawdor !' The third said, 'Hail Macbeth who shall one day (**some** time) be King of Scotland !

Si on se place d'un point de vue anglo-centré, c'est-à-dire où la graphie de l'anglais standard est l'étalon de la graphie, les séquences graphiques <qu> et <wh> et sont remplacées par <quh> (dans des mots tels que *when*, ou *Banquo*). Certaines caractéristiques du scots sont déjà visibles, comme les participes présents se finissant en <-and> comme *passand* ; il y a également une anaptyxe avec l'insertion d'un yod intervocalique dans le mot *said*, graphié <sayid> /ˈsejɪt/, qui va à l'encontre de la prononciation du prétérit /ˈsed/en anglais. Une graphie scots caractéristique est également présente dans des mots, comme *tym*(e), *thame* ou *thaim* ('*them*'). Tous ces mots furent régularisés à nouveau avec l'anglais jusqu'au 18ème siècle, où il était considéré comme une forme rustique d'anglais du nord, plus qu'une langue à part entière, à tel point que Ramsay écrivait de façon anglo-centrée, avec des

15. Les exemples sont tirés du travail de David Purves, auteur d'un manuel destiné à l'écriture en scots, « A Scots Grammar » (Purves & Saltire Society, 2002).

16. Traduction disponible dans l'activité autour du personnage de Macbeth dans *Scuilwab*, site qui propose du matériel pédagogique en scots (Chris, 2010)

apostrophes d'excuse : « By a' the rest it is confest, / By a' the rest, that she sings best. » (My Peggy is a Young thing) (Ramsay, 1737), où les apostrophes représentent les lettres manquantes du mot anglais « all ».

2.2.2.2 Standardisation contemporaine des graphies

La graphie moderne du scots est assez bien documentée, et présentée au locuteur actuel (de langue maternelle anglaise, ou plus habitué à la graphie standard de l'anglais) comme une simple substitution de séquences graphiques de l'anglais standard. Une graphie standardisée est loin de traduire les différences entre chacun des dialectes de l'écossais, selon certains. Dans (McClure et al., 1980), les auteurs affirment que l'utilisation d'une seule graphie standard pour les mots courants ne rendait pas compte des différences dialectales : par exemple, *guid* '*good*' a plusieurs prononciations suivant la variété de scots utilisée utilisé : [ˈgɪd] à Glasgow, [ˈgiːd] à Black Isle, [ˈgʷiːd] dans le nord-est du pays, [ˈgɛɪd] dans la région de Fife, voire [ˈgød] dans la région d'Angus et aux Marches écossaises [17]. Ces différences de prononciation ne font que confirmer le fait que la graphie <guid>, si elle ne recouvre pas toutes les prononciations possibles [18], est une indication graphique au lecteur, pour lui faire comprendre qu'il s'agit non d'un mot d'anglais standard, mais d'un scottisme dont la prononciation peut être différente de sa contrepartie en anglais standard, ou de sa prononciation dans d'autres régions. Les remarques apportées par Aitken sont intéressantes, et témoignent d'un souci pour la langue, mais elles ne sont pas forcément pertinentes. Je traiterai plus loin la question de l'unicité de la graphie, et des problèmes liés aux graphies alternatives non conformes à la prononciation.

Quand les mots ont la même prononciation, et qu'il n'existe pas de graphie en scots attestée et documentée, le mot s'écrit souvent comme en anglais. En revanche, il peut aussi être homographe en anglais et en scots, et se prononcer différemment (comme aunt, /ˈant/). Les questions de traitement graphique sont souvent subjectives, et soumises à la perception du locuteur de scots comme de l'audience à laquelle il adresse son discours (celles-ci seront évoquées en 2.2.2.3).

Le tableau suivant 2.4 n'est qu'un exemple parmi d'autres, tiré des recommandations graphiques pour écrire en scots. En effet, la présence de dialectes différents (notamment le *patter* de Glasgow, le dorique parlé au nord-est de l'Écosse, la variété du sud…) entraîne des prononciations différentes, qui sont parfois transcrites dans la langue écrite. La normalisation des productions écrites, et la gramma-

17. Les East Borders (ou Scottish Borders), composées des comtés du Berwickshire, du Peebleshire, du Roxburgshire, et du Selkirkshire.

18. A cause des différences temporelles de l'élaboration de l'écrit et de l'oral, aucune langue n'a un apport strictement graphophonématique parfait, à part les langues construites (Ong, 1982).

ticalisation de certains usages graphiques et, évidemment, lexicaux, relève d'une vision diachronique basée sur les graphies du scots, avant que son usage ne soit découragé. Les dictionnaires populaires et les travaux comme ceux d'Eagle [19], sans légitimité universitaire, sont de bons exemples d'autonomisation de la langue par réduction de l'hyperlangue, définie comme toutes les pratiques subjectives d'une langue, à l'oral ou à l'écrit, et par composition de recommandations graphiques (Mazière & Auroux, 2006). Les usages graphiques sont variables ; les recommandations de l'état concernant la graphie sont très vagues (Education Scotland, 2012) ; en conséquence, certains établissent des normes fondées sur des usages déjà constatés.

Son	Graphie	Exemples en scots
/a/	<a>	*aff* (‘*off*’) ; *tak* (‘*take*’)
/ɛ/	<e>	*ken* (‘*know*’), gled (‘*glad*’)
V̆	__ C(C)#	*bak* (‘*bake*’), wid (‘*wood*’)
V̆	__ CCe#	*birse* (‘*bristle, beard*’), carse (‘*fertile land*’)
/ɪ/	<i>	*bird, brig* (‘*bridge*’), kist (‘*chest*’)
/ɪ/ -> /ʌ/	/w(h)__	*whisper*, parfois écrit *whusper*
/ʌ/	/ /	*drumly* (‘*sad*’), *hunder* (‘*hundred*’)
/ə/	<a, e, i, o, u>	En S inaccentuées : *aboot* (‘*about*’)
V̄	__#	*wi* (/wi(ː)/), be /bi(ː)/
	__ Ce	*bane* (‘*bone, limb*’), hure (‘*whore*’)
	<au, aw, ae, ai, ay, ea, ee, ei, ie, oa, oo, ou, aa>	*Faw* (‘*fall*’), snaw (‘*snow*’), bairn (‘*child*’), dreich (‘*bad weather*’), hoose (‘*house*’), haar (‘*cold wind*’)
/ɑː/, /ɔː/, /a(ː)/	<au> init. ou med.	*auld* (‘*old*’), *cauld* (‘*cold*’), *haund* (‘*hand*’)
/ɑː/	<all> en ang. devient <aw> ; <aa> (Nord, scots insulaire	*aw* (‘*all*’), *caw* (‘*call*’)
	<a> #	*awa* (‘*away*’), twa (‘*two*’), wha (‘*what*’)
/ɛ/	<aCe>	*ane* (‘*own*’), ance (‘*once*’), hale (‘*haul*’)
	<ae> en position finale	*ae* (‘*one*’), brae (‘*hillside*’), *frae* (‘*from*’)
/eː/, /ɛ/	<ai>, surtout avant /r/	*aiple* (‘*apple*’), *airm* (‘*arm*’), *cairt* (/e/ ‘*cart*’)
	<ay>	*day, gray*
/i(ː)/, /e(ː)/	<ea>	*beard, beast, read*
	<ei> (/__ d#)	*heid, deid, leid, peir, weir* (quand ce sont des collocations de l'anglais)

Illustration 2.4 – Correspondances phonographématiques principales anglais-scots

19. Eagle est un amateur de langue écossaise, et tient un site fourni à ce sujet, mais il n'a pas de légitimité universitaire.

Son	Graphie	Exemples en scots
/ɛ/	<ea> ; <e>	*read*
/i(ː)/	<eCe>	*here*
	<ee>	*creep, feel*
	<ie> /__ <V> (sauf à l'initiale)	*chief, chield* (‘*young man*’), life (‘*beloved*’)
	<ie> /__ <l>	*shield* (quand ce sont des collocations de l'anglais)
	<ei> /__ /x/	*dreich, wheinch*
/e/ Centre	<y>	stany (‘*stony*’), *brawly* (‘*excellently*’), *sairly* (‘*badly*’)
/iː/ ; /ɪ/ (Nord, east Perthshire)		
/oː/	<oa>	Les deux phonèmes ont fusionné dans certains dialectes, mais si les graphies sont différentes, il faut dissocier les prononciations
/ɔ/	<o>	*hoast* ; *loan* (‘*lane*’) ; Dochter (‘*daughter*’), loch (‘*lake*’)
/u(ː)/	<ou> ; <oo>	*hoose* (‘*house*’), *soond* (‘*sound*’), *aboot*
/ʌu/		En position finale, dans le sud de l'Écosse
/u(ː)/	<uCe>	Dans les graphies attestées : *hure*, *dule* (‘*grief*’)
/ø(ː)/, /ɪ/, /e(ː)/, /i(ː)/	<ui>	*guid* (‘*good*’), *ruif* (‘*roof*’), *tuim* (‘*empty*’)
/ø(ː)/, /(j)u/, /(j)ʌ/	<eu>	*beuk* (‘*book*’), *deuk* (‘*duck*’), *eneuch* (‘*enough*’)
/ju/	<ew>	*Dew, few, new*
	Diphtongues	
/əɪ/, /aɪ/	<iCe>	*wife, knife, lice*
	<yCe>	*dyke* (‘*a stone hedge*’), *syne* (‘*next*’), *tyne* (‘*lose, fail*’)
/i/	<y>	*wynd, mynd*
/əɪ/, /aɪ/	<y# >, <ye# >, <ey# >	*aye* (‘*always*’), *ay* (‘*yes*’) *cley* (‘*clay*’)
	<# ey>	*eydent* (‘*industrious*’), eyster (‘*oyster*’)
/ʌu/	<owe>	*growe* (‘*grow*’), *thowe* (‘*melting snow*’)
	<ow>	*howff* (‘*enclosed yard, shelter*’)

Illustration 2.4 – Correspondances phonographématiques principales (suite)

Les diminutifs en scots sont formés avec le suffixe <-ie>, beaucoup plus employé qu'en anglais.

Ce morphème a une valeur affective dans les mots *grannie* (‘*grand-mère*’), *laddie* (‘*jeune homme*’) et *lassie* (‘*jeune femme*’), accompagné du redoublement de la consonne finale. Le nom ou adjectif *bonny/ie*, que l'on trouve dans l'expression « bonnie Scotland » [20] possède les deux graphies *bonny* et *bonnie* en anglais standard comme en scots. Il est très peu fréquent en anglais, et seules 199 occurrences de *bonnie* ont été trouvés dans le *British National Corpus* (BNC), contre 138 pour la graphie *bonny*, qui est plus fréquemment un adjectif qu'un nom propre. Ces mots sont souvent utilisés dans des noms propres (références à Bonnie et Clyde, ou adresses postales, voire à des personnages historiques comme le *Bonnie Prince Charlie*), et la graphie *bonnie* est privilégiée. De la même façon, cette terminaison peut être recatégorisante (comme le <-y> anglais) et transformer un nom ou un adjectif en adverbe. Un autre exemple est le surnom d'Édimbourg, *Auld Reekie*, le diminutif du nom *reek* (‘*fumée*’ en scots). Le verbe *yeuk* (‘*gratter*’) est recatégorisé en adjectif avec la terminaison <-y> (*yeuky*) mais sa graphie alternative « yuke » est recatégorisée en *yukie*.

La terminaison des bases verbales qui se finissent en <ie> se transforme en <y> lorsqu'ils deviennent des participes présents, comme *courie/ couryin* (‘*se pencher, se recroqueviller*’). Cette graphie, inspirée de la graphie anglaise (où les bases verbales ne finissent pas en <-ie>) est certainement plus facile à écrire que <*couriein> [21] ; et à penser pour des locuteurs de l'anglais. En outre, bien qu'aucune transcription phonétique ne soit disponible pour *couryin*, on peut penser que la prononciation ressemblerait à /ˈkuːrɪn/ (ou /ˈkɔrjn̩/ ; /ˈkʌrjn̩/ suivant les lieux). La palatalisation et l'introduction d'un yod, représenté par <y> graphiquement, semble nécessaire ; sinon, la prononciation ressemblerait à /ˈkuːrɪ(ʔ)ɪn/, et cette introduction d'une pause (voire d'une occlusive glottale) semble peu probable. Certains auteurs utilisent donc une seule graphie pour des raisons de cohérence.

20. de l'adjectif français « **bon**. »
21. une métathèse graphique fréquente en anglais standard est d'inverser les <e> et les <i> dans la séquence <ie>, comme « receive/*recieve ».

2.2.2.3 Typologie des écarts à la norme

La norme graphique est un concept sur lequel il y a peu de consensus en scots, mais les actants participant à l'élaboration d'une graphie du scots s'accordent plus ou moins à dire qu'il faut se dégager des conventions graphiques de l'anglais standard. En revanche, leurs points de vue divergent quant aux buts à atteindre : retranscrire la prononciation personnelle, s'éloigner simplement de la langue dominante…

Il existe tout de même un désir fort de s'éloigner des graphies de l'anglais standard : certains des mots en scots sont volontairement écrits différemment pour qu'aucune confusion ne soit possible (*aboot, hoose*) [22]. Je nomme ces graphies « graphies démarcatives », puisque leur but est de tracer une ligne de démarcation avec l'anglais, et qu'elles ne sont fondées que sur un désir d'éloignement des pratiques normées anglaises, et ainsi montrer la différence de prononciation avec l'Anglais Standard du Sud (GB) (ASS). Le terme « graphies discriminantes », autre candidat pour désigner cette caractéristique, me semblait trop chargé de connotations négatives, puisqu'il ne s'agit pas uniquement de discriminer le standard de la langue régionale.

Les graphies démarcatives montrent un réel effort non seulement pour se différencier de l'anglais standard, mais aussi pour faire émerger des standards personnels dans un scots qui en a parfois besoin.*The Scots Language* (McClure et al., 1980), ouvrage décrivant le futur de la langue de façon didactique, est d'ailleurs parfois pessimiste. Il semble que, dernièrement, des efforts pour faire à nouveau entrer le scots dans la catégorie des langues à part entière ont été réalisés, notamment avec la ratification du *Language Act* en 2005, ce qui éloigne le scots de la domination de l'anglais standard. Si les mots fréquents (qu'ils soient grammaticaux ou lexicaux) sont représentés de façon alternative, on s'éloigne de la caricature du scots pour s'approprier entièrement la langue. La graphie démarcative, même si elle ne représente pas toujours une différence de prononciation par rapport à l'anglais, code la possibilité d'une prononciation dialectale potentielle. Elle tend à harmoniser des textes en scots, en homogénéisant les graphies de mots contenant des prononciations identiques en scots.

Certains auteurs ont recours à la graphie en anglais normé, à condition que le mot soit prononcé

22. Dans certains guides d'écriture du scots, notamment(Eagle, 2001a).

de la même façon dans les deux langues (Purves & Saltire Society, 2002). Le traitement grapho-phonémique du scots est parfois problématique, comme dans ces recommandations, puisqu'il n'y a souvent pas de transcription phonétique pour soutenir les propos. Elles peuvent donc ressembler à des guides de conversion entre la graphie anglaise et la graphie écossaise, et être peu représentatives des systèmes phonétiques sous-tendant le scots.

Le ministère de l'Éducation écossais [23] a émis une feuille de style de recommandations graphiques pour écrire en scots dans un cadre scolaire (Education Scotland, 2012) arguant que l'absence de graphie standard permet aux enfants de développer leurs capacités scripturales (dans le sens où ils doivent réfléchir à leur prononciation pour écrire en scots). Les seules directives sont la cohérence graphique, et d'éviter l'utilisation d'apostrophes « paroissiales » [24].

Les apostrophes « d'excuse » pour signifier l'absence de lettres en scots, par rapport à l'anglais standard (dans des cognats comme *wi* '*with*') sont apparues au 18ème siècle. Le scots est la seule langue dans laquelle ces apostrophes sont apparues après la standardisation de la forme écrite. De nos jours, elles semblent être un effort conscient de hiérarchiser le scots en tant que sous-variété de l'anglais standard, et sont découragées, voire proscrites depuis la seconde moitié du 20ème siècle de façon unanime dans les guides d'écriture du scots. Des médias comme *Wikipedia* ont édité un guide d'écriture en scots (« Wikipedia », 2012), afin de permettre une standardisation de l'écriture. Celle-ci est fixée de façon prescriptive plutôt que descriptive par les administrateurs du site, et justifié ainsi :

> Here at Wikipaedia it's recommendit that fowk uises 'tradeetional'pan-dialect spellins. Awtho thir isna sae strict as in Inglis we ettle tae come up wi writin that's easy tae read an can be soondit bi readers in thair ain dialect [25].

Cette recommandation sert à uniformiser les diverses graphies issues des dialectes du scots et de n'en privilégier aucun. C'est une sorte de juste milieu entre une revendication exprimée au travers d'une graphie démarcative, et une liberté qui pourrait parasiter la compréhension si les représentations graphiques deviennent trop personnelles. En outre, les mots argotiques sont prohibés. *Wikipedia*

23. Sous la direction du Dr Alasdair Allan, en charge de l'éducation, des sciences et des langues d'Ecosse depuis 2011. Il est à noter qu'il a obtenu un doctorat ès langue scots en 1998.
24. Traduction de « apologetic/parochial apostrophe »
25. '*Here at Wikipedia, it is recommended that [people] use traditional pan-dialectal spellings. Although this isn't so strict as in English, we intend to come up with writing that is easy to read and which can be pronounced by readers in their own dialects*'(Traduction personnelle).

accompagne sa page « Spellin an grammar » de liens vers les feuilles de style et diverses recommandations académiques déjà évoquées dans cette partie.

Cette politique linguistique ressemble à celles d'autres langues minoritaires, comme le wallon : « certains usages graphiques à caractère phonétique sont parfois insérés dans les démarches adaptatives ou même dans quelques propositions orthographiques plus récentes » (Eloy, 2004). Pour écrire dans une langue minoritaire, il faut adapter la graphie de la langue dominante qui en est proche. La domination de la langue principalement parlée dans le pays se voit aussi dans les usages graphiques, et c'est cette langue enseignée à l'école qui sert de base pour les graphies du scots, qui ne sont pas uniformisées.

2.2.3 Syntaxe

La syntaxe du scots est comparable à celle de l'anglais standard. Le terme « grammaire » est d'ailleurs utilisé de façon abusive dans certains manuels de scots, et montrent la méconnaissance grammaticale de certains auteurs. La principale différence entre le scots et l'anglais standard est morphologique. Pour éviter la « fiction linguistique », les exemples de phrases sont tirés de plusieurs sources : les travaux de Stuart-Smith (Stuart-Smith, 2003), le corpus écossais de textes et de paroles ou *Scottish Corpus of Text and Speech* (SCOTS), '(Corbett et al., 2007) ainsi que des fils de discussion de *Scotster*.

2.2.3.1 Particularités verbales

Certains verbes ont gardé des désinences fortes alors qu'ils ont été régularisés en anglais standard ; certains verbes peuvent être forts en anglais standard et en scots, mais avec des formes inversées comme *see-seen-saw*, où le prétérit et le participe passé inversent leurs distributions en scots. Ainsi « A've saw that luk oan is face afore » (SCOTS doc 835) '*I've seen that look on his face before*'est correct, autant que « he seen the horrors o the Nazi camps » (SCOTS doc 164). Il existe une tendance en scots à utiliser le participe passé comme le passé, ce qui explique cette inversion (Corbett et al.,

2003).

Le verbe *kill*, dont le passé est *killt*, est resté régulier en anglais standard, mais ce verbe a conservé sa forme de *kill* + <-ed>, devenu, par amuïssement du /d/ à cause de la proximité avec le son /l/, la forme utilisée actuellement en scots. Certains verbes irréguliers en scots et en anglais standard comme *sell*, dont les formes passées et au participe passé sont respectivement *sellt* et *sauld* sont construits sur le même modèle que *kill* (Jones, 1997). L'anglais standard est passé à l'utilisation de désinences <-(e)d>, le scots garde une image de la régularité du verbe : la variante *sell, selled, selled* existe également, et le scots a conservé en partie une trace de l'anglicisation irrégulière de ce verbe dans la forme *sauld*.

Le système modal du scots est un peu différent de celui de l'anglais. *Shall, may* et *ought to* n'existent pas en scots (mais on en trouve des occurrences dans l'anglais écossais standard) et ils sont remplacés par d'autres verbes :

— *Shall* de permission « you shall go out tonight » est remplacé par *can, get to*, et *get* + gérondif : « you get going out tonight », « you get to go out tonight ».

— *Ought to* est très peu utilisé en scots, et est remplacé par *should* ; *want* recouvre aussi ce sens : « you ought to do something else » → « [You're kidding yourself and it's not like you.] You want to do something else. » (SCOTS doc 993).

— *Must* n'est utilisé que dans son sens épistémique « you must be tired », et peut être utilisé dans la reprise elliptique en fin de question : « It must have been a kind of sanatorium type camp, mustn't it ? » (SCOTS, doc 1432) et *mustn't* a le sens de *je conclus que ce n'est pas [X]* : « I mustn't have read the question properly » (p. 90 in Corbett et al., 2003). L'obligation n'est exprimée que par *have to* ou *need to*. L'obligation exprimée en *have to* est moins forte que celle de *have got to*, et *will have to* exprime une obligation plus faible (parfois émanant du sujet) : « I will have to go visit grandma (because it has been 6 months and I want to) ». On peut également l'exprimer en utilisant des tournures en *supposed to/meant to* « No. It's nae meant to be for eating » (SCOTS doc 1579) ; « But you're the shop lady. You're supposed to sell it to me. » (SCOTS doc 1582).

— *Need* a, comme en anglais standard, une valeur semi-modale, mais il n'est pas utilisé non plus en auxiliaire de question. La forme *needn't* n'existe pas en scots. Il n'exprime que l'obligation (émanant du sujet ou de l'extérieur) « You need to move that out of the way just now in case you trip over » (SCOTS, doc 1642).

— Les doubles modaux sont autorisés en scots : « This will maist like be the last year it will can perform at Riddles Coort » (SCOTS doc 1803) ; « she might can get away early ». *Might* est aussi compatible avec *would* et *should* . Ces séquences sont cependant en net recul (peu ou pas d'occurrences dans le SCOTS), surtout chez les jeunes. En revanche, elles sont toujours plus pratiquées dans les Orcades et par les locuteurs âgés (plus de 60 ans).

Les propositions résultatives sont également un peu différentes en scots. Elles sont appelées « clivées inversées »(‘*reverse cleft*’) (Corbett et al., 2003) et sont formées sur *that's/that was*, un syntagme nominal et un verbe au participe passé à valeur résultative optionel ; comme « that was me finished » (SCOTS doc 27) ‘*I was done* ’, ou même « We played some video games an that was us » (SCOTS doc 1384) ‘, exprimant le fait que l'activité marque la fin de la soirée.

Les interrogatives en scots, sont légèrement différentes. *How* remplace *why* « How did you do that next ? [26] » (SCOTS, doc 1547). L'adverbe *Whereabouts* [sic] remplace fréquemment *where* « So whereabouts in Edinburgh you stay anyway ? » (SCOTS doc 1386). Il est également possible d'en séparer les morphèmes : « Where in Edinburgh do you stay about ? ». L'expression *what time…at ?* remplace également *when*, et *what* couvre les deux sens de *what* et *which* : « what word was Aald Peter misusing, or where did he get the idea to produce his new word ? » (SCOTS, doc 1391).

2.2.3.2 Particularités nominales

Les pluriels en scots sont formés différemment, avec l'adjonction d'un morphème <-(e)n> : *tree* devient *treen* dans certaines formes archaïques. De la même façon que les noms peuvent être suivis de ce suffixe pour former des adjectifs de matière (comme *wood* → *wooden*), la forme *treen* peut également signifier ‘*fait de bois*’ . Les pluriels irréguliers de l'anglais, comme *wife* → *wives* ne sont pas

26. Dans le contexte, le *how* remplace un *why*. Ceci également visible avec l'adjonction de l'adverbe *next*.

utilisés en scots. Ainsi, *knifes, thiefs, dwarfs, loafs* ou *wolfs* contredisent les règles de voisement des fricatives quand le morphème <-s> les suit.

Les pluriels de l'anglais marqués par <(e)n> comme *child* → *children* sont, quant à eux, remplacés par le suffixe <-er> (*child* → *childer*). Ce suffixe, régulier en moyen anglais, est perdu en anglais standard, qui a ajouté au pluriel en <-er> un pluriel en <-en> [27], qui a rajouté à nouveau le suffixe <-en> de pluriel sur un mot au pluriel, avec une métathèse qui explique la forme en anglais moderne *children*. Les mots anglais qui ont soit un pluriel en <-en> comme *ox* (‘*bœuf*’), *brother*, voire qui utilisent la désinence <-en> pour signifier une plus petite version de la chose comme dans *cat* → *kitten* n'existent pas en tant que tels en scots. La forme fréquente de *brother* est *billie* (plur. *billies*) mais il existe tout de même la forme *brither* (‘*frère, égal*’) → *breither* prononcé ’brɪðər,’bridər] au singulier et [’bre :ðər, ’bri :ðər, ’bredər] au pluriel.

Maid a un pluriel régulier tout comme *maiden*. La forme en scots de *kitten* est *kittlin*, qui semble être un emprunt du français [28] datant du 14ème siècle. Le scots a ainsi le verbe *kittle*, ‘*donner naissance à des chatons*’. Il existe également une forme en *cat* → *cattie* pour désigner un petit chat. L'anglais et le scots ont donc des stratégies différentes pour traiter la suffixation en <-en>.

Le scots a également une forme de pronom personnel différent, pour la seconde personne du pluriel, comme certaines variétés de l'anglais. La forme de ce pronom peut être *yous*, ou *yous yins* [29] Ce pronom est utilisé dans des discussions informelles et par des locuteurs peu éduqués du scots. De la même façon, *us* se substitue à *me* quand il est employé avec des verbes comme *show, give, lend* (où on attend une interaction entre l'énonciateur et le coénonciateur). Les pronoms réfléchis *himself* et *themselves* → *hisself* et *theirselves* rétablissent, finalement, le rapport entre *I, my* et *myself*, qui est irrégulier dans *himself* et dans *themselves* [30].

Une particularité bien connue du scots, dans l'utilisation de l'article défini, est son utilisation plus fréquente, notamment devant des périodes de temps, les noms de certains lieux, et de maladies, où il n'y aurait pas d'article en anglais standard. *The now* a un sens plus large que *now*, puisqu'il peut

27. Ce pluriel en <-er> rappelle le pluriel de l'allemand *das Kind* → *die Kinder*.
28. Forme en vieux français : *chitoun*
29. Littéralement ‘*you ones*’.
30. Rapport que l'on peut trouver également dans la prononciation /me’səlf/, /mi’səlf/ qui n'est en fait que la monophtongaison de la première syllabe de myself.

désigner l'actuel au temps de la narration.

(1) « l'll do it **the now**… » (*Scotster*, 10ème message de SC3S1)
'*I'll do it **now***'

(2) « They didn't live for **the now** » (SCOTS, doc 547)
'*They didn't live for the moment*'

(3) « Did you hae a piece **the morn** ? » (SCOTS, doc 1594)
'*Did you have a piece this morning ?*'

(4) « For **the day** the wither is fair [31] » (SCOTS, doc 1416)
'*today, the weather is fair*'

(5) « it seemed to be the time […] where they were chucking out the guys that had been kept overnight in **the jail**. » (SCOTS doc 354). [= in jail]

De la même façon, les personnes et lieux très spécifiques peuvent être précédés non d'un pronom possessif, mais de l'article *the* : '*the lassie*', '*the bairns*' sont, implicitement dans certains contextes, la femme et les enfants de l'énonciateur. On peut également utiliser *the* devant les quantifieurs *most*, *hale* ('*all*'), *yin* ('*one*') et *both* :

(6) « the Old Dear's fussing plenty for **the both** of us » (SCOTS doc 596)

(7) « schools are shut, every one, in **the hale** o Moray and **the hale** o Aberdeenshire. » (SCOTS doc 1630)

31. Prévision météorologique d'Alex Smith

2.2.3.3 Prépositions et adverbes

(8) « Ghosts stae pit ; in the yin place…place where they **cannae** rest » (SCOTS doc 597) '*Ghosts stay put in the one place …place where they can't rest*'

(9) « There's **naebody** that wants the flu » (SCOTS doc 812)

La négation écossaise est un des points cruciaux utilisés pour parler de la langue : en effet, elle s'exprime, comme en anglais standard, par l'utilisation de l'adverbe *not* après le verbe, mais également par *no* et *nae*, adjoint aux verbes modaux ou auxiliaires : *cannae* ('*can't*') *dinnae* ('*don't*') rivalisent avec les autres formes de négation. Il est possible de donner une réponse courte, avec une reprise elliptique en choisissant la forme de négation que l'on préfère : « A : I could. B :Can no » (SCOTS doc 1550).

Si *no* est utilisé après un modal, il porte sur le syntagme verbal suivant :

(10) Ye **cannae** go to the party.

(11) Ye **can no** go to the party.

L'exemple 10 signifie « tu n'as pas le droit/tu ne peux pas aller à la fête » alors que 11 veut dire « tu peux ne pas aller à la fête ».

Le système prépositionnel écossais est un peu différent de celui de l'anglais. De plus, toutes les prépositions en scots ont leur graphie propre. Dans les tournures passives, la préposition *fae* ('*from*') est beaucoup plus utilisée que *by*. Les prépositions *off* ('*of*') et *wi* ('*with*') sont également utilisées en passif. *Off* (et non *from*) exprime généralement la source dont émane quelque chose ; la cause est exprimée par *from/fae* :

(12) « I was petrified **fae** worms ! » (SCOTS doc 27)
'*I was scared of worms*''

(13) « tea that haed juist been loaded **aff** a ship » (SCOTS doc 1770)
'*tea that had just been loaded off from a ship*'

(14) « his een got yaised **wi** the licht » (SCOTS doc 63)
'*his eyes got used to the light*'

(15) « Mam wisna ower chuffed **wi** a boquay in May.[32] »

'*Mom wasn't over-joyed by getting a bouquet in May*'

De la même façon, les autres prépositions, notamment celles de lieux (*down, up, in, out*) ne requièrent pas l'utilisation de *to*, après des verbes de mouvement : « My friend from down the street », mais il existe des formes telles que « I can see myself laying down the little roaring bundle on the futon » (SCOTS, doc 993), ou « He staggered up the bus » (SCOTS doc 1650).

Certains verbes prépositionnels changent légèrement en scots : *shout on someone* '*shout to*' comme dans l'exemple « this man was sort of shouting on us like, "Your child is like [laugh] //drowning" » (SCOTS doc 1575) ; through the post '*by post*' « somebody receiving an epistle through the post that was sparse in information » (SCOTS doc 1239) ; *think on something* '*think of/about*' : « Dinna think on the dule o the warld ! » '*don't think about the sorrows of the world*' (SCOTS doc 921) ; *wait on*[33] *wait for* « How long do I have to wait on you turning up » (*Scotster*, SCPP3). Les différences entre le scots et l'anglais sont systématiques, et posent des problèmes de compréhension et d'identité, notamment lors de la recherche de travail. Les recherches sur les particularités grammaticales du scots n'en sont qu'à leurs balbutiements, et n'ont commencé dans les années 1970. Jim Miller souligne le fait que les informants du corpus ECOSSE (Edinburgh Corpus of Spoken Scottish English) ont une cinquantaine d'années, et ceux du MTC (Map Task Corpus) ont une quarantaine d'années maintenant, mais que peu d'études sont faites sur les particularités grammaticales du parler des jeunes (Corbett et al., 2003).

32. Accessible à l'adresse <http ://www.scotster.com/today/judportia13.56897.html>. Dernier accès le 10/07/2013
33. Qui rappelle la structure de ce verbe (attendre + sur) des langues germaniques, notamment l'allemand *warten auf* et le néerlandais *wachten op*

Des efforts sont réalisés pour que la grammaire du scots soit comprise par ses locuteurs, et qu'ils comprennent que le scots n'est pas une forme abâtardie de l'anglais (Robinson & Dewar, 2012).

2.2.4 Phonologie

Les travaux de phonologie sur le scots ont surtout été réalisés par J. Stuart-Smith, qui a travaillé sur les relations entre le scots urbain (notamment de Glasgow) et les dialectes du scots.

Certains sons sont propres au scots, comme la fricative vélaire dévoisée /x/, présente en anglais uniquement à cause des emprunts au scots. Une certaine tendance à produire des t glottalisés (le son /ʔ/) après une voyelle ou en position intervocalique est également observable, comme dans les mots Scotland /ˈskɒʔlənd/ ou *jotter*, '*notebook*'(/ˈʤɔʔə/). La voyelle /ɤ/peut également se trouver à la place d'un /ɬ/ , en position finale pour figurer le digraphe <ll>.

2.2.4.1 Voyelles du Scots

Les monophtongues suivantes sont rencontrées en scots :

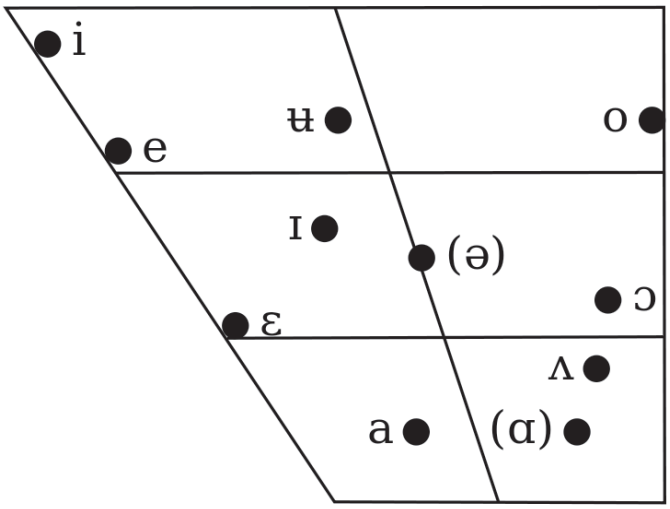


Illustration 2.5 – Voyelles en Scots (Scoobie, Gordeeva & Matthews, 2006)

Le tableau suivant est la comparaison des voyelles des classes d'isophones de Wells avec d'une part l'anglais écossais standard, et d'autre part, le scots urbain (tel qu'il est parlé à Glasgow). Il est centré sur les classes d'isophones de l'anglais écossais standard.

CLASSE D'ISOPHONES	ANGLAIS ÉCOSSAIS STANDARD	SCOTS URBAIN	CLASSE D'ISOPHONES	ANGLAIS ÉCOSSAIS STANDARD	SCOTS URBAIN
KIT	ɪ~ë	ë ~ë ~ʌ ~ɪ	CHOICE	ɔe	ɔe
DRESS	ɛ	ɛ	MOUTH	ʌu̯	u ~ʌu
TRAP	a	a	NEAR	i	i
LOT	ɔ	o ~ɔ	SQUARE	e	e ~ɛ̣
STRUT	ʌ	ä	START	a	ɛ̣ ~a̱
FOOT	u̯	ë ~u ~ʏ	NORTH	ɔ̣	o ~ɔ
BATH	a	a	FORCE	o	o
CLOTH	ɔ̣	o ~ɔ	CURE	ju̯	ju
FLEECE	i	i	HEAD	ɛ	i ~ɛ̣
FACE	e	e	AFTER	a	ɛ̣ ~a̱
PALM	a	a̱	NEVER	ɛ ~ë	ë ~ë ~ɪ
THOUGHT	ɔ̣	ɔ	STAY	e	əi~e
GOAT	o	o	STONE	o	e~o
GOOSE	u̯	u ~ʏ	STAND	a	ɔ̣~a̱
BIRTH	ɪ	ë~ä	OFF	ɔ̣	a̱ ~ɔ
BERTH	ɛ	ɛ̣ ~ë	DO	u	e~u
NURSE	ʌ	ä~ë	HAPPY	e	e~ë
PRICE	ʌi	əi	LETTER	ɪ~ʌ	ä
PRIZE	ae	ae	COMMA	ʌ	ä

TAB. 2.3 – Réalisation des voyelles d'après STUART-SMITH, 1999

Après ce recensement vocalique axé plus précisément sur le rapport entre l'AES et le scots urbain, il est possible de refaire la même opération du point de vue des voyelles du scots lui-même. Cette fois, le scots urbain théorique sera comparé à des relevés d'après des enregistrements (STUART-SMITH, 2003). La dernière colonne donnera tout de même des informations relatives à l'AES.

Les variantes des phonèmes sont marquées par une flèche bilatérale <⟷>.

Toutes ces données ne prennent en compte que la dichotomie partielle entre le scots urbain et l'anglais écossais standard. Elles ne rendent pas compte des différences dialectales au sein du scots même. Le dernier tableau des voyelles du scots (2.4) montre les différences entre les réalisations des mots des classes d'isophones entre le scots du centre, du sud et du nord de l'Écosse. Les relevés de

CLASSES D'ISOPHONES UTILISÉES POUR CETTE ENQUÊTE	SCOTS URBAIN	SCOTS URBAIN (RELEVÉS)	ANGLAIS STANDARD D'ÉCOSSE
MEET	i	i	i
BEAT	i	i	i
(DEAD)	i	i⟷ɛ̣	ɛ
MATE	e	e	e
(BOTH)	e	e⟷o	o
BAIT	e	e	e
PAY	əi	əi ⟷ e	e
BOOT	ë	ë ⟷ u	u̱
DO	e	e⟷ u	u̱
BIT	ë	ë	ɪ
BET	ɛ̣	ɛ̣	ɛ
OUT	u	u ⟷ ʌu	u ⟷ ʌ u
COAT	o	o	o
COT	o	o * ⟷ ɔ	ɔ̣
OFF	a̱	a̱⟷ ɔ	ɔ̣
CAT	a̱	a̱	a
(LONG)	a̱	a̱ ⟷ ɔ	ɔ̣
(WASH)	a̱	a̱ ⟷ ɔ	ɔ̣
HAND	ɔ	ɔ ⟷ a	a
START	ɛ̣	ɛ̣	a
CAUGHT	ɔ	ɔ	ɔ
(SNOW)	ɔ	ɔ ⟷ o	o
CUT	ä	ä	ʌ
(PULL)	ä	ä ⟷ u	u̱
NEW/DEW	ju	ju	j u̱
BITE	əi	əi	əi
TRY	ae	ae	ae
EYE	i	i ⟷ ae	ae
LOIN	əi	əi ⟷ oe	oe
VOICE	oe	oe	oe
LOUP 'JUMP'	ʌu	ʌu	(ʌu)

TAB. 2.4 – Voyelles et réalisations vocaliques du scots, comparé à l'anglais écossais standard

scots urbain sont ceux de J. Stuart-Smith (STUART-SMITH, 2003), et elle s'est servie de ses propres classes d'isophones pour cette étude.

2.2.4.2 Comparaison entre les différentes variétés de scots

Le scots est subdivisé en plusieurs dialectes, parlés dans diverses régions de l'Écosse, bien que les auteurs aient préféré écrire dans une forme centrale, un standard littéraire écossais qui ne reflétait que

très rarement ces différences (Kay, 1986). Kay en recense ainsi sept variétés : scots central (subdivisé en central, ouest-central, est-central), scots du sud, scots du nord, et scots insulaire, ainsi que la variété de scots parlée en Irelande, le scots d'Ulster.

Le scots central est parlé entre l'Angus et Galloway (jusqu'au scots d'Ulster) et les différences entre les scots centraux semble être plus volontiers une question d'accentuation que de dialecte. Le scots de Glasgow est un peu différent, et contient des diphtongues plus fréquentes (contre des allongements vocaliques dans les autres variétés de scots centraux).

Le scots du sud, parlé dans les régions de Roxburgh, Selkirk et le Dumfriesshire de l'est, est surnommé « dialecte de *yew an mey* » à cause de sa plus grande fréquence de diphtongues que dans d'autres formes du scots. Le scots du nord est parlé de l'Angus de l'est jusqu'à Caithness, et diffère grandement du scots central, surtout dans la réalisation des voyelles, à cause d'isolats où les pratiques linguistiques ont contribué à conserver un grand nombre de mots issus du scots du moyen-âge. Les diphtongues /ʊu/, /əu/ sont également beaucoup plus communes, et ceci est peut-être dû à l'isolement de ces scottophones au milieu de zones où le gaélique est parlé. Quant au scots insulaire, il est d'autant plus différent qu'influencé par le norrois (cf 2.2.1) et le norn ; les insulaires ont gardé un vocabulaire sensiblement plus influencé par le scandinave que les autres dialectes du scots.

À part l'aspect urbain du scots, et ses réalisations parfois différentes, j'ai choisi de ne garder que trois subdivisions de dialectes : le scots central, le scots du sud et le scots du nord (qui regroupe également le scots insulaire de B. Kay). Le tableau suivant montre la différence de réalisation entre les classes d'isophones dans ces trois grands dialectes du scots.

Les différences sont surtout remarquables entre le scots du nord et les autres variétés, étant donné que ces relevés incluent les variétés du scots parlées dans les Orcades et les îles Shetland. En revanche, les classes d'isophones bait, new, dew, meet, mate, out, coat divergent peu entre les différents dialectes. Cat est réalisé à peu près de la même façon dans les trois variétés, mais la fréquence des réalisations est différentes.

Classes d'isophones	Scots central	Scots du sud	Scots du nord
MEET	i	i	i, ɪi
TREE	i	ɛi	i, ɪi
BEAT	i	i	e, ɛi, i
MATE	e	e	e, i
BAIT	e	e	e
BOOT	ë	ë	i, e
DO	e	e	iː, ɪi, eː
BIT	ë	ë	ë, ɪ, ɜ
BET	ɜ̣	æ, a	e~ɛ
OUT	u	u	u, ü
COW	u, ʌu	ʌu	ʊu, u
COAT	o	o	ou
COT	o	o	ɔ, ɑ
CAT	a, ɑ, ɒ	ɒ, ɑ, a	ɑ, ɒ, ɔ, a
CAUGHT	ɔ	ɒ, ɑ, ɔ	ɑ, ɒ, ɔ, a
CUT	ʌ	ʌ	ʌ, ɐ, ɜ, ɔ
NEW	ju	ju, iu	ju
DEW	ju	ju	ju
BITE	əi, ëi	əi, ëi	ʌi, ɛi, əi
TRY	ae	ae, ɐe	ɑe, ae ~ɐe
LOIN	əi, ëi	oe	ʌi, ɛi, əi
VOICE	oe	oe	ʌi, ɛi, əi, oe, ɒi
LOUP 'JUMP'	əu, ʌu	əu	ɛʏ, əu

TAB. 2.5 – Différences dialectales dans la prononciation des classes d'isophones(STUART-SMITH, 2003)

2.2.4.3 Consonnes du Scots

Les consonnes suivantes sont présentes en scots :

Peu de consonnes changent entre le scots et l'anglais standard. Certains phonèmes consonantiques

	Biblabial	Labiodental	Dental	Alvéolaire	Post-alvéolaire	Palatal	Vélaire	Glottal
Occlusives	p^h b̥			t^h d̥			k^h g̊	(ʔ)
Fricatives		f v	(θ ð)	s z	ʃ ʒ	ʍ	x	h
Affriquées					ʧ ʤ			
Approximantes	w			ɻ	ɹ	j		
Battues				ɾ				
Nasales	m			n			ŋ	
Liquides				l ɬ		ʎ		

TAB. 2.6 – Répartition des consonnes (KORTMANN, UPTON, SCHNEIDER, BURRIDGE & MESTHRIE, 2008)

disparaissent *with* > *wi*, ou *give* > *gie*. Il existe un phénomène similaire au *TH-fronting* de l'AAV en

scots. Ceci est attesté dans la forme de certains mots comme *youdith* (‘*jeunesse*’) ou *youdfu’* (‘*youthful*’). En outre, dans le scots du nord, particulièrement dans les îles, le son /ð/ est réalisé [d] dans les pronoms personnels encore usités, comme *thy*, *thine*, réalisés [ðɛi], [ðɛin].

Les occlusives sont moins aspirées en anglais écossais standard et en scots. Les dentales et alvéolaires peuvent être produites librement entre ces deux points d’articulation. Certaines graphies comme <brotekin>, variante de *brodequin* (‘*bottine*’) suggèrent également un dévoisement de certaines consonnes en position intervocalique.

Le phonème /t/ est parfois produit comme un coup de glotte (t glottalisé), en position non-initiale (et surtout intervocalique) dans des mots comme *Scotland* /ˈskɒʔlʌn/. Ce trait est parfois exagéré par les imitateurs car c’est un stéréotype du parler du scots urbain, notamment dans la région de Glasgow. C’est un marqueur social, qui est plutôt réalisé par les gens d’une classe sociale moyenne ou défavorisée.

Les phonèmes /ʍ/ et /x/ sont polymorphes, et l’on ne les trouve pas dans l’anglais britannique standard ou les dialectes parlés en Angleterre. Ils sont parfois remplacés par /k/ pour /x/ et /w/ pour /ʍ/, et ces distributions diffèrent selon les villes. Ils sont moins fréquents à Glasgow, et à Édimbourg, /w/ est le phonème réalisé le plus fréquemment. Dans le nord de l’Écosse, /ʍ/ est remplacé par /f/.

Quant au /r/, qui donne au scots une identité particulière puisque c’est une langue rhotique, ses réalisations varient suivant l’entourage phonétique et la position dans le mot. Il est souvent réalisé comme un approximant, soit un post-alvéolaire [ɹ], soit un rétroflexe [ɻ]. On peut aussi l’entendre en tant que consonne battue /ɾ/. J. Stuart-Smith défend l’opinion que toutes ces variantes sont réalisées par chaque locuteur du scots [34], mais pas dans les mêmes conditions : les gens de la classe ouvrière réaliseraient ainsi une plus grande proportion de consonnes battues, surtout en lisant un texte ; et les adolescents des classes ouvrières réaliseraient également plus de /ɾ/ en parlant. Les informants des classes moyennes, eux, produiraient plus fréquemment des rétroflexes [ɻ] (Stuart-Smith, 2003). Graphiquement, certains préfèrent donc ne pas utiliser de <r>, ou utilisent des mots qui diffèrent de l’anglais standard à cause de leur absence de <r>. Par exemple, l’utilisatrice décrit ici le mot *fae*, qui a posé problème à un internaute : « ‘fae’ I think is Aberdonian or Dorric. ‘I say aw come frae ???’ meaning

34. Son étude portait sur une cinquantaine de locuteurs originaires de Glasgow.

from. » (SC2N59) [35]. La granularité des analyses pousse les utilisateurs de SC à considérer le scots comme une langue à part entière, faite de dialectes topolectaux.

Après cette présentation en détail du scots, la première langue régionale étudiée, l'afro-américain vernaculaire est mon deuxième point d'intérêt. Cette langue minoritaire a également des caractéristiques formelles, et se compose d'un ensemble de pratiques langagières qui le différencient de l'anglais américain standard, ou de l'Anglais Blanc du Sud (US) (ABS). À la différence du scots, l'AAV est utilisé dans l'ensemble du pays, par des locuteurs à prédominance afro-américaine. J'ai choisi le terme « topolecte »pour le désigner. La question de la visibilité de la minorité est évidemment centrale dans l'appellation AAV, mais ce topolecte n'est pas parlé que par des Noirs américains.

2.3 Présentation de l'Afro-Américain Vernaculaire

Les questions épistémologiques, ainsi que la caractérisation notionnelle des mots utilisés pour décrire locuteurs ainsi que langue minorée sont le premier point de ce portrait de l'AAV. Les spécificités formelles seront ensuite explorées, pour ensuite faire converger l'analyse des deux langues lors de la description des statuts sociolinguistiques et des représentations de ces deux langues et de leurs locuteurs.

2.3.1 Analyse critique des formes désignationnelles

Dans cette thèse, j'ai choisi d'utiliser la dénomination la plus moderne pour parler d'une langue appelée *Black English* (‘*anglais noir*’), puis *Ebonics* [36]. Considéré comme un continuum de pratiques langagières dans tout le pays, parlé non seulement par des afro-américains, mais par d'autres populations comme souligné très justement ici : « The social component is at least as important as the ethnic one. Black English is the speech of the socioeconomically lower class of the black population » (Edgar

35. À l'adresse <http ://www.scotster.com/forums/scottish-language-gaelic/Loss-our-language.74.html>. Dernier accès : 8 juillet 2013
36. Dans la *Linguist List*, la page qui agrège les contenus connexes à l'AAV s'appelle encore *Ebonics*.

W Schneider, 1989). Si je suis d'accord avec la première partie de la citation, la fin est plus problématique, et montre l'aspect relativement daté à la fois de l'appellation « anglais noir » et du focus de cette langue. Réduire l'AAV à une caractéristique ethnique est douteux, et l'émergence culturelle afro-américaine ainsi que le contact entre les différentes ethnies ont favorisé l'expansion culturelle de l'AAV. L'expression *anglais noir* est non seulement fausse (au même titre, d'ailleurs, qu'AAV, qui ne recouvre pas non plus les autres ethnies qui parlent cette langue) mais elle est en plus vaguement paternaliste et condescendante. En outre, l'expression *anglais noir* défend la vision de l'AAV comme étant une forme d'anglais. Certes, les locuteurs de l'AAV savent éliminer à volonté les pratiques langagières de l'AAV en fonction du contexte et du coénonciateur, et certes, l'AAV est basé sur un dialecte de l'anglais, mais ce prisme est réducteur. Je soutiens l'hypothèse que la capacité à utiliser et à comprendre l'AAV équivaut à utiliser une langue étrangère [37], étant donné que ses spécificités linguistiques le placent à part de l'anglais standard. L'intelligibilité est le critère choisi pour placer la ligne entre un dialecte et une langue.

Le nom *Ebonics* semble également être très lacunaire : amalgamation d'*ebony* '*ébène*' et *phonics*, il inclut également la composante ethnique et les caractéristiques phonatoires de l'AAV, mais semble faire l'impasse sur les différences graphiques ou lexicales. De plus, ce terme peut recouvrir n'importe quelle langue ou dialecte parlé par des populations noires, conséquences de la traite des Noirs, ou de la diaspora africaine (Baugh, 2000).

Le terme « Ebonics » renvoie également à la controverse d'Oakland, en décembre 1996, quand le bureau d'éducation en Californie a prononcé l'*ebonics* comme étant la langue native des enfants afro-américains de la région, et qu'il fallait ainsi l'inclure dans les programmes scolaires. Ceci a provoqué un tollé médiatique [38], et permis à la communauté linguiste (surtout la *Linguistic Society of America*, ce dont on trouve des traces dans la *Linguist List*) de statuer sur la validité de cette résolution, notamment W. Labov, J. Rickford et W. Wolfram qui ont abondamment commenté l'événement (J. R. Rickford, 1996 ; J. Rickford, 1998 ; Wolfram, 1998). Ceci a provoqué l'amendement de la réso-

37. Les commentaires de vidéos en ligne où les locuteurs d'AAV interviennent comprennent souvent des réflexions sur l'inintelligibilité de leur façon de parler.

38. Surtout à cause de l'utilisation de l'expression « genetically based », comprise par le grand public comme le fait que l'AAV était dû à une prédisposition génétique des Noirs américains.

lution, et la reconnaissance de l'AAV en tant que langue à part entière, méconnue ou mal jugée par les non-locuteurs, mais qui possède de réelles spécificités : « The systematic and expressive nature of the grammar and pronunciation patterns of the African-American vernacular has been established by numerous scientific studies over the past thirty years. Characterizations of Ebonics as 'slang', 'mutant', 'lazy', 'defective', 'ungrammatical'or 'broken English'are incorrect and demeaning. » (« Linguistic Society of America : Resolution on the Oakland "Ebonics" Issue », 1998).

L'expression 'afro-américain vernaculaire', au moins, rétablit les origines controversées de la langue, dans une tentative de la décrire comme parlée majoritairement par des locuteurs d'origine africaine, mais aussi par des Américains à part entière [39]. J'ai retiré volontairement le mot 'anglais' de la dénomination originale (*African American Vernacular English*) pour défendre le statut de l'AAV en tant que langue distincte (bien qu'elle soit utilisée comme un sociolecte dans ses formes les plus diluées). Le mot 'vernaculaire' est peut-être douteux, j'en conviens, mais il est utilisé pour montrer le statut minoritaire de la langue. Je me rapproche également de la dénomination 'langue afro-américaine' utilisée par certains linguistes (Baugh, 1999), et qui semble axiologiquement plus neutre qu'AAV. Ce nom est cependant moins connu, et pour des raisons de clarté, j'ai donc adopté le nom afro-américain vernaculaire.

Certaines caractéristiques de l'AAV sont connues par les non-locuteurs car cette langue est aussi exportée culturellement grâce à la musique, et ses célèbres interprètes, les jeux vidéo et certains films. Dresser une liste d'exemples est presque immédiatement condamné à l'anachronisme, mais des artistes influents comme Snoop (Doggy) Dogg, ou Jay-Z, des jeux vidéo comme *Parappa the Rapper* (Sony Computer Entertainment, 1997), qui fut l'un des plus vieux, mais le premier à mettre en scène la rythmique du rap comme mécanisme de jeu ; la série des *Grand Theft Auto* (GTA)(Take 2 Interactive, 2000) ou *Bulletproof* (Vivendi Universal Games, 2005) sont des exemples de jeux vidéo mettant en scène des locuteurs d'AAV.

39. On peut retrouver des caractéristiques de l'AAV dans les parlers d'immigrés récents venant du Nigéria par exemple.

2.3.1.1 Désignation des locuteurs

L'afro-américain vernaculaire est une langue voisine de l'anglais, issue de créoles (il est reconnu par certains comme un « post-créole » (Wells, 1982 ; Edgar W. Schneider, 2004). L'AAV est parlé dans toutes les régions des États-Unis, mais ce vernaculaire n'est pas commun à tous les Afro-Américains. Il est à noter que l'ethnonyme dont le nom de l'AAV s'inspire est l'un des seuls à s'inspirer du nom d'habitants de deux continents. Il accommode à la fois les origines africaines et le démonyme « américain ». L'afro-américain se distingue de l'*Asian American* ou de l'*Italian American* par un facteur primordial : l'aspect historique de l'arrivée des Afro-Américains en Amérique (bien avant les *Asian Americans* par exemple), et le fait que les identités plus fines (par pays) sont quasiment impossibles à retracer à cause des métissages entre divers pays africains. Le terme « African American » a commencé à être popularisé vers la fin des années 1980, et est une version plus politiquement correcte que *black* ou *negro*, voire *colored*, qui sont tous acceptés. *Negro* et *colored people* sont tout de même un peu archaïques, et sont utilisées par les personnes âgées. Des expressions comme *Negro English*, ou *Colored People English* n'existent pas en tant que telles. Elles n'ont été utilisées qu'avant les années 1960, et sont perçues comme injurieuses. C'est le cas de « Negro Nonstandard English » (Fasold, 1999). Seules *Black English* et AAV sont des dérivations usitées, et ces deux expressions peuvent être vues comme la corrélation du fait que les Afro-Américains se désignent plus volontiers par les expressions *black* ou *African American* (Goudet, 2010b).

2.3.1.2 Représentations de l'AAV et précautions méthodologiques

Certaines études linguistiques effectuées avant le mouvement afro-américain de droits civiques, considéraient l'AAV comme une pidgin ou un créole simple (C. M. Wise & H. Wise, 1965). L'AAV y était évoqué comme un sous-anglais issu d'un manque d'éducation, dans des perspectives racistes dégradantes.

Le début des recherches entreprises sur l'AAV en tant que langue à part entière coïncide avec

l'émergence des mouvements qui conduiront à la reconnaissance des droits civiques des Afro-Américains, pendant les années soixante, alors notamment au centre de débats portant sur l'apprentissage de la lecture par les enfants afro-américains (Baratz & Shuy, 1969). L'intérêt était surtout dirigé vers la détermination de l'aspect pathologique ou non du parler afro-américain, surtout focalisé autour des particularités syntaxiques de l'ethnolecte. Les études ont évolué vers son acceptation et son analyse, notamment dans les études de la culture hip-hop (Richardson, 2006), de sa rhétorique à la construction du personnage du gangster (Kermit Ernest Campbell, 2005).

Les explorations des mécanismes africains sous-jacents à certaines pratiques langagières suggèrent également une récupération positive de la culture afro-américaine, bien éloignée des stigmatisations des débuts de l'étude de l'AAV. Les descriptions du « Signifying Monkey » (Kermit Ernest Campbell, 1994 ; Gates, 1988) sont de bons exemples des études plus positives de l'AAV, conduites par des Afro-Américains, et qui témoignent de l'évolution et de la valorisation de la langue.

Il semble exister un biais mélanodermé dans la recherche, qui apparaît comme dicté par la couleur de la peau du chercheur même. Certains travaux peu scientifiques sont écrits par des locuteurs natifs de l'AAV, ce qui donne parfois des résultats un peu cavaliers comme « Among a vast majority of us, our form of English is recognized as bad, lazy English. » (C. J. Anderson, 2012), qui montre cependant tout à fait correctement la vision de certains (non-locuteurs) sur la langue. Ce qui est cavalier, en revanche, est l'utilisation du pronom « us », ou « our », qui semble aller de soi : le linguiste « sauvage » voit l'AAV comme pratiqué par lui-même, pour lui-même, même dans les travaux scientifiques [40] ; qui se retrouve également dans la recherche plus rigoureuse : « Like many of my Caribbean and African American colleagues I have devoted my entire carreer to the study of the languages of the New World Black Diaspora » (Ball et al., 2002). La phrase ouvrant son article montre qu'il y a un réel biais, autant des linguistes caucasiens (encore que ce serait nier les contributions de Labov ou Baugh à l'appréhension de l'AAV) que des linguistes afro-américains. Il existe effectivement une majorité de chercheurs afro-américains qui travaillent sur la langue (il suffit de regarder les programmes de colloques sur les cultures afro-américaines et hip-hop pour s'en rendre compte) ; et

40. Cette phrase est tirée d'un travail de 10 pages, posté sur *Scribd*. S'il n'est pas une référence rigoureuse, il est néanmoins un travail authentique et un bon point de départ pour la compréhension de l'AAV. Adresse : <http ://fr.scribd.com/doc/103783691/African-American-English-or-Ignorance> ; dernier accès : 10 mars 2013.

certaines thèses (Nurse, 2011) parlent d'être un individu noir dans un institution (éducative) majoritairement blanche [41].

L'interaction de deux identités, celle du chercheur et de l'individu noir [42] se retrouve d'ailleurs dans une citation de Malcolm X : il pose la question suivante à un parterre de scientifiques (dont un blanc qui l'énervait) : « comment appelle-t-on un homme noir avec un doctorat ? Un nègre » (X & Haley, 1992). Cette situation est une sorte d'apogée de violence (en l'occurrence, verbale) qu'il a vécue et qu'il projette sur l'indélicat curieux. Elle montre les rapports entre la couleur de peau, une institution à majorité blanche, et des reliquats d'une Histoire qui provoque encore énormément de tensions entre les ethnies qui l'écrivent.

Cette question est d'ailleurs un non-dit ou un non-vécu en Écosse : le scots est étudié par des locuteurs natifs, qui n'ont pas idée de définir qui ils sont en matière d'appartenance ethnique ou qui n'ont pas besoin de reprendre le pouvoir par rapport à une majorité dominante, encore que l'opposition avec l'anglais est une des raisons qui explique les graphies alternatives du scots, et ce désir de s'éloigner de la langue anglaise (Costa, 2008 ; voir 2.5.2). Pour en revenir aux Afro-Américains, leur histoire est très liée linguistiquement parlant—puisque c'est le sujet qui nous intéresse—au sud des États-Unis et surtout aux anciens états confédérés. L'AAV ne peut être étudié que lorsque ses rapports avec l'anglais blanc du sud (ABS) sont établis.

2.3.2 Interactions avec l'Anglais Blanc du Sud

Les conceptions rétrogrades de l'AAV expliquent son existence comme étant le produit d'un pidgin entre le parler des maîtres (qui parlaient majoritairement l'ABS) et celui des esclaves adultes, suivi d'un état de créolisation où les enfants ont appris un anglais incorrect par leurs parents (C. M. Wise & H. Wise, 1965). Keane, linguiste et ethnologue du 19ème siècle, décrit ainsi l'origine de l'AAV : « Thus were brought together members of every tribe […], speaking hundreds of different languages and dialects, and presenting every imaginable shade of the Negro type. Amidst this confusion of tongues

41. Traduction de *Predominantly White Institution* (PWI)
42. La tentation d'utiliser le mot « intersectionnalité » est grande, tant les différents types de domination (académique, ethnique) se complètent dans cette situation

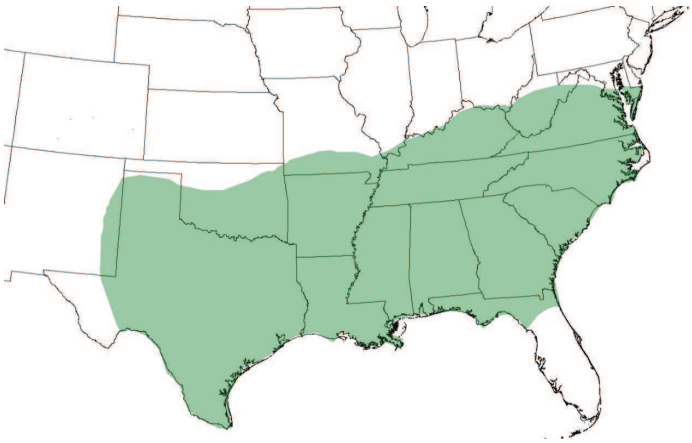


Illustration 2.6 – Aire où l'ABS est pratiqué (Labov, 2006)

and people, English, or at least a Negro-English jargon became the necessary medium of general intercourse amongst the descendants of the captives and the earlier arrivals from America » (Keane, 1895). Il souligne également l'existence d'une traduction de la Bible datant de 1829 et intitulée *Da Njoe Testament va wi Masra en Helpiman Jesus Christus* qu'il identifie comme étant écrite dans ce « jargon » [43], mais précise que cette version était si ridicule (« the result was a burlesque of the sacred text ») qu'elle a été vite remplacée par une version en anglais.

Les interprétations prenant l'AAV comme un créole inférieur à l'anglais ou un dialecte ne sont plus en vigueur depuis les travaux des années 1970 où le statut de l'AAV a commencé à évoluer vers celui d'une langue légitime. Les autres créoles nés de contacts entre les colons et les esclaves ont d'ailleurs un statut langagier similaire, et sont vus comme des langues à part entière (Calvet, 2004).

Il est cependant juste de remarquer que plusieurs traits de l'AAV se retrouvent dans l'ABS. Comme son nom l'indique, ce dialecte est parlé dans le sud des États-Unis, au sud-est (entre le Texas et la Virginie de l'Ouest et partage des caractéristiques communes (notamment phonologiques, comme la monophtongaison des voyelles tendues) avec l'AAV, à cause de raisons historiques, puisque le parler des propriétaires de plantations, qui utilisaient l'ABS, est la variété dialectale que les anciens esclaves ont donc appris en apprenant l'anglais. En outre, la proportion d'afro-américains dans ces régions est importante, ce qui explique également les interactions entre ces ceux dialectes. Il existe un parallèle entre l'émergence de l'afro-américains et de créoles (comme le créole haïtien) qui est défini comme

43. Il s'agit en fait d'un texte en sranan tongo, la langue créole encore parlée au Suriname.

une « langue africaine à vocabulaire français » (Sylvain, 1936). Les ressemblances sont cependant fines : les créoles se sont développés à partir de l'ensemble des langues africaines parlées par les esclaves et de la langue dominante, celle des colonisateurs et des propriétaires d'esclaves, mais les créoles du français ont pu se scinder de l'influence du français, alors que l'afro-américain est toujours en contact avec l'ABS ou l'anglais américain standard.

Le daté *Applied Phonetics* (C. M. Wise & H. Wise, 1965) fait état de plusieurs spécificités communes à l'ABS et à l'AAV :

— La séquence graphique <ar>, finale ou préconsonantale, et non précédée de [w] devient [ɔ] (comme hard /ˈhɔːd/, ou vraisemblablement ˈhɔrd/) plus fréquemment en ABS qu'en AAV, au moins dans le sud des États-Unis.

— La monophtongaison de diphtongues tels que /aɪ/ → /a, ɑ/ /_ l ; /eɪ/ → /ɛ/ /_k .

— Simplification des groupes consonantiques : le [d] final amuï dans des agrégats tels que /md, nd, ŋd, bd, zd, ld, vd, ʤd/ ; de la même façon, /l/ est omis lorsqu'il est devant /f, v, j, m/ *twelve*, *self*, *millin*. Avec le /m/ peut se produire un phénomène de svarabhakti (ou épenthèse vocalique), où un schwa peut se glisser entre le /l/ et le /m/ (*film* /ˈfɪləm/).

— Substitution de /θ/ avec /t/ en position initiale, avec /f/ en position finale ; /ð/ avec /d/ en initiale ou en finale.

— Vocalisation de /l/, qui devient /ə, ʊ, o/ [44] sont trois réalisations possibles de /ˈmɪlk/.

Ces remarques se retrouvent également dans le travail de Mufwene(Mufwene, 1998), qui cite non seulement l'AAV et l'ABS, mais aussi d'autres variétés de l'anglais. Ici, la vision est centrée sur l'AAV. Le tableau suivant (2.7) illustre non seulement les caractéristiques les plus saillantes de l'AAV, mais il montre également la fusion entre variété afro-américaine et dialectes du Sud (huit des quatorze traits présentés appartiennent à des variétés d'anglais parlées dans le sud).

44. Phénomène sensiblement similaire au remplacement des l sombres en écossais où [ˈmɪək], [ˈmɪʊk] et [ˈmɪok]

	Traits	Exemples	Portée
1	Réduction des groupes consonantiques finaux	*cold* → col_ [coul] ; hand → han_ [hæn]	La plupart des variétés de l'anglais ; plus fréquent en AAV
2	Amuïssement de syllabes (initiales/médianes)	about → 'bout [ˈbaut]	La plupart des variétés de l'anglais ; plus fréquent en AAV
3	Haplologie	*Mississippi* → [mɪˈsɪpɨ]	La plupart des variétés de l'anglais
4	Vocalisation de /l/ postvocalique	*bell* → [beʏ]	La plupart des variétés de l'anglais
5	Amuïssement de /r/ après C (θ et en σ non-accentuée	*throw* → [θou]	Autres vernaculaires, surtout dans le Sud
6	Labialisation de fricatives interdentales	*bath* → [bæf] ; bath → [bævz]	Autres vernaculaires, surtout dans le Sud
7	Alvéolarisation des fricatives	*those* → [douz]	Autres vernaculaires, surtout dans le Nord urbain
8	Alvéolarisation des fricatives interdentales	*tenth* → [tɪnt]	Autres vernaculaires
9	Métathèse des finales en /s/ [+stop]	ask → [æks]	ABS archaïque
10	Vocalisation ou amuïssement des /r/ intersyllabiques	hurry → [hʌɨ]	ABS archaïque
11	Vocalisation des /r/ accentués syllabiques	bird → [bɜd]	ABS archaïque
12	Vocalisation des /r/ postvocaliques	four → [foə]	ABS ancien
13	Vocalisation des /r/ inaccentués syllabiques	father → [fɑðə]	ABS ancien
14	Réduction de /ai/ avant consonne et en position finale	*tied* → [tɑːd] ; tie → [tɑɪ]	ABS

Tab. 2.7 – Traits phonologiques saillants de l'AAV d'après Mufwene, 1998

Ces particularités, surtout orientées autour des consonnes de l'afro-américain, sont communes à d'autres dialectes de l'anglais, outre l'ABS. Elles se retrouvent toutes dans l'afro-américain, qui concentre à la fois des particularités dialectales vernaculaires (éléments 5-8) et du parler du sud (éléments 9-14). L'aspect historique de l'utilisation de l'AAV est à prendre en compte pour expliquer la convergence de ces traits dans un dialecte parlé dans l'ensemble des États-Unis.

En revanche, s'il est porteur d'histoire en tant que dialecte parlé, l'ABS est une variété dialectale pure de l'anglais, avec peu de particularités lexicales : l'utilisation de « y'all » ou « you all » est commune aux deux dialectes, mais il existe un substrat culturel africain fort, qui justifie le fait que l'AAV se soit scindé de l'ABS, et qui explique certains apports lexicaux (que je vais traiter dans la section suivante). En revanche, les réalités afro-américaines contemporaines (qui sont vécues par n'importe quel groupe qui n'appartient pas à la majorité) ont apporté des champs lexicaux inédits. Quant à la syntaxe de l'ABS, c'est un sujet fort peu évoqué dans la littérature, ce que je ramène au fait que sa syntaxe est proche de l'anglais américain standard (contrairement à celle de l'AAV.

2.4 Examen des spécificités de l'AAV

2.4.1 Lexique

Le vocabulaire de l'AAV des premiers esclaves afro-américains s'est d'abord développé comme un langage secret utilisé par une population sans pouvoir politique, à la manière d'un argot, afin de partager des expériences communes. *Juba* signifie '*manger de la nourriture dont les maîtres ne voulaient pas*', et ils composaient des chansons à double sens afin qu'elles soient hermétiques aux propriétaires de plantation. Ainsi « Juba this, juba that / Juba killed a yella' cat » est une chanson évoquée par Bessie Jones [45] qu'elle paraphrase ainsi : « means giblet this [a little of this] and /Giblet that [a little of that] because they couldn't say mixed-up food might kill the white folks. They was afraid to say that because white folks'd kill them. » (Holloway, 2005). Cette chanson ne contient qu'un seul mot non-anglais (ce qui n'éveille pas les soupçons des propriétaires), les paroles en sont relativement transparentes si la valeur de moquerie cachée dans ces chansons est connue. Le mot *white* est soigneusement évité et remplacé par une autre couleur claire, afin de ne pas éveiller les soupçons.

De nombreux mots de langues africaines (principalement wolof et yoruba), sont entrés dans le vocabulaire de l'anglais standard, et semblent faire partie de grandes catégories : le vocabulaire culinaire (*gumbo*, mot bantou qui désigne l'ingrédient principal, l'*okra* en anglais ou '*gombo*' en français). Certains autres mots ont été réinjectés dans les argots parlés par les membres de gang. *Damu* est un mot désignant les membres du gang des Bloods [46], qui veut lui-même dire '*sang*' en swahili. Enfin, il existe des mots ayant la même forme que les mots anglais, mais qui viennent de langues africaines, comme *dig* (dans le sens '*apprécier*') qui vient en fait du wolof '*dega*', ainsi que *hip*, dans le sens '*à la mode*', également originaire du wolof *hipi(cat)*, '*qui a les yeux grands ouverts*'. Les noms de quelques styles musicaux et danses viennent aussi de langues africaines (*mambo, tango, jazz* [47]*, merengue, samba, funk*).

45. B. Jones (1902-1984) est une chanteuse de gospel, originaire de Géorgie. Elle apprit les chansons de son grand-père, ancien esclave.

46. Les Bloods et les Crips sont les deux gangs afro-américains les plus importants aux États-Unis ; certains membres autoproclamés sont présents sur BP.

47. Proche de *jasm* ou *jism*, mot d'argot du 18ème siècle signifiant « sperme ». Il désigne le style musical à partir de 1912 (Cohen, University of Missouri–Rolla & Department of Arts, 2005).

Certaines pratiques se sont exportées d'Afrique, ainsi que leur vocabulaire propre, comme l'action de « signifier ». La « signification » (*signifying*, en anglais) est une forme de joute verbale consistant à tourner son adversaire en ridicule, ou à se moquer d'une figure d'autorité à son insu, en utilisant la rhétorique à son avantage (Kermit Ernest Campbell, 2005). Ces joutes sont encore en usage dans des sous-cultures, notamment la culture hip-hop. Elles ont également entraîné l'attestation de formes graphiques et lexicales alternatives de mots américains. Par exemple, le pendant positif de la « signification », est la *rep* ou '*rep(résentation)*', c'est-à-dire la façon dont on se met en valeur et où on mobilise son charisme (*swag* [48]) afin d'être (re)connu dans son milieu. L'argot afro-américain regorge d'expressions et de lexèmes proches de l'anglais américain standard, mais ayant des étymologies différentes. Ainsi, *floss* ('*la vantardise*') est un emprunt au français '*floche*', tiré de la qualité de la soie (Bradford, 2001). Certaines de ces expressions seront étudiées en 2.5.2.

2.4.2 Graphie

Les graphies afro-américaines sont adoptées par certains locuteurs seulement. Elles ont été fixées très marginalement en diachronie, dans certains récits, ou poèmes du 17ème siècle ; les récits d'esclaves sont rédigés en anglais standard, mais à partir de la fin du 19ème siècle, on peut voir des utilisations de dialectes parlés par des afro-américains comme le gullah dans les œuvres de J.E. Campbell et un AAV écrit de façon relativement stable, dans les œuvres de J.S. Cotter, de P.L. Dunbar ou de L. Hughes (Sherman, 1997). La culture moderne hip hop, qui se fond avec la culture jeune (ainsi que les pratiques graphiques des nouvelles techniques de communication) ont énormément contribué au développement non seulement lexical, mais aussi graphiques pour transcrire l'AAV, grâce à la démocratisation de pratiques graphiques communes sur internet, média se prêtant peu à des pratiques prescriptives en matière de correction orthographique.

Les réductions et cliticisations sont encouragées, et elles se basent sur le modèle *got to* > *gotta*. Ainsi, des lexèmes comme *finna* ('*fixing to*'), *tryna* ('*trying to*'), *imma* ('*I am to*'), sont couramment

48. Charisme est une sous-traduction de *swag*, qui recouvre autant cette notion que celle de gouaille et de séduction. Le *swag* est une valeur extrêmement positive dans la culture hip-hop, et ce lexème a été récupéré dans l'argot jeune français contemporain.

utilisés en AAV. Contrairement au scots, les apostrophes marquant les épenthèses ne sont pas décriées, et l'AAV affiche clairement ses graphies démarcatives. La diphtongaison de certains sons (/o/ inter-consonantique > /aʊ, æʊ, əʊ/) est déjà représentée par l'utilisation du digraphe <aw> dans le poème *Bound No'th Blues* (Hughes, 1959) dans le mot *lawd* pour *lord*.

De la même façon, la labialisation des fricatives dentales /θ, ð/ est déjà montrée par des graphies démarcatives dans des ouvrages anciens : on peut trouver les graphies *mouf* < *mouth*, *de* < *the* dans The Don't-Care Negro (Cotter & Cosmopolitan Press, 1912). Les groupes consonantiques sont eux aussi réduits dans le même poème, avec une élision du <d> final : *min'* < *mind*, *boun'* < *bound*, *roun'* < *round*, *an'* < *and*.

En règle générale, plus les graphies sont modernes [49], et plus elles présentent des différences avec l'anglais standard. L'identification encore fragile de l'AAV en tant que langue à part entière explique cette flexibilité de figuration graphique : il n'existe pas de feuille de style ou de recommandation normative pour la graphie afro-américaine, contrairement à celle du scots. Cette liberté graphique –sans entrer dans des considérations onomastiques– est visible dans la composition des prénoms afro-américains, où il est possible de reprendre un prénom usuel, et de l'orthographier librement. Ainsi, dans les listes de suggestions pour des prénoms afro-américains, on peut trouver des graphies adaptées de prénoms usuels : *Makayla* pour 'Michaella', *K'rin* pour 'Karen' ou *Bartt* pour Bart(holomew). Certains prénoms sont également déclinés : Alicia devient *Eleasha, Eleesha, Elesha, Ellesha, Jalicia*, ou *Lakeysha* ; pour les hommes, le prénom 'John'est décliné en *Da-Jon, Da-Juan, Dajuan, Dejohn* et *De-John*. Le prénom féminin tiré de *Patricia* et courant dans les populations afro-américaines est 'Latrice'. Il possède de nombreuses variantes graphiques, comme *La-Trice, Latreece, Latreese, Latriese* (« African-American Baby Names », 2009 ; « Popular African American Names », 2013).

Tout ceci revient à dire que certains usages graphiques sont proches de la prononciation, d'autres sont extrêmement stylisés (écriture de gangs, autres blogs), d'où la confusion entre un afro-américain écrit et des pratiques discursives vernaculaires générales (l'*e-bonics* ou *typobonics*), deux néologismes au sens proche. Ce système est celui utilisé sur BP, où les graphies peuvent se rapprocher le plus pos-

49. À partir du 20ème siècle

sible de la prononciation, comme dans l'exemple « Look-a-here I'm fixta be skrate real wit cha I'm ghetto as Hell » '*Look here, I'm going to be straight (?) real with you I'm ghetto as Hell*' (BPPP29).

Certains utilisateurs de l'internet, notamment de l'UD, reconnaissent un certain type de communication en ligne, l'*e-bonics*. Le nom vient d'un jeu de mots sur le suffixe <e->, abréviation d'*electronic*, et d'*Ebonics*, l'un des noms les plus démocratisés pour désigner l'AAV. L'*e-bonics* est parfois appelé *typobonics*. Ce jeu entre *e-bonics* et *Ebonics* montre la déterritorialisation de l'entité culturelle (Deleuze, 1972), et la perception profane, qui consiste à actualiser le rapport perçu entre AAV et anglais, et à le mettre sur le même plan que le rapport entre graphies alternatives nées sur internet et l'anglais. Utiliser l'UD pour décrire ce phénomène est logique : dans une perspective écologique, il est normal de rechercher des descriptions de ce phénomène là où il a été créé, en ligne. Les auteurs des entrées de l'UD pour *e-bonics* et *typobonics* font bien la part des choses entre l'afro-américain et ces pratiques écrites, mais font le parallèle entre des pratiques écrites et orales vernaculaires. Voici les deux entrées les plus populaires pour *typobonics* (9 pour ; 2 contre) et *e-bonics* (54 pour ; 21 contre) respectivement :

> "The intentional misspelling or shortening of common words in order to sound cool in some circles. It is similar to ebonics in spoken english , but used, as the name implies in typed english . Typobonics is used almost exclusively on the internet, not uncommonly used in teenage chatrooms, blogs about boy bands, instant messages and e-mails between teenage girls about high school drama.
> Typobonics is considered by some to be lazy and low-class.
> Not to be confused with l33t, which is common with internet nerds, or the condensed written english associated with text messaging, which is shortened out of necessity of the nature of a text message.
> "My little sister always uses typobonics instead of normal english when she's instant messaging her dorky friends."
> Some typical typobonic word subsitutions are :
> 'cuz'instead of 'because'
> 'plz'instead of 'please'
> 'luv'instead of 'love'
> 'neva'instead of 'never'
> 'n'instead of 'and'
> 'da'instead of 'the'
> "u neva gunna looz weight cuz u always chattin on da internet. " [50]"

> "Electronic ebonics, or the practice of reducing normal, comprehensible sentences into a

50. Accessible à la page <http ://www.urbandictionary.com/define.php ?term=typobonics>. Dernier accès : 20 février 2013.

> few choice letters and symbols. Punctuation is optional and often discouraged. This form of communication began with instant messenging systems but has spread to other forms of dialogue, making it far more annoying.
> Guy 1 : omfg im n luv wit dis gurl
> Guy 2 : Are you retarded ? Gurl isn't even short for anything. You just spelled it wrong for the sake of being a dumbass. You need to ditch the e-bonics [51]."

On peut percevoir assez facilement un racisme sous-jacent (l'*ebonics* est à l'anglais ce que l'*e-bonics* est à l'anglais en ligne) ; et un rejet de la pratique parce qu'elle est obscure (« the practice of reducing normal, comprehensible sentences into a few choice letters and symbols. ») ou, au moins, difficile à comprendre pour certains—justement, les auteurs de ces entrées, qui se moquent de cette pratique. Ces surnoms pour un type de communication en ligne pratiqué par les jeunes est symptomatique de l'antipathie que certains ressentent envers l'AAV (voir 2.6.3.1). Dans le second exemple, l'expression *omfg* (*'Oh My Fucking God'* est un exemple de *netspeak*, ou expression née grâce à la communication écrite électronique non standard [52]. L'utilisateur montre donc la confusion entre e-pratique et *ebonics*, et la réception de ces écrits non standardisés.

Les deux expressions *typobonics* et *e-bonics* sont peu utilisées, il y a très peu d'entrées pour les définitions des deux lexèmes (4 pour *e-bonics* et une seule pour *typobonics*) et peu de votes pour les soutenir [53]. Cette définition montre que ces pratiques, ou, au moins, leur désignation, est rare. Elle se rapproche des tachygraphies attendues d'internet, qui font partie du netspeak (étudié dans la partie 4.4.2).

Ces noms, interchangeables, serviront à désigner les pratiques abusives de langage sur internet, voulant émuler l'AAV sans en être réellement. Par exemple, l'utilisation de la terminaison <-a> pour remplacer <-er>. Cette marque d'oralité est rencontrée dans certains lexèmes codifiés : *gangsta*, ou *nigga* représentent deux réalités afro-américaines. *Playa* '*player*' aussi, dans un contexte de séduction est rentré dans le vocabulaire courant, et le *hustla* '*hustler, dealer de drogues*' est une figure des milieux pauvres. Certains mots historiques comme *cracka* ('*cracker*', terme insultant) sont également écrits de cette façon sans que cela ne choque qui que ce soit (34 entrées pour la définition dans l'UD).

51. Accessible à l'adresse : <http ://www.urbandictionary.com/define.php ?term=e-bonics>. Dernier accès : 20 février 2013.
52. Ceci recouvre « langage SMS » et *Netspeak*.
53. Pour donner un ordre de proportion, le « Mot du jour » sur l'UD reçoit entre 7 000 et 2 000 votes (moyenne : 5 000) ; *Merica*, un mot peu populaire, qui n'a que trois entrées, totalise 2 330 votes cumulés.

Dans les faits, la terminaison est également utilisée dans des adjectifs comparatifs comme *longa* ('*longer*'), mais ils ne sont pas définis dans l'UD, et sont peu fréquents (8 500 occurrences de '*bigga*'sur BP ; 1 510 pour '*longa*'). La distribution de <-er> > <-a> porte essentiellement sur des noms.

En revanche, des utilisations exagérées comme *rappa(h)*, sont considérées ridicules. Bien qu'elles puissent figurer la prononciation, elles ne sont pas particulièrement entrées dans l'argot en tant que tel et n'appartiennent pas à une réalité ou à un ethos particulier. Bien que l'on puisse parler d'un style particulier de rap, le gangsta rap, *rappa(h)* ne recouvre rien, sinon la volonté de vouloir écrire les mots différemment. Les entrées pour *rappa(h)* soulèvent toutes le fait que c'est un mot du ghetto, tourné en dérision.

2.4.2.1 Authenticité des graphies scriptées

Il existe des scripts de « traduction » afro-américaine, proposés par plusieurs sites. Le principe est simple : l'utilisateur du site entre un texte, et le script le convertit immédiatement en « AAV ». Nonobstant le côté simpliste et normocentré de la chose, ces scripts présentent des graphies courantes en AAV, et le choix entre trois scripts (tirés de *Billism* (Krahmer, 2006), *Written Humour* (Brower, 2003) et *the Ebonics Translator* (Joel, 1999)) permet de dégager des propriétés communes entre les versions.

J'ai choisi d'utiliser les deux premières strophes de *The Tyger* (W. Blake, 1789/1977) comme cobaye pour montrer les transformations subies par le texte quand il « devient AAV graphique ».
Les mots dont les graphies ont été changées, ou les additions sont marquées en gras.

Ces scripts convertissent en échangeant les lettres : la séquence <th> est remplacée par <d> dans 1, 2 et 3 quand il est dans le mot <the>, qui fait partie des mots à forte fréquence. Ceci est dû à l'influence de la prononciation (voir 2.4.4.2 pour une explication sur le '*th-fronting*').Elle n'est cependant jamais remplacée dans *thine* ou *thy*, ce qui prouve que ces scripts ne remplacent pas cette séquence

Texte original	Billism (1)	Written Humour (2)	Ebonics Translator (3)
Tyger ! Tyger ! burning bright In the forests of the night, What immortal hand or eye Could frame thy fearful symmetry ? In what distant deeps or skies Burnt the fire of thine eyes ? On what wings dare he aspire ? What the hand dare sieze the fire ?	tyger tyger **burnin** bright in **da foests** of **da** night, what immortal hand or eye Could frame thy fearful symmetry ? in what distant deeps or skies burnt da fire of thine eyes ? on what **wins dis** he aspire ? what **da** hand **dis** sieze **da** fire ? **- aww yea foo.**	Tyger ! Tyger ! burning bright In **da** forests **o' da** night, What immortal hand or eye Could frame thy fearful symmetry ? in **wa'tch** distant deeps or skies Burnt **da** fire **o'** thine eyes ? On **wa'tch** wings dare **da nigga** aspire ? What **da** hand dare sieze **da** fire ?	**Yo** Tyger ! Tyger ! burning bright In **da** forests **o' da** night, What immortal hand or eye Could frame thy fearful symmetry ? In what distant deeps or skies Burnt **da** fire **o'** thine **peeps** ? On what wings dare he aspire ? What **da** hand dare sieze **da** fire ? **and git Sheniquah's ass back ova' heeah**

Tab. 2.8 – Comparaison entre les scripts automatiques d'afro-américanisation

automatiquement.

Certains changements sont curieux : le mot *forests*, graphié *foests* en 1, est unique, et d'autres <r> en position intervocalique (comme dans *aspire*) ne subissent pas la même transformation. Un /ʃ/ excrescent apparaît dans le mot *what*, dans la version 2, certainement à cause d'une automatisation malheureuse : un « what [do] you want » est transcrit par « wa'tch [do] ya'll want ? » ; la fricative postalvéolaire servirait donc à faciliter la prononciation entre le /t/ et le /j/ suivant, dans ce groupe, et l'adjonction d'une apostrophe sert à différencier le mot de « watch ». 1 est le seul à marquer la dévélarisation de /ŋ/ (*burning* > *burnin* ; *wings* > *wins*), mais les groupes consonantiques ne sont pas réduits. 2 et 3 utilisent également une simple modification de *of* en *o'*.

Les transformations sont donc de trois ordres : phonétiques, sur les mots grammaticaux fréquents (*the* > *da*) dans les trois scripts. La prosodie et les phénomènes de coarticuliation sont pris en compte, dans une certaine mesure (épenthèse malheureuse dans *wa'tch*). Ces transformations sont également lexicologiques, avec une permutation de mots lexicaux fréquents avec de l'argot (*eyes* > *peeps*). Enfin, stylistiquement, 1 et 3 choisissent d'ajouter des phrases stéréotypiques et explicitement afro-américaines (Sheniquah est un nom typiquement afro-américain ; *foo* '*fool*' est un vocatif fréquent en AAV.

Ces procédures automatiques s'appuient sur des pratiques linguistiques connues, et grossièrement mises en oeuvre systématiquement. Elles ne se rapprochent pas des graphies communaulectales de BP, à part « da » pour *the*, qui est plutôt utilisé chez les jeunes utilisateurs (une occurrence sera étudiée en 4.4.5.1) Ces éléments ne sont pas communaulectaux, puisqu'ils ne sont pas actualisés par des membres humains d'internet. Les clôtures de 1 et 3 placent ces transformations dans le pôle de l'humour normocentré envers les pratiques linguistiques orales d'autres groupes.

2.4.3 Syntaxe

Contrairement au scots, il n'existe pas de corpus d'AAV récent en ligne. Les exemples seront donc tirés de l'*Electronic World Atlas of Varieties of English* (eWAVE) pour en déterminer la fréquence [54] ainsi que du travail de Wolfram pour décrire la morphologie et la syntaxe de l'AAV(Kortmann, Upton, Schneider, Burridge & Mesthrie, 2008 ; Mufwene, 1998). Ils sont calculés en fonction de leur indice de diffusion [55] : les propriétés de rang A, B et C sont pondérées ($A = 1; B = 0.6 C = 0.3$) et divisées par le nombre d'occurrence du phénomène dans le corpus. La formule de calcul est :

A ×(nombre de traits en AAV, nombre de traits de rang A)+ 0.6 × B × (nombre de traits en AAV, nombre de traits de rang B) + 0.3 × C × (nombre de traits en AAV, nombre de traits de rang C)] ÷ [A ×(nombre de traits en AAV, nombre de traits de rang A)+ B × (nombre de traits en AAV, nombre de traits de rang B)+ C × (nombre de traits en AAV, nombre de traits de rang C)]

Étant donné que l'AAV possède 141 traits pour sa variété urbaine, et 164 pour la variété rurale, seuls les traits les plus fréquents (rang A) seront décrits. Certains exemples sont également pris de chansons interprétées par des artistes afro-américains.

54. L'eWAVE utilise un système de notation par rang, de A '*très prévalent ou obligatoire*', à C '*le trait existe, mais il est extrêmement rare*'.
55. *Pervasiveness index* en anglais, traduction personnelle.

2.4.3.1 Caractéristiques verbales

La syntaxe de l'AAV rappelle ses origines créoles, dans le sens où le système verbal est épuré et diverge de l'anglais standard. Une des caractéristiques syntaxique de l'AAV est l'utilisation (ou le manque d'utilisation) de *be*-copule et *be* dans la formation de l'aspect continu. Ce verbe est parfois évité, ou reste dans sa forme de base verbale. Ainsi *she nice* et *they acting silly* sont des formes tout à fait acceptées. *Be* seul est un marqueur utilisé devant un verbe non-fini. « We be burning » (Paul, 2005) en est un exemple dans une culture jeune, qui revendique l'utilisation de *be*-copule. On en trouve des exemples dans des chansons connues « Dr. Dre be the name still running the game » (« Doggystyle », 2001), ce qui a contribué à la connaissance de ce trait syntaxique essentiellement afro-américain et appartenant à la culture hip-hop.

L'utilisation de *be* habituel est l'une des constructions les plus caractéristiques de l'AAV. Certaines formes de *be* seules sont le résultat des amuïssements de la contraction de l'auxiliaire *will* ou *would*, prononcé /l/ dans une séquence /lbɪ/. On les retrouve néanmoins dans les reprises elliptiques ou dans les négations : « she be here in a minute, won't she ? She won't be here in a minute ». Pour l'aspect fréquentatif, en revanche, l'auxiliaire qui porte la négation ou qui est utilisé dans la reprise n'est pas *be* mais *do* : « sometimes they be playing tag »> « sometimes they don't be playing tag » ; « sometimes they be playing tag, don't they ? ». Cette utilisation est beaucoup plus courante dans l'AAV urbain, associé avec la culture jeune que dans les parlers ruraux et chez les locuteurs plus âgés.

Les formes du verbe $[\text{be to}]_V$ est surreprésenté en AAV, seulement à la première personne du singulier. Cette périphrase sert à exprimer le futur, dans l'expression *I'mma* ou *Imma*. Ce fait est peu documenté dans les travaux, mais l'utilisation et le nombre de recherches pour ce mot ont explosé en 2010, à la date de sortie du single « Imma be » [56](« The E.N.D. » 2009) qui a contribué à faire connaître cette expression.

L'AAV admet des doubles modaux, de la même façon que le scots. '*Could*' est le modal le plus utilisé dans les doubles modaux, mais des combinaisons différentes sont possibles comme '*might could*', '*may*

56. Plus de 3 millions de téléchargements légaux rien qu'aux États-Unis, le single s'est placé en tête du *Billboard Hot 100* pendant la semaine du 6 mars 2010. Les recherches les plus fréquentes sur l'expresion « Imma' » ont eu lieu en 2010, et les concordances lexicales contiennent « lyrics » et « be ».

can', '*should can*, ''*should oughta*'. En revanche, toutes les combinaisons ne sont pas acceptables, comme **may would* et **should could*. Les modaux peuvent aussi porter les mêmes marques de passé ou de présent, ou des marques différentes, comme dans la phrase « You might can step ya game up wit this »[57]. La négation peut se placer après le premier ou le deuxième modal. Chacun des deux modaux a une force distincte dans la phrase : par exemple, dans « she might could go to the show, couldn't she ? », le modal *could* est le modal-tête, celui qui peut être utilisé pour composer la reprise elliptique. *Might* peut parfois être interprété comme un remplacement de l'adverbe *possibly*, mais pas forcément à la même place (pour l'exemple précédent, « she [ ] could [possibly] go to the show », l'adverbe doit être mis après le modal).

Certaines expressions remplissent des rôles modaux en AAV : *come* indique l'indignation ou la colère ressentie par l'énonciateur « He come walkin' in here like he owned the damn place », mais l'ambiguïté sémantique avec des verbes de mouvement le rend difficile à repérer. De la même façon, *steady* intensifie les verbes à la forme progressive : « We be steady mobbin » (« We are Young Money », 2009). L'adverbe montre le côté insistant de l'action '*mobbin*' ('*être membre d'un gang*'). L'expression semi-modale *fixing to* est utilisé pour signifier '*être sur le point de, avoir l'intention de*'. Il est souvent écrit sous l'une de ses formes contractées, *finna, fixta, fitna, fidda* et est commun à l'AAV et à l'ABS.

Le participe passé *done* peut avoir le sens de complétion (il est également appelé *done* complétif). On le trouve dans l'ABS et d'autres variétés dialectales parlées aux États-Unis, ainsi que dans des créoles carribéens. Il n'apparaît qu'en position d'auxiliaire préverbal, et peut être utilisé pour intensifier un contraste : « I done tell you not to mess up » ; ou pour renforcer l'expression de la complétion : « they done used all the good ones ». *Be* et *done* peuvent également s'associer pour former une sorte de futur antérieur « My ice cream be done melted when we get there » mais il peut se teinter d'une valeur de futur inexorable issu d'une condition « if you love your enemy, they be done eat you alive in this society ». Ce type de structure est souvent associé avec la menace ou l'avertissement.

La morphologie des verbes est également différente, et la première personne du singulier est souvent codée par l'ajout de la désinence <s>. À l'inverse, le <-s> marquant la troisième personne du singulier disparaît régulièrement en AAV, surtout en milieu urbain. Dans certaines populations jeunes où on

57. Accessible à l'adresse <http ://www.blackplanet.com/forums/thread.html ?thread_id=1027255>. Dernière visite : 17 juillet 2013.

peut voir entre 75 et 100% d'utilisation de la troisième personne sans l'accord (Labov, 1968), alors que l'AAV rural accorde plus volontiers le verbe à la troisième personne du pluriel avec le sujet en lui ajoutant un <-s> « the dogs barks » (Cukor-Avila & G. Bailey, 2001), et cet accord peut également survenir avec un sujet à la première ou à la deuxième personne du pluriel.

2.4.3.2 Caractéristiques nominales

Les génitifs saxons sont peu fréquents en AAV, surtout chez les populations jeunes, mais les phrases gardent la même structure syntaxique « the man__ hat », « the cat__ ear ». Les pluriels sont également particuliers. Il existe une tendance forte à régulariser les pluriels irréguliers de l'anglais. Ainsi, *children* devient *childrens*, et certains incomptables ou noms qui n'ont pas de pluriel marqué peuvent se voir adjoindre un <s> final, comme « he saw three sheeps », ou « two corns ». En revanche, les syntagmes nominaux exprimant des mesures chiffrées, de distance ou de quantité « I got 50 cent__ » ou « It's 7 mile__ from here » ne s'accordent pas. C'est une forme partagée avec certains parlers vernaculaires du nord du pays, ainsi qu'avec l'ABS, mais elle est peu fréquente dans les parlers de l'AAV urbain.

Ces utilisations ressemblent à celles des pluriels associatifs, comme « Darius an 'em » ('*Darius and his friends*'), que l'on peut trouver dans des langues créoles. L'AAV (dans l'ensemble du pays) utilise, comme d'autres parlers vernaculaires, le pronom *y'all* en position sujet « y'all done now », en vocatif « y'all guys » ou en tant que pronom possessif « it's y'all ball » (que l'on peut rapprocher de l'absence de génitif saxon en AAV). Ce lexème code la deuxième personne du pluriel, est un trait que l'on trouve très fréquemment en ABS [58] , ainsi que dans de nombreuses variétés de l'anglais [59].

Quant aux démonstratifs, il n'est pas rare que le pronom personnel *them* remplace les déictiques *those* ou *these*, comme l'illustre la chanson *Graffiti* de C. Brown « she likes them girls » (« Graffiti », 2009).

Tout comme le scots, l'AAV a tendance à régulariser les pronoms réfléchis de la troisième personne ('*himself*'devient donc '*hisself*'). Le pronom personnel *me* est utilisé à la place de *I* dans les sujets

58. Il existe même un très sérieux magazine, « Y'all » à destination des gens du sud des États-Unis, bimestruel, dont le tirage est de l'ordre de 100 000 exemplaires ; il n'est malheureusement plus édité depuis 2010.
59. La forme en scots « youse » est relativement proche.

parataxiques, qui contiennent une coordination : la phrase « Him and me went to the movies » est correcte en AAV.

2.4.3.3 Caractéristiques adverbiales

Le placement et la distribution des négations est une des différences fondamentales entre l'AAV et l'anglais. La double négation, aussi appelée négation multiple ou pléonastique, consiste à utiliser au moins deux marqueurs de négation, sinon plus, pour exprimer une négation simple. En anglais standard, un seul marqueur morphologique de négation est nécessaire pour arriver au même résultat.

(16) « He **ain't** got **no** car. »

(17) « **Nobody** around here **ain't never** heard of him (**at all**). »
'*Nobody around here has ever heard of him (at all).*'

L'utilisation de la double négation « logique », où deux marqueurs négatifs forment une phrase positive est marquée par l'intonation et la prononciation. L'intonation peut porter des marques d'emphase sur l'une des négations (intonation montante) et ce mot a souvent un accent tonique contrastif, comme le montre la différence entre les deux exemples suivants :

(18) « I ain't say nothing –I just sat there. »

(19) « I didn't say **no˥thing** –I just said it very softly. »

Les marqueurs négatifs peuvent être distribués dans la proposition principale et dans les subordonnées. Il n'y a aucune contrainte de position des marqueurs, sinon celles de la grammaire de l'anglais standard : ils doivent se placer où se place une négation en anglais. Ces négations peuvent être des adverbes préposés : « **No way** you gonna get that outta here with **no** half-ton truck », les sujets ou les verbes auxiliaires ou modaux des phrases « nobody don't like that », et des syntagmes nominaux ou prépositionnels, en position objet direct :

(20) « he ain't got no genius. »

(21) « nobody saw nothing out there that looked like no bear. »

En revanche, si une expression indéfinie comme *anyone* ou *ever* est placée avant un verbe portant une négation, la polarité négative est contenue dans le sujet. La phrase « nobody got no gas » est donc correcte en AAV et en ABS, mais « *anybody ain't got no gas » ne l'est pas.

Le tableau suivant 2.9 compare les traits syntaxiques évoqués dans leur distribution en AAV rural et urbain, et entre les variétés d'AAV du sud et du nord des États-Unis, d'après des données de W. Wolfram, qui a contribué à l'eWAVE en documentant les entrées « Rural African American English », « Urban African American English » et « Southeast American enclave dialects » ; le sociolinguiste s'est servi de ses données diachroniques et synchroniques (il étudie Détroit depuis les années 1960, et a conduit une étude linguistique sur près de 700 locuteurs d'AAV), ainsi que de conversations spontanées pour eWAVE.

Le signe égal surmonté d'un point <≐> représente le fait que ce trait est stable et utilisé fréquemment dans le dialecte de l'AAV, et la croix dans un cercle <⊗> montre qu'il n'est pas utilisé. La flèche descendante <↘> indique qu'il est en déclin et la flèche montante <↗> que son utilisation est de plus en plus fréquente. Les parenthèses autour des différents symboles montrent que les traits sont peu usités. Les traits sont classés par type : d'abord ceux dont la fréquence est de plus en plus importante, puis ceux dont la fréquence décroît, pour finir par les traits utilisés de façon stable en AAV.

Structure	Exemple	AAV urbain	AAV rural	AAV du sud	AAV du nord
Be habituel et auxiliaire du progressif	*I always be playing ball.*	↗	(↗)	⊗	⊗
Be copule et d'emphase	*She be the diva.*	↗	⊗	⊗	⊗
Résultative *be done*	*She be done had the baby*	↗	⊗	⊗	⊗
Come d'indignation	*They come talkin' that trash.*	↗	⊗	⊗	⊗
Absence de <s> verbal 3è personne sg.	She run everyday.	↗	↗	⊗	⊗
Ain't remplacement de *didn't*	*I ain't go yesterday.*	↗	(↗)	⊗	⊗
Double modaux	*I might could do it.*	(↘)	↘	↘	⊗
Préfixe <a> devant les verbes à la forme continue	*I was a-huntin'.*	⊗	↘	(↘)	⊗
Accord <s> aux verbes à la 3è personne du pluriel	*The dogs barks.*	⊗	(↘)	↘	⊗
Absence de *be*-copule	*She nice.*	≒	≒	(≒)	⊗
Done à valeur complétive	*She done did that.*	≒	≒	≒	⊗
Négation pléonasmique	*We didn't do nothing.*		≒	≒	≒
Indéfini préverbal négatif	*Nobody don't like it.*	≒	≒	(≒)	⊗
Finna auxiliaire	*I finna paint the garage.*	≒	≒	≒	⊗
Omission du pluriel, noms de mesure	*Three mile*	≒	≒	≒	⊗
Absence de pluriel	*Three boy*	(≒)	≒	⊗	⊗
Régularisation des pluriels	Oxes	≒	≒	≒	(≒)
Absence de génitif saxon	*The girl hat.*	≒	≒	≒	⊗
Régularisation de *hisself*	*He shaved hisself.*	≒	≒	≒	≒
They possessif	*It's they books.*	≒	≒	⊗	⊗
Y'all 2è personne du pluriel	*Y'all should come.*	≒	≒	≒	⊗
Them demonstratif	*I love them books.*	≒	≒	≒	≒
An'em associatif	*James an'em will be there.*	≒	≒	(≒)	(≒)

Tab. 2.9 – Spécificités de l'AAV suivant la zone géographique d'après Kortmann et Lunkenheimer, 2008

Ces traits sont indicatifs, et peuvent évoluer dans les années à venir. Bien que les structures de l'AAV soient communes à d'autres dialectes, il est à noter que dans les zones métropolitaines et dans les villes en général, le parler AAV se démarque très nettement d'autres dialectes ou langues européennes pratiquées. Les marqueurs d'appartenance à un côté de la ville ou l'autre, *on the right/wrong side of the tracks* sont exprimées par la langue. La dichotomie traditionnelle et la configuration des quartiers dans la ville, répartis par appartenance ethnique explique des expressions comme *talking white*, ou sans marqueurs d'AAV, et *talking black*, avec plusieurs traits syntaxiques de l'AAV.

L'appartenance ethnique n'est pas le seul facteur à contribuer à l'utilisation ou pas de *white* ou de *black talk* : la géographie et les goûts personnels sont des facteurs importants, bien que la question de goût et d'attirance soit nettement plus délicate à traiter dans les enquêtes : Wolfram estime que les jeunes afro-américains refusent de « jouer au blanc » [60] et utilisent ainsi des tournures grammaticales plus proches de leur construction mentale de ce que l'AAV devrait être (Lanehart, 2001).

2.4.4 Phonologie

L'AAV diffère de région en région (de la côte ouest à la côte est, par exemple), mais certaines caractéristiques saillantes sont communes à toutes les variétés d'afro-américain. Tout d'abord, selon Dillard, l'allongement et la monophtongaison des voyelles tendues avant une consonne sourde sont des caractéristiques communes à l'afro-américain et à la phonologie de l'anglais parlé dans le Sud Profond américain (Dillard, 1972).

2.4.4.1 Réalisation des voyelles

Les voyelles de l'AAV sont réparties ainsi (illustration 2.7).

Table 1. Vowels of AAVE

	front	central	back
close	i	ɨ	u
	ɪ		ʊ
		ʊɪ	
close mid	e	əʊ	o
		ə	ɔ
open mid	ɛ		ʌ
	aɪ		
	æ		
open		a	ɑ

Illustration 2.7 – Voyelles réparties selon leurs traits articulatoires (Kortmann, 2004)

60. *Act white*, traduction personnelle.

Ces traits sont relativement proches de leur contrepartie américaine standard (approximativement, le *Midwestern English*) (Ladefoged, 1993). Les voyelles cardinales sont notées en noir, et les voyelles du *General American* (GA), en rouge dans l'illustration 2.8.

Les voyelles /e/, /a/ et /ɨ/ n'existent pas en GA, et la diphtongue //ɔɪ/ afro-américaine est en fait

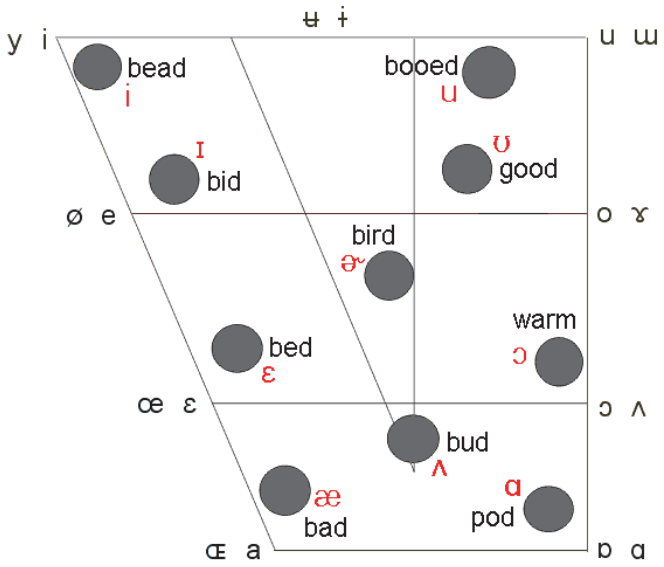


Illustration 2.8 – Monophtongues de GA (Ladefoged, 1993, 1999)

réalisée /ɔɪ/. Enfin, il n'y a pas de diphtongue ascendante comme /eɪ/ en AAV.

Le tableau suivant 2.10 (Kortmann, 2004) montre les réalisations AAV des classes d'isophones de Wells, ainsi que leur comparaison avec l'anglais standard américain (illustration 2.10).

Presque tous les sons sont réalisés différemment en AAV, à l'exception des classes d'isophones Lot, Foot, Goat et Goose, soit des voyelles d'arrière, ainsi que la diphtongue du set Face. L'AAV a aussi une tendance à diphtonguer les sons qui sont prononcés en américain standard avec une voyelle rhotique, ou suivie d'un /r/ comme c'est le cas pour Square, Start, Cure et Force. Dans ce cas, l'attaque de la diphtongue correspond à la voyelle réalisée en américain standard, et la diphtongue est centralisante (sa coda est toujours schwa).

CLASSE D'ISOPHONE	AAV	PHONÈME	ANGLAIS AMÉRICAIN STANDARD
PALM	/ˈpæm/	/æ/	/ɑ/
TRAP	/ˈtræɛp/~/ˈtræɛ/	/æɛ/	/æ/
BATH	/ˈbæɛθ/~/ˈbæɛt/	/æɛ/	/æ/
MOUTH	/ˈmæəθ/~/ˈmæət/	/æə/	/aʊ/
SQUARE	/ˈskwæə/~/ˈskæə/	/æə/	/ɛr/, /ær/
LOT	/ˈlɑt/, /ˈlɑʔ/	/ɑ/	/ɑ/
CLOTH	/ˈklɑθ/~/ˈklɑt/	/ɑ/	/ɔ/
START	/ˈstɑːt/	/ɑː/	/ar/
PRICE	/ˈprɑːs/	/ɑː/	/aɪ/
NEAR	/ˈnɪɤ/, /ˈnɪɻ/	/ɤ/, /ɻ/	/ɪr/
NURSE	/ˈnɤs/, /ˈn̩ɻs/	/ɤ/, /ɻ/	/ɜ̂/
KIT	/ˈkiət/, /kiɪt/	/iə/, ~/iɪ/	/ɪ/
DRESS	/ˈdrɛɪs/	/ɛɪ/	/ɛ/
FLEECE	/ˈflɪɪs/	/ɪɪ/	/i/
FACE	/ˈfeɪs/, /ˈfeis/	/eɪ/	/eɪ/
STRUT	/stʌt/, /stʌʔ/	/ʌ/	/ʌ/
FOOT	/ˈfʊt/, /ˈfʊʔ/	/ʊ/	/ʊ/
CURE	/ˈkʊə/	/ʊə/	/ʊr/
NORTH	/ˈnɔəθ/, /ˈnɔəf/	/ɔə/	/ɔː/
THOUGHT	/ˈθɔʊt/, /ˈθɔʊʔ/	/ɔʊ/	/ɔ/
FORCE	/ˈfɔəs/	/ɔə/	/or/
GOAT	/ˈgot/, /ˈgoʔ/	/o/	/o/
GOOSE	/ˈgus/	/u/	/u/
CHOICE	/ˈʧoɪs/	/oɪ/	/ɔɪ/

TAB. 2.10 – Réalisations vocaliques en AAV(KORTMANN, 2004)

2.4.4.2 Réalisation des consonnes

L'AAV utilise les consonnes suivantes, qui sont sensiblement les mêmes que celles de l'anglais standard. Elles sont présentées dans le tableau 2.11.

	Labiales Labiodentales	Dentales Alvéolaires	Palatales	Vélaires Glottales
Occlusives	p b	t d		k g
Fricatives	f v θ ð	s z		h
Affriquées		ʧ ʤ		
Nasales	m	n		ŋ
Liquides		l	r	
Semi-voyelles	w		j	

TAB. 2.11 – Consonnes de l'AAV classées par mode et lieu d'articulation

Si l'AAV utilise les mêmes sons que l'anglais standard, les consonnes peuvent être réalisées différemment ; et l'influence de l'ABS peut expliquer certains de ces phénomènes. Les codes suivants sont utilisés pour lire le tableau 2.12(adapté de KORTMANN, 2004), qui résume les distributions des usages phonologiques entre AAV et ABS :

— Quand la caractéristique n'est présente qu'en AAV, elle est symbolisée par un <⊗>.

— Quand la caractéristique est commune à l'AAV et à ABS et utilisée dans la même proportion, elle est marquée avec un symbole (<≡>).

— Quand elle est plus représentée en AAV mais qu'elle existe également en ABS, elle est représentée par une étoile (<⋆>).

Consonnes AAV	Exemple	Variantes	Réalisations	Parenté avec l'ABS ?
/t/, /d/ en coda syllabique	*Not, bad*	occlusives glottales /ʔ/	Not[ˈnɔʔ] Bad [ˈbæt], [ˈbæʔ]	⊗
		Dévocalisation de /d/ ou amuïssement total	Bid [ˈbɪt], [ˈbɪ] Good [ˈgʊt], [ˈgʊ]	⊗
/z/, /v/	*Isn't, business, seven, eleven*	*Fronting* des sons, réalisés en stops quand ils précèdent des nasales	Isn't [ˈɪdnt] ; Business [ˈbɪdnɪs] ; Seven [ˈsebn]	≡
/θ/, /ð/	*Thing, those*	En initiale et en finale, occlusion de ces C. À l'intérieur d'un mot, pouvant être réalisées comme des labiodentales /f, v/	Thing [ˈtɪŋ] ; Those [ˈdoz] ; With [ˈwɪt] ; Tenth [ˈtent] ; Bath [ˈbæf] ; Faith [ˈfeɪf] ; Mother [ˈmʌvə], [ˈmʌvr̩]	⋆
/ɹ/	*Floor, bird*	Réalisation vocalique ou chutes quand postvocalique (surtout après des V non centrales, non-accentuées	Floor [ˈfloə], [ˈflo] ; Bird [ˈbɣd], [ˈbɹ̩d] ; Record [ˈrekəd], [ˈrɛkɣd], [ˈrɛkɹ̩d] ; Carol [ˈkɑəl]	⋆
/l/	*Help*	Réalisation vocalique [w] après des V arrières ou chutes quand postvocalique et préconsonantal	Roll [ˈro] ; School [ˈskuw] ; Feel [ˈfil], [ˈfiw] ; Football [ˈfʊbɔw]	⋆
		Amuïssement après [e] mais rares après des V avant hautes.	Help [ˈhɛp] ; Bell [ˈbɛw]	⋆
/j/	*Computer*	Amuïssement optionnel du son	Computer [kəmˈputɪ] ; Beautiful [ˈbutɪfl̩]	⊗
/n/	*Man*	Nasales parfois omises en coda syllabique (nasalité transférée sur la V antérieure)	Man [ˈmæ̃n] ; Bang [ˈbæ̃]	⊗
/t/, /d/ en groupes consonantiques finaux	*Cold, left*	Le 2è son d'un groupe consonantique est souvent omis quand ils partagent le même trait [± voc], surtout dans des mots monomorphémiques, surtout s'ils sont suivis d'un mot commençant par une C constrictive	And [ˈæn] ; Left [ˈlɛf] ; Desk [ˈdɛs]	⋆
/s/ + stop	*Ask, grasp*	Métathèse dans certains mots	Ask [ˈæks] ; Grasp [ˈgræps]	≡
/k/, /t/ dans des groupes en /str/	*Street*	Dans certains mots, [t] est réalisé à l'arrière de la bouche et devient [k]	Street [ˈskrit]	⊗

Tab. 2.12 – Variations consonantiques en AAV

La plupart de ces variations ne sont plus communes à l'ABS, ou sont en net recul dans ce dialecte. Certaines réalisations, notamment celles des nasales, rappellent les différences de réalisations vocaliques et de classes d'isophones contenant des voyelles rhotiques.

2.5 Statut sociolinguistique des systèmes considérés

Après ces présentations formelles, il convient de voir comment la recherche scientifique considère l'AAV et le scots. Les dictionnaires (et leurs noms) seront d'abord envisagés dans ce qu'ils montrent de la langue, autant dans la rédaction des entrées et des sources utilisées. Les stratégies mises en jeu par le milieu universitaire et ses linguistes, philologues, ou lexicographes sont des démarches prescriptivistes ou descriptivistes, qui sont symptomatiques des traitements subis par les locuteurs de ces langues.

2.5.1 Dictionnaires et ouvrages de référence

La différence entre les dictionnaires d'AAV et ceux de scots semble se situer au cœur même du statut très différent de ces deux variétés de l'anglais. Le dictionnaire de Geneva Smitherman, *Black Talk : Words and Phrases from the Hood* [sic] *to the Amen Corner* ne se définit pas lui-même comme ouvrage de langue écrite, mais d'oralité, en grande partie non attesté par des sources littéraires mais par des utilisateurs de l'AAV (surnommées *sistas* par Smitherman). Une recherche sur *WorldCat* sur les dictionnaires et ouvrages de référence sur l'afro-américain renvoie à treize ouvrages.

Le titre de six d'entre eux contient une composante orale : soit vus sous l'angle de l'émetteur avec « talk » à trois reprises, dans des expressions telles que *Black talk* [61], *Afro- American Talk* [62], ainsi que *how they talk* [63], *jargon* [64], soit de la part du destinataire, avec le mot *ear* [65]. Certains retiennent, plus curieusement, l'aspect phonétique de l'afro-américain avec une écriture phonétique du titre du livre. Le dernier exemple, « The Third Ear » semble représentatif du problème de l'afro-américain : quelle écriture pour cette variété essentiellement orale, qui fait appel à la fois à une connaissance des codes de l'anglais américain standard et de codes socio-culturels d'origine ouest-africaine ?

L'ouvrage de Kearse, *Street Talk* (2006) qui est l'un des dictionnaires les plus récents regroupant le hip-hop, l'argot de rue et l'AAV, se subdivise en plusieurs catégories : l'auteur estime que les commu-

61. Smitherman, 2000.
62. Major, 1971.
63. Morton, 1964.
64. Claerbaut, 1972.
65. Roberts, 1971.

nautés qui utilisent cet argot sont les communautés hip-hop et urbaines « this book is about the unique words, metaphors and terms which make up the Hip-Hop and urban slang language. […] Part of understanding Hip-Hop and Urban « street » culture is understanding how amorphous and ever-changing the slang language is. » (‘*ce livre traite des mots, métaphores et termes uniques qui constituent l'argot du hip-hop* [66] *et de la Rue. […] Comprendre la culture hip-hop et urbaine « des rues », c'est en partie comprendre à quel point l'argot est informel et inconstant*’.) (p. xi de l'introduction). Il propose même un néologisme, « slanguage », afin de parler de ce sociolecte à mi-chemin entre langue et argot.

L'auteur fait même une distinction entre l'argot jusqu'à l'année 1995 incluse (argot *Old School*, ou vieille école) et l'argot contemporain, nouvelle école, entre 1996 et le moment de la rédaction de l'ouvrage (paru en 2006). La raison pour laquelle l'argot passe d'une période à l'autre n'est pas expliquée dans l'ouvrage, mais l'auteur estime le passage au hip-hop nouvelle école à l'année 1996 [67].

Les exemples et les entrées du dictionnaire sont également fortement teintées d'oralité : « a common phrase like “know what I mean ?” is transformed into “na'mean ?” », ‘*une expression courante comme « (tu) vois ce que je veux dire ? » devient « twa ? »* ’(page xi de l'introduction).

L'une des définitions du dictionnaire concerne le mot *please* (‘*s'il vous plaît*’), et présente cette graphie alternative : « po-leeze », dont la definition est :

> phrase (general sl.) old & new school 1. A playful or sarcastic way to say “please” in attempt to dismiss someone who is trying to tell you something unbelievable or ridiculous [usually in relations to a lie or excuse]. Ex : “You think I believe that, po-leeze” (p. 436).

La graphie <po-leeze> est un exemple d'allongement phonétique de mot, codée par un allongement graphique, avec scission du mot par un tiret. Le <o> de *po-leeze* représente une voyelle excrescente, et témoigne d'une resyllabification du mot (avec accent placé sur la seconde syllabe) : /pəˈliːz/ ou /pʌˈliːz/pourraient en être des transcriptions valides, qui montrent l'allongement du temps d'attaque vocale (*voice onset time*).

Le /i/ est éventuellement allongé avec la graphie <ee> au lieu de <ea>, qui code presque exclu-

66. Le hip-hop semble se différencier du rap par le fait que si le rap est un style de musique, le hip-hop est également le nom de la sous-culture qu'il représente.

67. Certaines classifications établissent le *real old school* de 79 à 86, l'ère classique de 87 à 90, l'âge d'or de 91 à 96, le *new school* de 97 à 2001, suivie par la période du *crunk* de 2002 à 2006, puis par l'ère de l'*autotune* de 2007 à 2012 (P. Walker & D. Walker, 2012). Les transitions sont dues à l'émergence de nouveaux artistes, ou genres musicaux.

sivement le phonème /iː/, tout comme <ea>. Le choix d'une graphie alternative pour ce segment est douteux parce qu'il n'apporte rien à la transcription, sinon l'éloignement avec la graphie standard. En revanche, la conservation du dernier <e> graphique, pourtant muet à la prononciation (sinon, la graphie aurait pu être « po-lee-ze », voire « po-lee-zee ») est illogique : si la graphie devait être entièrement différenciée de son pendant standard, il aurait certainement été ôté.

Geneva Smitherman préfère, elle, utiliser des graphies standards ou au moins prescrites, et ne s'en éloigne que pour utiliser des graphies déjà connues, et codifiées, *ya* pour 'you'ou *wit* pour 'with', qui illustre le TH-fronting dans la réalisation de /θ, ð/ tout en le rendant homographe avec 'wit', ('*l'esprit*'). Le contexte permet de faire la différence entre les deux mots. Ces graphies correspondent à la réalisation phonologique attendue. Le ratio d'apparition de *ya* par rapport à *you* est de 1 :11, alors que *wit* et *wid* ont le même ratio d'apparition comparés à *with*, 1 :104. Le principal problème d'une telle recherche est le fait que you et with apparaissent tous les deux dans la liste des mots les plus fréquents de la langue anglaise (Fry, Kress & Fountoukidis, 1993), respectivement en 18ème et 15ème position. *Google* bloque le nombre maximal d'occurrences pour chaque mot à 25 270 000 000, ce qui pose des problèmes pour l'exactitude des ratios d'apparition des formes graphiques alternatives.

Une recherche similaire sur *WorldCat* [68], en cherchant tous les dictionnaires de scots, renvoie à près de 400 références. Certains dictionnaires sont réédités, mais une plus grande proportion semble mettre l'accent sur le fait que le scots soit un dialecte (*Scots Dialect Dictionary* [69]).Le mot *tongue* est aussi utilisé pour parler du scots, mais les ouvrages sont moins souvent des dictionnaires que des études linguistiques. Une petite minorité met l'accent sur les usages langagiers actuels, comme *Scoor-oot : A dictionary of Scots words and phrases in current use* (Stevenson & Macleod, 1989). Les dictionnaires écossais ont souvent une vision historique de la langue, ou évoquent le gaélique ; il existe même un historique des dictionnaires (MacLeod & McClure, 2012) ; et certains se proposent comme des glossaires pour comprendre les auteurs écossais prestigieux qui s'exprimaient en scots : *Chambers's Scots dictionary : Serving as a glossary for Ramsay, Fergusson, Burns, Scott, Galt, minor poets, kailyard novelists , and a host of other writers of the Scottish tongue* (Warrack, 1979).

68. <http ://www.worldcat.org/search ?q=su%3AScots+language+Dictionaries.& qt=hot_ subject> dernier accès au 20 janvier 2010
69. *The Concise Scots Dialect Dictionary*, réédition d'un dictionnaire paru six ans auparavant(Warrack, 2006).

Cet aspect historique est complètement absent des dictionnaires d'AAV : ces deux langues sont à des niveaux de développement différents, et si les locuteurs du scots voient leur langue comme vieille et presque savante, inspirée d'une histoire longue, ceux de l'AAV envisagent surtout leur langue comme un parler actuel, dont la dimension orale est plus importante que les utilisations anciennes. Ceci peut expliquer le souci d'inclure des définitions écrites avec des graphies alternatives non standardisées, et la faible représentation des sources anciennes (récits, poèmes…). Ces deux approches s'opposent également sur la question du prescriptivisme ou du descriptivisme sous-jacent dans la construction des entrées et de la lexicographie.

2.5.2 Prescriptivisme ou descriptivisme

Les attitudes envers les deux langues sont différentes, suivant les linguistes et les lexicographes qui étudient les deux langues. Les dictionnaires d'AAV sont écrits à la fois par des locuteurs natifs, dont des non-linguistes (Kearse, 2006) et par des locuteurs universitaires. Les mots choisis dans cette étude illustrent par l'exemple des termes dont la compréhension est utile pour la suite de l'analyse. Les dictionnaires d'AAV contiennent des définitions en anglais standard, qui sont parfois parsemés de graphies alternatives : « Esseys : homeys of hispanic/Latino descent » (Smitherman, 2000). Les définitions des dictionnaires d'AAV et d'argot hip-hop sont plus colorées par des idiomes et des phrases écrites en anglais standard, mais qui appartiennent à l'AAV. Dans *Black Talk*, le mot *whore* renvoie à la définition de *ho* (ainsi qu'à *hoe*) ; et le mot *sister* n'est mentionné que dans la définition *sista* [70]. Il y a donc parfois un biais de lemmatisation autour d'une graphie jugée plus usitée. En revanche, le biais des lexicographes peut être plus grand, étant donné qu'ils peuvent choisir les définitions à incorporer, ou leurs styles. Les insultes, et mots tabous sont tous censurés dans Kearse alors que Smitherman parle librement de *shit*, *bitch* ou *nigger*, qu'elle définit, et ce mot fait partie d'une définition connexe, *nigger apple* (‘*pastèque*’). *Nigga* a sa propre définition, et dix entrées contiennent ce lexème [71].

Kearse n'évoque pas ces mots, et, plus étrange, alors qu'il consacre 22 pages de son dictionnaire,

70. Kearse ne définit pas *sister* ou *sista*, et ni *nigga*, ni *bitch* ne sont mentionnés.
71. « Bad nigga », « field nigga », « house nigga », « nigga mess », « nigga, please ! », « nigga rich », « nigga-toe », « niggamation », « niggas and flies ».

en annexe, à l'affixe <izzle> (voir 2.6.3.1.1), il n'inclut aucune entrée pour *nizzle*, l'un des deux mots forgeurs de cet affixe, dans les paroles *fo' shizzle my nizzle*(*'for sure my nigga'*). *Shizzle* n'est pas défini non plus, comme si cette expression n'existait pas. Cette carence peut être expliquée par deux faits : d'abord, Kearse est présenté comme un « slanguiste urbain » [72], un autodidacte de la linguistique. Il reconnaît en préface avoir omis les mots injurieux ou discriminants sur des critères de genre, d'appartenance ethnique, de préférences sexuelles, et de religion, tout en affirmant que l'argot dont il parle n'est pas caviardé. Cette attitude de prudence et de censure en amont est peut-être due à cette absence de légitimité (universitaire ?), ou une peur de choquer, comme on peut le voir avec le traitement des mots tabous, qui sont toujours censurés : *sh*** pour *shit*, ou *bit*** pour *bitch*. C'est le seul ouvrage sur la culture hip-hop (et non afro-américaine) qui soit censuré, Richardson et Campbell parlent librement de *nigga* (Kermit Ernest Campbell, 2005 ; Richardson, 2006) et ne censurent pas leurs lexiques, alors que leurs livres sont également tournés vers la génération ou la culture hip-hop. C'est le problème du politiquement correct ou de l'autocensure. Nonobstant les origines d'une telle pratique, ces formes de censure sont négatives sur le long terme puisque ces attitudes hypocrites entraînent un déficit d'informations pour les lecteurs, et un décalage avec les usages en cours dans les communautés afro-américaines. Le terme *sista* ou *sister*, fréquemment utilisé en AAV [73] car utilisé originellement par les églises et congrégations religieuses dont les membres sont majoritairement afro-américains [74] peut éventuellement être vu comme condescendant (surtout hors contexte religieux, surtout quand il est dit par un homme à une femme), mais c'est un terme de cohésion, dont l'absence du livre de Kearse est inexpliquée, puisque sa définition est positive dans *Black Talk*.

Smitherman adopte une attitude tout à fait différente, et inclut des mots discriminants (envers les afro-américains, ainsi qu'envers les caucasiens, comme en attestent les entrées *the man, devil, white white, whitey/honky, wigga/er*…). En introduction, elle décrit les évolutions entre les différentes déterminations de la langue et du peuple qui la parle. Cette approche est moins hypocrite, et Smitherman souligne elle-même les biais axiologiques qui existent dans la communauté afro-américaine, avec l'utilisation de deux expressions opposées, *good* et *bad hair*, qui désignent respectivement les cheveux

72. *Urban slanguist*, sur la quatrième de couverture.
73. J'ai moi-même été appelée *sister* pendant des discussions en ligne avec des afro-américains.
74. African Baptist Church, African Methodist Episcopal Church…

lissés artificiellement et les cheveux naturels ou crépus. Ces vues discriminatoires ne sont plus de mise de nos jours avec l'essor de l'« afrocentricité », ou mouvement visant à valoriser les Africains et les populations appartenant à la diaspora africaine dans le monde (Asante, 1988).

Smitherman prend donc l'aspect historique en compte : elle cite des phrases sans contexte, certainement construites, mais aussi des exemples livresque. Pour définir *good hair* (des cheveux qui ressemblent à ceux des Caucasiens, qui ne sont pas crêpus [75]), elle utilise la citation « good hair was also once refered as *nearer, my God, to thee* » (Hurston, 1938/1990). Richardson donne le nom des artistes ayant utilisé des expressions pour la première fois en chanson, ce qui revient non à identifier le forgeur inconnu [76], mais à donner une perspective historique et confirmée des mots qu'elle liste en annexe de *Hiphop Literacies*. Sa liste de vocabulaire est construite de façon ascendante : d'abord, elle donne des exemples attestés en musique, des citations de définitions provenant de quelques sources dont…Smitherman, ce qui renforce à la fois le sérieux de sa liste, et montre l'importance de l'interaction entre chercheurs qui travaillent sur le même sujet, sa propre définition si aucune n'existe, et enfin le nom de l'artiste et de la chanson. Par exemple, l'expression *get busy* est mise en contexte dans une citation du refrain de la chanson « get busy […] get jiggy/get crunked up perculate » de Sean Paul (Paul, 2003), les trois expressions *get busy*, *get crunked* et *perculate* sont définies par des citations de *Black Talk* (Smitherman, 2000), d'une émission intitulée *Slanguistics* (MTV2, 2004), et d'une définition qu'elle écrit elle-même.

Les dictionnaires d'AAV donnent plus d'exemples et de phrases en contexte que les dictionnaires de scots. SDD (Warrack, 2000) n'en contient aucune, tout comme le *Online Scots Dictionary* (OSD) (Eagle, 2002) qui liste tout de même les expressions relatives à chaque lemme, ainsi que chaque graphie alternative pour chaque mot. Ainsi, une recherche en anglais sur le mot *English* renvoie à plusieurs définitions différentes : *England, English/Ingles, Englifee, Sassenach, sooth* et ses sept dérivés. Cet ordre va du plus spécifique (pour parler de l'anglais, de l'Angleterre ou de l'anglicisation) jusqu'à la définition la plus vaguement reliée à l'Angleterre, puisque le mot *sooth* désigne le sud géographique, mais aussi îlien. Il n'est pas fait mention de l'origine des mots ou de la région dans laquelle ils sont

75. Adaptation personnelle de « Hair that is not naturally tightly curled […] hair akin to that of whites. »
76. Dont je rappelle la définition : il s'agit du premier utilisateur d'une expression, normalement inconnu puisque ces expressions sont essentiellement orales (Lecercle, 2004)

employés.

Ce n'est pas le cas du *Dictionary of the Scots Language* (DSL) (Skretkowicz, Rennie, Craigie & University of Dundee, 2001), qui, à cause de son aspect historique (et officiel), et son couplage avec un corpus de textes littéraires, contient des citations beaucoup plus orientées sur les attestations historiques que sur la rédaction des définitions. Si on reprend le mot *sassenach* du OSD, on voit que qu'il a plusieurs graphies : *Sassanoch, -enagh, Sasennach, -unnach*, une étymologie clairement identifiée (le gaélique *Sasunnaich* '*Saxons*)'et sept exemples d'attestation entre le 18ème et le 20ème siècle, uniquement dans son emploi en tant que nom commun. Il a une entrée différente pour ses emplois en tant qu'adjectif, assortie de sept autres attestations, qui sont toutes situées géographiquement en Écosse.

Les dictionnaires de scots sont implicitement des dictionnaires anglais-scots (cela est observable dans les dictionnaires en ligne, qui donnent des traductions de scots en anglais et vice-versa).Il n'est pas rare qu'un dictionnaire s'accompagne d'une série de conventions d'écriture, dont nous avons parlé en section 2.2.2, qui aborde la graphie à employer.

2.6 Analyse des représentations

2.6.1 Statut du Scots et de l'AAV

Il s'agit de langues régionales, mais cette définition est trop vague : l'AAV est panaméricain et le scots est parlé dans d'autres régions du monde, au Canada principalement, dans la province de Nova Scotia notamment. En revanche, il existe une stigmatisation de l'utilisation de l'AAV, perçu comme étant un dialecte substandard de l'anglais américain (voir 2.6.2)) et le scots est parfois tourné en ridicule ou amalgamé avec le gaélique, et la perception de l'accent peut changer diamétralement suivant le coénonciateur. La linguistique sauvage et les perceptions des langues varient énormément en fonction du lieu et de l'origine du commentateur.

Ce sont néanmoins des langues reconnues à part entières, mais elles ont un statut très ambivalent.

Ces langues et leurs locuteurs sont toujours en butte à des amalgames stéréotypants et stigmatisants. Le problème principal ne semble pas tant être le degré d'exposition des Britanniques au scots, ou des Américains à l'AAV, puisque des médias populaires exposent au moins des versions épurées des deux langues, grâce à certaines célébrités. L'accent de Sean Connery est certainement peu représentatif du scots, mais Ewan McGregor utilise encore certaines caractéristiques du scots après une carrière cinématographique d'abord britannique, puis internationale (« Ewan McGregor Interview », 2011). Quant aux célébrités afro-américaines s'exprimant en AAV, la liste est très longue, mais l'on peut citer un homme comme Ice T (nom de scène de Tracy Marrow), chanteur de hip-hop, qui marqua les esprits par le sulfureux titre « Cop Killer » lorsqu'il était chanteur dans un groupe de hard rock (« Body Count », 1992) puis devient acteur tout en gardant un parler typique.

L'antipathie certaine envers les locuteurs des deux langues étudiées peut être le fruit de leur statut de langue minorisée (ou langue de minorités) dans chacun des deux pays.

2.6.2 Portrait des locuteurs

2.6.2.1 Dialectologie perceptive

La dialectologie perceptive qualitative des accents peut être un bon point de départ pour estimer les opinions portées sur les locuteurs et donc les langues AAV et scots. Les points de vue des populations britanniques et américaines montrent certaines similitudes. Les accents jugés trop forts, ou caractéristiques de classes ouvrières sont les plus mal jugés, pour des raisons de compréhension et également de stigmatisation : l'accent [77] et la possibilité de moduler la fréquence des traits propres à l'AAV et au scots (ou n'importe quel dialecte ou langue étrangère) lorsque l'on s'exprime dans une langue-cible, est valorisé.

77. De façon empirique et perceptive, le concept « accent » représente non seulement la phonologie, mais également, dans une moindre mesure, les utilisations lexicales ou syntaxiques ressenties à l'écoute d'un discours oral.

2.6.2.2 Différences sociales et ethniques

Si certains dialectes de l'anglais comme le *cockney* ou le *scouse* de Liverpool ne jouissent pas d'une bonne réputation, les locuteurs du scots sont beaucoup moins stigmatisés : c'est une population plus âgée et plus respectée ; plus blanche, également [78]. Un accent écossais peut être synonyme de réussite sociale, ou au moins de qualités comme le sérieux au travail. Comme dit plus haut, beaucoup d'Écossais n'ont pas l'impression d'utiliser une forme dialectale différente de l'AES. De la part des étrangers qui pratiquent le scots, il y a souvent un lien personnel ou familial (par exemple, les membres de *Scotster* qui ont de la famille en Écosse veulent parler comme leurs ancêtres). D'une façon générale, l'engouement pour une langue écossaise d'étrangers ou de non-Écossais va très fréquemment se porter sur le gaélique, étant donné que c'est une langue dotée d'un certain prestige et d'une aura mystérieuse. Les gens attirés par le scots connaissent son existence, et des bribes de mots au préalable avant de vouloir l'apprendre.

Il y a toujours eu un sentiment anti-écossais en Angleterre (notamment des rumeurs selon lesquelles les Écossais auraient été cannibales au 17ème et au 18ème siècles, pendant les révoltes jacobéennes), qui s'est manifesté à nouveau lors de la dévolution des pouvoirs à l'Écosse et au Pays de Galles entre la fin du 20ème siècle et le début du 21ème siècle (H. Walker, 2007).

Certains régiolectes sont connus de la population générale, et n'ont pas le même statut que le scots, pris comme un ensemble de pratiques linguistiques. C'est le cas des régiolectes des grandes villes, le *patter* de Glasgow et la variété de scots parlée à Édimbourg. Les locuteurs du *patter* de Glasgow sont peu ou prou sur la même ligne que les afro-américains : ils sont jeunes, viennent de milieux urbains, et parlent un langage qui est peu compréhensible (un argot urbain et réel régiolecte similaire au *cockney* ou au *scouse*). Le *patter* est parlé notamment à Renfrew, Paisley et Clydebank, les banlieues à une dizaine de kilomètres de la ville. Étant donné que ce dialecte est moins connu que d'autres variétés dialectales parlées en Angleterre, par exemple, il ne fait pas partie des dialectes discriminés par les Britanniques. À Édimbourg, les questions régiolectales sont différentes : l'activité économique de la

78. Il serait hypocrite d'écarter la couleur de peau comme facteur atténuant la discrimination– les scottophones sont sans doute plus blancs que les locuteurs du *scouse*. C'est le cas, à en juger les photographies de profil de SC

ville entraîne un brassage important des populations en présence, et le scots qui y est parlé (l'*Embra*) a subi une érosion lexicale importante. La prononciation ressemble à celle du scots du sud du pays (Eagle, 2001b).

Les joutes et concurrences entre les Édimbourgeois et les Glaswegiens entraîne parfois des conflits ludiques entre les deux communautés[79] , par exemple à propos du fait que les Édimbourgeois se veuillent plus raffinés que les Glaswegiens, qui considèrent leur ville comme plus moderne. Les plaisanteries à propos du *patter* tournent autour d'expressions glaswegiennes comme *y'awright pal ? How's it gaun ?* '*hello, how are you ?*' ; *wit ur you aw aboot ?* '*what are you talking about ?*' qui occupent un chapitre du livre. Les attaques envers Édimbourg sont différentes. Puisqu'il n'y a pas de dialecte particulier à la ville, les seules références à la façon de parler sont très rares : « spoke in a strangely affected manner » est le seul exemple, et les tournures édimbourgeoises sont écrites en anglais standard.

L'AAV est décrié, parce que parlé par les Afro-Américains les moins favorisés. Il existe toujours un effort pour gommer son accent régional et ou ethnique. Se donner linguistiquement un air respectable ou sérieux, c'est gommer les traces d'une langue encore perçue comme un anglais substandard, et se rapprocher des règles syntaxiques et phonétiques d'un anglais standard[80]. D'ailleurs, même dans le corpus dans son ensemble (que ce soit *Trainspotting*, les forums de discussion ou *Precious/Push*), on retrouve le même mouvement : le lecte de connivence n'a cours que lorsqu'il n'y a aucune négociation de pouvoir en jeu : en présence d'un représentant de la loi[81], ou dans le cadre où il faut faire preuve d'autorité ou montrer son sérieux, l'alternance codique vers un anglais aussi standard que possible est de mise.

L'AAV peut être vu par certains de ses locuteurs comme une langue noble, surtout par les locuteurs qui ont à cœur la réappropriation du discours des anciens esclaves, dans le cas de méliorations postérieures. Il est porteur d'une histoire et accompagné d'un sentiment de fierté par les locuteurs qui le revendiquent et l'utilisent. En conséquence, des critères axiologiques fondés sur l'histoire peuvent

79. Ces disputes entre *Edinbuggers* et *Weegies* sont illustrées dans (Black, 2003)
80. Plus au niveau des métathèses, par exemple, que du système vocalique, qui dépend en grande partie de la zone géographique.
81. C'est une allusion à l'alternance codique manifestée dans *Trainspotting*. La première remarque de Renton lors de son procès, qu'il gagnera est « Naw. Eh, no, your honour » (alors que son comparse, Spud, qui ne maîtrise pas l'anglais standard, ira en prison : « That's spot on man…eh …ye goat it, likesay ».). Chap. Courting Disaster—Blowing it. Coïncidence… ?

émerger : celui qui refuse ses origines et le sang africain de ses veines devient un *Uncle Tom*, un vendu, ou un *Oreo* [82] (petit gâteau formé de deux biscuits noirs, fourrés d'une crème blanche) [83] dans un anglais plus moderne ; et le caucasien est un *Massa*, ou '*maître*', figure archétypique du propriétaire de plantation cruel et cupide).

2.6.3 Évaluation par les locuteurs d'autres dialectes

L'hostilité envers l'AAV est souvent palpable, et certains sites comme l'UD, qui sont tout de même reconnus et visités (plus de 2 000 entrées ou définitions sont rédigées chaque jour, pour un total de 6 million de définitions, et un trafic de 15 millions de visiteurs uniques par mois). Les visiteurs sont plutôt de jeunes adultes caucasiens, entre les âges de 18 et de 24 ans [84].

Pour chaque définition, les entrées sont publiées à la manière d'un wiki : elles sont envoyées par des rédacteurs anonymes (qui ne sont identifiables que par les définitions qu'ils ont écrites ou leur pseudonyme), et chaque mot peut avoir autant d'entrées que possible. Pour équilibrer, les visiteurs du site (même ceux qui n'ont pas de nom d'utilisateur) peuvent voter pour ou contre chaque entrée. Elles sont alors classifiées au *prorata* de votes positifs (Goudet, 2010a).

Ces entrées se composent souvent d'une description avec un exemple, qui ne provient pas d'une source attestée. Les descriptions et les exemples sont souvent inventés par les rédacteurs des entrées, mis à part les références à quelques chansons de hip-hop, dont *Ebonics*, par Big L, une chanson décrivant effectivement les lexèmes et idiomes afro-américains (« The big picture », 2000), ainsi que celle de Kelis, qui n'est pas particulièrement porteuse d'AAV, mais peut-être d'une certaine forme de culture populaire (Kelis, 2004). Il s'agit donc bien de fiction linguistique, qui ne repose sur aucune fondation sérieuse, mise à part la perception des contributeurs.

82. *Bounty* n'est pas utilisé aux États-Unis parce que cette barre chocolatée n'y existe pas, mais c'est un terme employé en France avec le même sens.
83. La comparaison avec la nourriture se retrouve aussi chez certains Asiatiques : *banana* est employé de la même manière pour désigner le fait d'être vendu à la cause des Blancs.
84. Cette section reprend des éléments de la communication présentée au colloque « Phonologie de l'Anglais Contemporain » (PAC) en septembre 2010 (Goudet, 2010a)

Les entrées de l'UD sont très utiles pour confirmer le sens d'un mot appartenant à l'argot des jeunes ou d'internet. Les notes renforcent la légitimité d'une entrée, puisqu'elles émanent d'autres internautes, qui émettent des jugements de pair sur leur légitimité.

2.6.3.1 *Ebonics* dans l'*Urban Dictionary* (UD)

Le mot *ebonics*, qui correspond à l'appellation archaïque et presque méprisante de l'AAV est défini dans 103 entrées du dictionnaire. À ce propos, la liste de mots connexes est édifiante : « slang, black, ghetto, **nigger**, nigga, **wigger**, rap, language, gangsta, hood, urban, **negro**, **stupid**, white, english, ebonic, **jive**, talk, ebony, **shit** [85] ». On se trouve donc bien dans une vision non afro-américaine, dans la perception d'autres populations hostiles ou au moins peu ouvertes à l'AAV. Les rapports ethniques sont prégnants, comme l'indiquent la présence de *white* et *black*, autant que de *nigger*, *negro* et *nigga* Quant à *wigger*, il s'agit de la valeur repoussoir, figure décriée dans les deux camps du jeune blanc qui se pare des atours ou de caractéristiques afro-américaines. Ce mot est un amalgame de *white* et de *nigger*, et il est communément accepté comme péjoratif et méprisant.

Ebonics s'inscrit dans la même lignée que *slang*, *language*, *english*, *jive* et *talk*, dans le sens où c'est effectivement un mot employé pour parler de cette langue à forte dimension orale.

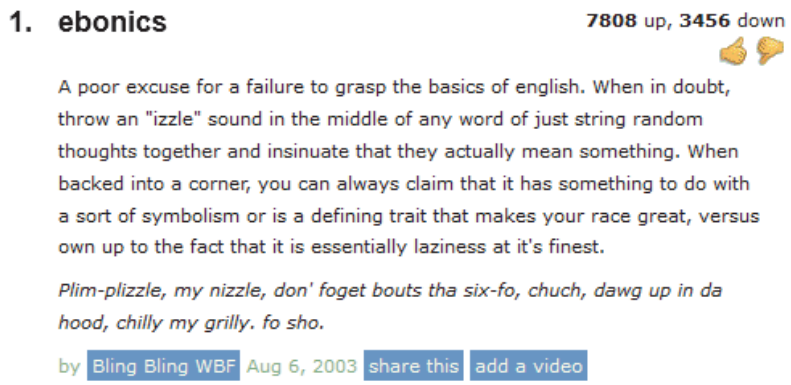


Illustration 2.9 – Entrée la plus populaire d'Ebonics avec nombre de votes, au 17 février 2013.

85. Les mots ont été mis en gras par mes soins. Je n'ai relevé que les mots utilisés comme des insultes, complété par les stéréotypes auxquels les Afro-Américains sont soumis comme *hood*, *gangsta* ou *ghetto*.

Cette entrée est découpée en un titre, une définition, et des exemples en italique. Elle est tout à fait représentative de la façon dont l'AAV est estimé par les non locuteurs. Le son « izzle » évoqué dans la définition est traité plus loin dans l'analyse (voir 2.6.3.1.1). L'exemple est un charabia incompréhensible, une sorte de patchwork entre des phrases et des idiomes AAV comme *fo sho* '*for sure*' (qui est une figuration graphique de la phonétique), ou *my nizzle* '*my nigga*', qui code clairement un AAV urbain, voire parlé par des membres de gangs. L'utilisation de mots comme *hood*, *chuch* (formule phatique utilisée à la fin d'une phrase assertive ou adverbe exprimant l'accord, utilisée dans l'argot de gangs), l'évocation de la *six-fo* ('*Chevrolet Impala 64*'), un modèle de voiture prisée par les membres de gangs pointent vers une connaissance parcellaire de l'AAV, certainement orientée par les éléments culturels apportés par la musique. *Chuch* est souvent utilisé par des artistes de hip-hop afro-américains, et il n'est pas difficile d'extrapoler et de mélanger des échantillons d'expressions. L'amalgame entre identité et linguistique est visible dans les chansons de Snoop Dogg, qui évoque souvent son Impala dans des chansons célèbres, et participe ainsi à la création de stéréotypes (« Still a g thang », 1998).

Les éléments intéressants sont dans les exemples donnés par les contributeurs. Certains sont dans l'exagération et la moquerie d'un style de vie qui s'apparenterait à un ghetto, d'autres sont relativement fidèles. Les phénomènes phonologiques décrits par les utilisateurs de l'UD sont de différents ordres. Il y a des marqueurs de coarticulation, et la transcription de l'AAV est souvent oralisée, comme décrit dans les exemples suivants. « Words that end in "oor" change to "oh". Words with "ing" change to "in' ". Words that are next to each other in the sentence are sometimes combined. » (entrée 17).

On y trouve des haplologies comme *somin* ('*something*'), *ignant* ('*ignorant*'), *muhfucka* voire le courant *mofo* ('*motherfucker*'). Les élisions sur des modèles courants comme *gonna avec sum colta* ('*something cold to [drink]*'), des réductions de groupes consonantiques comme *hurchoself likeat* ('*hurt yourself like that*') sont souvent décrits « Frequently involves droppin' the las' consonan' of a word. » (entrée 19). Les exemples font également mention des apocopes et aphérèses comme *'afore* ('*before*'), *bout* ('*about*'), *hood* ('*neighborhood*') qui fait effectivement partie du vocabulaire courant (ou au moins, des apocopes acceptées) et *nig* ('*nigger*'). Les choix graphiques sont également denses en exemples, comme le montre le tableau 2.13. Beaucoup de ces mots n'ont pas d'entrée propre, ils ne

sont souvent utilisés qu'à titre d'exemple graphique de l'AAV.

Mot	Transcription	Tokens	Sens
Yo	/ˈjoʊ/	47	You, your, yo (interjection)
Da	/ˈdə/	36	The
Dat	/ˈdət/, /ˈdæt/	27	That
Fo	/ˈfoʊ/	21	For
Shizzle/Shizzel	/ˈʃɪzl/	20	Sure (+ suffixe <–izzle>)
Kno	/ˈnoʊ/	15	Know
Dis	/ˈdɪs/	13	This
Mah	/ˈmæʰ/	9	My
Tha	/ˈdə/, /ˈðə/	9	The
Nizzle	/ˈnɪzl/	9	Nigga

Tab. 2.13 – Les dix mots les plus fréquents dans les exemples de l'UD

Certains vont jusqu'à faire des graphies alternatives représentant la phonologie *languigiz* (‘*languages*’), *exampul* (‘*example*’), *sho* (‘*sure*’) *rite* (‘*right*’). La même attitude discriminante que celle du scots est parfois utilisée : les mots prononcés de la même façon en AAV et en anglais sont orthographiés avec des graphies alternatives, pour se démarquer du standard : *like* est ainsi représenté par *lyk*, et ce changement n'est que stylistique. Ces graphies alternatives tentent de s'éloigner le plus possible de la graphie standard de l'anglais, et il n'y a pas une seule entrée pour le mot *Ebonics* qui ne contienne pas au moins une graphie alternative.

2.6.3.1.1 *Fo Shizzle my nizzle*

L'utilisation exagérée du suffixe <–izzle>, utilisé pour la première fois par Snoop Dogg [86], un artiste afro-américain célèbre dans l'introduction d'une de ses chansons, « What's My Name ? » (« Doggystyle », 2001), ainsi que par Jay-Z est surreprésenté dans le corpus. Dans cette introduction, Snoop Dogg tient un dialogue en utilisant uniquement des mots contenant <-izzle> et introduit la chanson en parlant sur la musique avec « Izzle kizzle, fo shizzle/my nizzle, what you sizzle/ fo' shizzle bizzle » (‘*Easy Kill, for sure/my nigga, what you say ?/ for sure bitch*’). Cette invention lexicale a frappé les auditeurs, et fait maintenant partie de la caricature de l'afro-américain—la personne—par

86. Anciennement connu sous le nom de Snoop Doggy Dogg ; le rappeur a changé plusieurs fois de pseudonyme : DJ Snoopadelic, Snoopzilla et Snoop Lion en 2013, mais il est revenu à Snoop Dogg en 2014. Je l'ai donc choisi pour désigner Calvin Cordozar Broadus, Jr. .

confusion ou stéréotypie avec le gangster, lui-même imité par certains artistes. Ce mécanisme langagier a une référence identitaire, et appartient à la structure du cliché

Je ne nie pas l'appartenance effective de certains rappeurs ou artistes de hip-hop à des gangs, qui écrivent des chansons à ce sujet comme 50 Cent, qui dirige certains chapitres du G Unit ; Snoop Dogg, qui est un Crip connu et respecté, ou Lil Wayne, un Blood–beaucoup revendiquent un passé de vendeurs de drogues. Les interactions entre le gang et l'industrie du rap sont nombreuses, et sont illustrées par les assassinats de personnages comme Tupak Shakur et Notorious B-I-G, sur des questions liées au banditisme (et au gang auquel leur label de musique appartenait) ; ni le fait que l'un des dictionnaires contenant une entrée sur <-izzle> ait été écrit suite à une peine de prison (Kearse, 2006). Je me permets simplement de souligner la récupération de l'ethos du gangster à la Stagolee par les artistes hip-hop.

Les entrées de l'UD contiennent 64 occurrences de mots contenant <izzle> dans 18 mots différents. C'est l'altération graphique la plus fréquente trouvée dans le corpus. Les utilisateurs réutilisent souvent les mots *shizzle* et *nizzle*. <-Izzle> est traité soit comme un infixe (réduit à <-iz->) *blizzock*, (‘*block*’) ; *bizzich* (‘*bitch*’), soit comme un suffixe, dans sa forme <-izzle>. Les utilisations d'<-izzle> ne sont pas répertoriées dans les entrées écrites par des utilisateurs ayant de vraies connaissances sur l'AAV : « izzle aint used alot in the hood, just cuz Snoop Dogg uses it it doesnt mean that its a part of the language » (entrée 76). <-Izzle> est donc une marque stylistique exagérée, tirée d'un cryptolecte inventé, mais représenté dans la culture hip hop. La part de parodie est non négligeable dans le corpus, autant dans les exemples que dans les définitions, et <-izzle> est une part intégrante de la façon dont certains membres de BP s'expriment. Il y a 13 400 occurrences de « nizzle » et 15 400 occurrences de « shizzle » sur le site, surtout dans les pseudonymes. Il n'est pas possible de voir tous les mots qui portent cet affixe sur BP à moins de faire une lecture humaine de toutes les pages, mais cela prouve que ce marqueur identitaire est actualisé sur la plateforme.

2.6.3.1.2 Description des Afro-Américains

Certains contributeurs donnent des explications en phonologie « sauvage » ou « profane » [87]. La linguistique sauvage est l'ensemble des discours épilinguistiques utilisés pour juger une langue, une variété dialectale ; ces discours sont spontanés et font montre d'une expertise ou d'une rigueur variable. Les jugements axiologiques ne sont pas polarisés autour de *good* → *bad*, mais plutôt autour de perceptions personnelles : *posh*, *dirty*, *ignorant*…

Ainsi, certains décrivent les phénomènes de *TH-fronting* comme des remplacements de <th> (ou /θ, ð/ par d'autres fricatives (/f/) « Words like "Teeth" are replaced with "Teefh" or "Teefhis." Most words with "Th" in the end of the them are replaced with "TF" (pronounced like Tee-Fifth) » (entrée 20) ; la figuration graphique est maladroite et témoigne de la volonté de vouloir garder le nombre de lettres du mot d'origine, dans la figuration graphique <teefh>, où le <h> n'a aucune raison d'être puisque le son prononcé est /f/. La deuxième figuration graphique est plus problématique : <teefhis> montre une absence de respect de la règle de l'alternance singulier-pluriel, qui serait présent en AAV : de /ˈtiːθ/, on passerait par une étape où le morphème <-s> de pluriel serait rajouté, en même temps que le /θ/ serait labialisé, donnant ainsi une transcription /ˈtiːfɪfθ/ (<tee-fifth>) peu convaincante.

La parodie peut se trouver soit dans la définition, soit dans les exemples, et j'ai utilisé les critères suivants pour estimer la présence de parodie. Les résultats sont :

— la surreprésentation de certains traits afro-américains ;

— l'utilisation de mots injurieux ou péjoratifs (*uneducated, hood*…)

— Utilisation de stéréotypes (prénoms jugés afro-américains, situations de trafic dans des gangs…)

87. Traduction personnelle de « folk phonology », lui-même basé sur le modèle « folk linguistics ». Le terme « linguistique profane » est emprunté à M.-A. Paveau (Paveau, 2007).

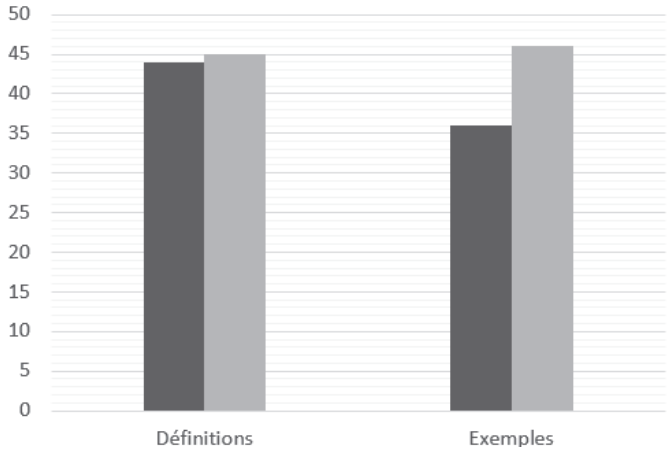


Illustration 2.10 – Proportion de définitions et d'exemples caricaturaux et neutres

La barre en gris foncé représente les utilisations caricaturales, et la barre en gris clair, les représentations non-caricaturales, neutres.

Le nombre total enregistré dans le graphique ne correspond pas au total des entrées sur l'UD puisqu'elles ne comprennent pas toutes des exemples. Les définitions les mieux notées sont souvent celles qui comprennent des éléments injurieux ou caricaturaux : sur les dix premières entrées, suivant leurs notes de popularité, seules trois étaient positives ou neutres. Certaines définitions (moins de 10%) sont positives ou montrent un savoir universitaire en AAV, comme le montre l'entrée 7 :

> African American Vernacular English (AAVE), also called African American English, Black English, Black Vernacular, or Black English Vernacular (BEV), is a type variety (dialect, ethnolect and sociolect) of the American English language. It is known colloquially as Ebonics (a portmanteau of "ebony" and "phonics"). With pronunciation that in some respects is common to Southern American English, the variety is spoken by many blacks in the United States. AAVE shares many characteristics with various Creole English dialects spoken by blacks in much of the world. AAVE also has pronunciation, grammatical structures, and vocabulary in common with various West African languages. Ebonics is not merely just the use of urban or "slang" words, but rather it is the manipulation and transformation of the English language (ie, In the sentence "I want to have sex with a chickenhead", a slang word is used, but the speaker is not using ebonics.)

Ici, l'utilisateur donne une caractérisation fine de l'AAV avec une typologie de la différence entre un dialecte/ethnolecte/sociolecte et l'argot. La connaissance de synonymes pour définir l'AAV ainsi que de ses liens de parenté avec d'autres pratiques langagières (ABS), cette définition est beaucoup plus scientifique que la plupart des autres entrées.

Peu de contributeurs sont donc afro-américains, mais ces rares lexicographes improvisés évoquent

l'aspect historique de l'utilisation de l'AAV : « Ebonics, represents that warm place in the hearts of many of us African-Americans, when we think about our Ancestors, who could not speak nor understand English, but struggled to speak a language they were not taught formally for hundreds of years. » (entrée 5)

2.6.3.2 *Scottish* dans l'*Urban Dictionary*

Quant à la perception du scots par les anglophones qui n'utilisent pas de dialecte, j'ai choisi, toujours sur le même site, la définition de *Scottish* (40 entrées) ainsi que des définitions connexes. J'explique le faible nombre d'entrées aux lexèmes *scots* et *lallans* par la méconnaissance de l'anglais écossais (et du scots) par les utilisateurs de l'UD–le terme *lallans* renvoie est presque toujours littéraire, et n'a jamais été utilisé par les locuteurs eux-même (Costa, 2012). Le mot *patter* ne renvoie pas du tout à des entrées connexes à des problématiques écossaises ; et la première définition, qui évoque un parler propre à une personne ou à un groupe, n'évoque l'origine écossaise qu'en passant, à la fin de la définition. Beaucoup plus de définitions de *scottish* contiennent des informations sur le scots, ou les langues parlées en Écosse.

La vision extérieure sur les Écossais est beaucoup plus positive, avec parfois un léger amalgame avec les Irlandais. Sur les 40 entrées, 19 sont positives (les mieux notées, d'ailleurs), 6 sont factuelles ou neutres, et 15 sont insultantes ou négatives. Dans les dix premières entrées, 7 sont positives et 3 sont neutres, alors que les dix entrées les moins bien notées sont presque toutes négatives (sauf une, plus neutre). « Dude those Scottish people are the shiznit » (entrée 2) montre une bien curieuse alternance codique, entre le vocatif dude qui code un parler jeune américain, plutôt californien, et l'utilisation de *shiznit*, ou *shit* auquel l'infixe <izz(le)> serait ajouté [88].

Certaines dimensions historiques et choquantes sont soulevées, et un contributeur parle de « negroes turned inside out » (entrée 12). C'est un des surnoms donnés par les Anglais aux Irlandais (et aussi aux Écossais dans une moindre mesure). Par conséquent, les populations noires étaient appelées « Irlandais/Écossais fumés ».

88. Cette utilisation a valeur d'emphase positive.

Les langues parlées en Écosse sont évidemment évoquées et la définition 5 est relativement bien documentée :

> A form of Gaelic, closely related to Irish and Manx, part of the Goidelic or Q-Celtic sub family of the Celtic language grouping. Not to be confused with 'Scots', 'Lowland Scots', 'Lallans' or 'Doric' which is a language/dialects of a language of Anglo-Saxon origin, formerly referred to by its speakers as Inglis/Ynglis, which assumed the title 'Scots' around the 15th century.

La définition 7 l'est un peu moins :

> a type of dialect. the scottish [sic] language was originally a form of gaelic [sic], from when the irish [sic] settled in scotland [sic]. That language is rapidly going away and so a form of english [sic] is more commonly used in scotland [sic]

Ce sont les deux seules définitions à qui parlent du gaélique, et la plupart des contributeurs qui évoquent la langue le font surtout dans les exemples. Certains évoquent des traits saillants phonologiques : « not every1 there rolls their R's and says occh I jimmy al the time » (entrée 7). La référence au /r/ roulé est certainement /ɾ/, et 'occh' est la figuration graphique de /ɔx/, une interjection écossaise. La lettre <I> est utilisée, elle, à cause de son homophonie avec aye 'yes' en scots (/aɪ/. Jimmy est le nom stéréotypique donné aux Écossais comme l'est Jock (A. Blake, 2005). Suivant le contexte, les deux peuvent être neutres à péjoratifs. L'utilisateur a ici voulu grouper trois marques écossaises, interjectoire, lexicale et onomastique, afin de créer un concentré de stéréotypes écossais.

D'autres contributeurs évoquent l'accent écossais comme un ensemble linguistique définissant la prononciation, et est perçu positivement ou négativement selon le contributeur : « a horrible accent » (définition 25) est plus négatif axiologiquement que « The Scottish accent is harsh and unpleasing to some people » (définition 40), qui a l'air moins subjectif, mais est tout autant négatif[89]. Certains avis subjectifs s'opposent à ces opinions : « The sexiest accent in the world » (définition 8). D'autres utilisateurs écossais, soulèvent les problèmes personnels qu'ils ont rencontrés à cause de leur accent : « I'm scottish as i was born in Irvine and had a Glasweigen accent . when i was 13 i moved to aberdeen and i kept my accent although i got bullied for it. » (entrée 21). Cela prouve l'antipathie ressentie par certains Écossais envers l'accent de leurs pairs[90], et peut-être les ressentis négatifs envers le patter de Glasgow (Irvine n'est qu'à une demi-heure de Glasgow, et est une ban-

89. Dans le contexte, qui contient, entre autres, une liste assez complète de stéréotypes.
90. Bien que l'âge puisse être un facteur aggravant de persécutions à l'école.

lieue pauvre où les habitants des classes sociales les moins favorisées parlent un mélange de scots et d'anglais écossais standard sans réussir parfois à faire d'alternance codique ou à distinguer entre les registres à utiliser).

Une définition évoque le dorique, forme du scots parlée au nord-est du pays ; et une autre propose une description relativement neutre pour de la linguistique sauvage, qui se caractérise souvent par un jugement subjectif, basé sur la perception et non sur la description ou la prescription (Paveau, 2007), qui repose souvent ici sur des stéréotypes. « The scottish accent may include words from the Scots language but they are not the same thing. Milk is pronounced "mulk" Fish is "fush" » (entrée 32). La figuration de la phonétique des deux mots d'exemples montre la réalisation plus postérieure des voyelles en syllabe fermée. Cette entrée ne donne pas d'illustration quant aux mots de scots évoqués. Cette définition n'étant pas celle de la langue écossaise, il y a peu d'exemples comme celui-ci. Celle de *lallans* contient une phrase d'exemple, qui est un patchwork « représentatif » du scots : « Aye baw-bag, awa oot. Aye workin. Ya ken ? Wir ain leid - Lallans. » ('*hey, stupid, get away. Always working. You know ? Our own language – Lallans*'). Celles de *scots* (6 entrées) ne contiennent pas d'exemples fictifs, mais un extrait de poème qui servit d'hymne national officieux, *Scots Wha Hae* « Scots, wha hae wi' Wallace bled, / Scots, wham Bruce has aften led, / Welcome to your gory bed / Or to victorie ! » (Burns, 1867). Certains ne perçoivent l'appartenance ethnique à l'Écosse qu'au travers de l'utilisation du scots : « If your from Glasgow then you'll use words like..maw..da..whit..aye..naw. ['*mom, dad, what, yes, now*'] If your from aberdeen then you'll use words like..awa..fit..fere..bide ['*away, going (fit like, how is it going ?, healthy, reside*'], If you don't use any of these words then your probably not Scottish. » [sic] (entrée 21).

2.6.3.2.1 Stéréotypie écossaise

Certaines définitions contiennent des stéréotypes à l'égard des Écossais. Étant donné la proportion de contributeurs écossais, plus les entrées contiennent des traits stéréotypiques négatifs, moins elles sont populaires. La première image associée à *Scottish* sur le site est légendée « he is one scottish mother fucker » [sic] (illustration 2.11).

Illustration 2.11 – Représentation stéréotypique individuelle associée à « Scottish » sur l'UD

Cette illustration montre bien la stéréotypie associée au peuple écossais : le côté ancien, le costume complet (kilt, bourse, chaussettes, chaussures), et le *stone put* [91] sont autant de marques d'affection réductrice appliquée au peuple écossais. Dans les entrées, on peut trouver des références au whisky et à l'alcoolisme, aux cornemuses, au haggis ainsi qu'aux violences conjugales, au racisme, à la haine des Anglais, ou à la drogue [92].

L'un des utilisateurs s'interroge d'ailleurs sur les visions trop stéréotypiques sur les Écossais : « wot is with all the other definitions ??? they are completely steriotyped, the people should either get a life or at least VISIT SCOTLAND and see for themselves that the scottish arent all whiskey-drinking, sheep-shagging, kilt-wearing, aggresive pigs » [sic] (définition 7). Cette entrée n'est pas la seule à décrier les points de vue réducteurs sur l'Écosse. Les rédacteurs des entrées sont plus souvent écossais que ceux de l'entrée *Ebonics*, ce qui explique certains règlements de compte entre les entrées, ce qui est moins souvent le cas pour l'AAV, où il n'y a pas de réflexivité ou de dialogue entre les contributeurs. Un utilisateur écossais va cependant jusqu'à défendre certains stéréotypes écossais : « Im scottish and drink

91. Forme de lancer de poids écossaise, encore pratiquée aux Highland Games.
92. Deux définitions évoquent même *Trainspotting*…Mais une seule évoque l'avarice.

whiskey, dont eat haggis, LOVE men in kilts and pipers ! […] am extremely proud of my heritage ! I embrace my steroetype ! » (entrée 27).

Plusieurs entrées contiennent également des confusions entre les Irlandais et les Écossais, un trope connu et documenté, « Scotireland » (« Scotireland », 2010)–voire ancien [93]. Cette confusion est présente dans 7 définitions sur 40, soit dans la perception du gaélique, soit dans celle du peuple : « a type of dialect. the scottish language was originally a form of gaelic, from when the irish settled in scotland. That language is rapidly going away and so a form of english is more commonly used in scotland. » (définition 7). La définition est nébuleuse, le gaélique n'est attribué qu'à l'Irlande, sans reconnaissance de la scission entre gaélique irlandais et écossais. Quant au scots, il est perçu comme « une forme d'anglais », ce qui lui retire son statut de langue minorée. L'appartenance identitaire à la *Celtic Connection*, et les considérations patriotiques contribuent à cette confusion, et l'expliquent en partie. Pour l'observateur extérieur, la langue parlée principalement en Écosse est le gaélique, effectivement importé par les Irlandais qui se sont établis en Écosse. Les initiatives de revitalisation du scots ne sont pas encore assez établies pour permettre une vraie prise en compte du scots, en Écosse même comme dans le reste du Royaume-Uni (Costa, 2008).

Ceci clôt l'analyse des points de vue perceptifs de locuteurs d'AAV et de scots, sur un site généraliste. Les pratiques langagières dépeintes sur l'UD sont souvent exagérées, et les stéréotypies associées y sont bien montrées (en tous les cas, plus que sur un site comme BP ou *Scotster*, étant donné que ces sites sont essentiellement destinés à des publics respectivement afro-américains et écossais). Pour l'AAV, le sentiment d'hostilité est puissant (les définitions les mieux notées sont racistes, méprisantes, ou réductrices). Pour le scots, l'américanocentrisme de l'UD et le sentiment identitaire écossais appuient les définitions les plus positives (Costa, 2009).

Les utilisateurs d'AAV et de scots ont une attitude plus mesurée, et leurs représentations des deux langues minorées sont bien moins caricaturales sur l'UD. Cependant, une certaine stéréotypie est attendue sur BP et *Scotster*, étant donné qu'elle viendra du centre de la communauté (de la part de locuteurs des deux langues minorées). Cette stéréotypie est utilisée, alors, de façon humoristique, sty-

93. Ce phénomène était déjà commenté il y a près d'un siècle : « The Scots (originally Irish, but by now Scotch) were at this time inhabiting Ireland, having driven the Irish (Picts) out of Scotland ; while the Picts (originally Scots) were now Irish (living in brackets) and vice versa. It is essential to keep these distinctions clearly in mind (and verce visa). » (Sellar, Yeatman & Reynolds, 1930).

listique…Cantonner l'utilisation de « fo' shizzle my nizzle » ou « och aye » au simple statut de curiosités linguistiques est réducteur, étant donné que certains locuteurs de l'AAV et du scots les emploient. La caractérisation sur l'UD sert donc à dresser un tableau exagéré, certes, mais basé sur de réelles pratiques langagières. Comme dit précédemment, la linguistique sauvage repose sur la perception de ses auteurs. Ceux-ci ne font pas forcément partie des locuteurs de l'AAV ou du scots, et plaquent des stéréotypes dans leur commentaires de l'AAV et du scots. Les stéréotypes qui portent sur les usages langagiers sont les plus faciles à vérifier dans le corpus. De fait, les utilisateurs de SC et de BP utilisent réellement ces mots, et la caricature se pare de traits attestés.

Conclusion du chapitre

Dans cette partie, j'ai délimité les cadres définissant le scots et l'AAV en tant qu'entités langagières utilisées en Grande-Bretagne et aux États-Unis, autant dans la littérature linguistique que dans des sites généralistes contenant des discours épilinguistiques (« Urban Dictionary », 2003), ou des sites à visée linguistique, qui proposent une vraie description de la langue (Eagle, 2001a). Ces deux langues minorées possèdent des caractéristiques phonologiques particulières et reconnaissables, qui se traduisent par des utilisations graphiques qui se démarquent de l'anglais standard des deux pays dans lesquels ils sont pratiqués.

Leurs particularités lexicales les font diverger le plus de l'anglais standard, soit à cause de cognats au sens sensiblement différents (*represent* ou *dig* pour l'AAV, *reek* ou *bra(w)* '*good*' pour le scots), soit à cause de mots inédits, qui ne sont jamais apparus en anglais standard, ce qui fait pencher la balance du côté d'une langue complète, avec des contraintes grammaticales et syntaxiques particulières. Seuls certains locuteurs sont réellement versés ou locuteurs actifs de ces langues, qui se diluent dans l'anglais (parlé tous les jours). En revanche, la majorité des locuteurs de ces lectes possède au moins un vocabulaire passif, plus ou moins important suivant leur exposition aux langues minoritaire (suivant également leur région, leur âge…).

Ces deux sociolectes sont utilisés en ligne, dans les conditions de production inédites, autant dans le sens de l'identité personnelle, que dans la liberté relative offerte par le cadre, ou à la possibilité de retourner au « forgeur inconnu » (Lecercle, 2004). Si *verba volens, scripta manens*, internet permet maintenant la documentation et la première attestation précise d'une pratique–si tant est qu'elle est née en ligne. Les spécificités des communautés virtuelles vont être examinées dans le chapitre suivant. Il est indispensable de comprendre l'écologie virtuelle et les contextes de production pour comprendre la façon dont les communaulectes fonctionnent et s'inscrivent dans les pratiques langagières en ligne.

Chapitre 3

Essai de caractérisation des pratiques discursives sur les forums internet

On the internet, nobody knows you're a dog.

P. Steiner - *The New Yorker*

Le chapitre aborde certaines caractéristiques discursives, en particulier sur les forums de niche qui sont le sujet de cette thèse. Je dresse également les modalités de l'écologie numérique, qui dresse également les conditions d'utilisation des graphies étudiées. Il faut inscrire les deux dialectes de l'anglais dans le contexte qui nous intéresse, c'est-à-dire la communication sur internet. Nous verrons d'abord ce qui fait d'internet une plateforme d'échange tout à fait intégrée dans la vie quotidienne, et les mouvements de création d'une communauté virtuelle. Les spécificités des sites (BP et *Scotster*), ainsi que des autres plateformes étudiées seront l'objet de la deuxième partie de ce chapitre. La constitution du corpus, envisagée sous tous ses aspects (juridique, éthique, technologique) conduira à l'étude plus fine du problème de l'identité en ligne.

Comme on le sait, l'accès à internet a connu une expansion incroyable durant les 15 dernières années, partout dans le monde. Les continents européens et nord-américains comptent le plus d'uti-

lisateurs par rapport au nombre d'habitants, mais l'Afrique et le Moyen-Orient ont connu des taux d'augmentation de la population numériquée record (3 500% et 2 600% d'expansion, respectivement, entre 2000 et 2012). Cette plateforme est connue universellement, et les communications en ligne sont devenues un objet d'étude scientifique depuis de nombreuses années. La recherche informatique, et le développement de transfert de données a été le premier sujet d'étude dès les années 70, et des recherches en Technologies de l'Information et de la Communication (TIC) et en Nouvelles Technologies de l'Information et de la Communication (NTIC) ont également fleuri. L'exploration menée sur des corpus numériques ou internet est apparue dans le paysage dans les années 90 seulement, avec la naissance de la Communication Médiée par Ordinateur (CMO), grâce à des études menées par S. Herring (Herring, 1996).

En juin 2012, *Internet World Stats* [1] a recensé 2.7 milliards d'utilisateurs d'internet ; les Américains (États-Unis) représentent 245 millions d'utilisateurs, soit 78% du pays, et 53 millions de Britanniques, soit 84% du pays, ont accès à internet. Ces statistiques officielles provenant du recensement des pays et des organismes de veille d'internet montrent la quasi-omniprésence du média dans la vie de la majorité des habitants des deux pays étudiés.

Type de contexte	Exemple
Mental	— État d'esprit du participant avant d'entrer sur la plateforme — État de santé ou de fatigue — Temps accordé à la visite de la plateforme
Physique	— Qualité de la vue — Lieu de consultation des plateformes (école, travail, domicile…)
Ergonomique	— Qualité de l'installation générale — Clavier : disposition des touches et qualité de l'outil
Numérique	— Disposition générale de la plateforme : éléments permettant l'accès à la discussion ou à l'édition des messages — Pression ou influence sociale positive ou négative (présence d'autres membres qui changeraient l'état d'esprit du scripteur sur un fil de conversation)

Tab. 3.1 – Éléments d'environnement influant sur le discours en ligne

1. Disponible à l'adresse <http ://www.internetworldstats.com/stats.htm> (dernière consultation le 12/12/12).

Le contexte écologique au sens large est très important lors de l'analyse de discours sur internet. Les productions des plateformes et réseaux sociaux sont régis par le contexte, plus ou moins large ou personnel. Ces catégories ne sont pas immuables, mais fluides, et des éléments peuvent influer sur plusieurs types de contextes à la fois. Certains de ces contextes ne sont pas observables en ne lisant que les messages, mais sont récupérables en utilisant des sites de métadonnées comme Alexa (« Alexa - The Web Information Company », 2000). Le problème peut être la fiabilité des données, qui sont indiquées en rapport avec la population générale d'internet.

Il arrive qu'un utilisateur fasse allusion à ce qui lui arrive dans « la vraie vie ». C'est une traduction de « In Real Life », ou *IRL*. Ce point de vue est dualiste et oppose un monde virtuel et un monde réel. À cette traduction, les termes « hors-forum » ou « hors-ligne » seront préférés. Il ne faut pas perdre de vue le fait que la communication sur internet au sens large (la CMO) implique une communication grâce à n'importe quel outil qui permet effectivement d'aller sur internet : smartphone, tablette numérique par exemple, sont autant de façons différentes de communiquer en ligne. En effet, si un clavier d'ordinateur n'a pas de correction automatique, attaché à toutes les activités de frappe (sur internet, sur un traitement de texte, dans une barre de recherche…), les claviers tactiles en sont pourvus. Si un utilisateur a une touche plus faible que les autres (barre espace, par exemple), la frappe demandera plus de force, et le scripteur corrigera ou non ses habitudes et ses écrits, en fonction des autres paramètres écologiques.

L'illustration 3.1 montre les différences de production entre un clavier de mobile et d'ordinateur. D'abord, la suggestion d'un mot et le soulignage rouge, qui montre la faute de frappe n'existe pas sur un ordinateur, sous cette forme. La convention sémiologique de la vaguelette de soulignement rouge vient du standard des corrections dans des logiciels de traitement de texte (type *Word* (Microsoft) ou *Writer* (Open Office), et même *Texmaker* (Free Software Foundation), éditeur de texte en LaTeX), elle même certainement influencée par les corrections manuscrites des professeurs sur les devoirs de leurs élèves. La suggestion d'un mot de remplacement dans une bulle en haut (sur Android et Windows Phone, les suggestions lexicales sont entre le clavier virtuel et la zone de texte) est inédite aux claviers virtuels de téléphones et de tablettes numériques.

Pour avoir accès aux caractères accentués, il faut presser plus longuement sur la touche en question.

Illustration 3.1 – Trois états du clavier de téléphone (iPhone 4S)

Il n'est pas possible de rajouter un diacritique au dessus d'une lettre qui n'en porte pas (comme un circonflexe au dessus d'un <r>, par exemple), ce qui limite les choix graphiques. La touche <123>, entourée en rouge, permet d'accéder à une première série de caractères extra-alphabétiques, et elle se transforme en <# +=> pour accéder au troisième état du clavier, qui contient les caractères non-alphabétiques les moins utilisés, comme les accolades et les symboles monétaires. Y subsistent les signes de ponctuation les plus courants. Ces claviers alternatifs se complètent par la touche <ABC>, pour rejoindre le clavier alphabétique (touche entourée en vert), et les trois états du clavier ont un signe supplémentaire, le pictogramme de la terre, pour sélectionner une autre disposition du clavier numérique, lorsque l'on écrit dans une autre langue par exemple. Le clavier présenté est un clavier AZERTY français, mais changer la langue d'écriture (et donc de dictionnaire) changerait la disposition

des touches virtuelles.

Ces trois états montrent l'optimisation de l'organisation de l'écrit dans des conditions de production plus ardues (avec moins d'espace pour placer son doigt sur une touche non matérielle). Ce qui facilite la production en ligne rend l'utilisation du leetspeak plus difficile [2], et ne facilite pas non plus l'utilisation de smileys créés au clavier, comme les smileys occidentaux < :-)> ou asiatiques <^__^>, à l'usage pourtant répandu en ligne, pour leur préférer des smileys iconographiques.

Pourtant, la tendance des réseaux sociaux est de figurer en tant qu'applications propriétaires, qui sont simplement des interfaces vers le site web, comme l'application BP [3]. Dans l'image 3.2, on peut voir des captures d'écran de BP sur mobile. Il est probable, bien que non-vérifié, que la navigation sur un téléphone portable change les conditions de production écrites : les corrections automatiques ou la taille même du clavier ne créent pas les mêmes conditions de production.

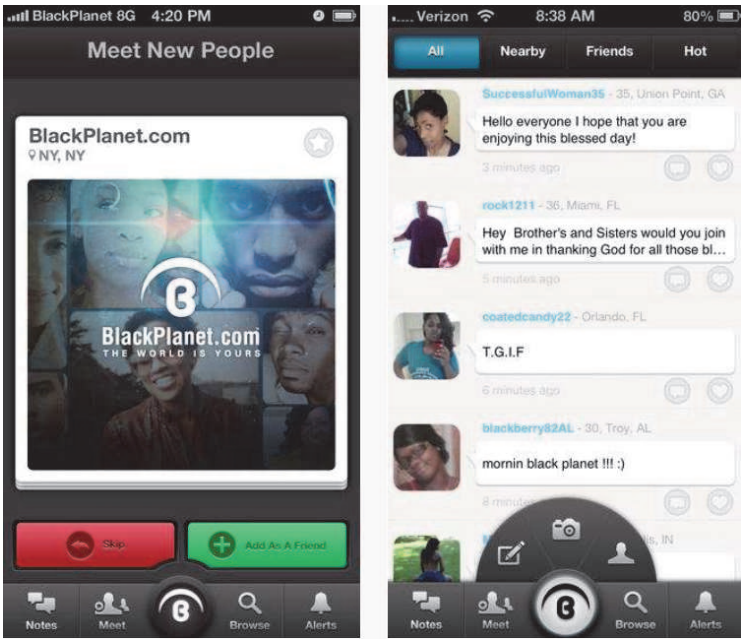


Illustration 3.2 – Captures d'écran de l'application BP (source : iTunes Store)

2. Le leetspeak, brièvement évoqué en 4.4.2 et sous-ensemble du netspeak, consiste principalement à remplacer les lettres d'un mot par des ambigrammes, ou caractères reconnaissables grâce à notre capacité à lire les graphèmes imparfaitement tracés. Par exemple, <4> pour symboliser <a>.

3. Disponible à l'adresse <https ://itunes.apple.com/us/app/blackplanet-meet-new-people/id583411580 ?mt=8>. Dernier accès : 20 mars 2014.

Les techniques de corps, pour reprendre la terminologie de Mauss (Mauss, 1936) doivent être remises à jour pour intégrer de nouveaux paradigmes humains et intégrer la théorie des communautés de pratique (Wenger, 1998) : les changements apportés par la CMO et les TIC sont tels que leur ampleur n'est pas perçue, même par leurs utilisateurs. Je prends pour exemple une informante de 65 ans, qui s'est plaint de l'absence d'informations sur les programmes de télévision sur l'écran, à cause d'un problème sur son réseau *Asymmetric Digital Subscriber Line* (ADSL), tout en proclamant qu'elle ne voit pas l'intérêt d'internet, d'un ordinateur…Les métadonnées mises à disposition dans tous les médias qui se nourrissent les uns des autres (comme les *hashtags* de *Twitter*, inscrits en temps réel à la télévision, dans certains programmes, les dossiers et blogs des journaux disponibles en ligne—l'exemple le plus frappant étant *Newsweek* qui a annoncé en 2012 le passage à une publication entièrement numérique) rendent l'impact d'internet tellement diffus dans la vie quotidienne qu'à tous les niveaux, le monde occidental est entré dans un postdualisme total : il n'y a plus de clivage fort entre une vie « virtuelle » et la « vraie vie » (*In Real Life* (IRL)), mais plutôt un brassage des deux, qui se complètent entièrement : si quelqu'un s'enregistre à un lieu en utilisant FourSquare, c'est qu'il se trouve physiquement à cet endroit. Á l'inverse, le *tweet* suivant (Chouine, 2013) est apprécié (78 partages, 27 favoris) parce qu'il joue sur les codes de la communication virtuelle (le « mode avion » est l'option des téléphones et ordinateurs portables qui consiste à couper tout réseau 3G, wifi et cellulaire) en les transposant dans sa vie hors-ligne (en l'occurrence, sexuelle).

Il est très symptomatique de la position adoptée dans cette thèse. Les communautés virtuelles sont

Illustration 3.3 – Tweet du 9 février 2013

des avatars des communautés humaines, où les conditions ne sont considérées comme inédites (voir 3.1) que parce qu'elles sont présentées sur un nouveau média ; nous pensons que les apparitions de

phénomènes langagiers auraient pu se développer de façon identique hors-ligne. En revanche, l'inédit se situe au niveau de la mémorialité d'internet, qui enregistre et laisse une trace de tous les échanges sur des plateformes asynchrones (tant que les utilisateurs ne les effacent pas d'eux-mêmes). Ici, la grammaticalisation des changements (Hopper & Traugott, 1993) est documentée, et la première occurrence d'une pratique graphique ou lexicale alternative est archivée, contrairement à la nature volage de la praxis orale.

3.1 Approche sémiologique de la matérialité des échanges dans les communautés virtuelles

Internet est une plateforme à l'utilisation grand-public relativement récente, datant des années 90 dans son utilisation commerciale et ouverte au public [4]. Chaque année, le nombre d'utilisateurs a augmenté de 20 à 50% et le trafic sur internet d'environ 100% vers la fin de la décennie 1990 [5], ce qui montre l'engouement immédiat du grand public vers cette plateforme, initialement conçue pour une utilisation militaire ou universitaire.

Les communautés construites sur internet présentent une forme inédite, puisqu'elles sont agrégées non autour d'un lieu géographique, mais autour d'intérêts en commun. Ceci constitue la définition d'une communauté de pratique, telle qu'elle a été envisagée par Wenger (Wenger, 1998).

Une communauté sur internet est une entité polymorphe, mais les caractéristiques générales d'accès et de communication restent peu ou prou identiques. Les modes de communication sont divers (synchrones ou asynchrones), et l'accès à la communication est plus ou moins ardu, plus ou moins privé. Certaines plateformes permettent le choix de l'interlocuteur (ou des interlocuteurs), mais l'accès à des salons de chat publics, par exemple, n'est soumis à aucune contrainte. Les utilisateurs d'internet ont donc une palette d'outils à leur disposition pour échanger de la façon la plus idoine.

4. Les premières communications entre deux ordinateurs datent de 1961, mais elles étaient limitées à des ordinateurs militaires, connectés au même réseau.

5. Aux États-Unis seulement (Coffman & Odlyzko, 1998).

Plateforme	Temporalité		Facilité d'accès		Nombre d'interlocuteurs		
	Synchrone	Asynchrone	Logiciel	Inscription	Unique	Multiple	Début d'activité
Courriels		⊗		⊗	⊗	⊗	1972
Listes de diffusion		⊗		⊗		⊗	1986
IRC		⊗		⊗		⊗	1988
Sites personnels		⊗		⊗		⊗	1990
Forums, message boards		⊗		⊗		⊗	1994
Blogs		⊗		⊗		⊗	1994-1999
Chat (ICQ, MSN…)	⊗		⊗		⊗		
MySpace		⊗		⊗		⊗	2003
Facebook		⊗		⊗		⊗	2004
Twitter		⊗		⊗		⊗	2006

Tab. 3.2 – Conditions de communication par type de plateforme virtuelle

Les données du tableau 3.2 ne concernent que les plateformes d'échange principalement écrit [6], et valent pour les réglages par défaut. Les résultats doivent donc être raffinées : les échanges sur *Twitter*, *Facebook* peuvent être privées, mais la valeur par défaut est « multiple ». Á l'inverse, il était possible de commencer une discussion avec de multiples interlocuteurs sur *Windows Live Messenger* (*Microsoft (Messenger Service) Network* (MSN)). *Twitter* et *Facebook* n'ont pas besoin de logiciel propriétaire pour être utilisés, mais ces logiciels existent afin de pouvoir être connecté plus facilement et rapidement que sur un navigateur, par exemple si on utilise un smartphone ou une tablette pour se connecter à internet.

Les courriels sont en fait la même chose que les listes de diffusion, qui utilisent ce biais, mais la différence principale est le fait qu'un courriel est envoyé à un seul individu, dont on connait déjà l'adresse électronique, alors que la liste de diffusion est centralisée par un propriétaire ou opérateur, qu'il faut s'y inscrire et que les messages envoyés sont ensuite transmis à tous les abonnés de la liste. C'est la raison pour laquelle les échanges sur les listes de diffusion et sur Usenet en général ont été les premiers corpus explorés et étudiés en linguistique : ces échanges étant répandus, et les corpus, faciles à récupérer (Anis, 1991, 1999).

La distinction entre les sites personnels et les blogs est presque entièrement stylistique. Le site per-

6. J'en exclus donc *Snapchat* ou *Instagram*, les *image boards* en général…

sonnel n'est pas construit de la même façon, et peut être entièrement autonome, c'est-à-dire sans retour de lecteur si le concepteur le désire. Le blog (contraction de *web log*, journal sur internet) est découpé, grâce à l'utilisation de logiciels et serveurs de blogs, comme *Blogspot* ou *Canalblog* en messages différents, émanant du ou des auteurs du blog, contenant une image et/ou du texte. Les internautes peuvent s'abonner à la liste de syndication du blog et choisir de répondre ou de commenter sur les billets publiés.

La communication décrite dans le tableau ci-dessus n'est pas limitée aux échanges écrits. Certaines plateformes permettent le partage d'image, de liens internet, de fichiers. C'est le point de pivot entre le web 1.0 et le web 2.0, comme l'explique D. DiNucci, qui a créé l'expression : « On the front end, the Web will fragment into countless permutations with different looks, behaviors, uses […] The Web will be understood not as screenfuls of texts and graphics but as a transport mechamism, the ether through which interactivity happens. » (DiNucci, 1999). DiNucci évoque surtout la multiplication des façons dont le web intégrera nos vies, hors-ligne également (dans les ordinateurs de bord et GPS des voitures, dans nos télévisions). Son expression, qui désigne le web 2.0 comme une prolongation d'internet dans les vies quotidiennes, ne prend pas entièrement en compte la façon dont la vie hors-ligne s'empare également du médium : partager des images de ses proches, de soi, comme les *selfies* et autres photographies de profil, comme cet utilisateur de SC qui documente son profil avec pas moins de 264 photographies, de sa barbe taillée pour les vacances (photographie 1), du contour de son tatouage (photographie 3), ou de l'achat d'une *claymore* pour décorer sa chambre (photographie 6).

La multitude des éléments partageables sur internet, dans les communautés en ligne, est propice à la création de liens interpersonnels forts entre les membres. Si ce membre diffuse autant de photographies, c'est parce qu'il reçoit des commentaires, qui encouragent et valorisent ses efforts de partage. C'est un cercle vertueux de la communication en ligne.

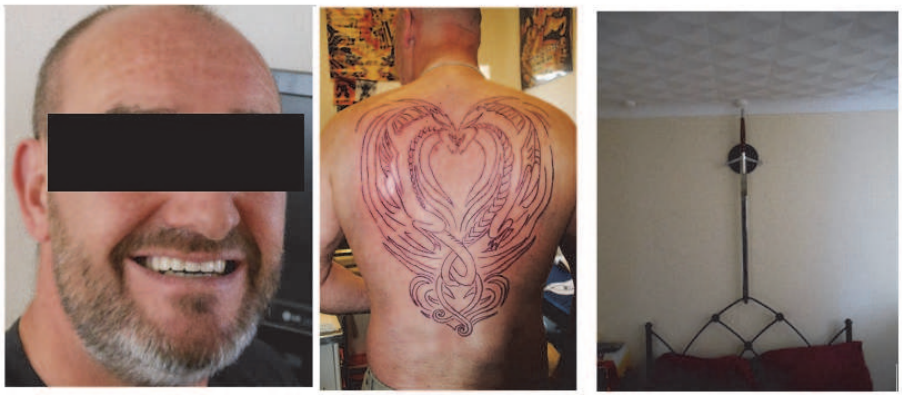


Illustration 3.4 – Photographies du profil d'un membre de SC (SCPH4)

3.1.1 Topographie des sujets

La porosité entre l'en- et l'hors-ligne est palpable, même dans les contextes qui ne semblaient pas attirer ce genre de phénomène : sur SC, une utilisatrice, Liz, a posté une photographie de son fils décédé [7]. L'illustration 3.5 présente le Corpus En Contexte (CEC), avec la photographie et les deux premiers commentaires. Le tableau 3.3 groupe les cinq premiers commentaires à cette photographie, qui est un photomontage. Le premier message cache des condoléances implicites ('*C'est un bel hommage*'), à laquelle la mère répond (message 2).

Cette mort a eu lieu bien avant l'inscription de la mère sur SC, en 1996, mais elle profite de son espace personnel sur SC pour en parler. Les événements hors-ligne ont donc leur importance sur SC. Même dans ce contexte, le sujet des commentaires change dans les deux derniers messages cités. Les messages 3 et 4 déplacent le point de vue du sujet vers la façon dont Liz a fait ce photomontage. Elle donne non seulement le nom du programme qu'elle utilise, mais elle commente aussi la difficulté de le remplacer par un autre programme. La discussion sort donc complètement du fond, pour parler de la forme de l'hommage rendu au fils : l'environnement numérique dans lequel elle évolue est presque plus important que la fin du message 3 (qui lui présente des condoléances calquées sur celles du message 1), et elle ne remercie pas son interlocutrice pour son compliment. Les commentaires afférents sont présentés dans le tableau 3.3.

7. Photographie accessible à l'adresse <http ://www.scotster.com/photos/Mathew.41017.html>. Dernier accès : 12 juin 2013.

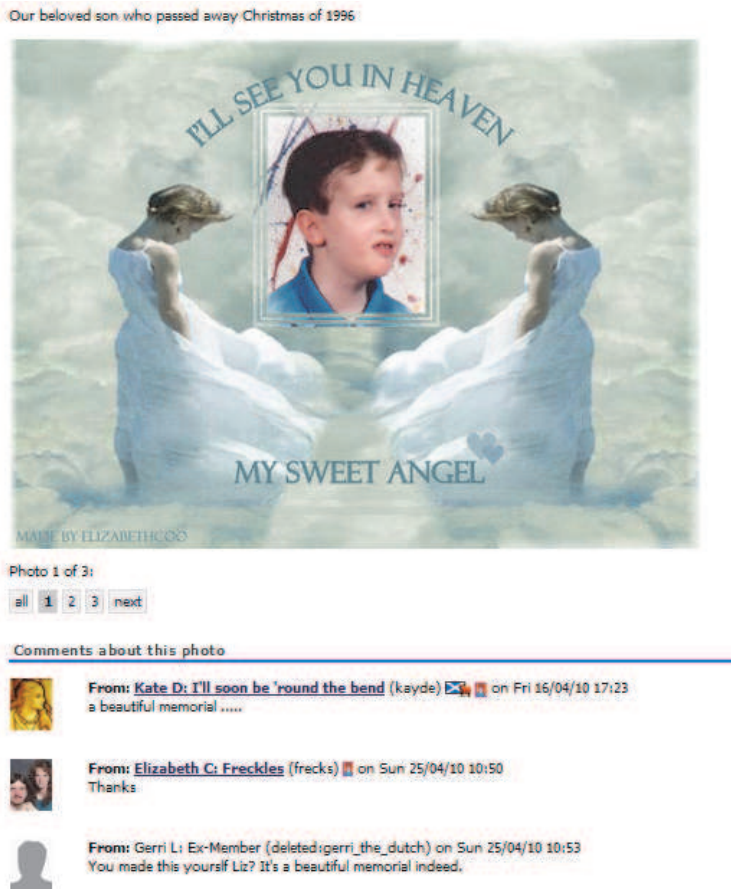


Illustration 3.5 – Capture d'écran du début de SCPH3, photographie et commentaires

Message	Corps
1	a beautiful memorial …
2	Thanks
3	You made this yourslf Liz ? It's a beautiful memorial indeed.
4	Yes with a program I had on the computer …the program wont work with newer computers .So I am searching for a new program…It was microsoft picture it, simple and easy to use.

Tab. 3.3 – Commentaires 1 - 4 de la photographie 3.5 (SCPH3)

Les utilisateurs peuvent parfois exprimer leur agressivité ou leur énervement vis-à-vis de la situation hors-forum comme le message 612 de *Random Thoughts*, sur BP. Ce message est très expressif,

avec des utilisations de majuscules, « has NO IDEA »), d'interjections de dépit, « UGH », « DANG » et de ponctuation trop abondante (15 points d'exclamation pour la plus longue séquence, trois pour la plus courte). Ce message est intégralement tourné vers des événements hors-forum, n'est pas en relation avec la conversation en cours. Il est placé ici à cause du caractère aléatoire du fil de discussion, qui encourage les conversations suivies autant que les messages ponctuels. Ce message n'est pas non plus évoqué ou même commenté par la suite du fil de discussion. Il ne trouve sa place sur le forum et le fil de discussion que parce que le forum est l'un des lieux où l'utilisateur estime qu'il peut exprimer sa diatribe à l'encontre de son travail. En revanche, il ne nécessite pas de réponse, et aucun des membres proches de l'utilisateur 2 n'y réagit ici.

Le forum n'est pas qu'un lieu de partage de la culture en ligne, ou d'agrégation de gens ayant

Message	Corps
612	Yo, my team at work has NO IDEA how difficult it is to schedule a videoconference between NY and LA at the last moment !!!! UGH !!!!!!!!!!!!!! […]. These folks are technologically incompetent, so I'm basically the team's IT support, fixing ALL conference room and software issues. […] DANG I deserve a raise !!! […] OK. Rage-filled rant over.

Tab. 3.4 – Message 612 de Random Thoughts 2011

les mêmes intérêts. C'est aussi un lieu où les événements hors-ligne sont partagés, dans le sens fort : ils sont relatés, et commentés par les autres membres. Le hors-ligne évoqué peut rester lettre morte (comme l'extrait où le membre se plaint de son travail), ou même dériver vers d'autres sujets (comme la photographie de l'enfant décédé). La forme conversationnelle ou technoconversationnelle [8] implique toutes les couches des interactions, et il n'est pas exclu que, même si un message n'a pas de réponse dans le fil de discussion, un autre membre y fasse allusion dans un autre espace (par message privé, dans un autre fil…).

Dans l'extrait présenté en contexte dans l'illustration 3.6, l'utilisatrice Elizabeth a écrit un statut sur son profil pour évoquer l'anniversaire de sa première inscription à SC. Félicitée par Erin (premier message présenté), Elizabeth en profite pour commenter une des photographies qu'Erin utilisait sur SC la dernière fois qu'elle s'y était connectée (les photographies des membres en ligne au moment

8. Pour une définition en profondeur des travaux sur la conversationnalité en ligne, se référer à Paveau, 2013a.

de la connection se trouvent dans le cadre rouge, en haut à droite). La photographie la montrait visant avec une arme à feu [9] Elizabeth profite donc de la présence d'Erin, matérialisée par son message, pour glisser deux commentaires, avant et après avoir été vérifier son profil « Erin, just checked out your photos », dans le message 3. Il y a six minutes de différence entre le moment où le message 2 et 3 ont été postés. Le déroulement chronologique et l'archivage de ces messages, échangés en 2010, sont parfaits. Cet échange cristallise le faisceau de mondes croisés sur les communautés virtuelles. Il est à la fois relié à l'hors-ligne, avec les photographies d'une arme à feu, qui est l'une des passions d'Elizabeth (« I'm a wee bit of a gun enthusiast. Well, a user, anyway. :-D Just curious. »), à divers moments de connexion à SC, parce qu'Elizabeth n'aurait pas vu la photographie d'Erin si elle n'avait pas été connectée en même temps qu'elle, et ne l'aurait donc pas commenté à sa connexion ultérieure. L'archivage des moments de l'échange est fait à la minute près, et la page restera accessible à son adresse tant que le site existera.

Ces incursions dans le monde hors-ligne expliquent pourquoi les utilisateurs empruntent librement

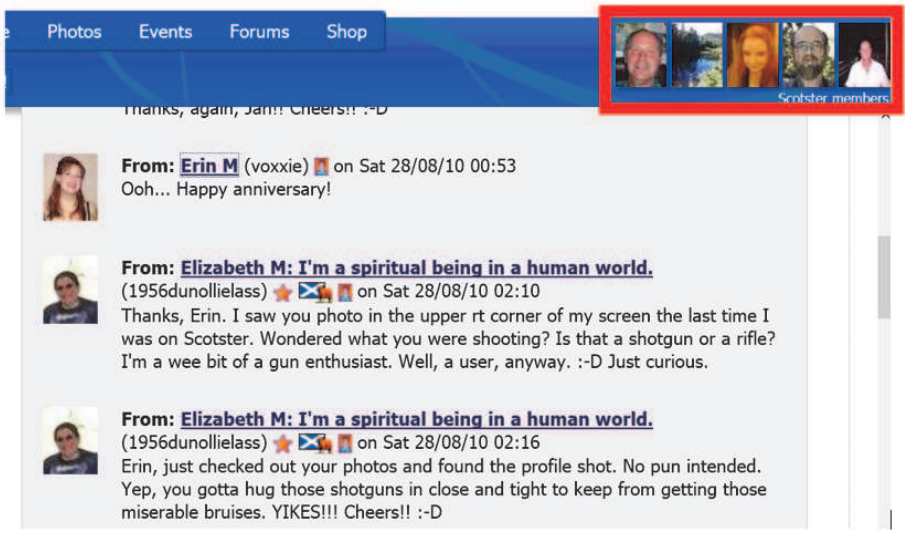


Illustration 3.6 – Extrait de CEC (SCP10)

les codes de l'hors-ligne, y compris les codes linguistiques des langues minoritaires qu'ils parlent. Les chapitres 4 et 5 s'intéressent plus particulièrement à ces échanges linguistiques communaulectaux, qui sont au moins en partie inspirés par l'hors-ligne. Les deux espaces cohabitent sur le même continuum.

9. Cette photographie ne s'y trouve plus au 20 juillet 2014.

En outre, le lieu virtuel reste accessible même lorsque les conversations sont achevées (le dernier message de SCP10 a été posté le 30 août 2010), ce qui le place hors de la péremption d'autres lieux virtuels, comme Snapchat [10].

3.1.2 Positionnement communicationnel

Parler en ligne, c'est utiliser des codes communs. On a le choix entre deux types majeurs de communication : ceux tournés vers l'humain même (comme les réseaux sociaux à la Facebook (FB)), où on communique de sa propre page personnelle à celle des autres, ou parler ensemble dans une zone commune, sur un sujet qui intéresse les membres de la communauté. Ce type de communication est plus proche de la communauté de pratique décrite par Wenger : « A community of practice is not merely a club of friends or a network of connections between people. It has an identity defined by a shared domain of interest. » (Wenger-Trayner, 2012). Cette définition concerne uniquement le positionnement communicationnel des membres, qui est le sujet de la partie suivante, mais pas la façon dont l'ergonomie des sites façonne les discussions qui peuvent s'y produire.

Chaque site étudié a sa propre identité, qui passe par des codes sémiotiques différents, imposés par les créateurs du site. Par exemple, en observant la page d'accueil de SC (illustration 3.7), on voit tout de suite les codes de couleurs utilisés, et l'organisation de la communication sur la page : SC utilise une charte graphique à dominante bleue et blanche, comme les couleurs du drapeau écossais. D'ailleurs, le drapeau est présent au dessus du nom du site (en haut à gauche), dans la case de recherche, sur la photographie dans le bandeau à gauche, et même après le nom d'un des utilisateurs, Don (cette icône signifie qu'il a payé pour soutenir SC pendant un an). L'imagerie écossaise typique est omniprésente, et cette icône s'accompagne d'une petite vache Highland qui est aussi un stéréotype.

La communication est organisée autour de la présence des membres : sous la bannière « Scotster is the photo sharing and social site for Scotland », qui rappelle la mission du site, les états de connexion des membres déroulent en temps réel, avec un lien vers leur profil. La communication du site est tour-

10. Snapchat est un programme qui permet de prendre des photographies et de les partager. La popularité de ce réseau social réside dans le fait que les photographies ne sont pas conservables une fois échangées–à moins de prendre une capture de son écran.

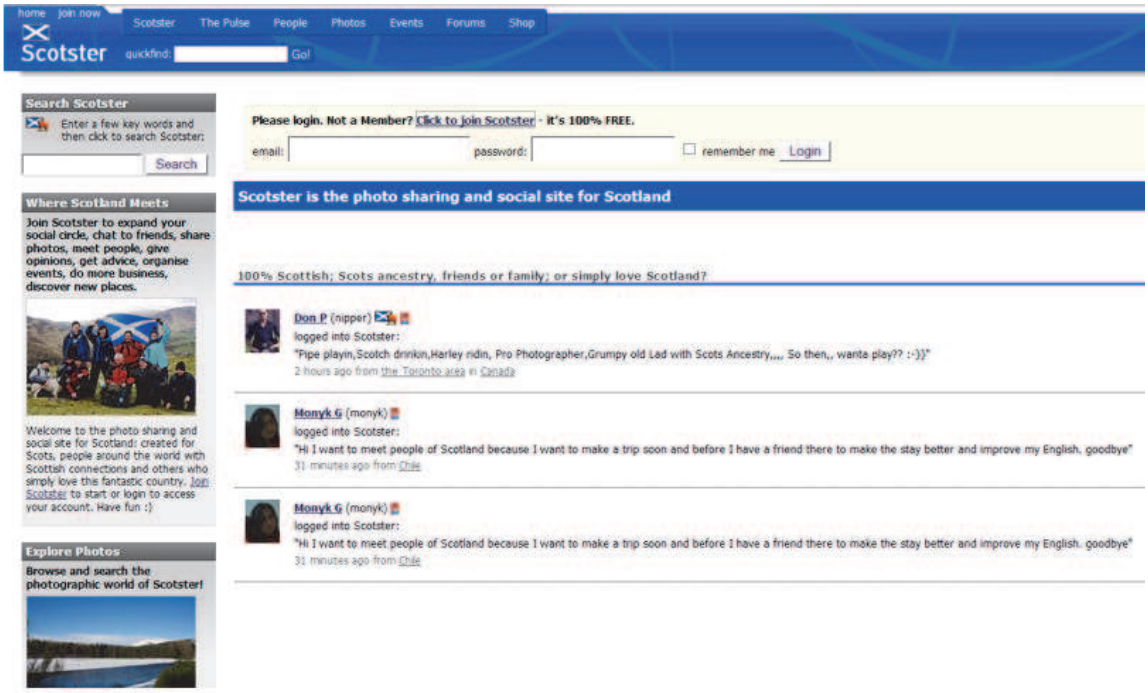


Illustration 3.7 – Page d'accueil de SC au 29 juillet 2014

née vers les activités qu'ils ont sur le site, pour donner envie de participer à son tour. C'est également la démarche des créateurs du site, qu'ils décrivent dans le manifeste de SC, que chaque membre reçoit à l'inscription [11]. Ce texte contient notamment le paragraphe de mission suivant :

> Our approach leads the way in social networking, successfully combining the web with the real world - whereas other sites may be about connecting you to your friends and family, we do that and more : we help you to expand your social circle and to meet new people on your wavelength. This is the new way to get to know people from the internet and you're far more likely to develop long-term friendships and great relationships.

Le positionnement est tourné vers l'inconnu, et la création de nouvelles amitiés plutôt que la préservation de cercles sociaux hors-ligne. Cela est renforcé plus loin : « Remember that our major difference is our approach to introducing you to other people - so move on from traditional sites and get up and do something different ! ». La définition d'un site communautaire standard, pour SC, n'est pas de découvrir de nouvelles personnes. Pourtant, lorsque un membre se connecte, la page d'accueil change. L'illustration 3.7 montre la page d'accueil pour les non-membres de SC ; mais la cartographie de cet espace numérique évolue pour les membres. Le second état de la page d'accueil est présenté dans

11. Le texte de ce manifeste est disponible dans les annexes, c'est le document A.6.1.1.

l'illustration 3.8. Tout le contenu de la page principale, à part le bandeau d'information à gauche,

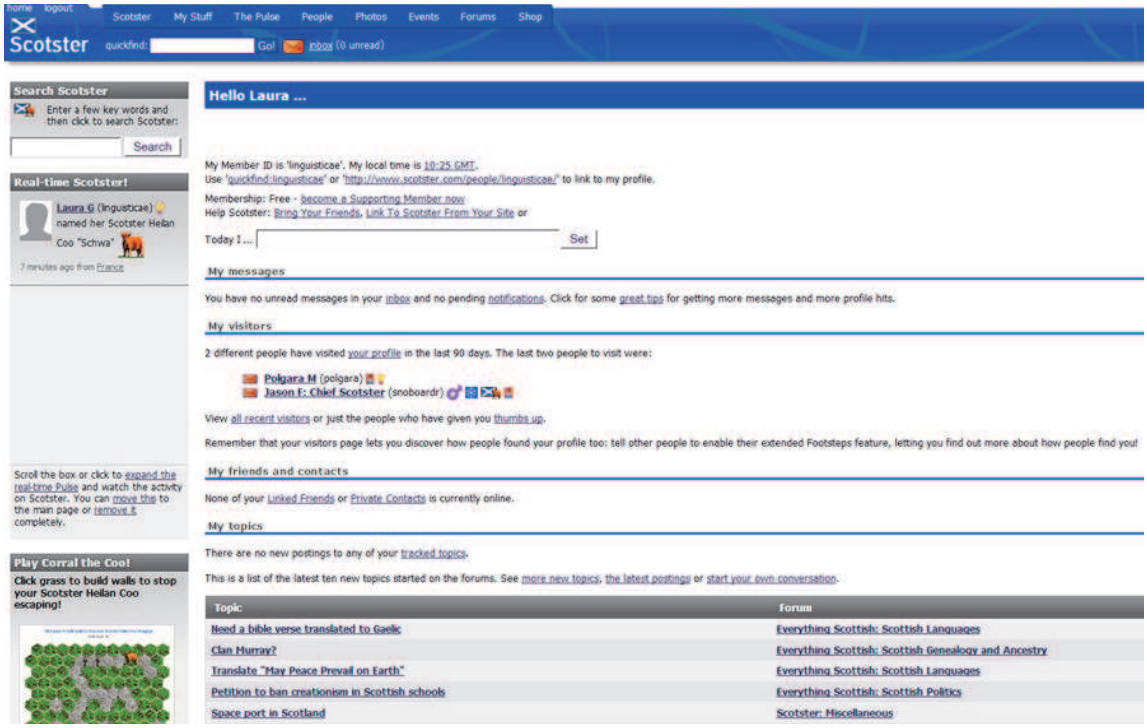


Illustration 3.8 – Page d'accueil de SC pour les inscrits, au 29 juillet 2014

devient une vitrine de l'espace personnel du membre inscrit. Sous la bannière « Hello [prénom du membre] », le site invite à mettre un statut (rubrique « Today I… »). Ensuite, les rubriques « my messages », « My visitors » ; « My friends and contacts » et « My topics » dessinent un tableau centré sur l'utilisateur et ses intérêts, et non sur la découverte de l'autre. En haut à droite, le macaron « Scotster is alive : click to feel the Pulse » donne le droit d'aller voir quels sont les derniers statuts postés par les autres utilisateurs.

La possibilité de suivre des fils de discussion dans la rubrique « My topics » invite à partager avec des inconnus ; et les deux rubriques sur les visiteurs et amis sont en fait centrés sur soi, avec les gens intéressés par le profil d'une part, et les autres membres connus de l'autre. Les visiteurs peuvent montrer leur appréciation (« View all recent visitors or just the people who have given you thumbs up »), et la possibilité de les trier sur ce critère semble unique à SC.

Sur BP, enregistrés ou non, les visiteurs du site ont accès aux statuts des autres membres (section « What's new ? »), mis en évidence au centre de l'écran (illustration 3.9) et, encore une fois, l'accent

est mis sur le compte du visiteur, avec une géolocalisation des membres connectés qui se trouveraient à proximité. La page d'accueil enjoint d'emblée à poster un statut, et donc à partager spontanément un statut.

Le forum et les parties où les utilisateurs peuvent communiquer ensemble sur un espace partagé sont signalés en haut de l'écran, dans les différents onglets (dont une description rigoureuse sera donnée dans la partie 3.2.1.1).

Illustration 3.9 – Page d'accueil de BP au 29 juillet 2014

Ces deux architectures sont centrées sur l'utilisateur et les outils pour partager à partir de son écran d'accueil. Ces outils sont plus accessibles que ceux pour partager dans la communauté même, sur les forums ou les salons de chat. Ceci s'explique par l'influence de médias comme Twitter et FB, qui

mettent plus volontiers l'accent sur l'individu que sur la communauté dans laquelle l'on partage. C'est un changement de pratique internet, qui a évolué de Usenet et des *billboards* de messageries, aux forums et blogs, vers le milieu des années 1990, pour ensuite favoriser les réseaux sociaux (J. Lee, 2012). FB et son ouverture au public en 2006 a changé la donne, même pour des sites plus anciens, comme BP, qui ont dû évoluer et le calquer pour rester attractifs.

L'émergence du soi construit en ligne [12] change donc le positionnement architectural des plateformes numériques, mais ce changement était prévisible. Sur les forums, les sections les plus détendues, où les discussions sont les plus générales, sont souvent celles qui ont le plus de réponses ou de fils de discussion : sur SC, la sous-partie « Scottish Humour » (SC2Y) est celle qui a le plus de fils de discussion, avec 289 fils (qui contiennent entre 4 et 50 messages chacun), suivi par la sous-section « Scottish Languages » (SC2N, 266 fils) et « Miscellaneous » de la section « Scotster » (SC1K) ; sur BP, ce sont les sous-sections « General » des sections « Health & Beauty » et « Relationships » et qui centralisent les plus grands nombres de fils de discussion, avec respectivement 384 854 et 48 925 fils de discussion respectivement [13]. Les utilisateurs préfèrent parler du plus général, ou partager du contenu tout fait (dans le cas de la sous-section sur l'humour écossais, qui est une anthologie de blagues souvent trouvées sur internet), ou d'eux-mêmes, comme dans la section « Relationships » de BP.

3.1.3 Caractérisation sémiologique des échanges communicationnels

La qualité des interventions sur les forums et les fils de discussion peut être analysée suivant les catégories définies par un groupe de chercheurs (Rourke, Anderson, Garrison & Archer, 1999) à propos des échanges dans des communautés de pratique sur internet :

— Affective : expression d'émotions, de sentiment et valeurs, d'anecdotes personnelles, et utilisation de l'humour. L'exemple de la photo hommage au fils disparu d'une intervenante de SC

12. Voir les travaux de L. Merzeau sur les traces numériques et leur ineluctabilité.
13. Ces chiffres ont été relevés au 29 juillet 2014, et peuvent changer de jour en jour avec l'adjonction et la suppression de fils de discussion. C'est pourquoi ils ne servent que de valeurs indicative, peut-être contredites par des relevés antérieurs ou postérieurs à cette thèse.

(illustration 3.5) vu précédemment donne un aperçu correct du partage d'émotions autour d'une anecdote personnelle. L'utilisation de l'humour est visible dans le morceau de CEC suivant, où les utilisatrices commentent les cheveux du bébé de Beyoncé, Blue Ivy. Les deux messages cités contiennent l'acronyme « LOL », caractéristique du netspeak (étudié dans le chapitre suivant, dans la section 4.4.2 ; voir aussi 4.4.3.1 pour une application de « LOL » en tant que nomenclature désignant un phénomène typique d'internet).

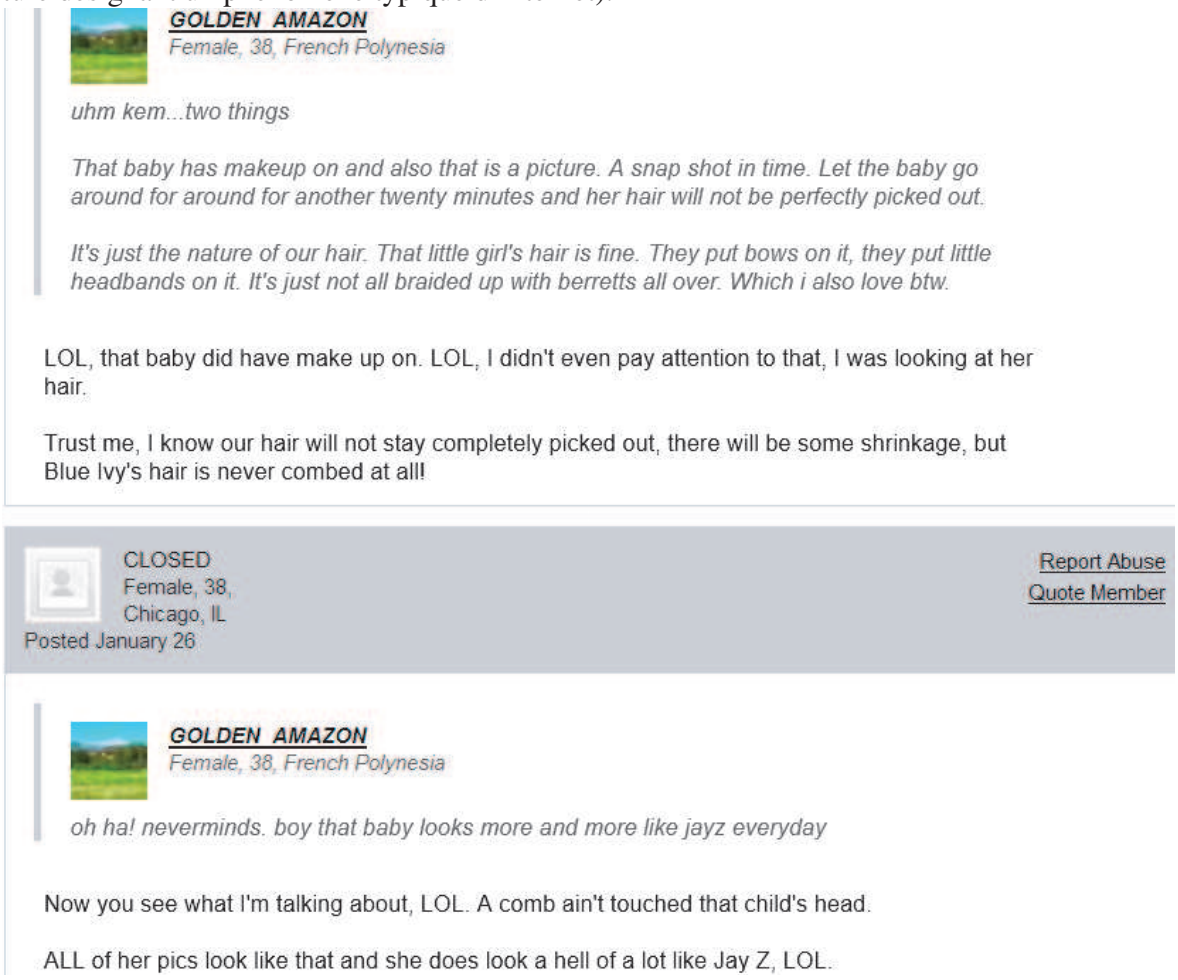


Illustration 3.10 – Extrait du fil « Random Thoughts MMXIV » (BP12E5582)

— Interactive : établissement et entretien de relations sociales : réponses sur les sujets de fils de discussion (par opposition à la création d'un autre sujet), citations d'autres messages, références explicites aux messages d'autres intervenants, et questions. La fonction interactive est vitale à un forum, et l'extrait de CEC illustrant la valeur affective de la communication contient une citation. Enchâssée dans le corps du premier message, le message cité n'a pas la même forme

que le message principal. La police est italicisée, la couleur du texte est en gris (une couleur moins forte que le noir de la police principale). Le retrait plus important accompagné d'une barre à gauche rappelle la mise en forme d'un courriel cité ; enfin, les données de l'auteure ouvrent la citation (pseudonyme et avatar, âge, sexe et lieu). Les citations sont faciles : il suffit de cliquer sur « Quote Member » en haut à droite de chaque message pour ouvrir une boîte de nouveau message, avec la citation insérée.

— Cohésive : établissement et entretien d'un sentiment d'appartenance au groupe. Ceci passe par des compliments et appréciation d'un message, expression d'accord avec un autre intervenant, utilisation de vocatifs, messages à l'intention de la communauté, références à la communauté en utilisant des pronoms inclusifs, et utilisation de salutations. Sur SC, une sous-section est particulièrement propice à l'observation de la fonction cohésive, la section des présentations « Welcomes and Introductions » (SC1A). L'auteur du premier message est accueilli avec des commentaires humoristiques (la fonction cohésive semble aller de paire avec la fonction affective), des injonctions à faire partie de la communauté activement « check out the forums and get your bit in ». Le nouveau membre, de Dundee, est aussi naturellement accueilli par une utilisatrice, qui vient de la même ville que lui, et qui a donc un point commun géographique avec lui.

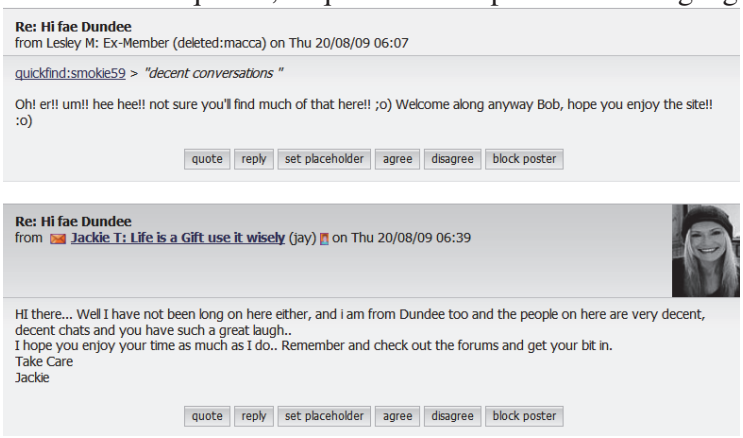


Illustration 3.11 – Extrait du fil « Hi Fae Dundee » SC1A82

Ces démonstrations d'inclusion et de création d'une identité virtuelle passe également par les démonymes virtuels, qui sont présentés dès le message automatique à l'inscription, qui commence par « Thank you for becoming a Scotster :) » (voir l'annexe A.6.1.1).

Nous pouvons comparer le modèle théorique de Rourke à celui de Jakobson (Jakobson, 1960) les fonctions du langage telles qu'elles sont décrites par Jakobson complètent celles de Rourke. Rourke ne s'attache qu'à la description de phénomènes se produisant lors de discussions en ligne (plus particulièrement, de discussions entre des étudiants et des professeurs, sur un forum de discussion éducatif), il est possible de créer des distinctions plus fines en joignant les deux schémas théoriques :

Rourke **Catégorie**	**Indicateurs lors d'une discussion virtuelle**	Jakobson **Fonction**
affective	Exprimer des émotions (dont utilisation de ponctuation, signes graphiques et typographiques…) Utiliser l'humour Donner des détails personnels et intimes	*expressive, poétique* *expressive*
interactive	Continuer un fil de discussion, citer ou faire des références directes à d'autres messages Poser une question Faire un compliment, exprimer son appréciation ; exprimer son accord avec un participant ou le contenu de son message	*référentielle* *conative ou métalinguistique* *expressive*
cohésive	Utiliser des vocatifs Utiliser des pronoms inclusifs pour désigner le groupe (‘*nous, notre, on*’…) Utiliser des salutations, des éléments phatiques	*conative* *expressive* *phatique*

Tab. 3.5 – Comparaison des théories de Rourke et Jakobson

Les deux types de communication ne se recoupent cependant pas toujours exactement, puisque Rourke se focalise sur la description des types d'échanges. Son modèle sert en fait pour l'analyse de discours observable sur internet, alors que les fonctions du langage sont bien plus générales. Dans le tableau 3.5, il n'y a presque pas de place pour la fonction poétique du langage, qui est bien plus prévalente dans les présentations personnelles que dans les discussions. Leur étude, section 3.4.4 et l'illustration afférente 3.36 sont de bons exemples de la fonction poétique du langage. Rourke s'intéresse peu à la fonction métalinguistique, puisqu'elle est en dehors de ce que la communication en ligne recouvre. En revanche, pour ce travail de thèse, c'est une question au centre des préoccupations des utilisateurs de SC, notammment.

Le modèle de facteurs inaliénables de communication décrit par Jakobson est également transposable aux forums, puisque chaque facteur communicationnel est relié à une fonction du langage.

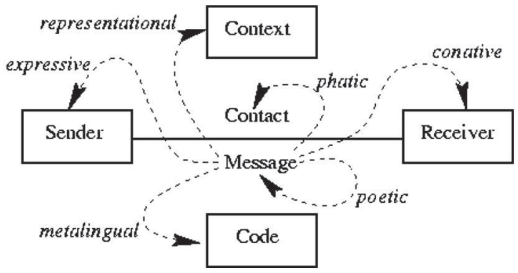


Illustration 3.12 – Facteurs communicationnels et fonctions du langage (Jakobson ; source : Gibbons, 1997)

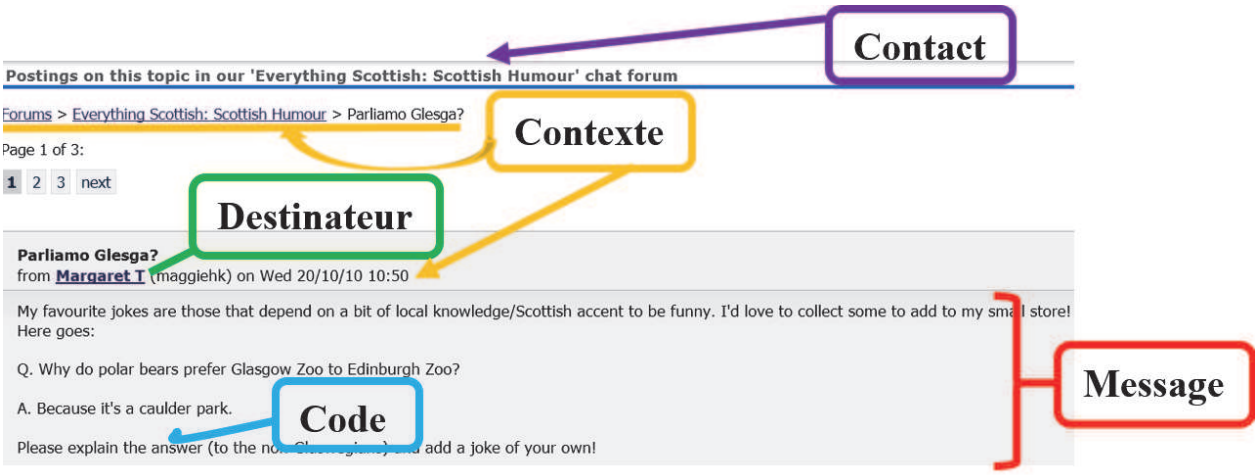


Illustration 3.13 – Schéma de la communication appliqué aux forums

Le message est l'élément le plus important. Sans message, pas besoin de communiquer sur la plateforme dans les espaces communs. Le message est aussi l'élément communiqué sur la présentation personnelle, dans les chats… Il est essentiel à la CMO.

Le destinateur est mis en exergue sur les forums. Sur SC, un lien hypertexte sur le prénom de l'utilisateur renvoie à sa page personnelle, et son avatar photographique figure à côté du pseudonyme.

Le destinataire est le seul élément absent de cette mise en forme du discours, parce que ce qui est écrit sur un forum est accessible à tous, inscrits ou non. Les messages sont des lettres ouvertes à tous, et n'importe qui peut répondre, pour peu que l'on corresponde aux critères de la plateforme [14].

14. Certains sites, comme *SlashDot*, ne nécessitent pas d'inscription pour commenter les messages. D'autres, comme *Something Awful* (SA), n'offrent l'accès aux forums qu'aux membres ayant acheté leur inscription sur le site.

Le contexte est rappelé à tous dans la mise en forme de la page virtuelle. La position exacte sur le forum, dans la section, sous-section éventuelle et fil de discussion est rappelée. Sur BP, beaucoup de sous-sections s'appellent « General » les discussions conduites ne sont générales que dans le cadre désigné par le titre de la section (« Health & Beauty », « Campus Life » …).

Le code commun aux membres du forum. Que ce soit l'anglais, le scots, ou des références communes, les utilisateurs partagent un langage communaulectal commun. Le mot pointé par la flèche est « caulder », jeu de mot sur la prononciation écossaise urbaine /ˈkʌul(d)/ de *cold* (source : OSD) et sur le nom du zoo de Glasgow, Calderpark, homophone de *cold*.

Le contact en tant que le site en lui-même est une interprétation large de « a physical channel and psychological connexion between the addresser and the addressee, enabling both of them to enter and stay in communication. ». Dans le cas d'un forum, il n'y a pas de fin de communication *in situ*. Les discussions asynchrones archivées permettent de pouvoir se connecter à nouveau, répondre et de se déconnecter : le forum est aussi un espace de *conversation permanente* (Bays & Encrevé, 2001), qui est le sujet de la sous-partie suivante.

Tous ces éléments se superposent sur les forums de discussion, et le chapitre suivant sera également l'occasion de voir comment les catégories de discours en ligne sont propices (ou non) à l'apparition de formes communaulectales (section 4.5).

3.1.4 Propriétés discursives de la communication asynchrone et des effets de citationnalité

Les forums de discussion ne suivent pas souvent le fil d'une pensée propre à la discussion sur d'autres médias et plateformes de discussion sur Internet. Il est tout à fait possible de répondre à un message plusieurs heures ou même plusieurs jours après sa parution dans le fil de discussion, et il est envisageable d'ignorer le dernier message, et de répondre à une intervention antérieure. Ceci est dû à plusieurs éléments : tout d'abord, les membres ont la possibilité de faire des citations de messages (comme c'est le cas dans toute autre plateforme de discussion sur internet, même celles permettant des échanges synchrones), avec l'avantage d'avoir une fonction de citation intégrée à tous les forums.

Les messages parus sur des fils de discussion n'ont souvent pas de récepteur particulier, chacun peut donc répondre à l'envi, par exemple pour préciser sa pensée dans un second message. Si le second message suit immédiatement le premier, il s'agit souvent soit d'une réponse *a posteriori* (par exemple, ayant envoyé son message en même temps qu'un autre intervenant, il répond à ce dernier dans un second message), soit d'une césure volontaire entre deux pensées différentes, voire entre deux conversations différentes sur le même fil de discussion. Il est possible de répondre à plusieurs interventions dans des messages séparés. Parfois, il peut y avoir des oublis dans le premier message, où une clarification est nécessaire.

Dans l'extrait présenté en contexte (illustration 3.14), deux utilisateurs (que j'appellerai Igno et Senor) se disputent à propos de la recrudescence de la violence. Cette partie est plus descriptive, mais elle montre les différentes strates discursives propres à internet, et le contexte qu'il faut savoir maîtriser pour naviguer dans une discussion réflexive. Le CEC montre deux messages consécutifs d'Igno, qui répond au même message de Senor. La situation conversationnelle est en fait plus complexe. Le schéma 3.15 résume la situation : Igno a écrit un premier message, que Senor a cité entièrement. Senor a cependant extrait la fin « In that particular case that you mentioned… » en y ajoutant des guillements, après la citation mise en forme par le forum, pour y répondre. Or, le message initial d'Igno (qui ne figure pas entièrement dans l'extrait CEC) est beaucoup plus long, et Senor décide de ne pas y répondre, en montrant qu'il trouve ce message stupide.

Le CEC est donc la reprise d'Igno de la fin de son propre message, de la réponse de Senor, qu'il réfute dans le premier message. Le second message d'Igno est une réflexion rajoutée après-coup, signalée par le « Also » initial. Il rajoute de nouveaux arguments, et la conversation peut ainsi repartir vers d'autres points de discussion.

Les citations et l'édition des messages permettent de montrer la sélection que l'on fait de chaque contenu lorsque l'on y répond. Le choix de la conclusion d'Igno par Senor montre une rupture avec la conversation en cours (il ne fait qu'émettre son avis sur l'intelligence d'Igno). Bien que l'illustration contienne à première vue deux messages, elle contient quatre messages différents, grâce aux citations enchâssées.

Les quatre messages cités ont été échangés le 18 juillet 2013, mais il n'y a pas de datation horaire

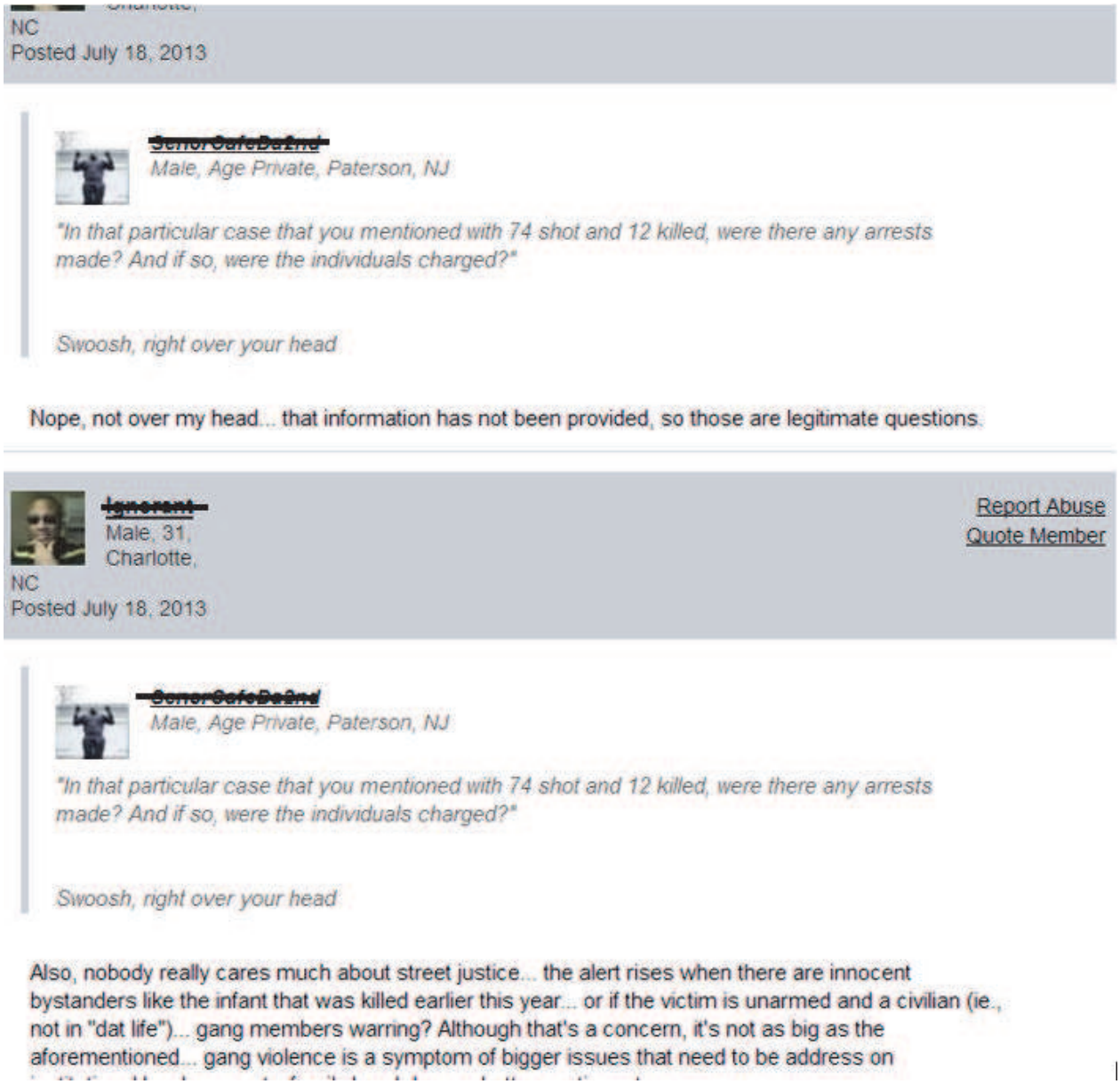


Illustration 3.14 – Extrait de « Random Thoughts 2013 » (BP12E17)

sur BP, ce qui empêche d'être plus précis concernant la façon dont les messages ont été échangés. Ces citations rompent avec la suite chronologique traditionnelle d'une conversation en face à face. Le tour de parole conventionnel s'arrête normalement avec la fin du message. Dans ce cas présent, Igno a même écrit à un autre utilisateur avant ses deux messages illustrés, en reprenant la citation pour

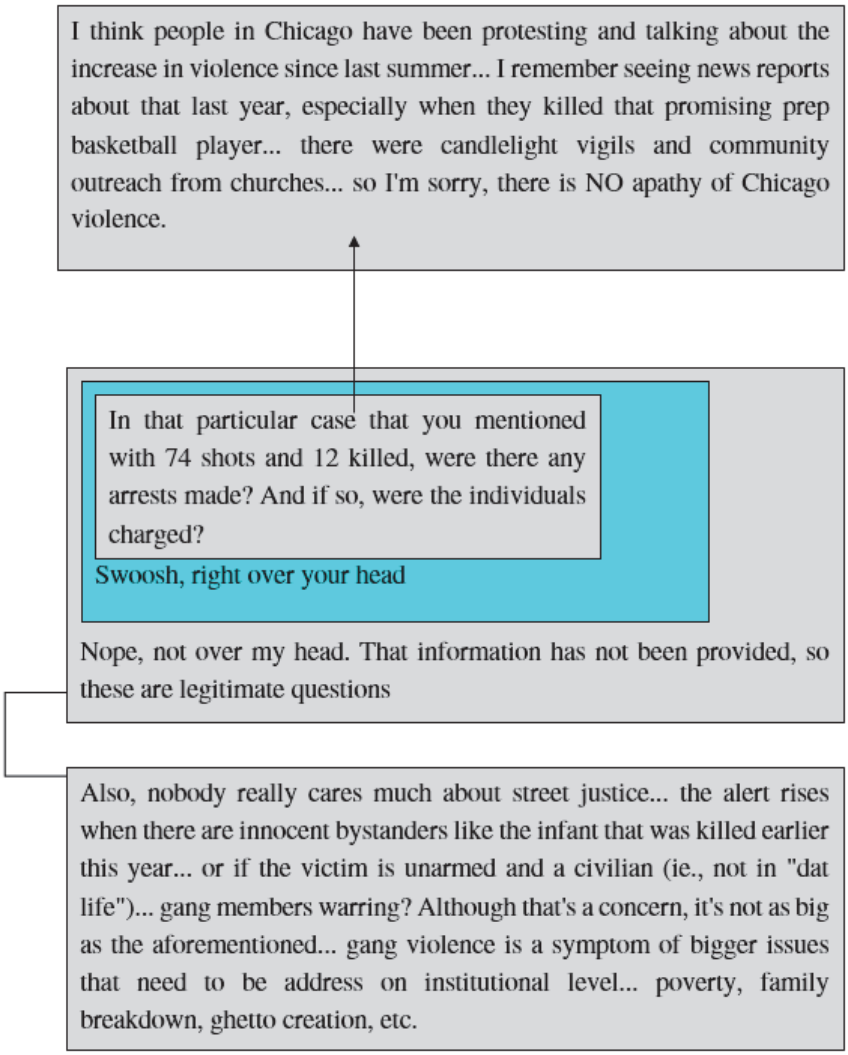


Illustration 3.15 – Schéma explicatif

montrer à qui il répondait. Les citations permettent de montrer à qui l'on s'adresse lorsqu'il y a plusieurs intéressés par la même conversation. Cette citationnalité est directe, et son inconvénient majeur est la césure du contenu écrit par une seule personne en plusieurs blocs. Cela rompt avec l'équilibre conventionnel du ratio un message pour un participant.

Sur SC, l'illustration 3.16 montre le message d'un utilisateur, Icanopit, qui répond à deux utilisateurs. D'abord Drake (dans le cadre rouge), qu'il ne cite pas directement mais qu'il appelle en début de sa première phrase. Il cite ensuite un site internet dont Drake avait donné le lien et lui redemande des précisions, avant de passer à la seconde partie (dans le cadre vert), pour répondre à Learmonth,

dont il cite directement un extrait du message avant de lui répondre.

La première réponse à Drake est une citation indirecte de son message. Sans l'intervention de

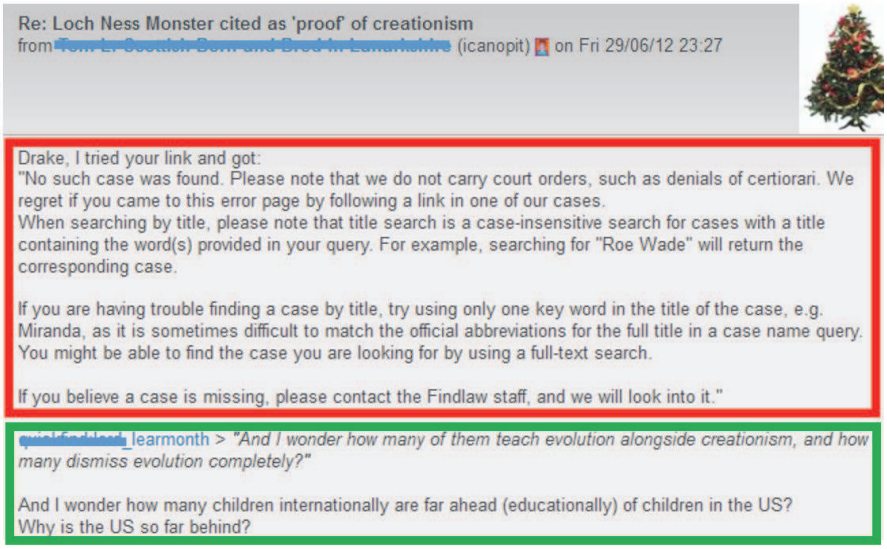


Illustration 3.16 – Extrait de « Loch Ness monster cited as "proof" of creationism » (SC11F16)

Drake, Icanopit n'aurait pas pu avoir l'adresse de ce site [15] Sa citation indirecte, introduite par un vocatif, ressemble plus aux conversations à voix haute, avec un *participant ratifié* décrit par Goffman comme étant un auditeur (ou destinataire, dans ce cas) vers lequel le destinateur « s'oriente de manière à indiquer que ses paroles [lui] sont particulièrement destinées » (Goffman, 1987). La seconde citation est directe, donc plus conventionnelle sur un forum.

L'asynchronicité des conversations sont propices à la création de nouvelles façons de communiquer. Les citations directes sont facilitées par les boutons offerts par la plateforme et par le contexte écologique de l'archivage perpétuel des messages déjà postés. Les citations indirectes ressemblent à la forme prise par la communication à plusieurs interlocuteurs, et leur forme est plus épurée que celles des citations directes. Le problème des citations directes est le volume qu'elles prennent, surtout pour les études textométriques. La lecture humaine prenant le contexte conversationnel en compte et le recueil de données hors-sol sont nécessaires pour trier les textes multipliés sur les forums. Il n'y a pas de citations d'un fil de discussion à un autre, mais la multiplication des résultats, notamment pour les

15. *Find Law*, site de compilation de précédents judiciaires américains, disponible à l'adresse <http ://caselaw.lp.findlaw.com/> (dernier accès : 14 juin 2014).

graphies alternatives, peut poser problème. La plupart des données recueillies dans le corpus final est nettoyée : les citations sont retirées pour ne garder que le corps du texte de la réponse, mais cela créée une lacune par rapport au CEC.

Quant aux données textométriques, elles sont recueillies par l'algorithme de Google (la section 3.3 abordera ces questions en détail), sans garantie que la même graphie alternative ne soit répétée dans une citation, à moins qu'il y ait une lecture humaine. Ce paramètre propre à l'étude des forums devra être pris en compte dans l'interprétation des résultats quantitatifs.

Une fois ces mises en garde faites, les spécificités des échanges dans les forums sont maintenant dégagées. Il est donc temps de présenter les deux forums principaux qui sont au cœur de cette étude, et de montrer les différences qui les séparent.

3.2 Analyse des systèmes écologiques

Les sites étudiés vont être scrutés dans cette partie, afin de comprendre les différentes façons dont les membres peuvent interagir. Ils ont peu de points communs, à part les outils offerts aux membres, et leur orientation vers des niches socio-culturelles (les Afro-Américains et les Écossais). Les stratégies mises en place pour rendre ces sites attractifs sont différentes, et les canaux de communication qu'ils mettent à disposition sont le sujet de cette section.

3.2.1 Présentation de *Black Planet*

BP est le plus gros site internet communautaire (réseau social) spécifiquement conçu pour cibler une population afro-américaine. Il a été lancé en 2001, ce qui le place parmi les sites communautaires les plus anciens d'internet. C'est le quatrième réseau social le plus visité, et il a été racheté en 2008 pour 38 millions de dollars par Interactive One, une entreprise basée à New York qui se spécialise dans les sites communautaires. Elle gère des sites ciblant particulièrement une communauté ethnique, et possède, par exemple, un site pour les asiatiques américains. Aucune statistique ou chiffre officiel

concernant le nombre de membre n'a été publié par Interactive One, Black Planet comptait un peu plus de 15 millions d'inscrits en 2008, et une estimation basse du nombre d'inscrits serait de 30 millions en 2014. Le matériel promotionnel du site parle de « plus de 20 millions de membres », et ce, depuis au moins les cinq dernières années.

La particularité de BP est le fait que ce site a su s'adapter, et évoluer afin de représenter un concurrent, sinon un complément à d'autres sites communautaires tels que Facebook. Constitué de quatre parties bien délimitées, BP se découpe en un site de rencontres, d'offres d'emploi, une partie informations sur divers sites satellites (*Hello Beautiful, News One*…) et la partie communauté en tant que telle. Il est possible d'utiliser son compte sur BP afin de commenter sur ces sites satellites ; cette transversalité rend donc ce site immersif et explique le grand nombre de membres.

La partie communautaire s'est enrichie au fur et à mesure du développement du site. Afin de communiquer, l'utilisateur du site dispose de moyens asynchrones, synchrones, publics, privés, unilatéraux ou de discussion. Les profils de membres comprennent un espace pour se décrire, un blog sur lequel les autres membres peuvent commenter, la possibilité d'écrire des sondages et de proposer des questions à choix multiples, et, depuis 2010, un espace où écrire une courte phrase visible par tous les membres, qui n'est pas sans rappeler *Twitter* ou le mur de *Facebook*. Les membres peuvent discuter en privé, soit par messages synchroniques sur une plateforme de clavardage, soit par messages ressemblant à des courriels.

Les membres peuvent également communiquer en groupe sur diverses plateformes : forums, groupes d'intérêt et salons de discussion type *Internet Chat Relay* (IRC) qui agrègent les utilisateurs autour de sujets ou de passions communes. Toutes ces possibilités de discussion sont accessibles grâce au tableau de bord en haut de toutes les pages du site. Le tableau 3.6 résume les différentes interactions possibles sur le site.

Les options de discussion sont plus souvent asynchrones, à l'exception du chat, qui est conversationnellement synchronique : bien que la réponse de chacun soit envoyée d'un bloc par l'auteur du message, il est envoyé rapidement, et est immédiatement lisible par les autres membres du salon de discussion. Cet élément de BP n'est pas recueilli dans le corpus, les messages périment rapidement (après

	Synchrone	Asynchrone	Public	Privé
Blog, sondages sur le profil		⊗	Appréciation laissée au membre	Appréciation laissée au membre
Chat	⊗		Sur les salons	Possibilité d'écrire à un seul membre
Forums		⊗	⊗	
Profils		⊗	⊗	
Statut (*What's new ?*)		⊗	⊗	
Commentaires sur les actualités		⊗	⊗	
Photographies		⊗	Appréciation laissée au membre	Appréciation laissée au membre
Groupes		⊗		
Messages privés		⊗		⊗

Tab. 3.6 – Caractérisation des interactions sur BP

une cinquantaine de lignes dans la fenêtre de chat), contrairement à ceux sur IRC qui se conservent depuis le début de l'entrée dans le salon de chat jusqu'au départ de l'usager.

La nature publique ou privée des messages ne pouvait pas être contournée, et, à moins d'être ami avec un membre [16]. Seuls les messages de blogs et les éléments de profils personnels comme les photographies étaient affectées par ces restrictions, qui n'ont donc pas posé de problème pour le recueil de corpus. De nombreux utilisateurs laissent leurs profils personnels sans restriction, et ce travail est principalement tourné vers les messages du forum lui-même, qui n'a pas de partie restreinte ou privée.

Selon Alexa, un site d'analyse du trafic sur les sites internet, les mots-clefs de référencement pour trouver BP sont répartis en plusieurs catégories :

Relatifs au lexème « African Americans » : 4 mots-clefs : african american, African American, african americans online, african american personals

16. Pour cela, il faut que l'utilisateur A envoie une invitation que l'utilisateur B accepte pour qu'ils soient mutuellement amis, comme sur FB.

Relatifs au fait d'« être noir » et contenant le mot *black* : 36 mots-clef : *Black Entertainment Television* (BET) [17] , black, Black, blacks, Blacks, black community, black culture, black history, black music, blacks online, blackplanet, black planet, blackvoices, black girls, black women, black men, black chat, black dating, black people meet, black personals, black singles, black sororities, black friends, black fraternities, black blogs, caribbean american, black classifieds, black community, black communities, find black people, black groups, black love, black matches, meet black people, black networking, local black singles, black social networking

Relatifs à d'autres sujets 17 mots-clefs dont les lemmes principaux sont la *soul* ou la musique hip hop [18].

Ces mots-clef sont centraux pour les outils d'optimisation de moteurs de recherche, qui décrivent la façon dont les internautes arrivent sur un site. Ils sont caractéristiques des variantes graphiques auxquelles on pourrait s'attendre sur les forums étudiés. Les mots-clefs représentent entre autres des recherches où il n'y a pas de césure de mots, comme pour chercher l'adresse du site (mot-clef « blackplanet »). D'autres mots clefs sont capitalisés (« African Americans », « Blacks »), comme ils le seraient hors-ligne. Ils montrent à quel point l'appartenance ethnique est primordiale, ainsi que le désir de communiquer (« find black people ; black people meet »), voire de trouver l'amour (« black dating ; black singles ; black love »). Les aspects communautaires passent aussi par la création d'un réseau personnel (« black communities ; black social networking »). Créés par les fondateurs du site, qui indexent ces mots-clef sur toutes les pages pour optimiser les recherches, ils sont complétés par les mots-clef qui ont permis aux internautes de trouver BP. Ceux-ci connaissent le site, puisque le mot-clef le plus populaire est « blackplanet » (20.59%), suivi par « black dating sites » (14.89%), selon Alexa.

17. BET est le nom de la chaîne de divertissement afro-américaine
18. Alexa est devenu payant en 2009, et la liste complète des mots-clés n'est plus accessible gratuitement.

3.2.1.1 Architecture de BP

BP est un forum très bien organisé, et construit autour de 18 sections principales, qui contiennent elles-mêmes entre une et 18 sous-sections. Chaque section traite d'un thème particulier, et la carte de ces sections est présentée dans l'illustration 3.17. Comme dit précédemment, c'est un forum de niche de taille importante. Selon l'illustration 3.18 qui montre la page d'accueil de toutes les sections de BP au 2 août 2014, le forum compte plus de deux millions de messages (2 050 576 au total), répartis dans environ un million de fils de discussion (1 009 765 fils). Sur internet, il y a des forums plus importants, comme *Gaia Online* [19] qui totalise plus de deux milliards de fils de discussion, et plus d'un million de messages par jour (Au, 2007). Pour comparer BP à d'autres sites de niche hors jeux vidéo, 4Chan est la plus grosse plateforme générale de discussions et contient au 2 août 2014, 1 626 430 977 messages dans les parties actives et non-archivées du site [20]. SA, autre forum et site d'informations, contient 3 253 923 fils de discussion pour 154 906 301 messages. BP est un petit site par rapport à ces concurrents, mais il n'est pas ancré dans la culture internet, contrairement aux deux autres forums, notamment *4Chan*, qui est le berceau du mouvement de pirates informatiques justiciers Anonymous. Ces deux plateformes mettent plus l'accent sur le partage sur forum que sur le profil personnel des membres ou la possibilité d'avoir un blog, et le forum de SA est un satellite d'un site d'informations sur la culture internet.

Comme souligné précédemment, sur BP, ce sont les sous-sections « General » des sections « Health & Beauty » et « Relationships » et qui centralisent les plus grands nombres de fils de discussion, avec respectivement 384 854 et 48 925 fils de discussion. Certains dépassent les 3 000 messages, comme « Why Are Black Men So Concern About Black Women's Hair ? » [sic] (BP12E2402) et ses 3 872 messages, ou « why do some black women opt to only date black males ? » (BP18A763), qui comporte 3 468 messages. D'ailleurs, ces deux messages sont promus en tant qu'« Hot Topics » sur BP, une rubrique qui figure sur la page recensant les sections du forum.

Le ratio entre le nombre de fils de discussion et le nombre de messages sur BP est symptomatique

19. Gaia Online est un jeu en ligne lancé en 2003. L'adresse du forum est <http ://www.gaiaonline.com/forum/> (dernier accès le 2 août 2014).
20. Il n'y a pas de statistiques concernant le nombre de fils de discussion.

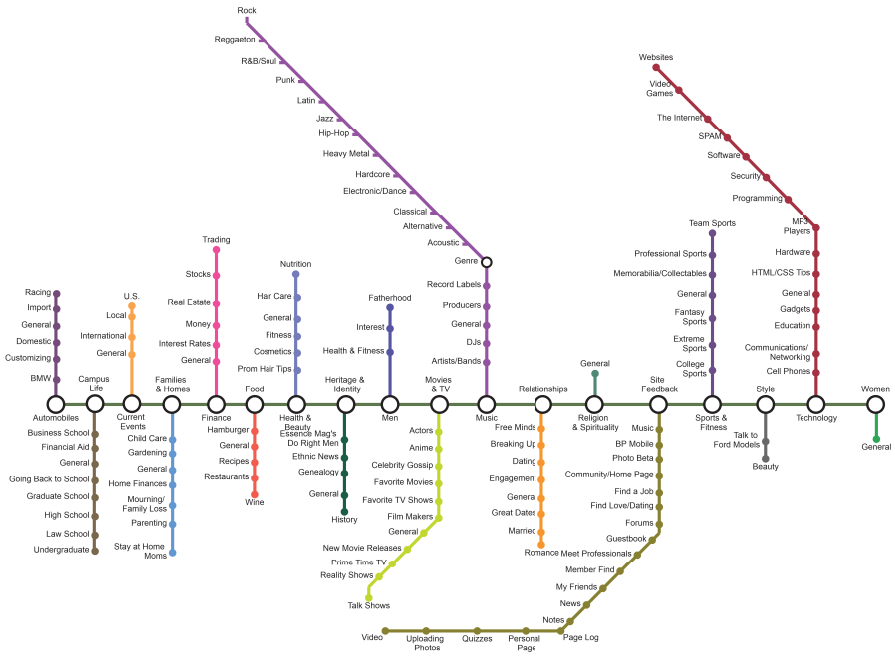


Illustration 3.17 – Architecture des sections et sous-sections de BP

d'un problème qui affecte les forums populaires. Le forum de BP est une plateforme régulièrement (mais non fréquemment) nettoyée par les administrateurs. Victime de vagues de spammeurs, et plus précisément de robots (appelés bots en ligne) qui postent des messages sans réponse, les 60 à 100 premiers messages de chaque sous-section du forum sont des publicités dupliquées par ces robots. La popularité de BP en est la cause : il est plus intéressant de cibler des sites fréquentés que des plateformes plus petites comme SC. Si la plupart des comptes créés à ces fins a été supprimée, certains continent à sévir sur BP, comme illustré par l'image 3.19. Ces fils de discussion ont des noms typiques du spam, ou courriel publicitaire, comme « #1 Cougar Hookup Site », et les graphies alternatives n'y sont pas utilisées puisque BP ne censure pas les mots-clef publicitaires, contrairement aux messageries type *Gmail*, qui savent déterminer quelles suites de mots sont à mettre directement dans le dossier « spam » [21]. Ces bots ont toutes les apparences d'humains, avec des pseudonymes contenant

21. Pour plus d'informations, voir l'analyse en profondeur de D. Crystal des courriels, chapitre « The language of e-mails » (Crystal, 2011).

Forum Category	Topics	Posts	Last Post
Automobiles	32,775	34,910	6 hours ago by: jack14381 View Post
Campus Life	45,242	45,695	6 hours ago by: jack14381 View Post
Current Events	32,008	99,960	6 hours ago by: jack14381 View Post
Families & Home	35,144	37,352	5 hours ago by: jack14381 View Post
Finance	33,274	35,069	6 hours ago by: jack14381 View Post
Food	23,634	26,025	6 hours ago by: jack14381 View Post
Health & Beauty	479,527	482,513	6 hours ago by: jack14381 View Post
Heritage & Identity	15,484	80,546	8 hours ago by: Dirty2DaCore View Post
Men	16,493	32,117	8 hours ago by: Dirty2DaCore View Post
Movies & TV	19,159	20,753	8 hours ago by: Dirty2DaCore View Post
Music	15,888	19,887	8 hours ago by: Dirty2DaCore View Post
Relationships	202,685	789,649	4 hours ago by: SenorCafeDa2nd View Post
Religion & Spirituality	19,423	260,792	8 hours ago by: Dirty2DaCore View Post
Site Feedback	10,739	11,665	July 30 by: jack14381 View Post
Sports & Fitness	5,798	16,000	1 hour ago by: bappyvalonai View Post
Style	5,367	6,814	July 28 by: jack14381 View Post
Technology	7,220	7,521	July 28 by: williamjame81 View Post
Women	9,905	43,308	6 hours ago by: jack14381 View Post

Illustration 3.18 – Organisation des sections du forum

des prénoms et des photographies en avatar, pour être au plus proche de l'humain et donner envie aux internautes de cliquer sur les fils de discussion. Ce ne semble pas être le cas, puisque même les messages qui ont plus d'un mois, comme ceux à la fin de la liste, n'ont aucune réponse. Les publicités n'ont pas été comptabilisées dans le corpus, et ne contiennent pas de graphies alternatives.

Hormis cette spécificité, aucune section du forum n'est vide, ou lacunaire. Les sous-sections ont été créées en fonction des besoins (voir 3.2.2.1), ce qui explique de curieuses disparités entre les sections : la section pour les hommes (Men, BP9) a trois sous-sections, « Health & Fitness », « Interest » et « Fatherhood » (respectivement BP9A, B et C), alors que la section pour les femmes, « Women » n'a que la sous-section « General » (BP17A). Il n'y a pas de nouvelle section ou sous-section du forum depuis le début de ce travail de thèse, ce forum a donc atteint sa maturité fonctionnelle à l'échelle du temps internet.

Le système de nomenclature des fils de discussions du forum sera discutée dans la section suivante,

All Forums > Automobiles > Customizing

Start a New Topic

Forum Topic	Posts	Topic Starter	Last Post
HELLO EVERYONE	1	July 09 by: corinta	July 09 by: corinta View Post
Sell CVV good (ICQ: 667027466) and fresh, track dumps(101,201) valid 100%	1	July 03 by: Osolug07	July 03 by: Osolug07 View Post
ICQ: 684239409/ Skype: jame william/ Cvv/Paypal/Bank login/ Transfer WU/ Sh	1	June 27 by: warker	June 27 by: warker View Post
#1 Cougar Hookup Site	1	June 14 by: EmmyBrady2002	June 14 by: EmmyBrady2002 View Post
#1 Cougar Hookup Site	1	June 14 by: EmmyBrady2002	June 14 by: EmmyBrady2002 View Post
Specialists in Black and White Dating services ...	1	June 11 by: kirchen	June 11 by: kirchen View Post
HACK FACEBOOK ACCOUNT of your lover	1	May 30 by: irina_mole	May 30 by: irina_mole View Post
HOW TO SPY ON TEXT MESSAGES of your lover	1	May 30 by: irina_mole	May 30 by: irina_mole View Post
LOCATE A CELL PHONE on google map for FREE	1	May 30 by: irina_mole	May 30 by: irina_mole View Post

Illustration 3.19 – Extrait de la première page de la sous-section « Customizing » de la section Cars

en 3.3. Après cette brève présentation des caractéristiques de BP, je vais présenter l'autre élément sur lequel repose cette thèse, le forum SC et ses caractéristiques.

3.2.2 Présentation de *Scotster*

Le forum SC a une envergure bien moindre, mais il se situe aussi au coeur d'un site mettant les informations des membres en exergue, comme BP. Son ambition de départ est d'être un réseau social

comme FB, pour fédérer des membres autour d'intérêts communs. Ces éléments sont essentiellement culturels, sinon géographiques. Ils ont trait avec le fait d'être écossais, ou d'aimer l'Écosse, comme l'indique l'une des accroches du site : « 100% Scottish ; Scots ancestry, friends or family ; or simply love Scotland ? ». La lettre d'accueil du créateur du site est sans appel à ce sujet : il décrit SC comme « the world's leading social network for Scotland, dedicated to bringing people together online and through local events. ». La volonté de réunir des Écossais ensemble n'a pas été testée dans ce travail, mais la sous-section « Event and Meetup Ideas » contient un fil de discussion où une rencontre pour le Bannockburn Live en juillet 2014 a été organisée et où quelques membres semblent s'être rencontrés (SC1H17).

Il n'y a pas de statistique fiable pour délimiter le nombre de membres du site, et l'information n'est pas non plus donnée par les fondateurs du site. On peut estimer, à partir de leurs propos, qu'il y a à peu près 7 000 membres en septembre 2014 [22]. La majorité de la population du site semble venir des États-Unis et du Canada, parce qu'il y a une population d'origine écossaise, parlant le gaélique dans la province de Nova Scotia, mais ces informations récupérées sur Alexa sont invérifiables. Le site est trop bas dans son classement, et seuls les 100 000 sites les plus visités ont un analyse en profondeur disponible. SC est classé bien plus bas, et ses fondateurs n'ont pas demandé à Alexa de faire une estimation certifiée, d'où ce manque d'information. Avant le rachat d'Alexa par Amazon, quelques informations avaient pu être recueillies en 2011. Les membres de SC sont plus vieux que ceux de BP, en général : comparativement à la population d'internet, les 45-54 ans y sont plus représentés (« Alexa - The Web Information Company », 2000). SC fait partie du premier million de sites visités dans le monde (919 414ème site le plus visité en octobre 2011 ; il est passé à 700 896ème le 27 juillet 2014, et un relevé le 2 août le place en 880 279ème position [23].

22. Il y aurait à peu près un à deux nouveaux membres par jour, soit entre 30 et 45 nouveaux membres par mois.
23. Il est impossible d'avoir des détails plus précis.

3.2.2.1 Architecture de *Scotster*

SC est un forum lacunaire : certaines sous-sections sont parfois vides, ou mal adaptées au contenu ; l'exemple le plus flagrant est la section 25, « Support » où, sur 28 sous-sections, 19 sont vides. C'est un défaut créé par les fondateurs du site, qui ont voulu créer artificiellement une demande alors que l'offre n'existe pas, et ainsi être le plus exhaustifs possible. Le tableau 3.7 complète l'illustration 3.22, qui est difficile à lire [24]. Les étiquettes de la carte représentées par des ronds vides symbolisent les sous-sections vides. La page d'accueil du forum contient à la fois les sections et sous-sections sur la même page, comme le montre l'illustration 3.20, ce qui ne permet pas une capture d'écran aussi facile que pour BP.

La particularité de SC est de permettre la navigation non par l'interface principale, mais grâce aux mots-clefs contenus dans un nuage sémantique pondéré par le nombre d'utilisateurs qui ont recours à ces mots-clefs. Le nuage du 2 août 2014 est disponible dans l'illustration 3.21. Le site invite à naviguer différemment, grâce aux affinités avec un sujet plus qu'avec un arbre de choix statique.

24. Une version en haute résolution sera disponible sur la clef USB d'annexes.

Illustration 3.20 – Extrait de la page « Browse Forums » de SC

Illustration 3.21 – Nuage de mots-clef du menu « Explore Tag » de SC (2 août 2014)

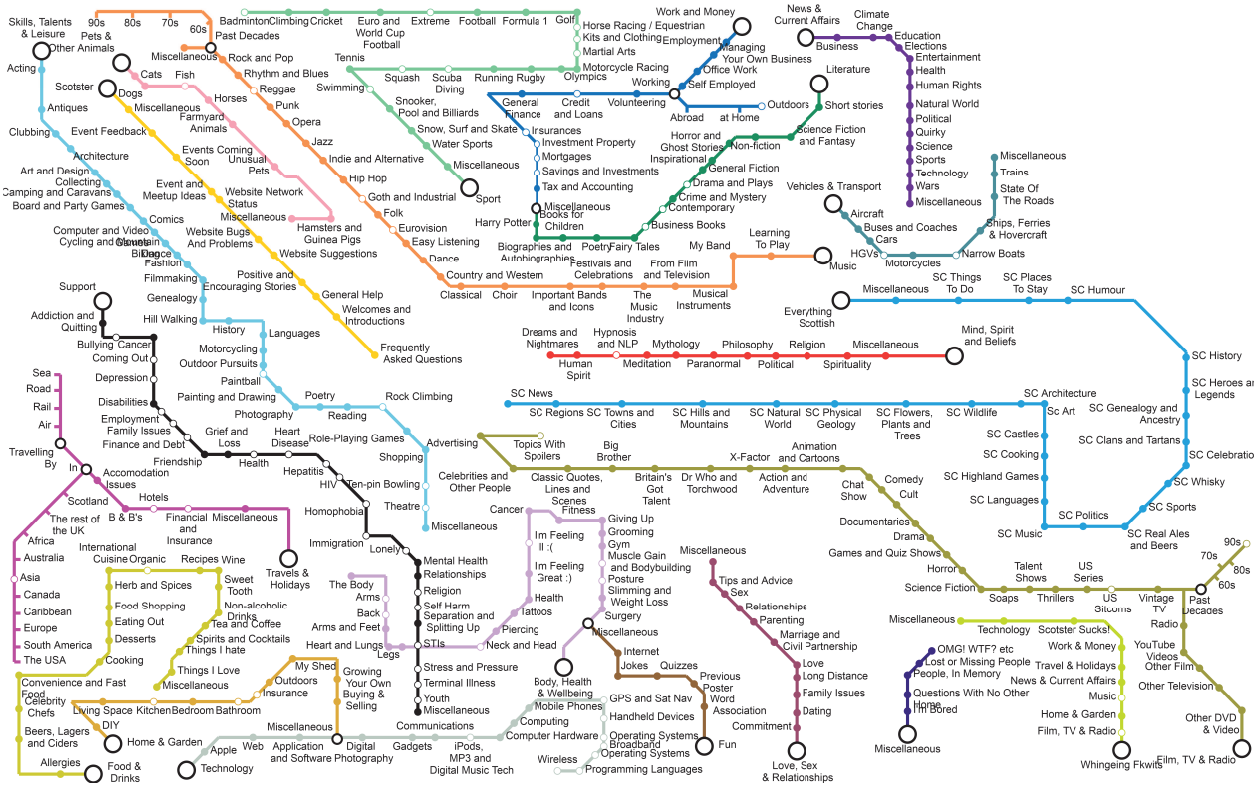


Illustration 3.22 – Carte des sections et sous-sections de SC

L'architecture des fils de discussion, ou architexture de SC n'est pas nettoyée, et les messages sont postés dans les bonnes sections : un fil de discussion « Euthanasia for or against ? » (SC1K16) a été posté dans une des sous-sections « Miscellaneous » sur le site (1SCK). C'est une faute imputable aux utilisateurs, qui décident de poster leurs messages sans observer la structure de SC, alors qu'une section du forum est plus adaptée (la section « Health »). Les administrateurs ne font pas non plus leur travail *a posteriori*, et ne classent pas les messages dans les bonnes sections. Ceci ne facilite pas la communication, ou les réponses qui pourraient être apportées par des utilisateurs qui seraient inéressés. Cependant, il est possible de retrouver ces messages, et les réponses grâce au *Pulse* de SC (voir 3.5.0.1). Certaines sections sont des redites : chaque section contient une sous-section de sujets divers (Miscellaneous), mais il y a aussi une section Miscellaneous. Ceci étaye l'hypothèse du désir d'exhaustivité décidée par l'administrateur [25].

25. En revanche, la présence d'une sous-section « Totally Miscellaneous » dans la section « Miscellaneous » est absconse.

Section	Nombre de sous-sections	Sous-sections vides	Nombre total de messages
1- Scotster	11	–	679
2- Everything Scottish	29	–	2130
3- Music	29	3	215
4- Film, TV & Radio	33	3	151
5- Literature	16	3	221
6- Food & Drink	21	2	180
7- Home & Garden	12	4	33
8- Skills, Talents & Leisure	31	3	130
9- Pets & Other Animals	8	1	63
10- Sport	22	8	129
11- World Interest	10	–	248
12- Travel & Holidays	25	4	84
13- Science & Nature	7	–	110
14- Body, Health & Wellbeing	24	8	75
15- Mind, Spirit & Beliefs	12	1	91
16- Work & Money	16	6	46
17- News & Current Affairs	15	–	202
18- Vehicles & Transport	10	2	24
19- Technology	17	7	40
20- Love, Sex & Relationships	11	–	76
21- Classifieds	8	1	44
22- Fun	15	1	614
23- Whingeing Fkwits	9	1	48
24- Miscellaneous	6	–	309
25- Support	28	19	11

Tab. 3.7 – Architecture de Scotster

Dans SC, l'esprit de cohésion est visible : contrairement à BP, SC est une petite plateforme, qui peut donc laisser les utilisateurs se présenter lors de l'inscription (pour se sentir bien accueillis, et oser participer à la communauté). Ces fils de présentations personnelles ont de nombreuses réponses : les nouveaux membres osent se présenter, mais surtout, les membres habitués sont réactifs, et accueillent les autres avec plaisir (ce qui explique ces messages nombreux : les habitués saluent individuellement chaque membre) [26].

26. En m'inscrivant, j'ai directement reçu un message d'une utilisatrice habituelle, qui a même tenté d'utiliser le français pour me saluer.

3.2.3 Justification du choix et des limites du corpus

Ces plateformes ont été choisies parce qu'elles incarnent chacune une spécificité du discours en ligne. Les membres plus jeunes de BP utilisent plus de jurons, qui seront étudiés notamment au chapitre 5. Les censures personnelles sur BP répondent à la nécessité de se censurer, puisque la plateforme possède une liste de mots-clef proscrits (un thème qui sera abordé en 4.4.5). Les utilisateurs de SC utilisent moins de jurons, mais ils se censurent également. L'origine géographique compte également : SC a une grande proportion de Britanniques et de Canadiens, alors que BP est à grande majorité peuplé d'Américains. Les deux forums utilisent des graphies et des langues minoritaires. Les membres de BP ne connaissent pas forcément le statut de l'AAV, mais ils ont recours à la création lexicale et graphique. Le vocabulaire du gang (étudié dans le chapitre suivant, en 4.4.5.1) en est un bon exemple, et est un mélange paradoxal de cryptographie et d'essais didactiques. À l'inverse, les membres de SC veulent souvent apprendre ou utiliser les mots de scots qu'ils connaissent.

La disparité entre les deux forums permet une analyse prenant en compte tous ces facteurs, et permet de montrer quelques caractéristiques discursives des sites de niche, basés sur l'appartenance culturelle ou ethnique plus que sur les intérêts purs. Les études de ces deux plateformes seront réparties entre analyse de discours, dans le chapitre 4, où les conditions d'apparition des graphies alternatives seront étudiées ; et la graphophonématique sera le sujet du chapitre 5, où les substitutions graphiques permettront de peindre le reflet des pratiques orales. Ce tableau sera lacunaire, parce qu'il n'est pas possible de savoir comment les membres parlent, mais il permet de confirmer ou d'infirmer quelques propriétés phonologiques sur la phonologie de ces langues minoritaires.

Les études des deux corpus sont légèrement différentes. Sur SC, plus de pages personnelles et de statuts seront étudiés, avec leurs commentaires. Pour BP, il y aura plus d'études de profils personnels et les « News Articles », entre les brèves d'actualité et les messages de blogs. L'étude de BP est un bon point de départ pour étudier l'intersection entre l'argot américain et son influence sur le netspeak, ou discours sur internet.

La variation graphique intra-scripteur ne sera pas étudiée, mais la variation inter-scripteur est un point important dans l'étude, notamment dans la partie 4.3.2, qui s'intéressera à la variation entre

<gurdle> et <girdle> dans le même fil de discussion. Il est impossible de retrouver un message d'un utilisateur à partir de son profil personnel, et le volume des deux forums est beaucoup trop conséquent pour faire une étude sans plus d'outils informatiques que ceux utilisés ici.

3.3 Positionnement épistémologique et collecte des données

Le corpus utilisé dans cette étude est double, voire dupliqué : il parait important de conserver le contexte écologique de chaque message, et de l'étudier en tant que part de cet écosystème (puisque les utilisateurs de forums perçoivent les messages et y réagissent au sein d'un contexte qu'on ne peut séparer de leurs productions). C'est la raison pour laquelle le fichier général de recensement de l'architecture des plateformes étudiées ne contient que le nom du fil et son adresse internet, et qu'il n'y a pas de corpus copié directement dans les annexes.

En revanche, certaines analyses statistiques ont, elles, dû être effectuées sur l'extraction des messages et de leur contenu. L'analyse textométrique des productions ne peut être faite sans enlever les références aux outils métalinguistiques du forum de discussion (commande « répondre », ou « citer », par exemple) ou à l'écologie complète de la structure. Un utilisateur répond à d'autres, en fonction de qui est présent sur la page, et des outils qu'il a a disposition. Le forum est une entité vivante à part entière, et il semble important de garder en tête l'écologie complète d'une page de fil de discussion, même lorsqu'elle est réduite à un suite de citations dans ce travail.

Le corpus le plus fréquemment cité est le corpus d'extraction, donc les recours au corpus en contexte seront précisés par l'utilisation de l'abréviation CEC. La nomenclature de l'architecture sera décrite plus loin dans cette partie. Elle est conservée dans un annuaire pour chaque site, créé en utilisant le logiciel Excel [27] afin de pouvoir conserver certaines données dont les URL, et les ordonner, ou pouvoir faire des requêtes par rapport aux noms des fils de discussion choisis. La constitution du corpus reposait initialement sur plusieurs bases : l'aspiration de sites et la sélection de fils de discussion sur divers critères, comme la fréquentation de la section du forum, et la longueur du fil de discussion. Si

27. Versions 2007-2010-2013.

des fils courts peuvent comporter des éléments intéressants et montrer des caractéristiques uniques, les discussions font apparaître des caractéristiques uniques, comme les vocatifs, et les marques d'interaction et de réponse. La justification et la méthode utilisée est discutée en 3.3.3.

L'analyse des éléments du corpus ne s'arrête pas aux fils de discussions seuls : les présentations de membres peuvent aussi avoir de l'importance dans la façon dont ils sont construits, puisqu'ils contiennent plus souvent des parties hautement stylisées (dans la mise en forme du texte, avec sélection de couleurs de police, par exemple, comme dans la façon d'écrire). Dans une moindre mesure, puisqu'ils sont anonymisés, les pseudonymes jouent un rôle important dans la présence en ligne, et ils seront également inclus au corpus, par le biais des données statistiques recueillies.

3.3.1 Méthodologie de recherche

Après avoir constitué le corpus, il faut non seulement le traiter, mais également créer une méthodologie de recherche. Certains problèmes sont été rencontrés lors de la constitution du corpus, ceci a guidé les choix effectués dans cette thèse. L'aspect éthique a joué un rôle dans la façon dont les informations sont traitées et conservées.

L'aspect hybride de ce corpus (images, vidéos, sous-titres, livres…), même au niveau du discours sur internet (page personnelle, échange sur forum, pseudonymes…) a posé moins de problèmes qu'attendu : les méthodologies étaient différentes, et comptaient sur des explorations sémiotiques diverses. C'est pourquoi cette thèse est abondamment illustrée d'exemples de mises en page.

Constituer un corpus et l'ordonner ont été les tâches les plus exigeantes lors de la conception de ce travail. La partie suivante est un récapitulatif des étapes qui ont donné naissance à la forme qu'a prise ce travail.

3.3.1.1 Méthodologie de la constitution du corpus

Dans cette sous-partie, j'aborde les raisons et limitations qui ont donné leur forme définitive aux différentes parties du corpus. Les problèmes les plus difficiles à circonvenir furent les problèmes tech-

niques d'une part, et les problèmes humains, qui m'ont poussée à changer l'orientation méthodologique de ce travail de thèse.

3.3.1.2 Problèmes lors de la constitution du corpus

3.3.1.2.1 Problèmes techniques

Le corpus internet sur lequel se base cette thèse forme un ensemble relativement fermé. Il représente le moment de captation, et devient figé à l'instant où le chercheur les aspire ou les collecte (bien que les analyses puissent être transposables à d'autres corpus). Cette aspiration s'est d'ailleurs révélée impossible : HTTrack (Roche, 2005) était le logiciel envisagé pour aspirer toutes les parties des forums, et les avoir en archivage local. BP interdit l'aspiration automatique de sites et bloque les connexions multiples qui auraient été nécessaires pour mener cette tâche à bien. L'adresse IP utilisée pour tenter l'aspiration a été bloquée par BP, donc il m'a été impossible de me reconnecter au site pendant plusieurs mois.

Pour SC, la plateforme est ancienne, et représente la seconde incarnation d'un site plus ancien qui s'appelait « Scot Clan ». Cette double construction a rendu l'aspiration beaucoup plus difficile, pour ne pas dire impossible : les redirections automatiques de l'ancienne adresse vers la nouvelle interdisent également l'aspiration. Cette méthode a donc été éliminée, parce qu'inutile. De plus, si le corpus était recueilli trop tôt, certaines formes n'auraient pas pu être observées. Il était impossible de faire des études parallèles et simultanées sur toutes les formes étudiées, et il y aurait toujours eu un décalage entre le corpus extrait et les forums. Ce décalage existe aussi avec les études faites sur le corpus en contexte, mais les fils de discussion sont toujours étudiés par rapport à leur URL et non par une captation, ce qui veut dire que les études se portent toujours sur la version la plus récente du site.

3.3.1.2.2 Problèmes humains

Dans la phase préliminaire d'exploration de sites pour créer un corpus, d'autres plateformes vir-

tuelles ont été sélectionnées, puis abandonnées : *Cocoa Lounge*, dont il reste quelques reliquats dans ce mémoire de thèse, était un forum afro-américain plus destiné aux femmes qu'aux hommes. Après m'être inscrite sur la plateforme et présenté mon projet de recherche, mon compte et mon adresse IP ont été bannis de façon permanente. Il y a deux façons de bannir un compte, le bannissement mécanique de l'adresse IP, comme ce qui était arrivé lors de l'aspiration de BP, et le « permaban », mot de l'argot internet, qui désigne le type de bannissement permanent d'une plateforme internet. *Cocoa Lounge* venait en 2009 de se remettre d'une attaque venue d'un forum raciste qui avait provoqué la fermeture provisoire du site, et n'avait aucune envie que les écrits de leurs membres soient lus ou étudiés. Je n'ai pas été notifiée de ce permaban avant qu'il n'arrive, et n'ai donc pas poursuivi les recherches sur *Cocoa Lounge*. Ce forum ne fait pas partie du corpus étudié.

Une étude de membres de BP était également envisagée, avec enregistrements sonores et comparaisons entre les prononciations et les graphies de ceux qui acceptaient de se prêter à l'expérience. Le séjour de deux ans à Austin en tant que lectrice à l'université du Texas entre 2008 et 2010 semblait idéal pour conduire cette partie de l'enquête. Malheureusement, poster de petites annonces sur le campus a été infructueux, parce qu'il n'y avait aucune compensation financière, ce qui n'est pas habituel pour les expériences scientifiques sur le campus. Les membres de BP qui m'abordaient sur ma page personnelle n'étaient pas du tout prêts à participer dans des conditions acceptables et sans contrepartie. J'ai donc abandonné cet aspect de la recherche, et c'est la raison pour laquelle l'exploration de corrélats graphiques a été abordée par le recours au corpus secondaire d'œuvres littéraires et de leurs adaptations en film [28].

3.3.1.3 Corpus

Le corpus est composé de deux parties principales, pour chaque forum. D'abord, un annuaire de l'architecture de chaque forum, qui attribue un code alphanumérique unique pour chaque fil de conver-

28. Une micro-étude sur le phénomène internet de l'*Accent Tag* est utilisée pour confirmer certaines remarques faites sur le corpus. Le corpus constitué est référencé au SLDR, à l'adresse <http ://sldr.org/sldr000830/fr>.

sation, y compris ceux qui ne sont pas étudiés dans cette thèse. J'ai recueilli les adresses de tous les fils de conversation dépassant les 5 messages suivant le système suivant :

$$\underbrace{BP/SC}_{\text{Forum}} \quad \underbrace{1\ldots}_{\text{Code de la section}} \quad \underbrace{A\ldots}_{\text{Code de la sous-section}} \quad \underbrace{1\ldots}_{\text{Code du message}} \quad (\quad \underbrace{-1\ldots}_{\text{Ordre du message}} \quad).$$

Ce système permet d'identifier rapidement les fils de conversation, mais il est bien plus adapté aux requêtes informatisées qu'à la lecture humaine. C'est pourquoi la clef USB d'annexes contient cette liste, avec les liens hypertexte renvoyant aux pages directement sur internet. Les annexes de cette version papier contiennent les tableaux A.6.2.1 et A.6.1.3, qui listent les codes pour chaque section et sous-section.

Il existe quelques cas particuliers : pour BP, les pages personnelles sont indiquées par le code « BPPP[numéro d'ordre] » et les messages du livre d'or de chaque membre sont indiqués par « BPGB[numéro d'ordre ». Les autres éléments de la plateforme comme les *news articles* sont indiqués avec le code « BPNA[numéro d'ordre » et les groupes de membres (un seul est mentionné dans ce travail de thèse) sont indiqués par « BPG[numéro d'ordre ».

L'architecture de SC est un peu différente, et au lieu de livre d'or sur les pages personnelles, la plateforme conserve les statuts de profil à la FB. Pour en prendre compte, j'ai utilisé une nomenclature particulière, reposant sur la forme de l'URL de la page. Les pages personnelles (dont l'URL contient <www.scotster/people/> sont notées avec le code « SCPP[numéro d'ordre] ». En revanche, les statuts personnels (adresse en <…/today>) sont marqués avec un <P> simple (« SCP[numéro d'ordre] ») et les commentaires sous les photographies sont indiqués avec le code « SCPH[numéro d'ordre] ». Le tableau 3.8 récapitule les nomenclatures choisies pour tous les éléments hors forum pour BP et SC. Les études sur pseudonymes ne sont pas incluses dans le corpus, puisque le pseudonyme permet une identification trop directe des membres, et que ces pseudonymes ne sont étudiés que dans ce chapitre, dans la partie suivante. Le corpus à proprement parler est un tableau composé pour chaque forum,

BP		SC	
PP		Page Personnelle	
NA	News Article	P	Profil
GB	Guest Book	PH	Photographie
G	Groupe		

Tab. 3.8 – Récapitulatif de la nomenclature des indices extra-forums de BP et SC

suivant des caractéristiques communes. Le nom des catégories est souvent une version abrégée pour faciliter l'exploitation avec R et autres logiciels de textométrie :

— Le lemme auquel appartient le mot, noté « Lemme ».

— Pour SC, une catégorie supplémentaire notée « Lemme SC » est utilisée pour les mots gaéliques ou ceux qui n'avaient pas de parenté graphique ou lexicale avec l'anglais.

— La graphie alternative, notée « AS ».

— La graphie standard, notée « SS ».

— Le nombre d'occurrences de la graphie alternative sur BP ou SC, notée « ASWebsite ».

— Le nombre d'occurrences de la graphie standard sur BP ou SC, notée « SSWebsite ».

— Le nombre d'occurrences de la graphie alternative sur *Google*, notée « ASInternet ».

— Le nombre d'occurrences de la graphie standard sur *Google*, notée « SSInternet ».

— Le type de graphie alternative [SCO], [VERN]…), tel que délimité dans le chapitre 1, section **??**. Ce facteur est noté « code ».

— La catégorie thématique à laquelle appartient le mot (*grammaire*, *famille*, *animal* …), notée « Catégorie »

Chaque corpus a son traitement spécifique. Bien que les catégories existent pour tous les mots, seuls certains ont eu des recherches plus poussées, qui sont également écrites dans le corpus. Les mots posant problème pour la recherche sur *Google*, et dont la requête n'était pas simple ont une catégorie supplémentaire appelée « Query ». Le *Corpus of Global Web-Based English* (GLOWBE) a également servi pour recueillir les donnes quantitatives pour certains mots des deux forums.

Pour BP, le nombre des occurrences a été comparé à celui dans le *Corpus of Contemporary American English* (COCA), alors que pour SC, certaines comparaisons ont été menées avec le BNC. Le corpus SC contient également certaines transcriptions tirées de l'OSD, dans une catégorie intitulée « Transcription OSD ».

3.3.2 Éthique de la recherche

3.3.2.1 Questionnements concernant les ethnies

En France, il est interdit de faire des recherches scientifiques portant sur l'ethnicité de participants, alors qu'il est possible de faire des recherches en sélectionnant les participants par nationalité. La question de la discrimination par origine ethnique est bien plus présente pour l'AAV, car ce sociolecte est plus volontiers parlé par les populations mélanodermes.

Il convient également de ne pas confondre ce que les utilisateurs annoncent comme étant leur localisation, et leur nationalité. La sélection d'utilisateurs par nationalité est restrictive, car elle peut mener à écarter des locuteurs expatriés, ou qui maîtrisent le dialecte malgré leur non-appartenance à la communauté en soi, comme c'est le cas pour les locuteurs canadiens du scots. Le problème des locuteurs caucasiens de l'AAV (qui peuvent l'avoir appris de diverses manières) est une réalité américaine, cristallisée dans l'expression péjorative « wigger » (amalgame entre « nigger » et « white »). Les utilisateurs caucasiens de BP sont accueillis par le site au même titre que les membres d'autres ethnicités et il en est de même pour *Scotster*.

Il semble donc plus judicieux de filtrer les utilisateurs par le seul critère valable : celui de l'appartenance à la plateforme, et de la rédaction de contenu sur les communautés virtuelles. Beaucoup d'utilisateurs ne renseignant pas le champ « location » du profil (ou s'en servant pour donner des renseignements qui peuvent s'avérer faux), cette méthode permet de n'écarter personne, et est conforme avec la législation française concernant la recherche sur des sujets humains.

3.3.2.2 Anonymat des données

Le corpus qui est utilisé dans ces recherches, et présenté en annexe, a été anonymisé afin de garantir la confidentialité des propos des différents intervenants de chaque communauté virtuelle, suivant les directives de la Commission Nationale de l'Informatique et des Libertés. Ainsi, aucune autorisation n'est nécessaire, puisque les données exploitées ne permettent pas d'identifier précisément les parti-

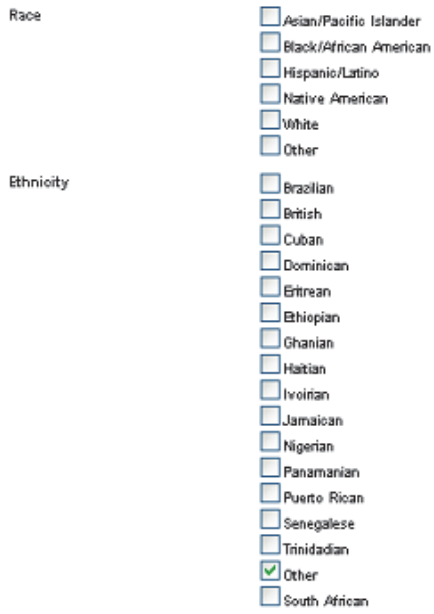


Illustration 3.23 – La page de choix de « race » et « ethnie » sur BP

cipants. L'accès aux ressources proposées par BP et SC est libre, les messages sont lisibles sans avoir besoin d'une inscription [29] et les utilisateurs de forums savent que tous les propos qu'ils peuvent tenir en ligne peuvent être retenus contre eux dans le cadre d'un procès par exemple. Les courriels et toutes les communications sur *Google* peuvent être cités lors d'un procès, comme indiqué dans les conditions d'utilisation des services. Les publications sur *Twitter* et FB sont filtrées et certains messages contenant des mots-clés incriminants sont conservés par le Département de la Sécurité Intérieure américain. L'affaire récente de la NSA et du type de données recueillies par les agences de sécurité intérieure ne sont que des illustrations de la porosité accrue entre vie privée et vie publique.

Certaines publications en ligne sont régulées, et la liberté d'expression américaine ne s'applique pas à tous les contenus présentés sur internet. Le *Communications Decency Act* (1996) protège les mineurs contre les contenus non protégés obscènes et pornographiques. La section 230 [30] est connue et souvent exploitée en procès pour défendre le fournisseur d'accès, et différentier sa responsabilité légale de l'hébergeur de contenu. Il n'existe pas de loi fédérale protégeant les données personnelles aux États-Unis, et les lois des états ne protègent que de la diffamation.

29. Le site BP a changé cette condition en mai 2014.
30. Code légal : 47 U.S.C. § 230

En Grande-Bretagne, diverses mesures ont été prises afin de réguler les messages sur internet (en particulier la pornographie et les contenus jugés comme choquants). Des « Lois d'Interception » de communications en vue de repérer les délits ont été passées en Grande-Bretagne en (1988 [31]), puis amendées et adoptées en l'Irlande (1993 [32]), à l'Île de Man (2001 [33]), en Écosse (1998 [34]). La loi sur la protection des données (*Data Protection Act* 1988), qui prévoit sensiblement la même protection de données que celle qu'offre la Commission Nationale de l'Informatique et des Libertés (CNIL), est complétée par la loi sur la communication (*Communication Act* 2003). Les données personnelles ne peuvent pas être exploitées à moins d'en avertir clairement les utilisateurs de services internet –entre autres– et ne doivent pas permettre l'identification précise des individus.

Les pseudonymes étudiés lors de la partie onomastique sur l'identité en ligne ne correspondent pas toujours aux pseudonymes des membres qui ont effectivement participé au forum, ou sont présentés dans le corpus. J'ai utilisé des graphies approchantes, et ai rebaptisé les utilisateurs en simplifiant leur pseudonyme. Un putatif « Lord_McKenzie » devient ainsi « McKenzie » ; « s0_niiiice_*life* » deviendrait « Life ».

3.3.3 Choix des fils de discussion étudiés

Au départ, les fils étudiés étaient soit choisis aléatoirement, soit choisis dans les sections ou sous-sections les plus fréquentées (jusqu'à plus de 3 000 réponses dans un seul fil de discussion). Ce choix était motivé par le fait qu'il semblait que les débats étaient plus intéressants dès lors qu'ils étaient entretenus par les autres membres du forum, et qu'il ne s'agissait donc pas d'interventions ponctuelles.

Plus un fil de discussion est long, et plus les membres participant s'égarent et quittent le sujet principal en s'entre-répondant. L'écueil principal, et certainement le plus craint peut être est le « point Godwin ». Énoncé par Mike Godwin, il concerne Usenet, une plateforme ancêtre des forums sur internet. Le « point Godwin »a été défini en 1990 : « Godwin's Rule of Nazi Analogies : As a Usenet discussion

31. Document de consultation disponible à l'adresse : <http ://www.cyber-rights.org/interception/ioca.pdf>
32. « Interception Act », version abrégée de « Interception of Postal Packets and Telecommunications Messages (Regulation) Act », dont le texte complet est disponible à l'adresse : <http ://www.irishstatutebook.ie/1993/en/act/pub/0010/index.html>
33. Texte intégral disponible à l'adresse : <http ://www.gov.im/lib/docs/infocentre/acts/ica2001.pdf>
34. Transfert des responsabilités aux Ministres d'Écosse.

grows longer, the probability of a comparison involving Nazis or Hitler approaches one. » (‘*Règle de Godwin des analogies avec les Nazis : plus une discussion Usenet est longue, plus la probabilité d'une comparaison avec les Nazis ou Hitler s'approche de un*’). C'est en fait un prolongement moderne de ce que l'historien Leo Strauss appela la « reductio ad hitlerum » dès 1953 dans son livre *Natural Right and History*(Strauss, 1953). Sur les deux forums, il y a en fait très peu de références aux *reductio ad Hitlerum* et d'occurrences de l'expression « point Godwin » : SC contient 25 références à l'expression, et il n'y en a que deux sur BP au 5 août 2014. Cet écueil n'est donc pas à craindre sur les deux forums étudiés.

Les analyses de phénomènes ont été ascendantes et descendantes : pour recenser les graphies alternatives, les fils de discussion les plus longs ont été choisis comme base d'analyse, et toutes les graphies alternatives ont été recensées. C'est une approche montante, du forum jusqu'au corpus. En revanche, pour étayer et approfondir ces premiers relevés, j'ai cherché les formes graphiques proches. Par exemple, pour tester l'existence d'une forme <airm> pour <airms> (‘*arm(s)*’), déjà trouvé sur SC, j'ai eu recours à une recherche sur Google du type « « airm » site :www.scotster.com ». Ces deux mouvements se sont complétés, et ont permis l'exploitation de fils de discussion initialement retirés de l'étude. Si une forme existe dans un fil de discussion qui a moins de cinq messages, ce fil est rajouté à la liste des fils exploitables. Après cette délimitation du type de corpus étudié et des conditions d'exploitation choisies, ce chapitre va maintenant s'intéresser à la création d'une identité en ligne, et d'une identité de membre de plateforme virtuelle, autant pour les membres de BP et SC que pour le travail de recherche et du chercheur. Le paradoxe de l'observateur et son application sur une plateforme numérique sera le dernier point abordé dans ce chapitre.

3.4 Essai de circonscription de l'identité en ligne

Se créer une identité virtuelle est important puisque les données circulent librement sur internet, et qu'il est possible de retrouver presque n'importe qui grâce à son empreinte virtuelle. Les utilisateurs de forums, comme ceux d'autres communautés virtuelles, ont donc à cœur de pouvoir compartimenter

leur vie réelle et ce qu'ils peuvent dire sur internet, dont ils se servent, au moins pour les utilisateurs de forums, comme un dérivatif. L'illusion de la conservation de son anonymat en ligne est puissante, et c'est ce qui pousse la création de communautés virtuelles sur la consommation de drogues (Barratt, 2011). C'est aussi la raison pour laquelle des membres de BP se revendiquent comme faisant partie d'un gang [35].

L'identité en ligne est un reflet de l'identité hors-ligne : elle en reprend parfois certains éléments (âge, lieu de naissance…si l'utilisateur décide effectivement de partager ces informations !) ; les goûts personnels et les opinions sont partagés, dans la conversation comme dans l'écriture de la page personnelle. Un consultant spécialiste des mondes numériques, F. Cavazza, a créé un tableau des traces qui constituent l'identité numérique (Fred Cavazza, 2006a), et qui est plus complet que ce qui est abordé dans cette section. Le tableau 3.24 recentre l'utilisateur au sein de toutes les traces qu'il laisse. Il note que « L'identité numérique d'un individu est composée de données formelles (coordonnées, certificats…) et informelles (commentaires, notes, billets, photos…). Toutes ces bribes d'information composent une identité numérique plus globale qui caractérise un individu, sa personnalité, son entourage et ses habitudes. ». Les traces formelle, à l'échelle d'une seule communauté virtuelle, seront au centre de cette partie.

La nature quelque peu anachronique de ce travail de thèse, consacré aux communautés virtuelles

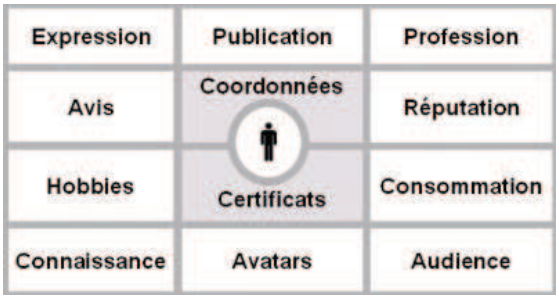


Illustration 3.24 – Panorama de l'identité numérique (Fred Cavazza, 2006b)

anciennes est bien plus microscopique que l'étendue des recherches modernes en CMO sur les nouvelles communautés, telles qu'elles sont vues par les chercheurs des années 2010 (Fred Cavazza, 2006a ; Merzeau, 2012 ; Deseilligny, 2012) sont trop complètes et ne correspondent pas à la façon

35. Pour reprendre une phrase mémétique reprise par la culture internet, « On the internet, nobody knows you're a dog. »

dont l'identité sur un forum est créée. C'est pourquoi je me suis appuyée sur des travaux plus anciens, sur l'identité sur les plateformes Usenet, comme les travaux de Donath sur l'élaboration d'une identité virtuelle (Donath, 1995) et du rôle d'expert, ainsi que sur des travaux de philosophie du langage (Kripke, 1980 ; Descombes, 2013). S'attribuer un nom et une identité en ligne est une étape primordiale.

Cette partie s'intéressera d'abord aux données socio-démographiques de membres de BP s'autoproclamant membres de gangs et d'utilisateurs de SC qui s'intéressent à la généalogie. Leur ville, leur âge et leur genre sont des éléments qu'il est possible de renseigner, mais les deux forums fournissent des informations peu fiables. Je ferai une courte étude de l'avatar pictural en tant que marqueur de l'identité, mais ce paramètre identitaire ne rentre pas en compte dans la création de graphies alternatives. Les pseudonymes seront étudiés dans la partie suivante [36]. Ensuite, la présentation personnelle, réelle vitrine de l'identité personnelle, sera abordée [37], tant dans ses caractéristiques fonctionnelles que dans ses propriétés discursives. La page personnelle est souvent l'espace où exprimer le plus volontiers son identité et à l'assigner par des graphies alternatives.

3.4.1 Fragilité des données socio-démographiques

Le lieu et l'âge renseignés par l'internaute sont souvent réels. Les localisations ne sont pas obligatoires, pas plus que de donner son âge (bien que la plupart des communautés interdisent l'accès des moins de 13 ans). Les lieux fantaisistes se repèrent immédiatement : lieux fantasmés, ou fictifs (dans le monde du Seigneur des Anneaux, par exemple). Les lieux fictifs sont impossibles à utiliser sur BP puisque les utilisateurs remplissent leur fiche de renseignements en indiquant leur pays dans un menu déroulant, puis leur code postal. Le site se charge ensuite de localiser l'utilisateur, et les données de localisation ne sont indiquées que si l'utilisateur le désire.

Sur SC, où l'accent est mis sur la fierté d'être écossais, les utilisateurs ont également l'option de localisation dans un menu déroulant, mais peuvent être plus précis en entrant le nom de la ville dans

36. Cette étude s'appuie sur un article en cours de resoumission chez l'African American Review (Goudet & Ballier, 2014).
37. Cette étude s'appuie sur un article concernant les membres de BP (Goudet, 2014).

laquelle ils se trouvent ou la zone ou banlieue de cette ville (deux cases « town/city » et « district/suburb », respectivement).

Les utilisateurs complètent leur profil avec des informations qui semblent souvent exactes, afin de montrer qu'ils veulent participer à la communauté, et les données personnelles qu'ils partagent sont multiples.

Pour BP, j'ai sélectionné un sous-corpus utilisé pour explorer la parole de gang qui contient les données socio-démographiques de 55 utilisateurs, sélectionnés parce qu'ils ont explicitement cité un gang (« Bloods » ou « Crip ») dans leur pseudonyme ou dans leur description personnelle. Le lieu renseigné par les utilisateurs correspond au moment de leur inscription, qui peut parfois avoir eu lieu il y a plus de 10 ans, sans remise à jour. Ceci est visible dans le type d'interface de la page personnelle, et dans le codage de l'URL utilisée.

Comme on peut le constater en regardant la carte 3.25, que j'ai créée sous *Google Maps* après un recueil de données sur 55 profils personnels, tous ces membres sont aux États-Unis. La majorité des membres se situent sur les côtes, surtout sur la côte est et la Californie, et le Texas représente une seconde zone de concentration. Ces membres ont des pages créées en moyenne en 2004, et leur moyenne d'âge est 24 ans en 2014. Le plus vieux a 37 ans, et son profil contient très peu de graphies alternatives : c'est un artiste qui a créé un profil pour promouvoir son disque, et la plus jeune a 14 ans au moment de l'écriture de son profil personnel.

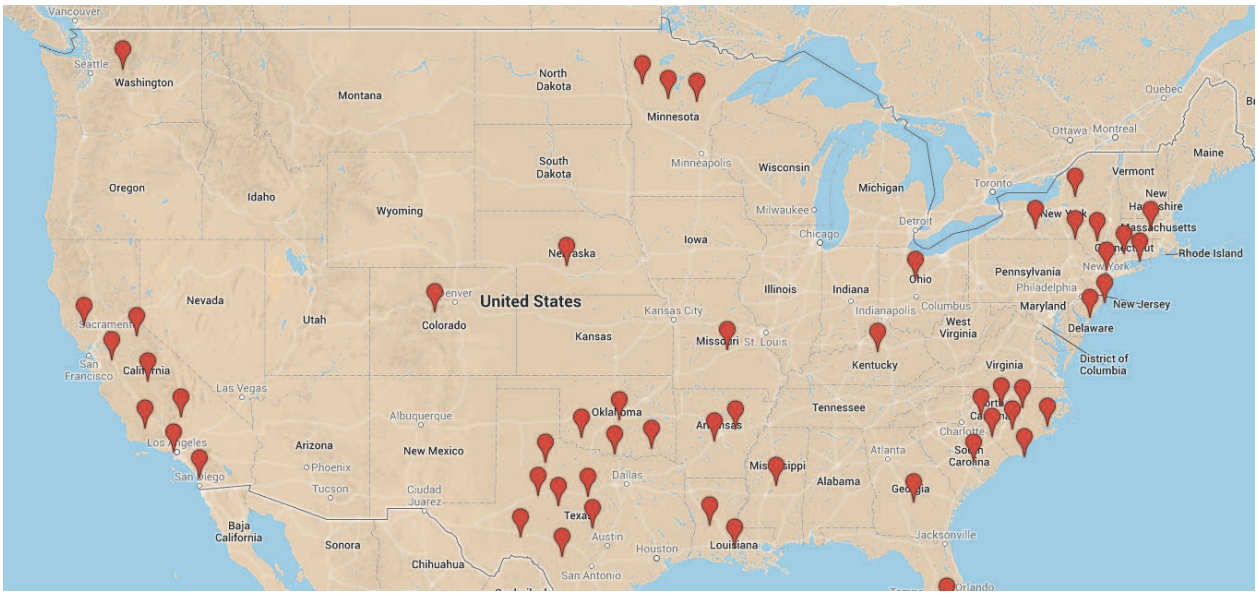


Illustration 3.25 – Répartition de 55 membres de gangs autoproclamés (BP)

L'analyse de ce sous-corpus est lacunaire, parce que ces membres ne semblent pas avoir remis leur page à jour. Certains sont membres depuis 2001, et les données curatées par le site peuvent être erronées. Ces utilisateurs disant faire partie d'un gang, mais ayant en moyenne 14 ans au moment de la création de leur page personnelle sont peut-être des reliquats d'une période plus ancienne de BP. Les gangs sont présent sur internet depuis les années 1990, et les groupes d'investigation de gangs comme le *National Gang Crime Research Center* s'intéressent à ces pages de près (Knox, 2011) depuis cette période. En revanche, cette captation diachronique nous permet de constater que les graphies alternatives les plus créatives se trouvent dans ces profils personnels, étudiés plus loin en 3.4.4. Ces membres ont recours à des graphies alternatives plus fréquentes, surtout dans leur profil personnel. Il n'est pas possible d'accéder aux messages qu'ils ont écrit sur le forum en partant de leur page personnelle.

Pour SC, j'ai sélectionné les utilisateurs ayant participé au fil de discussion SC2V13, qui parle de généalogie écossaise, et le fil SC2N59, où les utilisateurs discutent et échangent des mots en scots. Ces deux fils de discussion cumulent 42 membres dont le profil existe encore sur SC, et qui ont renseigné leur lieu de résidence, leur âge et leur genre, pour la plupart. Vérifier leurs données personnelles est étrangement plus difficile sur ce site : il faut avoir un compte utilisateur pour voir l'âge des membres,

et c'est la seule information ayant un tel prérequis.

La remise à jour des informations n'est encore une fois pas très fiable. Un utilisateur a renseigné son âge dans sa description personnelle « 31 year old man with a deep love of our wee Scotland » (Scott, participant à SC2N59), mais sa fiche remise à jour par SC indique 36 ans. Ceci va vers l'hypothèse de données peu fiables pour une analyse synchronique, mais qui peuvent être utiles pour examiner le profil moyen des utilisateurs en légère diachronie.

Les membres ayant participé au fil de discussion sur la généalogie écossaise (illustration 3.26) sont plus fréquemment Nord-Américains que Britanniques. La présence d'Australiens et d'un utilisateur Néo-Zélandais illustre la répartition mondiale de ces membres, bien que toujours circonscrite au monde anglophone. Les participants se sentent liés à l'Écosse par leurs origines, et c'est la raison pour laquelle on trouve moins d'Écossais que dans le second fil de conversation, SC2N59. Dans ce dernier, les membres discutent de la perte du scots et partagent des mots ensemble. C'est pourquoi ils sont presque tous situés en Écosse (j'ai éliminé de cette carte trois membres Américains pour avoir une vision plus nette de l'Écosse). Les membres se concentrent surtout autour de la ceinture centrale Glasgow-Édimbourg, et l'on peut imaginer que le scots qu'ils parlent est donc la variété centrale, le patter de Glasgow et l'argot de Leith déjà évoqués dans ce travail (en 2.2.4.2 et 1.3.1.1.3 respectivement). Ce fil de discussion concerne plutôt les locuteurs de scots, et c'est pourquoi la proportion d'Écossais est importante.

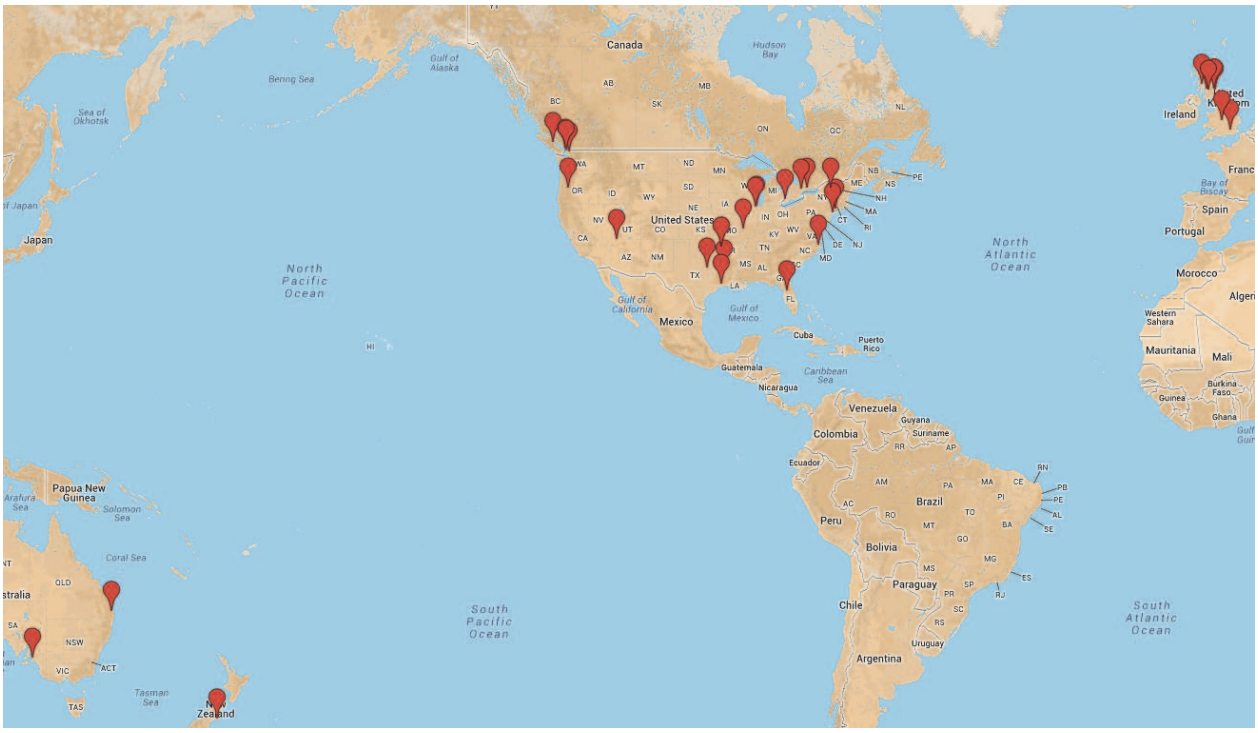


Illustration 3.26 – Répartition géographique des participants d'un fil sur la généalogie écossaise (SC2V13)

Illustration 3.27 – Répartition géographique des participants d'un fil sur les langues écossaises (SC2N59)

L'âge moyen des utilisateurs de SC semble plus élevé : le fil sur la généalogie rassemble 22 utilisateurs qui ont en moyenne 57 ans, et la plus jeune a 22 ans tandis que le plus âgé a 81 ans en 2014, selon les données de SC. Sur le fil linguistique, les utilisateurs ont en moyenne 53 ans, et la fourchette

de valeurs est un peu plus restreinte, entre 25 ans pour le benjamin et 75 ans pour le plus vieux. Ces données dressent un portrait tout à fait différent de cette communauté, dont les membres plus âgés n'utilisent pas des graphies alternatives néologisantes, mais des graphies historiques, en scots, comme « syne » (‘*next*’. La localisation et l'âge renseignés semblent logiques avec les propos tenus sur ces deux fils de discussion.

Circonscrire le public des deux communautés grâce à leurs intérêts est tout à fait possible, et ceci permettrait également de mener des études socio-démographiques sur les utilisations de graphies alternatives. C'est l'une des perspectives ouvertes par cette thèse, qui n'explorera pas cette piste, pour autant. Le recueil de données devrait être nettement amélioré, et l'étude devrait être faite hors-sol, ce qui échappe tout à fait à la perspective de ce travail.

3.4.2 Description fonctionnelle de l'avatar

L'avatar est la seconde façon principale de s'identifier sur internet. Il s'agit d'une vignette, le plus souvent carrée, qui est accolée à l'identifiant et sert d'illustration de l'utilisateur. Il la choisit soit dans une base de données fournies par le créateur du forum, soit en rajoutant une image de son propre ordinateur.

L'exemple le plus marquant de l'utilisation (ou de la non-utilisation d'un avatar !) se trouve dans le forum SA (Kyanka, 1999). Changer son avatar est payant, et ceux qui n'ont pas investi l'argent nécessaire se retrouvent avec une photographie représentant un bébé, légendé « Stupid Newbie ». Avoir un avatar, c'est avoir une identité virtuelle à part entière. Par conséquent, rares sont les utilisateurs de SA qui publient des messages fréquemment et n'ont pas d'avatar. Cette stratégie entre cynisme et marketing pousse à avoir un avatar, et montre à quel point incarner son identité sur internet et être un membre respecté d'une communauté passe par l'utilisation d'un avatar personnel. Si cette condition n'était pas si importante, un tel moyen ne serait pas mis en place par cette communauté.

Les utilisateurs de SC ne sont pas pénalisés lorsqu'ils n'ajoutent pas d'avatar, la case est simple-

ment omise du bandeau de présentation, et le pseudonyme remplace la photographie de profil, comme le montre les deux captures en 3.28, extraite des deux profils SCP11 et SCP2. Le profil ayant une image (SCP2) représente effectivement le membre inscrit, avec une photographie personnelle de sa famille. C'est aussi le biais utilisé par SC, qui encourage le partage et la navigation de photographies de membres (rubrique « Explore photos » dans le menu à gauche).

Les utilisateurs de BP se voient attribuer un avatar presque entièrement blanc représentant le haut du buste d'un pictogramme. Cela montre que l'avatar est censé jouer le rôle de photographie du membre, et en permettre une identification visuelle immédiate anthropomorphique. L'avatar est une extension de soi-même, de son identité.

Illustration 3.28 – Comparaison de profil avec et sans photo (SC)

Illustration 3.29 – Avatar neutre de membre sans photo (BP)

3.4.3 Caractérisation des pseudonymes

Les pseudonymes sont utilisés comme des prénoms sur les communautés virtuelles. Ils peuvent être composés de différents éléments : prénoms, lieux, thèmes ou intérêts personnels, et sont d'ordinaire uniques sur chaque communauté. Le pseudonyme est utilisé comme identifiant de connexion sur les sites internet, à moins que la plateforme n'autorise les messages d'utilisateurs invités, il est donc presque obligatoirement incontournable sur internet. L'accent est d'ailleurs souvent mis sur l'identification, et certains sites tournent en ridicule les utilisateurs anonymes.

3.4.3.1 Importance du pseudonyme

Je vais prendre *Slashdot* en exemple. C'est un grand agrégateur de nouvelles et d'informations spécialisées repostées par les internautes qui utilise un système de commentaire des articles intercommunautaire (Malda & Bates, 1997). Les commentaires des internautes se répondent en s'enchâssant les uns les autres, mais l'identification n'est pas necessaire. L'utilisateur anonyme se retrouve donc appelé « Anonymous Coward » (‘*lâche anonyme*’) par défaut, ce qui enjoint à la création d'un compte pour pouvoir répondre aux nouvelles. L'illustration 3.30 montre les diverses façons d'identifier un utilisateur de *Slashdot* : le premier message provient d'un utilisateur entièrement anonyme, le second

d'un utilisateur qui est identifié à la fois par son pseudonyme, et son adresse de courrier électronique, tandis que le troisième n'est identifié que par son pseudonyme. Il est intéressant de noter que chaque utilisateur enregistré se voit attribuer un nombre, qui correspond à son rang basé sur l'ancienneté sur la plateforme.

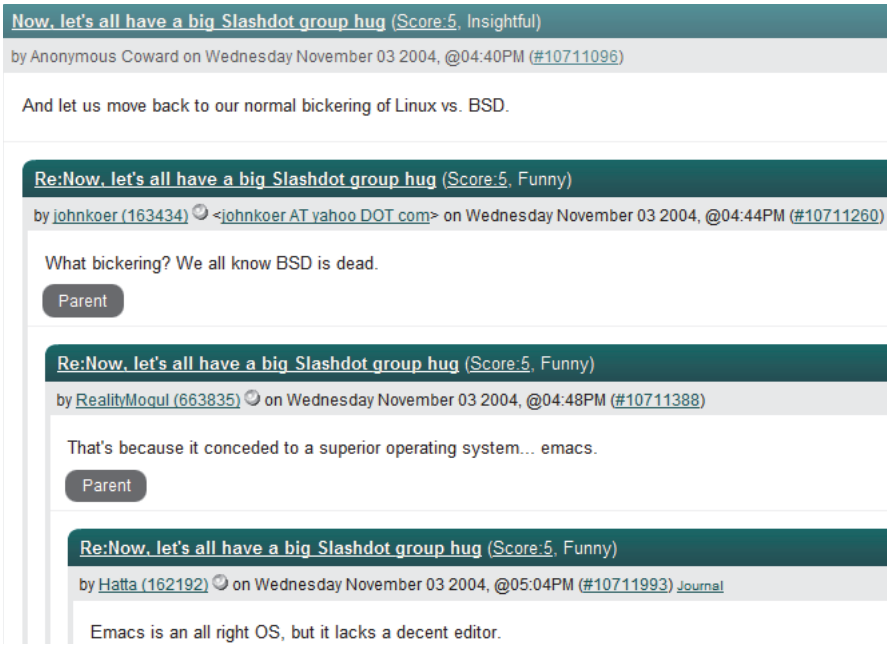


Illustration 3.30 – Commentaires enchâssés sur Slashdot

Le pseudonyme est donc le pilier garant de l'identité personnelle sur internet. Il est d'ailleurs fréquent que le même utilisateur ne change pas son pseudonyme de site en site, par des soucis pratiques autant qu'identitaires : même si la plupart des navigateurs internet peut retenir les mots de passe, ou enregistrer et conserver les cookies d'identification de chaque site, il est plus aisé de se servir du même pseudonyme sur internet [38].

38. Ainsi, mon navigateur internet révèle que, sur la cinquantaine d'informations confidentielles identifiant/pseudonymes pour une cinquantaine de sites, seuls deux pseudonymes sont employés : un pseudonyme professionnel, et un pseudonyme personnel.

3.4.3.2 Pseudonymes sur BP et SC

Le pseudonyme sur BP a un fonctionnement conventionnel, et est une information obligatoirement mentionnée sur les messages. Ainsi, l'illustration 3.31 montre trois cadres de données sociodémographiques, le premier appartenant à un membre dont le compte a été fermé sur BP. Les deux autres cadres contiennent un lien hypertexte qui renvoie vers la page personnelle de l'auteur, mais le mot « CLOSED » réfère à toute l'identité non-discursive du membre : les messages sur le forum sont les seuls éléments restants de ces membres fantômes.

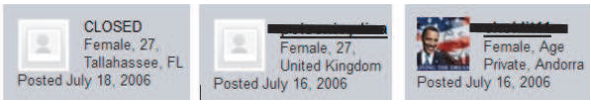


Illustration 3.31 – Message d'un compte fermé, d'un compte ouvert sans avatar et d'un compte ouvert avec avatar (BP)

Sur SC, le pseudonyme est traité différemment. Le site utilise un système hybride entre le vrai nom et le pseudonyme, et demande ces informations à l'inscription, un peu à la manière de FB [39]. Sur SC, les messages contiennent à la fois le pseudonyme et le nom réel du membre, sous la forme prénom et initiale du patronyme. L'illustration 3.32 est un exemple de la mise en forme des renseignements personnels sur les utilisateurs. Ici, le vrai nom est privilégié par rapport au pseudonyme : c'est lui qui porte le lien hypertexte qui renvoie vers le profil personnel. Le pseudonyme est écrit dans une police légèrement plus petite, en noir (alors que le nom de l'utilisateur est en bleu, ce qui indique ce lien vers une autre page). Comme le montre le troisième exemple, le membre a fermé son compte sur SC. C'est indiqué par la mention « Ex-member » après le nom et « deleted » avant le pseudonyme choisi.

39. La politique d'identification de FB est claire à ce sujet : « Facebook is a community where people use their real identities. We require everyone to provide their real names, so you always know who you're connecting with. This helps keep our community safe. ». (Zuckerberg, Saverin, McCollum, Moskovitz & Hughes, 2004)

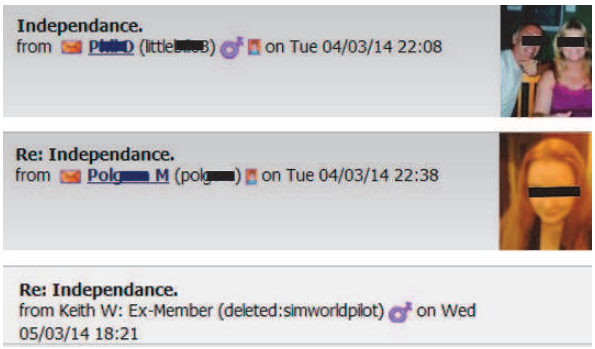


Illustration 3.32 – Trois exemples de renseignements personnels préliminaires à un message (SC)

3.4.3.3 'Etude de cas de deux matrices de pseudonymes sur BP

Certains utilisateurs choisissent des pseudonymes qui frappent par leur force, et j'ai choisi d'illustrer ce propos en traitant le cas des pseudonymes en *bitch* et en *nigga* sur BP qui sont à la fois des insultes et des mots repris de façon positive par les Afro-Américains. *Bitch* est repris par les femmes en général [40]. Il suffit de voir le cas de la phrase « it's Britney, bitch » utilisé par la chanteuse en 2007 dans son single « Gimme More » (Spears, 2007), devenu son surnom – qu'elle reprend dans son duo avec Will.i.am, « Scream and Shout » (Will.i.am et al., 2013). Quant à *nigga*, il est valorisé dans la culture hip-hop en particulier. Jay-Z l'emploie dans « Jigga my nigga », concaténation de son prénom et de « nigga » (« Jigga my n*** [nigga] », 1999).

Dans cette partie, je n'aborderai que les allographies employées pour composer un pseudonyme contenant l'un ou l'autre des mots, et ce travail lexical est une esquisse préliminaire au travail entrepris sur les variations allographiques en 5.5. Les mots *bitch* et *nigger* sont censurés sur BP, mais les utilisateurs veulent tout de même s'en servir pour composer des pseudonymes. Ils ont donc recours à des formes graphiques différentes, qui ont pour effet de modifier le sens du mot, au moins légèrement. En termes frégéens, cette modification du *Sinn* induit une modification du *Bedeutung* (Frege, 1892).

La réécriture de *bitch* en *beyotch* marque un changement d'investissement sémantique : *beyotch* ou *biatch* est souvent plus euphémisant que *bitch* (source : UD). Cette euphémisation passe par l'adjonction d'une voyelle épenthétique et par resyllabification. La graphie <bytch> de *bitch* est un excellent

40. La grande majorité de cette étude provient d'un article non publié (Goudet & Ballier, 2014).

exemple de substitution de la voyelle <i> par la voyelle <y> et constitue une variante graphique possible, identique du point de vue graphophonématique, mais aux connotations genrées fortes. Elle se rapproche d'une graphie alternative de *women*, <wimyn> ou <wymyn> qui fait intervenir l'<y> féminisant, tout en se détachant de <-man> (Ballier, 1997). La faible proportion de *nigga* écrit <nygga> corrobore cette hypothèse.

La réélaboration sémantique de *nigga* dépasse les frontières du forum, et cette graphie est amplement préférée sur internet (plus de 70 millions de résultats pour la recherche « nigga » sur *Google*, contre 14,5 millions pour « nigger »). Pour ce mot plus que pour *bitch*, les substitutions graphiques peuvent concerner les consonnes, comme <nigga> → <ni99a>, par ressemblance visuelle entre <g> et <9>. Le mot est tellement stable et connu que les consonnes peuvent en être changées. Ceci va à l'encontre de la tachygraphie ou des substitutions graphiques habituellement rencontrées sur internet, où les voyelles sont sensibles aux changements (Goudet, 2011).

Pour cette étude, j'ai sélectionné cinq graphies générales de *bitch* et *nigga*, et j'ai extrait les données de 1967 pseudonymes contenant l'une des graphies. Les résultats par graphies sont présentés dans le tableau 3.9. Le genre de chaque utilisateur sera représenté par la lettre <H> pour les hommes, <F> pour les femmes, et tous ces pseudonymes sont réécrits suivant le protocole dressé en 3.3.2.2.

Ces pseudonymes sont employés par des hommes ou des femmes, et les résultats quantitatifs sont

Bitch		Nigga	
<Bytch>	387	<Nigga>	894
<Btch>	72	<Ni99a>	475
<Biitch>	63	<Nygga>	32
<Biatch>	34	<Nigette>	6
<Beyotch>	2	<Nigress>	2
total : 558		total : 1409	

Tab. 3.9 – Répartition des pseudonymes en bitch et nigga par graphie

présentés dans les graphiques 3.33 et 3.34. Les pseudonymes peuvent être répartis en trois catégories : soit ils décrivent leur auteur (« Bitch Hater », H), soit ils sont composés de vocatifs (« Bitch I dare you », H), ou d'impératifs (« Nigga be my baby wipe », F), qui manifestent des jeux interlocutifs en miroir de ces deux insultes. Ils peuvent déjouer leur rôle attendu, la caractérisation identitaire

personnelle, car le vocatif singulier a vocation à interpeler l'ensemble des représentants du genre du pseudonyme. Les cas ne sont pas mutuellement exclusifs, mais des pseudonymes comme « nigga lover be my nigga » (F) sont doubles, dans la mesure où ils définissent à la fois la scripteuse en « nigga lover » et son coénonciateur, « be my nigga ». Ce pseudonyme complexe est un acte de parole interpellatif qui assigne les relations de genre dans un rapport hétérosexuel[41].

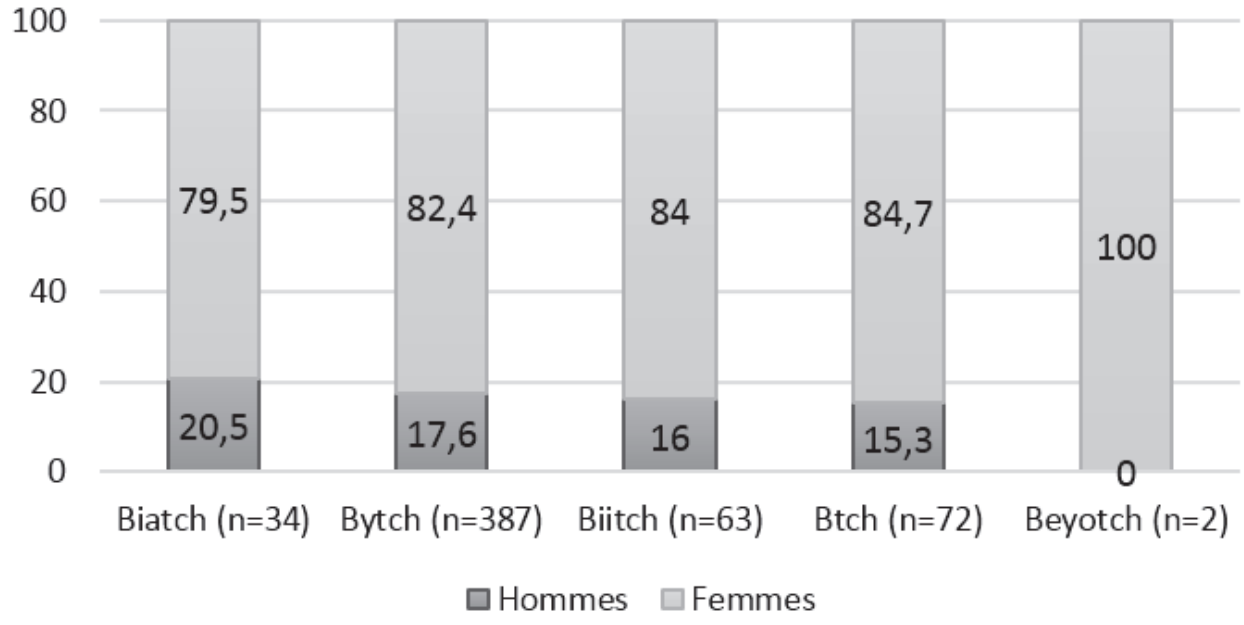


Illustration 3.33 – Distribution par genre des pseudonymes utilisant bitch

Les formes utilisées les plus fréquemment sont <bytch> et <btch> pour le graphique 3.33 et <nigga> et <ni99a> pour le graphique 3.34. Elles n'ont pas la même distribution par genre. Les hommes vont proportionnellement favoriser des formes simples, des substitutions ou ajouts d'un seul caractère lorsqu'ils utilisent *bitch*.

Les femmes préfèrent opter soit pour des formes uniques, avec adjonction morphémique (<-ette> ou <-ess>) qui reste leur apanage, ou, rester dans des formes employées par les hommes : il y a 179 et 178 utilisatrices pour <nigga> et <ni99a> respectivement, contre 2 utilisatrices de <nigress>, et 6 de

41. L'orientation sexuelle des membres de BP et SC ne sera pas discutée dans ce travail, mais l'exploration de ces pseudonymes montre une écrasante majorité de membres hétérosexuels.

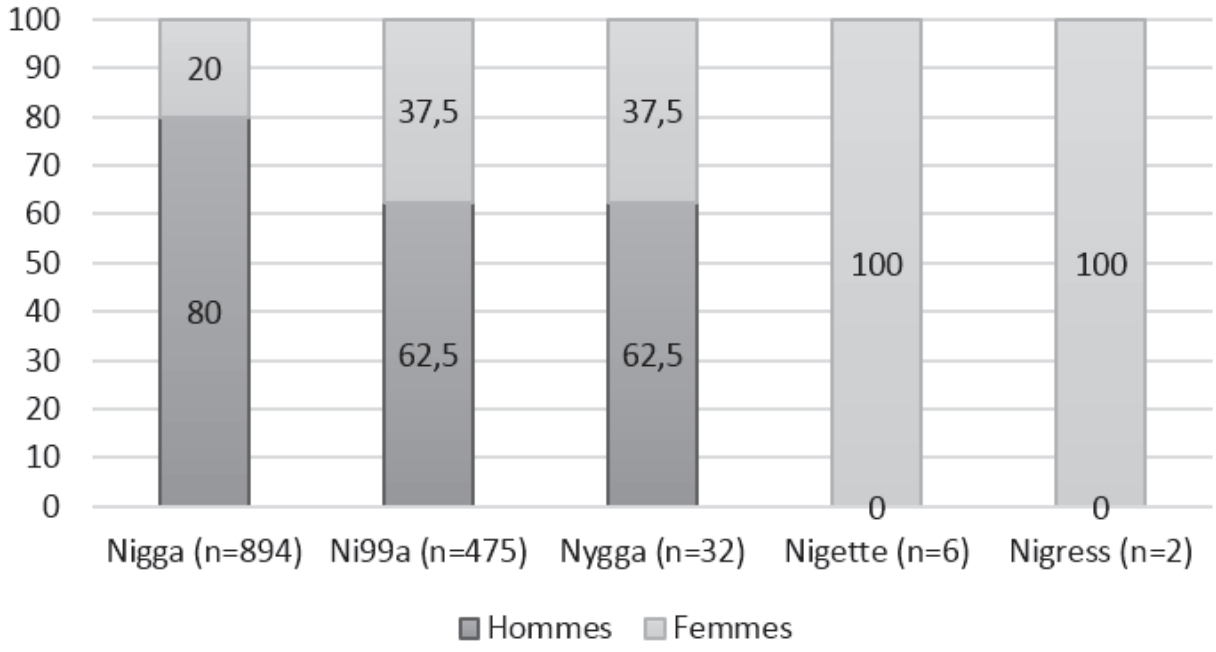


Illustration 3.34 – Distribution par genre des pseudonymes utilisant nigga

<nigette>.

L'utilisation genrée reste assez conventionnelle dans les deux cas : si les femmes ont tendance à utiliser plus de formes pseudonymiques que les hommes, et à épouser plus facilement un pseudonyme en *nigga*, toutes les formes de *bitch* sont plus souvent utilisées par les femmes, et vice versa pour *nigga*. Ces pseudonymes utilisant *bitch* ou *nigga* ne suivent pas forcément la répartition genrée attendue : Le marquage genré est absent de *nigger*, mais il représente souvent des hommes. Sur Google, une recherche sur la séquence [be-copule + nigger] renvoie des résultats disparates. « He's a nigger » enregistre 170 000 résultats, alors que « she's a nigger » n'en représente que 29 100, si on exclut l'appellation « nigger lover » (utilisé pour parler des femmes blanches attirées par les noirs)[42]. *Nigger* devient 'homme', un *nigga* (qui s'émancipe donc, grâce à l'investissement positif de la récupération du terme), voire un homme débrouillard, charismatique et dangereux, alors que le mot *bitch* voit son sens glisser vers 'femme noire', avec les qualités qu'on leur impute (la *sassiness* ou la façon impertinente de savoir faire valoir son bon droit, en protestant ou en n'acceptant pas ce qu'on leur impose).

Nigga/er n'est donc que masculin chez les Afro-Américains. Ceci pousse à l'étude du système lin-

42. Recherche effectuée le 25 novembre 2013.

guistique où tout fait genre. Un peu à la manière du Politically Correct où le <–man> de *chairman* est réanalysé en marque de genre masculin, le système linguistique se trouve réinvesti. Le non-marqué devient marqué dans une forme d'omnimarquage généralisé (Ballier, 1997).

L'identité « niggette » ou « nigga » féminine est plus assumée chez celles qui sont proches des milieux de gangs. En tant que membres d'une communauté qui utilise « nigga » comme terme valorisant pour les hommes, elles empruntent le terme pour montrer qu'elles en font également partie. Cette approche est loin d'un mouvement pour se dissocier de la *bitch* ou pour prendre le pouvoir dans le discours, par rapport aux hommes. Les utilisatrices de *bitch* se conforment à ce que les hommes qui les désignent ainsi attendent d'elles. La différence entre *nigette* et *bitch* se situe dans la façon dont la couleur de la peau est suggérée. *Nigette* rapproche la femme des *niggas*, donc des Afro-Américains les plus typés, et joue donc sur l'appartenance à un groupe, par la féminisation du mot. *Bitch* ne réfère à la femme noire qu'indirectement ; il est préféré à *nigette* sans doute en raison de l'importance des fantasmes leucophiles (attraction envers les femmes noires dont la couleur de peau se rapproche le plus possible de celui des blanches) héritées de la colonisation des Blancs (Richardson, 2006).

Le système bloque la femme dans ses représentations, en l'empêchant de s'auto-attribuer une appellation récupérée, parallèle à *nigga*. *Bitch* est simplement le mot utilisé par les hommes pour parler des femmes, et elles se contentent de le réutiliser tel quel. Il n'a pas la même dimension d'encapacitation linguistique que *nigger*, qui représente la repossession et l'oblitération au moins partielle du stigmate. *Bitch* apporte le couple genré *dawg*/*bitch*, malicieusement souligné par un mème internet.

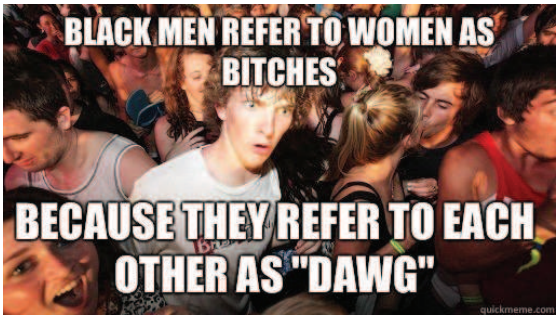


Illustration 3.35 – Mème « Sudden Clarity Clarence » illustrant le couple dawg/bitch

Ces utilisations montrent la perception et l'utilisation de ces deux lexèmes par des Afro-Américains : pour eux, il s'agit de reconstruire son identité à partir du détournement de l'insulte dont ils ont été (et sont toujours) affublés. La sexualisation de l'identité permettant une sortie de l'insulte, et sa forte ritualisation dans un spectacle qui assure la félicité du contre-transfert entre coénonciateur et auteur.e de ces pseudonymes. Les insultes sont rendues dérisoires grâce à la récupération des termes, qui déjoue et inverse les rapports de force vis-à-vis d'elles : un *nigga* s'émancipe de la domination blanche, et une *bitch* déjoue l'insulte en affirmant la force de son caractère.

Cette étude de cas détaillée est une étape nécessaire pour comprendre l'étude ultérieure des variantes allographiques de « nigga » et de « bitch », qui interviendra dans le dernier chapitre (partie 5.5). L'appropriation genrée de l'identité fait partie intégrante de ces jeux graphiques, qui concernent des mots dont l'orthographe est censurée sur BP.

3.4.4 Présentation personnelle et monologues

Les présentations personnelles, qui sont proposées sur tous les forums, sont une façon pour les autres utilisateurs de découvrir plus sur les membres (ce qui nécessite d'aller sur leur profil personnel). Les profils offrent un espace de liberté pour le partage et l'écriture de ce qui est censé représenter le membre, en ligne, donc l'originalité est de mise.

L'illustration 3.36 est tirée du profil d'une des membres de BP, et la traduction de ce texte en jamaïcain est « Contracte ton sexe sur mon corps comme ça/ Mon sexe te fait gémir / Alors saute (dessus) et ondule / Contracte ton sexe / et ondule de la taille / Mon sexe dans ton ventre et on s'amuse bien / Tu tournes j'adore comment / tu remues les fesses, ondules et bouges vite ». Ceci est la première section de sa description personnelle (et le seul extrait qu'elle ait écrit, son profil ne contient que des photographies ou des vidéos). Après le premier encadré, qui représente le titre, elle a copié les paroles d'une chanson, « Just Wine » de Jah Vinci. Les occurrences en ligne correspondent parfaitement,

Illustration 3.36 – Extrait de profil sur BP

graphiquement, à ce qui est copié [43]. L'originalité de ce profil tient dans le fait qu'il est mis en forme de façon tout à fait personnelle : l'utilisatrice a ainsi choisi de centrer le texte, et de changer la couleur de chaque vers de la chanson [44], une série d'opérations assez fastidieuse. Le résultat concentre l'originalité de son message, qui en aurait été privé si elle n'avait que copié les paroles de la chanson. La mise en forme graphique et les choix expriment la personnalité de ce membre (qui aurait pu choisir une alternance de deux couleurs seulement, par exemple) autant que l'aurait fait un texte personnel.

La seule partie du profil qui émane effectivement de cette utilisatrice est le titre, qu'elle a laissé en noir, « Godbless GAZA bunzz » [45]. Il y a trop peu de preuves en ligne pour pouvoir décider du statut de la graphie de ces éléments : on ne sait pas si *godbless* (absence d'espace entre *god* et *bless*) ou *bunzz* (doublement du <z> final) sont des fautes d'orthographe ou des graphies alternatives.

Les utilisateurs de SC sont souvent moins prolixes dans leur façon d'écrire sur leur profil, et cela est dû en partie à l'écologie numérique à leur disposition. Le profil d'un membre est toujours écrit dans la même police, aucun choix de mise en forme n'est possible. Les informations, comme la légende de la photographie de profil, le statut (dernière ligne du profil personnel), ou la rubrique « wanted/on offer » ne figurent sur les profils que si les membres s'en servent. Contrairement aux profils de jeunes membres

43. Au 31 mai 2011.
44. Les manipulations pour arriver à ce résultat sont : sélectionner le texte, ouvrir l'option des couleurs, choisir une couleur ou définir une couleur personnalisée puis valider. Ces changements doivent être réalisés ligne après ligne.
45. Ceci ne correspond pas au pseudonyme de l'utilisatrice.

de gangs de BP, le profil personnel n'est pas souvent un lieu pour utiliser des graphies alternatives. Sur 25 profils de membres actifs (qui ont posté une activité sur SC entre le 7 et le 8 août 2014), il n'y a que trois profils qui contiennent des scottismes. Ceux-ci sont souvent circonscrits à un seul mot : « Hi there was born and raised in Glasgow but now stay down in Ayrshire. Usually out and about at weekends visiting different parts of the country in my wee camper van. I am an avid Reader, learning Gaelic and like visiting historical sights in and around Scotland. Love going to music festivals with friends.. wanted : chat, friends, a **guid** laugh, honesty.. on offer : friendship, honesty, humour, a **wee blether** » (SCP4) ; « decided i **didny** want **tae** be grumpy **anymer** still retired,still love Loch Lomond and Scotland divorced but like it.. wanted : friendship.. on offer : honesty » (SCP13). Les scottismes sont circonscrits à des mots isolés, connus comme « wee » ou grammaticaux comme « tae ».

Illustration 3.37 – Profil d'un membre de SC (SCP12)

3.5 Identités causées

Être membre d'une communauté internet, c'est non seulement se représenter, mais aussi s'intégrer en tant que membre et se trouver une identité en tant que tel. Dans cette dernière partie, j'aborderai les marques linguistique de l'appartenance à une communauté virtuelle, en regardant d'abord les usages linguistiques créés par les fondateurs des sites étudiés. Cette identité et une courte partie méthodolo-

gique sur l'appartenance à une communauté virtuelle en tant que chercheur. Je traiterai également du paradoxe de l'observateur, tel qu'il est déplacé lors d'une étude sur un corpus numérique.

L'identité est performative dès lors que l'on indique que l'on fait partie d'une communauté virtuelle en utilisant des mots particuliers : l'on fait acte d'appartenance dans sa façon de s'intégrer à la population de la plateforme. Le cas particulier du chercheur sur la communauté virtuelle ainsi que le paradoxe de l'observateur sont aussi des circonstances particulières en tant que membre d'une communauté virtuelle. Le titre de cette partie provient de la théorie causale de la référence de Kripke, appliquée aux identités sur des communautés virtuelles (Kripke, 1980). L'identité personnelle sur SC et BP est créée par la connexion causale à soi-même et à la communauté, grâce à la communauté elle-même, qui perpétue ses propres codes. L'identité de la communauté par elle-même, celle du chercheur et de l'observation sont crées par son existence.

3.5.0.1 Néologismes des communautés virtuelles

Chaque communauté virtuelle a des mots particuliers pour désigner les éléments du forum, et pour le personnaliser à ses utilisateurs. Par exemple, sur BP, l'équivalent du mur de FB (le flux d'actualités composé par les textes, images, vidéos partagées par les « amis ») s'appelle « What's new ? » : c'est une page où l'on peut voir toutes les actualités du site, pas uniquement à l'échelle de ses amis, mais aussi à l'échelle du site. Ce fil d'actualité, présenté dans l'illustration 3.38 est aussi la page sur laquelle les membres arrivent après s'être authentifiés sur le site, comme sur FB. Sur SC, cette fonction s'appelle « The Pulse », et elle est présentée dans l'illustration 3.39. Cette section est légèrement plus complète que le « What's New ? », parce que les connexions de membres y figurent. Sur BP, il n'y a qu'un bandeau en haut de l'écran d'accueil qui indique les membres en ligne. Cette connexion s'accompagne de la mention du dernier statut personnel rédigé : « Simple man, simple life » par exemple (troisième mention sur 3.39). Ceci permet de présenter à nouveau le statut aux membres qui consultent le *Pulse*, et ainsi, d'obtenir de nouvelles réponses.

Le *Pulse* et *What's new ?* ont des utilisations tout à fait similaires. Ils donnent une identité à la plateforme même, en ne reprenant pas les noms utilisés par FB dont ils s'inspirent (‘*wall*’ou ‘*news feed*’).

L'unicité de l'identité de la plateforme est donc contenue dans les noms des types de communication, et non dans la référence, qui reste identique. Ces mécanismes de fil d'actualité mis à jour en temps réel sont aussi une façon de montrer la façon dont les utilisateurs investissent les lieux numériques. Le dynamisme inhérent à ces remises à jour fréquentes incitent à participer, en centralisant le contenu disponible et actualisé sur une seule page. Plus une plateforme est active, plus ses membres sont incités à y participer, parce que les contenus sont variés et peuvent potentiellement intéresser plus de gens. Le *Pulse* et *What's New ?* créent la communication, et le sentiment d'appartenir à une communauté vivante.

Les deux structures d'informations, qui donnent des renseignements quant aux actions des autres (et des siennes) permet la cohésion des membres du groupe. Voir leurs faits et gestes sur la plateforme permet de leur répondre (surtout dans le « What's new ? » de BP), donc d'initier le dialogue. Ces messages ressemblent à des petites annonces, ou à des messages de crieurs de places publiques. Certains demandent explicitement ou implicitement de communiquer avec eux, dans des petites annonces matrimoniales : « lookn for a single good looking woman […]hit me up with a note » (troisième message dans 3.38). Un de ces messages ressemble à une note de service adressée au site, et non aux membres : « Black Planet needs to do something about all the spam and porn on this site » (deuxième message dans 3.38), qui n'a pas de vrai coénonciateur. Néanmoins, cette critique pourrait provoquer des réponses d'utilisateurs. BP poste également de vrais messages de service, incitant par exemple à suivre l'activité de BP sur Twitter. Quant à la nature des échanges sur SC, elle est plus informative, et les réponses apportées à des fils de discussion pour permettre aux utilisateurs qui regardent le *Pulse* d'aller leur répondre. Le Pulse est plus ancré dans la vie du forum, et créée un pont avec les fils de discussion déjà crées : le dernier message de 3.39 montre que l'utilisateur est d'accord (l'équivalent de SC de « liker ») avec un des messages du forum.

Si *What's New ?* est une extension de la communication, et n'a pas de rapport avec le forum, puisqu'il n'y a pas de mentions de ce qui y est posté, le *Pulse* est un outil panopticonique virtuel qui permet de voir sur une seule page toutes les activités sur le site.

Illustration 3.38 – La page d'accueil « What's New ? » de BP au 28 juin 2013

La position de surplomb adoptée par le chercheur est facilitée par de tels outils. Elle est le sujet de la partie suivante, où une relecture adaptée aux nouveaux positionnements du chercheur par rapport à son objet d'étude permet de remettre à jour le paradoxe de l'observateur tel tel qu'il fut énoncé par Labov.

Illustration 3.39 – le « Pulse » de SC au 28 juin 2013

3.5.0.2 Paradoxe de l'observateur

Ce que Labov appelait « paradoxe de l'observateur » est tout à fait transposable à la recherche en ligne, avec des particularités dues à la plateforme : les échanges peuvent être étudiés comme si le chercheur n'était pas présent, donc sans incidence sur la qualité des échanges. Ce type de recueil de données permet l'observation systématique de données conservées au moins sur la plateforme, sinon dans des sauvegardes personnelles. Labov énonce ainsi le paradoxe : « the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed ; yet we can only obtain these data by systematic observation. » (Labov, 1973). C'est là le coeur du problème des études en sociolinguistique, et des recueils sur le terrain. L'effet Pygmalion (Rosenthal & Jacobson, 1968) s'en rapproche : par l'observation et l'attente d'un certain type de résultats, les

observateurs les obtiennent, dans une certaine mesure, et dévient l'observation neutre. Le vrai vernaculaire (« true vernacular ») de Labov n'était obtenu qu'en déviant l'attention des informants de leur production linguistique, soit en rendant la conversation informelle et détendue (« various devices which divert attention away from speech »), soit en leur posant des questions si chargées émotionnellement que leur attention se portait uniquement sur le contenu, et non sur leur production. C'est la question « danger de mort » (type « Avez-vous déjà été en danger de mort ? ») qui créée des tensions émotionnelles.

Ces deux solutions ne sont pas acceptables éthiquement, et peu efficaces, selon Labov lui-même. Il faut que les informateurs donnent des consentements éclairés pour que les données soient exploitées, et que leur bien-être mental et psychique ne soit pas affecté par les questions qui leur sont posées. L'observation de phénomènes langagiers sur internet échappe à ces deux solutions, mais crée un nouveau paradigme. Les internautes savent que toutes les données publiques qu'ils partagent sont accessibles à tous. Récemment, *Google* a permis le droit à l'oubli aux internautes désirant se débarrasser des résultats les concernant qui pouvait leur nuire, et plus de 70 000 cas ont été traités par le site après cette décision judiciaire (A. Preston, 2014). Les membres de forums savent que les autres peuvent accéder à leurs messages et à leur profil personnel, et, même à la clôture de leur compte sur SC et BP, leurs écrits subsistent sur les forums.

C'est une situation de *panopticon* de pratiques discursives unique pour le chercheur. Les données anonymes n'ont pas besoin de passer par l'étape du consentement éclairé pour être traitées, puisqu'elles sont disponibles–moyennant un compte d'utilisateur. C'est là où le paradoxe de l'observateur se déplace : le chercheur fait partie de la communauté, connaît les membres, lit leurs écrits. En m'inscrivant sur SC et BP, je devenais moi-même membre de ces deux plateformes, et me soumettais à ses règles. En tant qu'observatrice, je me devais de ne pas participer aux discussions, pour ne pas gonfler artificiellement le nombre de messages sur les forums, et avoir une existence hybride. J'ai tout de même décidé de répondre aux messages privés, qui ne sont pas du tout conservés ni dans le corpus ni dans l'étude. Les membres qui me voyaient passer sur leur page personnelle [46] savaient ainsi que je n'étais pas un *bot*, ou un escroc. Dans la partie suivante, 3.5.0.3, je vais décrire l'identité virtuelle que je me

46. C'est une rubrique obligatoire des profils sur les deux forums.

suis attribuée, et le positionnement adopté lors des prises de contact.

3.5.0.3 Positionnement du chercheur

Afin d'étudier internet et les forums, il est important d'assumer une identité virtuelle, pour avoir accès et participer aux forums, dans certains cas (poser des questions, se présenter). Un pseudonyme et une adresse de courrier électronique sont les deux prérequis indispensables dans ce cas.

Il était également important, pour des raisons éthiques, d'effectuer les relevés avec autant d'honnêteté que possible, et que ma présence soit connue. Pour m'inscrire sur BP et SC, j'ai donc décidé de ne pas me présenter dans un nouveau fil de discussion, mais de mettre un message explicatif sur ma page personnelle. J'ai choisi un pseudonyme composé d'un prénom, sans mention d'une passion ou d'un lieu, pour être la plus proche de mon identité hors-ligne.

J'ai rempli ma fiche personnelle avec honnêteté, et j'ai pris soin de remplir ma fiche de renseignement personnelle avec la nature exacte de ma recherche. Ainsi, le modèle de fiche remplie ainsi que le message de présentation est :

> Hello, I am a French PhD candidate, working on English linguistics. My dissertation is about African American Vernacular English and Scots, as they are used on internet forums.

Les premières réactions sur les forums sont négatives. Après m'être clairement présentée sur la partie du forum où il est obligatoire de dire quelques mots sur soi pour participer aux échanges, le site *Cocoa Lounge* a banni le compte créé pour la recherche, par peur d'attaques racistes [47], problème déjà abordé en 3.3.1.2.2. En conséquence, le monologue ne mentionne pas le fait que le compte a été créé afin d'étudier les échanges, bien que les participants soient prévenus de la démarche d'étude une fois le contact établi. Lors de discussions privées, je demandais ainsi si mon interlocutrice avait lu mon profil, et si elle pouvaient contribuer à ma recherche [48]. L'accueil est souvent négatif dans l'ensemble,

47. Lors de ces demandes, en avril 2009, le site venait d'essuyer des attaques racistes et avait dû fermer quelques temps, ce qui explique l'attitude stricte du forum, en phase de restructuration et de reprise de contact après la fermeture.

48. Je choisis le féminin pour parler de n'importe quel membre des forums, bien que j'aie été contactée par plus d'hommes que de femmes.

étant donné que les participants de BP notamment ne contactent les femmes connectées que dans une perspective de séduction, et deviennent agressifs ou fermés une fois que leurs attentes sont détrompées. Peu prennent la peine de lire les profils des femmes auxquelles ils parlent (ce qui a été vérifié à de maintes reprises, avec la question « Have you read my profile ? »). Cependant, certains envoient des messages privés pour entrer en contact directement avec moi « I would like to meet you to talk about the so called afro americans » mais ces messages n'ont pas débouché sur de vraies prises de contact.

C'est la raison pour laquelle les relevés ont été effectués en tout anonymat : les membres de BP ou de SC ne doivent pas être reconnaissables (ce qui est une gageure, en fait, lorsque l'on sait qu'il est possible de rentrer sous *Google* les contenus des messages copiés dans la présente thèse, et de retrouver ainsi les identités des membres). Ces questions éthiques ne trouvent pas de réponse, et je ne peux que demander de respecter l'anonymat des participants, et de ne pas chercher à les retrouver.

Le paradoxe de l'observateur numérique sur un forum l'inclut donc plus entièrement sur les plateformes observées, contrairement aux recueils de données sur *Twitter* qui peuvent être effectuées automatiquement. Subsiste un problème éthique : les données ne peuvent pas être rendues aux utilisateurs de BP et SC avant la fin de ce travail, et cette enquête de six années risque de créer des problèmes *a posteriori*. Le profil personnel, les discussions échangées sont peut-être considérées comme faisant partie du domaine du privé par les membres qui se reconnaîtraient. Me faire passer pour un membre et me déclarer comme chercheuse sur ces forums-ci de façon passive (par mon profil privé, et non dans un message actif dans le forum) est un aménagement de concession : sans me déclarer à tous, je ne risque plus le bannissement, mais je ne peux plus rendre les données aux internautes qui ont été observés pendant six ans.

Conclusion du chapitre

Internet offre des possibilités de communication, mais aussi et surtout des possibilités d'observation de faits langagiers inédits. Les deux gageures sont la facilité et la tentation de calquer des éléments discursifs observables sur d'autres médias (téléphone portable en tête, ou discussion en face à face)

sur la discussion sur internet. Le discours sur forums suit ses propres règles, prolongements de celles régissant les échanges sur *Usenet*. Le format de chaque forum, les liens entre les membres ou leur vitrine personnelle (leur profil) sépare les forums de cette ancienne plateforme. Les concepteurs de ces plateformes ont choisi de donner une identité visuelle, sémiotiquement forte, à leur création. Il ne s'agit pas de forums tout prêts écrits en phpBB, la langue informatique de référence pour créer un forum facilement. BP et SC ont su évoluer, se complexifier et s'aligner sur les réseaux sociaux. Les forums définissent les anciennes lignes sur lesquelles ces réseaux se sont développés, mais l'identité personnelle y est mise en avant.

Les forums sont en train de doucement disparaître, au profit de plateformes multimédias comme *Reddit* ou *Youtube*, dont les discussions suivent un autre type de patron. Il y est possible de voter pour un commentaire afin qu'il apparaisse plus haut après le message initial, et ainsi provoquer l'envie d'écrire un message qui sera populaire. Les réseaux sociaux sont plus axés sur la mise en relation de membres, que sur la communication en groupe, autour d'un même sujet. Les échanges sur SC et BP restent au centre de ces deux plateformes, et elles ont su incorporer les autres canaux de communication, comme le *Pulse* ou *What's New ?*. Ces services sont uniques aux deux communautés, et montrent l'investissement créatif de leurs fondateurs, qui font tout pour perpétuer l'activité sur ces sites.

La propriété commune à tous les types de communication sur internet est l'identité personnelle. Le pseudonyme ouvre la porte à la communication avec les autres membres. Paradoxalement, une fois que le compte, et donc l'identité, ont été supprimés, les messages postés restent sur les forums. Avoir un profil rempli n'est pas une garantie de participation à la plateforme virtuelle. Voici quelques unes des difficultés rencontrées lors de la constitution du corpus, et de la notion d'identité en ligne. Il n'y a aucune garantie que les données fournies par les utilisateurs soient exactes, ce qui fait passer le travail d'exploitation sociolinguistique derrière le travail d'analyse de discours. Les utilisateurs maîtrisent l'anglais et les langues minoritaires, et créent un communaulecte, qui sera le sujet du chapitre suivant. Ils ne cessent de changer leur position sur un axe entre la langue minoritaire qu'ils parlent et l'anglais standard, entre la monstration de soi narcissique, dans les profils personnels, la conformation aux normes alternatives, et le dialogue.

Les études sur des corpus internet ressemblent à celles sur les discours d'hommes politiques, à cause de l'accessibilité des données. Ce corpus internet est semi-public, beaucoup des requêtes ont été effectuées sur *Google* en cherchant certains mots-clef (et en étant, donc, tributaire des algorithmes du moteur de recherche pour les relevés quantitatifs). Ce chapitre montre à quel point cette thèse est inscrite dans un écosystème novateur, aux règles particulières qui ne sont pas réductibles à une CMO trop globalisante. C'est la raison pour laquelle la méthodologie, entre relevés hors-sol et exploitations de CEC est aussi novatrice, et fragile. Si ces forums de discussion venaient à être fermés, la majorité du corpus serait détruit, s'il n'est pas archivé dans des pages en cache de *Google*, ou sur des sites de curation virtuelle comme l'archive d'internet (Kahle, 1996). Cette thèse repose sur des analyses d'instantanés de l'état du forum, *hic et nunc*. Le danger de telles analyses est l'absence de stabilité, concernant le nombre de membres des forums, des données textométriques…Mais le forum est un espace mouvant, et ce choix de corpus reflète sa nature profonde. Les forums ferment chaque jour, les messages disparaissent ou sont réécrits, mais ceci fait aussi partie de la vie communautaire virtuelle. Les utilisateurs sont responsables de leurs messages, tout comme les modérateurs ou fondateurs du site, qui peuvent décider de censurer ou supprimer un message. Une attaque de pirates informatique est toujours possible, et l'exemple de *Cocoa Loung*el'illustre tristement. Le point fort de SC et BP est d'être restés relativement stables. Les sites de niche ethno-culturels se répandent de plus en plus sur internet, mais la force de l'AAV et du scots assure aussi une grande stabilité aux forums. Le contre-exemple est le site « Speak LOLspeak », qui rassemblait des membres passionnés des LOLcats [49]. Ce site n'a pas trouvé de public stable, et il a fermé au bout de deux ans d'activité, et le nom de domaine du site est retourné en vente, ce qui veut dire que la plateforme n'existe plus sur internet. Dans le cas des sites qui constituent le corpus, l'identité linguistique et culturelle semble assurer une certaine pérennité.

Cette identité linguistique, faite d'éléments en et hors-ligne sera le sujet du chapitre suivant : les communications sur les deux forums créent un communaulecte. Grâce aux outils à leur disposition, les utilisateurs peuvent créer un mode de communication unique à la communauté, qui repose sur le

49. Ces photomontages de chats accompagnés de légendes humoristiques seront abordés en 4.4.3.1.

partage linguistique de graphies alternatives et d'utilisations lexicales spécifiques.

Chapitre 4

Analyse de communaulectes

It is a shame that [Scots] is being 'modernized' and 'electronically' altered by the use of computerized cell phones, instant messages etc.

Charles B. - *Scotster*

Dans ce chapitre, les lignes directrices du rapport entre choix graphiques et pratique commune de membres de forums vont être examinées. Tout d'abord, une définition du terme « communaulecte », sera donnée, afin de fixer le cadre théorique dans lequel le chapitre est ancré. Cette définition s'appuie sur les travaux d'E. Wenger (Wenger, 1998) concernant les communautés de pratique, mais ne s'y cantonne pas. Les travaux de sociolinguistique sur les parlers jeunes (Bulot, 2004) ou les variations sociales langagières (Gadet, 2007) seront mis à contribution pour délimiter la façon dont les internautes composent les sociolectes plus ou moins cryptiques qu'ils utilisent en ligne.

La différence entre une faute d'orthographe et une graphie alternative, choisie, est le premier point de cette délimitation communaulectale, puisque les indices graphiques et lexicaux sont les plus remarquables, et les plus faciles à repérer lorsque l'on étudie la communication sur internet. Les graphies alternatives sont activement utilisées en ligne, dans des pratiques discursives multimodales comme les mèmes (étudiés en 4.4.3) qui mélangent graphies alternatives et images. Ces graphies alternatives

appartiennent au domaine vaste du netspeak, l'ensemble des pratiques linguistiques propres à internet. Les observations des usages non-communaulectaux permettent de tester et de délimiter encore plus précisément les règles d'usage des communaulectes. Connaître la culture internet, et ses pratiques graphiques relève d'un communaulecte à l'échelle mondiale, culturellement.

La qualité et la fonction des échanges (agonaux, cohésifs, humoristiques ou ludiques…) sont également convoquées pour expliquer l'apparition des pratiques communaulectales. L'hypothèse de ce travail est que la fonction cohésive (Rourke et al., 1999) dirige l'alternance codique communaulectale, sur BP et SC. Cette alternance est motivée discursivement, et transforme les graphies normées en graphies de transcription graphophonématiques, ou utilisations de netspeak. Je poursuivrai l'étude en montrant que le communaulecte n'est pas uniquement basé sur des indices graphiques, mais qu'il possède une composante discursive importante, qui le rend parfois plus difficile à cerner pour le lecteur non averti.

4.1 Définition de communaulecte

L'originalité des communautés internet réside dans le fait qu'elles sont constituées de membres actifs, qui s'agrègent tous autour de loisirs ou de traits communs. Ces communautés, théorisées par E. Wenger en tant que « communautés de pratiques » sont presque toutes fondées sur les mêmes critères. Les membres s'agrègent autour d'un intérêt commun, puis développent ensemble un ensemble de pratiques langagières qui se stabilisent autour de leur apparition sur la plateforme. Ce lecte formé par l'ensemble de leurs utilisations du langage correspond à un communaulecte, soit l'équivalent d'un sociolecte où les traits lexicaux et graphiques sont spécifiques, uniquement utilisé en ligne, sur cette plateforme, ou avec des utilisateurs de ces forums de discussion.

Le forum est donc l'espace de sédimentation où ces formes graphiques seront conservées. Elles peuvent être de différentes sortes : certains sont des remotivations à partir du signifiant, comme « Amerikkka », et sont plus fréquentes sur BP que sur d'autres communautés virtuelles, en raison des sujets abordés sur la plateforme, comme les tensions raciales. Les graphies alternatives qui parsèment le

corpus sont sans cesse réactivées par les membres qui y ont recours, et même par ceux qui savent les déchiffrer et y répondent. Ces traces communaulectales s'appuient également sur des choix lexicaux qui peuvent n'appartenir qu'à la communauté virtuelle. Sur BP, par exemple, la section où les utilisateurs peuvent poster des messages courts, lisibles par toute la communauté (comme les « statuts » de FB) s'appelle « What's New ? », et le livre d'or de chaque profil (où les utilisateurs peuvent envoyer des messages lisibles par toute la communauté) s'appelle le *guestbook*. Certains utilisateurs l'ont renommé le *G-spot* (‘*le point G*’), par jeu avec l'initiale. Ce choix lexical est dû à la plateforme et non aux utilisateurs, mais ces derniers détournent le nom choisi par BP pour se singulariser. Dans les fils de discussion, il n'est pas rare de voir des recours à des termes ou expressions qui font partie intégrante de la culture de la plateforme, ou du fil de discussion abordé (par exemple, les fils comme « Random Thoughts » de BP, qui contiennent plus d'un millier de messages pour 2009, 2011, 2013, voient l'apparition de jeux successifs, comme des détournements de pseudonymes pour imiter des personnages de *comics*…).

Les éléments communaulectaux sont éminemment interactionnels, et sont construits sur un patron mémétisant. Les mèmes sont des unités culturelles ou comportementales qui se transmettent et se dupliquent, à la manière d'un gène (Dawkins, 1976). Si l'on adapte cette théorie à l'écologie d'internet, on peut en dire qu'un mème est un sème reconnaissable qui se transmet en ligne. Internet est un excellent terreau pour les manifestations mémétiques, qui sont multimodales (images, graphies, vidéos…). Les mèmes sont également reproductibles à l'échelle d'une communauté virtuelle, c'est pourquoi je les inclus dans la lecture des pratiques linguistiques communaulectales. Celles-ci peuvent inclure les types d'échanges sur la plateforme à part entière (on peut, par exemple, s'exprimer différemment suivant que l'on écrive une réponse à un blog personnel sur BP, que l'on réponde à un sujet de fil de discussion ou bien encore que l'on signe le livre d'or…) Une utilisatrice peut avoir recours à des marques communaulectales dans un fil de discussion : « And also…Spicy **you a** mess…BUT…basically I'm taking it that you don't give a **dayum** if a White **bytch** is laid out or not… » (BP12E1024). Ici, elle utilise les graphies <dayum> et <bytch> pour contourner la censure du site, et omet *be*-copule. Dans son profil personnel, elle écrit de façon tout à fait normée. Le discours

s'adapte à son contexte.

> In real life I have a hump back, eat Starfish and Coffee, and have no qualms with running over a puppy. If that is not enough to get you off of my page from trying to get to know me, then come with me to church and let me tell you about the good news. See you in the forums.

Pour en offrir une définition en négatif, le communaulecte est l'ensemble des pratiques linguistiques qui ne relèvent pas du vocabulaire technique, et qui sont pourtant incompréhensibles immédiatement pour de non-membres de la plateforme. Ils sont utilisés naturellement et imités par les membres, à mesure qu'ils apprivoisent la communication sur cet espace-ci. Une connaissance des thèmes les plus abordés permet de comprendre ces traces communaulectales (les mots les plus employés ont tendance à être tachygraphiés, comme « BP ». On pourrait étendre la perception des communaulectes au vocabulaire que le lecteur de cette thèse partagera avec l'auteur après la lecture. Ceci représenterait l'équivalent du *lurking* sur un forum : lire sans avoir à écrire, mais en comprenant les raccourcis de(s) l'auteur(s), c'est intégrer progressivement les codes. Les lecteurs de *Push* et de *Trainspotting* ont aussi une marge de progression entre les premières pages du livre et la suite de leur lecture. L'adaptation à la lecture et l'acquisition de la littéracie semble identique en ligne et hors ligne, et la pratique de la « typoglycémie » , où l'ordre des lettres d'un mot est altéré, repose sur ce postulat. L'exemple de typoglycémie le plus connu est un texte commençant par « Aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. » [1] où les lettres internes du mot sont déplacées sont lisibles. Cette lecture a un coût (12% de vitesse de lecture en moyenne lorsque l'inversion touche les lettres internes du mot, 36% lorsque l'inversion touche les premières ou dernières lettres du mot) (Rayner, White, Johnson & Liversedge, 2008). Il est tout de même possible de lire les phrases typoglycémiques, tirées d'un contexte large, donc il semble que plus un lecteur est exposé à des pratiques linguistiques, et plus il s'y habitue, donc moins le coût de lecture en est important. Les membres de la « génération numérique » (Papi, 2012) sont aptes à alterner les codes avec plus de fluidité (Bolichowski, 2012). Les membres plus âgés de BP et SC ont une certaine maîtrise de la littéracie numérique, et peuvent adapter leurs discours et comprendre ceux que les autres membres ont produit, le tout créant un ensemble de

1. M. Davis, chercheur du Cognition and Brain Science Unit (Cambridge) a publié une lecture systématique du mème (Davis, 2012).

pratiques linguistiques appelées communaulecte. Le communaulecte est l'ensemble des affinités discursives, lexicales et graphiques développées en commun et en parallèle par les membres d'un forum ou d'une communauté internet. Elles reposent sur les références culturelles, conséquences des agrégats de scripteurs de la même communauté de pratique. Le communaulecte est surtout citationnel, y compris dans un rapport parodique d'écarts à la norme graphique ou grammaticale. Dans mon corpus, ces écarts sont nourris par la proximité de langues minoritaires avec la norme de référence.

La sous-partie suivante est une exploration du rapport entre les communaulectes et les langues minoritaires, qui se recoupent et se rejoignent dans les écrits sur BP et SC.

4.1.1 Communaulectes et langues minoritaires

Le choix de deux communautés réunies peu ou prou par leur appartenance ethnico-culturelle (ou nationale) rajoute une dimension à l'exploration mémétique (je donnerai une définition de ce que j'entends par « mème » en 4.4.4). Les membres de forums et de communautés virtuelles écossaises ou afro-américaines n'utilisent pas toujours la même proportion de pratiques langagières alternatives–ni le même type de discours. Les lectes peuvent s'ajouter les uns aux autres : le communaulecte peut intégrer des éléments de netspeak, et évidemment, des pratiques langagières typiquement afro-américaines ou écossaises (voir les mèmes communaulectaux en 4.4.4 et une étude en profondeur des mèmes en 4.4.3).

Certains préfèrent avoir recours au discours alternatif, en langue minorée, lors des discussions et donc dans des situations d'échange. D'autres, de façon purement personnelle, dans leur description de profil ou sur des blogs. Ces choix relèvent de la construction d'identité par le discours, et sont inconscients dans une certaine mesure. Ceci est illustré par un cas particulièrement visible d'alternance codique entre scots et standard de l'anglais :

> when yer talking tae somebody on the phone fae a different area ye hav [ː] to put a polite english accent on ive found, because they dont understand what your saying. (‘*Quand on parle au téléphone à quelqu'un d'une autre région, on doit [ː] utiliser un accent anglais poli je m'en suis rendu compte, parce qu'ils ne comprennent pas ce qu'on raconte.*’)

L'utilisateur scinde sa phrase en deux parties : celle où il utilise du scots (*yer, tae, fae* '*from*', *ye, hav*), et la seconde, où il évoque les difficultés des Anglais (ou des non-Écossais) à comprendre son accent. Le changement entre l'un et l'autre est montré par le signe <[:]> que j'ai rajouté à la phrase. Le scripteur ajuste certainement involontairement son discours, et donc ses graphies, à son propos. Les scottismes utilisés ne sont pas tant lexicaux que graphiques, mis à part *fae* qui est plus difficile à décoder pour ceux qui ne parlent pas le scots, à cause de la ressemblance avec « for » (nombre de caractères dans le mot, absence du groupe consonantique initial). Discursivement, la seconde partie, si elle n'est pas exempte de fautes d'orthographe (*ive* et *your* pour « I've » « you're »), correspond cependant à l'« accent anglais poli » dont l'utilisateur parle. Ces fautes semblent davantage dues à la paresse et au refus d'utiliser des apostrophes qu'à une ignorance de la graphie standard. Ici, l'identité écossaise (impolie ?) s'oppose au standard anglais, pour une raison simple : l'intelligibilité hors-ligne de son discours. Cette problématique se retrouve dans le corpus internet, et plus les utilisateurs écrivent dans leur propre domaine numérique, leur profil personnel, ou de choses proches d'eux, et plus ils ont tendance à recourir à des pratiques discursives tissées de langue minoritaire, parfois au détriment de l'intelligibilité.

Certaines de ces pratiques discursives sont plus discrètes que les autres : la tachygraphie est plus difficile à cerner, puisque l'utilisation de 3 ou 4 lettres maximum ne permet pas toujours de faire la différence entre une abréviation afro-américaine ou américaine standard. C'est pourquoi certains énoncés resteront toujours ambigus, et que les acronymes sont les éléments les moins faciles à identifier. SMH est à la fois « Shaking My Head » '*[je] secoue la tête*' (en signe de désapprobation) ou « So Much Hate » ('*Trop de Haine*'). La première signification est afro-américaine, et on ne peut les différencier qu'avec l'utilisation d'adverbes intensifieurs, comme dans l'exemple SMDH, « Shaking my Damn Head »). En revanche, SMFH est-il l'acronyme de « Shaking my Fucking Head » ou de « So Much Fucking Hate » ? C'est impossible à savoir sur BP. Une recherche rapide sur Google [2] montre une nette préférence pour « So Much Fucking Hate » (700 000 résultats contre 172 000), explicable par le fait que l'expression est beaucoup plus populaire. Cela n'est pas pour autant une preuve que tous les Afro-Américains utilisent « Shaking My Fucking Head ». En scots, on peut penser à l'utilisation du

2. Recherches sur les deux expressions « Shaking My Fucking Head »et « So Much Fucking Hate ».

suffixe <-ie>, affectif (et déjà évoqué en 2.2.2.2) qui peut se confondre avec…une utilisation similaire, non scottisante de ce suffixe, en anglais standard. Ainsi, « doggie »est une variante de « doggy », et est répertoriée dans plusieurs dictionnaires (Doggie, 2003 ; Gove & Merriam-Webster, 1993). Dans SC, la proportion de graphie alternative est plus élevée que sur internet (1 :25, contre seulement 1 :367, soit 551 graphies alternatives et 13 800 graphies standard sur SC, et 8 920 000 graphies alternatives et 328 000 000 occurrences de graphie standard sur internet.).

Les langues minorées sont cependant un élément cohésif des membres. Dans SC, les membres nord-américains sont heureux de pouvoir utiliser des léxèmes qui ne font *a priori* pas partie de ceux dont ils se servent tous les jours. Cet extrait provient de la page personnelle d'un utilisateur (SCPP1) :

> Convener fur Canada : Me Dad's Clan HOPE (Weir) ; nee Mum's Clan LOGAN (MacLennan). […] Own & operate me business noo. Proud tae sae, Canada is part of thee British Commonwealth ! Cheers, clansman Big Bren & me wee westie guard doggie Bentley.

Le scots utilisé par ce membre est relativement authentique (à part « fur », '*from*' qui s'écrit <fae>, <frae> en scots, et « me » '*my*' est orthographié <mae> ou <ma> en scots (Eagle, 2002), mais cette graphie peut être la graphie oralisante de la prononciation relâchée britannique [ˈmɪ]).« Doggie » a été traité plus haut, mais tous ces scottismes (*noo, tae, sae, wee, doggie*) sont utilisés par un membre non écossais, qui renforce sa légitimité sur le forum en utilisant du scots alors qu'il est canadien [3]. Ici, la langue minoritaire sert à prouver non seulement l'attachement à la « mère-patrie » et, par extension, l'appartenance à la communauté écossaise, en ligne. Les langues minorées sont en quelque sorte une carte de visite pour les utilisateurs. Quant aux utilisateurs de BP, s'ils n'ont ni à prouver leur origine, ni la couleur de leur peau quand ils participent, certains utilisateurs non noirs utilisent des codes afro-américains pour s'exprimer. C'est le cas des femmes blanches qui recherchent un partenaire noir, par exemple. Les deux extraits suivants proviennent respectivement d'un profil personnel, et de la description d'un groupe destiné aux femmes blanches qui aiment les hommes noirs :

> A LIL BOUT YA GIRL !!! […] Brothas, you have to be faithful,honest, and someone who likes to treat me like a princess.[…] I may be a white girl, but I've been through a lot with some brothas before. I'm not "ghetto" or a "hood rat".

3. Le fait que ses parents aient tous les deux un lien avec un clan écossais ne veut pas dire qu'ils ne sont pas tous les deux Canadiens—le plus grand groupe ethnique de Nova Scotia est, comme son nom l'indique, d'origine écossaise (28.3% de la population) (Government of Canada, 2007)).

> […] HOLLA BACK AT YOUR GIRL AND GET TO KNOW ME FOR WHO I REALLY AM (NOT JUST BASED ON THE PIC'S OR WHAT YOU READ) ! (BPPP3)
>
> this iz 4 all the sex-z black me nwho luv them sum white women !! since i am personally a white girl, i luv black men n everything they're about. so give ma group a holla if everything applies !!! (BPG1)

Ces deux extraits partagent des alternances codiques qui flirtent avec l'AAV : BPPP3 débute avec des figurations graphiques d'amuissements de phonèmes (<lil> où le /t/ est amuï ; et <bout> montre l'amuïssement de la première voyelle /ə/), observées dans l'AAV. Certaines graphies alternatives montrent une oralisation ou une graphie alternative stylistique, comme <ya>, <iz>, <luv>, <ma>, <n>, <sum> qui sont à la fois utilisées en AAV sur BP et par les internautes (dans le netspeak).

On peut également y voir des utilisations lexicales afro-américaines, comme « holla », « brothas » (qui sont dans un entre-deux graphique : les terminaisons <-er> écrites <-a> sont des tentatives afro-américanisantes, comme vu en 2.4.2). Les stéréotypes du « hood rat » et du « ghetto » renvoient à la culture afro-américaine stéréotypée : ces deux utilisatrices tentent d'émuler des thèmes et des graphies familières, d'utiliser des mots provenant de l'argot afro-américain pour faire valoir leur légitimité dans la recherche d'un partenaire noir.

La langue minorée, fantasmée ou non, est un élément incournable pour la cohésion indispensable à la création d'un sentiment d'appartenance. Cette appartenance est d'autant plus fragile que certains se considèrent à la marge de la communauté virtuelle, puisqu'ils ne font pas partie de la communauté (géographique) hors-ligne. Ils multiplient donc les marques de langue minorée. Les Écossais et les Afro-Américains utilisent également ces signes d'appartenance à leur communauté linguistique, et ces choix dépendent de l'identité qu'ils choisissent d'endosser (fervent nationaliste, amoureux du scots d'une part ; gangster ou habitant du *hood* de l'autre).

À présent que la définition du communaulecte est établie, autant dans son rapport avec le netspeak et internet qu'avec les langues minoritaires, les productions des utilisateurs de BP et SC, qui reposent en majeure partie sur les jeux graphophonématiques doivent être décrites. La partie suivante s'attache à créer une typologie des équivalents graphophonématiques en anglais standard, pour voir à quelles séquences graphiques ces scripteurs peuvent avoir recours pour leurs écrits.

4.2 Équivalence graphophonémique en anglais standard

Pour introduire la communication sur plateformes virtuelles, je rappelle la définition de N. Baron [4], qui délimite les échanges sur internet comme « une sorte de centaure linguistique, qui réunit des éléments discursifs de la communication écrite et du face à face, tout en étant plus que le mélange des deux ». Cette définition est un peu floue, et ne répond pas à la question de la communication sur internet : les pratiques lexicales, typographiques, et les interactions, dans le présent corpus, des dialectes et pratiques de langues minorées forment en fait un continuum linguistique qu'il est difficile de ne caractériser que comme un mélange qui est plus que l'écrit et l'oral ensemble.

Si la communication n'était qu'orale, il y aurait perte de codes, notamment intonatifs et appartenant même à la communication non verbale. Le ton et la production seraient tout à fait différents. Il y a une différence entre l'espace-temps de la communication face à face et la communication virtuelle. Il est à la fois impossible d'avoir une discussion avec autant de participants, commençant tous à répondre au même message de base en même temps à l'oral. Il est tout aussi impossible de conserver la discussion (jusqu'à être certain de la date et de l'heure précise de l'apparition de la première occurrence d'un phénomène) orale sans que les moyens de communication en soient affectés.

Si la communication médiée par ordinateur n'était pas écrite, pourquoi ne pas imaginer une écriture complètement libérée des codes de l'écrit ? Par exemple, sans les espaces entre les mots, et sans aucune importance accordée aux codes de l'écrit tel qu'il est enseigné ou standardisé. Certains mouvements (*Sayspel*, par exemple) prônent un abandon de l'écriture de l'anglais tel qu'il est enseigné car ils considèrent les équivalences grapho-phonémiques illogiques et trop nombreuses. En revanche, sur internet, on pourrait imaginer que chacun passe par un système non standard qu'il aurait conçus lui-même. Quels seraient les problèmes rencontrés dans la communication ? Ces écrits seraient impossibles à lire : le travail du lecteur, pour déchiffrer les mots, et le travail du scripteur, pour passer à son propre code personnel, seraient trop chronophages, avec des risques que personne ne lise les productions ainsi composées. La perte d'efficacité de la communication lui serait directement délétère.

4. « a kind of a linguistic centaur, incorporating features from both traditional written writing and face-to-face discourse but ending up being more than a simple amalgam of the two » (Baron, 2003).

Les utilisateurs d'internet ne s'y trompent pas : la liberté de la graphie en ligne n'est pas exploitée par tous au maximum, et ceux qui dérogent aux règles établies sont sanctionnés (au moins par de petites réflexions d'autres utilisateurs). Ceux qui ont recours aux majuscules doivent expliquer leur choix :

> THIS IS JUST LIKE ME TYPING IN ALL CAPS.....IT HAS BEEN SAID TO ME...."NLRSIDESTAR WHY ARE YOU SCREAMING".....NOW IM NOT DOING IT TO BE SCREAMING.....BUT.....I KNOW THAT TYPING IN CAPS ONLINE MEANS SCREAMING TO THE ONLINE COMMUNITY.....ITS CONSIDERED SCREAMING TO TYPE IN ALL CAPS.....NOW I DO KNOW THAT, YOU UNDERSTAND.....BUT I JUST LIKE TO TYPE IN CAPS BECAUSE I LIKE CAPS NOT BECAUSE I LIKE TO SCREAM.....BUT TO EVERYONE ONLINE WHO CHATS ONLINE AND THINGS OF THAT NATURE.....I AM DEFINITELY SCREAMING BY TYPING IN ALL CAPS.....does that make it any clearer as to what im saying ??????? (BP12E2188)

L'utilisatrice est consciente que sa graphie va à l'encontre de ce qui est attendu d'elle. Le standard en ligne, c'est d'écrire en minuscules, comme dans une production écrite hors ligne. L'interprétation normale de l'écriture en majuscule est de penser que l'utilisatrice crie (hors-ligne). Ce standard relève du *netspeak*, et c'est la raison pour laquelle elle doit faire une telle mise au point dans le fil de conversation. Ceci ne relève pas de la cohérence graphophonématique, mais de la pragmatique, et des règles de politesse en ligne.

La CMO est donc écrite, suivant des règles standards que l'on suppose avoir été inculquées par contact avec d'autres écrits en ligne, ceci faisant partie de l'écologie large de la communication sur ordinateur, mais les locuteurs peuvent avoir recours à un registre général oralisant, dépendant du contexte et de leur volonté.

4.2.1 Équivalences consonantiques

En anglais standard, il existe toujours plusieurs graphies pour retranscrire un seul phonème. Sans revenir aux causes diachroniques (ou étymologiques) qui ont causé ces différences de transcription, le tableau suivant (4.1) montre les sons et leurs correspondances graphiques afférentes. Elles sont classées par ordre de fréquence, pour les 26 sons consonantiques de l'anglais (si l'on considère que tous les graphèmes, y compris muets, servent à symboliser les sons). Certaines graphies sont très rares,

notamment dans les emprunts lexicaux à d'autres langues (c'est le cas pour « mho » qui est la seule occurrence de <mh> pour représenter un /m/).

Son	Séquences graphiques	Exemples
/p/	<p, pp, gh>	**p**ill, ha**pp**y, hiccou**gh**
/b/	<b, bb>	**b**u**bb**le
/t/	<t, tt, ed, pt, th, ct>	**t**ent, a**tt**ach, cropp**ed**, **pt**erodactyl, **th**yme, **ct**enoid
/d/	<d, dd, ed, dh, t, tt>	**d**ove, a**dd**, mail**ed**, **dh**arma, wai**t**er, fla**tt**er
/g/	<g, gg, gue, gh>	**g**ig, sta**gg**er, catalo**gue**, **gh**oul
/k/	<c, k, ck, ch, cc, qu, cqu, cu, que, kk, kh, q, x>	**c**at, **k**ey, ba**ck**, **ch**ord, a**cc**ount, li**qu**or, a**cqu**is, bis**cu**it, mos**que**, tre**kk**er, **kh**an, bur**q**a, e**x**citement
/m/	<m, mm, mb, mn, mh, gm, chm>	**m**ass, ha**mm**er, cli**mb**, hy**mn**, **mh**o, diaphra**gm**, dra**chm**
/n/	<n, nn, kn, gn, pn, nh, cn, mn>	**n**ice, fu**nn**y, **kn**ee, **gn**osis, **pn**eumonia, pira**nh**a, **cn**idarian, **mn**emonic
/ŋ/	<ng, n, ngue>	si**ng**, li**n**k, to**ngue**
/r/	<r, rr, wr, rh, rrh>	**r**ay, a**rr**ay, **wr**ong, **rh**yme, dia**rrh**ea
/f/	<f, ph, ff, gh, pph, u>	**f**ine, **ph**ysical, o**ff**, lau**gh**, sa**pph**ire, lie**u**tenant
/v/	<v, vv, f, ph, w>	**v**ine, sa**vv**y, o**f**, Ste**ph**en, **w**eltanschauung
/θ/	<th, chth, phth, tth, fth>	**th**in, **chth**onic, **phth**isis, Ma**tth**ew, twel**fth**
/ð/	<th, the>	**th**en, brea**the**
/s/	<s, c, ss, sc, st, ps, sch, cc, se, ce, z>	**s**in, **c**ity, me**ss**, **sc**ene, li**st**en, **ps**ychiatrist, **sch**ism, fla**cc**id, hor**se**, ju**ice**, citi**z**en
/z/	<s, z, x, zz, ss, ze, c>	ha**s**, **z**oo, **x**ylophone, bu**zz**, sci**ss**ors, bree**ze**, elec-tri**c**ity
/ʃ/	sh, ti, ci, ssi, si, ss, ch, s, sci, ce, sch, sc	**sh**uffle, na**ti**on, spe**ci**al, mi**ssi**on, expan**si**on, ti**ss**ue, ma**ch**ine, **s**ugar, con**sci**ence, o**ce**an, **sch**mooze, cre**sc**endo
/ʒ/	<si, s, g, z, j, ti, sh>	divi**si**on, lei**s**ure, **g**enre, sei**z**ure, **j**eté, equa**ti**on, Per**sh**ing
/ʧ/	<ch, t, tch, ti, c, cc, tsch, cz>	**ch**in, na**t**ure, ba**tch**, bas**ti**on, **c**ello, bo**cc**e, pu**tsch**, **Cz**ech
/ʤ/	<g, j, dg, dge, d, di, gi, ge, gg>	ma**g**ic, **j**ump, ju**dg**ement, bri**dge**, gra**d**uate, sol**di**er, Bel**gi**an, dun**ge**on, exa**gg**erate
/h/	<h, wh, j, ch>	**h**e, **wh**o, fa**j**ita, **ch**utzpah
/j/	<y, i, j, ll, e>	**y**es, on**i**on, hallelu**j**ah, torti**ll**a, **e**oarchean
/l/	<l, ll, lh>	**l**ine, ha**ll**o, **Lh**asa
/w/	<w, u, o, ou, wh>	**w**e, pers**u**ade, ch**o**ir, **Ou**ija, **wh**at
/hw/	<wh>	**wh**eel
/wr/	<wr>	**wr**en

Tab. 4.1 – Correspondances phonie-graphie des consonnes en anglais standard

Ces correspondances ne sont pas exactes, et ne concernent pas toutes les variétés de l'anglais. Pour les correspondances de /d/ avec certaines graphies (<t, tt>), c'est le cas dans certains dialectes de l'anglais (voire de l'anglais américain uniquement), tout comme <c> prononcé /z/.

Pour chaque phonème consonantique, il y a en moyenne 5,6 graphies possibles. Le son /k/ est celui qui a le plus de variations, avec 13 représentations graphiques ; les sons /hw/ (/ʍ/) et /wr/ qui ont le moins de variantes graphiques, ne sont reconnus comme des sons à part entière que dans certains dialectes de l'anglais.

4.2.2 Équivalences vocaliques

Les voyelles sont les éléments les plus prompts à changer. D'ailleurs, les équivalences graphiques des sons vocaliques sont beaucoup plus nombreuses que pour les consonnes.

Le tableau 4.1 représente ces équivalences entre les sons vocaliques et les graphies utilisées pour les représenter, des plus fréquentes aux moins courantes—la plupart de ces prononciations apparaissent en syllabes accentuées (il va sans dire que toutes les séquences graphiques représentant /ə/sont contenues dans des syllabes non-accentuées. Ces prononciations ne sont pas actualisées dans toutes les variétés de l'anglais.

Son	Séquences graphiques	Exemples
/iː/	<e, ea, ee, e…e, ae, ei, i…e, ie, eo, oe, ieCe, ay, ey, i, y, oi, ue, ey, a>	b**e**, b**ea**ch, b**ee**, c**e**d**e**, C**ae**sar, dec**ei**t, mach**i**ne, f**ie**ld, p**eo**ple, am**oe**ba (variante d'*ameba*), hyg**ie**n**e**, qu**ay** (*Received Pronunciation* (anglais britannique) (RP) ; GA : /eɪ/), key, ski, city, chamois, Portuguese, geyser (RP) ; GA : /aɪ/), karaoke
/ɪ/	<i, y, ui, e, ee, ie, o, u, a, ei, ee, ia, ea, iCe, ai, ii, oe>	b**i**t, m**y**th, b**ui**ld, pr**e**tty, b**ee**n (certains accents), s**ie**ve, w**o**men, b**u**sy, dam**a**ge, counterf**ei**t, carr**ia**ge, mil**ea**ge, medic**i**n**e**, barg**ai**n, sh**ii**take, **oe**dema
/uː/	<oo, u, o, u…e, ou, ew, ue, o…e, ui, eu, oeu, oe, ough, wo, ioux, ieu, oup, w, u>	t**oo**l, l**u**minous, wh**o**, fl**ute**, s**ou**p, j**ew**el, tr**ue**, l**o**s**e**, fr**ui**t, man**eu**ver, man**oeu**vre (variante britannique du précédent), can**oe**, thr**ough** (variante de "thru"), t**wo**, S**ioux**, l**ieu**tenant (GA), c**oup**, c**w**m, thr**u** (informel)
/ʊ/	<oo, u, o, ooCe, or, ou, oul>	l**oo**k, f**u**ll, w**o**lf, g**oo**s**e**berry, w**or**sted, c**ou**rier, sh**oul**d

Illustration 4.1 – Équivalences graphophonémiques pour les voyelles de l'anglais.

Son	Séquence graphiques	Exemples
/ə/	<a, e, o, u, ai, ou, eig, y, ah, ough, ae, oi>	**a**nother, anth**e**m, awes**o**me, atri**u**m, mount**ai**n, call**ou**s, for**eig**n, ber**y**l, Messi**ah**, bor**ough** (RP), Mich**ae**l, porp**oi**se
/ɛ/	<e, ea, a, ae, ai, ay, eaCe, ei, eo, ie, ieu, u, ue, oe>	m**e**t, w**ea**ther, m**a**ny, **ae**sthetic, s**ai**d, s**ay**s, cl**ea**ns**e**, h**ei**fer, j**eo**pardy, fr**ie**nd, l**ieu**tenant (RP), b**u**ry, g**ue**ss, f**oe**tid
/æ/	<a, ai, al, au, i>	h**a**nd, pl**ai**d, s**al**mon, l**au**gh (dans certaines variétés), mer**i**ngue
/ʌ/	<u, o, oCe, oe, ou, oo, wo>	s**u**n, s**o**n, c**o**m**e**, d**oe**s, t**ou**ch, fl**oo**d, t**wo**pence
/ɔː/	<a, au, aw, ough, augh, o, oa, oo, al, uo, u, ao>	f**a**ll, **au**thor, j**aw**, b**ough**t, c**augh**t, c**o**rd, br**oa**d, d**oo**r, w**al**k, fl**uo**rine (RP), s**u**re (dans certaines variétés), extr**ao**rdinary
/ɒ/	<o, a, eau, ach, au, ou>	l**o**ck, w**a**tch, bur**eau**cracy, y**ach**t, s**au**sage, c**ou**gh
/ɑː/	<a, ah, aa, i>	f**a**ther, bl**ah**, b**aa**, l**i**ngerie (GA)
/aɪ/	<ae, ai, aie, aille, ais, ay, aye, ei, eigh, ey, eye, i, iCe, ia, ie, ic, ig, igh, is, oi, ui, uy, uye, y, yCe, ye>	m**ae**stro, kr**ai**t, shangh**aie**d, can**aille** (RP), **ais**le, k**ay**ak, aye, h**ei**st, h**eigh**t, g**ey**ser (GA), **eye**, m**i**c, f**i**n**e**, d**ia**per, t**ie**, ind**i**ct, s**i**gn, h**igh**, **is**le, ch**oi**r, g**ui**de, b**uy**, g**uye**d, tr**y**st, t**y**p**e**, b**ye**
/eɪ/	<a, a…e, aa, ae, ai, aiCe, aig, aigh, al, ao, au, ay, e (é), eCe, ea, eg, ei, eiCe, eig, eigh, ee (ée), eh, er, es, et, ey, ez, ie, oeh, ue, uet>	b**a**ss, r**a**t**e**, qu**aa**lude, reggae, r**ai**n, coc**ai**n**e**, arr**aig**n, str**aigh**t, R**al**ph (RP archaïque), g**ao**l, g**au**ge, p**ay**, ukul**e**le (caf**é**), cr**e**p**e**, st**ea**k, th**eg**n, v**ei**l, b**ei**g**e**, r**eig**n, **eigh**t, matin**ee** (soir**ée**), **eh**, dossi**er**, dem**es**ne, ball**et**, ob**ey**, ch**ez**, lingeri**e** (GA), b**oeh**mite (ou /oʊ/), dengue (ou /i/), sobriquet (ou /ɛt/)
/ɔɪ/	<oi, oy, awy, uoy oyCe, eu>	f**oi**l, t**oy**, l**awy**er, b**uoy**, garg**oyl**e, Fr**eu**dian
/aʊ/	<ou, ow, ough, au, ao>	**ou**t, n**ow**, b**ough**, t**au**, T**ao**ism
/oʊ/	<o, oCe, oa, ow, ou, oe, oo, eau, oh, ew, au, aoh, ough, eo>	s**o**, b**o**n**e**, b**oa**t, kn**ow**, s**oul**, f**oe**, br**oo**ch, b**eau**, **oh**, s**ew**, m**au**ve, phar**aoh**, furl**ough**, y**eo**man
/ɑr/	<aar, ar, are, arre, ear, er, our, uar, arrh>	baz**aar**, c**ar**, **are**, biz**arre**, h**ear**t, s**er**geant, **our** (dans certaines variétés), g**uar**d, cat**arrh**

Illustration 4.1 – Équivalences graphophonémiques pour les voyelles de l'anglais. (suite)

Son	Séquence graphiques	Exemples
/ɛəɾ/	<aar, aer, air, aire, ar, are, ayer, ayor, ear, eir, er, ere, err, erre, ey're, e'er>	**Aar**on, **aer**ial, h**air**, million**aire**, v**ar**y, w**are**, pr**ayer**, m**ayor**, b**ear**, h**eir**, station**er**y (dans certaines variétés), wh**ere**, **err** (certaines variétés), part**erre**, th**ey're**, **e'er**
/ɪəɾ/	<ear, eer, eir, eor, ere, ers, e're, ier, iere, ir>	**ear**, b**eer**, w**eir**, th**eor**y (US), h**ere**, rev**ers**, w**e're**, p**ier**, prem**iere**, menh**ir**
/ɜr/	<er, or, ur, ir, yr, our, ear, err, eur, yrrh, ar, oeu, olo>	f**er**n, w**or**st, t**ur**n, th**ir**st, m**yr**tle, j**our**ney, **ear**th, **err**, amat**eur**, m**yrrh**, gramm**ar**, hors d'**oeu**vre, c**olo**nel
/juː/	<u, uCe, eu, ue, iew, eau, ieu, ueue, ui, ewe, ew>	m**u**sic, **use**, f**eu**d, c**ue**, v**iew**, b**eau**tiful, ad**ieu**, q**ueue**, n**ui**sance, **ewe**, f**ew** (dans certains dialectes)

Illustration 4.1 – Équivalences graphophonémiques pour les voyelles de l'anglais. (suite)

Il y a en moyenne 12,6 séquences graphiques possibles pour chaque phonème vocalique. Les <r> postvocaliques, qui provoquent une tension de la voyelle précédente en anglais RP, ne provoquent pas les mêmes tensions en scots, puisqu'il s'agit d'une langue rhotique. Les <r> sont donc écrits, puisque prononcés. En revanche, en AAV, les schwas colorés par les <r> (ə˞) sont souvent retranscrits avec des <a>, comme en *typobonics* (voir 2.4.2), évidemment en position finale dans les mots.

Les mots d'origine étrangère sont paradoxaux. À cause de leur origine, ils rajoutent des séquences graphiques pour retranscrire des sons adaptés à l'anglais, puisque les graphies sont différentes. C'est le cas de *myrrh*, *adieu*, *amateur*, ou *dengue*…Ces séquences graphiques se rajoutent donc à celles de l'anglais, mais ne sont pas des transcriptions courantes, pas plus que certaines graphies archaïques, conservées du vieil anglais ou du moyen-anglais.

4.3 Fautes d'orthographe

Le statut de la faute d'orthographe est délicat : tout dépend du curseur normatif, qui n'exerce pas la même influence sur les plateformes de discussion sur internet. La normativité des discours est moins graphique pour les modérateurs ou administrateurs des sites que discursives, puisqu'ils visent à la

communication non-injurieuse (ce qui est inscrit dans les chartes d'accès aux forums, à l'inscription). La normalisation est donc exercée différemment, et il faut être précautionneux lors de la constitution d'une liste de fautes sur internet.

Néanmoins, une figure importante de la culture internet est le *Grammar Nazi*, qui est en fait assez souvent un *Spelling Nazi*, si l'on peut dire : les fautes d'orthographe se localisent souvent autour d'homophones (*your* au lieu de *you're*, par exemple) qui recatégorisent la proposition, et ces défenseurs auto-proclamés de la normativité régulent l'orthographe de certains échanges. Plus d'information sur les fautes d'orthographe et leurs détracteurs seront explorées dans la section concernant la dimension normative (4.3.5), puisqu'elle n'est mise en exergue que par ses défenseurs.

Heureusement, il existe des outils de correction orthographique dont les utilisateurs peuvent bénéficier (parfois inclus dans les navigateurs internet). En outre, les avancées en NTIC marchent main dans la main avec les travaux sur les modélisations de fautes. Il en existe plusieurs types, pour les dictionnaires de traitements de texte (*Word* de Microsoft, ou *Writer* d'Open Office par exemple) ; les fautes sont prédites dans les algorithmes de correction automatique ou de saisie intuitive [5]. Les travaux de modélisation de fautes peuvent apporter un début de réponse quant aux fautes d'orthographe liées à la saisie sur clavier. La modélisation des erreurs sur laquelle cette partie repose (Deorowicz & Ciura, 2005), a été réalisée pour pallier les problèmes de graphies, qui interviennent à plusieurs étapes de la constitution du mot, dans l'optique de créer un correcteur orthographique plus performant.

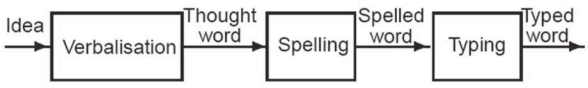


Illustration 4.2 – Modèle de Ciura & Deorowicz (Deorowicz & Ciura, 2005)

5. Les smartphones et les tablettes numériques offrent tous des corrections automatiques, plus ou moins heureuses, mais ces lexiques préconstruits savent acquérir les mots uniques (noms présents dans le répertoire des contacts, graphies alternatives) utilisés par le possesseur.

4.3.1 Verbalisation

L'étape de verbalisation concerne la construction du mot par adjonction de composants (affixes). Cette construction est déjà effectuée et dépassée pour les mots fréquents. Cette étape joue peu pour les locuteurs natifs d'une langue, puisqu'ils maîtrisent la sélection d'éléments corrects. Par exemple, **dérespectueux* est une construction erronée en français, mais *disrespectful* est correct en anglais. La sélection de préfixes et de suffixes (parfois synonymes) obéit aux lois de la langue et des utilisations grammaticalisées. Parmi tous les préfixes privatifs (a(b)-, dé-, iC- …), le locuteur natif sait choisir celui qui appartient à sa langue.

Beaucoup de mots sont aussi vus comme étant des ensembles sans affixe (par exemple, pour celui qui n'a pas réfléchi à la question, « orthographe » est vu comme un ensemble, et non comme une construction de [ORTHO]+[GRAPHE]. La grammaticalisation d'expressions figées court-circuite la réflexion autour de la création d'un mot, et propose des simplifications qui circulent par la parole, comme le marqueur quotatif « genre » (« like » en anglais) qui sont passés d'un statut d'interjections presque vides de sens à éléments grammaticalisés. Leur utilisation permet de remplacer des structures plus longues dans la phrase à valeur quotative équivalente (ROSSARI, RICCI & SPIRIDON, 2009).

La construction de la verbalisation est donc un phénomène actif chez le locuteur de L2, mais est relativement rare chez le locuteur natif, et le corpus contient relativement peu de fautes de ce type. Sur BP, la description personnelle d'un membre illustre ce phénomène :

> AM VERY OPEN GUY AND VERY MUCH UNDERSTANDIN ,STRONGLY I BELIEVE IN DIALOQUE,AM READY TO MATCH WITH ANY LADY AROUND THE WORLD IRRESPECT WITH DISTANCE COLOUR OR RACE .AM VERY SLIM GENTLEMAN IN AFGHANISTAN(WORKING) PLEASE IF U HAPPEN TO MAY U NEED SUCH A GUY JUST CLICK AND YOU WILL BE ANSWERED.

Ce membre (BPPP4) utilise ainsi « irrespect with » au lieu d'*irrespective of*, la forme attendue ici. Le style général de cette présentation personnelle montre que le locuteur semble avoir appris l'anglais en tant que L2, hypothèse soutenue par les non-respects de la grammaire anglaise (*to* + modal, l'omission du sujet). Les utilisateurs non-locuteurs natifs de l'anglais ne participent que très rarement aux fils de

discussion, ou ne le font que lorsque leur niveau de langue est bon, voire très bon. Ce type de discours alternatif est très rare dans le corpus général.

4.3.2 Graphie

Ce que Deorowicz et Ciura nomment les « misspellings », donc le non-respect de la norme graphique qui advient à l'étape de la graphie, sont les plus intéressantes. Pour retranscrire les sons, les utilisateurs ont un arsenal important à leur disposition, et peuvent choisir le mauvais candidat graphique pour retranscrire un mot qu'ils n'ont entendu…qu'à l'oral.

Les graphies incorrectes peuvent devenir le nouvel étalon d'échange sur internet, étant donné qu'aucune instance normative ne pèse sur la production graphique des utilisateurs. Dans le fil de discussion (*Anyone get tricked by a chick with a gurdle ?*, la graphie alternative <gurdle> est réutilisée par certains, qui avaient pourtant recours à l'orthographe <girdle> au début du fil de discussion. On touche ici du doigt la genèse du terme communaulectal involontaire. La graphie non-orthographique présentée est reprise par les autres, dans ce fil de discussion (4.2).

Message	Corps
1	Watching a conversation on facebook right now where a girl from HS apparently bought a **gurdle** to go out to the club last night because "it doesn't look cute if your stomach shakes more than your azz" […]
3	A **gurdle** looks uncomfortable as hell, if you hate your stomach that much then it's time to work out and lose weight and just be happy with what you have until that process happens. I know that a girdle has it uses, even thinner women wear them and some guys like to look of woman wearing a girdle but it seems like women are wearing as a replacement to losing weight.

Tab. 4.2 – Messages 1,3 de « Anyone get tricked by a chick with a gurdle ? » (BP12E5532)

L'étape de la graphie concerne activement le locuteur natif comme non-natif. Là, il n'est pas question de performance, ou de maîtrise de la langue, mais de pouvoir composer un mot par rapport à l'image sonore cognitive que l'on en a, si la graphie est inconnue. Sans entrer dans une opposition saussurienne et parler d'« image sonore du signifiant » entrant en conflit avec son image écrite, force est de constater qu'il est effectivement possible de trouver des problèmes dans la création graphique

d'un mot, surtout parce qu'il existe plusieurs concurrents graphiques pour représenter le même son pour qui ne sait pas écrire le mot.

Dans l'exemple 4.2, le mot « girdle » [6] ('*body*', dans le corpus) a retenu mon attention : en effet, il est écrit « gurdle », dans le titre du fil de discussion, et les utilisateurs vont alors osciller entre la graphie normée <girdle>, et la graphie alternative <gurdle>. La classe d'isophones de « gurdle » est NURSE pour le son /ɜː(r)/, et les graphies potentielles correspondant à cette classe sont <earC, yrC, ir, urC>. Les deux graphies cohabitent et *gurdle* représente 35% des occurrences dans le fil de discussion, soit 6 utilisations. Certains utilisateurs oscillent entre les deux graphies. Pourtant, chez le créateur du fil de discussion, c'est la seule graphie utilisée. Il y a donc un problème de construction graphique du mot, dû certainement à la proximité graphique avec « girl »et sa graphie familière, *gurl*. L'utilisation erronée du créateur du fil de discussion laisse penser que le terme est communaulectalisé. Une recherche sur Google permet de voir environ 12 800 résultats sur internet, hors forum, dont une entrée sur l'UD. Cette graphie non-standard est donc potentiellement réalisable par divers scripteurs, certainement pour des raisons similaires. Pour certains utilisateurs d'internet, il semble exister une libre variation paradigmatique sur la phonographématique des mots inconnus, qui sélectionnent les graphies qui leur semble la plus vraisemblable.

Ces graphies erronées sont sans doute la conséquence de la faible fréquence de ce mot. La fréquence relative des mots est un des facteurs les plus importants : plus un mot est rare, plus il sera difficile à orthographier (en revanche, plus un mot ou une expression est fréquent, et plus souvent il sera tachygraphié…ou mal orthographié s'il a des homophones).

4.3.3 Saisie sur clavier

Les problèmes de saisie sur clavier (qu'il soit physique ou virtuel) sont les plus fréquents, et les derniers à arriver dans la conceptualisation d'un message. Les fautes de saisie sur clavier sont faciles à repérer, puisqu'elles jouent principalement autour du redoublement, de la suppression ou de l'in-

6. Cette partie reprend des éléments de « Graphic conventions on niche websites : sedimentation of alternative spelling » (GOUDET, 2011).

version d'une seule lettre (par exemple, *menatl* < *mental*, (SC2V13) ; ou *grndfather* < *grandfather* (SC2V13)). Ces fautes peuvent être difficiles à percevoir, puisque l'omission d'une seule lettre n'est pas toujours un problème de saisie sur clavier : le mot *decendant* (< *descendant*, SC2V13) peut être une faute de graphie à cause de l'équivalence phonographématique /s/→<s>, <c>, <sc>.

Pour l'étude de fautes de saisie sur clavier, il faut bien sûr garder en tête la configuration des touches (question écologique, déjà abordée en 3.1, au chapitre précédent). Les fautes de saisie sont relativement courantes, et sont provoquées par le manque de soin de la relecture des membres des forums. Une autre raison évocable pourrait être le degré d'émotion ressentie par le scripteur (plus il est sous le coup de la colère, ou de la joie, et moins il relirait ses productions avant de les envoyer) : « I think you're so excited, you can't type…hahaha » (SCP3). L'utilisateur fait remarquer gentiment (le ton léger est justifié par l'utilisation de « hahaha ») à un autre que sa graphie est hésitante, et change certainement de ce à quoi il avait habitué les autres membres de SC. Parfois, la saisie sur clavier est rendue plus difficile par les distractions extérieures (« I can't write anymore right because I'm at work and it's pretty busy » (BPPP5)). L'utilisatrice qui écrit ceci dit que la qualité et la quantité de ses productions écrites est affectée par le fait qu'elle est au travail. L'écologie dans laquelle le scripteur évolue influence donc ses productions.

Certains refusent également de relire leur messages, problème évoqué dans la partie 4.3.5. Quelle que soit la source de l'erreur de saisie sur clavier, cette faute est souvent mieux perçue que le problème de graphie, qui joue parfois sur l'homophonie (<they're>, <their> [ðeə^{r}]. La faute graphique est souvent considérée comme stupide (cf le discours sur le *typobonics*, et celui des normativistes virtuels– « Things I dislike are […] uneducated people (if you can't type then don't write) (BPPP8 »), alors que la faute de saisie sur clavier est analysable comme une étourderie, allant jusqu'à la désinvolture.

Après cette petite typologie, inspirée par celle des concepteurs de correcteurs automatiques (Deorowicz & Ciura, 2005), et ainsi repéré le principe de faute graphique, il est temps de s'intéresser aux néologismes. Les néologismes et les mots déclarés comme inconnus suivent des règles qui leur sont propres, et ils sont plus difficiles à classifier que les mots dont la graphie est inconnue.

4.3.4 Néologismes et mots inconnus

Les néologismes et écritures hasardeuses de mots inconnus sont un terrain idéal pour tester les hypothèses concernant les graphies alternatives. Certains utilisateurs préviennent les autres lorsqu'ils pensent avoir fait une faute d'orthographe, en incise–mais malheureusement, ces tentatives sont rarement signalées, et presque uniquement visibles avec une lecture humaine du corpus.

Sur SC, la recherche « spelt wrong » renvoie 7 résultats ; sur BP, il y a 65 occurrences de l'expression. L'autre indice fréquemment utilisé afin de coder l'ignorance concernant la graphie d'un mot est l'utilisation du point d'interrogation, souvent doublé, entre parenthèses <(??)>. Si un seul point d'interrogation amènerait à s'interroger sur le sens de la phrase, le double point d'interrogation se place au niveau métatextuel. Le problème principal de ce genre de requête est de trouver de tels exemples, puisque BP et Google ne permettent pas de chercher des fils de caractères non alphanumériques.

> logerithums (probably spelt wrong) (SC22E1)

L'utilisatrice prévient les destinataires de sa faute potentielle : elle ne sait pas écrire « logarithm » : la graphie normée de ce mot n'est pas évidente. Il n'est pas composé de la combinaison graphique <-rhythm>, bien qu'il puisse se prononcer de la même façon /ˌlɒgəˈrɪðm̩/, /ˌlɒgəˈrɪθm̩/. La scripteuse a donc décidé d'écrire le mot de la façon la plus instinctive, en utilisant <u> pour figurer le /m̩/ ou /əm/. Sa représentation inconsciente de la consonne syllabique est donc une voyelle. Une recherche sur un mot qui pourrait avoir une graphie alternative relativement similaire, possum > <possim> n'est pas réalisable quantitativement à cause du trop grand nombre de résultats parasites. Outre le subjonctif latin, divers pseudonymes sont écrits avec cette graphie. Il est tout de même possible de trouver des graphies telles que <possim> dans des fils de discussions sur Amazon « how to get rid of possims »[7], et des images d'oppossums sur Google Images. L'interprétation de la consonne syllabique hésite donc entre <um> et <im>.

7. Dans la section du forum qui concerne les astuces du jardin, disponible à l'adresse <http ://www.amazon.com/forum/Patio%20Lawn%20and%20Garden%20Tips%20and%20Tricks ?_encoding=UTF8&cdForum=Fx27NQJFBYNQ3P6&cd
Dernier accès le 04/08/14.

> this is how some people in the medical field keep their jobs …making up new terms…in high school it was called socially awkward, nerds, mysfits(spelled wrong probably) (BP12E763)

Le discours étant relativement sérieux (sur l'utilisation du mot « autiste »), l'utilisateur essaye de trouver le plus de synonymes possibles pour ce qu'il estime être l'autisme. Le mot « awkward »est moyennement fréquent en anglais, selon le dictionnaire Collins (indice de fréquence : 3/5) (Crozier, 2008). Le COCA recense 5 817 occurrences, mais seul le classement des 5 000 premiers mots les plus utilisés en anglais est accessible gratuitement. Avec si peu d'occurrences, il n'en fait pas partie, c'est pourquoi le classement de Collins est employé.

En revanche, les réactions causées par les fautes d'orthographe sont parfois surprenantes, en raison de leur virulence. Beaucoup de membres de forums (ou d'utilisateurs d'internet) n'ont pas oublié la dimension normative de l'orthographe, et de ce qu'ils estiment être un discours « normal » (pour ne pas dire normé), en terme de cohérence, d'organisation ou d'intelligence [8].

4.3.5 Dimension normative

La normativité en ligne, et surtout ses fers de lance, ont été surnommés « grammar nazi » [9] (que je traduis par '*dictateurs de la grammaire*'). Le concept s'accompagne d'une iconographie fascisante, dérivée de leur intransigeance lors des conversations en ligne. Ironiquement, au moins discursivement, leur évocation entraîne forcément un point Godwin [10]. Les phénomènes discursifs numériques sont donc interdépendants.

L'étymologie populaire de *Grammar Nazi* est aussi certainement due à l'hostilité que ce personnage déclenche chez les destinataires, qui prennent les corrections comme étant pédantes et peu pertinentes. Mieux vaut être compris que de taper correctement, pour les détracteurs des ayatollahs de l'orthographe.

8. Il va sans dire que les allusions à ces critères ne servent qu'à améliorer le discours de l'autre, dans une critique toujours constructive
9. Le choix du mot « nazi » par les internautes américains pourrait s'expliquer par un certain manque de sensibilité à la Seconde Guerre mondiale, et à la méconnaissance du régime nazi.
10. La notion du point Godwin ou *reductio ad Hitlerum* a déjà été expliquée en 3.3.3.

Illustration 4.3 – Le logo non officiel des « Grammar Nazi »

Les dictateurs de la grammaire sont en fait polyvalents : ils défendent l'orthographe, conséquence de la syntaxe. Un exemple iconique de leurs corrections est la sélection de *your* au détriment de *you're*. *Your* est plus court, et un homophone, qui commence par les quatre mêmes lettres—bien que la ponctuation et la césure des mots soient différentes.

Un utilisateur de BP est le seul à utiliser l'expression dans le corpus, comme le montrent les illustrations 4.4 et 4.5. Le point soulevé dans ces messages ne concerne pas la graphie en soi, mais la cohérence des propos. Cette référence culturelle numérique est en revanche absente de SC.

Ce fil de discussion, et celui de SC où l'orthographe est évoquée (SCP1) partagent un point commun—encore une fois ironique, discursivement—la mention d'une faute d'orthographe, qui déroge donc à la normalité, s'accompagne presque toujours d'une faute d'orthographe. Les autres utilisateurs ne réagissent pas à l'utilisation de Grammar Nazi—certainement à cause de la méconnaissance de cette expression.

Sur SC, les utilisateurs qui annoncent ne pas être certains de la graphie d'un mot font preuve de précaution : plutôt qu'être pris à défaut, ils préfèrent être francs, parfois avec un smiley d'accompagnement (ou un point d'interrogation, pour montrer qu'ils tentent un bricolage graphique). Ceux qui y répondent en disant « your spelling was correct »sont bien vus, puisqu'ils valorisent l'autre, le féli-

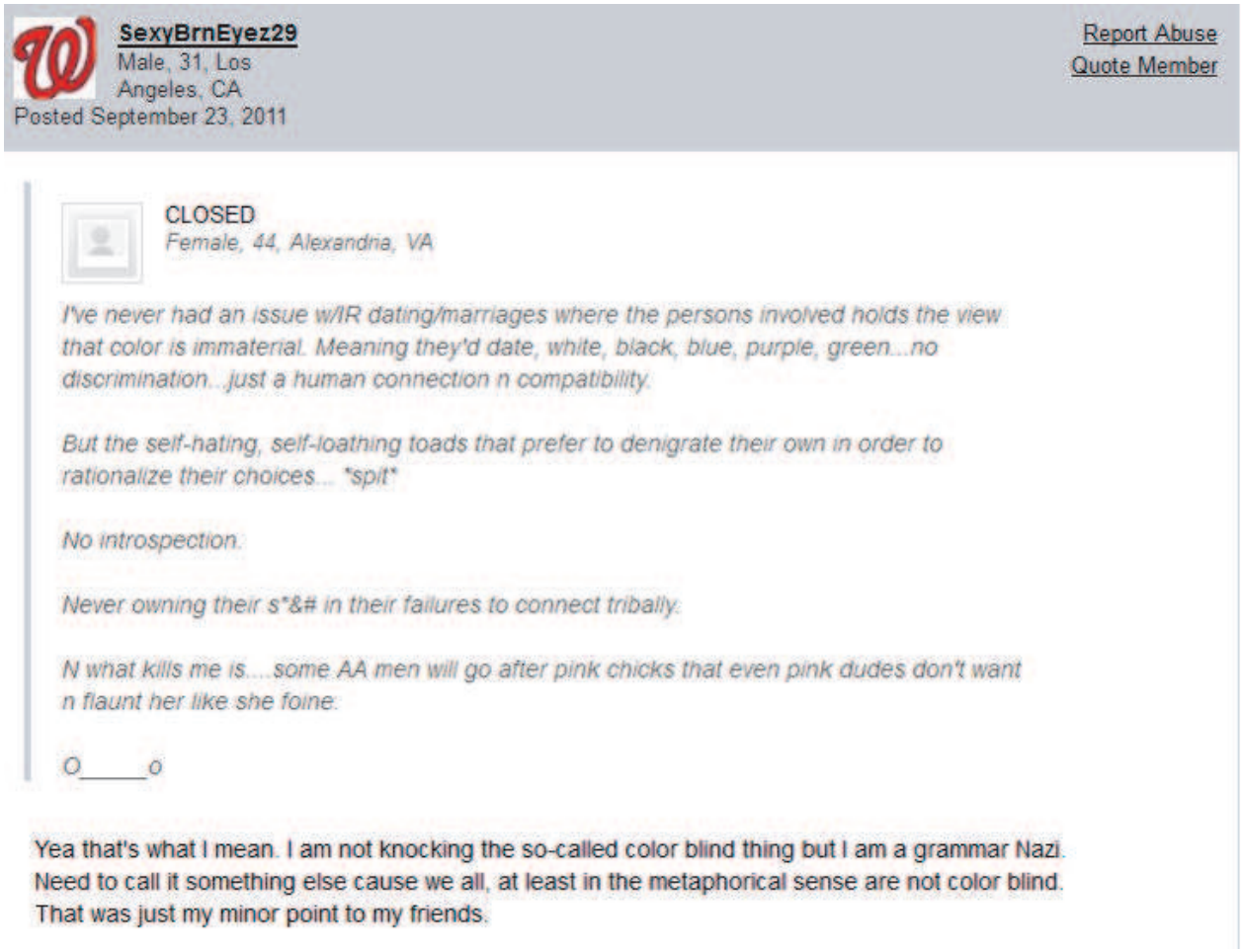


Illustration 4.4 – Utilisation 1 de « Grammar Nazi » (BP12E3357)

citent et assoient également leur position d'experts de l'orthographe.

La loi de Skitt [11] est donc vérifiée dans le corpus, et son interaction avec l'apparition de justiciers de la normativité et du point Godwin, au moins suggéré, montre l'interaction des phénomènes discursifs numériques. Ces justiciers de la normativité ne sont souvent que des utilisateurs, provoqués par d'autres : les modérateurs détiennent leur autorité de leur statut, alors que les dictateurs de la grammaire sont tournés en dérision à cause de la pédanterie ressentie de leurs corrections. Ce personnage conceptuel du web met en perspective la graphie alternative dans un contrat de liberté graphique, mais aussi de normativité. Sur SC, les échanges sont en général assez cordiaux parce que les administra-

11. Pour rappel, ce corollaire de la loi de Murphy est : « Tout message corrigeant les fautes d'un autre message en contiendra au moins une lui-même » (Don & Brad, 2008).

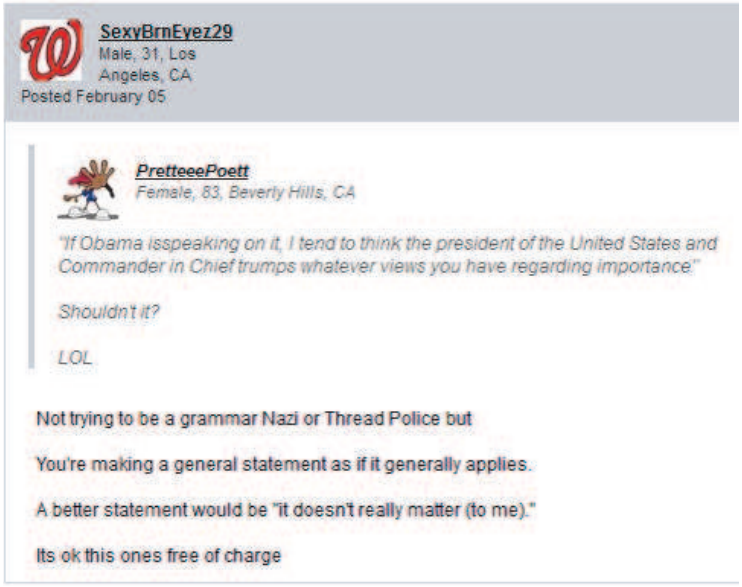


Illustration 4.5 – Utilisation 2 de « Grammar Nazi » (BP12E699)

teurs mettent l'emphase sur le dialogue avec les autres membres [12], ce qui explique aussi la liberté des utilisateurs, qui préviennent les autres que leur graphie est approximative, et qui se valorisent les uns les autres.

Dans l'exemple « haha i only posted twice cause in the first i had a spelling mistake…lmao…glad you liked it :) SC5M40 », l'utilisatrice réagit bien aux commentaires des autres (qui lui demandent pourquoi elle a posté la même chose deux fois, ce qui ne se fait pas non plus selon la Netiquette), mais l'échange est bon enfant, grâce au smiley et à *lmao* (‘*Laughing My Ass Off*’). Sa propre conscience l'a poussée à reposter le même message pour le corriger (elle ignore certainement que les messages peuvent être édités) : la graphie est plus importante pour elle que la continuité de la conversation.

De la même façon, les gens confrontés aux *Grammar Nazis* ont des théories bien personnelles les concernant : « a grammar nazi is someone that clearly doesn't get enough orgasms, whether self given or from someone else… » (BP12E17, message 308) : l'attention à la qualité de la langue serait proportionnelle à la frustration sexuelle. La réponse du berger à la bergère ne se fait pas attendre, et ceux qui se définissent comme Grammar Nazis associent qualité de l'anglais écrit et attraction : « grammar. Yeah, I'm a grammar nazi. I can excuse a "u" every once in a while, but don't write to me

12. À l'exception de la section sur la politique écossaise.

in that stupid chatroom code. Don't talk to me like, you have to, like, add, like, like, in every other, like, word. »(BP12B26, message 1). L'utilisateur a créé un fil de discussion sur les défauts rédhibitoires qui entraînent une rupture instantanée, et la sur-utilisation de tics de langage, ou le *typobonics* vu au chapitre 2, correspondent parfaitement à ce qu'il perçoit comme intolérable. Les fautes (d'orthographe, de vocabulaire…) sont parfois utilisées contre leurs auteurs pour les attaquer, comme on peut le voir dans l'extrait 4.3. Les utilisateurs s'y disputent sur le sens exact de « start », et s'envoient des liens vers un dictionnaire en ligne.

Message	Corps
62	Hmmm if your avatar picture is current, you're obviously not doing p90X unless you've just started. Second, Gold's is ok I rather like to lift during late hours. But um yeah I highly doubt you're doing P90X looking like that.
74	oh, and mr. thesis paper... you might want to slow your eyes down and work on those comprehension skills. 'Started' does mean beginning, which you didn't have to magically discover what i had already said... or did you ?
75	ummm...ya might wanna retract ur post here bcuz look at what it says here 4 the definition of start : http ://dictionary.reference.com/browse/start i mean DAMN...after THAT comment u left...ur pretty much made urself appear 2 b n the right 4rum
76	and look at the definition of beginning http ://www.thefreedictionary.com/beginning GOOD GOOGLY MOOGLY !!!
77	She's got issues with definitions.
78	I didn't realize I should be continuously mindful of punctuation and grammar on black planet, the website that has no complete relevance to myblife save to momentarily excite me. As for my thesis, if you must know its on "late onset of frontal-temporal dementia and hypersexuality in geriatric patients," so, what does my thesis have to do with the OP and punctuation and grammar ? Sub-Zero already corrected you so I wont echo the same point, but its obvious you have an affinity to be a grammar Nazi. Instead of worrying about my lack of editing perhaps you would be smart to not start threads that lack intellectual stimulation such as that Nivea thread which people unanimously had checked you. As far as P90X is concerned, I said "I don't know when you started in relation to that picture," just because you have P90X doesn't mean you're doing it. Just because you say you go to Gold's gym, doesn't mean you do. I am curious to see what you look like now.

Tab. 4.3 – Messages 62 - 78 de Stupid Black Women on Black Planet.....Ironic (BP12E3630)

Pour conclure cette partie sur les fautes d'orthographe, et ouvrir sur le pan des graphies alternatives, je voudrais évoquer un cas entre-deux, où les fautes d'orthographe sont parfois volontaires, ou au moins, non corrigées même si leurs auteurs les ont identifiées. Les justifications des utilisateurs varient, j'ai déjà étudié le cas de ceux qui, pour des raisons extérieures comme le manque de temps ou les sentiments puissants, ne peuvent pas relire leur message. En revanche, ce cas-ci est particulier : « To most of my English teachers you were not very good, so no spell checking, re-writing or editing in your honor. » (BPPP9). Il est donc également possible de faire des fautes d'orthographe…par vengeance. C'est une excuse originale pour le manque de relecture teintée d'humour, sinon de mauvaise foi, qui conclut une description de profil assez longue.

La faute d'orthographe volontaire ou non, excusée, décriée ou justifiée est donc présente dans tous les esprits, et dans bon nombre d'écrits internet. L'autre face de ce Janus de l'écrit est la graphie alternative, qui partage certaines caractéristiques formelles avec la faute d'orthographe. Cependant, les buts et causes des graphies alternatives sont tout autres, ce que je vais explorer dans la partie suivante.

4.4 Graphies alternatives

Les graphies alternatives sont nées du désir (ou des désirs complémentaires) des utilisateurs de forums de s'exprimer en changeant volontairement la graphie des mots, afin qu'ils servent plusieurs buts, dont je vais dresser une typologie dans cette partie.

Les graphies alternatives sont parfois évitées par certains membres, qui estiment que les autres ne comprendraient pas ces graphies inventées. C'est également un raccourci cognitif, dans le sens où les graphies standard sont maîtrisées : il n'est nul besoin d'y penser pour savoir écrire la plupart des mots courants, et les détracteurs du *typobonics* ou de l'*e-bonics* évoquent cet argument pour critiquer les graphies alternatives.

4.4.1 Utilisations propres à un forum

L'utilisation d'un style graphique personnel, mais surtout la tachygraphie peuvent être importantes dans les discussions en ligne : ni BP, ni SC ne sont normatifs quand il s'agit de l'orthographe, et les utilisateurs adaptent leurs graphies à leurs envies. Sur SC particulièrement, les utilisateurs peuvent adopter des graphies de leur choix pour se démarquer. En revanche, certains usages graphiques et lexicaux sont dus entièrement à la plateforme (comme « The Pulse », ou « What's New ? », déjà examinés en 3.5.0.1). Un souci stylistique qui se rapprocherait d'une calligraphie virtuelle semble se distinguer dans certaines communications sur Internet, notamment dans le cadre de blogs (surtout ceux des adolescents, qui, par exemple, en France, sont hébergés par *Skyblog*, ou, dans une moindre mesure, sur *MySpace* et autres plateformes qui offrent une opportunité de s'exprimer librement). En effet, la communication sur internet, bien souvent seul biais de rencontre entre les différents acteurs (qu'ils soient lecteurs, commentateurs, ou même participants à la même plateforme), est aussi une façon d'exprimer, par écrit, sa propre personnalité. C'est surtout un moyen de se forger une identité particulière, presque uniquement confinée au monde virtuel.

Une autre façon de se dire à travers ses écrits et ses choix graphiques est dans la récusation de la légitimité de l'autre, dans le cas de discours racistes par exemple. C'est la raison pour laquelle je vais étudier ce que j'appelle très généralement « parole blanche » et « parole noire » dans la sous-section suivante. Il n'est plus possible d'obtenir des données quantitatives sur les membres des sites étudiés, mais certains indices exposés dans la partie suivante laissent à penser que les auteurs des textes choisis sont Afro-Américains ou Blancs.

4.4.1.1 Etude de cas : la parole noire et blanche sur internet

Pour examiner les écrits blancs et noirs sur internet en général, j'ai choisi d'étudier des discours qui n'ont pas été créés sur BP, et qui relèvent donc d'opinions plus tranchées et moins modérées que celles sur le forum de niche choisi. La thèse que je voudrais expliquer dans cette étude de cas est la suivante : les tensions raciales, surtout dans les sphères américaines, sont contenues dans les graphies

alternatives employées par les victimes ou les responsables de ces tensions.

Pour cette étude, j'ai opté pour un extrait de *Topix*, un site hybride entre le site de questions (à la *Yahoo Answers*) et le forum de discussion, regroupant les informations par sections et thèmes. Il contient une sous-section afro-américaine, où les utilisateurs s'expriment librement. Les interventions choisies sont rédigées par des utilisateurs s'identifiant comme Afro-Américains, sur leur profil personnel. À l'opposé, après avoir choisi d'étudier Stormfront (forum de suprémacistes blancs, fondé par un ancien dirigeant de chapitre du Ku Klux Klan), j'ai plutôt opté pour une étude de contenu de l'*Encyclopaedia Dramatica* (ED) recopié sur BP, contenant du discours raciste [13]. Quelques requêtes seront néanmoins effectuées sur Stormfront, mais j'ai jugé qu'aller sur un site raciste aussi surveillé (donc, où le contenu écrit est scruté pour ne pas dépasser un cadre toléré par la liberté d'expression) n'était pas aussi représentatif graphiquement que l'« humour » très subversif de l'ED. Les caricatures sur *Stormfront* semblent plus distillées dans des textes haineux. J'ai écarté les discours extrêmistes afro-américains, et j'écarte aussi les discours extrêmistes des suprémacistes blancs, parce que ces discours n'apportent rien de plus à l'étude que les substituts que j'ai choisis. Les rhétoriques qui y sont utilisées nécessiteraient un autre type d'exploration en regard avec les théories qu'ils défendent. Le contenu racisé choisi est validé par BP, qui n'a pas effacé le profil, parce qu'il n'a certainement pas été lu par les autres membres.

Les lexèmes choisis pour cette étude sont *whitemanitus*, *kneegrows*, ainsi que les indices linguistiques disséminés sur chacune de ces pages (comme *massa*, ou l'analyse de l'illustration raciste sur la grammaire afro-américaine). Les membres de *Topix* écrivent presque tous en utilisant un style standard, qui ne dérive que marginalement dans l'écriture oralisée. Elle n'est teintée d'AAV que lorsqu'il s'agit d'un discours raciste, dirigé contre les Afro-Américains.

Dans le fil de discussion « Lets talk about white people » de Topix (16/03/2009), l'un des messages (#7) au contenu raciste contient au moins deux occurrences de mots vus comme étant afro-américains, « da »remplaçant « the »et « massa »pour « mister ». L'alternance de codes linguistique, utilisé pour insulter au maximum les afro-américains du site, démontre une conscience des autres locuteurs de

13. L'ED contient des pages telles que « There are no niggers on the Internet », une page acerbe et provocante qui liste les raisons pour lesquelles il n'y aurait pas de Noirs sur internet (tout comme il n'y aurait pas de femmes ou d'homosexuels). Elle contient aussi une liste de mèmes racistes. Adresse : <https ://encyclopediadramatica.es/There_are_no_niggers_on_the_Internet>, dernier accès le 14/07/14.

l'anglais vis-à-vis de ce sociolecte, exploité dans ce cas précis par souci de parodie et d'insulte :

> This whole site is for Negroes that FEAR white people, didn't you know that ?
> Listen to any Black "talk show", and what you will hear all day long is about, "racism" and da white man.
> Exactly who put the white man on a pedestal ?
> Negroes did !
> Which means..... whites still have power over Negroes without them even trying !
> "Whtemanitus" has affected the brains of many Negroes !
> I say, let's bring back slavery, so the Negroes will feel contented.
> Can you say "Massa" ?

Dans cet extrait figure une instance de néologisme, utilisé sur internet en assez grande proportion pour faire passer une idée plus efficacement. « whitemanitus », mis entre guillemets par le rédacteur du message, est en effet créé à partir de deux mots, *white* et *man*, auxquels le suffixe <-itus> a été accolé. Il semblerait que le scripteur parlait d'une maladie affectant les Afro-Américains, et que, par conséquent, le suffixe <-itis> était ce qu'il voulait signifier.

Une recherche sur *Google* du terme « whitemanitus »au 4 janvier 2010 ne donne que deux autres résultats, issus du forum afro-américain de *Topix* [14], et ce chiffre n'a pas évolué en septembre 2014. Les deux utilisateurs viennent d'états américains différents, « whitemanitus »n'est donc pas une utilisation argotique propre à une région [15]. Dans les deux cas, il s'agit encore de la description de cette affliction semblant toucher la population afro-américaine et leurs rapports avec la population blanche des Etats-Unis. Dans le premier extrait, écrit par « Dissident », qui vient de l'Ohio, le terme n'est pas mis entre guillemets.

> The "older generation" did what they had to do in order to make it better for YOU !
> Yet you whine about the damned white man all of the time, which leads me to believe that the white man has you under his control !
> it's very evident by only looking at the posting about black vs white.
> Some kneegrows,(and I'm being polite) seems to have whitemanitus of the brain !

Mais dans le second, écrit par Jawga Boy, le mot *whitemanitus* est mis entre parenthèses. Cela semble confirmer la nature néologique du terme, au moins dans le lexique de cet utilisateur :

14. Disponibles à http ://www.topix.com/forum/afam/TL1SIEKM2QQKKPP8P et http ://www.topix.net/forum/afam/-TH3LO2IP4DQKICSSQ/p3
15. Le premier à écrire ce terme vient de Chillicothe, dans l'Ohio et le second d'Atlanta, en Géorgie.

> Of course they are, just read the posting on this site ! It's always about the white man who they fear.
> Why else would any damned fool post sh !t every day about another group if they didn't FEAR the other group !
> They don't have ANY power at all, thus all the constant posting about,"you're a racist" nonsense !
> A classic case of "Whitemanitus" !

Afin de déterminer exactement ce que ce terme contient, et puisque l'UD n'a pas d'entrée pour ce mot, j'ai décidé de poser la question directement aux participants du forum Topix, dans un fil de discussion créé pour l'occasion [16] J'ai reçu 9 réponses au 7 janvier 2011 [17], dont une de Jawga Boy. Ces réponses sont réparties en trois catégories : comportements d'une population envers une autre, et maladie. Certaines réponses contiennent plusieurs possibilités :

Sentiment ou comportement provenant d'un Afro-Américain envers la population blanche	3
Sentiment ou comportement provenant d'un Caucasien envers la population afro-américaine	2
Infection ou maladie	5

Tab. 4.4 – Réponses au sondage sur « Whitemanitus »

Ce terme n'est pas clair, et semble contenir un concept et son inverse ; il est à noter que la majorité des réponses inclut une analyse du mot par son suffixe, <-itus>. Cela semble aller dans le sens de l'analyse consciente du néologisme. Il existe un autre exemple sur la plateforme de l'accolage d'un mot et du suffixe <-itis> : « dumbassitis ». La confusion entre <u> et <i> dans la fin de ce mot s'explique par sa position en syllabe réduite. Il n'y a pas d'exemple d'utilisation pour décrire un être humain (à la *Homo Erectus*…), et aucun des intervenants n'en a parlé.

L'importance de l'oralité sur internet est également mise en exergue dans cet extrait, grâce à la présence importante d'injonctions et d'éléments n'appartenant d'habitude qu'à l'oral, tels que « I say ». Les retours à la ligne fréquents peuvent également être interprétés comme des équivalents graphiques à des pauses plus longues à l'oral (ceci semble être le cas, surtout après les points d'interrogation, après ses questions rhétoriques).

Dans le titre des fils de discussion, la population de l'*African American Forum* de *Topix* semble

16. Disponible à l'adresse <http ://www.topix.com/forum/afam/TJSJ0R002AFNT9KF2>, dernier accès le 14/05/14.
17. Au 6 septembre 2014, il n'y a pas de nouvelles réponses.

se caractériser par des débats plutôt manichéens, au moins dans leur expression première. Ces titres (analyse faite en prenant les cent derniers fils de discussions commentés, au 26 mars 2009) laissent apparaître une grande proportion de questions, à l'inverse de nombreux forums, ou les sujets sont souvent posés comme des thèmes de conversation, construits dans des clause réduite. En prenant les cent premiers titres des sujets de fils de discussion, au 5 août 2009, on ne voit que 37% de questions, par opposition au recueil du 26 mars 2009 par exemple.

À l'inverse, certains utilisateurs [18] qui sont visiblement suprémacistes blancs, ou racistes, arrivent à garder un compte sur BP. Ils ne participent pas aux discussions et le profil choisi est fantôme (plus de connexions après la création). On peut imaginer que ce type de profil est créé pour aller sur BP, voir ce que les Afro-Américains s'y disent ponctuellement. L'absence d'activité, à part la création initiale, les rend invisibles aux administrateurs et modérateurs, qui n'effacent pas ce genre de profils de la plateforme.

Nigger Language
Since all %#&@$!s are too lazy to learn proper English, they are practically impossible to

Illustration 4.6 – Image accompagnant « Nigger Language »—originellement légendée « Black Honor Student »

understand. Their "language" is in fact the epitome of laziness. An example in modern ebonics is "ackin" for "acting" because its just too much effort moving the tongue to the roof of the mouth to make a "t" sound. Sometimes the %#&@$! is simply too **derro** to learn proper English. Example : "hello my name is jamie" converts to "YO YO YO wazzup %#&@$!es and %#&@$!ets me names jamie c, respectaz or ill pop a cap in yo' white %#&@$! foo' !" Typical %#&@$!'s **burns** consist of putting random words together. Examples include "yo ol' burnt skittle lookin azz" or "poptart frostin' bald head." After the%#&@$!s burn on

18. Il est difficile de rendre compte du degré de sérieux employé par l'ED—je pars du postulat, certainement naïf, que, si l'ED est humoristique, et joue sur la très fine limite de la provocation politiquement incorrecte, l'utilisateur qui recopie et s'attribue la citation considère sérieusement les Afro-Américains comme inférieurs (aux blancs).

> each other every other %#&@$! laughs as loud as possible for 20 minutes with their hands in front of their mouths in an idiotic way and then start shooting each other over drugs. The last %#&@$! standing will then take all the drugs in epic win. Hopefully, he'll kill himself after realizing his homies are dead, causing lulz. [‘*lols*’]
> **Nigger Speak**
> Not to be confused with English, for whites, Engrish, for Asians, El English, for Mexicans, Nigger Speak is it's own form of language, containing many profanities and numerous nonsense/useless slang words in a single sentence. On the internet, this stupidity is magnified at least 100 times and usually leaves all other races in a wtf moment. Unlike 13375p34k, where the words have minor purpose, Nigger Speak has no use whatsoever and is the Nigger's way of getting back at the white man. The normal conversation between two fluent in Nigger Speak goes something like this “Smh, lmlmlml, dat sht hd mi rolin on duh fukin flor” “Omkfc, ikr” [‘*Shaking My Head, Laughing Mad loud, that shit had me rolling on the fucking floor. –Oh my KFC, I Know Right.*’ (BPPP2)

Cette citation (ainsi que l'ensemble de ce profil a été tiré de l'ED) ne montre pas la façon dont les utilisateurs de BP s'expriment…dans une certaine mesure. Certaines délimitations linguistiques sont plutôt correctes, mais la satire est plus qu'exagérée. La caractérisation de certains traits phonologiques, comme la réduction de groupes consonantiques, est bien observée dans la description de « ackin » mais n'apparaîtrait qu'en fin de mot : si on peut imaginer une prononciation /æk/ pour *act*, *acting* est prononcé /ˈæktɪŋ/. Ce profil décrit également l'occlusion du phonème /ð/ (décrit dans le tableau 2.12) dans le mot « duh » (‘*the*’), et la dévélarisation de /ŋ/(*ackin, lookin, frostin', rolin, fukin*). Ces observations sont faciles à faire, dans le sens où ce sont des traits phonologiques saillants de l'AAV.

Discursivement et lexicologiquement, certaines expressions ne collent pas au tableau, dans cette description. Par exemple, *derro* est issu de l'argot australien. Ce mot désigne les jeunes les moins favorisés et c'est l'équivalent des *neds* écossais, des *chavs* anglais, ou des *trailer trash* américains. Aucun de ces mots ne désigne particulièrement les Noirs (au contraire, les *trailer trash* sont plutôt caucasiens).

Le texte évoque les *burn* des joutes verbales qui consistent à envoyer des attaques personnelles. Si le terme est un peu enfantin, voire inexact (on parlerait plus volontiers de *diss* pour désigner les joutes verbales afro-américaines), la caractérisation est plutôt correcte, et décrit ce que j'ai proposé d'appeler des « mégasyntagmes d'insultes » (Goudet, 2013c) utilisés par les Afro-Américains (mais pas uni-

quement eux), comme « yo ol' burnt skittle lookin azz ». Mais les termes utilisés sont incorrects (*ass* sert plus souvent comme intensifieur, « rich ass nigga »), la référence à la nourriture (*skittle, poptart*) est étonnante et plutôt absurde. Il y a une autre occurrence de référence plutôt raciste, « omkfc », qui joue avec OMG ('*Oh My God*') et le transforme en « Oh My KFC » [19].

Le stéréotype inclut également l'utilisation d'abréviations (« SMH », déjà évoqué en début de chapitre) et « LM » ('*Loud Mouth*') ou d'expressions typiques, comme *foo* ('*fool*', qui rappelle la phrase de Mr. T, « I pity the foo' ») et « pop a cap » ('*shoot a gun*'). La fin du texte et la conversation inventée sont absconses, et ne représentent pas du tout le type de conversation qui pourrait être échangée. La caricature emprunte à des éléments déjà caricaturaux (Mr. T a surexploité cette phrase dans ses divers rôles, et elle est devenue sa réplique culte (M. Lee, 2010)). Encore une fois, des éléments de réel, et des mots utilisables sur BP se retrouvent superposés dans un discours exagéré.

Le parallèle avec le *leetspeak* (« 13375p34k ») rappelle l'*e-bonics* et le *typobonics*, avec une nuance importante : selon ce texte, l'AAV serait inventé par rancœur envers les Blancs, et contiendrait des insultes inutiles. Mis à part le stéréotype et l'amalgame avec certains parlers jeunes, qui utilisent une plus grande proportion d'insultes et de jurons de toutes sortes (Allan & Burridge, 2006), cette réflexion montre en fait le fantasme de l'homme non-Noir sur les Noirs. Les tropes de consommation de drogue et l'imbécilité qui semble caractériser les Noirs, ici, ressemblent aux caricatures évoquées sur Topix. Le « Blackmanitus » de cet extrait, s'il repère correctement certains éléments discursifs, est caricatural : le discours sur BP est loin de ressembler à ces écrits.

Les Mexicains ne sont pas épargnés, avec l'évocation d'(*El English*, alors qu'on pourrait parler de *Spanglish*) ; l'*Engrish* (évoqué plus loin dans ce chapitre, en 4.4.3.2) est censé être le parler des Asiatiques, alors que l'expression désigne surtout des erreurs de traduction ou des problèmes phonologiques pour les locuteurs de japonais qui s'expriment en anglais. Ce sont donc des attributions erronées à des variétés et différents accents, plus que des pratiques linguistiques à part.

Internet sert à explorer ses fantasmes (nés de la peur et de l'inconnu) sur les autres : soit la peur de l'esclavagisme, soit celle des autres ethnies. La communauté de pratique, sur les sites afro-américains

19. Les chaînes de *fastfoods*, surtout celles qui servent du poulet, sont souvent utilisées pour ridiculiser les Afro-Américains : dans l'imagerie raciste stéréotypique, le poulet (et les pastèques) sont leurs aliments préférés.

(BP n'est pas épargné par le fantôme de l'esclavagisme) tourne souvent autour de la domination des Blancs sur les Noirs, et des Noirs qui se laissent dominer (les « kneegrows » qui parleraient au Blanc en utilisant « Massa »). La graphie alternative a un rôle de vecteur assertif, où les niveaux connotatifs évoqués s'articulent en plusieurs points. Il y a un détournement de la forme standard <OMG> → <OMKFC> ainsi qu'une imitation dévalorisante des parlers qui assigne son énonciateur supposé à la sphère de l'esclavage. Ces thèmes fréquents sont importants pour cerner la peur, et donc le discours exagéré et caricatural qui revient sans cesse sur la plateforme.

SC est épargné de ce type de discours. Les velléités indépendantistes ne se couplent pas avec un discours caricatural à l'encontre d'un certain groupe (pros ou antis ; voire contre les Anglais). Les arguments ne sont pas polarisés sur la question identitaire, mais plus souvent sur les arguments utilisés.

Dans cette partie, les pratiques exclusives à internet ont été évoquées, surtout ses acronymes (*SMH*…). La partie suivante s'attache à la description des phénomènes langagiers que l'on va trouver dans BP et SC, et partout sur internet. Il s'agit en quelque sorte de la *koiné* d'internet, dans le sens où ces pratiques sont reconnues comme standard dans les communautés internet.

4.4.2 Netspeak

Entre langues standards utilisées en ligne et culture internet, le netspeak se pratique partout, si tant est qu'on ait des affinités avec la culture internet (surtout pour les indigènes numériques [20], soit les gens de la génération Y qui ont grandi avec un accès à internet (Papi, 2012)). Le netspeak est un terme très général inventé par D. Crystal, qui définit les particularités des échanges en ligne comme étant l'ensemble de pratiques linguistiques (“mots, idiomes, particularités graphiques et grammaticales”), caractéristiques de la communication en ligne (McFedries & Logophilia Limited, 1995). Le netspeak se pratique en ligne parce que le discours n'y est pas aussi surveillé que dans les autres médias écrits. Selon D. Crystal, l'écrit est « typiquement circonscrit dans l'espace, construit, décontextualisé visuellement, communicationnel de fait, minutieusement structuré, révisable à merci et riche graphi-

20. « Digital natives » en anglais.

quement »[21]. Crystal compare le netspeak à l'oral, tout en le reconnaissant comme étant très différent (à cause de l'absence d'indication mimo-posturo-gestuelles (j'ai déjà parlé de la comparaison entre l'oral et l'écrit, et de leur place respective au chapitre 3). En revanche, il inclut dans sa définition la lenteur des réactions, beaucoup moins prévisibles dans des chats (à cause de soucis techniques, qui peuvent provoquer la déconnexion de la plateforme). En revanche, ces déconnexions sont accompagnées du sentiment de « conversation perpétuelle » (Bays & Encrevé, 2001). Les fils de discussion, tout comme les chaînes de chat telles celles d'IRC accumulent des réponses, auxquelles l'on peut réagir le lendemain, ou à n'importe quelle date ultérieure (dans certains forums français, on appelle la pratique de répondre à un fil de conversation qui n'a pas eu de réponse récente « déterrer » ce fil).

Les éléments les plus intéressants du netspeak sont les exceptions par rapport à la langue écrite non technodiscursive (Paveau, 2013c, 2013d). L'hyperexpressivité (« aaah » ou « hello ! !!!!!! !!!!!! !!! ») est caractéristique du netspeak : laisser son doigt sur une touche est beaucoup plus facile et spontané que serait d'écrire la même chose à la main.

Les « didascalies » sont importantes au discours en ligne, pour les mêmes raisons qu'elles sont importantes dans le théâtre : l'absence d'indices mimo-posturo-gestuels ou de corporéité physique permet l'apparition d'une spatialisation alternative. Par exemple « *rolls back into my corner* » (BP12E763) est la dernière phrase d'une utilisatrice, pendant une dispute. Elle indique un mouvement qu'elle ne ferait pas hors-ligne, mais qui, dans le contexte, signifie « j'arrête de parler ». Toujours dans le même fil de discussion, « *talking like Urkel* » est beaucoup plus direct : la phrase qui suit doit être lue mentalement avec la voix d'Urkel[22]. Ces didascalies s'inscrivent dans l'espace inédit de la communauté ; elles sont aussi orales et physiques que la communauté l'est. Elles font partie intégrante des écritures natives du web, dans le sens où, même si elles ne sont jamais actualisées hors-ligne, elles le sont dans l'esprit des co-énonciateurs.

21. Traduction personnelle de « typically space-bound, contrieved, visually decontextualized, factually communicative, elaborately structured, repeatedly revisable, and graphically rich » (Crystal, 2001, p.28).

22. Stephen Urkel, plus souvent appelé « Urkel », est un personnage de la sitcom « Family Matters » (‘*La Vie de Famille*’) diffusée dans les années 1990. Il représente le stéréotype du *nerd*, ou ringard inadapté social, très maladroit et sa voix est très nasillarde.

4.4.3 Mèmes

Les mèmes sont les plus petites unités de contenu culturel, et la théorisation d'un tel phénomène est apparue en 1976 dans « The Selfish Gene » de R. Dawkins (Dawkins, 1976). Ces mèmes se propagent et s'adaptent de la même façon que les cellules, en suivant les mêmes mécanismes. Ce modèle est particulièrement adapté à l'étude de phénomènes culturels sur internet, ainsi qu'à la détermination de communaulectes. La nature essentiellement protéiforme des mèmes contribue à leur succès et à leur propagation sur internet. Le mème se caractérise non par sa forme, mais par sa duplication, et le fait qu'il soit reconnaissable par les gens versés en culture internet. Le mème peut être une phrase qui semble ne pas avoir de sens, comme « I'm 12 and what is this ? » [23]. C'est un commentaire utilisé en réponse à des vidéos ou à du contenu jugé étrange ou surprenant (comme « wat » ou « I don't even. » [sic]). Cette phrase renvoie à 64 400 résultats (hors commentaires *Youtube*, puisqu'ils ne sont pas comptabilisés). La recherche sur *Youtube* uniquement renvoie, elle, à 396 000 résultats [24]. La survie du mème tient à son caractère essentiellement original qui pousse à le reproduire. C'est en cela qu'il se rapproche des graphies alternatives : ces graphies alternatives fréquentes sont les candidats les plus heureux de la sélection faite par les membres de SC et BP. C'est pourquoi comprendre ce pilier de la culture internet est primordial pour appréhender les graphies alternatives et la façon dont elles circulent sur les sites étudiés. L'institutionnalisation des mèmes et des communaulectes passe par un processus de validation par les utilisateurs des plateformes où ils sont employés. Ces constructions sont sanctionnées par l'appui apporté par les autres, symbolisé par leurs reprises des phrases ou éléments mémétiques.

Ils peuvent se présenter sous la forme de combinaisons de texte (jusqu'à la phrase fétiche ou slogan, déclinée sous de multiples formes), d'image, (jusqu'au choix de la police de caractères et du type de fond utilisé), ou vidéos (souvent, des arrangements avec musique et déformation de la voix des intervenants), voire de concepts communs (relecture de clips musicaux et autres avec chant de ce qui se produit à l'écran de façon littérale, créations de versions de chansons « mal entendues » (*misheard*),

23. Mème confirmé par des sites de surveillance de mèmes comme *Know Your Meme*. Adresse : <http ://knowyourmeme.com/memes/im-twelve-years-old-and-what-is-this>. Dernier accès le 14/07/14.
24. Requête « "I'm 12 and what is this" site :www.youtube.com » le 16 août 2014.

qui jouent sur les mondegreens et les à peu près).

Les mèmes graphiques partagent la même mise en forme : présence d'un personnage, animal, humain ou dessiné (*RageComics*) ; fond neutre ou roue de couleur, texte encadrant l'illustration écrit dans la police Impact (normalement en blanc, entourée de noir), réparti entre le haut et le bas de l'image afin de donner plus d'impact à la moquerie ou au trait d'humour. L'illustration 4.7 montre un petit échantillon des types d'illustrations, et fournit des patrons vierges à utiliser pour composer de courtes bandes dessinées humoristiques.

Dans cette partie, je vais parler des « mèmes internet » et des mèmes —selon leur acceptation par la Société Francophone de Mémétique, qui utilise une délimitation du phénomène qui me semble judicieuse. La différence entre l'acceptation populaire (ce que j'appelle « mèmes internet ») et les mèmes culturels généraux, comme chanter une version de chanson d'anniversaire connue par tous, parce qu'apprise au préalable, par exemple. En m'inspirant de la définition d'un méméticien, je délimite le mème comme étant un phénomène culturel (donc non-naturel ou inné) à forte dimension linguistique, reconnaissable, reproductible et imitable (Jouxtel, 2009). C'est l'ensemble de ce qui se fait au travers de nos pratiques : les techniques de corps de M. Mauss sont autant d'exemples de pratiques reproduites par imitation d'un pair : savoir se laver ou s'habiller relève de pratiques culturelles qui doivent beaucoup au *hic et nunc*. Savoir mettre un *hanbok* coréen ne relève pas des mêmes techniques de corps que de mettre un corset en Espagne à la Renaissance. Les mèmes sont transpersonnels, et événementiels. L'instanciation qui m'intéresse dans cette thèse est à la fois la pratique des mèmes internet, et, par extension, des mèmes. Ce que les utilisateurs de forums, d'internet en général indiquent dans leurs graphies ou leurs choix lexicaux (et qui seront développés ici) sont autant de traces d'exposition à un phénomène linguistique et donc mémétique, dont j'essaye de dresser une typologie.

Les mèmes internet sont des réplications *ad nauseam* d'éléments ludiques, amusants et publics, donc, par essence, non communaulectaux (dans le sens où le communaulecte montre l'appartenance à une communauté virtuelle en particulier). Les mèmes internet sont tout de même des marqueurs d'appartenance culturelle parce que leur maîtrise et leur propagation va de pair avec la connaissance d'internet, et de ses phénomènes discursifs. Les mèmes internet sont viraux, partagés dans les pra-

Illustration 4.7 – Extrait de la page de Know Your Meme qui recense les Rage Comics

tiques communautaires. Ils peuvent comporter un élément de multimodalité (audio, vidéo, image par exemple) qui rajoute à leur sémioticité, ils sont éminemment citationnels, et c'est là leur mode de propagation premier. Ils sont également industrialisables : les sites de curation de mèmes proposent une option de création d'un mème visuel à l'aide d'outils qui facilitent la mise en forme. Ici, il n'est point question de l'authenticité d'une œuvre, ou de son placement dans le temps et l'espace, mais au contraire, de la quantité de sa reproduction ou de ses déclinaisons (Benjamin, 1936). Les copies multiples du mème ne le dépouillent pas de son authenticité, à la manière d'une oeuvre d'art qui jouit

d'une « aura » ou d'un supplément d'âme. Les reproductions mémétiques restent dans le même espace (bien que le temps soit une caractéristique variable), et le mème est authentifié, acquiert cette aura par sa reproduction. Les mèmes à image présentés dans cette section ne sont qu'un sous-exemple des mèmes internet, mais ils sont typiques de la façon dont internet a modifié la création d'éléments de connivence, ou humoristiques. La typographie et l'iconographie des mèmes sont aussi codifiés, et ces contraintes formelles aident à les identifier immédiatement (comme les plaisanteries type « knock knock – who's there ? »). Les mèmes internet ne sont que des sous-catégories des mèmes culturels, mais dont la création et la multiplication sur internet rend protéiformes. L'imitation réciproque des membres de communautés n'a pas besoin de trope issu d'internet pour se produire régulièrement.

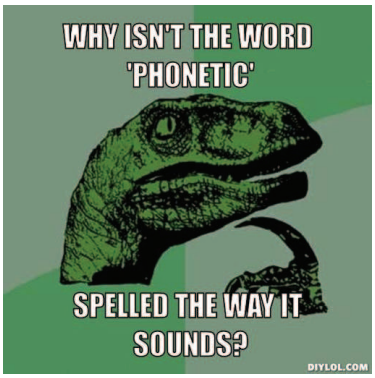


Illustration 4.8 – Un exemple du mème « Philosoraptor », auteur inconnu (Source : www.diylol.com

Les mèmes se disputent l'attention des utilisateurs d'internet, et de chaque concept important découle des centaines de déclinaisons autour de thèmes semblables. 'Philosoraptor', par exemple, est toujours montré en train de se poser une question existentielle, ou de soulever un paradoxe. Ce mème, qui date de 2008, a été consulté près d'un million de fois sur *Know Your Meme*, et la recherche « Philosoraptor » sur Google Image recense 5.3 millions de résultats (décembre 2011) [25].

Les mèmes peuvent être partagés par tous les utilisateurs d'internet, de façon transversale ou peuvent être spécifiques à une communauté, ce qui est le cas des communaulectes et des apparitions discrètes de graphies alternatives ou de néologismes propres à une communauté. Ils ne se cantonnent pas uniquement à internet, et peuvent provenir de n'importe quel contenu culturel (par exemple, Jay-Z, un

25. 3.4 millions de résultats en septembre 2014, non parce que ce mème a disparu, mais parce qu'il descend dans le classement algorithmique de Google Image

rappeur américain, a contribué en partie à l'utilisation du mot *nigga* sous cette forme de façon large au 20ème siècle, en l'utilisant dans ses chansons et en parlant de lui en utilisant ce léxème).

Les figures mémétiques prototypiques ont souvent un nom, ou appartiennent à une catégorie plus large, comme les *lolcats* [26], déclinés sous de multiples formes, et eux-mêmes faisant partie de la catégorie des mèmes motivant des animaux. Dans cette catégorie, on trouve également les *advice animals*, Doge (qui est un jeu linguistique qui utilise toujours des photos de chiens de la race Akita Inu) et les chats célèbres (Grumpy Cat, Lil Bub, Colonel Meow ont tous des déclinaisons mémétiques à part entière). Ces mèmes sont immédiatement reconnaissables par les internautes chevronnés, qui sont au moins capables de tracer certains éléments des images les constituant.

Les mèmes peuvent être basés sur des figures archétypales qui aident à identifier immédiatement le type de narration en jeu, ou être issus d'événements récents (capture d'images ou de vidéos de phrases choquantes ou étonnantes, réutilisées ou remixées avec ajout de musique). Ils exploitent souvent des thèmes provoquants comme la pédophilie, la violence ('*Insanity Wolf*'), ou s'inspirent de la vie réelle et tournent en dérision des événements triviaux et des sentiments simples auxquels les lecteurs peuvent s'identifier ('*Socially Awkward Penguin*', '*Always Alone*', deux exemples de personnages qui n'arrivent pas à s'intégrer à la société, désillusionnés, généralement malheureux et isolés, '*Foul Bachelor Frog*', la figure archétypale du *nerd*, à l'hygiène déplorable, intéressé essentiellement par les jeux vidéo et la masturbation).

La fréquence d'apparition est calculée en fonction du volume de recherches par mots-clefs sur *Google*, en utilisant la fonctionnalité « Google Insight for Search » qui documente le nombre de recherches d'un ou plusieurs mots-clefs et fournit un graphique qui montre l'évolution chronologique du nombre de recherches effectuées. Plusieurs sites répertorient les mèmes, et archivent les images et macros envoyés par les lecteurs du site correspondant à chaque mème (*Know Your Meme* [27], *Memebase* [28], *Memegenerator* [29]). Ces sites permettent aux visiteurs d'ajouter du contenu, à la manière

26. Un équivalent français du mot *lolcat* n'a pas été arrêté par la Délégation de la Langue Française (les efforts de WikiLF vont vers « chamusant » mais pour les propos de ce travail, la dimension culturelle du nom lolcat est conservée, et les traductions des noms de mèmes ne seront indiquées qu'à titre informatif.
27. Accessible à l'adresse <http ://knowyourmeme.com/>
28. Accessible à l'adresse <http ://memebase.com/>
29. Accessible à l'adresse <http ://memegenerator.co/>

d'un Wiki [30], en écrivant des articles sur de nouveaux mèmes, ou de générer leurs propres mèmes, en utilisant des photographies et du contenu populaire ou déjà exploité. Ils peuvent également contribuer au prestige d'un mème en votant pour ou contre lui (*Meme Generator* classe les concepts et macros en huit niveaux, du « God Tier » [31] où seuls 6 mèmes sont répertoriés, au « Fail Tier », qui compte les mèmes les moins intéressants, et les moins appréciés. Les niveaux contiennent entre 27 et 6 entités à part entière, sans compter leurs variantes et sous-entrées.

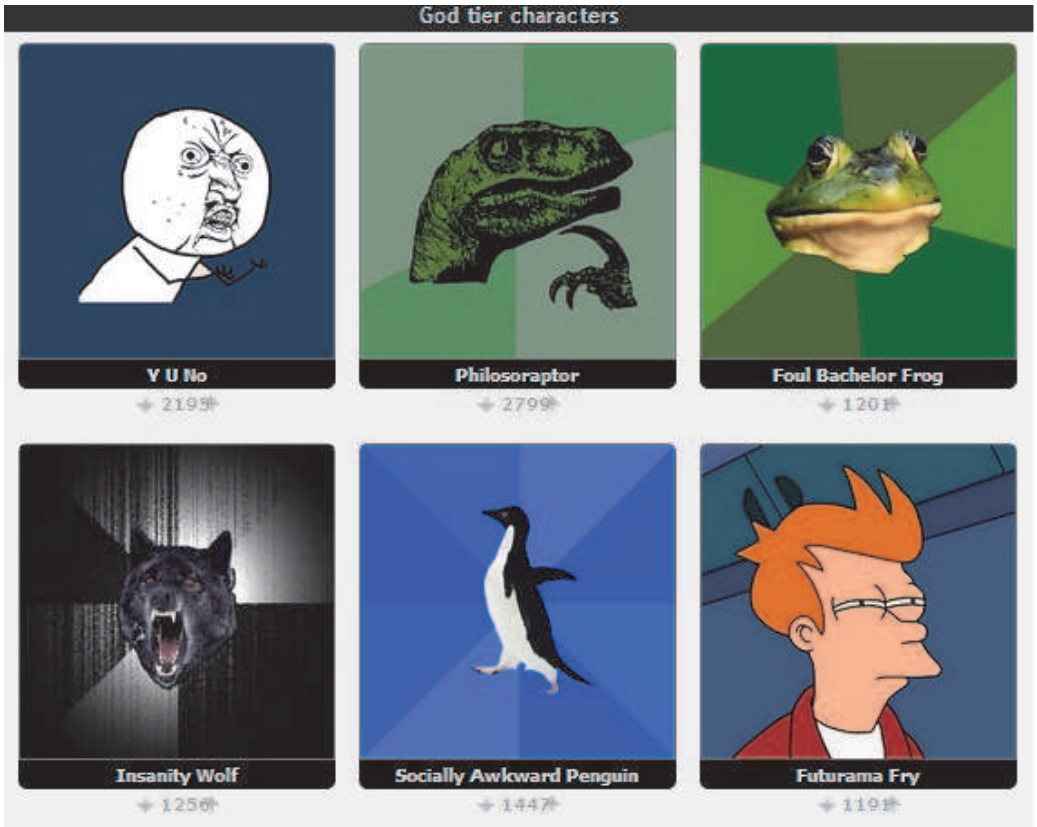


Illustration 4.9 – Les 6 mèmes du « niveau divin »de Meme Generator. Le nom et le nombre de votes positifs figurent sous chaque image.

Les six mèmes présentés dans l'illustration 4.9 le sont par les images de base, dites macros, et les internautes doivent cliquer afin de voir ces images transformées, et classées par popularité. Les six images du *God Tier* illustrent les intérêts des internautes, dont je vais tirer une courte typologie. Les

30. Toute structure participative sur internet, qui offre gratuitement le contenu généré par ses visiteurs.
31. Accessible à l'adresse : <http ://memegenerator.net/tier/god>. Dernier accès au 7 décembre 2011.

plus populaires sont celles des animaux, représentations archétypales à la manière des fables d'Ésope, dont le nom contient les caractéristiques formelles et l'identification de l'animal ([PHILOSO[PHY]] + [RAPTOR], [INSANITY] + [WOLF], [SOCIALLY AWKWARD] + [PENGUIN] et [FOUL BACHELOR] + [FROG]) : ces noms suffisent amplement à comprendre le type de sujets, ou plutôt, d'approches, de ces quatres archétypes, qui illustrent les exagérations des internautes. Être un célibataire sale, ou maladroit socialement, peut arriver aux internautes, dans une certaine mesure [32], et ils s'en amusent, en partie, par sympathie.

Les questions posées par Philosoraptor rejoignent les constatations de Fry de *Futurama* [33], et une illustration caricaturale de *RageComics*. Tous ces mèmes au fond neutre, qui est toujours une roue de couleur, se déclinent de la même façon : une proposition (protase) au dessus de l'illustration, et une autre sous l'illustration (apodose). Je prends d'ailleurs ce fond de couleur, toujours identique, comme un sous-mème, à part entière, tout comme cette disposition spatiale des éléments des phrases contenues dans les mèmes. Ce fond de couleur indique par sa présence qu'il s'agit d'un mème. Elle fait donc partie de la syntaxe sémiotique de leur représentation. Les outils comme *Meme Generator*, ou l'option *remix* de *Know Your Meme* permettent aux internautes de créer et de partager très simplement du contenu humoristique de façon rapide et percutante.

4.4.3.1 Étude de cas : les LOLcats

Le mème le plus connu, qui a donné naissance au plus gros engouement et qui est le plus universellement connu est celui des lolcats (ou LOLcats). À l'origine, ce mème est constitué de deux éléments d'identification : une photo de chat, et une légende écrite dans un anglais graphiquement et grammaticalement incorrects (développé ensuite en tant que *Lolspeak*), dans le but d'amuser le destinataire. Le nom même de ce phénomène est une fusion entre *LOL* (l'abréviation de '*Laughing Out Loud*') et de *cat*. L'utilisation de photos de chats de façon humoristique remonte au début du XXème siècle (voir

32. À ce sujet, le trope internet fréquemment rencontré « il n'y a pas de femmes (ou d'homosexuels) sur internet » est un corolaire de cette constatation.
33. Dessin animé populaire, réalisé par M. Groening (1999-2013).

l'illustration 4.10), mais ce mouvement a pris de l'ampleur à partir de 2007. Sur le forum *Something Awful*, Eric Nakagawa a publié l'illustration présentée en 4.11, qu'il avait composée en imitant le style de *4chan*, une autre communauté virtuelle d'images au contenu souvent humoristique mais provocant.

Le lolcat est la figure de proue de la culture internet, mais n'en est évidemment pas la seule repré-

Illustration 4.10 – Photographie de 1905 (Harry Whittier Frees)

sentation. À eux seuls, les lolcats symbolisent le communaulecte à l'échelle d'internet (donc l'anti-communaulecte), qui est partagé par tous les internautes [34]. Les usages linguistiques communaulectaux sont partagés dans une communauté à part entière, mais ils peuvent aussi avoir été créés simultanément ailleurs. La graphie <gurdle> (‘*girdle)*’) n'est pas inédite à BP, mais elle y est plus fréquente dans un seul fil de discussion sur ce site que sur la plupart des sites internet en général. De la même façon, les manifestations mémétiques ne se concentrent pas uniquement sur une seule plateforme in-

34. Je suis persuadée que, sans forcément connaître le nom « lolcat », presque tous les internautes ont été confrontés au moins une fois au phénomène.

ternet, ou dans un seul type de communication, comme les *imageboards* [35] à la *4Chan*. L'existence et la propagation des usages communicationnels alternatifs du communaulecte et du mème les rapprochent cependant. Un communaulecte a besoin d'être mobilisé par ses locuteurs pour exister, tout comme une figure mémétique tombe dans l'oubli si elle n'est pas utilisée en ligne. Le lolcat est le précurseur de la communication par utilisation d'images aux mêmes caractéristiques formelles, dont l'usage est répandu dans certaines niches de la culture internet.

Illustration 4.11 – Le premier lolcat, créé par Cheezburger (Eric Nakagawa)

Cette image a connu tant de succès que E. Nakagawa a créé un site afin de publier des images de lolcats, *I Can Has Cheezburger*. Ce site enregistre environ 1.5 million de visites par mois, et est à l'origine de nombreuses déclinaisons mémétiques, autour de thèmes différents (*FAILblog, Know Your Meme*…). Les lolcats reposent sur le principe simple d'une image de chat (principalement, bien que le site accepte des photographies d'autres animaux), et d'une légende en anglais grammaticalement et

35. Forums basés principalement sur la communication par images, opposés aux *textboards* étudiés dans cette thèse.

graphiquement incorrect, qui sont des conditions obligatoires dans la création d'un lolcat.

Créer des lolcats a donné naissance à une forme d'argot internet utilisé par les amateurs de ce phénomème. Cet argot s'appelle *LOLspeak*, et détourne des règles élémentaires de grammaire, comme l'omission ou mauvaise conjugaison des verbes : utilisation de la terminaison <-s> pour les verbes à la première personne, ainsi que d'utilisations lexicales figées : des phatiques comme *ohai* et *kthxbai* (‘*OK, thanks, bye*’), qui sont des reprises d'utilisations, notamment dans le langage SMS (<k> pour « OK », <thx> pour « thanks »). L'originalité de kthxbai tient dans le fait que les trois mots ([OK] + [THANKS] + [BYE] sont joints, et utilisés comme formule de clôture. C'est une nouvelle convention accompagnant la connaissance des lolcats.

Speak LOLspeak[36] était un site dédié aux gens aimant assez les lolcats et leur univers pour tenter de « communiquer » comme eux. S'en est suivi le développement d'un communaulecte fragile, avec des utilisations lexicales ou grammaticales ressemblant celles des lolcats. Ces occurrences uniques de mèmes ont donc donné naissance à des graphies alternatives, et à des pratiques communaulectales, spécifiques à *Speak LOLspeak*. Malheureusement, la communauté était très petite (moins d'un millier de membres), et beaucoup d'entre eux ne participaient pas sur le forum. Le site a disparu en 2013, mais les relevés ont été faits en 2010-2012 pour un article (GOUDET, 2012). Dans le fil de discussion « Who actually chooses to talk like this », les messages se concentraient en trois pôles : soit les utilisateurs n'employaient pas de lolspeak (22 % des messages), soit une quantité moyenne, entre 40% et 75% de mots écrits en lolspeak (la majorité, avec 58%), soit des messages presque entièrement en lolspeak (20% du fil de conversation). En revanche, certains utilisent le *LOLspeak* pour montrer une divergence avec l'anglais (ou avec la situation énoncée) :

> I did notice that whenever anyone here wants to be understood and taken seriously they use real, intelligible English. Mai dikshunerry wil b so rleeved !! ;)

L'utilisateur scinde son message entre la partie sérieuse et une plaisanterie, en LOLspeak, à la fin de son message. Le décalage est d'autant plus grand qu'il utilise deux points d'exclamation (hyperexpressif) et un émoticône pour appuyer le côté ludique de son utilisation du *LOLspeak*. Le communaulecte sert donc à renforcer la ludicité, le recul pris dans son message—et il souligne très justement la distance

36. Anciennement accessible à l'adresse <http ://speaklolspeak.com>, dernier accès le 23 juillet 2012.

entre pratiques transgressives et performativité des graphies alternatives. Pour être pris au sérieux, il ne faut pas avoir recours à un communaulecte. Le rapport à la norme est inscrit dans les graphies communaulectales, et leur transgression revient à rentrer dans une dimension discursive ludique. À cause de ses caractéristiques formelles (erreurs grammaticales, comme le montre la phrase emblématique « I can has cheezburger », graphies alternatives…) le LOLspeak se disqualifie en tant qu'argot utilisable dans des circonstances sérieuses. Il est par essence le langage drôle des lolcats, et émule leur parler fantasmé. Ce communaulecte est emblématique de ce que l'on trouve également sur SC et BP, et des alternances codiques en fonction du type de messages (ce qui sera évoqué en 4.5).

Un autre phénomène langagier est l'*engrish*, qui a donné naissance à des utilisations entre le mème internet et le communaulecte, entre des internautes qui maîtrisent et aiment ce code.

4.4.3.2 *Engrish : All your base are belong to us*

L'*engrish* est essentiellement une pratique scripturale alternative, née hors-ligne mais dont l'exploitation est indigène au numérique. Cette étude de cas est emblématique entre communaulecte et graphies alternatives, car la substitution <l>/<r> dans <Engrish> renvoie à un transfert phonétique erroné. Les exemples hors-ligne sont également agrégés en ligne, dans certains sites spécialisés. L'exploration des caractéristiques de l'engrish à travers le cas d'*All your base [are belong to us*] (AYB) permet de voir la façon dont un mème essentiellement linguistique peut dépasser son cadre initial (ici, l'introduction d'un jeu vidéo) et être transposé dans des médias audio, vidéo, visuels. Dans ce contexte, certaines phrases ont plus de succès que d'autres, et permettent également d'identifier immédiatement le mème dont on parle.

AYB est un excellent exemple d'*engrish* : il s'agit de la première erreur de traduction dont l'utilisation a généré des récupérations en ligne (chansons, montages images…). AYB est un étalon de l'*engrish*, non seulement dans son origine, mais par ses dérivations numériquées [37].

L'*engrish* est originellement la mauvaise retranscription de termes en anglais, à partir de la phoné-

37. J'utilise ici la nomenclature numérisé-numérique-numériqué pour décrire les types de textes (Paveau, 2012).

tique (« bad lomanizashon » [38] selon le wiki de *Transformers* ; ce type de pseudo-anglicisme est appelé *wasei-eigo* ; littéralement, '*l'anglais fabriqué au Japon*' en japonais). Le terme « engrish » est une démonstration de ce phénomène, la confusion entre les deux consonnes alvéolaires /l/ et /r/, qui entraîne une mauvaise graphie de mots anglais : ces deux phonèmes de l'anglais ne sont pas distinctifs en japonais. Les traductions littérales—potentiellement pleines d'erreurs grammaticales et syntaxiques, comme le titre de cette section le suggère, et d'expressions curieuses, sont un bon avatar de ce phénomène. L'illustration 4.12 est un exemple de la confusion fâcheuse entraînée par l'*engrish*, à cause de l'homographie et de la sémantique.

Illustration 4.12 – Image postée sur engrish.com en 2008, tirée d'une bannière d'élection de Konomiya (Japon)

L'engouement sur internet autour des formules en *engrish* a pris son essor avec le même « All your base (are belong to us) », un des extraits de la séquence d'introduction d'un jeu vidéo, *Zero Wing* (« Zero wing », 1991). Le jeu a été produit en 1989, mais les dialogues particulièrement mal traduits à partir de la version japonaise, ont été popularisés dans le monde anglophone en 1998. Des utilisateurs du forum *Classic Gaming Powered by IGN*, anciennement connu sous le nom de *Rage Games*), ont créé des images animées et de vidéos parodiques en 1998, puis le mème a connu un essor ouvert au plus

38. Accessible à l'adresse <http ://tfwiki.net/wiki/Engrish#Types_of_Engrish>

grand public d'internet au début des années 2000. Le tableau 4.5 présente l'incipit de l'introduction du jeu. Cats est le nom du chef des forces antagonistes, les ZIG sont des vaisseaux de combat et les citations sont découpées selon leur apparition à l'écran. La syntaxe de tout l'extrait est erronée : les questions ne sont pas construites correctement, de « What happen ? » dont le verbe devrait être au prétérit, les phrases sont concaténées, comme « You have no chance to survive make your time » qui colle arbitrairement deux phrases ensemble, et l'ensemble est beaucoup plus court que la version japonaise. Les choix lexicaux sont mauvais : « take off » veut effectivement dire « décoller », mais pas « déployer ». Cette traduction est à la limite de l'absurde, et n'est même pas proche de la version japonaise initiale. Pour « all your base are belong to us », la traduction littérale de la phrase originale est « Cats a reçu toutes vos bases ». C'est l'absurdité de l'ensemble qui a attiré les internautes.

Traduction en anglais	Traduction correcte en français à partir du japonais
Captain : What happen ?	Capitaine : Qu'est-ce que c'était ?
Mechanic : Somebody set up us the bomb.	Ingénieur : Quelqu'un a placé une bombe dans le vaisseau.
[…]	[…]
Operator : Main screen turn on.	Opérateur des Communications : Je passe le message entrant sur l'écran principal.
Captain : It's you !!	Capitaine : V…Vous !
Cats : How are you gentlemen !!	Cats : Vous avez l'air occupé, monsieur.
Cats : All your base are belong to us.	Cats : Avec l'aide des forces de la Fédération Gouvernementale, toutes vos bases sont passées sous le contrôle de CATS.
Cats : You are on the way to destruction.	Cats : Votre vaisseau court également à sa perte.
Captain : What you say !!	Capitaine : C'est…c'est ridicule !
Cats : You have no chance to survive make your time.	Cats : Nous vous remercions de votre coopération. Chérissez les derniers moments de votre vie.
[…]	[…]
Captain : Take off every 'ZIG' !!	Capitaine : Je vous ordonne de déployer toutes les unités ZIG !
Captain : You know what you doing.	Capitaine : Nous n'avons pas d'autre choix.
Captain : Move 'ZIG'.	Capitaine : Ce sont nos espoirs pour notre avenir.
Captain : For great justice.	Capitaine : Nous comptons sur vous les ZIG !

Tab. 4.5 – Extrait de l'introduction de Zero Wing *traduite en anglais.*

Les données pour la recherche sur le mème « all your base » ne sont disponibles que depuis 2004 [39], et le pic de recherche en utilisant cette expression a été atteint en mars 2004, comme le montre l'illustration 4.13. Les recherches ont été de moins en moins fréquentes à partir de début 2006, ce qui prouve l'obsolescence des mèmes, qui gardent un fond de popularité malgré leur ancienneté. La formule plus courte « all your base » suffit à reconnaître le mème, et elle enregistre plus de résultats que sa contrepartie non abrégée. Ce phénomène mémétique n'est donc pas obsolète immédiatement, comme le montrent les résultats de Google Insight. Cette plateforme recense les recherches effectuées-qui contiennent l'un des termes de recherche (soit « all your base », soit « All your base are belong to us »). Les valeurs ne sont pas absolues, et elles sont normalisées sur une échelle de 0 à 100, en prenant la valeur la plus élevée, et en divisant les autres valeurs par ce nombre avant de les multiplier par 100. Ce n'est pas parce que le mème n'est plus recherché sur internet qu'il n'est pas connu ou n'a pas intégré les pratiques discursives des internautes. Son interprétabilité reste stable, malgré le faible nombre de recherches.

Sur BP, il y a 40 occurrences de la phrase « All your base are belong to us », dans des profils personnels, qui jouent sur les références culturelles partagées. La plupart de ces utilisateurs listent un lien vidéo vers l'introduction du jeu dans leurs liens préférés (rubrique « Favorite Links » de leur page personnelle).

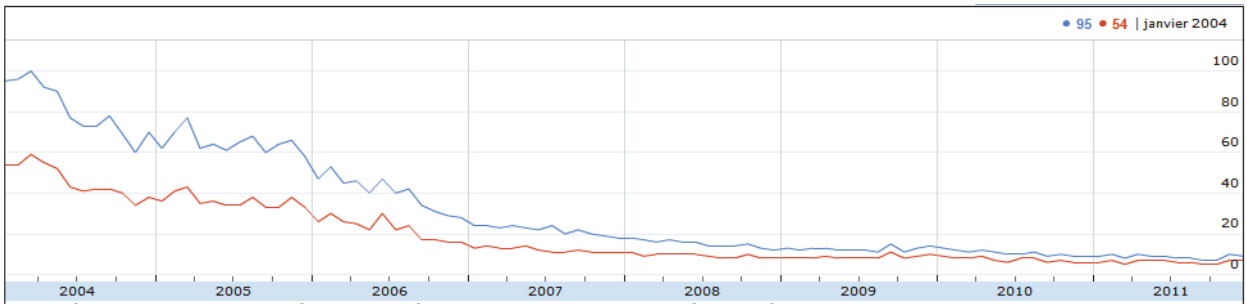


Illustration 4.13 – Intérêt et nombre de recherches pour « All your base » (en bleu) et « All your base are belong to us » (en rouge) entre 2004 et 2011.

Elle contient la faute de grammaire la moins importante de la réplique (l'erreur d'accord de « base »),

39. Source : <http ://www.google.com/insights/search/#q=All%20your%20base%2Call%20your%20base%20are%20belong-%20to%20us&cmpt=q >. Dernier accès : 15/04/11. Une recherche similaire en 2014 donne les mêmes résultats, pour les pics d'utilisation.

mais la citation, tronquée, ressemble à une apocope. Cette troncation n'est pas problématique, parce que les internautes ne cherchent pas tous les termes : ils connaissent le contenu auquel ils veulent accéder. L'acronyme AYB peut être vu comme la désignation de l'introduction complète de *Zero Wing*. La citation de Cats dépasse son propre cadre, comme la citation d'un vers de poème peut être une allusion à toute l'œuvre. La première faute de grammaire est un indice suffisant pour faire la distinction entre un syntagme écrit en anglais standard « all your bases » et le mème en question.

Même si les internautes n'utilisent que le début de la citation, les requêtes sur *Google* montrent une très faible disparité entre le ratio de présence de « All your base » et « All your base are belong to us », 1 :5,7. Cela signifie que les internautes qui effectuent cette recherche savent ce qu'ils vont trouver, et n'ont besoin que de l'introduction de la citation pour la trouver.

D'autres répliques ont eu un succès plus modéré « For great justice », par exemple, est également exploité, mais ne revient que très rarement sur internet. Son ratio d'apparition par rapport à « All your base are belong to us » n'est que de 1 :541,7. « You have no chance to survive make your time », ou « somebody set us up the bomb », d'autres citations connues, n'ont pas eu le succès de AYB. Ces citations sont l'occasion de créer d'autres parodies et de citer le texte de l'introduction. Par exemple, la définition la plus populaire de « for great justice » de l'UD est « Primary reason for to take off every ZIG. » ; celle de « make your time » est « This expression is to be used when you have no chance to survive. It originated in AD 2101, when war was beginning. To successfully "Make your time", you must take off every 'Zig'. You know what you doing. » [40].

La référence est plus qu'évidente pour les connaisseurs, et cet humour de répétition et d'allusions au texte intégral de l'introduction, entretient le phénomène cohésif entre les utilisateurs ou les initiés de ce mème. Un clip vidéo, avec musique et ajout de voix électroniques a été produit en 2001 par l'internaute Bad_CRC, et la version disponible sur *Youtube* depuis 2006 [41] a été visionnée plus de cinq millions de fois en décembre 2011. Ce clip, qui commence par la vidéo et le doublage en voix électroniques de l'introduction du jeu pendant les premières 45 secondes, est ensuite suivi par un remix audio (« Invasion of the Gabber Robots » par The Laziest Men on Mars (« Invasion of the Gabber Robots »,

40. Ces définitions ont été écrites entre 2003 (pour celle de « Zero Wing » et de AYB) et 2005.
41. Accessible à l'adresse <http ://www.youtube.com/watch ?v=qItugh-fFgg>

2000)) de cette introduction. Les images qui le constituent sont des détournements de panneaux ou de paquets de cigarette et autres éléments de culture populaire (fenêtre d'avertissement de Windows, reprise de publicités…) dont le texte a été systématiquement remplacé par « All your base are belong to us ». Cet exemple est symptomatique de la reprise de concepts afin de créer des mèmes, avec ajout d'information à but humoristique sur une base neutre, ou inattendue.

Pour résumer, l'exploitation de ce mème illustre plusieurs phénomènes de leur utilisation :

— Le passage à des médias et à des formes différentes (d'un dialogue de jeu vidéo mal traduit du japonais, à une exploitation en image puis doublée, puis à un clip vidéo qui exploite, à son tour, des images tirées de la vie courante).

— La connaissance d'une unité plus longue que le titre qu'elle représente, ce qui favorise le nombre de recherches.

Le choix est donc presque aléatoire de ce qui constitue un mème performant. Ses formes concurrentes sont tirées du même environnement, et ne connaissent pourtant pas le même succès. L'*engrish* est également utilisé dans les profils personnels (reprise d'images, de contenus humoristiques) pour amuser l'auteur et les destinataires éventuels. Il y a 24 résultats sur BP pour le terme « engrish ». Presque tous renvoient à des mèmes internet, à part quelques citations, et il est évoqué sur BP, pour se moquer d'un autre membre : le mot *engrish* est utilisé ici comme désignant une sous-forme de l'anglais, qui serait trop difficile à comprendre pour l'autre :

> Is that all yu got ? try again. it must be really tough to read simple Engrish. (BP13A1001)
>
> She speaks velly good engrish !!!! (BPNA1)

Dans les deux citations, l'utilisateur qui utilise ce mot se moque de l'autre. Dans la première citation, tirée d'un message de la section sur la religion, c'est l'intelligence de la coénonciatrice qui est remise en question : si elle ne peut pas lire l'anglais, l'*engrish* aussi doit poser problème. Le reste du message comprend aussi des insultes, et des attaques à l'encontre du vocabulaire utilisé par la coénonciatrice (« A question is not a claim you effing moron »). Ce mot ne peut pas être une faute d'orthographe, puisque les touches <R> et <L> ne sont pas proches, sur un clavier en QWERTY. C'est

donc un choix conscient de la part de l'utilisateur du mot. Dans le second, c'est l'accent anglais d'une Philippine dont la vidéo est le sujet de l'article qui est tourné en ridicule. « Engrish » appartient à la même catégorie que « velly », ce sont deux jeux sur l'interversion des phonèmes /r/ et /l/. L'hyper-expressivité des points d'exclamation multiples, la nature lapidaire du message montrent l'ironie du locuteur, et la valeur d'autant plus péjorative du léxème « engrish », pour lui. Le terme *engrish* sur BP qualifie l'ensemble de pratiques langagières syntaxiques erronées sur internet. L'*engrish* est une sous-catégorie du *netspeak*, bien qu'il soit né hors-ligne, dans des traductions douteuses de l'anglais au Japon (dont le *wasei-eigo* est un exemple). C'est sur internet qu'il vit ses grandes heures, où il a été baptisé et vulgarisé. En revanche, le mot en soi désigne un anglais de mauvaise facture, soit parlé par un étranger (surtout asiatique), soit pratiqué ou lu par des gens considérés comme peu intelligents. La pression normative force à considérer l'infraction à la norme comme un défaut d'intelligence. Le mot « Engrish » est employé dans des phrases condescendantes (exemple de BP13A1001). Ici, l'utilisation fautive de l'anglais est prise dans son sens le plus péjoratif, mais l'*engrish* d'internet—celui exploité par AYB et autres—est un outil essentiellement humoristique.

Dans la partie suivante je vais explorer quelques mèmes et leur fonctionnement lorsqu'ils sont au carrefour entre une utilisation communaulectale (c'est-à-dire dans créés sur une communauté) et mémétique, où ils sont tirés de leur contexte initial pour appartenir à la culture internet générale.

4.4.4 Intersection entre mème et communaulecte

Les mèmes sont donc des patrons multimodaux, qui sont mis à la disposition de tous les internautes, pour exprimer des idées déjà connues, renforcer des archétypes existants (*socially awkward penguin*, l'animal représentant la gaucherie), ou même donner lieu à des confessions permises par internet, de façon anonyme (grâce au *confession bear*, ours à l'air penaud). Ces sortes de « patrons » numériques ont un but double : pour exprimer ses idées et en mobilisant ses connaissances en culture internet, dans une forme compréhensible par ceux qui la partagent. L'intersection entre un mème et un communaulecte amène l'idée d'un « communaulecte à l'échelle d'internet ». C'est une entité identifiable,

partageable et potentiellement virale. La contradiction entre la notion d'identification entre membres d'une même communauté et l'échelle d'internet est résolue quand on prend la maîtrise (ou, au moins, la connaissance) des pratiques discursives d'internet comme appartenant à une communauté à part. Les gens qui fréquentent les hauts lieux culturels d'internet comme *4Chan* y trouvent des pratiques langagières nouvelles qui apparaissent tout le temps. Par exemple le *greentext*, qui consiste à raconter une histoire très personnelle, et un moment d'embarras ; parfois complété par d'autres mèmes internet, comme le *bait-and-switch* « spaghetti », où la conclusion absurde de l'anecdote est invariablement le moment où des spaghettis tombent des poches de l'auteur). Dans l'illustration 4.14, le moment de pivot entre l'anecdote et le bait-and-switch est « palms are sweaty ». Elles sont extraites d'une chanson d'Eminem dont les deux vers suivants font souvent partie de ce mème « His palms are sweaty, knees weak, arms are heavy / There's vomit on his sweater already : mom's spaghetti » (Eminem, 2002).

Illustration 4.14 – Une des histoires en greentext publiée sur 4Chan

Ceux qui se documentent dans des sites comme *Know Your Meme* et autres sites de propagation de mèmes (*9Gag*, où les images mémétiques sont sans cesse partagées sur d'autres sites (FB, par exemple) ont également accès à cette culture dans une moindre mesure. Cet accès est secondaire, dans le sens où il ne s'agit souvent que de consommation des mèmes, et non de création. En revanche, ces utilisateurs d'internet connaissent les codes, et sont aussi à même de les employer à l'écrit. Dans cette partie,

je vais explorer les expressions snowcloniques (Pullum, 2004), qui sont des adaptations matricielles propres à l'internet, à ce communaulecte à l'échelle du média.

4.4.4.1 Concept exprimé en un seul mot, une seule expression

Certaines expressions, comme « Do want/ Do not want », « Fail », « cool story bro », sont figées, grâce à leur utilisation sur internet, qui les a fait passer de réponse standard à expression grammaticalisée. Elles sont issues de divers supports, comme « Do not want », qui provient des sous-titres d'une version pirate chinoise d'un DVD de *Star Wars*, « La Revanche des Sith ». La scène finale où Dark Vador apprend la mort de Padmé et crie « No ! » y est sous-titrée « Do not want. ». Toutes ces expressions sont répandues sur internet, surtout quand elles illustrent une image. Elles servent de vecteur à des expressions stéréotypées.

« Fail » est un exemple emblématique de ces utilisations détournées de mots afin de s'intégrer à un argot ou à une culture internet. Encore une fois, « fail » vient d'une traduction malheureuse d'un jeu vidéo japonais, où le message de fin de partie après avoir perdu était « you fail it » au lieu de « game over ». La capture d'écran du jeu, ainsi que diverses vidéos ont commencé à circuler sur internet dès 2004 sur des sites tels que *Something Awful* ou *4Chan*, mais il faut attendre 2007 pour que son utilisation atteigne son apogée dans les recherches d'images.

Ce verbe qui signifie « échouer » a été utilisé comme commentaire sur des vidéos ou des illustrations de situations d'échec, mais il a subi une recatégorisation grammaticale à mesure que son utilisation sur internet s'est développée : de verbe, il est passé à une forme nominale servant d'interjection ou de commentaire sur l'action en cours. Il est assez répandu sur BP (l'expression « epic fail » apparaît 1060 fois ; sur SC, seulement 3 fois). Sur SC, les utilisations concernent des photos ou des vidéos d'événements pitoyables (une chirurgie esthétique ratée, un homme ivre qui plonge tête la première dans un lac glacé et fait un plat). « Epic fail » sert à catégoriser la vidéo, et son contenu. Sur BP en revanche, il sert à ponctuer des attaques inter-personnelles « I would like for you to explain to class how am I angry when I'm laughing AT you and your epic fail ? » (BP12E2220) ou « Your comeback is neither original nor witty, epic FAIL » (BP12E5579). L'expression « epic fail », dans les deux cas, endosse

le rôle de la situation en tout ce qu'elle a de plus ridicule et ratée, du point de vue de l'énonciateur d'un tel jugement. Dans la première citation, le « fail » représente la dernière réponse, agressive, du coénonciateur. Dans la seconde, il s'agit plus d'un jugement général sur la réplique de l'interlocuteur, distanciée du reste de sa phrase par la virgule, et appuyée par la capitalisation des lettres de <fail>. C'est donc une utilisation anaphorique, à référent plutôt large, à l'échelle du discours plus qu'à celle d'un syntagme.

« Cool story bro » est une expression figée utilisée pour montrer son mépris et son manque d'intérêt pour une histoire racontée par un autre membre d'une communauté ; cette expression ambivalente, qui peut être utilisée soit seule, soit en légende d'une image macro [42] ajoutée au fil de discussion est une déformation de la fonction cohésive de la discussion en ligne. Sur BP, l'expression est employée dans plusieurs fils de discussion (33 occurrences sur le site), et uniquement concernant des sujets polémiques (le mariage des homosexuels à l'église, le révisionnisme, le créationnisme…). Un utilisateur en particulier aime à utiliser l'expression. Ainsi, il souligne le côté hors-sujet d'une anecdote racontée par un autre, tout en continuant le débat :

> Cool story bro.
> On Topic : Okay. But do we agree or disagree that homosexuality, or at the very least, homosexual sex occurs in nature absent human beings ?

« Cool story bro » englobe la totalité de la réponse hors-sujet, et souligne qu'elle était dispensable, tout en montrant qu'elle a été lue : pas de réponse ou de réaction à son message n'aurait pas été méprisant de la même façon. Ici, souligner l'existence de l'histoire en lui répondant une phrase stéréotypée constitue une attaque de la face du coénonciateur (Kerbrat-Orecchioni, p.d.).

Ces deux expressions contiennent le même genre de jugement discursif : leur grammaticalisation constitue un raccourci discursif agressif, compréhensible sans pour autant connaître les codes d'internet ou les avoir lues auparavant (« fail » et « cool story bro » n'ont pas particulièrement changé de sens). Leur utilisation permet de court-circuiter le débat en lançant une phrase lapidaire, à laquelle les autres utilisateurs ne répondent pas. Sur internet, de nouveaux mots sont également apparus, et ont

42. Les macros sont les patrons visuels utilisés pour créer un mème logo-visuel comme un lolcat.

soit remplacé un équivalent sémantique, soit rempli une case lexicale vide. Ces expressions sont des néologismes propres aux indigènes numériques qui maîtrisent ces pans de la culture internet.

4.4.4.2 Créations Lexicales : *Derp*

Le netspeak a également vu la création d'unités lexicales à part entière, fondées sur des hapax marquants. C'est le cas de « derp », qui a été utilisé dans le film *Baseketball* (1998) en tant qu'interjection par Matt Stone, un des créateurs de *South Park* (série d'animation américaine, qui met en scène des enfants dans la ville fictive de South Park. Elle est connue pour son humour douteux et provocant, et les situations subversives qu'elle présente). Ce mot a été réutilisé ensuite dans divers épisodes de *South Park*, en nom de personnage (1999, The Succubus, saison 3 épisode 3) et en interjection dans d'autres épisodes.

Le terme est également utilisé en remplacement humoristique de mots, comme dans la bande annonce parodique de « Derp » dont le script est : « Rob Schneider, derp dee derp ! Derp dee derpittee derpee derp ! Until one day, a derpa derpa durba derp. Derp dee derp, – da teetley tum ! From the creators of 'DER' and 'TUM TA TITTALY TUM TERPA DERP' Rob Schneider is – DA DERP DEE DERP DA TEETLEY DERPEE DERPEE DUMB. Rated PG-13. ». La facilité d'adaptation de cette syllabe, ses caractéristiques phonatoires avec l'attaque plosive et la coda contenant /r/ et /p/ en font un mot court mais marquant. Le côté absurde de cette bande-annonce est dû à cette sur-utilisation lexicale (ainsi qu'à la parodie des bandes-annonces, ou au personnage de R. Schneider, réccurent dans *South Park*).

Le site *4chan* s'est alors emparé du terme « derp » et divers macros sont parues à partir de 2006. Sur *Youtube*, les bandes-annonces de films populaires (*The Avengers*, *Iron Man*, *Star Wars*…) sont détournées en utilisant le mot *derp* (73 000 résultats en recherchant « derp trailer »), non plus pour changer les dialogues, mais pour faire de nouvelles bandes-annonces, faites des bêtisiers et des erreurs de tournage. *Derp* est ainsi devenu, d'interjection neutre, un mot qui sert de remplacement d'informations galvaudées (comme les péripéties d'un film d'action générique), puis un synonyme d'idiotie [43].

43. Ce mot peut être mélioré, et *derp* peut ne représenter que la maladresse.

Pour signifier l'imbécilité d'une personne, ses paroles sont remplacées par « derp », ou son comportement est décrit comme « derpy » (le dérivé « derpiest »existe, mais son nombre d'occurrences ne dépasse pas le million sur *Google*, à l'inverse des deux autres formes du mot).

Dans le corpus, l'utilisation de *derp* est assez faible (39 occurrences, toutes sur BP), mais elles sont cohérentes avec les utilisations du lexème sur internet en général. Dans le fil de discussion « Coincidence ? Or…lightskinnedracism ??? » (BP12E932) se tient une curieuse discussion à propos de cette expression.

Message	Corps
8	I know your trying to be literate but all I see is "derp, derp a herp a derp.".
10	Justinman…WTF is a derp, derp a herp a derp
11	It's the sound he makes when he's bent over/right when he feels like gv is about to ejaculate, into his rectum.
12	LOL…check out the South Park episode with Mr. Derp. […]

Tab. 4.6 – Messages 8 - 12 de Coincidence ? Or…lightskinned-racism ??? (BP12E932)

Les utilisateurs se moquent tour à tour de celui qui ne maîtrise pas le code (la discussion de 4.6 est plutôt litigieuse, à cause de contentieux antérieurs entre plusieurs intervenants). Ils réutilisent les attaques « rituelles » de ce fil de discussion, les allusions à l'homosexualité, (utilisées depuis le début du fil). Le seul à répondre correctement à la question est l'utilisateur initial de « derp » (peut-être est-il le seul à en connaître le sens ?). « Derp » remplace ici (et sur internet, en général) n'importe quel contenu discursif jugé comme idiot, et tellement peu reconnaissable qu'il sonne comme « derp » (le parallèle avec « blah blah blah » est évident, et leur utilisation est identique, à l'exception que *derp* peut se combiner à *herp* pour casser la répétition de l'onomatopée). Le mot peut remplacer une phrase complète, et créée des euphonies plaisantes, qui aident à son passage vers des mèmes audio (comme la parodie de bande-annonce de *South Park*).

Après avoir brossé ce panorama des graphies alternatives nées sur internet et parfois reprises sur les forums étudiés, je vais examiner le cas des graphies nées d'une volonté d'échapper à la censure des sites. Il s'agit de graphies qui contournent les diktats des plateformes numériques, et créent ainsi de nouvelles pratiques linguistiques contournatives.

4.4.5 Censure et détournement

La graphie alternative sert d'argot sur internet : « pron » est l'une des premières occurrences d'une graphie alternative utilisée comme vecteur transcommunauté. La métathèse « pron » pour « porn » est tellement courante qu'elle est devenue un mot-clef à part entière pour aller sur des sites pornographiques. Les sites se servent donc également de « pron » comme mot-clef mentionné dans leurs pages (ainsi que de « pr0n », avec un zéro, qui rappelle plus volontiers le leetspeak et les substitutions graphiques avec des caractères ressemblants). Il y a 23 fois plus d'occurrences de « porn » que de « pron » [44]). La subversion et l'illégal sont au cœur des utilisations censurées, et les mots sont donc détournés graphiquement.

Certains mots sont censurés soit à cause de la volonté de la plateforme, soit à cause d'un désir individuel de l'utilisateur de mettre à distance certains mots en les camouflant graphiquement. J'ai choisi d'étudier la cryptographie de gangs, qui mêlent opérations illégales et caractéristiques langagières d'une société secrète, et désirent donc que les mots utilisés soient obscurs pour les néophytes, ainsi que la censure dans les messages de BP et SC, afin de voir comment les utilisateurs gèrent leurs discours injurieux. Les graphies alternatives y sont utilisées pour contourner des réalités perçues comme taboues (le nom du gang opposé, principalement). Il s'agit d'un cas particulier de la censure exercée par ses propres auteurs, pour porter atteinte à la face de l'autre gang tout en étant moins repérables pour les modérateurs de BP, qui ont à cœur de ne pas encourager les activités illégales sur leur site [45].

4.4.5.1 Études de cas : la cryptographie de gangs

Les gangs afro-américains (notamment les Bloods et les Crips) se retrouvent parfois sur internet, soit dans des plateformes générales comme BP, soit dans des forums ou des réseaux destinés à leur rencontre [46]. Les membres (autoproclamés) de gangs utilisent des graphies différentes et jouent avec

44. 376 000 000 et 16 200 000 liens, respectivement

45. Extrait de leur charte d'utilisation : « Company reserves the right to investigate and to involve and cooperate with law enforcement authorities, and to pursue a civil lawsuit or criminal prosecution for any alleged or actual harmful or illegal activities involving any of the Company Sites. »

46. Cette partie reprend des éléments publiés dans un article de 2014 de la revue *Itinéraires*, numéro « Textualités Numériques »(Goudet, 2014)

l'espace discursif à leur portée. Ils stylisent les graphies et rendent leurs profils personnels dynamiques (bannières défilantes, textes illisibles à moins de les surligner à la souris…). Dans ces conditions, on peut imaginer que les activités de gangs qui vont être décrites devront obéir à certaines règles.

Le membre de gang doit remplir son monologue de présentation personnelle avec certains éléments qui ne vont subir presque aucune variation : une défense de son gang, de son sous-gang éventuel et de ses amis, une attaque contre l'autre gang et son rang au sein de son organisation. En plus de ces codes rhétoriques, il doit se soumettre à une stylisation virtuelle importante : l'omission de certaines lettres ou l'ajout d'autres, en fonction du gang auquel il appartient. Ces codes stylistiques font partie de son identité virtuelle.

Les contraintes graphiques sont nombreuses : il existe par exemple un tabou tacite autour de l'initiale du nom de l'autre gang, qu'il convient toujours de neutraliser. Ceci est schématisé dans le tableau 4.7.

Caractère interdit - Bloods	Substitution	Exemple	Caractère interdit - Crips	Substitution	Exemples
<c>	<CK, ¢, ø>	HO A$$ RABS, cKRIPS	b	bk	Ki77a slobk [Killer Slob]
(bk)	—	—	<ck> (‘*crip killer*’	<cc>	Wuz craccin [what's cracking ?]
Crip	crab	just a ni55a who dont like cKrabs	Blood	slob, slobk	Kill Slobz When I See Slobz

Tab. 4.7 – Expressions et graphèmes tabous pour les membres des Bloods et des Crips

La cryptographie du gang passe également par l'exploitation de mèmes culturels, qui ne sont en fait que le reflet de l'agonalité, et sont adaptables à volonté au gang ennemi. Ainsi, de courts poèmes ou de courtes phrases-choc peuvent être utilisées, pour ritualiser sa détestation du gang ennemi, et ainsi flatter la face de son gang (par extension, la sienne).

Ashes 2 Ashes
Dust 2 Dust.
In Bloods we trust,
and ckrips we bust.
Kill a ckrip win a prize,

Kill a Blood ur whole family dies.
Red to tha head flu to da shoe.
Flip a klip,
hit a ckrip
a watch that tru color drip.

Ces codifications linguistiques sont constituées de reprises « Ashes to ashes / Dust to dust »est un paradigme snowclonique–et religieux, utilisé pour ouvrir ce court poème : les répétitions en début et en fin de vers de « ashes »et « dust »impriment un rythme à la suite de ce message. Le paradigme « in X we trust », qui ressemble à la devise américaine (‘*In God we trust*’) renforce cet effet de litanie.

Le pendant crip existe (extrait de CA4) :

ASHES TO ASHES
DUST TO DUST
IN CRIPS WE [trust]
IN SLOBS WE BUST
U KILL A SLOB U WIN A PRIZE
U KILLA CRIP YO WHOLE FAMILY DIES

Ce poème (et ses dérivés) fait donc partie de la rhétorique de gang au sens large : les phrases courtes font partie intégrante de ces courtes formules, faciles à apprendre, mais qui font partie des productions écrites stéréotypiques du membre de gang (autoproclamé) sur internet. Ici, la censure porte quasiment uniquement sur le gang ennemi, et pas du tout sur le type de discours. Parler de meurtres ou de vengeances (« your whole family dies » ; « kill a ckrip, win a prize ») est tout à fait acceptable, dans un anglais reconnaissable par tous. En revanche, parler de l'autre gang passe par l'utilisation de graphies alternatives, pour neutraliser la force potentielle des mots qui les désignent. Neutraliser un caractère comme <B> en lui adjoignant un <-k> revient à écrire « Blood Killer », en abrégé (et *vice versa* pour les Crips). La mise en scène de la censure diverge de ce qu'elle touche habituellement les thèmes tabous (Allan & Burridge, 2006). L'illustration 4.15, toujours tirée d'un profil personnel de membre de gang, contient deux mots potentiellement censurés sur BP, et dont la graphie est alternative (<bulls**t> pour « bullshit » et <a$$> pour « ass »).

Ces graphies sont stratégiquement logiques : les mots restent compréhensibles, il a recours à des permutations déjà employées (l'astérisque, ou <$>, un signe visuellement proche de <s>). Les mes-

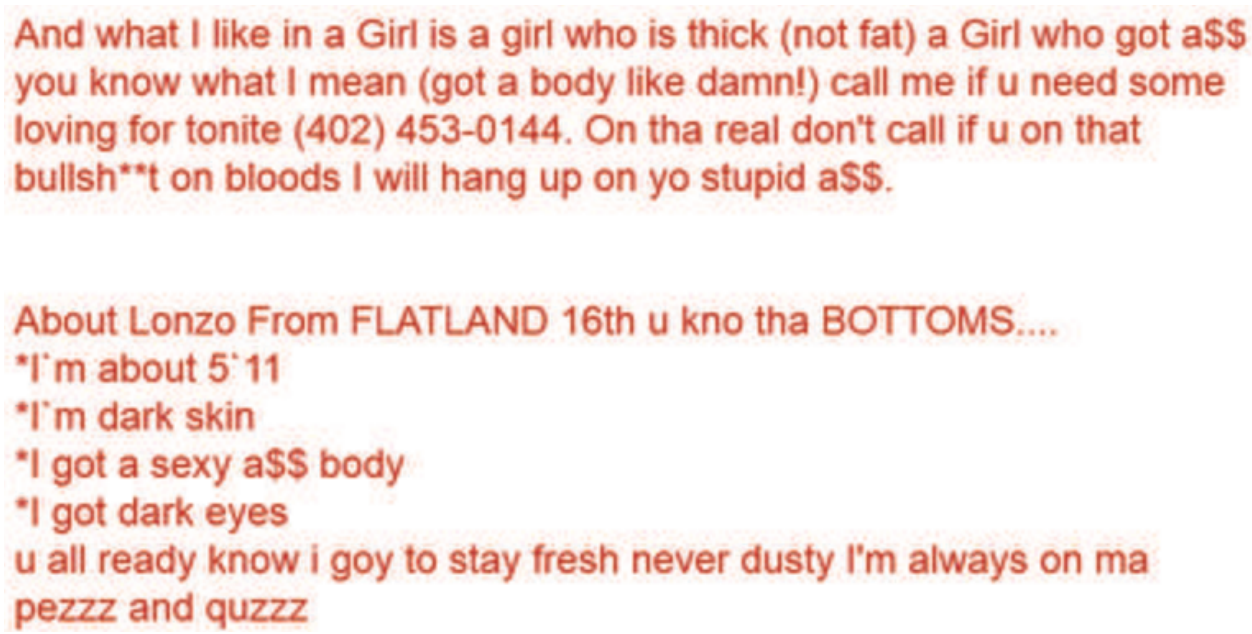


Illustration 4.15 – Extrait d'un profil personnel

sages de membres de gangs, s'ils répondent invariablement au cahier des charges concernant l'appellation de l'autre gang, peuvent aussi contenir des censures personnelles plus couramment trouvées sur internet. Il s'agit du choix de l'utilisateur, qui emploie aussi des tachygraphies (<u> pour « you ») et des graphies alternatives fréquentes (<yo> pour « your », <ma> pour « my », <tha> pour « that »). Sa connaissance d'internet et de la plateforme (il a changé la couleur de fond et la couleur de police de son profil) explique peut-être son désir d'écrire ces mots, habituellement censurés sur BP sans qu'ils le soient.

4.4.5.2 Censure dans les messages

La censure apparaît de deux façons dans le corpus : une censure personnelle, et une censure automatique[47]. La censure est personnelle quand l'utilisateur ne veut pas se résoudre à utiliser un mot tabou, une insulte, une injure, et a donc recours à une graphie euphémisante (suppression des voyelles, remplacées par une astérisque prosthétique, remplacement de toutes les lettres par des astérisques, à l'exception de l'initiale et de la finale…).

47. Certains éléments de cette partie sont tirés de l'article (Goudet, 2013a).

Les censures peuvent également être automatiques, et BP dispose d'une liste de mots censurés, qui sont toujours remplacés par la même suite de caractères, <%#&@$!>. Ainsi, la liste de mots censurés de BP contient des termes sexuels (*ass, cunt, fuck, pussy* et même *vagina* et *penis*), injurieux (*bitch, faggot, whore, nigger*) ou faisant partie des indicibles (*shit*). En revanche, les mots tolérés, qui ne le sont pas forcément sur d'autres plateformes [48] : *damn, dick, dyke, jap, hell, piss* ou *suck* ne causent pas de problème sur BP.

Les utilisateurs réagissent à la censure quand ils ne s'y attendent pas, et censurent donc leurs messages par réaction. Souvent, ils ne comprennent pas la raison de la censure, *a fortiori* quand elle ne concerne pas des mots immédiatement tabous, qui ne sont pas injurieux. L'article BPNA2 évoque la censure incompréhensible de « vagina », et a reçu sept commentaires, qui tournent tous autour de ce thème :

> BP is kinda crazy with the censors. I can see censoring the slang terms but not the scientific terms. i remember when I first started on BP I used to be in the chatroom alot. There was a person in there called wristwatch and the "%#&@$!" (t w a t just in case it gets censored) in wristwatch was bleeped out.

Cette utilisatrice déplore la censure automatique, cite un pseudonyme qui aurait posé problème au script de remplacement (ce qui n'est pas le cas dans son message, puisque <twat> tout seul est censuré, alors qu'inséré dans le mot « wristwatch », il est accepté dans cette citation). Elle a même recours à une onomatopée sonore (« bleep ») pour désigner la permutation de caractères. La censure est donc comprise comme étant habituellement sonore (étant donné qu'elle est utilisée à la télévision, et dans la musique) ; mais elle devient arbitraire quand elle s'applique à un mot de registre standard. La sémantique est plus importante que le registre, pour la censure, et elle s'exerce dans certaines conditions seulement (« twats » n'est jamais utilisé sur BP, et la forme « twat » n'a qu'une seule occurrence dans le forum (BP12E5580)). Ce qui est sanctionné dans <twat> dans le mot « wristwatch » est donc exactement l'inverse de ce qui est sanctionné dans la censure de « vagina » : sémantique et forme graphique s'affrontent donc dans les listes de censure automatique mal calibrées. L'utilisatrice, sensibilisée au problème de censure graphique et aux moyens d'y circonvenir, écrit pour la compréhension de tous <t w a t>, où les lettres, séparées par des espaces, ne sont plus reconnaissables pour le script. La censure automatique déclenche donc une série de stratégies discursives. Cet enrichissement linguistique forcé

48. Sur *SlashDot*, site de nouvelles concernant l'informatique, la plupart d'entre eux est interdite.

par la plateforme n'existe pas sur tous les sites, et dépend du scripteur quand le site n'a pas de censure automatique.

Sur SC, il n'y a pas de censure automatisée, mais les utilisateurs eux-même ont tendance à se censurer et à écrire l'initiale d'un mot et à la faire suivre par deux ou trois astérisques, par exemple. Le fil de discussion SC2Y86, dans le sous-forum sur l'humour écossais, est une description de « Jack Shit » (et un jeu sur l'expression). Le titre du message est censuré (« shit » est graphié <sh*t>) et contient aussi un avertissement « (adult content », toujours dans le titre. Ce message a eu 15 réponses. Les réactions divergent, mais le ton humoristique pousse les autres à continuer à plaisanter–et à utiliser des graphies alternatives, pour censurer « shit » (ou non !). Cinq utilisateurs écrivent « shit » sans utiliser de graphies alternatives de censure, et six ont recours à des stratégies différentes [49]. La censure utilise toujours le même mécanisme : la neutralisation du milieu du mot (en particulier, de la voyelle <i>), par des caractères non-alphabétiques. Deux utilisateurs ont recours à des astérisques, en écrivant <s**t> (l'un d'entre eux utilise <sh**>) ; deux utilisent des points, et écrivent <sh..>, et un utilisateur censure son message avec un tiret bas <_> (<sh_t>). Ces graphies que je nomme contournatives évitent aux utilisateurs d'écrire des injures, tout en préservant le mot original. Le nombre de caractère de chaque mot est préservé (il reste toujours quatre graphèmes à <shit>, même censuré), et l'initiale et la finale sont plus promptement conservées pour comprendre immédiatement de quoi il retourne.

Les contraintes qui s'imposent pour la compréhension des graphies détournées sont d'abord la prise en compte de l'écologie du site. Les utilisateurs identifient d'abord les mots appartenant à la liste censurée, puis adaptent les graphies en remplaçant principalement les voyelles par des caractères non-alphabétiques (<*> ou <.>), et en changeant subtilement la forme du mot grâce à des espaces supplémentaires qui ne gênent pas la lecture humaine.

49. En fait, un utilisateur écrit deux fois le mot dans son message, l'orthographiant la première fois « shit », puis « shite » et finit son message par « birds**t ».

4.5 Comparaison d'utilisation

La distribution des choix graphiques est essentiellement quantitative. Je pars du postulat que les types d'interaction guident les utilisations communaulectales, et connivences éventuelles. Les conditions qui peuvent favoriser les utilisations communaulectales diffèrent, c'est la raison pour laquelle j'étudierai d'abord les types d'échanges qui peuvent favoriser ou au contraire inhiber l'utilisation de communaulectes. La dimension ludique, la complicité me semblent appartenir à une seconde catégorie que je vais analyser dans la seconde sous-partie de cette section. Les membres de communautés virtuelles réagissent parfois aux utilisations presque cryptiques des termes communaulectaux, et cela est perceptible quand ils demandent aux autres des explications ou des éclaircissements.

4.5.1 Types d'échanges

Pour étudier les échanges, et voir quelles conditions provoquent l'utilisation de communaulectes, ils sont séparés en deux grandes catégories. D'une part, Le discours négatif, qui sert à attaquer l'autre ou à poursuivre une dispute, et le discours valorisant, subdivisé en compliments et en discours empathique dans une seconde partie.

4.5.1.1 Insultes et disputes

Lorsqu'un différend apparaît sur SC, en l'occurrence une dispute concernant le côté archaïque ou non des mots de scots enseignés aux non-locuteurs, le lexique scots devient une arme pour se battre. Dans un fil de discussion concernant les biais historiques enseignés à l'université écossaise et anglaise (SC2X50), les divergences d'opinion sont résolues assez calmement, avec des phrases toutes rédigées en anglais standard : « I dont know Tam I think it may work both ways. ». Même chez les gens qui déclarent être très affectés par ces sujets, la discussion reste courtoise (bien que parfois, appuyée par de multiples points d'exclamation, comme « Can you tell I didn't have a good parents evening yet !!!! »). L'utilisation de communaulectes, surtout d'expressions à connotation identitaire, ne passe donc pas

forcément par le type de sujet évoqué ou le degré de colère (ou d'enthousiasme) ressenti.

Dans un autre fil de discussion, le scots est laissé à part, et encadre la phrase en anglais, qui porte l'argument (pour l'utilisation du scots et du gaélique) :

> It's been no hauf dreich in Embra the day. I'll need tae get ma scairf oot o' the cubby hole. [*'There's been very [= no half] dreary [weather] in Edinburgh today. I need to get my scarf out of the cupboard'*
> We all need English to communicate but that should be no detriment to Scots (or Gaelic) for that matter.

Les deux parties de la réponse n'ont pas de rapport : l'ouverture en scots est presque ancillaire (et répond aux exigences du fil de discussion, où les uns et les autres parlent de leur vie quotidienne) et se poursuit par un débat plutôt convenu, sur l'importance du scots. Le communaulecte est utilisé comme marque d'appartenance personnelle, identitaire, alors que le débat en anglais (qui en souligne l'importance, contrairement aux messages d'autres internautes qui mettaient surtout l'emphase sur le scots) est plus général, et insiste peut-être plus sur la légitimité de l'énonciateur. Il finit tout de même son message avec une traduction en moyen-scots de la déclaration d'Arbroath de 1320 : « Sae lang as bot ane hunnert o' us yet abide, ne'er sall we, upon ony conditioune, be subjecktit tae the dominoune o' ye Inglis. It be in truithe, nocht for honourre, nor glorie, nor gowd we be fechting, bot for libertie alane. Thon whit nae guid mon gies up, bot wi' his life itseel. » (*'As long as a hundred of us remain alive, never will we on any condition be subjected to the lordship of the English. It is in truth not for glory, nor riches, nor honours that we are fighting, but for freedom alone, which no honest man gives up but with life itself.'* [50] Au scots quotidien suit l'anglais normatif, puis le scots affectif, qui appelle à la fois à l'histoire nationale, à celle de la langue (le scripteur dit que c'est une traduction personnelle), et au jeu sur la domination de l'anglais (et de l'Anglais). La déclaration d'Arbroath est un texte bien connu par les Écossais et la convoquer ici appelle les sentiments nationalistes et patriotiques écossais. C'est un argument-massue, qu'il n'est pas possible de contrecarrer. Les diverses strates linguistiques se superposent donc pour donner une réponse panachée, où l'alternance codique remplit des rôles bien établis.

Le discours agonal revient souvent dans les communautés, même à cause de broutilles : soit les

50. Traduction officielle de la déclaration, *National Archives of Scotland*.

utilisateurs ont des problèmes les uns avec les autres hors-ligne et les amènent en ligne, soit ils sont crées directement sur le forum. D'ailleurs, les oppositions peuvent se retrouver dans plusieurs fils de discussion : quand on doit discuter avec une personne que l'on n'aime pas, le biais négatif est déjà fondé, et il est difficile de parvenir à des rapports harmonieux. Il existe plusieurs types de disputes :

— Les différences d'opinion, par exemple sur la sauvegarde du scots et du gaélique.
— Les atteintes à la face de l'autre : là, l'opposition ne porte pas tant sur un problème extérieur, mais sur la personne en face de soi. Les communautés internet sont des lieux où il n'est pas possible d'identifier l'autre, sinon à travers ses dires. C'est pourquoi la négociation de la face de chacun est primordiale dans les échanges en ligne. Ce concept s'adapte parfaitement à ce type de discours.
— Les disputes ludiques : ces disputes sont en fait teintées de cohésion : à l'écrit, il est facile d'émettre des insultes et autres attaques sans les penser (l'ironie est impossible à décerner à l'écrit).

Sur BP, sans rentrer dans des études qualitatives, les plus grosses concentrations de messages se trouvent dans le forum religieux. Il n'y a pas de vraie sous-section dans cette subdivision du forum (Religion & Spirituality > General), et elle contient 15 000 fils de discussion. C'est la sous-section qui contient le plus de messages, après la partie « générale » du sous-forum Relationships. Les oppositions entre membres appartenant à des religions différentes, ou qui prônent des points de vue différents se multiplient. Parfois, ces disputes passent d'une divergence d'opinion à une divergence personnelle, où l'enjeu n'est plus que l'atteinte de la face de l'autre.

Les insultes et les discours agonaux peuvent apparaître dans plusieurs cas de figure : soit des dissensions et des différences d'opinion au sein d'une discussion constructive. Dans certains fils de discussion, un avertissement est adjoint au titre : « College or War for our children ? (come in peace) » (BP3A331).

Certains fils ne sont que des attaques gratuites, surtout sur BP : c'est le cas de « You porch monkeys are just subhuman, degenerate rapists and murderers » (BP3D22). Dans le fil présenté dans le tableau 4.8, certaines graphies alternatives apparaissent. Dans cet échange, qui commence par des insultes racistes envers les membres de BP, l'utilisatrice auteure du message 1 n'utilise presque pas de

Message	Corps
1	The rest of the world is getting sick of your crap. You are subhuman and always will be. You thugs and lowlifes will never achieve what the superior Aryan race has achieved. We built civilization. You destroy it. While the white man splits the atom you **coons** can split the watermelon. Did I hurt your pathetic little feelings ? Eat some fried chicken, it'll make you feel better, apes.
3	**u** got your nerve **u** baboon infested pale genetic recessive **bytcch** go back to your caves where **u** belong that's why **bytches** like **u** are on my page trying to get at me. Your the real apes walking on all fours like dogs and **fukin** dogs and cats. Eating raw meat and anything else to survive **bytch**.
4	lol at this **dam** porch money **hoe**.
5	for the record **u** dumb **bytch arayans** are not caucasian its a name your raggedy **azz** pale people stole from the hindu's who are reddish brown **skinn** origin. So much for education huh ? looks like **u** didn't make past kindergarten **u** illiterate **bytch**
6	Ignore these silly White boys who come here to post nonsense. They are usually 15 or 16 year olds, bored playing X-box so they come on this site to express their frustration of being less than blacks.[…]
9	**Lol** wow and **wat** was ted bundy (serial killer who mudered white women)[sic]. Serial killers tend to be white, heterosexual males in their twenties and thirties who are sexually dysfunctional and have low self-esteem. Their methodical rampages are almost always sexual in nature. Their killings are usually part of an elaborate fantasy that builds to a climax at the moment of their murderous outburst. who the **%#&@$! r u too** judge us ? if **yall had off** left us the **%#&@$!** in africa you wouldn't feel this way. white ppl **r** the most ignorant **ppl** ever. Christopher **c**olumbus was supposed to be going to **i**ndia how he got to north **a**merica **cause** he **%#&@$!ing** retarded.

Tab. 4.8 – Extraits de You porch monkeys are just subhuman, degenerate rapists and murderers (BP3D22)

graphies alternatives, à part *coon*, aphérèse de *raccoon*. En revanche, l'auteur des messages 3 et 5 connaît l'autocensure de BP, et sait la contourner : il écrit donc *bitch* <bytch>, régulièrement (sauf une occurrence, où il ajoute un second <c> au mot—sans doute une faute de frappe, étant donné que c'est une exception dans son système graphique). De la même façon, il circonvient à la censure de *ass* en écrivant <azz> (une graphie alternative connue sur BP). Une autre graphie connue est utilisée dans le message 4 : *hoe* pour « whore », qui ressemble plus à une transcription de la phonie du mot : /hɔː/. Cette transcription comporte un <e> final qui montre l'allongement de la voyelle. Elle ne suit pas la règle habituelle qui consiste à retranscrire autant de lettres graphiques à la graphie alternative qu'il en existe dans la graphie standard, ce que l'on trouve dans « bytch ». Elle est donc plus proche

de « fucking » retranscrit <fukin> dans le message 3 (qui a le double avantage d'être plus court, et de contourner la censure automatique de BP).

Le message 9 s'accompagne d'une faute de grammaire : « had off » est certainement prononcé /ˈhædʌv/, et transcrit « had have ». C'est un calque sur le conditionnel en « could have », et cette transcription est inspirée par la prononciation rapide et relâchée. L'auteur de ce message est très en colère, ce qui peut expliquer l'utilisation d'insultes censurées automatiquement (toutes évitées par les autres scripteurs de ce fil de discussion). Il confond également « too » et « to », et ne met pas de majuscule aux noms propres (mis à part « Christopher » au début de la dernière phrase). Ce message a certainement été écrit très rapidement, sans relecture ou soin particulier apporté aux conventions graphiques. Ici, il s'agit réellement de fautes d'orthographe et non de graphies alternatives.

Le message 1 est celui qui contient le moins de graphies alternatives (mis à part *coon*, pour *raccoon*) ; les autres membres évitent presque tous la censure, sauf l'auteur du message 9. Pour contourner la censure, il faut réfléchir non seulement à ce que l'on exprime, mais à la façon dont on l'exprime. Le premier message, s'il est nettement plus agressif que les autres dans le contenu, n'est pas aussi émotionnellement chargé que le neuvième, ce qui lui permet de ne pas être censuré. En revanche, les réactions sont presque toujours guidées par la volonté de contourner la censure (d'où l'utilisation de« bytch », « dam » et « hoe »), mais l'émotion passe par les autres fautes d'orthographe : les messages 3 et 5 contiennent aussi des graphies non-standard à cause de l'agacement : « bytcch » montre un redoublement de la consonne <c> qui n'est pas cohérent avec les autres occurrences du mot dans son message, et le message 5 flirte entre la graphie alternative et la faute d'orthographe. « Arayan » pourrait être une transcription de /ˈæriən/, d'autant plus que les touches <A> et <Y> ne sont pas proches sur un clavier. La graphie « skinn » n'est pas une transcription logique de /ˈskɪn/ à cause de la lettre finale répétée. En revanche, l'utilisateur a tout de même pensé à retranscrire « ass » en <azz>, ce qui montre qu'il a tout de même apporté un certain soin à son message (ou qu'il a internalisé les graphies alternatives sur BP, et qu'il les utilise malgré sa colère).

L'hyperexpressivité ne se manifeste pas toujours, même si les sentiments sont forts. L'agression pure ne déclenche pas la même réaction que les dissensions interpersonnelles, peut-être parce que

l'injure est envoyée de façon générale, et que le groupe est un parapluie qui protège chaque individu des atteintes de face, et donc qui les pousse à réagir avec plus de retenue. Mais ce lynchage ne vaut que s'il est partagé par tous, et l'exemple suivant montre des attaques de plusieurs utilisateurs contre un membre perçu comme étant un mouton noir de SC.

Dans le fil de discussion « Peter D. is single and looking » (SC20E5), l'auteur du fil de discussion est un membre connu pour ses attaques virulentes envers les autres membres, et pour ses opinions politiques nationalistes. Lorsqu'il poste un lien vers son site, qui contient une page matrimoniale sous la forme d'une très longue description dans la section « Love, Sex and Relationships », les railleries pleuvent (le message initial a 57 réponses, ce qui représente une bonne proportion de commentaires sur SC).

Message	Corps
2	Think we are well aware of your posts…Good luck with this latest crazy idea but I think the quality of intelligence on Scotster is way out your league !!!!
3	Ian, is this for real. It just can't be. LMAO
5	My mates Jack Russell at this time of year willl 'love' anybody should I pass on the contact details ? lol
6	the peer wee dog, Ian ! That's cruelty !
8	Ach ! It's got tae be a spoof - he's havin us on !
23	"*But I'm safe - I'm past the age to have his children !*" OMG !!! You mean he could breed !!!!! Think I'll put my brekkie aside till later !!!! lol

Tab. 4.9 – Réactions au message initial de Peter D. is single and lookin (SC20E5)

Les réponses oscillent entre expressivité (points d'exclamation multiples dans la réponse 2 et 26), abréviations usuelles d'internet, qui dénotent toutes le rire (« LMAO, lol » '*Laughing My Ass Off ; Laughing Out Loud*') ou l'étonnement (« OMG » '*Oh My God*'). L'étonnement est d'ailleurs le seul discours du message 3 et 8, où les utilisateurs ne croient pas à la véracité de cette page matrimoniale. En revanche, le scots est utilisé dans deux des réactions que j'ai choisies (pour leur brièveté). Les auteurs des messages 6 et 8 utilisent un code alternatif (« the peer wee dog » ; « Ach ! It's got tae be a spoof »), qui renforce le lien affectif qu'ils ont avec leur discours. Dans le premier, le discours est plus affectif et empathique avec le chien (comme si l'utilisatrice croyait à la plaisanterie du message précédent), alors que dans le message 8, il s'agit plus d'incrédulité par rapport à l'auteur du fil de

discussion. L'interjection « ach » renforce l'étonnement (tout en étant très scottisante), et les phrases de ces deux messages sont toujours ponctuées avec des points d'exclamation. Ici, les plaisanteries et les railleries sont le ciment des membres (l'auteur du premier message n'a plus réagi sur le fil de discussion), qui n'utilisent le scots que lorsqu'ils veulent envoyer un message expressif : « the peer wee dog », qui renforce le plaidoyer humoristique par exemple. La connivence de la raillerie ne passe donc pas tant par le communaulecte propre à SC que par une expressivité accrue et par l'ironie.

Les utilisateurs se répondent les uns aux autres (les messages 2, 5 et 23 ont été rédigés par le même utilisateur, Ian), s'appellent par leur prénom ou citent les messages précédents (début du message 23). Cela accentue le lien entre ceux qui rient, et surtout, le décalage avec Peter D.. Le fil de discussion est d'ailleurs mort après que nombre d'utilisateurs aient dit aux autres de ne plus lui faire de publicité, et de ne plus répondre au fil de conversation (« Move on people. Nothing to see here..... », message 41). Ignorer l'autre, et ne pas lui laisser de place discursive, est une bonne tactique sur internet. Peter D. répond à nouveau à son fil de discussion quelques mois plus tard, sans prendre en compte les messages précédents, et en disant qu'il cherche toujours une compagne (réponse de novembre, après avoir créé le fil en avril 2010–et message 45 du fil de discussion), et son message est suivi du même type de railleries.

4.5.1.2 Discours valorisant

La valorisation peut être exprimée de diverses façons, et se jouer entre un membre du site et la plateforme (pour féliciter les administrateurs d'avoir créé une section qui leur plaît, ou remercier pour la qualité du contenu mis en ligne, par exemple), ou entre un membre et un autre (en quasi face-à-face). La valorisation est également codée discursivement : lorsque l'on utilise un terme ou une tournure communaulectale, on fait simultanément une demande d'inclusion à la communauté de pratique, et l'on offre une preuve de son inclusion dans le groupe social. On valorise également celui qui comprend le message sans avoir besoin d'explications. Les pratiques communaulectales sont donc valorisantes intrinsèquement.

Le communaulecte est donc valorisant, et positif à utiliser. Dans les fils de discussion, surtout sur

SC, la langue minoritaire est mise en valeur : dans SC2X50, les utilisateurs parlent du scots, et, toujours pour reprendre la citation de la déclaration d'Arbroath, l'utilisateur souligne le fait qu'il est l'auteur de sa traduction : « My own translation into Middle Scots. Hope you like it. Feel free to use it. ». Non seulement il se met en valeur, parce qu'il peut traduire en moyen scots, mais il la partage avec les autres, en espérant qu'ils l'apprécient. Il est donc positif pour soi (et pour la communauté) de pouvoir s'exprimer en scots. De la même façon, une utilisatrice, Genevieve, écrit un roman où les dialogues sont en scots du 17ème siècle. Les autres lui apportent donc de l'aide, qui est toujours très appréciée. Dans l'extrait 4.10, elle discute avec Michael et Julia. Michael est écossais, et spécialiste de l'histoire picte alors que Julia, fraîchement installée en Écosse, connaît un dialecte qu'il ne maîtrise pas (le scots de l'Aberdeenshire) :

Message	Corps
43	Hi there Genevieve. Dougal might say " Ah dinnae ken whit am daein,foryet (forget) that ! I a dae ken,here we sit talking but I'm havering, which drove oot a ma heid whit ah wis supposed tae tell ye." Just an idea whit it might soond lie in Scots. Guid luck wi it aw.
46	Thanks a ton, Michael …as usual !
50	Just catching up on the forum and noticed Genevieve was asking about words for kids. Well in aberdeenshire where i am the call the girls quines and the lads loons. Think thats how you spell it onyways. Learning to understand it is one things learing to spell it is something else but oh this is definitely the best topic on scotster.
51	Hi there Julia,I didnae ken whit the bairns were called up Aberdeenshire wi , but I hud a feeling it wisnae bairns or weans. Cheers fer the info.
53	Thank you Julia ! Isn't that funny - I thought "quine" was some expression for cow, but maybe that's just because I know "kine" was cattle …And loons, well, that's just funny because most of the lads I know ARE loonie …You are so right about this topic. I'm loving it. Thanks again.

Tab. 4.10 – Messages 43 - 53 de Loss of our Language (SC2X50)

La dimension communautaire est très importante dans cet échange. Michael répond en scots (message 43, 51), et donne à la fois une traduction pour le livre de Genevieve, et des commentaires sur le scots d'Aberdeen. C'est lui qui utilise la plus grande proportion de scots dans ses écrits. Julia (message 50) n'utilise des scottismes que lorsqu'elle donne le vocabulaire à Genevieve (« quines ; loons »), et laisse échapper un « onyways » ('*anyway*'), un vieux mot de scots. Ni elle, ni Michael ne séparent

les scottismes avec des guillemets (dans le cas de Michael, quand il parle, il n'a pas besoin de séparer les mots, ce qu'il a fait quand il traduit le dialogue pour Genevieve). Genevieve, étant l'utilisatrice la moins à l'aise avec le scots, utilise le plus de marques pour le séparer de son discours (pour « quine » et « kine »), et ne pratique pas de vraie alternance codique dans son message. Pour elle, il s'agit moins de communaulecte que d'une langue minoritaire, qu'elle met à part dans ses productions (elle est canadienne, et n'entend pas et ne pratique pas le scots hors-ligne).

Ces utilisateurs multiplient également les remerciements et les marques affectives dans leurs messages : « guid luck » ; « thanks a ton » ; « cheers for the info » ; « Thank you […] thanks again ». Ils se valorisent les uns les autres, non seulement sur leur maîtrise du scots (Geneviève remercie les deux autres participants, Michael est content d'apprendre de nouvelles choses sur la langue), mais ils valorisent également tout le fil de discussion (« definitely the best topic on scotster »), avec un superlatif de comparaison et même des mots de jugement affectif forts (« I'm loving it »). Leur entraide mutuelle, autour de la langue minoritaire rejaillit sur tous, ce qui se voit dans leurs choix de remerciements différents, car « thanks » n'est presque jamais utilisé tout seul, et leur utilisation de vocatifs dans les messages renforce l'interpellation et donc l'importance de l'autre dans cette conversation. Le communaulecte est donc complice, outil et vecteur du sentiment d'appartenance à la communauté. Être heureux d'avoir des informations grâce à deux utilisateurs revient à être heureux d'être sur la plateforme qui permet de tels échanges.

4.5.1.3 Discours empathique

Le communaulecte est partagé dans la connivence. Mais cette connivence doit être positive, joyeuse : les fils de discussion sur la perte d'êtres chers ne contiennent presque pas de graphies alternatives (bien que des fautes puissent apparaître). C'est le cas de BP4E2, dans la sous-section « Mourning and Loss », où l'auteur du fil de discussion écrit une lettre ouverte à sa femme, décédée un an auparavant. Ce fil contient 16 messages lisibles [51]. Voici le premier message de ce fil de discussion.

51. Sur 25 initialement, mais ils ont été effacés, et remplacés par « The content you are trying to view has been removed. ».

> […] i will raise our baby right i promise honey. you know, she is in high school already. and you missed her first date. she looked so pretty baby. and the young man **who i** threatened regularly about talking to my daughter in church, is the one that took her out. he was **kinda** scared to stand next to her on the home comming pictures tho. but he really turned out to be very nice young man after all. but i am still watching him tho.
> she is continueing her education, not to be a chef like her dad, grand father, great grand father and her great, great grand father. but to be a veterenarian. you know how much she loves annimals.[…]
> baby i am so lonley now, and i miss you so much. but i only cry when our daughter is not around. i will be ok tho. and maybe if i am lucky, i will find someone in my life. not to take your place, but to fill this huge void of lonliness i am feeling. you gave me 20 of the best years of my life. and there is really nothing anybody could say to make it better. but you god i thank, for letting me borrow on of your most precious angels.
>
> i miss you baby. and i dedicate my life to you.

Les tachygraphies, et effets prosodiques de la parole sont les graphies non-orthographiques les moins courantes. L'utilisateur commet quelques fautes d'orthographe : *home comming*, *continueing* ; *veterenarian* et *lonley/lonliness* (par méconnaissance de la graphie, et calque de la prononciation). Il semble ne faire qu'une seule faute de frappe, *annimals*. Les réponses conséquentes suivent ce modèle, et ne contiennent pas ou très peu de graphies alternatives communaulectales.

Dans ce fil de discussion, il n'y a aucun smiley, aucune marque graphique d'expressivité (multiplication de ponctuation), à part le message d'une utilisatrice, « I love this so much,,,,,,,this was so nicely said !!!!!!!!!!!! » (BP4E2-21). Cette graphie personnelle (confusion entre le point et la virgule, sur la même touche, hyperexpressivité de la ponctuation) évoque la stylisation personnelle ainsi que la valorisation de l'autre par cette expressivité affective.

Sur SC, la section du forum dédiée à l'entraide des autres (section Support, SC25) contient très peu de fils de discussion : sur 28 sous-sections, il n'y a que 11 fils de discussion. Un fil de discussion assez similaire a été ouvert (extrait 4.11) :

Message	Corps
1	I am taking a short break from caring for my ailing wife. She is in hospice care and the Hospice nurse told me she has 24-48 hours left. I have told a handful on the board about what my wife is facing and I wanted to thank all of you for being my friend, chatting with me when I was feeling down and for being there when I needed a friendly ear.
2	Best wishes to you LT, i'll stay with you through this whole thing too if needed.
3	Loving thoughts for you from across the sea. Kia Kaha.Arohanui.May you be strong.Much love.
7	Thanks everyone for the kind words. However, this morning at 0800 Eastern (US) my wife of almost 27 yrs passed away. I am so sad but i have no regrets she knew I loved her with all my heart. I still can't believe she is gone.....if you belive in god, please say a prayer for me and my family that we will have the strength to get through this.
19	Sherwood, I am sue that our Lord will bless you and pull you through this sad time.
20	for sue read sure

Tab. 4.11 – Messages 1 - 20 de With Sadness (SC25K1)

Ici aussi, il n'y a pas d'utilisations de communaulecte, et tous les utilisateurs sont tournés vers l'auteur du premier message (qui répond aussi en 7). L'utilisatrice 3 a recours au maori (« Kia Kaha.Arohanui »), une langue qu'elle doit parfois entendre (elle est originaire de Nouvelle-Zélande), qu'elle rend plus solennel avec l'utilisation de points (qui ponctuent aussi la traduction, « May you be strong.Much love. »). Les discours empathiques sont tournés vers celui qui souffre, et ceci court-circuite l'aspect communaulectal, pour privilégier la communication centripète, autour de lui. Les messages 19 et 20 ont été écrits par le même utilisateur, qui veut simplement corriger sa faute d'orthographe (« sure » écrit <sue> dans son premier message). L'écriture communaulectale, qui nécessite une réflexion hors-émotions pour l'utiliser, ne semble pas avoir sa place dans les discours empathiques de ce corpus.

4.5.2 Fonction complice du terme communaulectal

Les graphies alternatives, et autres communaulectes, représentent des actes de connivence et celle-ci doit être positive. Les fils de discussion où les émotions négatives (peur, deuil, empathie…) sont évoquées entraînent une proportion plus basse d'utilisations communaulectales.

Les agressions s'accompagnent de tachygraphies ou de graphies libératoires, contournant la cen-

sure. Les attaques envers un membre, effectuées par un groupe, contiennent plus de graphies alternatives : les liens interpersonnels entre les agresseurs se retrouvent renforcés par le fait qu'ils soient tous d'accord sur le fait que l'agressé mérite l'agression, comme le montrait la petite annonce de Peter D. tournée en ridicule par tous les autres intervenants de SC. Ces échanges ressemblent à la connivence sur les émotions positives (fils laudatifs, joie, soulagement…), où les utilisateurs montrent plus d'expressivité.

Le style personnel est toutefois une donnée importante, qui peut influencer les interactions : l'identité passe également par les fautes, la tachygraphie ou un style qui démontrent l'expressivité de chacun de manière différente. Les graphies alternatives et autres pratiques langagières se diffusent de manière virale, d'un utilisateur à un autre. Certains termes communaulectaux renforcent par leur essence même les liens entre les utilisateurs. C'est le cas des démonymes virtuels, ou noms donnés par les membres à l'ensemble des membres de la plateforme, comme *Scotties* ou même *Scotsters* sur SC (9 000 occurrences) , ou « Black Planeteers » sur BP. Ces noms sont directement dérivés des noms de ces plateformes, et sont attribués spontanément aux autres membres.

« Black Planeteers »renvoie à 7 700 résultats sur BP, principalement sur les pages personnelles et les légendes des photographies postées par les utilisateurs. Cette expression est souvent associée à des termes cohésifs, comme « to my Black Planeteers », ou « Hello Black Planeteers ». Si les membres les utilisent dans leurs pages personnelles, que l'on imagine rédigées avant leur participation au forum et outils de communication, ce sont donc des créations personnelles, spontanées, qui renforcent et créent l'identité communautaire. Ces mots se doublent de renforts grammaticaux : « my » (pronom inclusif en diable) ; ou même « fellow » (« Hello to all my fellow Black Planeteers out there and thank you for taking the time to visit my page. »). Le remerciement, en plus de l'association du démonyme, du pronom inclusif et de l'adjectif *fellow*, prouve la forte volonté de complicité exprimée par l'utilisatrice. C'est l'ouverture de son profil personnel, et l'utilisation d'un vocatif, qui montrent sa volonté de s'intégrer à ces « planeteers » peu définis. L'expression se retrouve aussi dans les livres d'or : « Just dropping a line to say welcome to BP and that I hope you enjoy your stay here on the Planet. I see I am the first to hit your G-Spot so keep it real and spread some luv to the rest of our REAL Planeteers. Drop by

and see me sometimes » (BPGB1). *Planeteer* est donc utilisé non seulement pour se sentir accueilli.e, mais aussi pour accueillir les autres et désigner la communauté (encore que l'expression « real planeteer », surtout avec la capitalisation de <real>, connote l'existence Planeteers non authentiques (?)). Les autres occurrences de l'expression de la plateforme sont rares (2 résultats supplémentaires), mais remontent à 2001, date de la création du site. Peut-être alors est-ce une façon de dire qu'il y a des utilisateurs réels, actifs, sur une jeune communauté—qui devait être attaquée ou se sentir peu légitime au vu de son ancienneté. Une occurrence plus récente, datant de 2004 emmène sur une autre piste : « Will the real planeteers please stand up ? » (BPPP11). L'emprunt à la chanson d'Eminem (« Will the real Slim Shady please stand up ? », (Eminem & Dr. Dre, 2000)). Encore une fois, il s'agit d'une ouverture de profil (qui correspond à l'introduction de la chanson), et l'expression n'a qu'une fonction phatique : il est sous-entendu que, si l'utilisateur ne veut parler qu'aux « real planeteers », il en fait aussi partie.

La pratique démonymique est un peu différente sur SC, parce que l'appartenance communautaire a tendance à être exprimée dans les équivalents des statuts (dans le *Pulse*) : « Today I am wishing my fellow scotsters happy St Andrews day ! » (SCP5) ; « Today I wish all my fellow Scotsters a happy Burns Night, have a wee dram for me ;) lol. » (SCP6). Ces messages sont toujours positifs, et concernent souvent les fêtes et célébrations diverses (que l'on se doit de fêter avec les gens de son cercle proche, donc de son cercle sur un réseau social). Il sert également dans les titres de fils de discussion (« Gay Scotsters, Friends and Family » (SC1A98) ; « Meeting up with Scotsters at The Gathering » (SC1H15) ; « Crafty/Talented Scotsters » (SC8AE9)), et ces démonymes servent d'autant mieux à réunir ensemble les membres de la communauté qui partagent les mêmes caractéristiques. Le mot « planeteer » est absent des titres de messages sur BP, mais 14 fils de discussion de SC portent le démonyme.

Il est aussi employé…par les créateurs du site eux-même, qui se servent du mot dès la page de création d'un compte sur la plateforme (cf image 4.16).

Illustration 4.16 – Page d'inscription au forum Scotster

4.5.2.1 Réaction aux utilisations

Les éléments communaulectaux, à leur paroxysme, peuvent être utilisés comme des mots-code, c'est-à-dire des termes iconiques. Ils ne sont connus que par certains utilisateurs. Ceci sert à les différencier des non-connaisseurs (les *noobs*[52], ou débutants). Sur SC, ces utilisations sont rares, mais le vocabulaire écossais peut troubler certains néophytes. Dans le fil de conversation SC2N59, une utilisatrice non écossaise se retrouve soudain dépassée par le nombre de mots en scots utilisés dans la conversation (exemple 4.12).

Message	Corps
84	Foo's yer doos ? aye pecking.
85	Oh, come on ! Help a poor Cdn out - "foo's yer doos ? aye pecking" ??????? Ye ha' tae be makin' that one up ! lol
87	Na, na ! He's nae makiin' thon up. Foos yir doos ? Peckin' weel ? Quite a common phrase nor o' the border.

Tab. 4.12 – Messages 84 - 87 de Loss of our language (SC2N59)

Ce qui commence comme un simple échange de mots de vocabulaire se transforme en cauchemar pour l'utilisatrice. Après l'intervention 84 (où, en fait, l'utilisateur utilise du vocabulaire dorique, et dit « Comment ça va ? [= How are your doves ?] Very good [=always pecking] », elle croit à une blague.

52. Anciennement appelés « newbies ».

Le troisième message, s'il n'est pas explicatif du sens de la phrase [53], la rassure quant à l'authenticité de la phrase, sans lui donner plus d'information. Le message 87 montre une certaine supériorité, de la part de ceux qui maîtrisent le code. Sans expliquer, l'utilisatrice en donne une variante fonctionnelle, et rajoute que la phrase est utilisée au nord de la frontière, c'est-à-dire pas en Angleterre, d'où elle imagine la rédactrice du message 85. Connaître le scots, c'est ne pas être anglais, et donc avoir une certaine supériorité linguistique. Cependant, elle n'est pas remise en cause, et les interventions ultérieures ne feront plus mention de cette expression, ou d'une gêne linguistique (au point de penser que l'expression n'existe pas) ; et les utilisateurs en présence continueront à converser comme si de rien n'était.

Une recherche sur « don't understand » n'est pas concluante, parce que les utilisateurs ne comprennent souvent pas les arguments qui leur sont avancés, plus que les écrits des autres : « I don't understand why you took personal offence on the Iranian Authorities renaming one of their own street after a an IRA member. »[sic] ; « I don't understand your point here. » (SC11A23). En revanche, certains prennent des précautions pour parler aux autres (les non-Écossais qui s'intéressent à la culture, par exemple), et, lorsque le ton est bon enfant, certains proposent leurs services pour « traduire » aux autres locuteurs. C'est surtout le cas de SC, où le scots est réellement perçu comme une langue étrangère pour certains. Dans l'exemple 4.13, l'auteur du fil liste des blagues en scots, censées n'être compréhensibles que par les « natifs ».

Message	Corps
3	I'm not hanging out with enough Scots to get these. Where are my translators ?
10	Any words you don't understand please do not hesitate to ask, we at Scotster are only too willing to help

Tab. 4.13 – Messages 3 et 10 de Scottish Jokes (SC2Y29)

Même si les message 3 et 10 n'ont pas de suite, ils montrent la bonhommie des membres les uns envers les autres [54]. Les autres utilisateurs continuent la discussion naturellement, mais ce n'est pas toujours le cas dans toutes les discussions, surtout lorsque le vocabulaire est au cœur des enjeux discur-

53. Sens que l'on ne lui donnera jamais dans le fil de discussion !
54. Ce qui va complètement à l'encontre de l'échange de 4.6.

sifs. Les rapports au pouvoir et à la culture communaulectale/communautaire sont facilement négociés sur SC. Sur cette plateforme, il est moins question de dominer les autres avec ses connaissances que de les faire partager, ce qui court-circuite la valeur agressive ou dominatrice de la maîtrise de la connaissance.

4.6 Le communaulecte, entre graphie et discours

La graphie est l'indice de l'utilisation communaulectale : le décalage par rapport à la normalité du discours, que l'on peut trouver ailleurs (hors plateforme) est en fait souvent contenu dans la graphie ou l'utilisation lexicale alternative. C'est la raison pour laquelle la graphie est le premier niveau d'étude, beaucoup plus facile à repérer que les pratiques communaulectales discursives, qui peuvent être écrites de façon standard. C'est d'ailleurs là où l'analyse de discours rejoint la phonétique—la graphophonématique, pour être plus précis— et où les deux types d'utilisation convergent. La conjonction d'indices graphiques et discursifs permettent de voir où les communaulectes apparaissent en ligne.

Au niveau discursif, le communaulecte apparaît de façon encore plus frappante. Beaucoup de membres décident non pas d'utiliser des graphies alternatives, mais de créer de nouveaux mots (démonymes virtuels, jeux de mots sur les pseudonymes…).

4.6.1 Originalité du discours

Pour comprendre l'originalité du discours en ligne, il faut délimiter ce qui est considéré comme « normal », ou « standard », en terme de production graphique.

La définition de la normalité est une gageure, mais la pondération d'apparition des formes graphiques originales dans le corpus, par rapport aux autres productions sur internet [55], dresse un tableau quelque peu différent des résultats d'étude dans le corpus, ce que l'on peut voir dans les études quantitatives du chapitre suivant, et dans l'annexe A.6.3 et A.6.1.2.

55. Qui sont, elles-même, parfois issues d'autres communaulectes…La boucle est sans fin.

La production moyenne sur internet est donc normalisée, lissée, ce qui explique que l'on trouve les valeurs maximales du moteur de recherche pour les lexèmes les plus fréquents de la langue anglaise [56].

Le communaulecte se situe aussi discrètement, dans des marques discursives qui flirtent entre graphies alternatives et graphies standardisées. Ainsi, sur SC, se joue un curieux changement lexical, qui illustre la différence de perception entre le premier et le second locuteur, dans l'extrait suivant (tableau 4.14), encore une fois tiré de « Loss of our language »

Message	Corps
51	Huv ye started writing in **Scots** yet ? I hae tae watch masell fer ma Uni work cos I have a habit aye writing hings the wi they soond then having tae chinge it later. I think i'll mak ma spellchecker learn Scots,Ha ha ha ['*Have you started writing in Scots yet ? I have to watch myself for my university work because I have a habit of writing things the way they sound, then having to change [them] later. I think I'll make my spellchecker learn Scots*'].
52	Lol Michael. Its going tae tak me a wee whiley to learnt tae spell in **scottish** but am picking up understanding more of the written stuff. While I did nae get all of what you wrote above I did get the jist of it.

Tab. 4.14 – Messages 51 - 52 de Loss of our language (SC2N59)

Les deux utilisateurs ne entendent pas du tout le scots de la même façon dans ces deux messages [57]. Alors que le premier parle d'écriture en scots, la seconde parle de graphier (différence entre « write » et « spell ») en écossais. Il sait que sa façon d'écrire n'est pas acceptée dans son travail universitaire, mais il se pose en garant de la langue quand il se tient au courant des progrès de l'autre membre du forum. Il pourra l'aider, et elle reconnaît la supériorité de Michael dans sa réponse : « I did nae get all of what you wrote above ». Le scots, la façon de l'appeler aussi (de simple « écossais » vague, à langue minoritaire à part entière) jouent sur l'appartenance communautaire. L'utilisatrice fait des efforts pour inclure des mots de scots dans sa réponse (« tae, tak, wee, whiley, nae »), alors qu'il dit au contraire avoir du mal à ne pas écrire en scots dans des contextes hors-ligne. Ce qui est standard pour lui (écrire en scots) explique sa supériorité linguistique sur SC, mais lui pose des problèmes hors plateforme, ce qui est certainement l'inverse pour elle, qui commence à utiliser et à comprendre le scots.

56. Prise comme un tout standardisé, entre l'anglais britannique et américain, sujet abordé dans (Crystal, 2011).
57. Le lecteur attentif aura reconnu un message de Michael et une réponse de Julia, déjà cités plus haut.

4.6.1.1 Communaulecte graphique

Les graphies alternatives obéissent à des règles particulières, qui sont issues de tous les utilisateurs des forums. Certains utilisent leurs graphies propres, qui ricochent à leur tour sur les graphies des autres utilisateurs (cf 4.3.2).

D'autres sont plus souvent utilisées sur le site à cause des sujets traités (*Amerikkka*, par exemple), et la communauté de pratique s'accompagne d'une communauté de sujets traités (l'agglomération autour de sujets communs, comme décrit par Wenger dans sa description de la communauté de pratique, présentée dans le chapitre 3). Ces sujets sont connexes à la création lexicale et graphique, et, par extension, responsables en partie des communaulectes.

Je défends également l'idée que le communaulecte graphique soit une utilisation née d'initiatives personnelles, individuelles, copiées ou adoptées simultanément dans les communautés virtuelles. Par exemple, il est facile d'imaginer plusieurs forgeurs inconnus pour la graphie <Amerikkka>, ou plusieurs membres décidant de ne pas relire leurs écrits ou d'adapter leurs graphies pour des raisons différentes, mais avec des résultats similaires. Plus retors, on peut très bien imaginer une communauté de membres dont le communaulecte coïncide avec la norme–ce qui les rapproche entre eux, et éloigne les gens utilisant des graphies alternatives. Le discours anti graphies alternatives est d'ailleurs monnaie courante sur internet, comme le montre l'illustration 4.17 en français. Il ne me semble pas exister une telle véhémence sur les plateformes en ligne en anglais—à l'exception, évidemment, de la figure du Grammar Nazi, qui est une figure solitaire. Il existe cependant de nombreuses variations sur le fait que le discours sur internet doive plus ressembler à du langage écrit standard qu'à des SMS.

> Yea I noesum uv ma wurds iz spelt wrong…I did dat for a reason…I'M GHETTO so HATE IT or LOVE IT (BPPP1)
> '*Yeah, I know some of my words are spelt wrong…I did that for a reason…I'm ghetto so hate it or love it*'

L'utilisatrice instaure une relation de causalité entre sa graphie stylisée, et le fait qu'elle soit « ghetto », donc une personne qui pratique un anglais vernaculaire, et sur laquelle la dimension normative de

Illustration 4.17 – Variante de « Ceci est un forum, pas un portable »

l'orthographe n'a volontairement pas de prise sur ses écrits sur internet. La déviance par rapport à la graphie standardisée fait partie intégrante de la présentation de soi, de son identité, en ligne.

De leur côté, les normativistes virtuels ont également pris le parti d'éviter les tachygraphies et autres écarts par rapport à la normativité, pour montrer leur identité en ayant recours au respect de la langue, ou d'une certaine idée de la langue.

Conclusion du chapitre

Dans ce chapitre, j'ai peint le portrait des communaulectes, dans leur contexte d'utilisation (ou en extraction) dans les deux communautés virtuelles. J'ai également exploré les différences et les ressemblances entre les fautes d'orthographe et les graphies alternatives, articulées autour des équivalences graphophonématiques, dans les parties 2 à 4. Les utilisateurs ont recours à des graphies alternatives à cause de la complexité de ces relations entre le son et la graphie. Dans la cinquième partie, j'ai examiné le rapport entre ces utilisations alternatives, et le contexte (positif, agonal …), et j'ai montré l'importance de la cohésion entre les membres de la communauté—qui est la première composante pour un

communaulecte riche et varié. Enfin, cette cohésion est explorée dans la dernière partie du chapitre, où la notion de communaulecte est étendue au discours dans le sens large ; le communaulecte est aussi constitué d'unités discrètes, de phrases cryptiques à la graphie ou à la syntaxe irréprochables qui s'y inscrivent tout de même.

Le communaulecte n'est pas seulement discursif ou graphique, il se situe entre les deux niveaux, ce qui le rend d'autant plus difficile à cerner. Pragmatiquement parlant, les traces communaulectales sont plus dures à repérer puisqu'elles sont plus diluées, mais elles sont plus fréquentes que les graphies alternatives. Les graphies alternatives, d'ailleurs, sont classables en plusieurs catégories, que je décris ainsi :

— Les graphies contournatives qui servent à se débarrasser de la censure, sur BP, ou à ne pas écrire entièrement un mot jugé vulgaire (surtout sur SC),

— Les graphies démarcatives que les auteurs utilisent pour se différencier de l'anglais standard, consciemment ou non. Elles peuvent parfois servir à prouver son appartenance à la communauté, et passent par des mots-clefs faciles à reconnaître,

— Les graphies identitaires individuelles, issues du désir de se démarquer personnellement par rapport aux autres (utilisation de majuscules, de graphies récurrentes non imitées), et qui sont très proches de la catégorie suivante,

— Les graphies culturelles dictées par une sous-norme (non le standard officiel, mais un standard de la culture à laquelle l'on appartient). Ainsi, un protestant écossais qui se fait appeler « prod-die », ou un membre de gang autoproclamé qui recopie des phrases rituelles, utilise un discours culturel.

Les pratiques graphiques sur les communautés virtuelles sont idiosyncratiques : les mots s'affrontent, se créent et s'effacent au gré de leur viralité. Les pratiques numériques et identités linguistiques dans les communautés virtuelles sont volatiles, comme le montre le syntagme « real planeteer », qui ne fut employé qu'à la genèse de BP.

Les pratiques numériques en présence ne sont en fait pas toutes utilisées en même temps : chacun s'exprime en fonction de son style (qui peut être une combinaison de tous les types de graphies dont le

standard grammatical et lexical). La subjectivité est très importante, tout comme le contexte au sens large, et ceux qui utilisent des insultes sur BP sans utiliser de graphie contournative sont condamnés à être censurés par le script de censure. Les stratégies linguistiques employées, même si elles ont des buts très différents, sont souvent similaires : le discours est mis en forme, de la simple utilisation de majuscules pour appuyer son propos, jusqu'à la modélisation phonétique du discours (ce qui sera abordé dans le chapitre suivant), en passant par l'hyperexpressivité et les diktats des tabous et autres pratiques identitaires individuelles et culturelles. La seule frontière à ne pas dépasser lorsque l'on utilise des graphies communaulectales est la compréhension : l'autre doit pouvoir comprendre. Le cas échéant, s'il montre son inexpérience sur la plateforme, il court le risque d'être raillé. De plus, il est évident qu'être compris par tout le monde (tant qu'il n'y a pas de langue minoritaire en jeu) est très important sur des réseaux sociaux censés fédérer leurs utilisateurs.

Les enjeux de SC et BP sont différents. Le statut du scots en tant que langue minoritaire, reconnue par la plupart des membres de la plateforme est tout à fait différent des graphies alternatives nées sur la communauté afro-américaine [58]. L'âge moyen des membres, qui est plus élevé sur SC que sur BP explique la différence de virulence lors de disputes sur les deux communautés. Les batailles identitaires, passant par l'affirmation de soi par son écriture, n'existent pas sur SC, et laissent plutôt place à l'utilisation d'un scots adouci (seuls quelques mots-phares sont employés, comme « wee », par exemple) et compréhensible par tous, ce qui va totalement à l'encontre des copiés-collés massifs de phrases rituelles chez les membres de gangs autoproclamés sur BP. L'attitude générale entre les membres d'un petit forum communautaire pour passionnés d'une nation est plus clémente, tandis que BP a trop de membres pour que les scripteurs actifs se connaissent tous entre eux (excepté ceux qui postent des messages dans les mêmes sections), ce qui peut provoquer une certaine méfiance à l'encontre de ceux que l'on ne connaît pas, noyés dans le flot d'informations de *What's New ?* par exemple.

Les pratiques communaulectales sont donc valorisantes intrinsèquement : on prouve que l'on fait partie de la communauté à part entière puisque l'on en maîtrise les codes—et l'on prouve par ceci son désir d'y appartenir. Le coénonciateur n'est pas laissé pour compte, puisque s'il comprend le message sans problème, cela veut dire qu'il fait partie de la communauté, des connaisseurs du site—et qu'il ne

58. Celle qui fréquente le site, au moins, et celle que j'ai pu étudier aussi.

fait pas partie des *noobs*. D'ailleurs, quand bien même serait-il un noob, qu'il aurait le droit (surtout sur SC, et hors discussion agonale) de demander des explications linguistiques. Ceci montrerait sa faiblesse communaulectale, mais le rapprocherait du ou des membres qui voudraient bien répondre à sa question, afin de le faire entrer dans la maîtrise du communaulecte.

C'est un cercle vertueux dans lequel certains internautes se (re)trouvent, autant en tant que membre du groupe valorisé qu'en valorisateur des autres. Ceci explique la multiplication des marques communaulectales, des plus discrètes (pronoms personnels, ton familier avec certains autres membres), qui sont souvent tournées vers l'autre, aux plus extrêmes et stylisées, qui prouvent surtout le désir de personnalisation et d'unicité de certains.

Dans le chapitre suivant, je vais donc pouvoir examiner les liens entre écriture orale de sa langue minoritaire (l'écriture de ce que l'on prononce) et les écrits graphiquement alternatifs sur les deux plateformes. Je postule l'existence d'une grille d'équivalences mentales entre les phonèmes et les graphèmes, ce que j'ai effleuré dans ce chapitre, qui servirait assez systématiquement à remplacer des séquences de lettres dans les mots, afin de faire épouser plus parfaitement la graphie à la phonie.

Chapitre 5

Graphophonématique des communaulectes

« Am I just to be like a clockwork orange ? »
I didn't know what made me use those slovos, brothers,
which just came like without asking into my gulliver.

A. Burgess - *A Clockwork Orange*

Dans le chapitre précédent, j'ai délimité le communaulecte et ses conditions d'apparition, mais particulièrement dans ce corpus, il comprend aussi l'adaptaton graphique de la phonologie des langues minoritaires. Les deux utilisations se complètent dans des écrits sur internet, et il est parfois difficile de les séparer. Ce qui m'intéresse ici, ce sont les indices phonologiques de l'oralité dans les écrits, qu'ils proviennent d'une langue minoritaire ou d'une pratique orale sociolectale.

Les conditions d'apparition des écrits purement dialectaux seront étudiées, ainsi que les variations graphiques autour des consonnes et des voyelles : si certaines sont dialectales, d'autres sont adoptées à cause des particularités des plateformes (par exemple, présence de scripts de censure ; mais aussi surnoms de membres, ou noms de parties de *Scotster* et BP, comme la rubrique « Heilan Coo » de SC). Certaines graphies peuvent être des fautes d'orthographe qui se retrouvent réemployées par les autres membres, atteignant ainsi un statut communaulectal.

Ainsi, il existe des variations graphiques à substrat oralisant. Certaines variations appartiennent à la langue minoritaire, et le scots est une variété aux caractéristiques polylectales, parce qu'il est scindé en plusieurs dialectes, dont le système phonologique varie. Ainsi, le mot *after* est un exemple qui possède plusieurs variantes fonctionnelles. Selon un dictionnaire (Eagle, 2002), en scots, il peut se traduire par soit *efter*, soit *ahint*. Le DSL offre une définition plus précise, et circonscrit l'utilisation de *ahint* aux îles du nord (Orcades), ainsi qu'à des régions du sud est (Roxburghshire, Fife et Angus). *Efter* semble utilisé partout ailleurs. Les deux dictionnaires reconnaissent cependant pléthore de variantes (pour le dictionnaire d'Eagle, *ahint* est une variante d'*efter*) : <efter> est la graphie la plus commune, mais il existe *eftir*, *aftyr*, *aftir*. Pour *ahint, ahin, ahen, ahent* et *ahaint* sont répertoriées. Presque toutes ces variantes se trouvent sur *Scotster*, comme l'indique le tableau 5.1.

Ces utilisations diverses penchent pour une utilisation réaliste, plus commune d'*efter*, mais les deux

<Efter>	<Eftir>	<Aftyr>	<Aftir>
78	0	0	3
<Ahint>	**<Ahin>**	**<Ahent>**	**<Ahaint>**
7	5	4	0

Tab. 5.1 – Nombre d'occurrences d'[efter] et [ahint] sur Scotster

mots sont en concurrence sur le site (notamment parce qu'*ahint* signifie également '*derrière*' « I'll help oot ahint the bar » (SCP4) '*I'll help out behind the bar*'). La graphie préférée correspond à la plus proche de l'anglais, et sa distance de Levenshtein (exploitées en 5.5) par rapport à l'orthographe n'est que d'un (comme les graphies <aftyr> et <aftir>). En revanche, cette graphie est la meilleure pour retranscrire la prononciation /ˈɛftər, ˈeftər/, qui est logique pour la graphie <efter>. Le choix d'une voyelle finale plus haute graphiquement (dans les variantes <aftir>/<aftyr> à la place de <e>) est certainement influencée par la rhoticité et la prononciation [ɘɾ] finale, dont la consonne provoque une fermeture de la bouche plus importante qu'un schwa final.

Dans cette partie, j'étudierai donc la graphophonématique telle qu'elle est présentée dans les ouvrages théoriques de phonologie (Deschamps, Fournier, Duchet & O'Neil, 2004 ; Giegerich, 1992). je m'appuie également sur les travaux plus orientés vers la graphie et l'orthographe (Catach, 1978). Le grand problème des travaux sur l'orthographe est leur inclinaison souvent prescriptiviste,

souvent dans le but de travailler sur l'acquisition du langage (University of Texas System, 2009), ce qui n'est pas la dimension du travail de N. Catach. Comme précisé dans le chapitre 2, les recommandations de l'Écosse concernant l'écriture sont les plus ouvertes (Education Scotland, 2012). Nous pouvons prendre comme point de départ réflexif cette citation assez longue de Chomsky et Halle, lorsqu'ils commencent à appréhender le système accentuel de l'anglais :

> The fundamental principle of orthography is that phonetic variation is not indicated where it is predictable by general rule. […] Orthography is a system designed for readers who know the language, who understand sentences and therefore know the surface structure of sentences. Such readers can produce the correct phonetic forms, given the orthographic representation and the surface structure, by means of the rules that they employ in producing and interpreting speech. It would be quite pointless for the orthography to indicate these predictable variants. […] [Conventional orthographic systems] are designed for the use of speakers of the language. […] English orthography, despite its often cited inconsistencies, comes remarkably close to being an optimal orthographic system for English. Correspondingly, it would not be surprising to discover that an adequate theory of the production and perception of speech will find a place for a system of representation not unlike orthography, though there is, for the moment, little evidence that phonemic transcription is a "psychologically real" system in this sense.
> It should be also observed that very different dialects may have the same or a very similar system of underlying representations. (p.49) (Chomsky & Halle, 1968)

Les « variations prévisibles » des langues minoritaires sont justement au cœur des communaulectes : certes, ce n'est pas l'orthographe qui compte, mais la beauté de la communication en ligne, lorsqu'elle mélange codes endogènes et exogènes, c'est-à-dire codes de la communauté virtuelle et d'internet en général, et variations graphiques issues de la volonté de retranscrire son parler. Ces graphies alternatives pallient le manque d'optimalité de retranscription des langues minoritaires à fort substrat anglais. Les valeurs des voyelles de l'anglais, assez stables suivant leur contexte phonologique, ne valent plus en scots ou en AAV : un <r> postvocalique ne colore pas les voyelles de la même façon dans une langue rhotique.

Les scripteurs peuvent produire les formes phonétiques correctes, ou standard, à l'oral, tout en choisissant d'écrire à l'envi la forme graphique qui reflète le plus ce qu'ils pensent. Cette forme graphique peut s'inspirer de celles qui sont attestées dans la plateforme qu'ils fréquentent. Utiliser des graphies alternatives est plus que bienvenu sur SC, où aucun *Grammar Nazi* ne viendrait corriger la graphie scottisante ; et les graphies oralisantes sur BP, notamment celles des diphtongaisons des voyelles (<dayum> pour « damn »), ne sont pas sanctionnées non plus.

Le système de représentations sous-jacentes que j'étudie ici diverge de l'anglais, et de l'orthographe, parce qu'il n'y a pas qu'une seule représentation pour un mot donné. Les variations allographiques de « after » représentent en fait les transcriptions phonémiques (et variations allophoniques) dont Chomsky et Halle parlent. Elles représentent un système psychologiquement proche de la vision mentale que les locuteurs ont de leur langue. La grande difficulté de cette étude est la disparité de ces variations allographiques. Les échanges sur internet ont une forme éminemment libre, et les pratiques présentées sont exclusivement standardisées par la communauté : les variations stylistiques sont incorporées dans les strates du communaulecte.

J'étudierai donc les lexèmes dont la forme graphique s'éloigne du standard et comparerai les résultats aux travaux théoriques sur la phonologie des langues minoritaires, pour voir si ces études sur terrain graphique donnent les mêmes résultats que leurs contreparties basées sur des productions orales. Les différences graphiques sont souvent dues à la prononciation des locuteurs, surtout en scots, où des dictionnaires listant les variantes graphiques sont disponibles. Ces variantes graphiques ne sont pas toujours connues par les locuteurs, mais leur façon de représenter un mot est souvent influencée par la prononciation. Je me pencherai sur les variations de la graphie des consonnes par rapport au standard orthographique, réparties par leur mode d'articulation, afin de comparer différences suivant le degré de marque phonologique (Zwicky, 1972), et je ferai la même chose pour les voyelles brèves d'abord, les représentations du schwa, puis les diphtongues.

Je décalerai ensuite la perspective d'étude du phonème au mot entier. Je me tournerai donc vers l'inventaire des variations graphiques de mots à forte indexicalité, ou à forte iconicité dans la communauté afro-américaine. Je traiterai également les problèmes d'homophonie, et de conflits graphiques éventuels, ce que j'appelle les « limites allographiques » du communaulecte. Parfois, les graphies alternatives ne sont pas compréhensibles, à cause de l'homographie (qui suggère une homophonie). Ces homographies peuvent freiner le travail de constitution du corpus, surtout pour les mots courts, qui représentent presque toujours une acronyme. Pourtant, la plupart des graphies alternatives arrivent à être compréhensibles en contexte, à refléter la phonologie des langues minoritaires, ou au moins, un accent ou une emphase particulière–quand elles ne sont pas utilisées pour censurer ou euphémiser un

mot. La flexibilité des variations graphiques n'éclipsent pas la compréhension, dès que la phonologie est mise à contribution pour déchiffrer les transcriptions de la prononciation.

5.0.2 Remarques méthodologiques

L'analyse et le tri des données sur les mots courts est beaucoup plus difficile ; les mots jusqu'aux trigraphes peuvent être des abréviations, des acronymes ou même des mots étrangers, surtout avec les lexèmes communs avec les langues germaniques. L'interjection « och » est aussi un acronyme désignant divers organismes (Olympic Chess House, Off-Campus Housing, Ottawa Community Housing…). Cette interjection existe dans le corpus SC, mais les relevés sur internet sont faussés par le volume de mots à écarter. Pour chercher « Ben », dans le sens de *montagne*, il faudrait éliminer tous les résultats portant sur le prénom Benjamin, ce qui est impossible à moins de lire tous les résultats, qui représentent parfois des dizaines de milliers de pages. Les confusions avec des toponymes ou des noms propres peuvent également arriver avec les autres mots, comme Bealin, nom de famille courant. Sur SC, <bealin> veut dire '*en colère*', et il est très difficile d'écarter toutes les références aux Bealin. Il n'est pas toujours possible d'écarter tous les résultats non pertinents, parce que Google restreint son champ de recherche à 50 critères. La recherche sur la graphie <thouless> (SC), '*thoughtless*' devait écarter des prénoms, des patronymes, et autres éléments. La recherche sur Google était donc « thouless -david -robert -michael -kosterlitz -brian -john -nightingale -mark -angela -theorem », où les <-> devant les mots éliminent toutes les pages qui les contiennent. Sur BP, la recherche sur <wit>, *with* était « wit -intelligence -institute -FM -movie -world -studio -antoni -associes -women -optimal -github -jan -jacob -plumbing -technology -internet -delight -ISO ».

Les caractères non-alphanumériques sont aussi problématiques, parce qu'ils sont considérés comme un ensemble, et pas en eux-mêmes. C'est pourquoi les apostrophes d'excuse (ou paroissiales) repérées dans le corpus ne sont pas traitées à part : à moins de lire les fils de discussion sans outil de textométrie, il n'est pas possible de faire des études quantitatives sur leur utilisation.

Dans le corpus BP, il y a de nombreux remplacements graphiques de lettres par des chiffres et autres

caractères extra-alphabétiques. Ces remplacements repoussent la limite de la graphie au clavier, pourtant si codifiée (il n'est pas possible de choisir la police de caractère, ou le nombre de caractères sur un clavier. La matérialité est très importante, et l'environnement écologique des échanges les influencent (voir le tableau 3.1). Dans certaines graphies alternatives, notamment dans les études de cas de variances allographiques, les lettres sont remplacées par des caractères alphanumériques. Le tableau 5.2 présente tous les remplacements observés dans le corpus.

Voyelles			
<a> <@>	<e> <3>, <#>	<i> <y>, < !>, ^, <'>	<u> <v>
Consonnes			
<c> <(>	<g> <6>, <9>	<l> <1>	<s> <z(z)>, $

Tab. 5.2 – Exemples de substitutions alphanumériques dans le corpus

5.1 Production en langues minoritaires

Les écrits entièrement rédigés en langue minoritaire sont rares, et souvent fruits de volontés particulières de la part des scripteurs : les exemples qui suivent dans cette partie en attestent, surtout parce que l'alternance codique est tellement intégrée qu'elle est pratiquée naturellement, à l'écrit comme à l'oral [1].

Sur SC, ces écrits sont souvent pédagogiques, soit de la part d'utilisateurs qui aident les autres à acquérir du vocabulaire, soit de la part d'apprenants, qui testent leurs connaissances en scots. Le scots est aussi utilisé à des fins humoristiques, ou stylistiques dans certains cas. Ils vont ainsi complètement à l'encontre de l'idée de l'alternance codique, qui permet un passage d'un sociolecte à l'autre, et ainsi le détachement cohésif (et émotionnel) apporté par l'alternance.

1. Les tiraillements d'alternance codique entre le scots et l'anglais sont décrits diachroniquement par J. Costa, qui décrit le statut de « full language » du scots suivant le temps. On peut transposer cette discussion au statut du scots dans l'alternance codique parlée synchroniquement, et donc utilisée en ligne (Costa, 2009).

> I love using my own Glaswegian on here but usually write in English because that is what I was taught and that's what most folks would rightly understand. (SC5M9)

Le message défend l'idée de la compréhension la plus partagée pour ne pas utiliser une langue qu'il aime utiliser. L'attache affective personnelle passe après le partage au plus grand nombre. Il ne faut pas oublier ce point de vue sur les langues minoritaires, et sur les graphies alternatives. L'intelligibilité est un facteur important, et les scripteurs de forums sont conscients que l'utilisation de graphies alternatives peut être un obstacle pour les autres. Dans cette partie, et surtout dans la sous-partie suivante, je parlerai de cas d'inintelligibilité entre membres de forums.

La définition d'un message en langue minoritaire peut varier. Le scots, en tant que langue minoritaire reconnue comme telle par ses locuteurs, sera utilisé pour créer les repérages. Il ne faut pas oublier que certains fils de discussion sur Scotster sont écrits en gaélique, mais la proportion de fils et de messages écartés de l'étude est négligeable, parce que peu d'entre eux sont en gaélique. 'A titre d'illustration, il n'y a que 932 occurrences de « failte » ('*bienvenue*'), et pourtant cette formule s'inscrit tout à fait dans le folklore linguistique –et touristique–écossais. Les membres de SC ne s'embarrassent donc pas de formules stéréotypées, et préfèrent demander des traductions de l'anglais au gaélique, surtout dans la partie SC2N, « Scottish Languages »

De rares fils de discussion proposent d'utiliser le gaélique, comme « Practise your Gaelic here ! Beginners welcome too ! » (SC2N64), qui annonce clairement la couleur. Même les messages de ce fil sont parsemés de traductions ou de compléments en anglais, et seuls 17 messages sur les 100 premiers sont entièrement rédigés en gaélique, et n'ont été écrits que par trois utilisateurs du forum, gaélophones confirmés, soit une minorité des 13 intervenants. En revanche, les connaissances gaéliques permettent à ces trois membres d'occuper la plupart de l'espace discursif, et les trois intervenants, Isobel (créatrice du fil de discussion) a 34 messages à son actif, Pòl, 31 messages et Tam 22, soit 87 messages sur les cent premiers.

Sur BP, une caractérisation similaire de la compétence des scripteurs est beaucoup moins évidente. Les pratiques discursives oscillent entre mélange d'anglais, formes grammaticales afro-américaines, graphies alternatives ambiguës, entre netspeak et AAV. Les utilisateurs possèdent plusieurs niveaux de maîtrise linguistique, plusieurs littéracies, ou compétences pour écrire. On peut citer les usages du

netspeak, qui est beaucoup plus répandu chez les membres de BP ; les utilisateurs qui ne maîtrisent que le communaulecte de la plateforme, comme les membres de SC, et qui ont très peu recours aux initialismes propres à internet, par exemple. Les strates graphiques du scots, dont certains mots sont connus (le morphème <-ae> de négation pour les auxiliaires et modaux ; « wee » ; « ken », mais aussi « radge » '*jeune femme*' ; « dreich » …). Sur BP, certains utilisateurs n'utilisent ni mot en langue minoritaire, ni netspeak, mais changent subtilement la graphie d'un mot (ou concatènent deux mots ensemble) pour ajouter un faisceau sémantique supplémentaire. Le remplacement de <c> par <kkk> est obligatoirement conscient, dans un mot comme <Amerikkka> ; et les <coonservatives> sont certainement des « coons » '*terme dégradant pour les Noirs*'conservateurs. Certains messages, comme la citation ci-dessous, alternent fautes d'orthographe (« hosptial »), marques grammaticales de l'AAV comme la double négation, « aint » et marques prosodiques, comme la compression de « cuz ». La logique narrative afro-américaine est perceptible dans le syntagme « lazy fat azz people », qui ajoute les adjectifs les uns derrière les autres, sans ponctuation.

> I don't feel like writing this report on improving hosptial cleanliness…
> why I gotta do this…
> the hospital aint clean…it aint never gonna be clean…cuz the nursing and EVS staff are all lazy fat azz people that don't work because they have union support…
> dayum unions…(BP12E763)

Pourtant, cette citation est aussi écrite en grande partie en anglais standard, et reste très compréhensible pour les non-locuteurs de l'AAV. À l'inverse, la cryptographie de gangs étudiée au chapitre précédent reste entièrement dans son sociolecte fermé ; mais ce sociolecte n'est presque employé que dans les pages personnelles. Plus le discours est ouvert et lisible par les autres, plus il est général, et moins les alternances codiques vont être importantes. Pour un autre sociolecte d'internet, le lolspeak, la majorité des messages du forum se situaient dans une fourchette entre 40% et 70% de graphies alternatives (Goudet, 2012). Même lorsque le but est d'utiliser les pratiques discursives alternatives, les messages n'arrivent pas à 100% de graphies alternatives.

Comme le souligne W. Ong, « The process of putting spoken language into writing is governed by consciously contrived, articulable rules : for example, a certain pictogram will stand for a certain

specific word, or *a* will represent a certain phoneme, *b* another, and so on. » (Ong, 1982), et cette réflexion est tout à fait juste, d'autant plus que la forme moins formelle des échanges sur internet permet une certaine oralité. Les phonèmes et sèmes rajoutés sont contenus dans l'adjonction d'une lettre, et l'acte d'écriture est sans cesse conscient. La conventionnalité des sons représentés par les graphèmes permettent des combinaisons non-standard, plus évocatrices de transcriptions phonologiques que de graphie standardisée.

5.1.1 Étude des limites perçues du scots

Les conditions de production sur les deux forums doivent être examinées. Tout d'abord, existe-t-il des écrits purement dialectaux à 100% ? La faible fréquence de tels messages s'explique par l'alternance codique entre l'anglais standard, le netspeak, la langue minoritaire… Écrire en anglais standard est normalement automatisé, et fait partie d'un système maîtrisé ; rajouter des éléments de netspeak montre la maîtrise de quelques pans de la culture internet. De plus, les utilisateurs n'ont souvent pas d'autre occasion de pratiquer à l'écrit ce qu'ils pratiquent à l'oral, et l'alternance codique est fortement pratiquée à l'oral, surtout dans le cas du scots. Trudgill souligne la différence entre l'alternance codique, et le changement de registre linguistique propre au scots, qui n'est pas un anglais (parmi les anglais du monde [2]) : « Native speakers of Lowland Scots dialects may switch, in relatively formal situations, to standard English (spoken with a Scots accent of course). » (Trudgill, 1983). Ces observations sont beaucoup plus faciles à conduire sur le scots que sur l'AAV, et cette partie sera entièrement consacrée au premier. L'AAV n'en est qu'aux balbutiements d'une reconnaissance en tant que langue à part entière, comme vu au chapitre 2, et l'absence de dictionnaires ou de travaux universitaires sur les graphies de l'AAV freine ce type d'étude.

Les écrits rédigés en langues minoritaires sont parfois problématiques pour les locuteurs mêmes : le scots est utilisé de façon utilitaire dans certains fils de discussion (pour l'apprendre, ou pour entretenir son niveau) et les utilisateurs peuvent débattre quant à la qualité des mots proposés dans les fils de discussion didactiques, ce qui freine les utilisateurs moins confirmés, qui sortent de la conversa-

2. Traduction maladroite de « World Englishes ».

tion en cours pour demander des précisions lexicales. La variété des parlers du scots (des Lowlands jusqu'aux Orcades) explique aussi la difficulté que certains ont à comprendre les membres venant d'autres régions. J'ai trouvé un fil de discussion où les utilisateurs se disputent sur les utilisations lexicales, SC2N27 (échanges dans le tableau 5.3).

Message	Corps
71	I once bought a book by Billy Kay, I think and there was a TV program about it as well and it was hell to read. Mainly because he wanted to be so serious about the structure of the language, you'd need a diploma in English to understand the grammatical terminology. I aslo bought a Scots Dictionary, and in it, there was none of the words I was looking for. It seemed more intent on recording all the old words when the whole of Scotland spoke Scots, as well as our leaders. These words, in my opinion, are no longer relavent to modern Scots and unless you're a history buff, there's no point in learning them. Folk don't use the terms and words written by Burns. Anyway, here's a word I've not heard for a long time, I don't even know how to spell it, so I'll just spell it as it sounds. Hoodjamadoodjary(hood-ja-ma-dood-ja-ray) This seems to be used these days when referring to some machinery gizmo. I remember using it mainly when talking about a musical wind instrument.
72	[…] Is it true that the words we have been posting are completely irrelevant ? That would be disheartening. But at least we would have a nice broad Scots vocabulary !
76	The words we have been posting are NOT irrelevent ! Although certainly some of them are less common in some regions than others. Ayrshire dialect and words are entirely different to Glasgow, Borders, Edinburgh, Dundee, Aberdeen, Inverness or the Western Isles ! In particular, the Western Isles have a very strong tradition of Gaelic, and quite often 'sprinkle' their sentences with the odd Gaelic word ! On the other hand, Aberdeen, Elgin, Buckie area has a strong inclination towards Doric, which has a quite unique vocabulary ! If you wanted to 'speak like a native Scot' perhaps this 'Vocab List' might be more useful if each word was accompanied by the Region(s) and/or city in which it is still commonly used. Although I've lived in Kirkintilloch, Glasgow, Stornoway, and Elgin, as well as spending several days in Edinburgh, Aberdeen, Dundee, Perth and Ayrshire, and had no trouble at all in understanding and being understood !
77	I'm not talking about the words from various regions, but ancient words that are out of use. I thought that was clear, but mabye I'm as 'Thick as shite in the neck o' a bootil'. = Very stupid.
78	"The Scots Language is not a unified language. There are words they use in Ayrshire that are never used in the Lothians, and the same can be said for Tayside and Grampian, Grampian being more based on the Doric." (Citation de l'auteur des messages 71 et 77)

Tab. 5.3 – Messages 71 - 78 de Weekly Scots Vocab List (SC2N27)

La dispute continue sur les messages suivants, et dégage la problématique intéressante de la conservation des langues minoritaires par leur utilisation (et l'exposition aux lexèmes–par exemple, « stramach », '*victoire*' est utilisé par un des membres de SC parce qu'il l'a entendu à la télévision) ou, au

contraire, la relégation de ces mots aux « history buffs » et autres philologues et autres dictionnaires, comme le DOST. Le livre évoqué dans le message 71 s'avère être, après lecture du fil entier, Scots – The Mither Tongue (Kay, 1986). L'auteur n'y fait pas d'inventaire de mots en scots. Le scots archaïque, rejeté comme étant celui de R. Burns, n'est pas le même que celui d'I. Welsh. Il n'y a pas de référence à cet auteur ou à tout autre auteur écossais contemporain comme I. Rankin pour comparer un scots littéraire synchronique et un scots archaïque.

Les utilisateurs passent tout de même par l'anglais standard pour commenter le scots. Les plus sérieux passent à l'anglais pur, le retour à la discussion légère est codé par l'utilisation du scots, qui sert, en fait de communaulecte cohésif en tant que tel : l'instigateur de cette dispute tente de faire diversion en réutilisant du scots (« thick as shite in the neck o' a bootil »), mais la discussion se fait surtout en anglais, la langue du prescriptivisme et de la raison. L'auteur du message 72 souligne la différence entre les types de scots, et surtout, les langues qui l'influence (le gaélique et le dorique). Les Écossais ont une connaissance accrue de leur langue, de leur accent et de la différence avec ceux des autres régions, et cette conscience se retrouve dans leurs écrits. C'est pourquoi l'argument d'autorité utilisé dans les messages 76 et 78 consiste à citer des régions ou lieux différents d'Écosse, et de dire que les usages linguistiques y sont différents.

C'est une constante sur le site, et les membres de SC justifient souvent le vocabulaire qu'ils emploient par une mention de leur appartenance géographique :

> Just catching up on the forum and noticed Genevieve was asking about words for kids. Well in aberdeenshire where i am the call the girls quines and the lads loons. Think thats how you spell it onyways. Learning to understand it is one things learing to spell it is something else but oh this is definitely the best topic on scotster. SC2N59-50

Plus les thèmes sont proches d'eux, dans leur environnement quotidien, et plus le recours au scots est facile. L'alternance codique est donc de mise pour les utilisateurs de forums, qui peuvent changer de style d'un mot à l'autre. Sur SC, les locuteurs du gaélique écossais introduisent parfois des éléments en gaélique, non à cause d'une demande de traduction, mais parce que c'est le seul mot qu'ils utilisent pour décrire certaines réalités locales, comme la nourriture, ou des éléments affectifs. Dans la liste thématique des éléments décrits sur SC, la nourriture représente une proportion assez importante de

mots en scots.

Message	Corps
135 136	they've got different words in Lewis (marag = black pudding) marag = pudding (Gaelic). A black pudding is (strictly speaking in Gaelic) marag dhu.

Tab. 5.4 – Messages 135 - 136 de Loss of our language (SC2N59)

Message	Corps	Traduction
28	my mither used tae say ah wis thouless , an ma da used tae tell me an ma wee sister tae stop cyangalin ,an when ma hauns were durralin wi the caul he telt me tae pit them under ma oxters.	My mother used to say I was incompetent, and my dad used to tell me and my little sister to stop fighting, and when my hands were numb with cold he told me to put them under my armpits.
29	Hi Jean,the only wan I got wis aboot yer sister pittin her haunds under her oxters tae keep them warm !! maybe a need tae go bile ma heid. anither wird is peelywally, fer no weil. keep them coming the mair wirds I learn the better.	Hi Jean, the only one I got was about your sister putting her hands under her armpits to keep them warm ! Maybe I need to go boil my head. Another word is *peelywally*, for *sick, unwell*[= Not well]. Keep them coming the more words I learn the better.
30	hiya michael, ken when yer hauns are fair sair wi caul ma mither used tae sae they were durralin , she wis frae embra,ma da wis frae galloway an cyangalin wis fichtin,if it wis yin o thon thunery days it wis mawkin, if ye wur that tired ye could lie doon ye were fair furfochin,……lol	Hi Michael, you know that when your hands are fairly numb [sore] with cold my mother used to say they were 'durralin', she was from Edinburgh, my dad was from Galloway and *cyangalin* was fighting, if it was one of those thundery days (?) it was *mawkin*, if you were so tired you could lay down you were fairly *furfochin* [=exhausted]…lol

Tab. 5.5 – Messages 28 - 30 de Loss of our language (SC2N59)

L'extrait 5.5 est particulièrement dense, à cause d'un mélange entre nostalgie (avec les répétitions de « used tae sae ») et volonté pédagogique : en partageant ces anecdotes familiales, l'utilisateur montre sa maîtrise du scots, reproduite grâce au contact avec ses parents scottisants. Le message 29 récompense cette culture et enjoint à continuer à partager du vocabulaire : « the mair wirds I learn the better. ». De la même façon, ne pas connaître ou ne pas comprendre est négatif : « the only wan I got wis aboot yer sister […] maybe a need tae go bile ma heid. ». L'utilisateur s'admoneste tout seul avec l'expression « Go boil one's head » est une vieille expression écossaise, qui signifie en fait « aller se faire voir ». Tout en ne comprenant pas tout, il maîtrise tout de même plusieurs expressions écossaises, qu'il partage à

son tour. Le ton général de l'échange est rendu cordial grâce à l'utilisation de formules phatiques qui ouvrent le message 29 et 30. L'échange de vocabulaire en scots est une activité qui rapproche, parce qu'elle permet à la fois de partager sa vie personnelle, et une culture commune. Tous les participants s'y enrichissent.

Dans les deux extraits du fil de discussion (tableau 5.4 et tableau 5.5), les utilisateurs font référence à la ville dont ils proviennent (« she wis frae embra » ; « in Lewis… »). Le placement géographique est primordial dans la justification non seulement de l'utilisation, mais aussi de la connaissance de mots en scots–et fait partie de l'iconicité des mots eux-même. Plus ils sont faciles à rattacher à une origine géographique, plus ils sont porteurs d'identité culturelle indexicale. Pour reprendre l'exemple de B. Johnstone qui travaille sur l'anglais de Pittsburg, « a person who sums up a situation in which someone corrected her grammar by saying 'That's a Pittsburgh thing' has clearly formed an indexical link between the non-standard form she used and her place of origin. » (Johnstone, 2011).

Dans le premier exemple (le tableau 5.4), le premier participant assimile le mot *marag* à Lewis, ce qui le détache potentiellement de l'Écosse continentale (puisque Lewis et Harris sont de grandes îles à l'ouest de l'Écosse). Le second intervenant le corrige, en grandissant le cadre du registre non pas à l'île, mais à tous les habitants gaélophones, sinon habitant dans des régions gaélophones.

5.2 Variation allographique de la représentation des consonnes

Les consonnes sont plus ou moins marquées, suivant différents traits phoniques. La marque phonétique est organisée autour d'une hiérarchie de sonorités (Zwicky, 1972). Les sons plus sonores vont avoir tendance à être conservés, autant à l'écrit qu'à l'oral, alors que les sons plus bas dans l'échelle (et les séquences graphiques qui les représentent) vont avoir tendance à chuter. Le schéma de la marque phonétique standard (le tableau 5.6) recoupe assez correctement celui des variations trouvées dans le corpus. Les consonnes les plus souvent écartées sont les moins sonores. Dans le corpus, les variations consonantiques suivent relativement bien les hiérarchies décrites par Zwicky. Les phénomènes de prosodie interviennent dans une certaine mesure dans les graphies adoptées par certains utilisateurs, et

Hiérarchie de la sonorité phonémique	
Voyelles	[-CONS] > [+CONS]
Glissées	[-OBSTR] > [+OBSTR]
/r/	[+VOC] > [-VOC]
/l/	[-ANT] > [+ANT]
/n/	
/m/	
/ŋ/	
Fricatives	
Occlusives	[+ CONT] > [-CONT]

TAB. 5.6 – Hiérarchie des sons (ZWICKY, 1972)

montrent l'amuïssement total d'une syllabe dans un mot (« Embra » pour *Edinburgh* ; ou « 'cos », voire « cuz » pour *because*, dans les deux corpus [3]).

Je vais sélectionner dans les deux corpus les mots dont les consonnes ne sont pas écrites (disparition du ou des graphèmes), ou ceux dont les consonnes subissent un changement de placement dans le mot, en m'appuyant sur le concept d'archigraphèmes, tel qu'il est exposé dans l'ouvrage de N. Catach. Prenons ces graphèmes comme « forme maximale stable » (KLINKENBERG, 1994) pour représenter un phonème, quelle que soit sa position dans le mot. L'équipe du CNRS HESO (Histoire Et Structure de l'Orthographe) a délimité une liste d'archiphonèmes et d'archigraphèmes, soit une liste qui regroupe des éléments ayant les mêmes traits, pour identifier les types de fautes d'orthographe, entre autres. Ainsi, l'archigraphème <O> représente les graphies <o>, <au> et <eau> en français ; et l'archiphonème /O/ représente /ɔ, o/, qui ne sont pas allophones : leur distribution est différente dans la plupart des variétés du français, et le /o/ fermé ne sera réalisé ainsi qu'en fin de mot, par exemple.

Les critères utilisés pour déterminer quelles lettres représentent les archigraphèmes standards de l'anglais sont :

— La fréquence d'utilisation de l'archigraphème pour représenter un son donné. Je m'appuie sur des études d'acquisition de l'écrit, qui donnent des inventaires de ces fréquences. Les archigraphèmes peuvent évidemment être des compositions de deux graphèmes.

— Le degré de stabilité : la séquence <ck> n'apparaît qu'en coda de mot ou de syllabe ; <cc> uniquement entre deux voyelles. Le représentant le plus stable de /k/ est donc <c>. <K> apparaît

3. Et dans l'anglais vernaculaire également.

peu de manière isolée, et bien qu'il représente toujours le son /k/, on rencontre plus souvent <c> en anglais.

— Le degré de pertinence phonologique : <ph> est un digraphe qui représente /f/, mais sa graphie grecque n'est ni la plus fréquente, ni la plus simple pour représenter ce son.

— La « rentabilité linguistique » : les terminaisons ou morphèmes de l'anglais me semblent être des éléments moins probants de sélection, mais par exemple le digraphe <ng> représente /ŋ/ dans les terminaisons, notamment celles des participes présents. Ce digraphe est donc conservé dans mon inventaire d'archigraphèmes.

— La créativité linguistique–c'est peut-être le concept qui sera le moins invoqué, puisque le corpus n'est qu'une présentation de créativité linguistique. L'exemple de « parking » de N. Catach, qui parle de l'introduction de mots dont la terminaison est <-ing> en français ne tient évidemment pas en anglais ; mais je ferai un tableau d'équivalences entre le leetspeak utilisé dans le corpus et les archigraphèmes correspondants.

Ces recherches graphophonématiques utilisent des travaux croisés (Blevins, 1998 ; Deschamps et al., 2004 ; Carr, 1999) pour tirer un résumé des archigraphèmes, ou représentation optimale d'un phonème dans une langue donnée. Dans les graphies alternatives à visée phonologique, les lettres représentent un son donné. Ces équivalences graphophonématiques servent à déchiffrer la phonologie contenue dans les graphies alternatives. Certains critères pris en compte par N. Catach, comme le doublement des graphèmes (et la différence entre <-il> et <-ille> en français) ne s'appliquent pas à l'anglais.

Ce tableau complète, en outre, celui situé dans le chapitre précédent, qui traite de toutes les équivalences graphophonématiques de l'anglais [4]. Dans le tableau présent, les phonèmes seront répartis selon leur mode articulatoire, comme dans les explorations suivantes. Comme N. Catach, j'utilise les majuscules comme notation des archigraphèmes.

<CK> est candidat graphique retenu pour le son /k/ parce que, malgré sa faible correspondance (6% du son /k/), la combinaison des trois graphies <c> + <ck> + <k> représentent 92% de la distribution de /k/ (Blevins, 1998). L'archigraphème représente donc les trois combinaisons graphiques.

4. Le tableau des équivalences des consonnes dans 4.2.1 est le tableau 4.1.

Phonème	Archigraphème	Fréquence
	Plosives	
/p/	P	96%
/b/	B	97%
/t/	T	97%
/d/	D	98%
/k/	CK	92%
/g/	G	88%
	Fricatives	
/f/	F	78%
/v/	V	99.50%
/θ/ et /ð/	TH	100%
/s/	S	73%
/z/	Z	23%
/ʃ/	SH	26%
/ʒ/	J	22%
/h/	H	98%
/x/	CH	100%
	Affriquées	
/ʧ/	TCH	97%
/ʤ/	DGE	93%
	Nasales	
/m/	M	94%
/n/	N	97%
/ŋ/	NG	59%
	Approximantes	
/l/	L	91%
/r/	R	97%
	Glissées	
w	W	92%
j	Y	44%

Tab. 5.7 – Archigraphèmes des consonnes de l'anglais

<V> est le seul candidat pour /v/, mais les /v/ en fin de mot sont toujours écrits <-ve>. Les omissions du <e> graphique seront donc traitées à part, le cas échéant. Contrevenir à la graphie sans changer la prononciation théorique des mots peut tout de même être un signe sémiologique intéressant à étudier. La double attribution de <TH> à /θ/ et /ð/ n'est pas problématique, parce que leur distribution se répartit respectivement entre les mots lexicaux et les mots non-lexicaux (Carr, 1999). Comme N. Catach dans sa composition de la liste des archigraphèmes du français concernant le voisement de <s> en position intervocalique, je suis confrontée au problème de /z/ : /z/ est bien plus fréquemment écrit <-s> à la fin d'un mot (64% de la fréquence de /z/). Le graphème <z> ne représente que 23% de la fréquence totale mais il n'a que la valeur phonétique /z/, et la forte fréquence de <s> s'explique par sa

fréquence morphologique. J'ai donc retenu <Z> pour représenter ce son, parce que ce graphème ne présente aucune ambiguïté phonologique. Le son /ʃ/ pose également problème : il est bien plus souvent représenté par la séquence graphique <ti>, dans le suffixe <-tion>. Cet élément morphologique dépasse de loin son concurrent, <sh>. En revanche, <sh> s'utilise en début ou en fin de syllabe (« dish, shard ») et a donc été sélectionné, malgré sa faible fréquence (26%), tout comme <J> pour /ʒ/, qui est moins fréquent que d'autres concurrents, comme le suffixe <-si_>, mais qui est intuitivement plus juste pour représenter ce son. J'ajoute aux fricatives le son /x/ du scots, et propose l'archigraphème <CH> pour le représenter.

Les affriquées présentent quelques problèmes, à cause de leur articulation particulière. Pour les représenter, j'ai choisi d'opter pour des combinaisons graphiques : <TCH> est un montage entre les trois graphies <CH> (fréquence : 55%), <TCH> (fréquence : 11%) et <T> en attaque d'affriquée (fréquence : 31%, dans des mots tels que « denture, feature »). La combinaison <DGE> ne présente pas d'ambiguïté de prononciation, et représente les combinaisons <DGE> (qui ne représente que 5% de la fréquence) et <GE> représente, avec <GI/Y>, 88% de la fréquence totale. En outre, leur trigraphie et l'évocation graphique de la première consonne qui les compose (/d/ et /t/) justifient ces choix d'archigraphèmes.

Je pourrais rajouter les phonèmes /hw/ et /kw/ à l'inventaire, et leur attribuer les valeurs <WH> et <QU> respectivement, mais ces deux combinaisons ne sont pas reconnues comme étant des sons de l'anglais à part entière. Ils ne sont pas utilisés dans cette étude, et la valeur de /hw/ revient à une combinaison des archigraphèmes H et W. Le son /kw/ n'est pas convoqué dans cette étude. En outre, les archigraphèmes ne sont pas un inventaire des formes allographiques, et la différence entre /l/ sombre et clair n'est pas pertinente pour le repérage des archigraphèmes.

Dans les sous-sections à venir, il s'agira de faire la part entre les 26 archigraphèmes consonantiques de l'anglais (plus <CH>) que je viens de dégager, c'est-à-dire un résumé des formes graphophonématiques les plus stables, et les pratiques réellement en cours sur les deux communautés. Dans chaque sous-section correspondant à un mode articulatoire, je ferai un inventaire des formes les plus fréquemment trouvées lors des récupérations des forums, ou des formes les plus intéressantes pour l'AAV et

pour le scots. Si possible, j'exploiterai les données textométriques pour chaque cas qui me semblerait intéressant, et tirerai des conclusions aussi générales que possibles sur les changements graphiques, en m'appuyant sur les modes articulatoires des consonnes de l'anglais standard.

5.2.1 Traitement des occlusives

5.2.1.1 Variantes graphiques des plosives sur SC

Les changements subis par les occlusives plosives par les utilisateurs de SC sont présentés dans le tableau 5.8. Quand les sons changent de mode articulatoire, ils sont présentés dans la sous-section concernant le nouveau mode d'articulation choisi, si cela est plus pertinent à l'étude.

Dans « thoughtless » → <thouless>, c'est tout le groupe consonantique dont le <t> est le seul gra-

Son/mode	Variantes	Exemple
/d/	<d#> → Ø	da, '*dad*'
/t/	<t> → Ø	thouless, '*thoughtless*'
[± voisé]	<t> → <d>	caw**d** '*caught*'
[± voisé]	<d> → <t>	hunner**t**, '*hundred*' ; bles**t**, '*blessed*'

Tab. 5.8 – Variation des graphies des occlusives en scots

phème ayant un équivalent phonologique qui disparaît dans la graphie alternative. Il est donc considéré comme un bloc cohérent, mais la prononciation que cette graphie alternative suggère, /ˈθɔːʔləs/, avec un éventuel remplacement du /t/ par une occlusive glottale, n'est pas représentée. Ceci me permet de parler des occlusives glottales, qui sont souvent représentés par une apostrophe en milieu de mot (comme c'est le cas, d'ailleurs, dans la graphie <Hawai'i>). Dans le corpus, ces apostrophes sont rares, en grande partie à cause de la difficulté de trouver ce type de mot en effectuant une simple recherche sur Google. Aucune occurrence de « Sco'land » n'a été trouvée (alors que Google renvoie 25 800 résultats). L'apostrophe pour remplacer un stop glottal n'est donc pas une marque communaulectale de SC.

Dans le couple graphique <hundred> > <hunnert>, la figuration de la consonne finale avec un <t>, qui indiquerait un dévoisement, apparaît avant « o' », '*of*', ou « years ». Ce phénomène n'est pas dû

à une coarticulation, mais à l'influence de formes anciennes en scots et surtout, à cette citation de la déclaration d'Arbroath : « Sae lang as bot ane hunnert o' us yet abide, ne'er sall we, upon ony conditioune, be subjecktit tae ye dominioune o' ye Inglis. » (Great Britain, Scottish Record Office & Scotland, 1320). Sur SC, la forme <hunnert> n'existe qu'avec un groupe prépositionnel en « of » qui effectue une opération d'extraction, et seulement un tiers des occurrences est un extrait de la déclaration d'Arbroath. Cela veut dire que les utilisateurs ont instinctivement distribué leurs utilisations d'<hundred> et d'<hunnert> entre un <hundred> + nom, qui sert d'adjectif cardinal de quantification simple, et un <hunnert> + groupe prépositionnel en *of*.

5.2.1.2 Variantes graphiques des plosives sur BP

Les occlusives en AAV présentent les mêmes types de transformations, avec des affinités particulières pour la démotion de <t>, l'un des sons consonantiques les moins marqués en anglais (Zwicky, 1972).

Les élaborations graphiques comme *tryna* sont courantes sur BP, et témoignent d'une utilisation du

Son/mode	Variantes	Exemple
/t/	<t> → Ø	tryna '*trying to*' ; ish '*shit*'
/t/	<t> → <'>	lil', '*little*'
/b/	<s> → <b>	Jeebus, '*Jesus*'
[± voisé]	<(c)k> → <gg>	fugg '*fuck*'

Tab. 5.9 – Variation des graphies des plosives en AAV

snowclone des quasi-modaux *wanna*, *gonna*…Certaines formes ont déjà été évoquées en 2.4.2, comme *fixna* pour remplacer « fixing to ». Ces formes dérivées des prononciations relâchées sont beaucoup plus courantes en AAV et sur BP que sur SC et témoignent de leur productivité. Ce mot appartient surtout aux discours les plus vernaculaires, comme l'extrait de profil personnel 5.1 le montre :

Les nombreuses graphies alternatives et le ton général de ce début de profil, destiné à être une petite annonce matrimoniale renforcent l'idée de l'utilisation de *tryna* comme marqueur du registre familier, utilisé entre pairs, d'où le refus de l'utilisatrice de rencontrer des hommes plus vieux (hors de

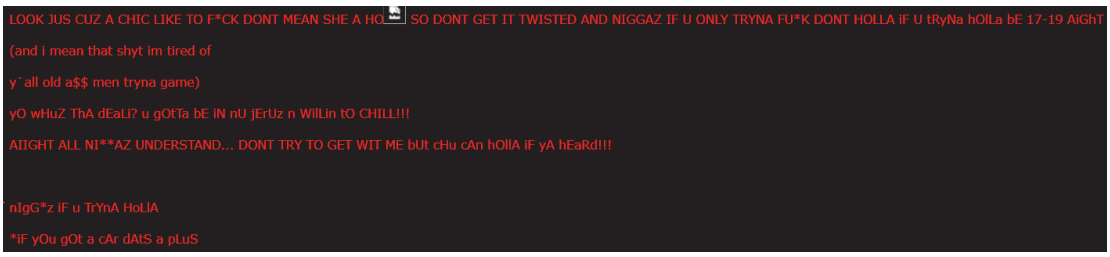


Illustration 5.1 – Extrait de profil personnel (BPPP17)

la fourchette 17-19 ans qu'elle a fixée). Malgré sa brièveté, l'extrait contient 3 occurrences de *tryna*, dans deux types de contextes différents (« tryna holla » et « tryna fu[c]k »), ce qui montre qu'il n'est pas confiné à un seul verbe et qu'il n'est pas utilisé dans une expression figée. On peut aussi prendre en compte l'influence de la musique en ce qui concerne *tryna*. Cette forme est utilisée dans la musique (et parsème le refrain d'une chanson populaire de 2006, *U and Dat*, qui culmina en treizième position au hit-parade américain Hot 100 (« My Ghetto report card », 2006). Les paroles de cette chanson sont copiées dans le profil d'un utilisateur (BPPP16), et la popularité de cette chanson peut expliquer à la fois la fréquence de la prononciation–si elle est fréquente en AAV, son exploitation musicale par un rappeur afro-américain n'est pas étonnante–et sa fréquence explique sa réinjection dans les écrits. Les paroles copiées dans BPPP16 contiennent cette forme graphique, d'où sa popularité. De plus, c'est le candidat graphique le plus vraisemblable : il contient cinq lettres (comme les modèles *wanna, gotta, gonna*) et finit par un <a>.

Certaines graphies témoignent de la co-incidence d'une métathèse et d'une inversion, à cause de la censure. La graphie *ish* montre l'influence de la musique et de la culture hip-hop sur les productions écrites des membres de BP. Cette graphie est à la fois une interversion des deux sons consonantiques de *shit*, et la chute de la consonne /t/. Elle est influencée par la censure des chansons de rap, qui peut être réalisée de deux manières :

— Un blanc remplace les mots incriminés : par exemple, dans « Drop it like it's hot », les références au gang de Snoop Dogg (les Crips) disparaissent de la version censurée : « I keep a [blue] flag hanging out my backside / But only on the left side, yeah that's the [Crip] side ». Les paroles de Pharell Williams contenant « fuck », « shit », « nigga » sont effacées de la piste vocale.

Les allusions sexuelles, même voilées, comme « that's whiter than what's [spilling] down your throat » disparaissent également (Snoop Dogg, Williams & Jelly Roll, 2004).

— Les paroles peuvent être inversées : « shit » devient ainsi un « ish » mécanique, comme on peut entendre dans la chanson « Just a lil' Bit » de Too Phat (« Classics : 'the collection of Phat tracks' », 2004).

Les chansons sont rarement censurées avec des bips, comme c'est le cas pour la télévision [5]. L'habitude d'inverser les paroles est la méthode de censure la plus agréable à l'écoute, parce qu'elle préserve en partie le rythme de la chanson, et attire moins l'attention de l'auditeur. Pour « shit », il est facile de le prononcer /ɪʃ/)malgré la censure, et certains rappeurs l'ont directement intégré à leurs chansons, pour contrecarrer la censure tout en restant très compréhensibles. C'est le cas de Soulja Boy, dans sa chanson « Eazy » (Soulja Boy & Gucci Mane, 2008). Cette forme du mot est naturellement entrée dans le vocabulaire populaire, et le mot « ish » a 144 entrées dans l'UD, la plus populaire étant celle décrivant ce phénomène de remplacement pour « shit ». Sa présence sur BP n'est pas étonnante, si l'on prend en compte l'influence de la musique populaire.

La chute de la fin du mot dans <lil'> est un phénomène connu de transcription de prononciation vernaculaire. Il est utilisé à la fois pour des pseudonymes (à la fois sur le forum, et pour les pseudonymes de rappeurs cités, comme Lil Wayne, ou Lil' Kim), et dans le corps de textes. C'est l'une des graphies alternatives les plus courantes sur BP (1 440 000 occurrences), mais elle l'est aussi hors plateforme, avec pas moins de 297 000 000 occurrences de « lil['] » en ligne. Ce n'est donc pas un indicateur communaulectal en soi.

Certaines graphies évoquent un potentiel changement de mode articulatoire, ou de voisement, par rapport au standard graphique. La graphie <Jeebus> pour *Jesus* pourrait presque suggérer une prononciation [β]. Malgré le changement du nombre de lettres dans le mot, il reste très reconnaissable, certainement parce qu'il n'y a qu'un changement au milieu du mot (<s> → <eb>). Il n'est pas indexical de la communauté des locuteurs de BP, puisqu'il s'agit d'une référence à un épisode des *Simpsons* (E15S11, première diffusion le 20 février 2000), où Homer prononce *Jesus* /ˈʤiːbʌs/ [6]. Il y a plus d'un

5. La chanson « Hot Dog » de Limp Bizkit contient plus de 50 occurrences de « fuck ». Des versions censurées à coups de bips existent, comme <http ://bit.ly/1zhlv11> et sont évidemment très désagréables à écouter.

6. Extrait consultable à cette adresse : <http ://vimeo.com/4787742>. Dernier accès : 05/07/2014.

millier d'occurrences sur BP, ce qui pointe vers une connaissance partagée (le poids culturel d'une telle série n'est pas négligable), éventuellement imitée. En plus de créer de la connivence entre les connaisseurs de la série qui lisent le mot, le changement consonantique évite le blasphème, puisqu'il ne s'agit plus exactement de Jésus. Il est possible de voir ceci comme une censure morale inscrite dans une axiologie chrétienne, où citer le nom de Dieu en vain est interdit [7] (Exode, 20 :7, Bible de Jérusalem).

Quant à « fugg », il sera traité plus en avant dans la partie sur les variations allographiques de fuck, dans la section 5.5.3. Il s'agit du seul jeu sur le voisement consonantique dans le corpus BP, avec la réélaboration de <ratchett> en <ratched>.

5.2.1.3 Récapitulation des variantes graphiques des plosives dans le corpus

La récapitulation de toutes les stratégies répertoriées pour traiter les occlusives sur SC et BP est présentée dans le tableau suivant (5.10) :

Dans les deux corpus, les consonnes les moins marquées (/t, d/) subissent le plus grand nombre

	SC	BP
<d#> → Ø	×	×
<t> → Ø	×	×
<s> → <b>		×
/t,d/ [± voisé]	×	
/k,g/ [± voisé]		×

Tab. 5.10 – Récapitulation des changements graphiques impliquant des occlusives dans les deux corpus

de changements, notamment des chutes, surtout en position coda (<fugg> pour fuck, <blest> pour blessed) . Les deux seuls couples de consonnes exploités ici sont /t,d/ et /k, g/, que les locuteurs reconnaissent bien comme étant phonologiquement proches, ce qui explique la facilité avec laquelle ils sont intervertis. Cette interversion n'existe pas avec /p, b/ par exemple, et c'est peut-être parce que ces lettres sont moins fréquentes que les autres plosives. Selon D. Crystal, qui a calculé la fréquence de chaque lettre dans la Cambridge Encyclopedia of the English Language (7.5 millions de lettres en tout), <p> et <b> sont assez bas dans le classement par fréquences cumulées : <p> est à la

7. Voir « Tu ne prononceras pas le nom de Yahvé ton Dieu à faux, car Yahvé ne laisse pas impuni celui qui prononce son nom à faux. »

seizième position, et <b> à la dix-huitième, alors que <t> et <d> sont plus fréquents en anglais (troisième et onzième position, respectivement) (Crystal, 1995). L'estimation pour /k/ et /g/ est plus difficile, puisque /k/ a plusieurs candidats graphiques pour le représenter.

Je vais maintenant examiner les graphies des fricatives, et leurs changements graphiques, qui annonceraient des différences phonographématiques.

5.2.2 Traitement des fricatives

5.2.2.1 Variantes graphiques des fricatives sur SC

Les altérations des fricatives dans SC sont nombreuses, d'autant plus que le son /x/ est pris en compte dans cette partie. Il s'agit principalement de scottismes qui influencent les choix graphiques des utilisateurs :

Son	Variantes	Exemple
/ð/	<th> → <h>	hings, '*things*'
/s/	<ss> → <sh>	pish, '*piss*'
/x/	<ght> → <cht>	micht, '*might*'
/v/	Ø → <v>	div, '*do/did*' ; oer, '*over*'

Tab. 5.11 – Variation des graphies des fricatives en scots

La graphie <hings> pourrait passer pour une simple faute d'orthographe, et une omission de lettre. Sa faible fréquence (16 occurrences dans le corpus) le place tout naturellement dans cette catégorie, mais il s'agit d'une prononciation attestée dans l'OSD, /hŋ/, /hɪn/, sans précision géographique. La débuccalisation de ce son semble courante, surtout vers Glasgow et dans le centre de l'Écosse (Burov, 2012 ; Lawson, 2014), mais elle est troublante à cause de sa proximité graphique avec la forme graphique qui correspond à la prononciation standard. Le corpus ne contient pas d'autre mot qui subisse une débuccalisation de /ð/, mais le dictionnaire précise que, dans les Îles Shetland, le son est délabialisé dans des mots lexicaux. D'ailleurs, dans cette région, le son /θ/ subit le même phénomène :

thought est prononcé /toxt/ et *thunder*, /ˈtʌnər/, sans distinction [8]. Ces deux mots viennent du corpus de graphies alternatives de SC, mais le digraphe <th> n'a pas été affecté par quelque changement.

La dégradation de /v/ et la chute de sa représentation graphique dans <oer> est un indice qui pointe vers les prononciations attestées dans l'OSD, [ˈʌu(ə)r, ɔu :r, o :r, u :r], mais dont la seule graphie attestée est <ower>. Le DSL accepte les trois graphies <over>, <ower> et <oer>. Le parcours graphique suggère (sans confirmer) un amuïssement progressif de la consonne, devenue semi-consonne : /v/ (prononciation anglaise) → /w/ (/ˈʌu(ə)r/) → Ø (/o :r, u :r/).

La forme de *do* est normalement une forme interrogative utilisée dans le nord de l'Écosse [9], selon l'OSD. Dans le corpus, il est plus volontiers utilisé pour remplacer un *do* d'emphase. On le trouve notamment dans un message qui retranscrit une blague écossaise : « 'At's true, I div hae a gairden. […] 'Aye, I dee hiv a hoose.' […] 'I div hae a wife !!' » '*Yes, it's true, I do have a garden. […] Yes, I do have a house. […] I do have a wife !*' (SC2N59). Cet exemple vient cependant d'une blague disponible sous la même forme sur internet, et ne représente pas toutes les utilisations sur SC, comme « Div ye look like the mannie on the porridge box ? » (SC24C17) ou « Aye, I div » (SC22E1), où cette forme est effectivement utilisée pour une interrogative, et comme tag de réponse.

La graphie <pish> existe aussi dans *Trainspotting* (Livre) (TR-L), et elle est attestée dans le DSL, grâce à des exemples remontant au 18ème siècle mais seulement dans le sens « uriner », alors que dans TR-L et *Trainspotting* (Film) (TR-F), c'est aussi un mot utilisé pour parler d'alcool : « ah'm no even pished yet » (II, 2). Sur SC, à part certaines utilisations dans des blagues, il est surtout utilisé comme synonyme de « shit » : « Well that's the kind of pish I was taught anyway » (SC2X21). L'utilisation d'un mot désignant des réalités physiologiques considérées comme taboues (les excréments, en l'occurrence) remplace une autre, et l'on peut y voir un parallèle étonnant entre cette utilisation et sa paire minimale afro-américaine, <ish>. Je n'irai pas jusqu'à lier les deux ensemble, mais la préférence pour un mot court, qui finit par un /ʃ/, rappelle les interjections de dégoût ou de mépris comme « sheesh » ou « pfft ».

La séquence <cht> est courante dans le corpus, pour représenter /xt/ en position finale (<licht>,

8. Je n'ai pas trouvé de variante graphique affectant /θ/ dans le corpus SC.
9. Et au nord de l'Angleterre.

<thocht>, <micht>), ce qui confirme le statut de <CH> pour représenter le son /x/. La prononciation en /x/ est confirmée par l'OSD. Cette graphie est un compromis avec l'anglais, en ne substituant qu'une lettre au groupe consonantique original (<ght> → <cht>). La longueur du mot est ainsi préservée, mais les utilisateurs de SC savent parfaitement de quel mot il s'agit, tout en utilisant une graphie courante dans un autre mot en scots, *loch*. Ainsi, même sans connaître la graphie du scots, l'extrapolation est facile, puisque les utilisateurs ont un mot-patron à disposition.

5.2.2.2 Variantes graphiques des fricatives sur BP

Dans le corpus BP, les fricatives présentent de nombreuses variations de retranscription, qui sont présentées dans le tableau 5.12.

La forme <outta> est une concaténation de deux prépositions, ce qui dépasse le patron verbe +

Son/mode	Variantes	Exemple
/f/	<f> → Ø	outta, '*out of*'
/v/	<v> → Ø	shoulda, '*should have*' ; iolence, '*violence*'
[± voisé]	/s/ → /z/	azz, '*ass*' ; coolezt '*coolest*'
/ð/	<th> → <d>	dem, '*them*'
/ð/	<th> → <t>	wit, '*with*'
/ð/	<th> → <v>	bruvas, '*brothers*'
/ð/	<th> → <f>	wif, '*with*'
ts	<nce> → <nts>	differents, '*difference*'

Tab. 5.12 – Variation des graphies des fricatives en AAV

<to> des quasi-modaux. De la même façon, <shoulda> est un assemblage fait d'un verbe de modalité, « should », et du verbe « have ». Encore une fois, les graphies obtenues ont cinq lettres, ressemblent aux patrons (avec le <-a> final), et surtout, provoquent la chute des sons les moins marqués ou les moins importants, c'est-à-dire ici les consonnes finales. Ces concaténations sont populaires dans les écrits au registre relâché sur internet.

Le mot <iolence> a peu d'occurrences sur le forum, mais il est surtout utilisé en ligne (sur BP et sur internet en général) comme graphie stylisée du très reconnaissable « violence ». Il est souvent utilisé dans des jeux linguistiques, où « love » et sa variante graphique « luve » sont utilisées comme

acronymes. On trouve ce phénomène dans « “LOVE” L ET O UR V IOLENCE E ND » (BPPP18) ; « E-nternal [sic] - through L-adies U-nder V-iolence E-ven HATARS !!!! » [E.-L.U.V.E.] (BPGB2). Ce n’est donc pas un mot communaulectal à proprement parler, mais il s’inscrit dans des jeux sur les mots ritualisés sur BP : cette volonté acronymiste se place dans le même spectre que la volonté de cryptographie (utilisée notamment par les gangs), et la didactisation de certains profils, qui proposent des acronymes et leur signification, surtout dans le discours des gangs, avec des acronymes tels qu’A.D.I.D.A.S, « All Day I Diss All Slobs » (BPPP19). Le mot « violence » est très intégré au discours, et bien que le concept soit rejeté dans les acronymes L.O.V.E. et E.-L.U.V.E., il est tout de même évoqué. Le positif est toujours joint au négatif, et violence et « haters » font souvent partie des éléments phatiques des profils personnels, par exemple.

Certaines graphies suggèrent des voisements par rapport au standard orthographique, et cela est plus fréquent sur BP. La graphie <azz> est représentative d’un mouvement qui associe <s> et <z>. La proximité phonologique est évidente pour les utilisateurs, mais surtout, cette graphie ressemble beaucoup à la graphie standard. Cette substitution n’apparaît pas que dans les mots censurés sur BP, comme l’atteste la graphie <coolezt>. Cette substitution semble être née d’une volonté d’échapper à la censure, et la faible fréquence de la lettre <z> alliée à une différence phonologique et visuelle faible avec <s> la rend idéale pour ces remplacements. Ils sont surtout effectués par des utilisateurs ayant recours à des formes écrites argotiques, comme ces trois extraits le montrent. Ils sont légèrement contextualisés :

> how iz u doin..juzs had to come by n give u some love..your de coolezt trini..LOL.hahaha.yaezs datz no lie now..lol.okey anywayz hunz tek it ez dont strezz ur poor lil trini brain out okey..maddd love from your cake mzs rummy o3 (BPGB3)
> “How are you doing.. Just had to come by and give you some love…You’re the coolest Trini…LOL. Hahaha. Yes that’s no lie now…LOL. OK anyways honey take it easy don’t stress your poor little Trini brain out, OK ? Mad love from your cake Miss [?] Rummy [smiley satisfait]”

> I had to make sure I signed your G-Book, you one of tha coolezt n cutezt girl I‘ve met. Muaaaaaaaah [10] !!! XXXtremely Happy to have YOU in MY LIFE. I‘m out for now, see you around. Muaaah !!!! (BPGB4)

> alyssa\my sis stay coo and gutta u kno how we get down stay lovin katrell thomas person fa lyfe jackie :my g 4 sho she da coolezt person u can hang wit (BPPP20)

10. Onomatopée imitant le bruit d’un baiser.

'Alyssa : my sister stay cool and ghetto you know how we get down stay loving Katrell Thomas Person [11] *for life. Jackie : my g[angster] for sure she is the coolest person you can hang with'*

La substitution de <s> par <z> semble être un bon indicateur du degré de changements graphiques dans les messages. Chacun des trois extraits contient au moins un autre graphie alternative, même si le second extrait est un peu moins fourni que les deux autres. Le premier extrait montre un remplacement quasiment systématique des <s>, et une volonté d'écrire en transcrivant la prononciation : ainsi, l'utilisatrice emploie des antériorisations de /ð/, « datz », « de » et $[be]_V$ à la troisième personne du singulier alors qu'elle devrait utiliser la seconde personne. Ce sont des caractéristiques d'un discours afro-américanisé (d'autres différences syntaxiques sont décrites dans le tableau 2.9). Dans le troisième extrait, la graphie <gutta> pour « ghetto » est très vernaculaire. Il n'a que 14 entrées dans l'UD, mais elles évoquent toutes le ghetto, le « hood » ou être un gangster. Ces caractéristiques discursives placent donc ces utilisations dans le registre familier urbain afro-américain.

Dans le même esprit, les phénomènes de TH-fronting affectant /θ,ð/, dont la graphie <th> est alors retranscrite <t, d> ne sont pas forcément très révélateurs quant aux pratiques discursives de BP, parce qu'il s'agit d'une caractéristique phonologique connue de l'AAV. J'ai déjà commenté les antériorisations et labialisations de l'AAV dans la section 2.4.2, et elles sont associées à ce que certains internautes (hors BP) appellent "typobonics", soit l'utilisation de certaines graphies alternatives [12]. Ces traits sont tournés en dérision par ceux qui ne les utilisent pas, et ils sont donc connus et reconnaissables comme n'étant pas neutres (non stigmatisées), mais ils restent tout de même utilisés par les membres de BP. La graphie <th> peut également être remplacée par <f> ou <v> pour le son /ð/, en position coda de syllabe pour <bruvas>, ou <wiv>, auquel la graphie <wif> est préférée (1 600 occurrences de <wiv> contre 19 000 <wif>). De la même façon, la distribution de <brufa>, où le son /ð/ est intervocalique est très différente. Il n'y a que 8 occurrences de <brufas>.

En prenant le corpus BP, j'ai créé le tableau 5.13, qui expose les phénomènes de TH-fronting (antériorisation, plosivation). Pour les mots grammaticaux, où <th> est souvent en première position dans

11. Une recherche sur Katrell Thomas Person renvoie à plusieurs affaires judiciaires dans le Wisconsin où un homme portant ce nom est impliqué pour des délits mineurs. Vraisemblablement, il s'agit du petit ami de la sœur de la scripteuse.

12. « The intentional misspelling or shortening of common words in order to sound cool in some circles. It is similar to ebonics in spoken english », pour reprendre une définition donnée sur l'UD.

le mot, la graphie <d> est privilégiée, et n'a pas de concurrents (pour des graphies telles que <da>, <dat>, <dem>, <dey>, <dis>, <dat>…). Le mode d'articulation change, et l'attaque est plosive. C'est un faisceau de graphies alternatives très connu, et utilisé par les Afro-Américains. Le forum dédié aux gangs afro-américains *Da Blocc* [13] utilisait cette graphie alternative ; de nombreuses chansons populaires comme « In Da Club » de 50 Cent (« Get rich or die tryin' », 2003), rappeur déjà évoqué dans la partie 2.6.3.1.1, utilisent cette graphie [14]. Elle est donc intégrée au sociolecte, et naturellement utilisée par les membres de BP. En revanche, elle n'est pas utilisée pour les deux autres positions (intervocalique et en coda de syllabe) dans le corpus. Pour ces deux positions, les graphies /v/ et /f/ sont préférées. La graphie <v> est généralement intervocalique, et une consonne voisée est phonotactiquement plus facile à prononcer entre deux voyelles. En revanche, en position coda, une lettre correspondant à une consonne dévoisée comme <t> ou <f> est préférée parce que la fin du mot correspond à un dévoisement général. La graphie <wit> est bien plus utilisée que <wif> (à peu près dix fois plus d'occurrences dans le corpus), et <wiv> est marginal. Le choix des utilisateurs est donc relativement logique, ce qui prouve leur conscience phonologique et phonographématique, couplée à l'influence de la culture populaire.

<#th_>	
Graphie préférée	Graphie secondaire
<d>	–
<_th_>	
Graphie préférée	Graphie secondaire
<v>	<f>
<_th#>	
Graphie préférée	Graphie secondaire
<t>	<f, v>

Tab. 5.13 – Distribution des changements graphiques de <th>

La graphie <differents> est un cas à part, parce qu'elle présente une réélaboration graphique erronée. L'adjectif « different » est beaucoup plus courant que le nom : dans le COCA, il y a 186 692

13. Il a malheureusement disparu, après avoir été attaqué par le site *Something Awful*.
14. Rihanna et son single « Pon de Replay » ou « Pon de Floor » de Major Lazer viennent aussi à l'esprit, mais les deux interprètes sont respectivement barbadienne et jamaïcain (pour Vybz Kartel, le chanteur de « Pon de Floor »).

résultats pour ce mot, contre seulement 52 613 pour le nom afférent, « difference ». C'est en partant de la base de l'adjectif que les utilisateurs arrivent à une transcription presque phonétique du mot (/ˈdɪfrəns/ → /ˈdɪfrənts/ en ajoutant un simple <s>, qui déclenche la prononciation du <t> final de la graphie source. Si l'intuition phonologique est correcte, elle créé une graphie alternative logique, qui n'est pas imputée à un seul utilisateur, vue sa fréquence, mais à une même façon de penser la morphologie. Ces adjectifs ne sont pas pluralisés, et n'ont pas de valeur plurielle, ils deviennent des noms par l'adjonction du morphème <s> qui transforme la prononciation du suffixe en /ts/.

La même réélaboration existe avec <violents>, que l'on pourrait croire être une transcription de « violence », mais elle est surtout utilisée dans une parodie de poésie. C'est en fait un jeu de mots sur « violets/violent » : « Roses are red / Violents are blue / Sorry Boo I'm bumping [15] you » (BPPP21) ou « […] VIOLENTS ARE CORNY / WHEN I THINK OF YOU / OOH BABY I GET SO HORNY » (BPPP22). Le seul changement est « violents », et cette lettre supplémentaire signale le changement du schéma habituel avant qu'il n'ait eu lieu dans les deux derniers vers.

5.2.2.3 Récapitulation des variantes graphiques des fricatives dans le corpus

	SC	BP
<f> → Ø		×
<v> → Ø		×
Ø → <v>	×	
/s/ <ss> → /ʃ/ <sh>	×	
<ght> → <cht>	×	
<th> → <h>	×	
<#th> → <d>		×
<th#> → <t>		×
<_th_> → <v>		×
<_th_> → <f>		×
<-nce> → <-nts>		×

Tab. 5.14 – Récapitulation des changements graphiques impliquant des fricatives dans les deux corpus

Le tableau 5.14 synthétise les changements graphiques subis par les fricatives dans les deux corpus. Il n'y a pas de recoupements dans les façons dont les graphies changent autour des fricatives. La débuccalisation de /ð/ en scots n'a pas d'équivalent dans la plosivation ou dans l'antériorisation

15. Certainement une confusion avec « dumping ».

en AAV. Ces variations, qui ne se recoupent pas, sont une façon de voir les points importants de divergence entre les deux langues minoritaires, et leur évolution respective à partir d'un même son, /ð/. Si les Écossais passent systématiquement à une graphie écossaise, avec <cht> qui remplace <ght>, les membres de BP réinventent également un standard graphique logique par rapport à l'adjectif, avec la réutilisation de <-ents> pour symboliser /nts/.

5.2.3 Traitement des affriquées

5.2.3.1 Variantes graphiques des affriquées sur SC

Étonnamment, aucune affriquée ne semble être affectée par les graphies alternatives, ou écossaises. L'exemple de <radge>, mot pour lequel l'OSD donne la transcription [radʒ], utilise logiquement la graphie <dg> pour retranscrire le son ʤ. Ce mot ne semble exister que dans le nord de l'Angleterre (en Geordie) et en scots, et provient certainement d'un emprunt au français « rage », ou d'un mot gitan, « raj », qui aurait le même sens (source : DSL). Dans les deux cas, la consonne finale se serait donc affriquée avec son passage en scots, ce qui a provoqué le changement graphique indiqué dans le dictionnaire, et sur le forum.

Son	Variante	Exemple
/ʤ/	<dge>	radge, '*fou, sauvage*'

Tab. 5.15 – Variation des graphies des affriquées en scots

5.2.3.2 Variantes graphiques des affriquées sur BP

De la même façon, en AAV et surtout dans le corpus BP, il y a peu de changements graphiques affectant les affriquées (cf le tableau 5.16). Les deux seuls mots trouvés suivent deux chemins différents. La graphie <madjesty> est très marginale (seulement sept occurrences dans le forum), et la réélaboration graphique <dj> pour le son /ʤ/. Le corpus ne contient pas de /ʤ/ à l'initiale d'un mot.

Par exemple, <djesture> existe sur internet, à la fois comme jeu de mot sur *DJ* et *gesture*, et comme graphie de « gesture » « I will try to try and use the 2 finger djesture to enable it » [16]. Le mot *majesty* n'est pas très courant en anglais, il n'y a que 2 175 occurrences dans le COCA et il ne figure pas dans la liste des 5 000 mots les plus fréquents en anglais. Cette graphie n'est pas non plus un jeu de mots sur *mad* et *majesty*, parce qu'elle est surtout utilisée comme pseudonyme d'une femme qui se met en valeur dans sa page personnelle (BPPP32).

Le cas de la graphie <bishes> est plus intéressant, parce qu'il s'agit d'un choix graphique pour

Son	Variantes	Exemple
/ʧ/	<tch> → <sh>	bishes, '*bitches*'
/ʤ/	<j> → <dj>	madjesty, '*majesty*'

Tab. 5.16 – Variation des graphies des affriquées en AAV

changer le mot *bitches*. Bien que l'allographie de ce mot soit étudiée plus loin, dans la partie 5.5.2, c'est la seule graphie qui suggèrerait une simplification de l'affriquée. Le son /ʧ/ correspond à la graphie <tch>, mais la graphie <sh> lui est préférée dans 1 150 occurrences sur BP. La réduction des groupes consonantiques, surtout en fin de mot, est fréquente chez les Afro-Américains, mais l'amuïssement porte souvent sur le second son, ce qui n'est pas le cas ici. Il serait possible de penser que l'entité /ʧ/ est perçue comme/t+ʃ/, donc comme deux sons [17], et que le premier, moins marqué et donc moins important que la fricative, pourrait chuter. Pour vérifier cette hypothèse, j'ai recherché des chansons contenant le mot « bish ». La seule chanson sortie dans le commerce et contenant ce mot dans le titre est « Yass Bish », le nouveau single de Nicki Minaj où elle chante en duo avec Soulja Boy (Minaj & Soulja Boy, 2014). Dans cette chanson, l'accroche chantée par Soulja Boy est effectivement prononcée /jɑs bɪʃ/. Si la graphie n'avait qu'affecté le <t>, la prononciation aurait été plus floue : la séquence <-[i]ch> se prononce /k/ dans « Michael », par exemple, ou /tʃ/dans « lichee » ou « rich ». La prononciation /ʃ/ne semble apparaître que dans des mots d'origine étrangère, comme « niche » ou « lychee », et c'est pourquoi la graphie <sh>, moins ambiguë, est choisie pour signifier le changement consonantique. Quant à la graphie <bitsh>, elle est très marginale dans le corpus BP (seulement 140

16. Trouvé à l'adresse <http ://grokbase.com/t/gg/eyes-free/1376nhwwxt/crashing>, dernier accès le 12/05/2014.
17. Cela a déjà été prouvé par les phonéticiens (Roach, Hartman, Setter & Jones, 2006), mais je parle ici de la perception des locuteurs.

occurrences). La séquence graphique <tsh> est très rare en anglais, et ne correspond qu'à des composés ou des dérivés, comme « outshot », « saltshaker » ou « courtship », où le <t> appartient à la fin du premier mot.

5.2.3.3 Récapitulation des variantes graphiques des affriquées dans le corpus

L'inventaire général des changements concernant les affriquées est assez modeste, parce que ce type de son semble plus rare. Dans les corpus, aucun mot commençant par <ch->, la graphie la plus fréquente pour représenter /ʧ/à l'initiale, n'a été affecté par une graphie alternative. La fréquence des phonèmes semble donc influencer le nombre de variantes possibles. Moins un phonème est mobilisé, et moins il a de variantes à disposition, parce que les scripteurs ne pensent pas aussi souvent à leur trouver une variante graphique. Mis à part la graphie <dj> pour /ʤ/, qui montre l'équivalence mentale phonographématique des utilisateurs du forum BP, aucun phénomène graphique ou phonologique ne se croise dans les deux forums.

	SC	BP
/ʤ/ = <dg>	×	
<j> → <dj>		×
<tch> → <sh>		×

Tab. 5.17 – Récapitulation des changements graphiques impliquant des affriquées dans les deux corpus

5.2.4 Traitement des nasales

5.2.4.1 Variantes graphiques des nasales sur SC

Les nasales ne présentent pas de nombreuses formes différentes dans le corpus SC, mais on peut y noter certains changements légers. Les dévélarisations de ŋ de BP et SC seront traitées ensemble, entre les deux parties de cet inventaire de formes.

La fin du prénom *William* est changée dans le surnom *Wullie*. La dernière consonne est plus faible,

Son	Variantes	Exemple
m	<m> → Ø	Wullie, '*William*'
m	<m> → <w>	maw, '*mom*'
ŋ	<ng> → <n>	missin, '*missing*'

Tab. 5.18 – Variation des graphies des nasales en scots

à cause de sa prononciation nasale (par rapport à la consonne latérale précédente). La forme irlandaise *Liam* est une aphérèse de *William* alors que le scots retire la dernière syllabe et la remplace par le suffixe affectif <-ie>. Cette particularité écossaise a déjà été vue dans la partie 2.2.2.2, et elle est surtout adjointe à des mots existants, comme « scotties » (scot + ie), « auntie » (aunt + ie), « lassie » (lass + ie)…Ni la fricatives /s/, ni la plosive /t/ ne sont affectées par cette adjonction, et elles sont pourtant moins marquées phonologiquement. Cette forme est donc certainement née d'une apocope du prénom William en Will, puis d'une adjonction de <-ie> ultérieure [18].

Le <m> (et le son /m/) a également tendance à s'éroder et à devenir une consonne glissée en scots dans le corpus. C'est un phénomène qui touche également /l/ (et que je traiterai dans la partie 5.2.6), et que j'ai déjà abordé avec <ower>, la graphie alternative d'<over>. Le phonème /m/ est fragile en scots : dans la forme écossaise de la préposition « from », *fae* (parfois graphiée <fy> ou <frae>) il disparaît complètement. En finale de mot, les consonnes ont tendance à disparaître, et la graphie <maw> pour *mom* (voire la forme écossaise *mam*) représente une première étape vers une vocalisation de la fin du mot, grâce à l'utilisation d'une spirante après l'amuïssement de /m/. Ce <w> graphique a deux avantages : conserver la longueur du mot, pour éviter la confusion avec la graphie <ma>, par exemple, pour *my* (observée dans TR-L et dans SC) ; et montrer un allongement compensatoire après la chute du phonème /m/.

5.2.4.2 Variantes graphiques des nasales sur BP

Son	Variantes	Exemple
/n/	<n> → <kn>	kneegrows, '*negros*'
ŋ	<ng> → <n>	puttin, '*putting*'

Tab. 5.19 – Variation des graphies des nasales en AAV

18. Je ne parlerai pas ici de la transformation de la voyelle <i> en <u>.

Les changements n'affectent pas le phonème /m/ dans le corpus BP, et /n/ n'est altéré qu'à l'initiale. La graphie <kneegrow> est un agencement graphique qui ne change pas la prononciation du mot. Les scripteurs ont rajouté des lettres, qui brouillent les repérages automatiques de la censure. « Nigger » étant déjà censuré, le mot « negro » pourrait appartenir à la liste de mots interdits. Cette graphie alternative dépasse le cadre de remplacement d'une lettre par une autre et prend l'apparence d'un composé, en mobilisant deux mots dont la prononciation combinée est homophone à « negro » (/ˈniːgrəʊ/. C'est un cas rare, voire inexistant ailleurs dans le corpus, parce qu'il implique à la fois des phonèmes sans beaucoup de représentations graphiques différentes (/n/ est principalement représenté par <n> ou <kn> en début de mot), et des éléments lexicaux combinables. La différence sémantique est faible, et cette graphie sert surtout à atténuer la valeur péjorative du mot. La graphie <kneegrow> est utilisée dans des pseudonymes, ou dans des adresses phatiques : « Wassup to all my kneegrows in here » (BPPP23), où il a un sens positif, autant que dans des contextes négatifs, comme « when i first watched jack city, i cried as nino brown (the evil field kneegrow) gunned down g money (the loyal house niglet)..... » [sic] (BP12E4199) ou dans les deux sens, pour certains : « I'm no longer wasting it on undeserving kneegrows ! […] My family on here, I luv yall crazy kneegrows.. » (BPPP24). Cette graphie ne change donc pas le sémantisme de *negro*, qui s'inscrit dans le même cadre sémantique que *nigga*. Cette graphie est préférée à <negrow> (seulement 36 occurrences, soit dix fois moins que <kneegrow[s]>). Cette graphie ne changerait que la fin du mot, et accolerait le mot « grow » à un morphème inexistant, <*ne>. La solution préférée est celle qui ressemble à un composé de deux mots. La graphie alternative relève potentiellement de la réanalyse morphologique.

La disparition de <g> à la fin de <ng> est tellement systématique que les données recueillies à l'œil ne sont pas fiables. Il y a certainement beaucoup plus de tokens sur les deux forums que dans le corpus recueilli. On peut extrapoler à partir des éléments trouvés dans le corpus la moyenne des remplacements, mais pas tous les mots sur lesquels ils proviennent. On les trouve également dans les corpus secondaires *Push* et TR-L. En outre, ces remplacements ne sont pas communaulectaux, puisqu'ils sont pratiqués dans le parler vernaculaire en général. Sur SC comme sur BP, plusieurs graphies alternatives maintiennent le <-ing> final, comme <naething>, <fechting>, '*fighting*'(SC) ; ou <cumming>, '*éja-*

culer', <splaining>, '*explaining*' et pas moins de quatre graphies alternatives de « fucking », sur dix tokens. La dévélarisation symbolisée à l'écrit est souvent la seule entorse à la graphie standard, dans le corpus, et, dans de rares cas, la chute du <g> graphique peut s'accompagner de la chute du <i>, comme dans <lookn>, « looking » (BP) ou <deen>, « doing » (SC), qui prouve la conscience de la valeur syllabique de /n/. Le corpus SC contient plus de <-in> que le corpus BP, rien que dans les scottismes proches de l'anglais, avec 34 occurrences, soit 7.8% contre 22 sur BP, ce qui y représente 7.3% des mots recueillis. Cette dévélarisation symbolisée est donc importante dans les deux corpus, et très fréquente dans le monde anglophone en général.

5.2.4.3 Récapitulation des variantes graphiques des nasales dans le corpus

Les traits graphophonémiques affectant les nasales sont résumés dans le tableau 5.20. Encore une fois, les traits ne se croisent pas, et les amuïssements consonantiques n'affectent ici que le scots. Au contraire, les réélaborations graphiques afro-américaines de <n> en <kn> montrent qu'il n'y a aucune différence dans la réalisation sonore des nasales. La dévélarisation de <-ing> est un trait partagé dans tout l'espace anglophone, attendue puisque le phonème /n/ est plus courant que /ŋ/, et plus simple à prononcer ou « plus naturel » (Neuman, 2009) en phonétique articulatoire.

	SC	BP
<m> → Ø	×	
<m> → <w>	×	
<ng> → <n>	×	×
<n> → <kn>		×

Tab. 5.20 – Récapitulation des changements graphiques impliquant des nasales dans les deux corpus

Le remplacement d'une consonne par une glissée, et par <w> en particulier est bien plus fréquent en scots qu'en AAV. Bien qu'il existe dans les deux langues minoritaires, ce remplacement ne touche pas les mêmes phonèmes, et donc les mêmes représentations graphiques. Pour le corpus SC, la plupart de ces changements seront examinés dans la sous-partie traitant des glissées (5.2.7. Dans le corpus BP, ce remplacement touche plus volontiers les <r>, étudiés dans la section suivante 5.2.5).

5.2.5 Rhoticité et /r/

La rhoticité des langues en présence est un problème épineux, parce que si la variété RP de l'anglais est non-rhotique, ce n'est pas celle que parlent les Écossais. Le GA est rhotique, mais les représentations graphiques de l'AAV dans le corpus tendent à le peindre comme étant de moins en moins rhotique. De plus, le phonème /r/ est affecté par un phénomène proche du rhotacisme : il est souvent remplacé en scots et en AAV par un son vocalique (parfois semi-consonantique), et l'orthoépie doit parfois corriger les déplacements de prononciations des « r faibles » [19], qui se dégradent en /w/ (Neuman, 2009). Tous ces phénomènes sont perceptibles dans le traitement de la rhoticité dans les deux corpus en présence.

5.2.5.1 Variantes graphiques des <r> sur SC

Son/mode	Transformation	Exemple
/r/	<r> → Ø	fae, '*from*'
/r/	<Vr> → <ur>	thurty, '*thirty*', hur, '*her*'

Tab. 5.21 – Variation des graphies des /r/ en scots

Dans la partie dédiée à la phonologie du scots, et plus particulièrement dans la partie dévolue aux consonnes (2.2.4.3), j'avais abordé les réalisations différentes [ɾ, ɻ, ɹ], mais le problème est l'absence de lettres qui pourraient représenter ces sons, tout en n'étant pas <r>. L'amuïssement de /r/ est rare, et n'intervient que dans « from », graphié <fae> ou <fy> dans le corpus. Cette préposition est la seule à avoir une telle forme, comparée à sa contrepartie anglaise, et elle est attestée dans des textes depuis le 17ème siècle selon le DSL. L'OSD lui donne la transcription phonétique /fe/, et la forme <frae> est tout de même plus courante. La chute de /r/ est facile à comprendre. C'est un phonème peu marqué [20], surtout lorsqu'il est accolé à une consonne fricative. De plus, cette séquence dans un mot grammatical et donc non-accentué a certainement provoqué l'amuïssement de /r/.

La tendance à regraphier [21] les séquences <_r> avec <ur>, comme le montrent les exemples <hur>

19. Traduction personnelle de « weak r ».
20. D'autant moins facile à prononcer s'il est réalisé comme une consonne battue ou comme une consonne rétroflexe approximante.
21. Soit adopter une graphie alternative observable dans les variations inter-scripteurs.

et <thurty> tentent de normaliser la position vocalique avant un /r/, qui se prononcent peut-être non plus /ɜː/ comme cela aurait été le cas en anglais RP, mais une prononciation plus proche de la classe d'isophones de NURSE en scots, /ɤ̈r ~ ër/. Malgré le changement vocalique, le /r/ est conservé, quelle que serait sa prononciation. Cela montre qu'il s'agit d'un son important en scots, qui reste bel et bien une variété rhotique, et où le /r/, même en frontière de mot, ne subit pas d'altération.

5.2.5.2 Variantes graphiques des <r> sur BP

SON	VARIANTES	EXEMPLE
/#r/	<r> → <i>	iight, '*alright*'
/#r/	<r> → Ø	aiight, '*alright*'
ɚ ~ɜˠ	<er> → <a>	gangsta, '*gangster*'
ɚ ~ɜˠ	<er> → <ah>	sistah, '*sister*'
/r/	<r> → <y>	scayed, '*scared*'
ɔːr	<or> → <aw>	Lawd, '*Lord*'

TAB. 5.22 – Variation des graphies des /r/ en AAV

Les graphies alternatives qui affectent les <r> graphiques sur BP sont présentées dans le tableau 5.22. L'AAV, en tant que variété rhotique, partage néanmoins des caractéristiques avec l'ABS, comme la réalisation vocalique ou la délétion de /r/ quand il est postvocalique ou lorsqu'il est avant une voyelle centrale ou non accentuée (comme déjà exposé dans le tableau 5.22 de la partie 2.4.4.2). Ces indications phonologiques sont utiles pour examiner le cas de <er> → <a(h)> dans <gangsta> et <sistah>. Ce changement graphique est plus que répandu, dans la culture urbaine hip-hop ou dans l'imitation des Afro-Américains (voir la partie 2.4.2, où les changements graphiques perçus comme étant afro-américains sont décrits). Non seulement le <r> final disparaît, mais en plus, il peut être remplacé par un <h>, qui suggère le passage de /ɚ/ à une vocalisation plus proche de « ah », soit /ɑ(ː)/. Cette vocalisation de fin de mot se retrouve aussi dans les chutes de <r> en début de mot, dans les deux graphies <aiight> et <iight> pour « alright ». Cette graphie émule une prononciation, apparemment répandue dans l'état de New York (UD), où les consonnes approximantes /l/ et /r/ laisseraient la place à la diphtongue aɪ et à une répétition du /a/. <Iight> se prononce donc /aˈaɪt/ [22] La perte consonantique a donc

22. On entend bien cette prononciation dans le refrain de la chanson « Imma be aight » de Lil Wayne (« Tha Carter III », 2008). L'audio de la chanson est disponible à l'adresse <http ://j.mp/theseLG>, dernier accès le 11/07/2013.

provoqué un allongement compensatoire de la diphtongue sur sa droite, et les répétitions graphiques de <ii> n'en sont qu'une transcription, basée sur la valeur libre de la lettre <i>.

Les vocalisations de /r/ sont parfois symbolisée par l'utilisation de <w> ou <y>, qui prouvent une étape peut-être intermédiaire de réalisation du phonème /r/. Dans une position intervocalique, comme dans <scayed>, le /r/ dégradé en <y> semble provoquer non pas la prononciation /skeərd/ ou, en /r/ faible, /skeəwd/ mais une diphtongue plus centrifuge, /ske(j)ɪd/. De la même façon, la graphie <Lawd> déroge à l'allongement de la voyelle avant le /r/, au contraire, elle suggère l'apparition d'une diphtongue centrifuge laʊd, avec assimilation de la classe d'isophone FORCE avec MOUTH. Ce n'est pas le seul exemple dans le forum ou le corpus : la graphie <yawd> pour « yard » y est attestée, et d'autres mots sont réécrits avec une séquence <aw>, surtout interconsonantique, comme <dawg> pour « dog » ou <dayum> pour « damn ». Ces graphies évoquent toutes une aperture et un allongement du mot, qui passe par une diphtongaison. Ces deux dernières graphies seront examinées dans la partie sur la variation graphique affectant les voyelles, mais elles montrent toutes l'intérêt pour les transformations impliquant des séquences graphiques plus rares en anglais, qui transcrivent les diphtongaisons afro-américaines (SIDNELL, 2002).

5.2.5.3 Récapitulation des variantes graphiques de <r> dans le corpus

Les deux langues minoritaires n'ont pas du tout la même attitude concernant la rhoticité (ce que l'on peut voir dans le tableau 5.23), mais leur tendance phonologique semble aller à l'inverse de ce qui est pratiqué dans l'anglais standard de leur pays. Ainsi, malgré la rhoticité du GA, l'AAV semble évoluer vers une chute des /r/ postvocaliques, qui ne servent qu'à créer des diphtongues centrifuges ou à rallonger des diphtongues préexistantes, comme dans <iight>, alors que le scots préserve les <r> graphiques.

Les voyelles autour de ces consonnes sont affectées par leur prononciation, c'est la raison pour laquelle le scots se tourne vers la graphie <ur> pour représenter /ɤ̈r ~ ër/ et la conservation du /r/ final. On pourrait même y voir la fusion des classes d'isophones LETTER et NURSE, étant donné que même un mot grammatical, non-accentué comme « her » porte la même graphie que la première

voyelle, accentuée de « thirty ». À l'inverse, l'AAV adopte le nouveau standard graphique <aw> pour représenter <Vr>, afin de montrer la dégradation de la consonne en glissée, voire en voyelle.

L'étude des latérales continue avec la section suivante 5.2.6, où /l/ et ses allophones seront observés.

	SC	BP
<Vr> → <ur>	×	
<r> → Ø	×	×
<or> → <aw>		×
<er> → <a[h]>		×
<r> → <y>		×

Tab. 5.23 – Récapitulation des changements graphiques impliquant des /r/ dans les deux corpus

Le mouvement de conservation de la latérale /r/ en scots s'accompagne d'un mouvement inverse de dégradation de /l/ et de ses allophones en glissée.

5.2.6 Traitement de /l/ et ses allophones

5.2.6.1 Variantes graphiques des <l> sur SC

Son	Variantes	Exemple
/l/	<l#> → <'>	sma', '*small*'
/l/	<l(l)#> → Ø	fitba, '*football*'
/l/	<l#> → <w>	aw, '*all*'
/l/	<V +l> → <ye>	ayeweys, '*always*'

Tab. 5.24 – Variation des graphies des latérales en scots

Les /l/ et ses allophones sont potentiellement plus intéressants en scots qu'en AAV, puisque la prononciation de /ɬ/ est plus vocalique. Le son /l/ est souvent vélarisé voire vocalisé, ce qui explique les changements graphiques présentés dans le tableau 5.24. Les /l, ɬ/ sont souvent amuïs : la graphie <sma'> préserve le souvenir de <ll> avec l'apostrophe, ce qui ne veut pas dire que le son /l/ est encore prononcé. La graphie <smaw> n'existe pas dans SC, ce qui suggère qu'il n'y a pas d'allongement compensatoire de la voyelle précédente, ou de vocalisation de /l/, ce qui est contredit par l'OSD, qui lui donne la prononciation /smɑː/. Si la voyelle est longue, et qu'elle respecte la loi d'Aitken des voyelles écossaises, qui ne sont longues qu'avant /r, v, z, ð/ ou une frontière de mot ou de morphème (Trudgill, 1984), cela suggère que le phonème /l/ est peut-être rayé de la prononciation. L'environnement

graphique standard <all> est prompt à éliminer le /l/, comme c'est aussi le cas dans <fitba> pour « football ». Cette forme graphique est attestée dans TR-L, bien que l'OSD ne cite que « fitbaw », sur le modèle « baw », ce qui le rapproche de la graphie <aw> pour « all ». Tous ces mots qui portent la graphie standard <all> sont prononcées avec /ɑː/, /aː/, /ɔː/. Le <ll> n'est remplacé par une lettre que dans le cas de <aw>, qui est aussi graphié <aa> ou <au> dans le corpus SC. Ce maintient d'un mot de deux lettres est certainement dû à la peur d'une confusion avec l'article indéfini « a » ; mais il sert aussi à montrer l'allongement de la voyelle (qui provoque la transformation de /l/).

La graphie <ayeweys> montre un changement total de la prononciation de /l/ en glissée. On peut imaginer une prononciation qui incluerait le mot « aye » à « always », comme /ˈaɪwəz/. L'OSD propose la transcription douteuse /ˈəɪˌwəz/ [23], qui est bel et bien une concaténation du mot écossais « aye », '*toujours*' et de l'« always » anglais. Cette graphie est la seule où le <al> potentiel anglais deviendrait une diphtongue, allongeant encore une fois le <a> préconsonantique pour compenser la perte du phonème /l/, et cela s'explique peut-être par sa position coda de syllabe, mais pas coda de mot.

5.2.6.2 Variantes graphiques des <l> sur BP

Son	Variante	Exemple
/l/	<VlV> → Ø	soley, '*solely*'

Tab. 5.25 – Variation des graphies des liquides en AAV

Les graphies de BP ne présentent aucune différence pour le phonème /l/, à part des chutes très marginales, comme on peut constater dans le tableau 5.25. Le seul exemple dans le corpus est la graphie <soley> (2 750 occurrences sur le forum), pour remplacer « solely ». Cette graphie est utilisée dans les pseudonymes (trois résultats dans les dix premiers résultats d'une recherche sur le mot), ainsi que dans les messages, comme « Is Bp Soley for African americans and pple living in the US ? » (BPNA4) ou dans cet extrait de poème sur une page personnelle, « Oh, how I wish you would commit with soley me » (BPPP25). Si cette graphie est répandue, c'est peut-être à cause d'une prononciation trop rapide du mot, qui géminerait le premier et le second /l/. Cette hypothèse semble plus vraisemblable que la

23. Douteuse à cause d'un accent secondaire sur une syllabe non accentuée, qui est peut-être une simple faute de frappe de la part de l'auteur.

faute de frappe généralisée, parce que les lettres <e>, <l> et <y> sont assez éloignées pour permettre d'alterner d'une main à l'autre, si l'on se fie au diagramme de la répartition standard des doigts sur un clavier (illustration 5.2). Il n'y a pas d'exemple musical de l'utilisation de ce mot, qui est peu courant (6 155 occurrences dans le COCA). L'absence d'autres mots dont le morphème <-ly> perdrait le phonème /l/ pointe plus vers l'hypothèse d'une faute d'anglais généralisée que d'une observation phonologique.

5.2.6.3 Récapitulation des variantes graphiques de <l>

Lorsqu'il est affecté par des changements, donc dans le corpus SC, le son /l/ a tendance à disparaître. Aucun autre son n'est remplacé par cette latérale, qui est souvent vocalique. Le tableau 5.26 montre que seul la chute du <l> graphique est partagée par les deux corpus, bien que la proportion en soit négligeable sur BP qui n'a pas d'autre exemple de ce phénomène. Il y a 2 750 occurrences de la graphie <soley> sur BP.

Ce panorama des consonnes se finit naturellement par le mode articulatoire qui se place entre les

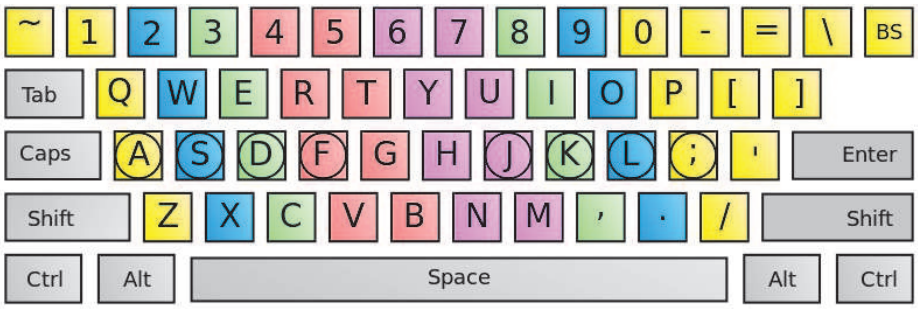


Illustration 5.2 – Position et zone des doigts typiques sur un clavier QWERTY (source : Wales et Sanger, 2001)

	SC	BP
<l#> → <'>, Ø	×	×
<l#> → <w>	×	
<V +l> → <ye>	×	

Tab. 5.26 – Récapitulation des changements graphiques impliquant des fricatives dans les deux corpus

consonnes et les voyelles : les phonèmes glissés /w/ et /j/. Bien que j'aie déjà traité certains cas de

phonèmes consonantiques devenus des glissées, certaines glissées sont également affectées par des changements traduits dans les graphies alternatives.

5.2.7 Traitement des glissées

5.2.7.1 Variantes graphiques des glissées sur SC

Les glissées sont les consonnes les plus souvent réinjectées dans la graphie, et leur statut de semi-voyelle les rend plus marquées, tout en permettant des transformations plus naturelles en glissées pour les autres consonnes. De plus, le graphème <y> est hybride, lui aussi, entre la consonne et la voyelle. Sa grande fréquence pour les diphtongues de l'anglais, comme dans la séquence graphique <ay> (/eɪ/), tout comme son utilisation en attaque de syllabe, comme dans « year » en fait un graphème particulièrement polyvalent. Le graphème <w> est également souvent présent dans des digraphes à valeur vocalique (<ew>, <ow>, <aw>), tout comme il est employé avec une valeur consonantique, notamment dans des mots grammaticaux ; la graphie <w> représente 92% des occurrences du son /w/ en anglais.

En scots, particulièrement, les glissées sont écrites avec <w> pour /w/ et <y> pour /j/. Les utilisateurs de SC optent pour des caractères montrant la différence avec l'anglais standard. La prononciation de <Glesca> est ˈgleskə, selon l'OSD. Le dictionnaire propose trois autres variantes de prononciation, /ˈglɛzgə, ˈglɛskə, ˈglezgə/ qui ont toutes le point commun de ne pas inclure de diphtongue, à la différence de la prononciation standard /ˈglɑːzgəʊ/ (Longman). On peut faire remonter cette prononciation au proto-brittonnique **glas cu*, '*vallon vert*' [24]. La monophtongaison, ou plutôt l'absence de diphtongue explique la chute de ce <w> vocalique. À l'inverse, le graphème <w> est réinjecté dans la

Son	Variantes	Exemple
/w/	<w> → Ø	Glesca, '*Glasgow*'
/w/	/w/ → <w>	wan, '*one*'
/g/	<g> → <y>	foryet, '*foryet*'

Tab. 5.27 – Variation des graphies des glissées en scots

24. Cette appellation est souvent traduite plus plaisamment par « the Green Dear Place », surnom de Glasgow et titre d'une série qui se déroule dans un parc de la ville.

graphie lorsqu'il est prononcé : c'est le cas de la graphie <wan> de « one », qui suit plus fidèlement la prononciation écossaise /wan/ (prononciation en anglais standard : /wʌn/). Ce mot est également écrit <yin>, avec l'utilisation de l'autre consonne glissée /j/, ce qui constitue une variante phonétique proche, transcrite /jɪn/ dans l'OSD.

Dans « foryet », c'est la forme écossaise qui prime : selon l'OSD, c'est une prononciation du nord de l'Écosse ([fər'jɛt]). L'origine de cette prononciation pourrait être l'influence du /r/, qui influence la prononciation de la consonne suivante, mais ce n'est qu'une conjecture personnelle. Historiquement, le <y> figure ici le yogh moyen-anglais (<ȝ>), qui est préservé en scots, mais absent des cognats concurrents (*vergessen* allemand, *forġietan* vieil-anglais). Ceci montre que la prononciation est évidemment plus importante que la graphie. Certains historiens ont recours au <3> pour remplacer <ȝ> pour des raisons pratiques, ce graphème n'était pas disponible dans toutes les polices. Les scripteurs de SC n'ont pas forcément cette culture diachronique, étant donné qu'aucun message du forum n'utilise un <3> pour remplacer yogh, mais ils tiennent tout de même à changer la graphie standard pour représenter le /j/. Il n'y a pas de graphie concurrente, les recherches pour <forjet> et <foriet> ne renvoient à aucun résultat alors qu'elles sont utilisées sur internet, principalement dans les archives de livres en scots. On trouve la forme <foryet> dans le SCOTS et le DSL, et cette transcription semble donc être la plus récente. Les deux autres graphies sont surtout conservées dans des noms propres, et n'ont pas d'existence en scots contemporain. Le graphème <y> est donc optimal pour retranscrire le yod prononcé en scots moderne [25] : son statut est semi-consonantique, et sa prononciation n'est pas ambiguë dans cet environnement.

5.2.7.2 Variantes graphiques des glissées sur BP

Son	Variante	Exemple
/j/	<an> → <ay(u)m>	mayn, '*man*', dayum, '*damn*'

Tab. 5.28 – Variation des graphies des glissées en AAV

25. Le son correspondant à <ȝ> était plus proche d'un /ɲ/, qui a progressivement évolué en yod.

En AAV, le graphème <y> sert à regraphier certains mots, et code notamment l'allongement des voyelles. Son utilisation dans les graphies <mayn> et <dayum> (« man » et « damn », respectivement) rappelle celle de <scayed> (« scared »). Les voyelles dans les deux mots sont allongées et diphtonguées. Les voyelles de ces mots appartiennent à la classe d'isophones TRAP (réalisée en AAV /æɛ/), ce qui les rapproche quelque peu de la classe d'isophones de FACE. La diphtongue de FACE est représentée soit par la séquence <aCe>, qui ne correspond pas aux deux placements en syllabe fermée, donc les scripteurs de BP ont recours à la séquence graphique <ay>. La voyelle est tellement allongée dans <dayum> que le <y> peut même faire figure de borne syllabique : le mot est parfois prononcé [ˈdeɪjəm] [26], ce qui montre la resyllabification subie par ce mot. Son allongement est parfois aussi montré par les multiplications de lettres, comme <daaayum> ou <dayummm>, plus esthétiques et expressives qu'indicatrices d'un changement ultérieur de prononciation.

5.2.7.3 Récapitulation des variantes graphiques des glissées dans le corpus

	SC	BP
<an> → <ay(u)m>		×
<g> → <y>	×	
<m> → <w>	×	
<w> → Ø	×	

TAB. 5.29 – Récapitulation des changements impliquant des glissées dans les deux corpus

Beaucoup de sons consonantiques perdent leur marque en devenant <w> ou <y>, et certains sons vocaliques sont également signifiés par cette graphie, comme on peut le voir dans le tableau 5.29. Ces deux semi-consonnes sont donc correctement perçues par les locuteurs comme des jokers phonologiques, entre indicateurs de diphtongaison (comme dans <mayn>) et témoins d'une resyllabification (<dayum>). Le graphème <y> peut représenter une regraphie contenant le son /j/, comme <yin> pour « one ». Il existe une plus grande propension en scots à passer d'une consonne à une glissée (comme le montrait la graphie <ower> pour « over »), voire à une voyelle (<aw> pour « all »), dans un entre-deux entre l'anglais (ou l'AES) et le scots pur. Elle s'explique certainement par le tâtonnement des

26. Voir à ce sujet la vidéo qui a provoqué des remix mémétiques intitulés « Oh my dayum » disponible à l'adresse <http ://-bit.ly/1mIMmNR> (dernier accès le 10/07/2014). À 5 :03-5 :09, l'auteur lance trois « damn » particulièrement expressifs, où la resyllabification est manifeste (« Five Guys Burgers and Fries Review », 2012).

locuteurs, qui composent avec leurs connaissances de la graphie de l'anglais standard, autant que de leur perceptions phonologiques.

5.2.8 Réduction de groupes consonantiques

Son/mode	Transformation	Exemple
/ld/	<ld> → <l>	ol, (BP), spennin, '*spending*' (SC)
/st/	<st> → <ss>	massa, '*master*' (BP)
/lf/	<lf> → <l>	masel(l) (SC) ;

Tab. 5.30 – Variation des groupes consonantiques

Dans cette sous-partie, AAV et scots seront traités simultanément, tant les simplifications des agrégats consonantiques sont similaires dans les deux corpus. Cette partie complète celle sur les occlusives, qui font souvent les frais de la réduction de groupes consonantiques. L'agrégat consonantique qui tend à se simplifier le plus promptement est le <ld>. Ce groupe est toujours en position coda dans une syllabe, et le /d/ a tendance à disparaître de la graphie. Cette disparition est connue en scots, notamment, dans des mots tels que « hieland » (trouvé sous la forme <hielan> dans le corpus). Selon l'OSD, ce mot est prononcé ˈhilən(d), -lənt], donc plus souvent avec une prononciation dévoisée de la dernière consonne, quand elle est prononcée. Le phénomène se produit également sur BP, bien qu'il n'y en ait pas autant d'exemples. La graphie <ol> voire <ole> pour « old » n'appartient pas qu'à ces deux langues minoritaires, et la différence de marque phonétique entre /l/ et /d/ explique l'amuïssement du second phonème, même en milieu de mot (comme dans « spennin »). On retrouve également cette graphie dans l'expression américaine « good ol' boy » [27], pour laquelle la graphie sans <d> à « old » a un tiers d'occurrences de plus sur Google que son concurrent graphique standard « good old boy ». Cette tendance vers la graphie <ol> pour « old » dépend du contexte et de l'expression : « old boy network » [28] est plus volontiers orthographiée <old[-]boy['s]> que <ol' boy> ou <ole boy> (sans le <d> d'« old »). Il y a 249 000 occurrences des variations orthographiques, contre seulement 73 300 résultats pour <ole boy> et 67 700 occurrences de la graphie <ol' boy> [29]. La variation graphique sur

27. Traduction : homme blanc du Sud des États-Unis, qui en accepte les valeurs traditionnelles.
28. Parfois orthographiée « old-boy network » ou « old boys' network ».
29. Recherche effectuée le 10 juillet 2014.

ce mot dépend donc de son utilisation, mais elle est connue de tous, et donc potentiellement actualisable à l'écrit.

La graphie <masel> est le seul exemple de la disparition de /f/ dans la séquence /lf/. Cette graphie existe également dans le corpus TR-L, et est confirmée par l'OSD, qui indique également que le /f/ n'est pas prononcé. Les autres pronoms réfléchis ne sont pas très courants dans le corpus, et les graphies qui effacent le <f> sont des hapax legomena, comme « yersel », « yarsel », ou ont très peu d'occurrences, comme « hersel » (10 résultats dans SC) et « himsel » (17 résultats dans SC). Ce n'est pas le cas du COCA, qui liste « myself » derrière « themselves » (450ème position) et « himself » (370ème position), en 663ème position. Ce n'est pas le cas dans le corpus SC, où il y a 133 occurrences de « masel », certainement parce que les scripteurs parlent d'eux-mêmes, et utilisent donc ce pronom réfléchi plus souvent.

Massa est une prononciation historique de « master », associée à l'esclavagisme américain. On le trouve dans la littérature, surtout en tant que transcription de la prononciation des esclaves. Une recherche des Ngrams sous Google [30], qui consiste à rechercher dans un corpus littéraire de textes entre 1800 et 2000 donne le résultat présenté dans l'illustration 5.3, pour les trois types de graphies possibles (<Massa>, <massa> et <MASSA> [31]). Le pic d'apparition est en 1863 pour les deux formes <massa> et <Massa>, avec une nette préférence pour <Massa> (en rouge). Le candidat graphique <MASSA> est quasiment absent des résultats, ce n'est donc pas un acronyme qui pourrait changer les résultats de la requête. Les occurrences consultables sont toutes composées de discours directs d'un esclave à son maître (ce mot étant utilisé comme vocatif), et on retrouve également ce mot en pidgin camerounais, où il est graphié <masa> mais y signifie « patron » (Bellama, Nkwele, Yudom, Peace Corps (U.S.) & Cameroon, 1983). Cette simplification du groupe consonantique pourrait être historiquement liée à la phonologie des langues ouest-africaines. Elle n'est utilisée sur BP qu'avec ironie, comme dans ce titre d'article : « can a black conservative serve his massa and still own his soul » (BPNA3). Ce mot et ses implications historiques sont complètement absents du corpus SC, qui ne contient aucune réduction de <st> en <ss>, même dans d'autres mots.

30. Disponible à l'adresse <https ://books.google.com/ngrams/>, dernier accès le 01/05/2014.
31. Comme il est impossible de rechercher les mots qui contiennent des caractères extra-alphabétiques, la graphie potentielle <m.a.s.s.a> n'est pas étudiée. Elle n'a été observée nulle part dans le corpus.

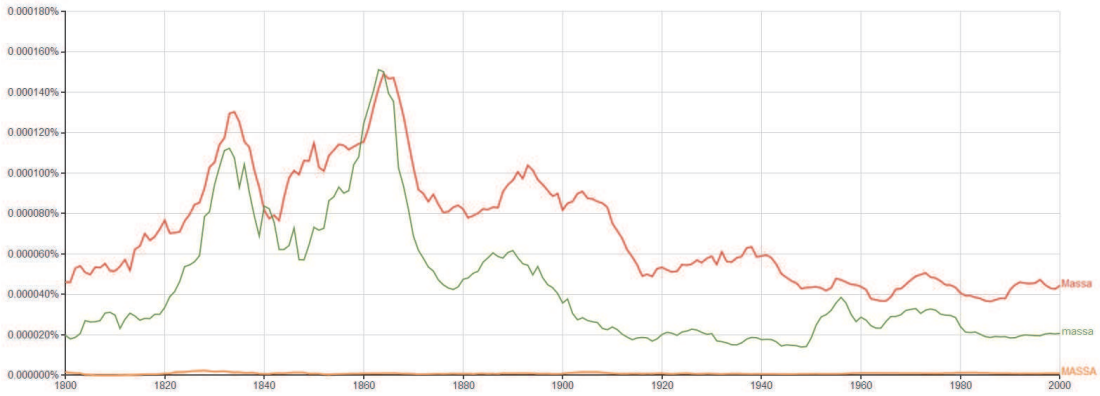


Illustration 5.3 – Occurrences de <massa> dans la littérature (Google Ngrams, résultats du 23 août 2014)

Le statut du digraphe <-ng> a déjà été traité précédemment, dans la partie 5.2.4, et il n'est groupe consonantique que dans sa transcription orthographique. Phonologiquement, ce digraphe ne montre que la vélarisation de /n/ et le passage à /ŋ/. Représenter un seul phonème avec deux graphèmes alors que chacun a une valeur phonologique, où <n> = /n/, et <g> = [+vélaire] pousse la graphie vers une simplification, reflet de la prononciation. Ce n'est pas le cas pour des digraphes comme <ch> qui représente /x/ en scots ou /ʃ/ dans les mots d'origine française, par exemple.

Plus le phonème est marqué, plus il subit de différences graphiques. Cela ne signifie pas qu'il n'y a pas de différences pour les occlusives, mais qu'elles ont simplement tendance à chuter, alors que les autres consonnes ont tendance à être remplacées par un équivalent consonantique (ou semi-consonantique). Le modèle chamito-sémitique de préservation des consonnes est moins important que la retranscription des sons tels qu'ils sont prononcés. À présent que l'inventaire de la variation des consonnes est réalisé, je vais aborder l'inventaire des graphies alternatives qui affectent les voyelles.

5.3 Variation allographique des représentations des voyelles

Les voyelles sont les éléments les plus prompts à évoluer en fonction des utilisations de chacun. La phonologie des deux langues minoritaires, traitée dans le chapitre 2 grâce notamment aux travaux de

J. Stuart-Smith et J. Dillard (Dillard, 1972 ; Stuart-Smith, 2003) sera maintenant confrontée aux graphies adoptées par les membres des forums.

Les voyelles représentent le plus visiblement les utilisations dialectales, notamment grâce aux graphies démarcatives (évoquées dans le chapitre précédent, et utilisées pour s'éloigner de l'anglais standard). L'approche des réélaborations graphiques des voyelles sera donc différent, puisqu'il s'agira de comparer les productions au standard. Je partirai donc de la graphie des forums pour revenir à la graphie-source de l'anglais. Ce travail sera facilité par les transcriptions phonétiques de l'OSD pour le scots, et par d'autres sources pour l'établissement des graphies alternatives, notamment dans les noms de chansons afro-américaines. Par exemple, dans le titre « Hot in herre » de Nelly, la prononciation homophonique de *here* avec *her*, et donc la rhoticisation et le changement de la valeur vocalique en /hɜˠ/ est signifiée par la graphie <herre> (« Nellyville », 2002).

Les archigraphèmes des voyelles de l'anglais sont beaucoup plus difficiles à dessiner, à cause du changement inhérent de prononciation suivant la position de la voyelle dans une syllabe, et de la présence éventuelle du /r/, qui change le son. De plus, <u> a déjà deux prononciations en syllabe fermée depuis le foot-strut *split*, ce qui complexifie d'autant plus le propos. La fréquence d'utilisation sera le guide principal pour déterminer les archigraphèmes vocaliques de l'anglais. Le tableau 5.31 donne encore une fois les propositions d'archigraphèmes de l'anglais. Au lieu de n'utiliser que le phonème comme point de départ, j'utiliserai également la classe d'isophones de Wells pour tenter de combler les différences entre la phonologie de l'anglais, de l'AAV et du scots. Le problème principal est le nombre de symboles pour représenter les voyelles de l'anglais : il n'y a que six lettres, <y> inclus, pour représenter les 24 classes d'isophones de l'anglais [32]. Il y a également moins de consonnes que les 24 sons consonantiques de l'anglais, mais la différence est moindre.

Dans la description de phonèmes, s'il figure deux sons différents, le premier est en RP et le second, en GA. J'ai préféré garder les oppositions telles qu'elles étaient présentées (en adaptant Wells, 1982), afin d'être la plus exhaustive possible. Cet inventaire est rendu d'autant plus difficile que les deux variétés d'anglais, RP et GA se confondent (dans thought, north et force) en RP. Cet inventaire n'est pas aussi performant que celui des consonnes, loin de là, et ne prend pas vraiment en compte le pla-

32. Sans compter les trois sons de voyelles non accentuées.

Phonème	Classe d'isophone	Archigraphème	Fréquence
	Voyelles brèves		
/æ/	TRAP	<A>	96%
/e/	DRESS	<E>	91%
/ɪ/	KIT	<I>	66%
/ɒ, ɑ/	LOT	<O>	79%
/ ɔ/	CLOTH	<O>	79%
/ʌ/	STRUT	<U>	86%
/ʊ/	FOOT	<U>	69%
	Voyelles longues		
/ɜː, ɜr/	NURSE	<IR>	40%
/ɑː, æ/	BATH	<AR>	89%
/ɑː,ɑ/	PALM	<AR>	89%
/iː, i/	FLEECE	<EE>	6% (<e> = 70 %)
/ɔː, ɔ/	THOUGHT	<O>	97%
ɔː, ɔr	NORTH	<OR>	97%
ɔː, or	FORCE	<OR>	97%
/ɑː,ɑr/	START	<AR>	89%
/uː, u/	GOOSE	<OO>	38%
	Diphtongues centripètes		
/ɪə, ɪr/	NEAR	<ERE>	—
/ɛə, ɛr/	SQUARE	<ARE>	—
/ʊə, ʊr/	CURE	<URE>	21% (<u> = 69%)
	Diphtongues centrifuges		
/eɪ/	FACE	<A(Y)>	45%
/əʊ, o/	GOAT	<O>	73%
/aɪ/	PRICE	<I>	37%
/ɔɪ/	CHOICE	<OI>	62%
/aʊ/	MOUTH	<OU>	56%
	Voyelles non accentuées		
/i/	HAPPY	<Y>	41%
//ɚ/	LETTER	<ER>	40%
/ə/	COMMA	–	–

Tab. 5.31 – Tentative d'inventaire des archigraphèmes des voyelles de l'anglais

cement de la voyelle en syllabe fermée ou ouverte, voire l'influence d'un <r> postérieur. Il est donc uniquement présenté pour tester la robustesse du système, et rendre compte de la difficulté d'attribuer une graphie symbolique à une classe d'isophone. C'est aussi le cas des utilisateurs de BP et SC, qui se confrontent à des graphies alternatives qui se rapprochent le plus des prononciations en langues minoritaires. Pour s'en éloigner, on peut s'inspirer de la répartition des sons vocaliques de l'anglais (Deschamps, 1994), et la rapprocher des classes d'isophones.

Il y a donc deux tableaux pour les voyelles de l'anglais. Le premier donne des équivalences graphophonémiques en fonction du point d'articulation, et le second tableau récapitule et classe les classes

Voyelles brèves (/V̆/)	
/æ/	TRAP
/e/	DRESS
/ɪ/	KIT
/ɒ, ɑ, ɔ/	LOT, CLOTH
/ʌ/	STRUT
/ʊ/	FOOT
Voyelles brèves modifiées par <r> (/V̆ + R/)	
/ɑː, ɑr/	START
/ɜː, ɜr/	NURSE
/ɔː, ɔr/	NORTH
Voyelles longues (/V̄/)	
/eɪ/	FACE
/iː, i/	FLEECE
/aɪ/	PRICE
/əʊ, o/	GOAT
/uː, u/	GOOSE
/aʊ/	MOUTH
/ɔɪ/	CHOICE
Voyelles longues colorées par <r> (/V̄ + R/)	
/ɛə, ɛr/	SQUARE
/ɪə, ɪr/	NEAR
/ɔː, or/	FORCE
/ʊə, ʊr/	CURE
/aʊə/	—

Tab. 5.32 – Voyelles et classes d'isophones à partir de Deschamps, 1994

d'isophones en fonction du type de voyelle dont il s'agit. En anglais, la valeur de la voyelle dépend de son contexte, hors quelques exceptions, et le système est relativement cohérent, sinon, l'idée même de graphématique ou de correspondance systémique entre graphie et phonie ne se poserait pas. Les différences vocaliques en scots doivent également prendre en compte la règle de longueur des voyelles ou règle d'Aitken (Trudgill, 1984) :

V → [+ long] / _{[+voisée, + fricative], /r/#}.

Cette étude vocalique sera donc concentrée autour des cinq voyelles de l'alphabet, dans une perspective graphématique. Les mots contenant des voyelles représentées par un seul graphème, lui-même différent de la graphie standard, sont placés sur un trapèze vocalique, où les phonèmes correspondants sont eux aussi indiqués.

5.3.1 Variations graphiques des voyelles brèves en AAV

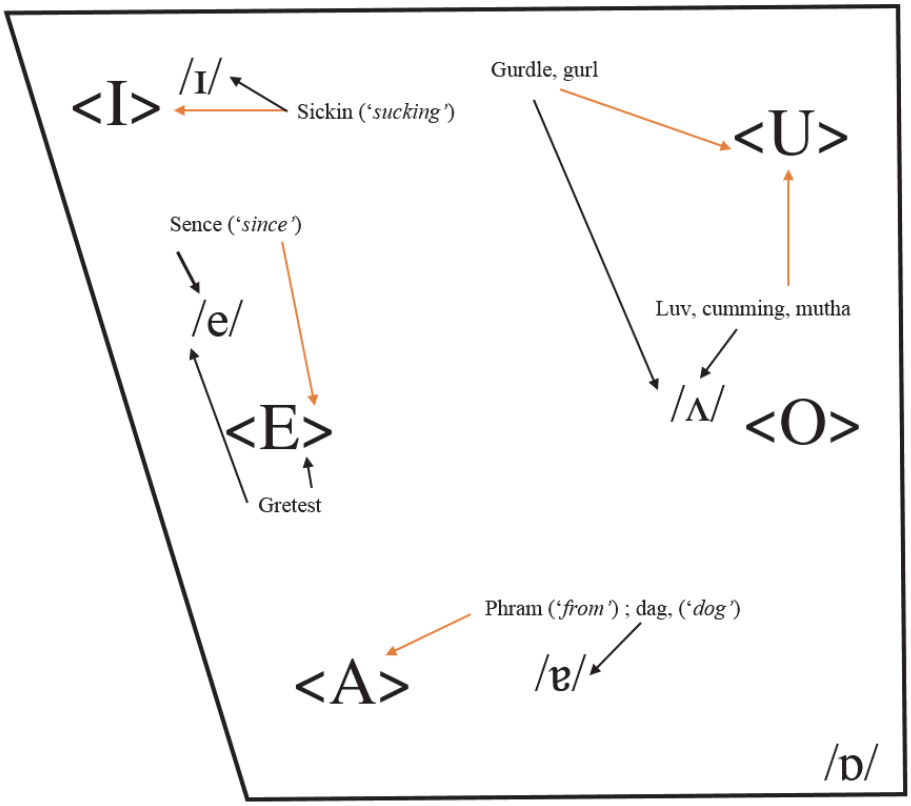


Illustration 5.4 – Exemples de graphies alternatives sur voyelles brèves en AAV

Les graphies afro-américaines donnent deux indications importantes : elles confirment la prononciation et régularisent la graphie. Un grand nombre de graphies, comme <gurdle>, <gurl>, <luv>, <cumming> et <mutha> utilisent le <u> (donc en environnement <_+r> ou <_CC>), qui correspond à un /ʌ/ et à la classe d'isophone STRUT, ce qui est confirmé par un dictionnaire de prononciation (*Longman Pronunciation Dictionary* (LPD)). Paradoxalement, ce retour logique à la graphie qui correspond le plus au phonème la rend incorrecte. Certains de ces mots, comme « love », « come » ou « mother » sont trop courants pour que leur orthographe soit inconnue. Il s'agit donc d'un effort conscient pour que la graphie mime la prononciation. C'est aussi le cas de <gretest> pour « greatest ». Le digraphe standard représente le standard /eɪ/ en RP et en GA, qui est monophtongué en AAV. La classe d'isophone *dress* est plus souvent représentée par un <e> graphique, et il semble qu'il s'agisse encore d'une normalisation de la graphie par la phonie.

Le changement vocalique afro-américain est représenté dans l'illustration 5.5. Les flèches marrons représentent les changements vocaliques phonologiques suggérés par les graphies alternatives. Ce schéma peut expliquer en partie les graphies <sence> et <sicking>. S'il existe une confusion entre DRESS et KIT dans le changement vocalique afro-américain, les voyelles graphiques qui les représentent changent également. Le glissement de DRESS vers KIT est illustré par <sence>, où la valeur de la voyelle est remise en question, et rapprochée de /e/. La graphie <sicking> suggère un nouveau type de changement, mais cette fois de KIT vers STRUT, ou au moins une plus grande porosité entre les deux classes d'isophones.

Quant à <phram> pour « from » et <dag> pour « dog », ils illustrent le glissement de LOT vers

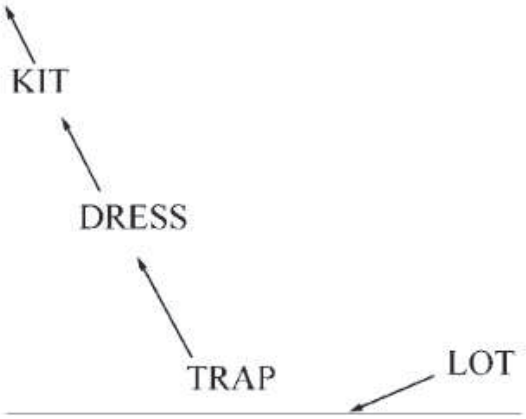


Illustration 5.5 – Changement vocalique afro-américain (THOMAS, 2007)

TRAP, ou au moins vers un son vocalique intermédiaire, peut-être un /ɐ/, voire un /æ/ à part entière. Encore une fois, ces mots sont très courants, d'autant plus que l'un d'entre eux est une préposition, et leur orthographe n'est pas inconnue aux membres de BP. La substitution d'une consonne simple <f> pour un digraphe à valeur phonique équivalent, <ph>, indique une graphie alternative volontaire, qui déroge au principe d'économie ou de préservation du nombre de lettres dans un mot. Le <a> graphique sert à marquer une plus grande aperture, dans une prononciation vernaculaire du mot « from » qui se rapproche de la réalisation plus ouverte de la voyelle en jamaïcain. La volonté de se séparer le plus possible de la graphie normée indique une prononciation non-standard. En l'occurrence, il s'agit d'un quasi-hapax (seulement 6 occurrences) dans le corpus BP, et l'utilisateur tente d'imiter un accent

jamaïcain : « For all those who come “phram yawd” » (BP12E5581). « Yawd » est une des graphies observées dans la partie 5.2.5, où l'allongement de la voyelle /æ/ avant un /r/ est une diphtongue, et ce mot (ou sa forme orthographique « yard ») est utilisé pour désigner la Jamaïque. C'est une synecdoque avec les « government yards » de Trenchtown, en Jamaïque. Le terme a même traversé l'Atlantique en devenant « yardie », qui décrit les délinquants d'origine jamaïcaine (ou simplement noirs) dans la Grande-Bretagne des années 1950. C'est donc une imitation, comme celles de Sick Boy/Simon, transcrites dans TR-L (et dont la prononciation est confirmée dans TR-F), où la graphie tente de se rapprocher le plus possible de la phonie. En revanche, ce n'est pas le cas de <dag>//dæg/, qui est rentré dans le vocabulaire afro-américain, en tant que terme d'adresse entre hommes. La graphie <dawg> (148 000 occurrences) est nettement plus fréquente que <dag> (22 400 occurrences), et dépasse même le nombre d'occurrences de la graphie standard (61 900 résultats pour <dog>). Ces deux variantes graphiques montrent le passage de la classe d'isophones LOT à TRAP, voire à un allongement diphtongué qui correspondrait à MOUTH (/daʊg/). L'antériorisation est attestée et confirmée par la graphie, qu'elle aille vers une antériorisation pure (de LOT à TRAP) ou vers une diphtongue (de LOT à MOUTH). La première partie de la diphtongue fermante est le premier élément à changer : LOT → TRAP → MOUTH.

Les graphies afro-américaines des voyelles confirment les observations phonologiques conduites précédemment (THOMAS, 2007). Certaines graphies permettent de reconnaître des groupes sociaux (les allusions à la Jamaïque avec « yawd », par exemple), mais la plupart d'entre elles sont personnelles, et semblent provenir du désir de rapprocher le plus possible la graphie de la phonie. Aucune graphie alternative étudiée ici ne transforme la voyelle brève en <o>, ce qui confirme le mouvement général de montée et d'antériorisation des voyelles.

5.3.2 Variations graphiques des voyelles brèves sur SC

Quant aux changements affectant les voyelles monographiques sur SC, ils sont également nés d'une volonté de se rapprocher de la phonie. Leur représentation sur un trapèze vocalique figure dans l'illustration 5.6. Pour ce corpus, il est plus facile d'avoir des confirmations de la prononciation et des gra-

phies, grâce aux transcriptions phonologiques l'OSD.

Les graphies écossaises témoignent de deux regroupements, autour de l'anglais d'une part, et du

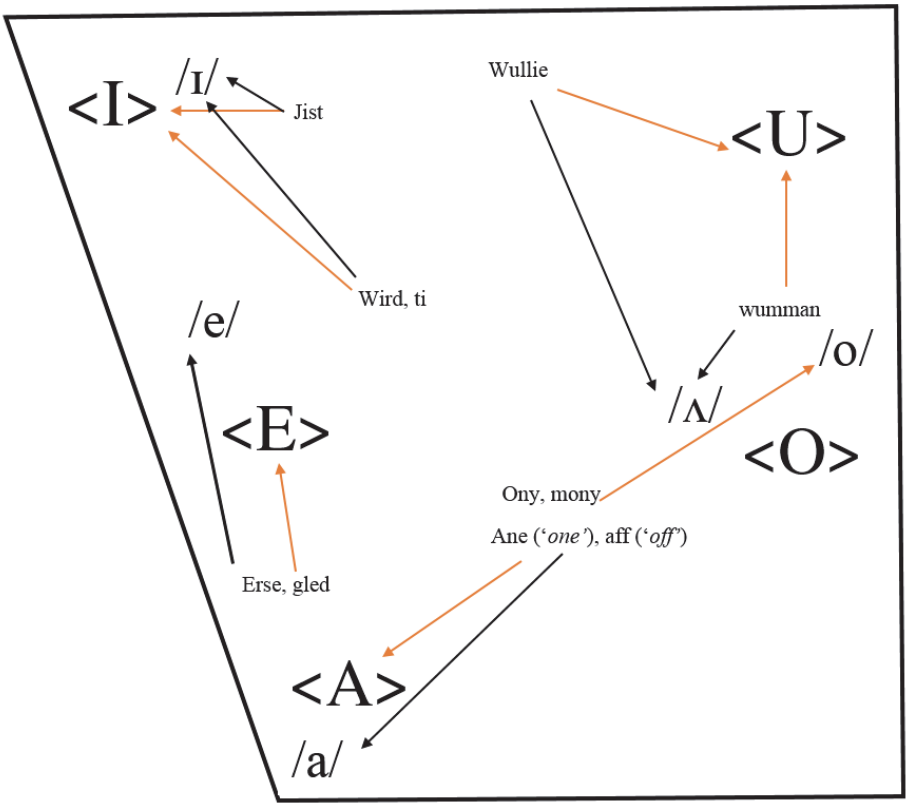


Illustration 5.6 – Exemples de regraphies sur voyelles brèves en scots

scots de l'autre. Comme dans le corpus afro-américain, ces graphies se manifestent par un réalignement graphique qui vise à harmoniser l'écriture et la prononciation. J'ai baptisé ce concept réélaboration graphique ou re-graphie. Par exemple, le changement graphique de « woman » en <wumman> rapproche la voyelle graphique de sa prononciation (/ʌ/, soit la classe d'isophone STRUT). Ce mot est aussi le plus proche de l'anglais, dans le champ sémantique « femme » en scots, selon l'OSD : le dictionnaire liste quatre traductions possibles pour le « woman » anglais (« dame », « hen », « wife » et « wumman »). Le corpus SC contient également « quine ». Ces mots s'éloignent tous de la sémantique de l'anglais : une « wife » n'est pas forcément mariée, en scots ; et une « dame » n'a pas de titre honorifique, c'est plus souvent une jeune femme qui travaille dans une ferme. <Wumman> est le plus proche du sens anglais, qu'il embrasse totalement, mais la graphie s'harmonise avec la prononciation. Il s'agit

certainement d'un emprunt tardif à l'anglais, le mot n'étant pas défini dans le DSL. La regraphie de « wumman » s'attache à le garder proche de l'anglais, d'où l'absence d'un mot du champ sémantique écossais plus éloigné, tout en le pliant à la prononciation /ˈwʌmən/.

Le changement vocalique écossais semble suivre le changement afro-américain, dans le mouvement d'antériorisation et de réduction de l'aperture. Les autres mots cités dans cette étude ont permis la création du schéma 5.7. Il récapitule les évolutions des réutilisations vocaliques que les regraphies de SC permettent de retracer. Les classes d'isophone LOT et TRAP, voire BATH semblent se confondre dans le corpus SC, et l'échange des lettres <a> et <o> y est fréquent. L'OSD propose la prononciation /a/ pour <ane>, <aff>, sans jamais proposer une réalisation vocalique /ɒ/, /ʌ/ ou /ɔ(ː)/ pour les deux mots. Au contraire, la prononciation d'<ane> peut être /en/, soit une antériorisation et une réduction de l'aperture, vers la classe d'isophone DRESS. L'évolution de TRAP en DRESS est confirmée par les graphies <erse> et <gled> pour « arse » et « glad » : les voyelles courtes antérieures ont une aperture moindre en scots, ce que l'OSD montre également. Le dictionnaire ne propose aucune transcription utilisant un /a/ ou un /æ/, pour ces deux mots dont les entrées sont « gled » et « airse ».

Ce n'est pas le seul changement en chaîne vocaliques dans le corpus : les voyelles centrales, accentuées ou non (correspondant au classes d'isophones COMMA et NURSE) trahissent des pressions vers une antériorisation fermée (vers la voyelle de KIT). En position non-accentuée, comme dans /tɪ/, /tə/, le mot anglais « to » est graphié <ti>. La forme écossaise standard est <tae>, et celle-ci est plus fréquente dans le forum (7 800 occurrences contre 1 600). La prononciation /tɪ/ a donc influencé la réélaboration graphique, et c'est également le cas pour le /ɜː, ɜˠ/ dans le mot anglais « word ». La transcription de l'OSD /wɪrd/ est confirmée par la graphie écossaise <wird>, attestée dans le dictionnaire et sur SC. Cette tendance touche également les mots contenant un <a> dans la graphie anglaise standard (classe d'isophone TRAP), qui se rapprochent de *kit* en scots. C'est le cas de « change », dont la prononciation écossaise est /ʧɪnʤ/ (OSD), et qui est regraphié <chinge> sur le forum. C'est la même chose pour le schwa de « was », qui est graphié <wiz> sur SC.

Les mots dont la voyelle appartient à la classe d'isophones DRESS en anglais standard, comme « get » voient leur voyelle devenir un /ɪ/, comme <git>, attesté dans TR-L et dans le corpus SC et dont la pro-

nonciation écossaise est /gɪt/. Le mouvement ne s'arrête pas là, et les mots contenant un /ɪ/ en anglais standard peuvent voir leur prononciation évoluer vers un /ʌ/, comme <Wullie> pour « Willum » (graphie écossaise de « William », comme vu en 5.2.4). Ce glissement est peu commun (un seul exemple dans le corpus), mais peut suggérer un mouvement ultérieur vers l'arrière, pour que toutes les voyelles brèves ne saturent pas l'espace entre DRESS et KIT.

Les déplacements de voyelles brèves en scots et en AAV semblent aller vers une antériorisation des

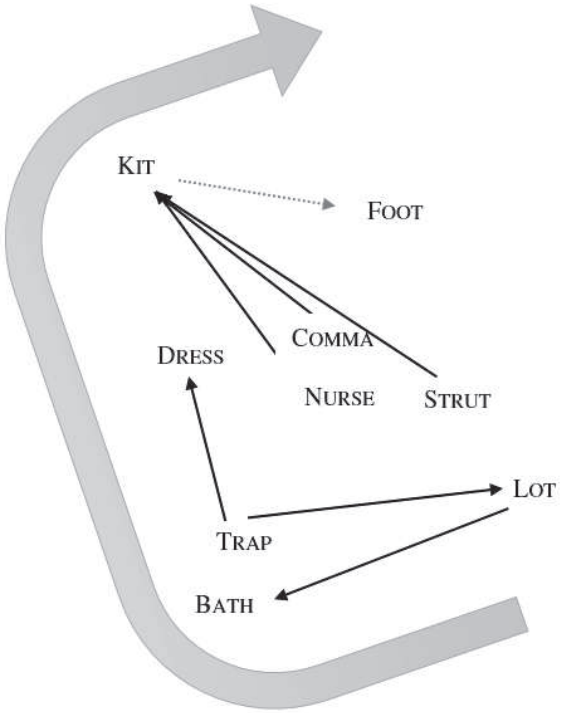


Illustration 5.7 – Évolutions vocaliques suggérées par le corpus SC

voyelles. Il y a une confusion entre TRAP et LOT, et donc entre les lettres <a> et <o>, qui sont parfois utilisées à la place l'une de l'autre dans les deux corpus, et cette tendance est plus nette dans le corpus SC, alors que dans BP, le mouvement ne va que de LOT vers TRAP. En revanche, le glissement de KIT vers DRESS afro-américain va à l'inverse du courant de SC qui fait remonter la voyelle de TRAP vers DRESS, et de DRESS vers KIT. Les deux corpus présentent un mouvement postériorisant la voyelle de KIT vers STRUT, voire FOOT.

5.3.3 Multigraphies pour voyelles inaccentuées

Les syllabes brèves et accentuées des deux parties précédentes permettent des observations pertinentes pour les voyelles monographiques. Les lettres représentant les sons correspondant à la classe d'isophones COMMA (le schwa) sont sujettes à des transformations étonnantes. Dans les deux corpus, j'ai sélectionné des mots dont les schwas sont représentés avec des graphies alternatives, pour voir quelles lettres étaient le plus souvent sélectionnées pour représenter les sons non-accentués, suivant leur position dans la syllabe (en syllabe ouverte, fermée, et en position coda). La position dans le mot est surtout importante pour le traitement de la voyelle, et sa perception par les scripteurs des forums. La répartition des graphies alternatives sur les schwas phonologiques est présentée dans le tableau 5.33. Ce tableau ne représente que les variations de certains scripteurs du corpus.

SYLLABE OUVERTE		
	SC	BP
<A>	afore (‘*before*’)	hypathetical (‘*hypothetical*’)
<E>	logerithum (‘*logarithm*’)	—
SYLLABE FERMÉE		
	SC	BP
<A>	—	yaself (‘*yourself*’)
<E>	yersel (‘*yourself*’)	—
<I>	intrist (‘*interest*’)	fagitry (‘*faggotry*’)
<O>	—	simpothy (‘*sympathy*’)
POSITION CODA		
	SC	BP
<A>	Glesca (‘*Glasgow*’)	yella (‘*yellow*’)
<E>	—	yeller (‘*yellow*’)
<I>	—	pusci (‘*pussy*’)
<U>	—	niggu (‘*nigger*’)

TAB. 5.33 – Distribution graphique des représentations de COMMA LETTER et HAPPY dans les deux corpus

Toutes les positions de schwa n'ont pas de graphie alternative,et toutes les voyelles alphabétiques occupent ce rôle, tout comme en anglais standard : le problème réside dans le fait qu'il n'y ait pas de voyelle qui aurait une valeur de schwa orthographique. Les choix graphiques alternatifs sont parfois involontaires, comme dans graphies <hypathetical> ou <logerithum>. Ces deux mots sont relativement

peu courants, et ces graphies sont certainement influencées par la prononciation. La seconde voyelle d'<hypathetical> semble indiquer une prononciation plus postérieure, et plus arrière, alors que l'inverse pourrait être vrai pour <logerithum>. La place de l'accent primaire est importante, et, par comparaison avec la première syllabe qui porte l'accent 1, la seconde de <logerithum> peut sembler d'autant plus faible et centrale. Cette faiblesse est représentée par un <e> graphique. La variation libre entre COMMA et KIT, constatée dans le corpus SC, semble attestée dans le corpus BP. C'est pourquoi les deux graphies <intrist> et <fagitry> [33] figurent la voyelle inaccentuée centrale avec le graphème <i>. Quant à la confusion entre <a> et <o> dans <simpothy>, elle montre que le passage de LOT à TRAP s'accompagne d'une variation avec les voyelles centrales sur le chemin entre ces deux points d'articulation. Le schwa pourrait être prononcé /ɐ/. Ces graphies sont très marginales dans les corpus principaux, et ces constatations ne peuvent que renforcer les conclusions antérieures sur les changements en chaîne afro-américains et scots qui rapprochent TRAP de LOT.

Les utilisations du <e> pour figurer un schwa suivent un chemin différent sur SC et BP. Les graphies alternatives de « yourself », la différence de choix pour la figuration de <our> est due à la rhoticité, et à son traitement différent dans les deux langues : l'AAV privilégie une aperture plus importante, ce que l'on retrouve dans les morphèmes finaux en <-er> graphiés <a>. Le scots rapproche <our> de <e>, certainement à cause du <r> postvocalique, ce qui prouve une variation entre NURSE et COMMA, qui eux-mêmes convergent vers KIT (comme je l'ai montré dans la partie précédente).

Les deux graphies afro-américaines de « yellow », <yella> et <yeller> sont particulières, parce que la monophtongaison garde une trace inexacte de l'ancienne longueur de la syllabe. La diphtongue inaccentuée semble conceptualisée comme un <er> final, comme le montre l'une des graphies. L'autre, avec <-a>, rapproche le mot de tous les mots dont la terminaison <-er> est écrite avec cette lettre en AAV. Dans les deux cas, la diphtongue finale semble monophtonguée, et elle est certainement perçue comme une voyelle moins ouverte.

La graphie écossaise <Glesca> s'en rapproche parce qu'elle finit par la même diphtongue standard /əʊ, oʊ/, également traitée comme un <a>, auquel correspond un schwa, ce qui est confirmé par

33. La graphie <fagitry> rappelle l'un des membres de cette famille et ses graphies alternatives <faggit>/<phaggit>, qui totalisent chacune plus de 460 occurrences sur BP.

l'OSD. Cela rappelle le cas d'<afore>, dont la graphie est écossaise (et attestée dans l'OSD et le DSL). Les mots <afore> et <before> n'ont pas la même origine, et sont respectivement construits sur [on] + [fore] et [by] + [fore]. L'évolution du préfixe écossais le laisse non-accentué, libre d'être graphié avec un <a>, et appartenant à la classe d'isophones COMMA [34], mais il est possible que la longueur vestigiale potentielle permette une prononciation /ɐ'foːr/.

L'utilisation d'un <u> pour figurer un schwa est plus rare, et le cas de <niggu> est un peu à part. C'est l'une des seules représentations de schwa qui implique un <u>, et elle est très marginale par rapport aux autres représentations graphiques de la dernière voyelle de « nigga ». Elles seront étudiées dans la partie 5.5.1. Un <u> final est rare en anglais, à part dans les emprunts aux langues comme le français (« adieu » ; « aperçu ») ou le japonais « tofu » ; « jiujitsu »). Le mot « nigger » et sa graphie « nigga » ne sont inconnues d'aucun utilisateur de BP, et cette graphie surprenante pour un schwa en position coda est certainement due à la volonté de choisir des caractères graphiques originaux, peu courants dans ces positions. C'est aussi le cas de la graphie <pusci>, où le <y> final laisse la place à un <i>, qui est encore une fois plus couramment trouvé dans des mots d'origine étrangère qu'en anglais. On retrouve cette volonté d'originalité graphique dans la représentation de la consonne /s/ par <sc>, qui est moins fréquente que <ss> [35].

Les classes d'isophones LETTER et surtout COMMA permettent des graphies aussi libres que possible, même si elles sont statistiquement rarement attestées dans l'anglais standard, comme <niggu>. Ces graphies permettent surtout de pouvoir vérifier les liens entre la perception d'un son et sa représentation graphique. Le schwa, censément central, peut se calquer graphiquement sur le caractère qui se rapprocherait le plus de sa réalisation : <e> semble représenter un schwa légèrement trop antérieur ; <a>, un schwa à l'aperture plus importante.

La dernière étude des voyelles se consacre aux diphtongues, et surtout aux phénomènes de sédimentation graphique de certaines diphtongues du scots, et des représentations phonétiques et leur interprétation à l'écrit.

34. L'OSD donne une transcription /ə'foːr/.
35. Le tableau des équivalences graphophonématiques est peu précis concernant la proportion de la graphie <sc> par rapport à <ss>, mais les <ss> sont comptés dans les 73% d'occurrences de <s> pour représenter /s/.

5.3.4 Diphtongues

Deux mouvements opposés existent : les deux corpus présentent des monophtongaisons et des diphtongaisons, suivant les règles phonologiques des deux langues minoritaires. Dans le tableau récapitulatif des monophtongaisons, je pars de la prononciation standard afin d'examiner les changements graphiques qui peuvent s'interpréter comme un changement phonologique. Les deux corpus présentent des changements inverses : le scots privilégie les monophtongaisons, alors que l'AAV a tendance à diphtonguer les phonèmes.

Les changements graphiques du scots sont présentés dans le tableau 5.34. La diphtongue /aʊ/,

	SC
/aʊ/→ /u/	<oo> = aboot, '*about*' ; <ou> = broun, '*brown*'
/əʊ/→ /e/	<ai> = almaist, '*almost*'
/eɪ/→ /i/, /e/	<ie> = eniwie, '*anyway*'<a> = awa, '*away*', <i> = chinge '*change*'
/aɪ/→ /i/	<ei> = heilan, '*highland*'

Tab. 5.34 – Changements graphiques qui suggèrent des monophtongaisons dans le corpus SC

souvent écrite <ou> ou <ow> en anglais standard est prononcée /u/ en scots. Plusieurs exemples l'attestent, selon l'OSD : <aboot> (/əˈbut/), <cooncil> (/ˈkunsəl/), <coo> (/kuː/ …Cette monophtongaison est intégrée par les utilisateurs qui ont recours au scots. Les graphies ne sont pas tout à fait stabilisées, et les dictionnaires ne proposent pas les mêmes représentations. L'OSD propose toujours une graphie en <oo> pour un mot dont la graphie anglaise est <ou>, et toujours la graphie <ou> pour les mots dont la graphie anglaise est <ow>, ce qui n'est pas le cas du DSL qui ne systématise pas ces graphies.

SC lui-même a adopté une graphie particulière pour l'un de ces mots, « coo », dans ses animations autour des vaches de la race Highland. Le site comporte un petit jeu gratuit (illustré en 5.8), et la possibilité d'avoir une vache sur son profil personnel (illustration 5.10) est vantée dans une des rubriques sur le site (illustration 5.9). À travers cette exposition, les utilisateurs savent peut-être quelle est la forme graphique préférée par la plateforme pour le son /u/, et s'en servent : il y a 25 500 occurrences de <coo> contre 2 940 de la graphie standard <cow>. En outre, le son /u/ est le pendant de la diphtongue, parce qu'aucun de ces mots ne correspond à une voyelle longue, suivant la règle d'Aitken [36].

36. Les voyelles écossaises ne sont longues qu'avant /r, v, z, ð/ ou une frontière de mot ou de morphème.

On constate donc la même réélaboration graphique, avec quelques variantes, pour les graphies

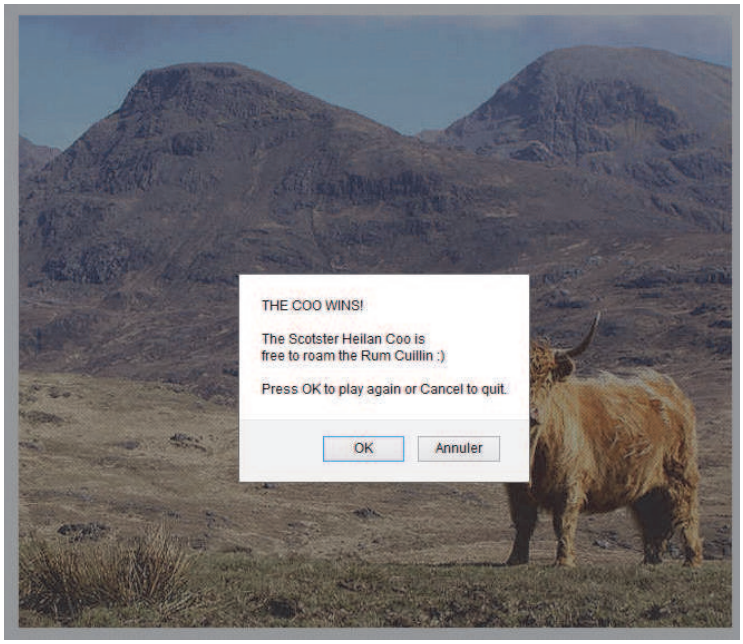


Illustration 5.8 – Écran de fin du jeu « Coo Game »

Illustration 5.9 – Rubrique du site SC

correspondant au son anglais /əʊ/, qui se prononce /e/ ou /ɛ/ dans les mots écossais correspondants, comme dans « almost », dont les transcriptions dans l'OSD sont /ˈɑːmest/, /ˈɑːmɛst/. La graphie <ai>

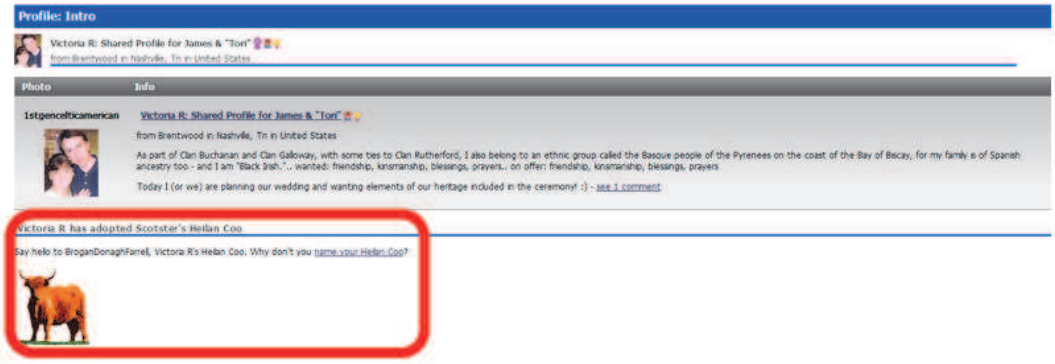


Illustration 5.10 – Extrait de profil comportant la rubrique « Heilan Coo »

est une norme écossaise utilisée pour retranscrire ces sons, même dans les mots qui n'avaient pas de diphtongue, comme dans <gairden> ou <airm>, repectivement prononcés /ˈgeːrdən, 'gɛr-/ et /eːrm, ɛrm/. Une seule convention graphique tend donc à harmoniser la graphophonématique.

Les deux diphtongues de l'anglais /eɪ/ et /aɪ/ sont plus complexes à traiter, parce que leur prononciation peut se confondre. Elles subissent une monophtongaison, qui peut privilégier l'attaque de la diphtongue, comme <awa> pour « away » (prononcé /əˈwɑː/), ou la seconde, comme c'est le cas pour <chinge> ([tʃin(d)ʒ]). La direction de la monophtongaison dépend peut-être de son environnement en syllabe ouverte ou fermée : la syllabe ouverte engage peut-être le déplacement de la diphtongue dans l'appareil phonatoire. Lorsque l'on observe les deux prononciations de <wie>, '*way*'selon l'OSD, [wəi, N. wae], on constate la scission entre le scots, qui conserve la diphtongue en AES, et le scots du nord qui abaisse la diphtongue : au lieu de passer de /ə/ à /ɪ/, elle passe de /a/ à /e/ suivant qu'elle soit en syllabe ouverte ou fermée.

Pour tester cette hypothèse, et voir si la graphie <ie> à la fin d'<eniwie> indique une monophtongaison, ou au moins un abaissement de la diphtongue, j'ai utilisé les transcriptions de vidéos de l'« Accent Tag »[37]. Une vlogueuse[38] de Glasgow ne diphtongue pas la fin du mot « anyway » et le prononce [ˈeniwe]. De surcroit, elle rappelle le suffixe affectif écossais, <-ie> (prononcé /i/ ou /e/ dans « bonnie » selon l'OSD), qui se place évidemment en position coda, comme il l'est dans <eniwie>.

37. Je m'appuie sur des travaux antérieurs ayant conduit à la constitution d'un corpus de parole afro-américaine et écossaise sur *Youtube*, transcrit orthographiquement et partiellement aligné (Goudet, 2013b).
38. Responsable d'un « vlog », ou blog vidéo.

La graphie <heilan>, '*highland*', est une graphie scots [39]. Le mot se prononce [ˈhiləɴ(d), -ləɴt] , sans diphtongue. L'entrée dans le DSL liste la graphie <heland> en tant que lemme, et cela interroge la conservation de <ie>s sur SC pour symboliser la monophtongue de la première syllabe. Les autres variantes graphiques du DSL comportent également un digraphe, comme <heyland>, <heeland>, <hyeland>…Cet allongement du mot est cosmétique, et sert certainement à pallier la disparition de <gh> à la fin de la première syllabe.

	BP
/i/ → /eɪ/	<ay> = bootay, '*booty*'
/ɒ/→ /aʊ/	<aw> = dawg, '*dog*'

Tab. 5.35 – Changements graphiques qui suggèrent des diphtongaisons dans le corpus BP

Les changements graphiques qui suggèrent des diphtongaisons ne sont pas tant afro-américaines qu'américaines. L'allongement de fin de mot en <ay>, prononcé /eɪ/ sert à changer le sens du mot. Selon l'UD, un « bootay » est une paire de fesses, souvent généreuse et attirante, qui provoque de l'enthousiasme chez le locuteur [40]. Le remplacement de fin de mot en <-y> par <-ay> existe aussi dans le mot <partay>. L'enthousiasme, et la qualité « supérieure » s'inspirent de la prononciation des emprunts au français, perçu comme une langue plus élégante, ou plus prestigieuse. Ainsi, les mots « touché » et « buffet » se prononcent respectivement /ˈtuːʃeɪ/ et /ˈbʊfeɪ/ [41]. La diphtongaison finale confère un caractère plus élégant au mot lui-même, sans gêner la compréhension, puisqu'il n'y a pas d'homophonie possible [42].

La tendance à diphtonguer les mots pour en changer légèrement le sens, ou pour des raisons d'expressivité accrue est commune aux allongements de /ɪə/ (observés plus loin en 5.5.2) et /aʊ/. L'allongement compensatoire en /aʊ/ après l'amuïssement de /r/ dans « Lawd » ou « yawd » ont déjà été traités dans la partie 5.2.5. Il existe aussi dans la graphie <dawg>, mais dans ce cas, il n'y a aucune question de rhoticité. La fusion de COT-CAUGHT explique le point de départ de cette diphtongue, et la vitesse

39. Et elle est adoptée par les créateurs de SC, qui l'utilisent dans les références aux vaches Highland.
40. Cette définition est un condensé des dix premières entrées de la définition de ce mot sur le site de l'UD.
41. Pour le repas et non pour le verbe, régulier suivant les règles phonologiques de l'anglais (source : Longman).
42. Pour cette raison, certains Américains parlent humoristiquement de Target, grande chaîne de supermarchés à la Walmart, en utilisant la prononciation francisante /ˈtɑːʒeɪ/.

d'élocution de l'ABS et de l'AAV sont connues pour être plus lentes que l'anglais parlé au Nord des États-Unis (Jacewicz, Fox & Salmons, 2007). Cette vitesse moindre provoque donc la diphtongaison constatée dans « dawg ».

Les évolutions phonologiques touchent presque les mêmes diphtongues dans les deux corpus, alors que les changements sont inverses. Elles sont toutes centrifuges, donc pointent vers des glissements de classes d'isophones déjà repérés dans la partie précédente. Ces changements ont un lien fort avec la prosodie de chacune de ces langues. Le scots privilégie les voyelles courtes, et a moins de sons que l'anglais standard ; alors que l'AAV montre au contraire des allongements aussi stylistiques que prosodiques. Dans le corpus, il est frappant de remarquer à quel point le paradigme de domination

	SC	BP
Monophtongaison de /aʊ/	×	
Monophtongaison de /əʊ/	×	
Monophtongaison de /eɪ/	×	
Monophtongaison de /aɪ/	×	
Diphtongaison en /eɪ/		×
Diphtongaison en /aʊ/		×

Tab. 5.36 – Changements impliquant des diphtongues dans les deux corpus

de la langue écrite est renversé. La vision de l'orthographe en tant qu'entité supérieure est remise en cause par ces choix graphiques, qui semblent presque toujours témoigner d'un changement phonique.

Les réélaborations graphiques adoptées par les membres de communautés virtuelles a plusieurs points communs avec celui des corpus d'écrits fictionnels. La culture de l'écrit chez les membres de forums leur fait respecter le doublement de consonnes qui n'ajoutent rien à la prononciation (<Wullie>, <wumman>), et sauf à vouloir être expressif, le triplement ou le quadruplement de consonnes est évité. Écrire son parler, ou un certain parler [43] revient à harmoniser les différences entre l'orthographe de l'anglais et les prononciations ayant cours dans les langues minoritaires.

43. Il est impossible de savoir dans quelle proportion les utilisateurs de ces forums parlent comme ils écrivent, ou adoptent les codes de leurs pairs.

5.4 Tests de la spécificité des mots sélectionnés

La proportion d'utilisation de ces formes est souvent bien plus importante sur SC et BP que sur internet, même s'il s'agit de nombres d'occurrences très faibles sur les deux forums. Une interprétation statistique des fréquences d'apparitions, calculée en divisant le nombre d'occurrences de graphies alternatives par celui de la graphie standard révèle que ces caractéristiques militent effectivement en faveur de l'hypothèse de l'existence d'un communaulecte.

5.4.1 Méthode

J'ai utilisé toutes les représentations commentées dans ce chapitre dans les analyses quantitatives. Pour calculer la différence d'apparition entre les forums et internet, j'ai employé le test de Wilcoxon, pour calculer la moyenne des deux échantillons, et déterminer si la moyenne de chacun présente une grande différence. En l'occurrence, l'hypothèse défendue consiste à dire que, si les mots sont communaulectaux, la moyenne de ratio graphie alternative :graphie standard devrait être plus basse dans les corpus de forums. Le corollaire, l'hypothèse zéro, est donc que la moyenne de ratios graphie alternative :graphie standard est supérieure dans les corpus de internet.

Pour vérifier la normalité des données, c'est-à-dire leur distribution gaussienne, sans laquelle ces tests statistiques ne seraient pas fiables, j'ai utilisé le test de Shapiro-Wilk. Malgré la distribution gaussienne, le test du χ^2, pourtant designé pour traiter les petits corpus, ne donne pas de résultats probants.

La différence de taille entre les forums et internet d'une part, et la disparité des données malgré leur distribution gaussienne, lorsqu'une seule variable est prise en compte oriente le choix vers le test de rangs signés de Wilcoxon. Les données peuvent y être liées, mais comme il s'agit de mesures prises sur des corpus tout à fait différents, Wilcoxon est un meilleur choix [44].

44. Le test du χ^2 fonctionne tout de même sur les données, mais la valeur p flirte avec la validation de l'H_0. Par exemple, 0.1448 pour le test t des corrélations du scots est statistiquement signifiant pour un indice de confiance à 95%, mais de peu.

5.4.2 Résultats

Voici le résultat de la console de R, pour les deux variables (d'abord, les proportions sur internet, et les proportions sur le forum, pour les deux corpus) :

```
> shapiro.test(ScotCalc$ProportionSc)
Shapiro-Wilk normality test
data : ScotCalc$ProportionSc
W = 0.3903, p-value = 7.55e-12
> shapiro.test(ScotCalc$ProportionNet)
Shapiro-Wilk normality test
data : ScotCalc$ProportionNet
W = 0.2436, p-value = 2.984e-13
> shapiro.test(AAVCalc$ProportionBP)
Shapiro-Wilk normality test
data : AAVCalc$ProportionBP
W = 0.2376, p-value = 6.26e-14
> shapiro.test(AAVCalc$ProportionNet)
Shapiro-Wilk normality test
data : AAVCalc$ProportionNet
W = 0.1376, p-value = 8.431e-15
```

La valeur p est très inférieure à 0, ce qui signifie que les données suivent une distribution qui permet des exploitations statistiques. Pour lire les résultats de la valeur p, l'exposant est marqué <e> dans la console. Par exemple, le test de distribution de la proportion des utilisations sur SC renvoie à une valeur p de 7.55×10^{-12}. Ces valeurs sont dans une fourchette entre le pico et le femto, ce qui confirme bien la distribution normale de chaque échantillon, pris indépendamment.

Pour montrer qu'il existe bel et bien une corrélation entre les formes prélevées et le communaulecte, je vais donc regarder les différences de valeurs. L'hypothèse zéro est la suivante : il n'y a aucune corrélation entre les graphies alternatives et le lieu d'écriture, et la différence de moyenne entre les

deux variables est inférieure à 0. Les résultats du test de Wilcoxon sont présentés en annexes, en A.8. La valeur p du corpus SC est 3.349×10^{6}, et celle du corpus BP est 0.003558.

5.4.3 Discussion

Les valeurs p sont bien en dessous de l'indice de confiance [45], et les différences entre les proportions trouvées sur internet, et celles trouvées sur les forums sont statistiquement significatives.

Il y a 83 rangs positifs, dont les valeurs sont au-dessus de 0, dans le corpus SC, c'est-à-dire que les valeurs s'échelonnent sur 83 niveaux. La valeur p à 3.349×10^{-06} est bien plus basse que la valeur p pour le corpus AAV (valeur p : 0.003558), mais les deux résultats sont significatifs. Le nombre de rangs (résultat V, vérifiable en A.8) est bien plus élevé pour BP. C'est un corollaire de cette valeur p si supérieure : les amplitudes de rangs entre les résultats de BP sont bien plus vastes, ce qui explique ces chiffres.

Pour confirmer les résultats, j'ai utilisé les médianes [46] des écarts des rangs obtenus par chaque couple (graphie alternative et graphie standard) pour chaque corpus. Je les ai comparées à l'intervalle qui correspond à l'indice de confiance à 95%. Pour le corpus SC, la médiane devait se trouver entre $-\infty$ et -3 217.134 pour qu'elle confirme l'hypothèse défendue, et elle est effectivement dans la fourchette, à -6 349.581. De la même façon, la médiane pour le corpus BP devait se trouver entre $-\infty$ et -43.39062 pour être recevable, et elle l'est bel et bien, parce qu'elle est à -109.988.

L'intervalle de confiance à 95% permet de tirer des constatations plus larges, sur la population totale des graphies alternatives, en l'occurrence. Ces échantillons de populations sont fiables pour tirer la conclusion que ces graphies alternatives sont plus fréquentes sur les deux forums que sur internet, et que les utilisateurs ont développé leurs pratiques discursives propres à leur plateforme, sans qu'elles ne soient inédites sur internet : par exemple, il y a un ratio $1 : 2\,430$ pour les occurrences de <wummin> par rapport à <woman> sur SC contre un ratio de $1 : 61\,338$ sur internet.

Les résultats afro-américains tranchent moins avec les résultats trouvés sur internet, à cause des em-

45. Traditionnellement, en dessous de 0.05 quand l'indice de confiance est à 95%.
46. La médiane est la différence entre la valeur la plus haute et la plus basse d'un ensemble d'échantillons, et non la moyenne des résultats trouvés

prunts culturels divers : le netspeak s'inspire de la culture américaine, de ses graphies, et les emprunts graphiques à la culture afro-américaine (le suffixe <-a> qui remplace <-er> en tête) dilue la spécificité de ces usages. La musique, les titres de chansons et les noms d'artistes de hip-hop permettent une interface entre une certaine culture afro-américaine, et des calques graphiques subséquents. Par exemple, la graphie <nigga> est proportionnellement plus fréquente sur BP que sur internet, et son ratio d'apparition par rapport à <nigger> est $1 : 0,03$. Cette valeur est proche du ratio qu'on trouve sur internet, à $1 : 0,02$. La graphie <nigga> est beaucoup plus fréquente que <nigger> sur BP et internet, ce qui explique ces résultats.

5.5 Études de cas lexicaux

Les particularités phonologiques des deux dialectes expliquent la plupart des graphies alternatives. Il existe des phénomènes extérieurs qui sous-tendent la variation des phonèmes consonantiques et vocaliques. En revanche, lorsque l'on sort de la transcription de sa phonologie dans les graphies alternatives, pour se censurer ou utiliser des mots connus de tous dans une forme euphémisée, les graphies n'obéissent pas aux règles vues précédemment. C'est pourquoi je vais analyser des lexèmes dont les graphies contournatives visent à contre-carrer la censure et dont la variation allographique est forte. J'ai retenu la distance de Levenshtein pour mesurer ces écarts dans les productions des scripteurs. La distance de Levenshtein est adaptée à ce type d'analyses, parce qu'elle mesure l'écart à la forme graphique initiale. La pertinence symbolique de cette exploration devra faire l'objets de commentaires *a posteriori*. J'utiliserai une représentation en dendrogramme pour grouper et hiérarchiser les graphies en fonction de leur degré de ressemblance. Plus précisément, le calcul de la distance de Levenshtein, fournit l'écart pondéré entre deux formes graphiques en fonction des permutations, des éliminations ou des rajouts de lettres. Par exemple, <niger> a une distance de 1 avec <nigger>, et <nigga>, une distance de Levenshtein à 2.

D'abord, j'ai relevé le pourcentage de similarité entre les formes graphiques, et une distance de Levenshtein équilibrée, qui prend en compte les permutations de lettres dans la distance relative entre

deux formes du même mot. La combinaison des deux types de données (pourcentage de similarité et distance pondérée de Levenshtein) dressent un tableau exhaustif des phénomènes de permutation et de quelques jeux graphiques autour de formes iconiques, ou représentations de mots sur-représentés sur BP en l'occurrence, et porteurs d'un sens tabou, et dont les graphies alternatives n'ont que peu de rapport avec la phonologie. Les mots choisis pour les tests de variation sont tirés des lemmes les plus fréquemment utilisés dans les deux corpus. Pour BP, les lemmes [BITCH], [FUCK] et [NIGGER] ont été sélectionnés. Les graphies alternatives se concentrent sur certains mots particulièrement tabous, ou représentatifs pour une communauté donnée. Ils donnent surtout des indications discursives concernant les utilisations de voyelles tests et de consonnes postiches. Ces variations graphiques sont à la confluence de l'analyse de discours et du déplacement sémantique, et de la graphophonématique, afin de rester compréhensibles tout en ajoutant une couche sémantique aux mots. La différence entre <Amerikkka> et <America> est un bon exemple d'ajout sémantique, tout en s'appuyant sur l'homophonie entre la séquence <kkk> et <c>.

Ces analyses sont essentiellement celles des comportements graphiques des utilisateurs de BP, confrontés aux mots censurés sur le site, mais apparemment indispensables, puisqu'ils y sont tout de même écrits, avec plusieurs variantes graphiques par phonème. Cette partie tire les observations de Labov sur la logique de l'anglais non-standard, parlé par les locuteurs afro-américains vers des analyses des formes utilisées (LABOV, 1972). L'AAV a une logique interne intégrée par ses locuteurs, et que l'on peut reconnaître à partir d'indices grammaticaux et lexicaux, dont des formes particulièrement saillantes. Ces formes iconiques, reconnaissables et même attendues sont plus courantes chez les Afro-Américains que chez les Écossais, à cause de leur âge (les utilisateurs de BP sont plus jeunes, et utilisent plus de mots tabous, d'où l'étude de [FUCK] ; ils sont également aux prises avec le racisme, et les stratégies d'empowerment passent par l'utilisation de [NIGGER]. Le mot [BITCH] est ambivalent. Il oscille entre l'insulte, parfois hyperexpressive, avec l'allongement de la première voyelle ; et la valorisation d'une femme indépendante. Il n'existe pas de telles variations dans SC, parce que ces tensions y sont absentes.

5.5.1 Variations graphiques de *nigger*

Le mot *nigger* contient plusieurs éléments phonologiques intéressants. La prononciation du dernier phonème–voire des derniers phonèmes est le point principal de réélaboration graphique. La graphie standard de <nigger> correspond au découpage phonographématique présenté dans le tableau 5.37. Si

Phonème	Classe d'isophone	Description	Graphie standard
/n/	—	C. alvéolaire nasale	<n>
/ɪ/	KIT	V. pré-fermée antérieure non arrondie	<i>
/g/	—	C. vélaire plosive sonore	<gg>
/ɚ ~ɝ ~ər /	LETTER	voyelle centrale rhoticisée	<er>

TAB. 5.37 – Décomposition phonologique et graphique de nigger

l'on prend en compte toutes les graphies du lemme [NIGGER], qui n'est affecté que par le passage au pluriel, l'on obtient 23 graphies, qui se distinguent plus ou moins de la graphie standard du mot. La variation allographique de *nigger* est présentée dans le tableau 5.38, sans compter les deux variations qui contiennent des suffixes de marque de genre, <niggress> et <niggette>. Les fréquences d'apparitions sont également importantes, c'est pourquoi l'histogramme présenté dans l'illustration 5.11 classe les variantes en fonction de leur fréquence sur BP.

Les consonnes, surtout la consonne initiale, ont beaucoup moins de variantes alternatives que les voyelles. Ce mot permet des substitutions alphanumériques, et c'est le seul des trois cas d'étude lexicale. Le <n> ne ressemble à aucun chiffre, et la substitution n'est donc pas possible, contrairement à <i> (<1>) et <g> (<9>). Avec ces deux chiffres, la hauteur de casse n'est pas respectée, mais comme seule compte la ressemblance graphique, la bouma ou forme générale du mot chez les typographes et les cognitivistes qui s'intéressent à l'acquisition du langage est moins prioritaire que la censure. La perception ne change donc pas.

	/n/	/ɪ/	/g/	/ɚ ~ ɝ ~ ə/	/r/	/z/	
k	n	e ee i’ i i. l y !	cc g gg g.g. 99	a ah e a. u #	h r	z zz	s

Tab. 5.38 – Variation allographique de <nigger>

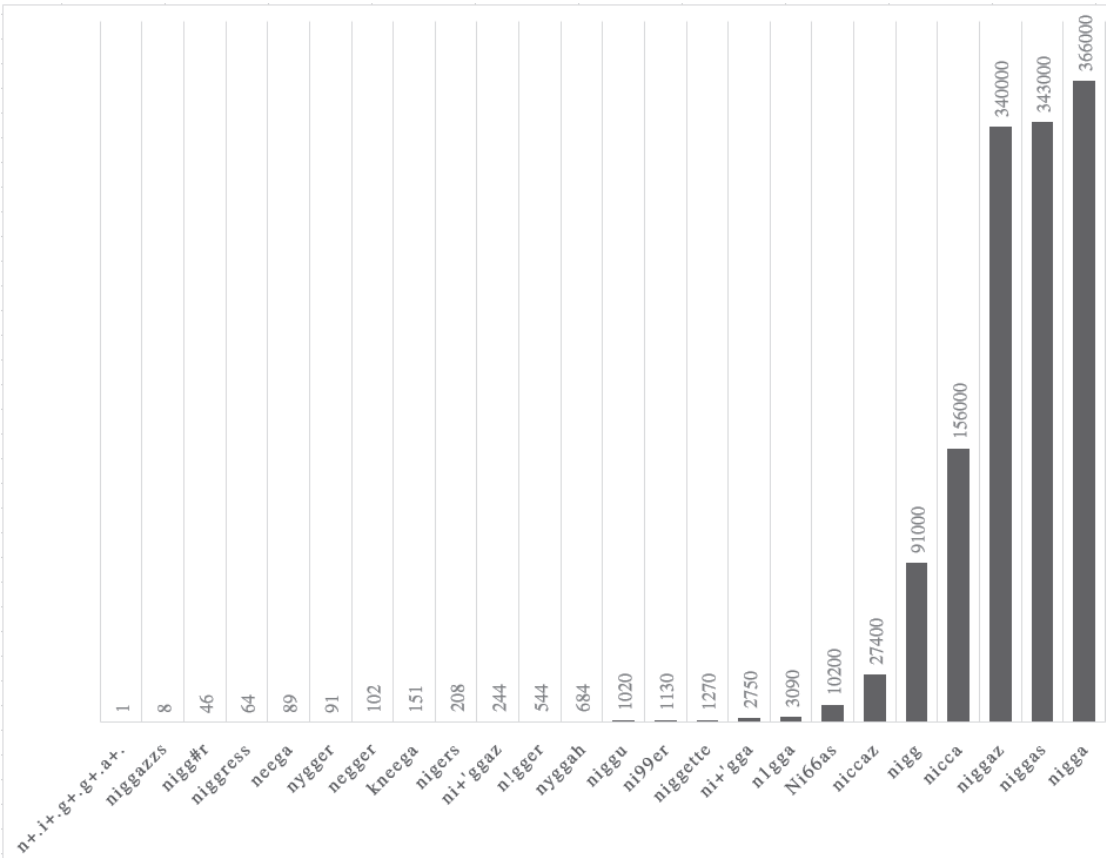


Illustration 5.11 – Classement des formes de <nigger> en fonction de leur fréquence sur BP

Les trois graphies les plus fréquentes sur BP sont <nigga> (366 000 occurrences), les formes au pluriel <niggas> et <niggaz> (343 000 et 340 000 occurrences, respectivement) et <nicca> (156 000 résultats). Cela assoit le statut de <nigga> comme lexème à part entière, tant il est fréquent : il est même plus fréquent sur internet que <nigger>, qui ne renvoie qu’à 14 400 000 résultats sur Google, tandis que

<nigga> compte 70 100 000 occurrences [47], soit cinq fois plus de résultats. Le glissement sémantique effectif, entre valorisation et insulte euphémisée semble confirmé par ces indications numériques. La graphie alternative est aussi exploitée dans le pluriel, où le morphème <-s> est en concurrence avec <-z> : le voisement est injecté dans la graphie, même si la forme est parfois tautologique, comme la graphie <niggazzs> le montre. Cette forme est très marginale (seulement 8 occurrences dans le corpus), et les utilisateurs préfèrent utiliser soit un <z> unique, soit un <s> orthographique.

Reste la forme <nicca>, qui reste très proche phonologiquement de la matrice « nigga ». Visuellement, le <c> est doublé (comme le <g> du mot original) pour conserver la longueur du mot. En outre, l'arrondi du sommet de la lettre ressemble à la partie supérieure de <g>, et, à phonie égale, il est donc préféré à <k>. La séquence <nicka> n'aurait pas pu convenir, parce que ses graphèmes s'éloignent trop de la bouma de <nigga>, qui contient le même nombre de lettres mais qui n'a pas du tout la même forme générale. Le troisième phonème est soit représenté par une lettre unique (qui est alors uniquement <g>), soit par un graphème doublé. Malgré le nombre de formes, elles suivent des règles systémiques et logiques. La substitution n'est pas qu'une question de valeurs phonologique, et la forme et la taille générales du mot interviennent dans la sélection de candidats de remplacements graphiques.

Les voyelles ne s'éloignent jamais des sons attendus : soit le /ɪ/ est transcrit par un <i> orthographique, soit par le graphème le plus proche phonologiquement, <CyC> (ce qui sera aussi le cas pour <bitch>/<bytch>), soit par son équivalent phonologique allongé, <ee>, qui représente /iː/. La lettre <e> seule dans <negger> évoque le glissement de DRESS vers KIT, et n'est utilisée que dans cette graphie (seulement 109 occurrences dans le corpus BP).

Des caractères extra-alphabétiques sont aussi convoqués, à cause de leur ressemblance avec les lettres qu'ils remplacent. Le point d'exclamation ressemble à un <i>, grâce à son point, et la symétrie axiale horizontale permet de retrouver la forme initiale de <i>. Le corpus BP contient des substitutions alphanumériques qui ressemblent au leetspeak. Le remplacement de <i> par <1> est typique de ce jeu linguistique en ligne. Dans le cas de < !> et de <1>, seule compte la forme verticale et la relative fréquence des caractères extra-alphabétiquues qui servent la substitution. Le caractère <|>, présent sur

47. Résultats obtenus le 13 mars 2013.

le clavier QWERTY est placé tout à droite, sur la première ligne, après deux touches de caractères non-alphabétiques, et, pour l'obtenir, il faut presser sur la touche SHIFT [48]. À l'inverse, produire <1> est très facile, puisque c'est la première touche au dessus de <Q> et que c'est la valeur par défaut de la touche. Le point d'exclamation est sur la même touche, et en est la valeur secondaire [49].

Le <i> orthographique est la lettre qui subit la plus grande variation allographique, mais ces variations restent circonscrites autour d'un espace phonologique très réduit, dans un triangle dont /i/, /ɪ/ et /e/ sont les sommets. Les représentations de /ɚ ~ ɝ ~ ə/ sont plus variées.

Dans la majorité des cas, soit 13 graphies sur 23, la dernière voyelle est représentée par <a>. Cette voyelle sert à resémanticiser le mot, c'est pourquoi elle est si représentée, par rapport à <er>, contenu dans seulement quatre graphies. En tant que voyelle non accentuée, elle est aussi représentée par <u>, qui tire le /ə/ vers /ʌ/. La lettre <e> n'est jamais présente sans un <r> qui la suit, et ne sert donc pas pour représenter le schwa sans coloration par /r/.

À titre expérimental, je propose une analyse automatique de classification des variantes allographiques sous forme de dendrogramme, afin d'évaluer les familles selon leur degré d'étrangeté ou de familiarité à la forme matrice. Pour lire un dendrogramme, ou regroupement hiérarchique phylogénétique, il faut savoir que la hauteur relative des branches et de leurs recoupements est en fonction de la différence de Levenshtein de chaque forme. Les dendrogrammes sont découpés en groupements de mots, et chaque branche (ou clade) qui arrive dans le tiers supérieur du graphique est une famille, notée A, B, C…Les sous-familles, ou regroupements inférieurs seront notés A^{α}, A^{β} …. La hauteur à laquelle les branches se rejoignent est primordiale pour comprendre la logique derrière la cladistique, ou groupement spatial en fonction de caractéristiques identiques. La forme la plus différente des autres les rejoindra donc le plus haut dans l'arbre ; logiquement, les graphies standard seront vers 0, au milieu d'une sous-branche, en tant qu'inspiration graphique des autres variations.

Le graphique 5.12 présente les distances entre les variantes graphiques de <nigger>. Ces distances prennent en compte toutes les distances de Levenshtein de toutes les formes les unes par rapport aux

48. Sur le clavier AZERTY, taper <|> reviendrait à taper <£>. En outre, il y a une touche non-alphabétique de plus sur la première rangée sur le clavier QWERTY.
49. L'équivalent sur un clavier AZERTY est <&> et <1>.

autres [50].

Le dendrogramme regroupe les graphies en trois familles distinctes ; plus les formes sont proches

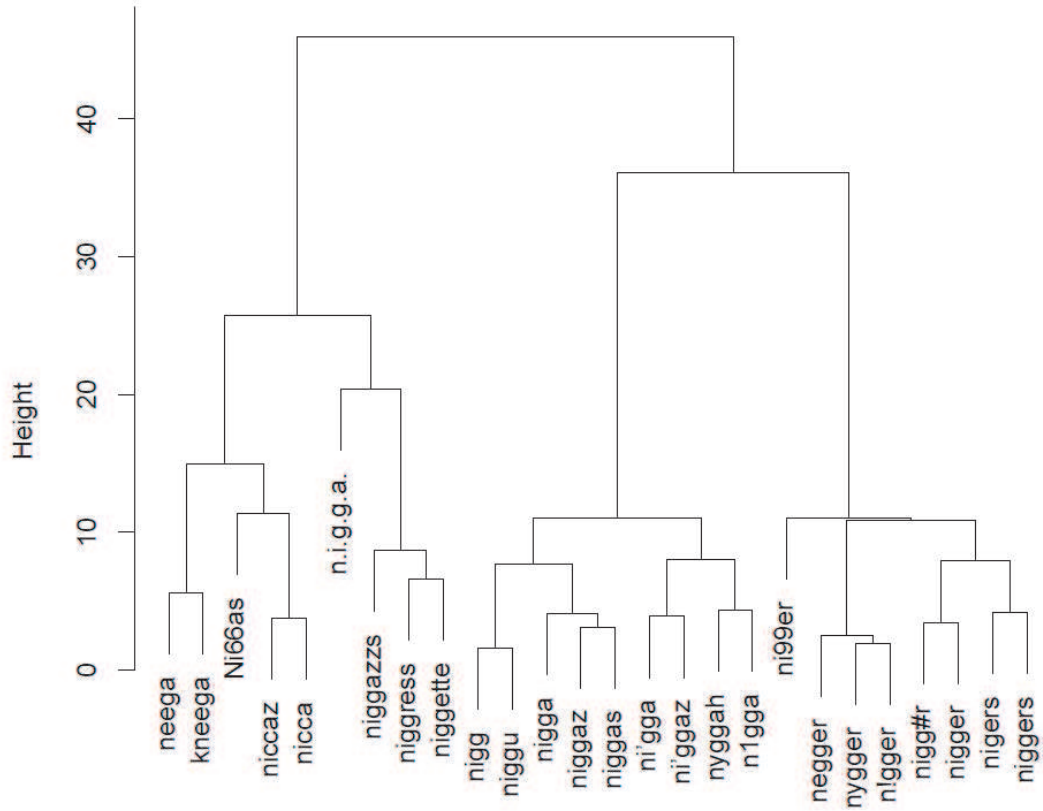


Illustration 5.12 – Dendrogramme des distances entre les graphies de <nigger>

et plus bas sera leur noeud. La clade A, de <neega> à <niggette>, regroupe tous les éléments les plus éloignés de la forme <nigger>, où les <g> sont remplacés par un <c>, et les deux variantes qui portent un morphème de marque de genre (<niggette> et <nigress>). Dans cette famille se trouve également la forme <n.i.g.g.a.>. On pourrait s'attendre intuitivement à ce que cette forme se rapproche de <nigga>, à cause de la lecture humaine qui traite les points comme un élément facultatif, qui ne gêne pas la compréhension. Pour le traitement automatique, cette forme est en fait très éloignée, puisqu'elle est graphiquement la plus longue des variantes, avec 10 caractères. Cette différence quantitative explique son rattachement tardif à la feuille qui contient <niggazzs>, <niggette> et <nigress>, autres variantes

50. Le jeu de données est une matrice de recoupements des distances de Levenshtein de chaque variante graphique. Il est décrit dans les annexes en A.7.1 et figure sur la clef USB d'annexes.

de sept à huit caractères. Même la graphie <ni66as>, pourtant écrite avec des caractères alphanumériques, rejoint <nicca(s)> plus vite dans la hiérarchie. Il est à la même hauteur que <ni99er>, qui appartient à la famille C.

La famille B et la famille C se rejoignent lorsque leur seule différence réside dans la terminaison <-a> ou <-er>. La clade B regroupe aussi les variantes <nigg> et niggu>, parce qu'elle regroupe les variantes qui commencent par <nigg>, avec une tolérance à l'introduction de caractère comme <'>, ou la substitution de <i> par <y>. La clade B est donc centrale parce que c'est celle qui regroupe les éléments retrouvés le plus souvent dans le corpus. En termes quantitatifs, les trois graphies les plus courantes (<nigga>, <niggaz> et <niggas>) sont dans cette clade. Les trois formes sont très proches graphiquement, ce qui suggère que plus la variation est faible par rapport au standard et plus elle va être employée.

La clade C regroupe les mots dont les graphèmes finaux sont le plus souvent <er>. Si cette famille a autant de membres que la famille B, ce sont aussi les variantes graphiques les moins populaires sur BP. La constatation la plus étonnante est de voir que la forme <nigger> n'est en fait pas au milieu de la branche, et qu'elle ne forme pas un couple avec <niggers>. On touche à la limite du calcul automatique des distances de Levenshtein, qui rapprochent plus volontiers <nigg#r> à <nigger> que ce dernier à son pluriel, parce que le nombre de lettres est identique, et qu'il n'y a qu'une substitution graphique.

Les regroupements phylogénétiques des distances de Levenshtein ne regroupent pas du plus lisible au moins lisible, parce que toutes les formes sont compréhensibles. En revanche, l'éloignement relatif de la branche A par rapport aux branches B et C est révélatrice de ses formes plus inédites, par rapport aux familles qui se regroupent autour de leur suffixe. Les formes qui sont les plus éloignées du centre des clades sont toujours les formes les plus marginales, par rapport aux centres qui partagent plus de caractéristiques avec leurs voisins. Ces formes centrales sont aussi les plus courantes dans le corpus. L'effet problématique de ce traitement automatique est le regroupement sur le critère de la longueur : les formes féminines <niggette> et <nigress> se rapprochent artificiellement de <niggazzs>, par exemple. Le pluriel n'est pas traité de manière satisfaisante sémantiquement.

La partie suivante va traiter d'un mot qui est sémantiquement féminin, « bitch » et ses variations

allographiques. Sa forme n'est jamais masculinisée, ce qui assure une plus grande stabilité des formes. Les pluriels sont plus fréquents que pour « nigger », ce qui permet de développer les réflexions faites dans cette section.

5.5.2 Variations graphiques de *bitch*

La décomposition phonologique de [BITCH]$_N$ a les mêmes caractéristiques, au moins au niveau consonantique, que *nigger*, mais ce mot est monosyllabique dans sa forme standard singulière.

Les formes de <bitch> sont réparties suivant leur fréquence dans l'histogramme 5.13, et les va-

Phonème	Classe d'isophone	Description	Graphie standard
/b/	—	C. bilabiale plosive	<b>
/ɪ/	KIT	V. pré-fermée antérieure mi-brève	<i>
/tʃ/	—	C. postalvéolaire affriquée	<tch>

TAB. 5.39 – Décomposition phonologique et graphique de bitch

riantes graphiques elles-mêmes sont présentées dans le tableau 5.40. Les formes les plus fréquentes sont un peu trompeuses : la graphie <b^ches> représente toutes les graphies contenant un caractère non alphanumérique (<b.tches>, <b'tches>, <b tches>…), qui sont impossibles à rechercher sur *Google*. Le moteur de recherche ne prend pas en compte ces caractères individuellement, même en ajoutant des guillemets aux termes de recherche, et les traite tous comme <b[caractère non-alphanumérique]tch>. C'est une des limites de la technologie au service de la linguistique de corpus. Aucun serveur personnel[51] n'aurait la puissance de calcul pour traiter plusieurs centaines de millions de pages d'un site, en ne prenant que BP. Quant à un tel traitement géré par des expressions régulières sur l'entièreté d'internet sur ce type de requête, c'est absolument impossible, même avec un parc personnel d'ordinateurs très puissants. Les coûts en temps, en énergie et en puissance dont il faudrait disposer rendent l'indexation impossible à moins d'utiliser un moteur de recherche, et donc de se limiter à des technologies

51. Une configuration d'ordinateur personnel à la maison contient au maximum 36Go de mémoire vive, ce qui permet des calculs parallèles…Mais à quelle vitesse pour ces volumes de requêtes ?

offertes au grand public.

Les formes les plus fréquemment rencontrées dans le corpus sont <bytch>, avec 124 000 occurrences sur le forum (suivi par son pluriel <bytches>, qui compte 45 400 occurrences), puis 28 400 <b[]tch> et 14 800 résultats pour <biatch>. Les graphies suivantes sont nettement moins représentées, avec 6 190 occurrences de <bish>.

Le mot <bitch> a moins de variantes allographiques (seulement 12 formes étudiées) que les deux autres cas traités dans cette section, et bien qu'il soit plus long que <fuck>. Ses variations se regroupent surtout autour de la voyelle /ɪ/ et de l'affrication de /ʧ/. Tout se concentre donc autour de la substitution la plus évidente, entre <i> et le caractère <y> : phonologiquement, ils ont la même valeur en syllabe fermée (par exemple, « myth », /mɪθ/). En outre, la lettre <y> est plus rare : sa fréquence n'est que d'1.974% d'apparition en anglais, contre 6.966% pour <i> [52]. L'exotisme de la lettre <y> et ses propriétés phonologiques en font une candidate toute désignée pour remplacer le <i>, et préserver la valeur phonographématique du mot en voyelle fermée accentuée.

Encore une fois, le point d'exclamation est utilisé pour remplacer <i>, à cause de leur ressemblance visuelle. Les caractères non-alphabétiques sont quasiment tous utilisés pour remplacer la voyelle, et ainsi préserver la structure consonantique du mot, à l'exception de la parenthèse ouvrante <(>. Cette dernière remplace le <c>, grâce à leur ressemblance, mais ce choix n'est pas très populaire, et <b !t(hez> est relégué à la dernière place des variations allographiques.

Les substitutions peuvent s'éloigner de la phonologie standard du mot « bitch », en utilisant les mêmes stratégies que pour « nigger », comme l'allongement vocalique de /ɪ/ en /iː/, symbolisé par la graphie <ee>. L'allongement de la voyelle résulte en une diphtongaison de /ɪə/, symbolisée par <eo> et <eyo>. C'est le même phénomène que celui de la graphie <biatch> (qui totalise 14 800 occurrences sur BP). L'allongement de la voyelle correspond à un désir d'expressivité de la part des utilisateurs : « Biatch is another way to say the word bitch. The small tweak in pronunciation gives the word emphasis or an air of humor depending on situation. » (UD, entrée 6 [53]). Ces graphies transcrivant des diphtongaisons sont souvent utilisées dans des contextes d'insulte : « Yall know if any beyotch mess

52. L'histogramme des fréquences d'apparition des lettres en anglais se trouve dans les annexes, dans la partie A.5. Il repose sur des travaux de cryptographie (Lewand, 2000).

53. L'adresse de la page est <http ://www.urbandictionary.com/define.php ?term=biatch>. Dernier accès le 20 juin 2014.

around thats not going down » (BPPP26) ; « BROTHAZ :DO YOU BEAT YO BEOTCH UP, WOULD YOU SLAP THE SHEOT OUT OF A BEOTCH ?[54] » [sic] (BPNA5), ou de valorisation, comme le pseudonyme DA-BADD3ST-BEOTCH [sic] (BPPP27) le montre.

Le statut de /ʧ/ est décomposé par trois variantes allographiques : en utilisant seulement <sh> ou <ch>, sans le <t>, les utilisateurs rendent à /t/ sa valeur phonologique pleine. La séquence <tsh> est très rare, et la graphie <bishes> a déjà été traitée dans la section 5.2.3. Sa forme au singulier, <bish>, est beaucoup plus courante (6 190 occurrences contre seulement 1 130), mais la chute de l'affrication passe après la simple substitution de <i> en <y> et la diphtongaison en <biatch>.

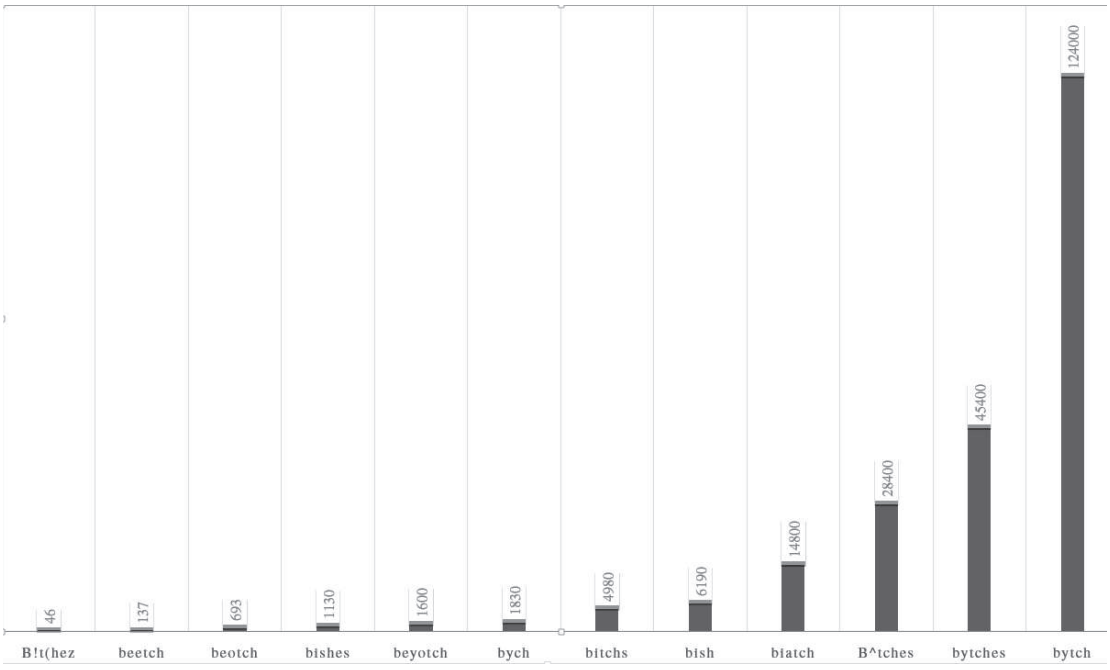


Illustration 5.13 – Classement des formes de <bitch> en fonction de leur fréquence sur BP

54. Le nom surprenant de cet article provient des paroles d'une chanson de Suga Free, « I'll never hit a woman/ but I'll slap the shit out a bitch. Why you bullshyttin' ? » (Suga Free, 2011)

/b/	/ɪ/	/t/	/ʃ/	/ɪz/
b	i ! ^ y ee ia eo eyo	t	ch sh (h	es s

Tab. 5.40 – Variation allographique de <bitch>

Quant au dendrogramme des variations graphiques de <bitch>, il présente des disparités moins marquées que celui de <nigger>. Les variations sont regroupées en deux clades principales, la famille A qui va de <beyotch> à <biatch>, et la famille B qui contient <B !t(hez> jusqu'à <bish>. La feuille A^{α} contient trois éléments dont le point commun est de changer la prononciation de /ɪ/, soit avec un allongement de la voyelle, soit avec une diphtongue. L'algorithme a également bien éloigné les deux graphies <bishes> et <bish> des autres (de <B !t(hez> à <bitches>), dans la feuille B^{γ}.

Ces graphies s'éloignent phonologiquement de <bitch>, qui est groupé dans la même feuille que <biatch>. La distance entre la graphie standard et son pluriel, qui ne se trouve même pas dans la même clade, s'explique par la différence de nombre de lettres. <Bitches> est donc plus proche de <bytches>. L'algorithme attribue ce degré de proximité important, alors que l'interprétation phonétique, ou morphologique en divergerait. Phonologiquement, <biatch> devrait se rapprocher de la feuille A^{α} (qui contient entre autres les formes <beyotch>, <beotch>) à cause de la diphtongaison de la voyelle. La graphie vocalique <ia> est donc la plus efficace pour changer le son du mot, à la lecture humaine, sans trop l'altérer. Le rejet de <bitches> dans la clade B, alors que celle-ci contient les formes les plus inédites est logique, si l'on prend en compte le fait que de nombreuses graphies s'éloignent du standard, et remplacent <i> (9 graphies, dont <biatch>) ; la digraphie du morphème de pluriel repousse encore plus cette variation.

Dans ce dendrogramme, les formes <bitch> et <bytch> sont centrées, dans la feuille A^{β}. Contrairement aux graphies de <nigger>, qui étaient surtout des dérivés de <nigga> et où la graphie standard était rejetée à la marge, les variantes s'organisent réellement autour de <bitch>. La forme la plus em-

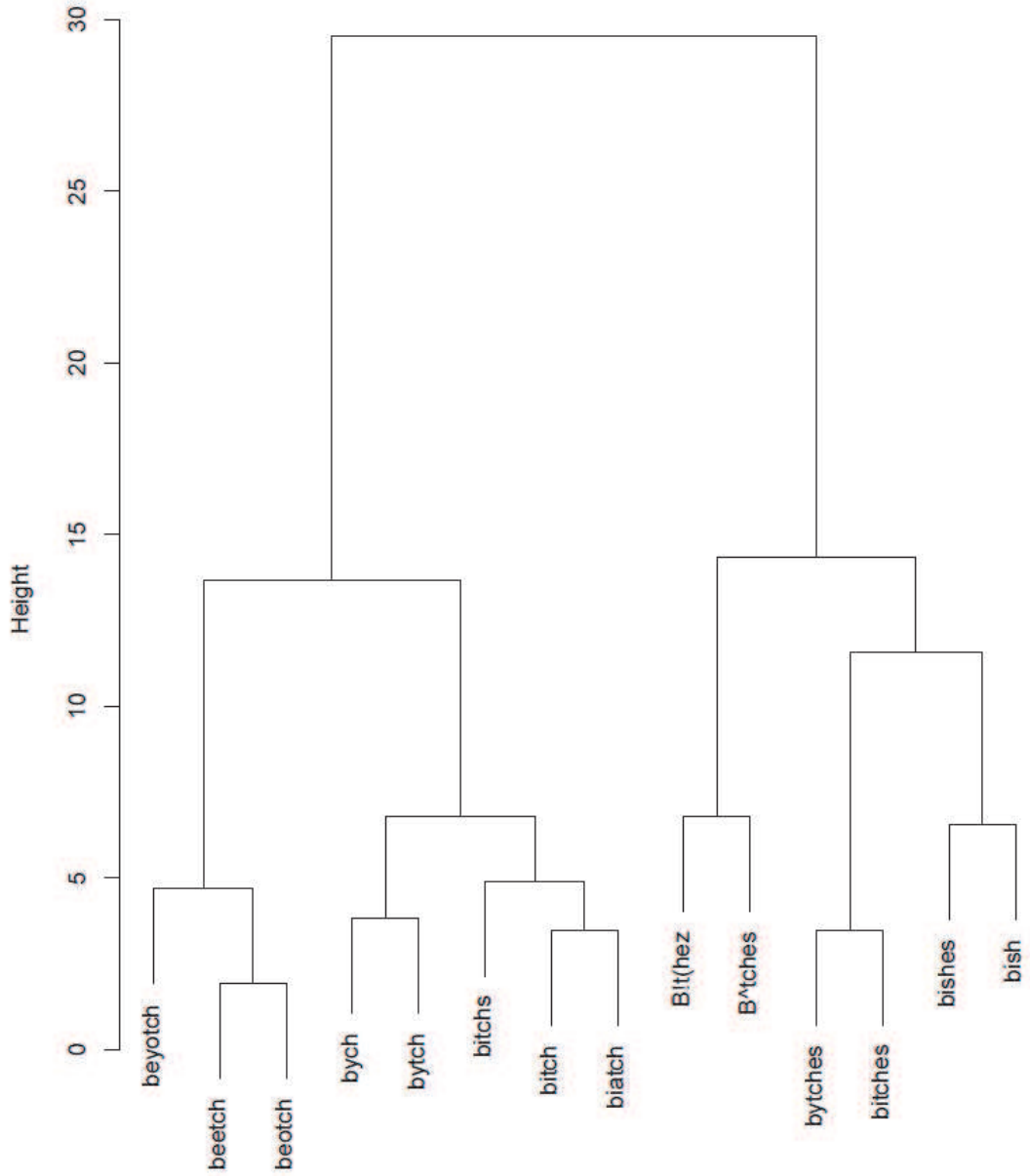


Illustration 5.14 – Dendrogramme des distances entre les graphies de <bitch>

ployée, <bytch>, est également centrale, mais elle ressemble à peu d'autres formes. C'est certainement la raison pour laquelle elle est préférée dans le corpus, et sa proximité avec le mot matrice explique pourquoi elle est centrale dans le dendrogramme. Je vais finir cette analyse des variations allographiques des mots iconiques avec l'étude de <fuck> et ses variantes simples.

5.5.3 Variations graphiques de *fuck*

Le mot *fuck* n'a pas de phonème en commun avec les deux autres mots étudiés jusqu'ici, et n'a pas la même catégorie grammaticale que les deux noms communs vus jusqu'ici, donc les variations graphiques peuvent changer : on peut imaginer qu'il n'y aurait pas de substitution de la voyelle <u> par les mêmes lettres, par exemple. La description phonologique formelle de l'infinitif est présentée dans le tableau 5.41.

Phonème	Classe d'isophone	Description	Graphie standard
/f/	—	C. labiodentale fricative	<f>
/ʌ/	V. mi-ouverte postérieure non arrondie		<u>
/k/	—	C. vélaire plosive sourde	<ck>

Tab. 5.41 – Décomposition phonologique et graphique de fuck en AAV

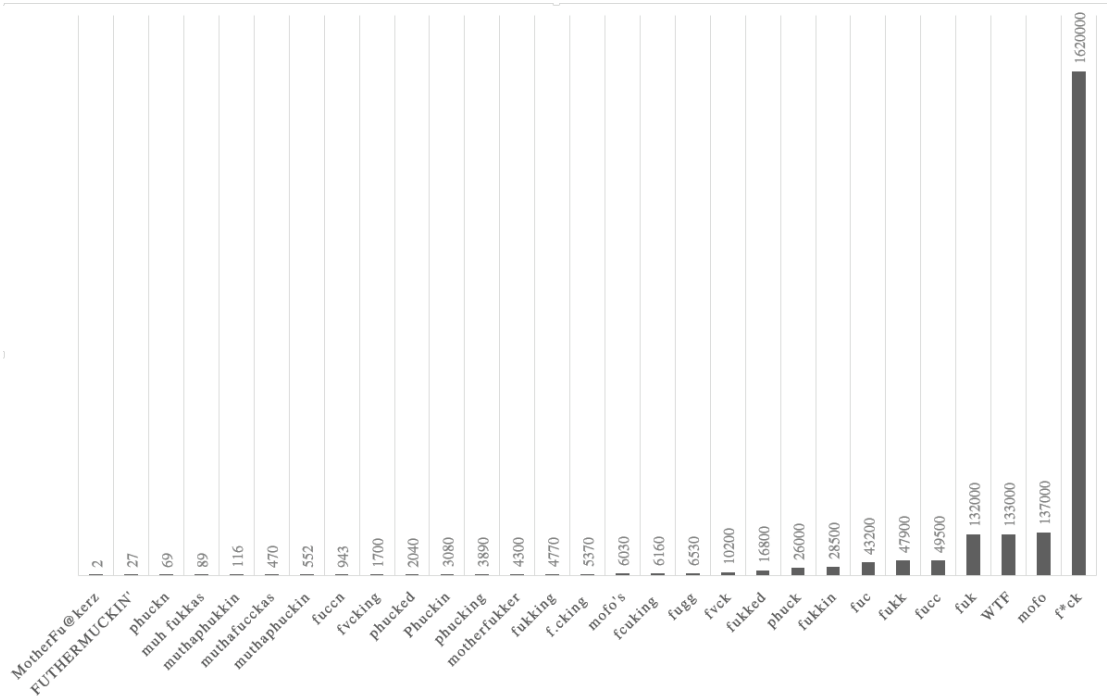


Illustration 5.15 – Classement de toutes les formes de <fuck> en fonction de leur fréquence sur BP

L'étude de [FUCK]$_V$ s'avère plus difficile, à cause de son utilisation adjectivale, verbale, ou nominale, et de son composé principal, *motherfucker*. Les distances de Levenshtein des variantes du lemme sont présentées dans le dendrogramme 5.20. L'histogramme 5.15 présente à la fois <fuck> et le sous-ensemble des *motherfucker*. L'analyse traitera surtout de *fuck*, sans l'acronyme « WTF » (‘*What The Fuck*’). Il est typique de la culture internet, et de l'argot militaire [55], mais ne fait pas partie à proprement parler, des composés ou des dérivés de *fuck*. L'étude se focalisera donc sur les formes présentées dans l'histogramme 5.16, qui reprend les informations du premier histogramme.

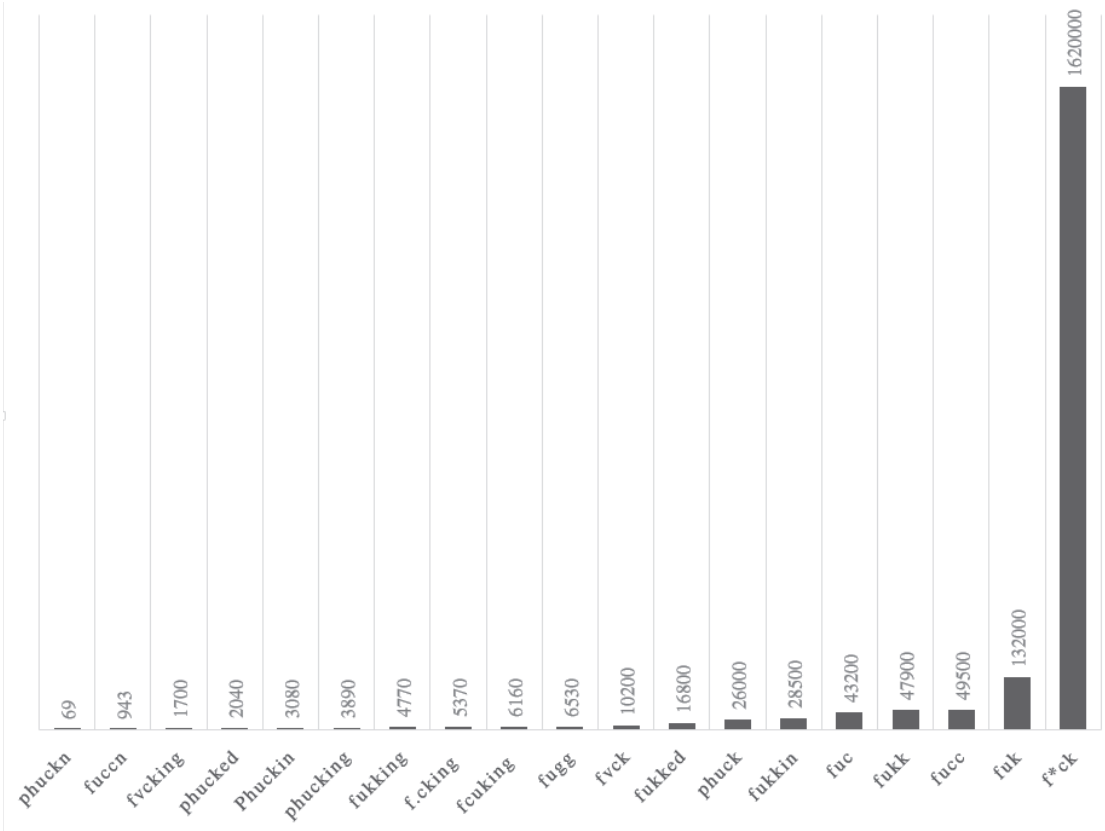


Illustration 5.16 – Classement des formes de <fuck> étudiées, en fonction de leur fréquence sur BP

Les formes les plus fréquentes du mot sont <f[caractère non-alphabétique]ck>, avec 1 620 000 occurrences, qui surclasse les autres variantes graphiques. La seconde plus fréquente est <fuk>, avec

55. Sous la forme « Whiskey Tango Foxtrot », utilisé notamment par un officier du Pentagone (LIBERMAN, 2009). Le blog Language Log avait déjà couvert les sigles associés à « fuck » (ZIMMER, 2009).

132 000 occurrences, et la troisième, <fucc>, avec 49 500 occurrences. Pour ce mot, les changements semblent plus souvent concentrés autour des consonnes que de la voyelle. Ce n'est pas parce que le mot est monosyllabique, sinon, <bitch> aurait subi les mêmes transformations.

/f/	/ʌ/	/k/	/d/, /ŋ/
f	u	ck k c	ed
ph	v . *	cc kk gg	n in ing
fc			

Tab. 5.42 – Variation allographique de <fuck>

Les utilisateurs censurent plus souvent la voyelle, et, par manque de voyelle de substitution ou d'un caractère extra-alphabétique qui lui ressemblerait, choisissent un caractère non-alphanumérique pour remplacer la voyelle <u>. Il n'y a pas de conflit de compréhension. Les mots « fack » (‘*l'un des quatre estomacs des ruminants*’) et « feck » (‘*voler ou jeter*’, argot irlandais) sont très rares et plutôt cantonnés à l'anglais britannique. Les mots **fick* et **fock* n'existent pas en anglais. Un mot de quatre lettres à la structure <fVck> est donc forcément « fuck ».

Les changements qui portent sur le son /k/ interviennent surtout quand la séquence <ck> est en position finale, sauf pour les graphies <fuccn> et <fukkin>. La forme <fuccn> peut être reliée au discours de gang : le livre d'or d'une utilisatrice contient des attaques envers les Bloods (illustration 5.17 donne une capture d'écran du CEC). Elle ne dit pas directement appartenir à ce gang, mais on peut inférer son son appartenance par des indices sémiotiques (utilisation de la couleur rouge dans son profil, pourtant vide), ou par ses amis sur BP, en particulier un homme dont le pseudonyme contient 5Star[56]. Sa page est donc sujette à une attaque d'un Crip, que j'appellerais Gangsta, dont les deux messages sont cités et traduits en anglais :

> [Blood Killer] all day kill a slob [= un Blood] bitch !! C[rips] [up] B[loods] down 6 [= l'étoile à six branches] stacking 5 [=l'étoile à cinq branches] cracking ; Kill a slob when I see a slob fuck all [of] y'all.

56. L'étoile à cinq branches est le symbole des Bloods, celle à 6 branches, celle des Crips.

> Blood killer all day bitch respect real G[angsters] you ain't shit I'm [an Original Gangster[57]] I get respected ! Whores like you still try to get fucked and you ain't official you ain't no slob cuzz [=cousin] ! I would blow your fucking head off you and your homies [=friends] ; and remember crips run all, Bloods (Crabs[58]) We them [= we are those] niggas fucking smut. C[rips] up ; bitch Blood cab slob on a dick. (BPGB5)

Il ne convient pas ici de commenter les formes à part les utilisations de <fucc>, qui est la seule graphie acceptable pour un Crip, qui n'a pas le droit d'écrire <ck> : cette séquence graphique représente « Crip Killer ». Le scripteur fait usage de <bk> à plusieurs reprises dans son message, ce qui veut dire qu'il connaît ce standard. Son discours est fortement influencé par son gang, et les formes <fucc> et <fuccn> sont régies par le tabou graphique des Crips.

Ce n'est pas une généralisation cependant, parce que certains profils contiennent aussi cette forme

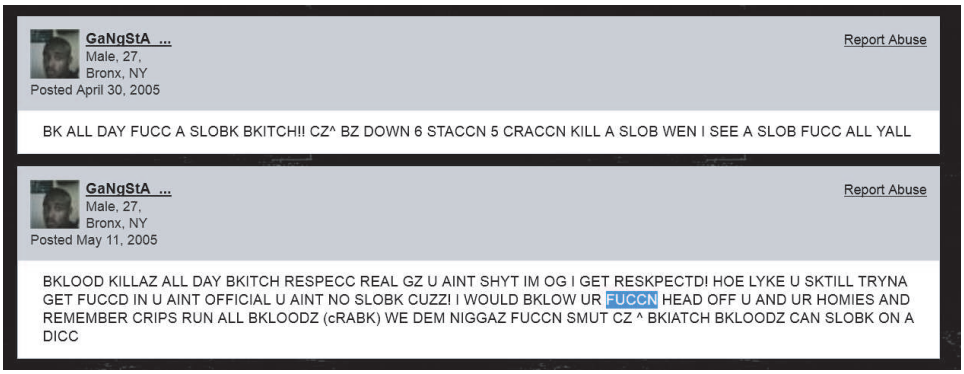


Illustration 5.17 – Extrait de livre d'or d'un profil contenant <fucc> et <fuccn> (BPGB5)

sans pour autant que leur auteur ne soit affilié à un gang. La capture d'un extrait de profil (illustration 5.18) le prouve–à moitié. L'extrait est montré en contexte alors que la qualité de l'image ne permet pas de déchiffrer l'ensemble des codes sémiotiques mobilisés. Cette utilisatrice mêle des codes que l'on pourrait assimiler aux Crips, comme le texte en couleur bleue, à des éléments que l'on pourrait assimiler aux Bloods. La photographie de fond, montre un jeune homme en rouge, ce qui est incompatible avec l'imagerie Crip. Pourtant, l'utilisatrice utilise également <bkullshit> et <abkout>, reprise de « Blood Killer ». C'est un paradoxe que rien n'explique sur son profil, d'autant plus qu'une autre photographie la représente également avec un t-shirt rouge. La seule interprétation plausible est un emprunt linguistique à la culture du gang, qui reste partiel et qui ne fait pas partie de l'identité réelle

57. Un nom désignant les membres les plus hauts dans la hiérarchie du gang.
58. Récupération du terme censé insulter les Crips par un Crip.

de l'utilisatrice. Les figurations graphique du gang ont déjà fait l'objet d'une étude en 4.4.5.1. La séquence graphique <bk> ne repose que sur l'insertion d'une valeur sémiologique supplémentaire à <b>, parce que cette lettre représente le nom des ennemis (pour les Crips). De même, la séquence <cc> de <fuccn> ne semble motivée que par la mise en avant de la lettre <c>, l'initiale de « Crip ».

Une autre influence culturelle sur les graphies de « fuck » est la marque de vêtement French Connec-

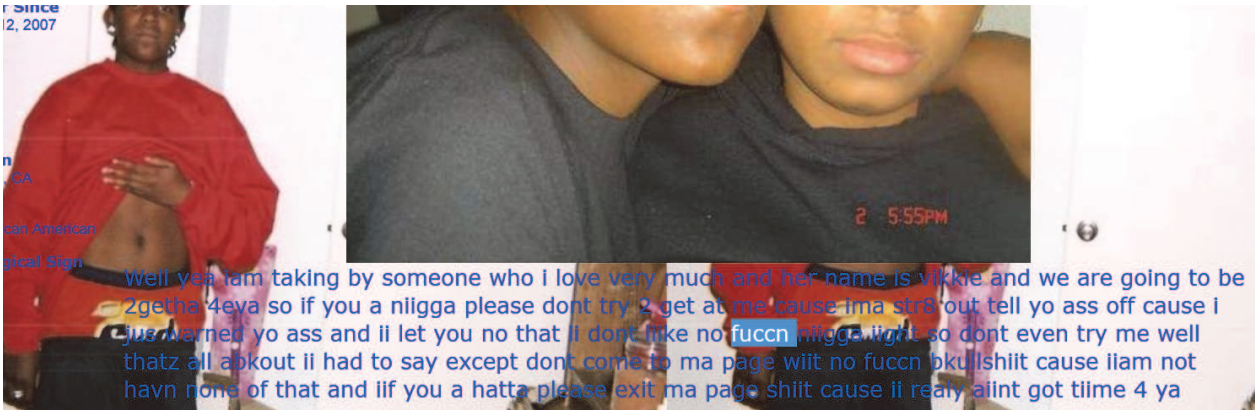


Illustration 5.18 – Extrait de profil contenant <fuccn> (BPPP28)

tion–UK, qui a basé ses campagnes de publicité sur l'acronyme « FCUK », comme le montre l'illustration 5.19. Cette marque a influencé une des graphies alternatives étudiées dans ce corpus, <fcuk> et cette graphie n'est pas exclusive à BP. La métathèse n'est donc pas une faute de frappe, ou une néographie idiolectale.

Les changements graphiques qui suggèrent un changement phonologique sont plus rares : le di-

Illustration 5.19 – Publicité de la marque FCUK (1997)

graphe final <ck> est parfois représenté par <gg> (dans le mot <fugg>, en dixième position du classement par fréquence). Le changement évoque un changement en [+sonore] qui n'est pas attesté à l'oral, mais il assoit surtout la reconnaissance phonologique de /g/ en tant que contrepartie de /k/. En revanche, la graphie <*fug> n'est pas attestée, peut-être à cause de l'éloignement trop grand avec le mot

initial : ne pas conserver le nombre de caractères et sonoriser la dernière consonne le ferait trop ressembler à d'autres mots de trois lettres, comme <pug> ou <bug>. Quand les variantes allographiques de *fuck* ne comportent que trois caractères, la consonne finale /k/ est soit représentée par <c>, soit par <k>, et ces deux stratégies sont préférées à la sonorisation de la dernière consonne (<fuk> est en seconde position, et <fuc>, en cinquième).

La consonne initiale ne devient jamais une consonne sonore dans les variations allographiques : le <f>, qui correspond au son /f/, n'est remplacé que par <ph>, et ce remplacement est moins populaire que celui de la voyelle, ou de la consonne finale. Aucune autre substitution n'intervient à l'initiale, et aucun utilisateur ne double la première lettre, en écrivant <*ffuck>, par exemple.

L'illustration 5.20 est une représentation spatiale des distances de Levenshtein, qui groupe les formes graphiques en trois familles, suivant leur proximité graphique. L'échelle de séparation est différente des dendrogrammes de *nigger* et *bitch* à cause de l'unité utilisée, qui est un pourcentage de similarité de Levenshtein, et non les données pondérées, mais la différence n'affecte que cet élément et pas les groupements.

La clade A regroupe toutes les formes où <f> a été remplacé par <ph>. Excepté la forme <phuck>, les membres de cette famille ne sont pas très populaires dans le corpus, et les formes dérivées ressemblent d'autant moins à la graphie standard. C'est pourquoi elles sont rejetées dans la feuille A^{α}, soit à gauche du graphique. Les formes dérivées occupent exclusivement la clade C, et ce sont des formes marginales, comme celles qui remplacent <ck> par <kk> (comme la feuille C^{β}) ; ou celles qui altèrent la voyelle (feuille C^{α}).

La clade centrale B contient la forme <fuck>, dans la feuille B^{β}, ainsi que les formes populaires comme <f*ck> et <fuk>. Cette famille contient les formes les plus courtes, donc les plus proches de la forme prototypique du mot. La famille B contient donc toujours la forme orthographique, et la distribution des autres formes autour de cette graphie montre qu'elles s'en inspirent toutes, ou plutôt, qu'elle a forcément un point commun avec les graphies alternatives.

Les dendrogrammes constitués ne montrent pas que la distance entre les variantes allographiques et la forme standard, mais la distance entre toutes les formes les unes par rapport aux autres. Malgré

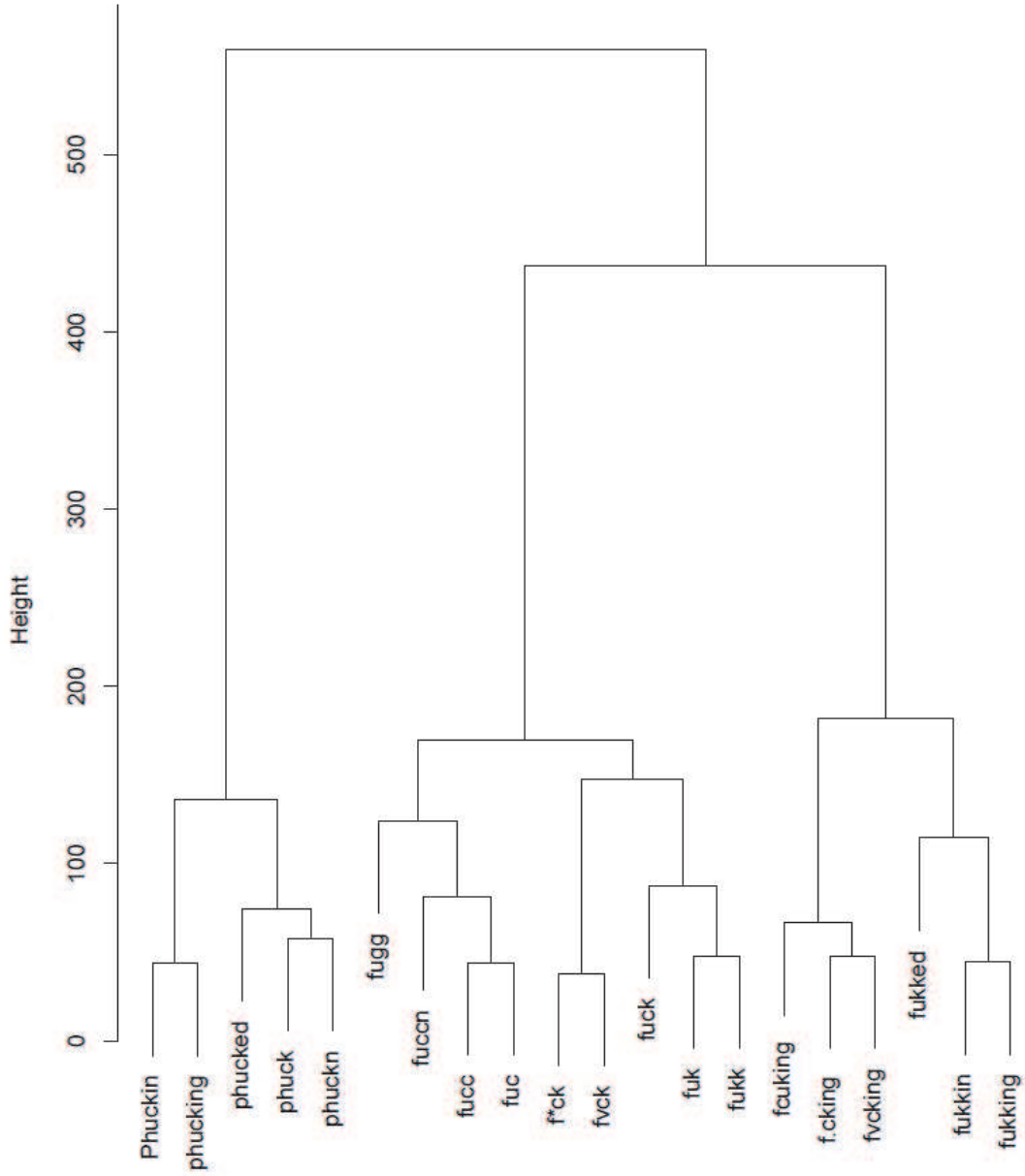


Illustration 5.20 – Dendrogramme des graphies de <fuck>

cela, les formes standard (mais pas leurs dérivés) sont toujours au milieu de la famille centrale. Dans le cas de *nigger*, le dendrogramme montre même que la forme initiale, qui inspire toutes les autres, est en fait <nigga>. Le glissement de sens est actualisé par les utilisateurs, qui ont souvent recours à cette forme pour se désigner, par exemple dans les pseudonymes. Les trois analyses de <nigger>, <bitch> et <fuck> démontrent que les scripteurs ont recours à des variantes allographiques surtout concentrées

sur les voyelles. Elles sont parfois des retranscriptions de prononciations exagérées, comme la diphtongaison de <biatch> (ou <beotch>, <beyotch>). Ces graphies restent très proches du mot orthographié, et les changements de consonnes servent souvent à substituer une représentation d'un son pour une autre, comme <ph> pour le <f> de « fuck ». Les consonnes peuvent également être remplacées par un voisin phonologique (comme dans la graphie <nikka> pour <nigga>), où seul le voisement est affecté.

La partie suivante est une présentation rapide d'autres cas de variances graphiques pour des lexèmes moins représentatifs d'une communauté, et achèvera ce chapitre.

5.5.4 Autres phénomènes de variances graphiques

Je fais ici la description de phénomènes similaires sur SC, sur des variantes allographiques beaucoup moins nombreuses. Certaines graphies utilisées sont prosthétiques, et montrent l'absence d'un équivalent graphique stable pour les prononciations écossaises.

Dans le corpus SC, la graphie <noe> peut représenter à la fois « now », et « not ». Elle n'a en outre rien à voir avec le système de votes (des *ayes* et *noes*) britanniques, et <noes> n'existe pas dans le corpus SC. Lorsqu'il s'agit de « now », la faute s'explique simplement par la proximité des touches <E> et <W> dans un clavier QWERTY. Pourtant, cela peut être ambigu, comme le montre l'illustration 5.21.

En lisant le message, (une réponse à une utilisatrice qui déclare qu'elle va aller en Écosse) la

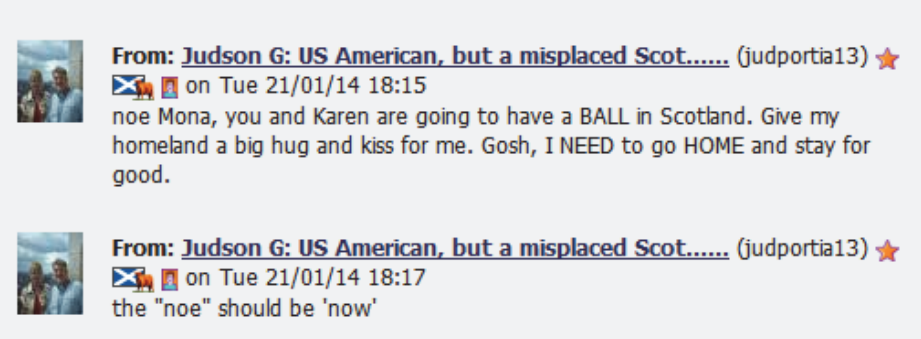


Illustration 5.21 – Extraits de réponses à un statut contenant <noe>, 'now' (SCP9)

phrase pourrait très bien se lire « No Mona, you and Karen are going to have a BALL in Scotland. ».

La correction ultérieure, deux minutes après avoir posté le premier message, change le sens de cette graphie, et montre qu'il ne s'agit que d'une faute d'orthographe. Ce cas est d'autant plus ambivalent que d'autres utilisateurs écrivent *not* <noe>, peut-être pour symboliser l'absence de <t> prononcé /t/, qui est certainement remplacé par une occlusive glottale.

> Aye ! yell hiv ta be carefull its **noe** a woman, check first and see if he/ she has bells and give them a good shoogle jist ta make sure [59]. (SC2A79)
> My name is Angela I was born in Glasgow,Lived Aberdeen for 26 years and then shipped to Florida. I am married to a yank. I have 4 grown up children apple of my eye.Also a beautiful 4 year old grandson. 3 sons 1 daughter. I miss hame its jist **noe** the same here. The sun is braw thats about aw. [=I miss home it's just not the same here. The sun is great, that's about all.] (SC1K49)

Dans les deux citations, <noe> est bien *not*, sans autre interprétation possible. Ces graphies ressemblent aux apostrophes d'excuse [60], mais sont beaucoup plus faciles à observer. Il n'est pas possible de faire une étude des apostrophes d'excuse avec les recherches sur Google, parce qu'elles sont traitées comme n'importe quel caractère non-alphanumérique. Les graphies prosthétiques comme le <e> dans <noe> peuvent servir à contourner le problème d'homographies.

Les utilisateurs peuvent se tromper sur l'interprétation de mots, surtout hors contexte. Dans le fil de conversation SC2N59 et l'extrait proposé dans le tableau 5.43, les utilisateurs échangent des mots en scots. L'auteur du message 106 propose « wids », mais le mot est interprété comme « would/wood », deux homophones (/wʊdz/), voire « weds » (/wedz/). Pour « would », on pourrait même proposer la forme non-accentuée /wəd/. Elle est cependant écartée tout de suite, étant donné qu'il n'est pas possible de rajouter un <s> à un modal.

La confusion entre /ɪ/, la valeur phonologique attendue pour la lettre <i>, et les propositions « woods », « would » et « weds » (/ʊ/, /e/, /ə/) montre la confusion entre les classes d'isophones en scots : le rapprochement de STRUT, DRESS (et COMMA) de KIT est permis par l'utilisation de graphies alternatives.

59. Le contexte de ce message est une diatribe contre David Starkey, un animateur de télévision homosexuel qui a déclaré que l'Irlande, le pays de Galle et l'Écosse étaient de « ersatz dérisoires de nations » [*'feeble little nations'* , traduction personnelle).
60. Apostrophes utilisées pour symboliser la suppression d'une lettre dans une graphie alternative, présente dans la forme orthographique. Voir la partie 2.2.2.1.

Message	Corps
106	[…] Heres anither wan “Wids”. […]
108	Well Michael, it’s either- Woods, or would ?? Am I right ?
109	I thought it was ’weds’
111	Am wi you, Donna, i think a bear shites in the wids :-)
113	Hi all been busy walking the dugs in the wids takin photos,got fair clag-gered in mud daein it tae. [=Hi all been busy walking the dogs in the woods taking photos, got fairly covered in mud doing it too.]

Tab. 5.43 – Messages 106 - 113 de Loss of our Language (SC2N59)

Ceci se rapproche de la constatation faite plus haut : plus le mot est court, plus il est difficile de le comprendre dans sa graphie alternative, à cause des multiples candidats qui pourraient rentrer dans le patron CVC proposé.

L’auteur de cette proposition, dans le message 106, clarifie donc le sens du mot avec une phrase en contexte, dans le message 113. Aucun des autres participants ne commente la réponse, et ils continuent à échanger des mots : le mystère de la graphie alternative a provoqué plusieurs réactions, mais une fois levé, les membres de SC ne s’y attardent pas.

Conclusion du chapitre

Ce chapitre est le plus important pour l’analyse graphophonématique des pratiques discursives en ligne. Le chapitre précédent traitait des conditions d’apparition des graphies alternatives, mais celui-ci tente de dresser un portrait modeste des types de graphies trahissant un changement de prononciation. Les graphies alternatives étudiées dans ce chapitre recoupent effectivement les graphies démarcatives du chapitre 4, car les variantes allographiques de <nigger>, <bitch> et <fuck> ne seraient pas si nombreuses si ces mots pouvaient être écrits ainsi sur BP. Les graphies alternatives ne s’éloignent pas des figurations de la prononciation, et confirment l’hypothèse du premier chapitre. Certaines graphies de *Push* et de TR-L sont effectivement trouvées sur le corpus BP et SC, comme <dat> pour « that » ou <airm> pour « arm », respectivement.

L’analyse est certes centrée sur des graphies dans SC et BP, mais elle est généralisable. Les graphies

alternatives et les variations allographiques ne se concentrent pas toujours sur la voyelle. La consonne, pourtant vue comme porteuse d'identité de mot, à cause de la plus grande régularité des distributions entre graphie et phonie pour les consonnes, peut aussi subir des transformations, autour de la prononciation de sons moins forts (chute de <T> dans le groupe consonantique /st/, dans <massa>). En revanche, les transformations des voyelles sont plus spectaculaires, à cause des déplacements phonologiques des classes d'isophones. Les transcriptions en graphies alternatives obéissent souvent à des désirs de transcriptions de la prononciation. C'est pourquoi les voyelles, surtout les voyelles non accentuées, et schwa en tête, sont transcrites avec n'importe quel caractère. Les fautes d'orthographe, comme <logerithum>, donnent aussi des indications sur la façon dont les locuteurs perçoivent ce qu'ils prononcent.

Dans ce chapitre, j'ai tenté de dégager les propriétés du communaulecte, par une analyse quantitative de certaines graphies alternatives relevées. Bien qu'une graphie soit employée très rarement sur les forums, elle l'est souvent proportionnellement plus souvent que sur internet. Les éléments communaulectaux ainsi dégagés ne sont pas pensés ou conçus par les fondateurs de la communauté (à quelques exceptions près, comme l'utilisation de la « Heilan Coo » en tant que mascotte de SC, dont la graphie lue sur presque toutes les pages du site influence certainement les pratiques graphiques des utilisateurs du forum). Les utilisateurs ne parlent pas des artefacts linguistiques créés par leurs expériences graphiques : le <wids> '*woods*' n'est pas commenté ni par son forgeur, ni par les membres qui ont tenté de deviner de quel mot il s'agissait, et il demeure un hapax dans le corpus. Les utilisateurs limitent d'ailleurs les questions épilinguistiques à leur identité : le parler « ghetto » sur BP, ou l'endroit où ils ont été élevés, comme pour les utilisateurs de SC, qui listent parfois les lieux où ils ont vécu. Ces questions épilinguistiques sont parfois plus éloignées, comme les imitations de l'accent jamaïcain qui passent par l'alliance d'une expression stéréotypée et d'une graphie démarcative (« phram yawd ») ; ou l'imitation du parler des esclaves, avec « massa ».

Les pratiques communaulectales identitaires passent également par l'utilisations de termes-icônes, stéréotypiques. Sur SC, la graphie <highland> est celle de l'anglais, et elle est refusée par les fondateurs du site ; les graphies diverses de *nigger*, *bitch* et *fuck* sur BP sont dues à la censure aveugle

pratiquée par les scripts automatiques. Alors qu'elles sont surtout utilisées pour subvertir l'ordre imposé par BP, ces graphies permettent de dégager la prévalence sémantique de *nigga* par rapport à *nigger*. Ces graphies ne sont cependant pas utilisées uniquement dans ces communautés virtuelles, et l'influence de l'anglais américain sur internet et le netspeak explique pourquoi le communaulecte de BP est moins unique que celui de SC, comme les valeurs des tests statistiques l'ont montré. Des sources extra-communautaires, comme la marque FCUK normalisent la métathèse et la mettent à disposition des membres qui veulent censurer le mot *fuck*.

La différence de graphie entre les deux langues minoritaires tient aussi à leur différence de statut : les locuteurs du scots ont conscience que leurs usages sont réellement scindés de l'anglais, et le statut écrit non-standardisé de la graphie du scots est visible dans les dictionnaires, comme le DSL. Les libertés graphiques offertes par le scots correspondent aux différences d'accents (et l'OSD offre plusieurs variations phonologiques). L'attachement à la langue explique les variations graphiques comme <aa>, <au> et <aw> pour *all*, par exemple. Pour l'AAV, la difficulté principale résidait dans la proximité avec le *netspeak*, et à l'utilisation de caractères non-alphabétiques pour contrecarrer la censure, comme la graphie <b !t(hez> pour *bitch*. La volonté de styliser le discours avec des graphèmes plus rares (substitution de <i> par <y>, ou <f> par <ph>) et l'influence culturelle de certains groupes comme les gangs, poussent les utilisateurs à puiser dans plusieurs sources pour personnaliser leurs graphies. Dans les deux corpus, la proportion d'erreurs orthographique n'est certainement pas négligeable mais, à moins qu'elles ne soient corrigées par leur auteur (comme la graphie <noe> pour *now*), elles se confondent avec les graphies alternatives désirées.

L'influence de la prononciation sur la graphie est manifeste, et le désir d'écrire au plus près de ses réalisations phonologiques est prégnant dans les deux forums. Il n'est pas possible de s'éloigner du standard orthographique en ne suivant pas un nouveau standard. Il est soit essentiellement graphique, voire presque ambigrammique [61], comme le leetspeak, qui repose sur la capacité à reconnaître un caractère imparfait : la lettre <l> est remplacée par <1>, sans presque de changement visuel. Pour le remplacement de <a> par <4>, on joue sur la représentation en capitale (<A>). Ces jeux se basent

61. Un ambigramme est une lettre ou un mot qui garde le même sens quel que soit le sens dans lequel on le regarde, ou qui peut même avoir un nouveau sens, comme le logo du groupe AƁBA ou de Nine Inch Nail (<NIИ>), le logo de la marque New Man…L'ambigramme repose sur la capacité à comprendre une lettre imparfaitement tracée et à en extrapoler du sens.

également sur la coopération du lecteur, qui devra déchiffrer ces jeux. Le nouveau standard qui dirige la graphie peut être sémantique, mais il se cantonne à très peu de mots : écrire <Amerikkka> ou <yawd> transforme le sens en '*Amérique raciste et dirigée par le Ku Klux Klan*' et '*la Jamaïque*', respectivement. Ces graphies ne sont comprises que par les initiés, qui connaissent le KKK [62] ou l'accent jamaïcain. Le dernier critère qui m'intéressait tout particulièrement dans ce chapitre est le standard phonologique. Créer les graphies alternatives suivant sa prononciation donne des indications inédites sur la façon dont sont perçus phonèmes, graphèmes, et leurs relations. Le forum est un espace permissif qui permet la création de formes artefactuelles graphiques. Ces graphies alternatives en appellent d'autres, ce qui explique pourquoi les utilisateurs peuvent donner spontanément plusieurs candidats pour une graphie comme <wids>, qui n'est pas comprise immédiatement. Personne ne s'offusque de cette nouvelle règle de la lecture : normalement, si un mot écrit n'est pas compris, c'est qu'il est inconnu, pas que sa graphie est inconnue. Dans une utilisation communaulectale, ce n'est pas le cas, et les utilisateurs ont intégré la remotivation proposée par un autre membre, qui comporte souvent un substrat phonologique.

Dans ce chapitre, j'ai délimité certaines propriétés graphophonématiques dérivables des usages des membres de BP et SC. Il est temps de refermer ces analyses, et d'en tirer des conclusions générales concernant la nature et l'utilisation du communaulecte.

62. Il serait surprenant mais pas impossible qu'un anglophone ne connaisse pas cette organisation.

Conclusion

Real Gs move in silence like lasagna.

Lil' Wayne - *6'7'*

Récapitulatif des problématiques traitées

La rencontre entre la phonologie, plus particulièrement la graphophonématique, et l'analyse de discours a permis d'arriver à plusieurs constatations. Le forum est un espace de conservation des graphies alternatives, à l'exploitation risquée, tant il est facile de perdre les données. Celles-ci sont volatiles, peuvent disparaître d'internet à tout moment. L'étude en contexte de ces forums permet de n'écarter aucun aspect de leur écologie, et de privilégier l'exploration verticale du forum, en respectant les lieux où les liens hypertextes conduisent, par exemple. Ce type de recherche était primordial pour l'analyse de discours, mais va à l'encontre des méthodes automatiques de la textométrie, qui imposaient au contraire un recueil des données rigoureux, et des calculs quantitatifs bien éloignés du travail qualitatif. Les forums sont potentiellement des espaces infinis, alors que les corpus qui en sont tirés sont des espaces essentiellement clos, et circonscrits au moment du recueil de données.

L'éthique de l'observation est elle-même dans une zone grise, puisque ces discours sont recueillis sans que les internautes n'aient à le savoir. Le recueil est facilité grâce à l'archivage automatique essentiel aux forums, où les messages sont conservés dans les fils de discussion dès qu'ils sont publiés.

L'identité et la reconnaissance de soi, et de soi en la communauté est primordial pour la bonne marche de forums de niche comme BP et SC. L'activité scripturale est la pierre angulaire à la vie de la communauté virtuelle, et l'interaction est maîtresse lors de la création de différents types de graphies. Certains utilisateurs influencent les autres, comme le scripteur de <gurdle> pour « girdle » sur BP ; certains apprennent et utilisent les mots enseignés par les autres utilisateurs en scots.

Les graphies contournatives sont nées du contexte de censure de BP, les graphies démarcatives sont des marques de la distanciation vis à vis de l'anglais standard, et les identités personnelles sont aussi exprimées dans ces graphies qui contournent les tabous. Le contexte de production dicte le type de graphies adoptées. Les graphies sont aussi indicatives des différences phonologiques les deux langues minoritaires et la norme de référence de leur pays. Le scots a su préserver des graphies anciennes, remises au goût du jour par les utilisateurs de SC. Ces graphies sont conformes avec les recommandations scolaires quant à l'utilisation du scots (Chris, 2010). L'AAV possède un lexique disparate, mêlé d'emprunts aux langues africaines ou de récupérations de mots anglais, comme le mot « nigga ». Les figurations fictionnelles graphiques de ces langues minoritaires rejoignent les formes trouvées dans le corpus fait d'écrits authentiques. Le corpus secondaire et le corpus principal partagent les mêmes caractéristiques, où la graphie est indicatrice d'une appartenance identitaire. Prouver son appartenance à la communauté virtuelle, c'est utiliser les graphies alternatives qui lui sont propres. Ces graphies sont partagées par les scripteurs, et elles sont plus fréquentes dans BP et SC qu'en ligne.

Les écrits sur internet sont évidemment un discours à part. Les plateformes communautaires ont évolué, BP en premier, pendant cette analyse, avec l'implémentation de plusieurs fonctionnalités le faisant ressembler à FB. Ce lieu numérique n'a rien à voir avec un lieu où le pouvoir linguistique n'est partagé que par les dominants, où le *skeptron* discursif ne serait donné qu'à la norme. Le mot « forum » est tout à fait juste pour décrire cet espace semi-public, où tous les utilisateurs ont la parole à part égale, tant est qu'ils peuvent se conformer à ses règles. La graphie est encore plus porteuse de l'identité du forum que les utilisations lexicales spécifiques, parce que les graphies alternatives peuvent s'immiscer dans le moindre mot, et le rendre hors-norme : la graphie <bytch> entraîne des changements sémantiques, par rapport à « bitch ». Le lexème <bytch> peut être valorisant pour une femme forte

et indépendante, montrer une prononciation différente, changer le genre préconstruit en masculin ou être utilisé en tant que censure, selon les entrées de l'UD [63]. Les utilisations lexicales iconiques sont des indices d'appartenance à la communauté, qui ne sont pas mutuellement exclusifs des graphies alternatives : la <Heilan Coo> de SC n'est pas une graphie unique à ce site, mais elle y est très représentée, au détriment de « Highland Cow ». La nature américanisante du netspeak (Crystal, 2011), qui emprunte de nombreuses graphies à l'anglais américain, créée l'illusion que le communaulecte de BP est moins spécifique à cette communauté que celui de SC [64], mais les deux communaulectes sont statistiquement différents du netspeak, de l'anglais standard, ou même des langues minoritaires, et ils acquièrent une identité propre en s'inspirant de toutes ces sources. Ces données sont labiles, mais leur changement n'entraîne pas de changements drastiques : certaines recherches effectuées en 2010 renvoient le même nombre de résultats, comme « whitemanitus » (3 résultats en 2010 et en 2014).

Positionnement épistémologique

Cette thèse s'inscrit totalement dans la troisième révolution de la grammatisation (Auroux, 1994), soit l'émergence du traitement automatisé des langues, et plus précisément l'émergence du traitement automatique de pratiques linguistiques propres au discours sur internet. Les pratiques scripturales alternatives en ligne échappent aux dictionnaires, qui tentent de s'adapter aux nouveaux mots créés sur internet : les acronymes « LOL » (‘*Laughing Out Loud*’), « OMG » (‘*Oh My God*’) et même le symbole « ♥ » (verbe, glosé en « to heart » dans la définition) sont rentré dans les dictionnaires Oxford en 2011. Leur pratique est surtout numérique, bien que <♥> vienne des phrases du type « I ♥ NY ». Ce symbole est réalisable sur la plupart des plateformes virtuelles, et la combinaison < <3> crée un cœur (par exemple, sur FB). Il a donc été intégré aux pratiques numériques en ligne, dont les plus populaires sont intégrées aux dictionnaires. Le discours numérique est constitué de créations personnelles d'artefacts graphiques souvent éphémères, qui contribuent à un code qui diverge des normes de référence

63. Disponibles à l'adresse <http ://www.urbandictionary.com/define.php ?term=Bytch>. Dernier accès le 23 août 2014.
64. Je rappelle les valeurs p pour BP et SC : 0.003558 et 3.349×10^6, respectivement (voir 5.4.2).

établies des langues parlées. Les usages grammaticaux, lexicaux et graphiques sur internet pourraient y créer de nouvelles langues, dont le netspeak est le représentant général. Les limites de la technologie sont repoussées, et d'autant plus visibles sur ce corpus que dans des réseaux sociaux ou forums généralistes. Le netspeak est la toile de fond sur laquelle se jouent ces usages. Le besoin d'étrangeté de graphies alternatives (comme la simple substitution de <i> par <y>, lettre moins fréquente) semble né du besoin de se distinguer, et ce besoin est primordial dans la constitution d'un communaulecte, qui distingue les pratiques d'une communauté virtuelle de toutes les autres.

Comme le montre l'épigraphe de cette conclusion, au premier abord mystérieuse, l'important est la façon dont les utilisateurs perçoivent la langue qu'ils manipulent : que le /g/ de « lasagne » disparaisse réellement, ou qu'il cause une controverse, l'important est de faire passer le message – ici, un jeu phonologique sur la comparaison des « Gs » (‘*gangsters*’) et d'un caractère graphique qui n'exprime pas tout à fait sa valeur phonologique prototypique. Pour utiliser une comparaison un peu plus classique, le héros de la nouvelle « A Sound of Thunder » (Bradbury, 1965), ne se rend compte qu'il a créé une réalité parallèle qu'en lisant un panneau au mur : « Tyme Sefari Inc. Sefaris tu any yeer en the past. Yu naim the animall. We taekyuthair. Yu shoot itt. ». La langue écrite est donc une garante de l'espace culturel partagé.

Les pratiques linguistiques sur internet sont moins un arrangement chimérique entre l'oral et l'écrit, « a kind of linguistic centaur, incorporating features from both traditional writing and face-to-face discourse but ending up being more than a simple amalgam of the two. » (Baron, 2005) que les deux faces d'une vitre, qui se perçoivent l'une l'autre sans pour autant se mélanger tout à fait. On est capables de voir la langue orale, inscrite dans les marques graphiques, mais les tentatives oralisantes, comme l'utilisation de « LOL » ne sont pas des marques d'oralité. Les conversations en ligne n'ont pas la même structure de tours de parole, à cause notamment des effets de citationnalité. Le reflet de l'oral dans la vitre de l'écrit montre qu'elles sont proches, que les réélaborations graphiques tendent vers une transcription de l'oral, mais qu'il en sera toujours absent. Il n'est pas possible de savoir comment tous les scripteurs de <biatch> prononcent ce mot, et s'il est différent de <beyotch>, ou même, s'il est perçu différemment par les énonciateurs et les destinataires. Cette étude phonographématique est

donc frustrante, parce qu'il y manquera toujours l'aspect phonétique, mais la graphie tend à réduire la différence entre norme orthographique et prononciation, sans tout à fait y parvenir.

L'aspect discursif engage l'agencement collectif d'énonciations :

> L'unité réelle minima, ce n'est pas le mot, ni l'idée ou le concept, ni le signifiant mais l'agencement. C'est toujours un agencement qui produit les énoncés. Les énoncés n'ont pas pour cause un sujet qui agirait comme sujet d'énonciation pas plus qu'ils ne se rapportent à des sujets comme sujets d'énoncé. L'énoncé est le produit d'un agencement toujours collectif qui met en jeu en nous et dehors de nous des populations, des multiplicités, des tentations, des devenirs, des affects, des évènements. (Deleuze & Guattari, 1980)

Cette vision deleuzienne de l'agencement résume parfaitement le but qu'avait ce travail de thèse. Sans énonciation collective dans le contexte des forums, il n'y aurait pas d'émergence de ce type de graphies alternatives. Les forums mettent le collectif au centre des communications, contrairement aux réseaux sociaux plus centrés sur l'identité personnelle. Cette identité, allant jusqu'au narcissisme et à la « pornographie émotionnelle » (Ippolita & Milani, 2013) pour ses détracteurs, ne permet pas une telle identification à la plateforme, à cet agencement à la fois personnel et identitaire, qui renvoie autant à soi qu'au forum fréquenté.

Perspectives de recherche

Ce travail de thèse ne saurait satisfaire des exigences d'exhaustivité, mais il est le point de départ de recherches ultérieures, tant en phonologie qu'en analyse de discours numérique.

Les graphies en tant que marqueurs sémiotiquement pertinents peuvent aussi être exploitées autrement. Il faudrait étudier les discours transversaux et les variations intra-scripturales, afin de voir quelles pratiques discursives les internautes adoptent ou déposent suivant le contexte d'énonciation. Les utilisateurs développent des pratiques communaulectales, qu'ils imitent et exportent certainement dans leurs autres écrits virtuels. La péremption de certaines pratiques, ou l'émergence de pratiques discursives issues d'autres plateformes est une extension de ce travail, et on peut observer la perméabilité des sphères internet par les études textométriques. Par exemple, la graphie <surfbort> de *surfboard* est une adaptation de la prononciation de ce mot par Beyoncé, et il apparaît sur BP, depuis janvier

2014, soit depuis qu'il a été détourné graphiquement. Sur SC, les utilisateurs sont moins réactifs au netspeak, mais ils le pratiquent également, et une deuxième étude de la proportion de netspeak dans les deux communautés pourrait être envisagée.

Les analyses graphophonématiques pourront également donner lieu à la création d'un logiciel qui permettrait de prédire les graphies alternatives à partir d'un mot orthographié, pour des applications pratiques d'indexation de mots-clef, à la façon de *FatFingers* (« Fat Fingers », 2004), qui n'effectue que des recherches de graphies alternatives sur *eBay*. Les algorithmes de classification phonétique, comme soundex ou son cousin plus efficace, NYSIIS (New York State Identification and Intelligence System) pourraient servir de base d'application, non seulement pour les raffiner, mais aussi pour les adapter à l'AAV et au scots, voire à d'autres langues régionales et minoritaires.

Bibliographie

Linguistique générale

Allan, K. & Burridge, K. (2006). *Forbidden words : taboo and the censoring of language*. Récupérée à partir de http://bit.ly/1lJJhwm

Anis, J. (1999). *Internet, communication et langue française*. Paris.

Auroux, S. (1994). *Histoire des idées linguistiques*. Liège.

Bakhtine, V., Mikhail Mikhaïlovitch (Volochinov. (1977). *Le Marxisme et la Philosophie du Langage : Essai d'Application de la méthode sociologique en linguistique*. Paris.

Ballier, N. (1997). *Les écoles françaises de linguistique anglaise (1967 - 1992)*.

Bellama, D., Nkwele, S., Yudom, J., Peace Corps (U.S.) & Cameroon. (1983). *An introduction to Cameroonian pidgin*. Yaounde, Cameroon.

Benveniste, E. (1966). *Problèmes de linguistique générale*. Paris.

Blake, N. F. (1981). *Non-standard language in english literature*. London ; Boulder, Colo.

Blevins, W. (1998). *Phonics from a to z : a practical guide*. New York.

Brown, E. K. & Anderson, A. (2006). *Encyclopedia of language & linguistics*. Boston.

Bulot, T. (2004). *Les parlers jeunes : pratiques urbaines et sociales*. Rennes.

Burov, I. (2012). *Les phénomènes de sandhi dans l'espace gallo-roman*.

Calvet, L.-J. (2004). *Essais de linguistique : la langue est-elle une invention des linguistes ?* Paris.

Carr, P. (1999). *English phonetics and phonology : an introduction*. Malden, Mass.

Catach, N. (1978). *L'orthographe*. Paris.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York.

Croft, W. & Cruse, D. A. (2004, janvier 22). *Cognitive linguistics*.

Crozier, J. (2008). *Collins dictionary*. Glasgow. Récupérée à partir de http://www.contentreserve.com/TitleInfo.asp?ID=%7B3E145529-0095-4C78-B477-999D1493D9DF%7D&Format=900

Crystal, D. (1995). *The cambridge encyclopedia of the English language*. Cambridge ; New York.

Cukor-Avila, P. & Bailey, G. (2001). The effects of the race of the interviewer on sociolinguistic fieldwork. *Journal of Sociolinguistics*, 5, 254–70.

Davis, M. (2012). *Aoccdrnig to a rscheearch at cmabrigde uinervtisy.*. [Cognition and brain sciences unit]. Récupérée à partir de http://www.mrc-cbu.cam.ac.uk/people/

Dehaene, S. (2007a). *Les Mécanismes Cérébraux de la Lecture*. Cours du Collège de France. Récupérée à partir de http://www.college-de-france.fr/site/stanislas-dehaene/p1347276757294_content.htm

Dehaene, S. (2007b). *Les Neurones de la Lecture : la Nouvelle Science de la Lecture et de son Apprentissage*.

Deleuze, G. & Guattari, F. (1980). *Mille plateaux*. Paris.

Deschamps, A. (1994). *De l'écrit à l'oral et de l'oral à l'écrit : phonétique et orthographe de l'anglais*. Gap ; Paris.

Deschamps, A., Fournier, J.-M., Duchet, J.-L. & O'Neil, M. (2004). *English phonology and graphophonemics*. Paris.

Doggie. (2003). In The free dictionary.

Eloy, J.-M. (2004, septembre 1). *Des langues collatérales : problèmes linguistiques, sociolinguistiques et glottopolitiques de la proximité linguistique -*.

Fry, E., Kress, J. E. & Fountoukidis, D. (1993). *The reading teacher's book of lists*. Englewood Cliffs, NJ.

Gadet, F. (2007). *La variation sociale en français (2è édition)*. Gap, France.

Gambier, I. (1996). *Les transferts linguistiques dans les médias audiovisuels*.

Gibbons, D. (1997). *Functions of language*. Récupérée à partir de http://coral.lili.uni-bielefeld.de/Classes/Summer98/Functions/functions/functions.html

Giegerich, H. J. (1992). *English phonology : an introduction*. Cambridge ; New York.

Goffman, E. (1987). *Façons de parler*. Paris.

Goudet, L. G. (2013b). *Phonologie et lexicologie sauvages : l'Accent tag sur Youtube*. Congrès de la SAES. Université de Bourgogne–Dijon.

Gove, P. B. & Merriam-Webster, I. (1993). *Webster's third new international dictionary of the English language (unabridged)*. Springfield, Mass.

Hogg, R. M., Blake, N. F., Lass, R., Romaine, S., Burchfield, R. W., Algeo, J. & McClure, J. D. (1992). *The cambridge history of the english language*. Cambridge ; New York, NY, USA.

Hopper, P. J. & Traugott, E. C. (1993). *Grammaticalization*. Cambridge ; New York, NY, USA.

Jacewicz, E., Fox, R. A. & Salmons, J. (2007). Vowel duration in three american english dialects. *American speech*, 82(4), 367–385. Récupérée à partir de http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2829779/

Jakobson, R. (1960). Closing statement : linguistics and poetics. *Style in language / ed. by Thomas A. Sebeck*, 350–377.

Johnstone, B. (2011, novembre 1). Dialect enregisterment in performance. *Journal of Sociolinguistics*, 15(5), 657–679. Récupérée à partir de http://onlinelibrary.wiley.com.ezproxy.unr-runn.fr/doi/10.1111/j.1467-9841.2011.00512.x/abstract

Kerbrat-Orecchioni, C. (nodate). L'impolitesse en interaction : aperçus théoriques et étude de cas. *Studii si Cercetari Filologice : Seria Limbi Romanice*, 1. Récupérée à partir de http://screcherche.univ-lyon3.fr/lexis/IMG/pdf/Lexis_special_2_-_Kerbrat-Orecchioni.pdf

Klinkenberg, J.-M. (1994). *Des langues romanes : introduction aux études de linguistique romane*. Louvain-la-Neuve.

Kortmann, B. (2004). *A handbook of varieties of english vol. 2, morphology and syntax*. Berlin.

Kortmann, B. & Lunkenheimer, K. (2008). *The world atlas of varieties of english - online*. Récupérée à partir de http://www.ewave-atlas.org/home

Labov, W. (1973). *Sociolinguistic patterns*. Philadelphia, PA.

Labov, W. (2006). *The atlas of north american english : phonetics, phonology, and sound change : a multimedia reference tool*. (). Berlin ; New York.

Ladefoged, P. (1993). *A course in phonetics*. Fort Worth.

Ladefoged, P. (1999). *Handbook of the international phonetic association : a guide to the use of the international phonetic alphabet : american english*. Récupérée à partir de http://www.archive.org/details/rosettaproject_eng_phon-4

Larson, K. (2004). *The science of word recognition or how i learned to stop worrying and love the bouma*. Récupérée à partir de https://www.microsoft.com/typography/ctfonts/WordRecognition.aspx%5C#m1

Lecercle, J.-J. (2004). *Une philosophie marxiste du langage*. Paris.

Lee, M. (2010). *A brief history of fool-pitying : origins of the phrase "I pity the fool"*. [Overthinking it]. Récupérée à partir de http://www.overthinkingit.com/2010/05/04/i-pity-the-fool-origin/

Liberman, M. (2009). *Language log » annals of bowdlerization : whiskey tango foxtrot*. [Language log]. Récupérée à partir de http://languagelog.ldc.upenn.edu/nll/?p=1941

Mauss, M. (1936). *Les techniques du corps*. Chicoutimi. Récupérée à partir de http://dx.doi.org/doi:10.1522/cla.mam.tec

Mazière, F. & Auroux, S. (2006). Introduction : hyperlangues, modèles de grammatisation, réduction et autonomisation des langues. *Histoire Épistémologie Langage*, 28, 7–17. Récupérée à partir de http://www.persee.fr/web/revues/home/prescript/article/hel_0750-8069_2006_num_28_2_2878

MTV2. (2004). *Slanguistics*. [Slanguistics].

Neuman, Y. (2009). *L'influence de l'écriture sur la langue*. Récupérée à partir de https://www.academia.edu/1435958/Neuman_2009._Linfluence_de_lecriture_sur_la_langue_Ph.D._dissertation_Sorbonne_Nouvelle_

Ong, W. J. (1982). *Orality and literacy : the technologizing of the word*. London ; New York.

Paveau, M.-A. (2007, mars 1). Les normes perceptives de la linguistique populaire. *Langage et société*, n° 119, 93–109. Récupérée à partir de http://www.cairn.info/article.php?ID_ARTICLE=LS_119_0093

Preston, D. R. (1982, juillet 1). 'ritin' fowklower daun 'rong : folklorists' failures in phonology. *The Journal of American Folklore*, 95(377), 304–326. doi :10.2307/539912

Quirk, R. (1963). *The use of english*. New York.

Rayner, K., White, S. J., Johnson, R. L. & Liversedge, S. P. (2008). *Raeding wrods with jubmled lettres : there is a cost*. Récupérée à partir de http://hdl.handle.net/2381/3897

Roach, P., Hartman, J., Setter, J. & Jones, D. (2006). *Cambridge english pronouncing dictionary.* Cambridge ; New York.

Rossari, C., Ricci, C. & Spiridon, A. (2009). *Grammaticalization and pragmatics : facts, approaches, theoretical issues*. Bingley, UK.

Saussure, F. d. (1917). *Cours de linguistique générale*. Paris.

Sorlin, S. (2010). *La défamiliarisation linguistique dans le roman anglais contemporain*. Montpellier.

Sorlin, S. (2012). *Langue et autorité : de l'ordre linguistique à la force dialogique*. Rennes.

Sylvain, S. (1936). *Le créole haïtien : morphologie et syntaxe*. Wetteren.

Trovato, V. (2010). *Parlécriture, une méthode pédagogique humaniste*.

Trudgill, P. (1983). *Sociolinguistics : an introduction to language and society*. Harmondsworth, Middlesex, England ; New York, N.Y., U.S.A.

Trudgill, P. & Hannah, J. (2007). *International english : a guide to varieties of standard english*. London.

Wells, J. C. (1982). *Accents of english*. Cambridge ; New York.

Wise, C. M. & Wise, H. (1965). *Applied phonetics*. Englewood Cliffs, N.J.

Zimmer, B. (2009). *Language log » the inherent ambiguity of "WTF"*. [Language log]. Récupérée à partir de http://languagelog.ldc.upenn.edu/nll/?p=1766

Zwicky, A. (1972). Note on a phonological hierarchy in english. *Linguistic Change and Generative Theory : Essays*,

Afro-Américain vernaculaire : sources primaires

Daniels, L., Fletcher, G., Paul, D., Grigorov, M., Sapphire, Sidibe, G., … Kravitz, L. (2011). *Precious*. Paris.

Sapphire. (1996). *Push : a novel*. New York.

Afro-Américain vernaculaire

African-American Baby Names. (2009). Récupérée à partir de http://www.babynames.org.uk

Anderson, C. J. (2012). *African-american english or ignorance ?* [Scribd]. Récupérée à partir de http://www.scribd.com/doc/103783691/African-American-English-or-Ignorance

Asante, M. K. (1988). *Afrocentricity*. Trenton, N.J.

Ball, A., Makoni, S., Smitherman, G., Spears, A. K., Thiong'o, F. b. N. w. & Winford, D. (2002, décembre 19). *Black linguistics : language, society and politics in africa and the americas*.

Baratz, J. C. & Shuy, R. W. (1969). *Teaching black children to read.* Washington.

Baugh, J. (1999). *Out of the mouths of slaves : african american language and educational malpractice*.

Baugh, J. (2000). *Beyond ebonics : linguistic pride and racial prejudice*. Oxford ; New York.

Bradford, J. (2001). *African american slang*. Récupérée à partir de http://faculty.fullerton.edu/jsantucci/african_am_slang.pdf

Brower, K. (2003). *Ebonics translator*. [Written humor]. Récupérée à partir de http://www.writtenhumor.com/ebonics.html

Campbell, K. E. [Kermit E]. (2007). There goes the neighborhood : hip hop creepin' on a come up at the u. *College Composition and Communication*, 58, 325–344.

Campbell, K. E. [Kermit Ernest]. (1994). The "signifying monkey" revisited : vernacular discourse and african american personal narratives. *Journal of Advanced Composition*, 14, 463–73.

Campbell, K. E. [Kermit Ernest]. (2005). *Gettin' our groove on : rhetoric, language, and literacy for the hip hop generation*. Detroit, Mich.

Claerbaut, D. (1972). *Black jargon in white america,* Grand Rapids.

Classics : 'the collection of Phat tracks'. (2004). Selangor, Malaysia.

Cohen, G. L., University of Missouri–Rolla & Department of Arts, L., and Philosophy. (2005). *Jazz revisited : on the origin of the term : draft #3*. Rolla, Mo.

Cotter, J. S. & Cosmopolitan Press. (1912). *Negro tales*. Récupérée à partir de http://catalog.hathitrust.org/api/volumes/oclc/4893591.html

Dillard, J. L. (1972). *Black english : its history and usage in the united states*. New York.

Eminem & Dr. Dre. (2000). *The real slim shady*. Santa Monica, Californie.

Fasold, R. W. (1999). *Ebonic need not be english*. Washington, DC.

Five Guys Burgers and Fries Review. (2012, avril 26). Récupérée à partir de http://www.youtube.com/watch?v=zGkHRa64sDY&feature=youtube_gdata_player

Garratt, S. (2010). Gabourey Sidibe interview for *Precious*. Récupérée à partir de http://www.telegraph.co.uk/culture/film/6990259/Gabourey-Sidibe-interview-for-Precious.html

Gates, H. L. (1988). *The signifying monkey : a theory of afro-american literary criticism*. New York.

Get rich or die tryin'. (2003). Santa Monica, CA.

Goudet, L. G. (2010a). *Grapho-phonemic representations of african american vernacular english on the internet*. Université Montpellier III.

Goudet, L. G. (2010b). *Identity and self-description in african american communities online*. University of Texas at Austin.

Goudet, L. G. (2013c). *''that foo just talks too much" : légitimité et économie du pouvoir dans un forum de gang afro-américain*. Corte, Corse.

Goudet, L. G. (2014). Discours de gangs afro-américains sur internet : dialectique du monologue de présentation personnelle. *Textualités Numériques*, Itinéraires. Littérature, Textes, Culture.

Goudet, L. G. & Ballier, N. (2014). Subversion genrée des pseudonymes en bitch et en nigga. Non Publié.

Green, L. J. (2002). *African american english : a linguistic introduction*. Cambridge, U.K. ; New York. Récupérée à partir de http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=125093

Hémart, L. (2001). *A study in african american English : Clareence Precious Jones's dialect in Push*.

Hickey, R. (2004). *Legacies of colonial english : studies in transported dialects*. Cambridge, England ; New York.

Holloway, J. E. (2005, août 30). *Africanisms in american culture*.

Hurston, Z. N. (1990). *Their eyes were watching god : a novel*. New York. (Date de première publication 1938)

*Jigga my n*** [nigga]*. (1999). New York.

Joel. (1999). *Ebonics translator - translate english into ebonics with the pimp translator*. Récupérée à partir de http://joel.net/EBONICS/Translator

Kearse, R. (2006). *Street talk : da official guide to hip-hop & urban slanguage*. Fort Lee, NJ.

Knox, G. W. (2011). Gang members on facebook : should we look the other way ? National Gang Crime Research Center. Récupérée à partir de http://www.ngcrc.com/gangface.html

Krahmer, B. (2006). *English to ebonics translator*. [Billism]. Récupérée à partir de http://www.billism.com/eng_to_ebon_trans.html

Labov, W. (1968). *A study of the non-standard english of negro and puerto rican speakers in new york city. cooperative research project no. 3288*. New York.

Labov, W. (1972). *Language in the inner city ; studies in the black english vernacular.* Philadelphia.

Labov, W., Wolfram, W., Baugh, J., Torbert, B. & Rickford, J. (1996). *African american vernacular english on the linguist list*. Récupérée à partir de http://linguistlist.org/topics/ebonics/

Lanehart, S. L. (2001). *Sociocultural and historical contexts of african american english*. Amsterdam ; Philadelphia. Récupérée à partir de http://public.eblib.com/EBLPublic/PublicView.do?ptiID=623117

Linguistic society of america : resolution on the oakland "Ebonics" issue. (1998). *Journal of English Linguistics*, 26, 174–176.

Major, C. (1971). *Black slang : a dictionary of afro-american talk.* London.

Miethaner, U. (2012). The blur (blues lyrics collected at the university of regensburg) corpus : blues lyricism and the african american literary tradition. *Current Objectives of Postgraduate American Studies*, 2. Récupérée à partir de http://copas.uni-regensburg.de/article/view/64

Miethaner, U. & Schneider, E. (1997). *BLUR–blues lyrics collected at the university of Regensburg*.

Minaj, N. & Soulja Boy. (2014). *Yass bish*.

Morton, W. H. (1964). *How they talk : one thousand misused words and idiomatic expressions*. Durham.

Mufwene, S. S. (1998). *African-american english : structure, history, and use*. London ; New York.

My Ghetto report card. (2006). Burbank, CA.

Nellyville. (2002). New York, NY.

Nurse, L. C. (2011). *Being black : exploring my caribbean diaspora identity through purposeful living, education, leadership and a divine plan*. Récupérée à partir de https://library.uvm.edu/jspui/handle/123456789/290

Popular African American Names. (2013). Récupérée à partir de http://www.babycenter.com/0_popular-african-american-names_10329236.bc

Richardson, E. B. (2006). *Hiphop literacies*. London ; New York.

Rickford, J. R. (1996). *Ebonics notes and discussion*. Récupérée à partir de http://www.stanford.edu/~rickford/ebonics/EbonicsExamples.html

Rickford, J. (1998). The ebonics controversy in my backyard : a sociolinguist's experiences and reflections. *Language in Society*, Récupérée à partir de http://www.stanford.edu/~rickford/papers/EbonicsInMyBackyard.html

Roberts, H. E. (1971). *The third ear : a black glossary,* Chicago.

Schneider, E. W. [Edgar W]. (1989). *American earlier black english : morphological and syntactic variables*. Tuscaloosa.

Sherman, J. R. (1997). *African-american poetry : an anthology, 1773-1927*. Mineola, N.Y.

Sidnell, J. (2002). *African american vernacular english*. Récupérée à partir de http://www.hawaii.edu/satocenter/langnet/definitions/aave.html

Smitherman, G. (2000). *Black talk : words and phrases from the hood to the amen corner*. Boston.

Snoop Dogg, Williams, P. & Jelly Roll. (2004). *Drop it like it's hot*.

Soulja Boy & Gucci Mane. (2008). *I, Soulja Boy Tellem*. Santa Monica, CA.

Suga Free. (2011). *Why u still bullshittin' ?–the best of Suga Free*. Suburban Noize.

Sullivan, C. (2009). *Jamaican patois*. [Reggae movement]. Récupérée à partir de http://web.archive.org/web/20090410050601/http://www.reggaemovement.com/History/patois.htm

Switzer, K., Brasseau, M. & Youngblood, K. (2010). *Aae : difference or disorder ?* College Station, Texas. Récupérée à partir de http://www.txsha.org/%5C_pdf/pdf/Switzer,%20Kathy-AAE.pdf

Tha Carter III. (2008). New York.

Thomas, E. R. (2007). Phonological and phonetic characteristics of African American Vernacular English. *Language and Linguistics Compass*, 1, 450–475.

Wolfram, W. (1998). Language ideology and dialect : understanding the oakland ebonics controversy. *Journal of English Linguistics Journal of English Linguistics*, 26, 108–121.

X, M. & Haley, A. (1992). *The autobiography of malcolm x*. New York.

Scots : sources primaires

Boyle, D., Macdonald, A., Hodge, J., Tufano, B., Welsh, I., Miramax Films, … Buena Vista Home Video (Firm). (2004). *Trainspotting*. Burbank, CA.

Welsh, I. (2011). *Trainspotting*. London. (Date de première publication 1993)

Scots

A dictionary for the dialects in "Trainspotting". (2007). Récupérée à partir de http://www.zoklet.net/totse/en/media/televisionary_film_vidiots/transpot.html

Alexander. (2012). *Learn scottish gaelic via skype*. [Learn scottish gaelic via skype]. Récupérée à partir de http://gaelicviaskype.webs.com/

BBC. (2008). *Bbc - alba - beag air bheag*. Récupérée à partir de http://www.bbc.co.uk/alba/foghlam/beag_air_bheag/

Black, I. (2003). *Edinbuggers v weegies (and vice versa !) ; weegies v edinbuggers (and vice versa !)* Edinburgh.

Blake, A. (2005). 'if boyo is racist so is jock'. Récupérée à partir de http://www.walesonline.co.uk/news/wales-news/2005/08/26/if-boyo-is-racist-so-is-jock-50082-15897870/

Boece, H. (1527). *Scotorum historiae a prima gentis origine cum aliarum et rerum et gentium illustratione non vulgari*. Parisiis.

Boece, H. & Bellenden, J. (1941). *The chronicles of scotland*. Edinburgh. (Date de première publication 1536)

Burns, R. (1867). *Poems and songs of robert burns*. London ; New York.

Chris. (2010). *Scuilwab - for teachers an young people fae nursery tae university*. Récupérée à partir de http://www.scuilwab.org.uk/

Corbett, J., Anderson, W., Beavan, D., Edmonds, F., Sweeney, L., O'Hare, C., … Stuart-Smith, J. (2007). *Scots project - scottish corpus of texts and speech*. Récupérée à partir de http://www.scottishcorpus.ac.uk/

Corbett, J., McClure, J. D. & Stuart-Smith, J. (2003). *The edinburgh companion to scots*. Edinburgh.

Costa, J. (2008). Langue et construction identitaire en écosse. Récupérée à partir de http://halshs.archives-ouvertes.fr/halshs-00376848

Costa, J. (2009). Language history as charter myth ? scots and the (re)invention of scotland. *Scottish Language*, 28. Récupérée à partir de http://hal-ens-lyon.archives-ouvertes.fr/ensl-00632432/en/

Costa, J. (2010). *Revitalisation linguistique : discours, mythes et idéologies. une approche critique de mouvements de revitalisation en provence et en écosse*. Récupérée à partir de http://tel.archives-ouvertes.fr/tel-00625691

Costa, J. (2012). Langue et nationalisme en écosse : trois langues pour une nation. *Bulletin d'Histoire Plitique*, 21, 96.

Durand, J. (2004). English in early 21st century scotland : a phonological perspective. *La Tribune des Langues Vivantes*, 36 English Pronunciation : Accents and Variation, 87–105. Récupérée à partir de http://w3.erss.univ-tlse2.fr:8080/index.jsp?perso=jdurand&subURL=EnglishinScotland.pdf

Eagle, A. (2001a). *Wir ain leed - aspects of scots spelling*. [Wir ain leed - a innit tar modren scots]. Récupérée à partir de http://www.scots-online.org/grammar/spelling.htm

Eagle, A. (2001b). *Wir ain leed - edinburgh scots*. Récupérée à partir de http://www.scots-online.org/grammar/edinburgh.htm

Eagle, A. (2002). *The online scots dictionary*. Récupérée à partir de http://www.scots-online.org/dictionary/engscots.htm

Education Scotland. (2012). *Scots spelling - writing in scots - knowledge of language*. Récupérée à partir de http://www.educationscotland.gov.uk/knowledgeoflanguage/scots/writinginscots/scotsspelling/index.asp

Europe. (1992). *Ste no. 148 - charte européenne des langues régionales ou minoritaires*. Récupérée à partir de http://conventions.coe.int/treaty/fr/Treaties/Html/148.htm

Ewan McGregor Interview. (2011, octobre 1). Récupérée à partir de http://www.youtube.com/watch?v=zlc01vH8kLU&feature=youtube_gdata_player

Flaws, M. T. & Lamb, G. (2001). *The orkney dictionary*. Kirkwall, Orkney.

Fòram na Gàidhlig. (2007). Récupérée à partir de http://www.foramnagaidhlig.net/foram/index.php

General Register Office (Scotland). (2005). *Scotland's census 2001 : gaelic report*. Récupérée à partir de http://www.gro-scotland.gov.uk/press/news2005/scotlands-census-2001-gaelic-report.html

Government, W. (2012, décembre 11). *2011 census : first results on the welsh language*. Récupérée à partir de http://wales.gov.uk/topics/statistics/headlines/population2012/121211/?lang=en

Great Britain. Registrar General Scotland & Scotland. General Register Office. (2001). Scotland's census laid before the scottish parliament pursuant to section 4(1) of the census act 1920 population report. *Scotland's census ... : laid before the Scottish Parliament pursuant to section 4(1) of the Census Act 1920 Population report.*,

Groundskeeper Willie Defines U.S. Image of Scots. (2007). Récupérée à partir de http://www.impactlab.net/2007/09/20/groundskeeper-willie-defines-us-image-of-scots/

Jenkins, M. (1996). 'Trainspotting' made easy - for americans [The independent]. Récupérée à partir de http://www.independent.co.uk/news/uk/home-news/trainspotting-made-easy--for-americans-1349197.html

Jones, C. (1997). *The edinburgh history of the scots language*.

Kay, B. (1986). *Scots : the mither tongue*. Edinburgh.

Kay, B., Orr, M., Horne, A., Milne, C., Souter, J. C. & Tayside (Scotland). Regional Council. (1995). *Tae be yersels : investigating the scots language*. Dundee.

Lawson, R. (2014, janvier 22). *Sociolinguistics in Scotland*.

Macafee, C. & McGarrity, B. (1999). Scots language attitudes and language maintenance. Leeds studies in English. Récupérée à partir de http://sfx6.exlibrisgroup.com/sherbrooke/sfx%5C_local?issn=0075-8566%5C&ctx%5C_enc=info:ofi/enc:UTF-8%5C&ctx%5C_ver=Z39.88-2004%5C&rfr%5C_id=info:sid/sfxit.com:azlist%5C&sfx.ignore%5C_date%5C_threshold=1%5C&svc.fulltext=yes

Macdonald, K. (2010). *Irvine welsh about "trainspotting" (1995) filmed on hi-8*. Récupérée à partir de http://www.youtube.com/watch?v=NnHTJkDd2ZM&feature=youtube_gdata_player

MacLeod, I. & McClure, J. D. (2012). *Scotland in definition : a history of Scottish dictionaries*.

McClure, J. D., Aitken, A. J. & Low, J. T. (1980). *The scots language : planning for modern usage*. Edinburgh.

McPake, J. (2002). *Mapping the languages of edinburgh : final report*. Edimbourg.

Mountford, L. (2005). *Yer spraffin' mince ! the Leith slang dictionary online*. Récupérée à partir de http://www.le.ac.uk/ee/glossaries/leith/essay.htm

Murison, D. D. (1977). *The guid scots tongue*. Edinburgh.

Poulter, S. & Chorley, M. (2013). *Revealed : the language map of england which shows where up to 40% of people say english is not their mother tongue*. [Mail online]. Récupérée à partir de http://www.dailymail.co.uk/news/article-2270638/Revealed-The-language-map-England-shows-40-people-say-English-mother-tongue.html

Purves, D. & Saltire Society. (2002). *A scots grammar : scots grammar and usage*. Edinburgh.

Ramsay, A. (1737). *The gentle shepherd a scots pastoral comedy. by allan ramsay.* Récupérée à partir de http://galenet.galegroup.com/servlet/ECCO?c=1&stp=Author&ste=11&af=BN&ae=N032605&tiPG=1&dd=0&dc=flc&docNum=CW111777734&vrsn=1.0&srchtp=a&d4=0.33&n=10&SU=0LRK&locID=ureginalib

Ross, D. (2001). *Scottish place-names*. Edinburgh.

Russell Powell, F. (1996). *Dazed and confused–danny boyle interview*. Récupérée à partir de http://www.fionarussellpowell.com/interviews/dannyboyle.htm

Scoobie, J., Gordeeva, O. & Matthews, B. (2006). *Acquisition of scottish english phonology : an overview*. Récupérée à partir de http://eresearch.qmu.ac.uk/149/1/wp-7.pdf

Scotland. (2011). *Specimen questionnaire of scotland's 2011 census*.

Scotland. Scottish Government. (2010). *Curriculum for excellence building the curriculum 5 : a framework for assessment : understanding, applying and sharing standards in assessment for curriculum for excellence : quality assurance and moderation*. Edinburgh. Récupérée à partir de http://www.scotland.gov.uk/Publications/2010/10/12154100/0

Scotsman, T. (2006, mai 19). *Scots language, if you ken whit ah mean ?* Récupérée à partir de http://www.scotsman.com/news/scots-language-if-you-ken-whit-ah-mean-1-466495

Shepherd, W. R. (1956). *Historical atlas.* New York.

Shorto, R. (2007, août 22). Simpsons quotes enter new oxford dictionary. Récupérée à partir de http://www.telegraph.co.uk/news/uknews/1561004/Simpsons-quotes-enter-new-Oxford-dictionary.html

Skilling, M. (2012). *Edinburgh past and present*. Récupérée à partir de http://edinburghpastandpresent.com

Skretkowicz, V., Rennie, S., Craigie, W. A. & University of Dundee. (2001). *Dictionary of the scots language dictionar o the scots leid : dsl*. Récupérée à partir de http://www.dsl.ac.uk/dsl/

Stevenson, J. A. C. & Macleod, I. (1989). *Scoor-oot : a dictionary of scots words and phrases in current use*. London ; Atlantic Highlands, NJ.

Stuart-Smith, J. (1999). Glasgow : accent and voice quality.

Stuart-Smith, J. (2003). The phonology of modern urban scots.

Stuart-Smith, J., Timmins, C. & Tweedie, F. (2007). 'talkin' jockney' ? variation and change in glaswegian accent. *Journal of Sociolinguistics*, 11, 221–260. Récupérée à partir de http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9841.2007.00319.x/abstract

"Treehouse Of Horror V" The Simpsons. (1994).

Trudgill, P. (1984). *Language in the British Isles*. Cambridge ; New York.

Walker, H. (2007, décembre 3). Scottish mps voice concern over increase in anti-scottish sentiment. Récupérée à partir de http://www.journal-online.co.uk/article/2747-scottish-mps-voice-concern-over-increase-in-antiscottish-sentiment

Warrack, A. (1979). *Chambers's Scots dictionary : serving as a glossary for Ramsay, Fergusson, Burns, Scott, Galt, minor poets, kailyard novelists, and a host of other writers of the Scottish tongue*. Edinburgh.

Warrack, A. (2000). *The scots dialect dictionary*. New Lanark, Scotland.

Warrack, A. (2006). *The scots dialect dictionary*. New Lanark, Scotland.

Welsh, I. (2007). *Irvine welsh : in the past 15 years dublin has gone from being calcutta to paris*. [Independent.ie]. Récupérée à partir de http://www.independent.ie/life/travel/irvine-welsh-in-the-past-15-years-dublin-has-gone-from-being-calcutta-to-paris-26438333.html

Wikipedia : spellin an grammar. (2012, octobre 28). Page Version ID : 299864. Récupérée à partir de http://sco.wikipedia.org/w/index.php?title=Wikipedia:Spellin_an_grammar&oldid=299864

Numéralia

Anis, J. (1991). Texte et ordinateur : les mutations du lire-écrire.

Au, W. J. (2007, avril 22). *Move over MySpace, gaia online is here*. [GigaOm]. Récupérée à partir de http://gigaom.com/2007/04/22/move-over-myspace-gaia-online-is-here/

Baron, N. S. (2003). Language of the internet. *The Stanford handbook for language engineers*, 59–127. Récupérée à partir de http://www.american.edu/cas/lfs/faculty-docs/upload/n-baron-language-internet.pdf

Baron, N. S. (2005). Instant messaging by american college students a case study in computer-mediated communication. In American association for the advancement of science annual meeting. Récupérée à partir de http://www.angel.fenre.org/Year%202%20-%20Thesis/Lecture/Baron%20-%20Instant%20Messaging%20by%20American%20College%20Students.pdf

Barratt, M. J. (2011). Discussing illicit drugs in public internet forums : visibility, stigma, and pseudonymity. In Proceedings of the 5th international conference on communities and technologies (p. 159–168). C&#38 ;T '11. New York, NY, USA. Récupérée à partir de http://doi.acm.org/10.1145/2103354.2103376

Bays, H. & Encrevé, P. (2001). *Échanges conversationnels sur internet une analyse sociolinguistique d'un nouveau mode de communication*.

Bolichowski, J. (2012). *OMG ! kids conversant in both netspeak and english*. [Niagara falls review]. Récupérée à partir de http://www.niagarafallsreview.ca/2012/01/18/omg-kids-conversant-in-both-netspeak-and-english

Cavazza, F. [Fred]. (2006a). *Qu'est-ce que l'identité numérique ?* [Fredcavazza.net]. Récupérée à partir de http://www.fredcavazza.net/2006/10/22/qu-est-ce-que-l-identite-numerique/

Cavazza, F. [Fred]. (2006b). *Qu'est-ce que l'identité numérique ?* [Fredcavazza.net]. Récupérée à partir de http://www.fredcavazza.net/2006/10/22/qu-est-ce-que-l-identite-numerique/

Cavazza, F. [Frédéric]. (2014). *Panorama des médias sociaux 2014*. [MediasSociaux.fr]. Récupérée à partir de http://www.mediassociaux.fr/2014/05/19/panorama-medias-sociaux-2014/

Coffman, K. & Odlyzko, A. (1998). *The size and growth rate of the internet*.

Crystal, D. (2001). *Language and the Internet*. Cambridge, UK ; New York.

Crystal, D. (2011). *Internet linguistics : a student guide*. Abingdon, Oxon ; New York, NY.

Deorowicz, S. & Ciura, M. G. (2005). Correcting spelling errors by modelling their causes. *International journal of applied mathematics and computer science*, 15, 275. Récupérée à partir de http://sun.aei.polsl.pl/~mciura/publikacje/spelling.pdf

Deseilligny, O. (2012). La mémoire appareillée : dispositifs numériques et écriture de soi. *ESSACHESS – Journal for Communication Studies*, 95–105.

DiNucci, D. (1999). Design & new media : fragmented future. *New York*, 53(4), 32–35.

Don & Brad. (2008). *Skitt's law*. Récupérée à partir de http://knowyourmeme.com/memes/skitts-law

Donath, J. (1995). *Identity and deception in the virtual community*. Récupérée à partir de http://www.academia.edu/554594/Identity_and_deception_in_the_virtual_community

Goudet, L. G. (2011). Graphic conventions on niche websites : sedimentation of alternative spelling. *Kosiče : Skase*, Language, Literature and Culture In Present-Day Context : Contemporary Research Perspectives In Anglophone PhD Studies.

Goudet, L. G. (2012). Typologie des usages langagiers à partir de trois communautés virtuelles. In Actes du colloque des doctorants de l'école doctorale ERASME à l'occasion du 40ème anniversaire de paris 13-nord.

Goudet, L. G. (2013a). *Alternative spelling and censorship : the treatment of profanities in virtual communities*. In Aspects of linguistic impoliteness (Cambridge Scholars Publishing).

Haraway, D. J. (2009). *A cyborg manifesto : science, technology, and socialist-feminism in the late twentieth century*.

Herring, S. C. (1996). *Computer-mediated communication : linguistic, social, and cross-cultural perspectives*. Amsterdam ; Philadelphia.

Ippolita & Milani, C. (2013, septembre 1). *J'aime pas Facebook*.

Jouxtel, P. (2009). *Société francophone de mémétique –mémétique, tentative de définition par un méméticien*. Récupérée à partir de http://www.memetique.org/2009/01/memetique-tentative-de-definition-par-un-memeticien/

Kahle, B. (1996). *Internet archive*. Récupérée à partir de https://archive.org/

Keegan, K. M. (2006). *''i submitted this with a funnier headline" multiliteracies and communities of practice at fark.com*. Récupérée à partir de http://proquest.umi.com/pqdweb?did=1192196931&sid=14&Fmt=2&clientId=39334&RQT=309&VName=PQD

Lee, J. (2012). *How we talk online : a history of online forums, from cavemen days to the present*. Récupérée à partir de http://www.makeuseof.com/tag/how-we-talk-online-a-history-of-online-forums-from-cavemen-days-to-the-present/

Lievrouw, L. A. & Livingstone, S. M. (2002). *Handbook of new media : social shaping and consequences of ICTs*. London ; Thousand Oaks, Calif.

Marcoccia, M. (2004, novembre 1). L'analyse conversationnelle des forums de discussion : questionnements méthodologiques. *Les Carnets du Cediscor. Publication du Centre de recherches sur la didacticité des discours ordinaires*, (8), 23–37. Récupérée à partir de http://cediscor.revues.org/220

McFedries, P. & Logophilia Limited. (1995). *The word spy*. Récupérée à partir de http://www.wordspy.com/

Merzeau, L. (2012). Les enjeux de l'identité numérique. In Acte de la journée académique des professeurs documentalistes de l'académie de rouen, 5 décembre 2012. Université de Rouen. Récupérée à partir de http://eduscol.education.fr/cdi/actualites/LM-jpro-rouen-idnum

Nowson, S. (2006). *The Language of Weblogs : a Study of Genre and Individual Differences*. Récupérée à partir de http://hdl.handle.net/1842/1113

Papi, C. (2012). Les « digital natives », un public particulier ? *EUTIC'*, Publics et pratiques médiatiques.

Paveau, M.-A. (2012, juillet 28). *Ah. un "vrai" livre numérique ?* [Technologies discursives]. Récupérée à partir de http://technodiscours.hypotheses.org/189

Paveau, M.-A. (2013a, mai 10). *Analyse discursive des réseaux sociaux numériques [dictionnaire]*. [Technologies discursives]. Récupérée à partir de http://technodiscours.hypotheses.org/431

Paveau, M.-A. (2013b, février 4). *Environnement [dictionnaire]*. [Technologies discursives]. Récupérée à partir de http://technodiscours.hypotheses.org/311

Paveau, M.-A. (2013c). Technodiscursivités natives sur twitter. une écologie du discours numérique. Récupérée à partir de http://hal.archives-ouvertes.fr/hal-00859064

Paveau, M.-A. (2013d). *Technologie discursive [dictionnaire]*. [Technologies discursives]. Récupérée à partir de http://technodiscours.hypotheses.org/277

Preston, A. (2014, août 3). The death of privacy. *The Guardian*, Récupérée à partir de http://www.theguardian.com/world/2014/aug/03/internet-death-privacy-google-facebook-alex-preston

Pullum, G. K. (2004). *Snowclones : lexicographical dating to the second*. [Language log].

Rourke, L., Anderson, T., Garrison, D. R. & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14, 50–71.

Urban Dictionary. (2003). Récupérée à partir de http://www.urbandictionary.com/

Wenger, E. (1998). *Communities of practice : learning, meaning, and identity*. Cambridge, U.K. ; New York.

Wenger-Trayner, E. (2012). *Communities of practice : a brief introduction*. Récupérée à partir de http://wenger-trayner.com/theory/

Autres références

Alexa - The Web Information Company. (2000). Récupérée à partir de http://www.alexa.com/

Atwood, J. & Spolsky, J. (2008). *Stack overflow*. Récupérée à partir de http://stackoverflow.com/

Bickerstaffe, I. (1775). *The padlock : a comic opera. as it is performed by his majesty's servants, at the theatre-royal in drury-lane*. Dublin.

Blake, W. (1977). *Songs of experience*. (Date de première publication 1789)

Body Count. (1992). New York, NY.

Bradbury, R. (1965). *The golden apples of the sun*. London.

Chouine, H. (2013). *Ma vie sexuelle est en mode avion.* [Twitter]. Récupérée à partir de https://fr.twitter.com/Hervetechouine/status/300192888933974016

Dawkins, R. (1976). *The selfish gene*. Oxford ; New York.

de Clermont-Ferrand, A. (2011). *La leçon d'écriture*. Récupérée à partir de http://ien.aurillac3.free.fr/spip.php?article73

Deleuze, G. (1972). *L'anti-OEdipe*. Paris.

Descombes, V. (2013). *Les embarras de l'identité*. Paris.

Doggystyle. (2001). Beverly Hills, CA.

Eminem. (2002). *Lose yourself*. Germany.

Fat Fingers. (2004). Récupérée à partir de http://www.fatfingers.co.uk/

Frege, G. (1892). *Über sinn und bedeutung*. Récupérée à partir de http://www.gavagai.de/HHP31.htm

Golding, P. & Bailey, A. (1999). *Itc guidance on standards for subtitling*. Récupérée à partir de http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/standards_for_subtitling/index.asp.html

Government of Canada, S. C. (2007, mars 13). *Statistics canada : 2006 community profiles*. Récupérée à partir de http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-591/details/page.cfm?Lang=E&Geo1=CSD&Code1=1209034&Geo2=PR&Code2=12&Data=Count&SearchText=halifax&SearchType=Begins&SearchPR=01&B1=All&Custom=

Graffiti. (2009). New York, N.Y.

Gregorio, M. D. C. (2006, mai 1). *L'enseignement de la lecture et de l'écriture au portugal : (1850-1974) - trois facettes d'un rituel*.

Hughes, L. (1959). *Selected poems.* New York.

Invasion of the Gabber Robots. (2000). YouTube.com. Récupérée à partir de http://www.youtube.com/watch?v=XRzBVv2CSE0&feature=youtube_gdata_player

Keane, A. H. (1895). *Africa*. London.

Kelis. (2004). *Milkshake*. England.

Kincaid, J. P., Fishburne, J., Robert P, Rogers, R. L. & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel.* Ft. Belvoir.

Kripke, S. A. (1980). *Naming and necessity*. Cambridge, Mass.

Kyanka, R. C. (1999). *Something awful*. Récupérée à partir de http://www.somethingawful.com/

Levitt, S. D. & Dubner, S. J. (2005). *Freakonomics : a rogue economist explores the hidden side of everything*. New York.

Lewand, R. (2000). *Cryptographical mathematics*.

Malda, R. '. & Bates, J. (1997). *Slashdot*. [News for nerds. stuff that matters.]. Récupérée à partir de http://slashdot.org

Needham, N. (1990). *Internet movie database*. Récupérée à partir de http://www.imdb.com/

of America, S. S. (2014). *Popular baby names*. Récupérée à partir de http://www.ssa.gov/cgi-bin/babyname.cgi

Ontario. (2006). *Guide d'enseignement efficace de l'écriture de la maternelle à la troisième année*. Récupérée à partir de http://www.atelier.on.ca/edu/resources/guides/GEE_Ecriture_M_3.pdf

Paul, S. (2003). *Get busy*. Australia.

Paul, S. (2005). *We be burning*. Récupérée à partir de http://www.youtube.com/watch?v=uUGFhVWPOFs&feature=youtube_gdata_player

Roche, X. (2005). *Httrack website copier*. Logiciel. Version 3.48-19. Récupérée à partir de http://www.httrack.com/

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom : teacher expectation and pupils' intellectual development*. New York.

Schneider, E. W. [Edgar W.]. (2004). *A handbook of varieties of english vol. 1, phonology*. Berlin.

Scotireland. (2010). Récupérée à partir de http://tvtropes.org/pmwiki/pmwiki.php/Main/Scotireland

Sellar, W. C., Yeatman, R. J. & Reynolds, J. (1930). *1066 and all that : a memorable history of england, comprising all the parts you can remember, including 103 good things, 5 bad kings and 2 genuine dates*. London.

Sony Computer Entertainment. (1997). *Parappa the rapper*.

Spears, B. (2007). *Blackout*. New York, New York.

Stewart, L. (1999). *Essays about computer security*. Cambridge, Grande-Bretagne. Récupérée à partir de http://www.cl.cam.ac.uk/~mgk25/lee-essays.pdf

Still a g thang. (1998). Los Angeles, CA.

Strauss, L. (1953). *Natural right and history*. Chicago.

Take 2 Interactive. (2000). *Grand theft auto*.

The big picture. (2000). Los Angeles, CA.

The E.N.D. (2009). Santa Monica, Calif.

University of Texas System, T. E. A. (2009). *Alphabetic understanding, phonics, and spelling*. Austin, Texas. Récupérée à partir de http://www.esc19.net/documents/programs/alphabetic_understanding_phonics_and_spelling.pdf

Vivendi Universal Games. (2005). *50 cent bulletproof*. Fresco, CA.

Wales, J. & Sanger, L. (2001). *Wikipedia*. Récupérée à partir de http://www.wikipedia.org/

Walker, P. & Walker, D. (2012). *Old school vs. new school hip hop*. [Empower network]. Récupérée à partir de http://www.empowernetwork.com/freedombacknow/blog/old-school-vs-new-school-hip-hop/

Wasserman, L. (2004). *All of statistics : a concise course in statistical inference*. New York.

We are Young Money. (2009). New York.

Will.i.am, Simons, E., Spears, B., Brown, C., Cyrus, M. & 2NE1 (Musical group). (2013). *#WILLPOWER*. Santa Monica, Californie.

Williams, J. R. (1998, octobre 1). Guidelines for the use of multimedia in instruction. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(20), 1447–1451. Récupérée à partir de http://pro.sagepub.com/content/42/20/1447

Zero wing. (1991). Japon.

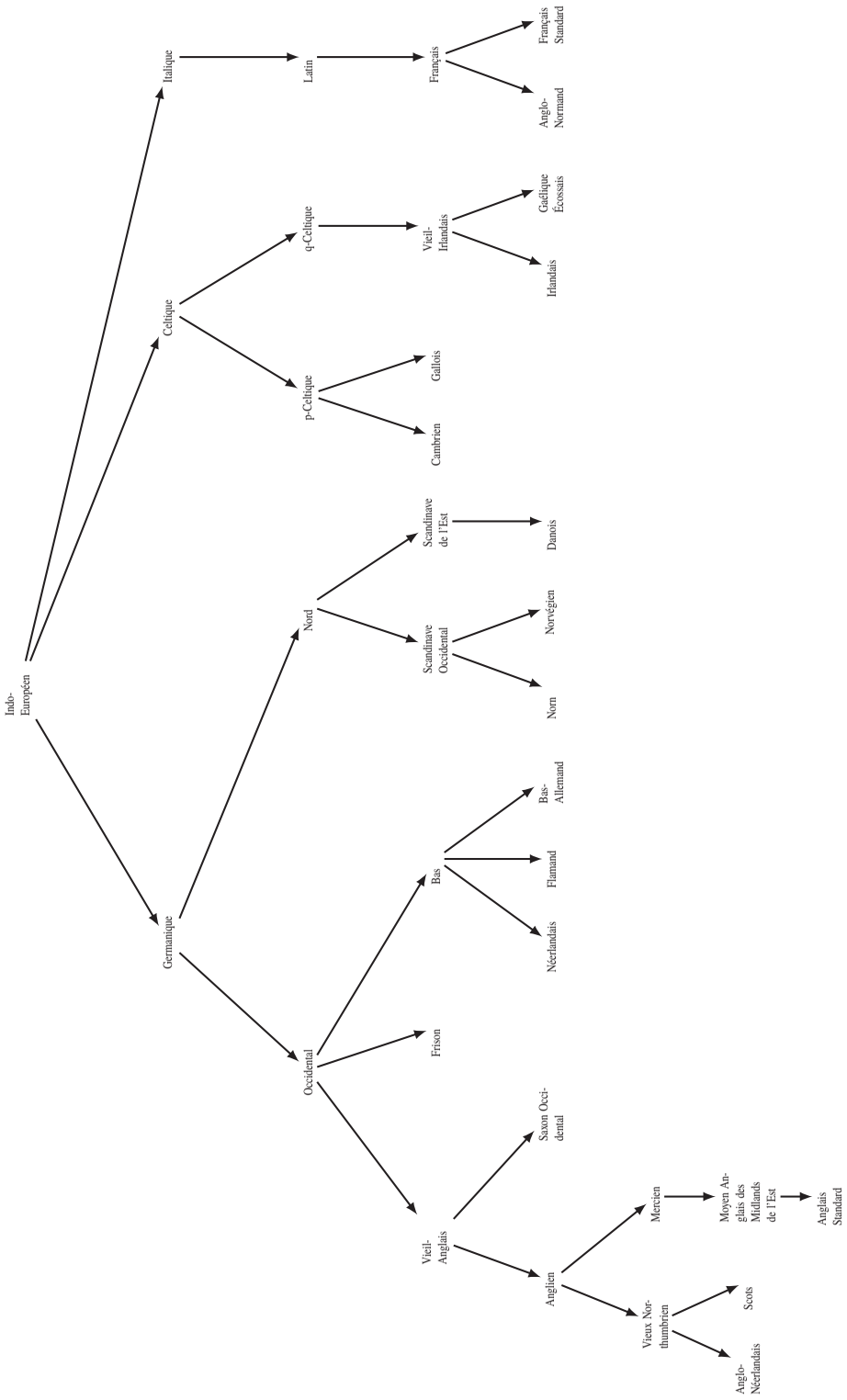
Zuckerberg, M., Saverin, E., McCollum, A., Moskovitz, D. & Hughes, C. (2004). *Facebook*. Récupérée à partir de https://www.facebook.com

Annexe A

Annexe

A.1 Arbre de parenté des langues

Ce tableau a été fait d'après Skretkowicz et al., 2001.


Indo-Européen
Italique
Latin
Français
Français Standard
Anglo-Normand
Celtique
q-Celtique
Vieil-Irlandais
Gaélique Écossais
Irlandais
p-Celtique
Gallois
Cambrien
Germanique
Nord
Scandinave de l'Est
Danois
Scandinave Occidental
Norvégien
Norn
Occidental
Bas
Bas-Allemand
Flamand
Néerlandais
Frison
Vieil-Anglais
Saxon Occidental
Anglien
Mercien
Moyen Anglais des Midlands de l'Est
Anglais Standard
Vieux Northumbrien
Scots
Anglo-Néerlandais

A.2 Precious

A.2.1 Inventaire des équivalences graphophonématiques des voyelles simples

Lettre	Son représenté	Exemple	Transcription phonétique (LPD)	Orthographe
<a>	/ɑː/	d**a**ter	/ˈdɑːt̬ər/	daughter
	/æ/	d**a**t	/ˈðæt/	that
<e>	/iː/	t**e**cher	/ˈtiːʧɚ/	teacher
	/e/	**e**rly	//	early
	/ɝː/	l**e**rn	/ˈlɝːn/	learn
	/ə/	ign**e**r**e**nt	/ˈɪgnərənt/	ignorant
<i>	/ɪ/	pr**i**tty	/ˈprɪti/	pretty
	/e/	forg**i**t	/fərˈget/	forget
	/ə~ʊ/	thankf**i**ll	/ˈθæŋkfʊl/	thankful
<o>	/ɑː/	n**o**ddin	/ˈnɑːdɪŋ/	nodding
	/ɒ/	c**o**ncreet	/ˈkɒŋkriːt/	concrete
	/ʌ/	h**o**ngry	/ˈhʌŋgri/	hungry
	/ɔː/	m**o**r	/ˈmɔːr/	more
<u>	/uː/	impr**u**v	/ɪmˈpruːv/	improve
	/ʊ/	g**u**d	/ˈgʊd/	good
	/ə/	app**u**l	/ˈæpəl/	apple
	/ɝː/	all-t**u**r-nah-tive	/ɔːlˈtɝːnət̬ɪv/	alternative
	/ʌ/	s**u**mebuddy	/ˈsʌmbədi/	somebody
	/ɑː/	anyb**u**ddy	/ˈenibɑːdi/	anybody
	/ɒ/	whutevr	/wɒt̬ˈevər/	whatever

Tab. A.1 – Équivalences phonographématiques des voyelles simples <a, e, i, o> dans les erreurs graphiques de Precious

A.2.2 Entrées du journal de Precious dans *Push*

Les données avec les corrections de Blu Rain, ainsi que la page de l'édition de travail sont disponibles dans la clef USB d'annexes.

A.2.3 Casting du film

Nom du personnage	*Acteur*	*Provenance*
Precious	Gabourey Sidibé	Harlem, New York
Mary	Mo'Nique	Baltimore, Maryland
Blu Rain	Paula Patton	Los Angeles, Californie
Cornrows	Sherri Shepherd	Chicago, Illinois
Rhonda	Chyna Layne	Jamaïque et New York
Rita	Stephanie Andujar	Manhattan, New York
Jermaine	Amina Robinson	Philadelphie, Pennsylvanie
Joann	Xosha Roquemore	Los Angeles, Californie
Consuelo	Angelic Zambrana	New York
Mrs Weiss	Mariah Carey	Long Island, New York
Nurse John	Lenny Kravitz	Manhattan, New York

Tab. A.2 – Casting de Precious et provenance des acteurs

A.3 Trainspotting

A.3.1 Casting du film

Seuls les personnages principaux du film sont cités dans cette liste. Lorsque les acteurs ont déménagé pendant leur enfance ou leur adolescence, ils sont marqués sur la carte à l'endroit où ils ont passé le plus de temps (informations trouvées sur leurs biographies en ligne, sur l'IMDB ou dans leurs CV en ligne).

Nom du personnage	*Acteur*	*Provenance*
Renton	Ewan McGregor	Crieff – Perth and Kinross, Écosse
Spud	Ewen Bremner	Édimbourg, Écosse
Sick Boy	Johnny Lee Miller	Londres, Angleterre
Tommy	Kevin McKidd	Elgin – Moray, Écosse
Begbie	Robert Carlyle	Glasgow, Écosse
Dianne	Kelly Macdonald	Glasgow, Écosse
Swanney / Mother Superior	Peter Mullan	Peterhead, Aberdeenshire ; Glasgow, Écosse

Tab. A.3 – Casting de Trainspottinget provenance des acteurs

Illustration A.1 – Carte de l'origine géographique des acteurs de Trainspotting

A.4 Classes d'isophones - Récapitulation

Ce tableau est une récapitulation des informations trouvées dans les travaux de J. Stuart-Smith et de B. Kortmann dans leur description du scots et de l'AAV. Les informations sur les classes d'isophones sont citées dans le chapitre deux, mais ce tableau résume et compare les classes d'isophones ensemble, entre l'afro-américain vernaculaire et le GA, ainsi que le scots urbain et le RP.

Scots Urbain	RP	Mot-clef	GA	AAV
ë ~ë ~ʌ ~ɪ	ɪ	KIT	ɪ	/iə/, ~/iɪ/
ɛ	e	DRESS	ɛ	/ɛɪ/
a	æ	TRAP	æ	/æɛ/
o ~ɔ	ɒ	LOT	ɑ	/ɑ/
ä	ʌ	STRUT	ʌ	/ʌ/
ë ~u ~ʏ	ʊ	FOOT	ʊ	/ʊ/
a	ɑː	BATH	æ	/æɛ/
o ~ɔ	ɒ	CLOTH	ɔ	/ɑ/
ä~ë	ɜː	NURSE	ɜr	/ɣ/, /ɻ/
i	iː	FLEECE	i	/ɪɪ/
e	eɪ	FACE	eɪ	/eɪ/
ạ	ɑː	PALM	ɑ	/æ/
ɔ	ɔː	THOUGHT	ɔ	/ɔʊ/
o	əʊ	GOAT	o	/o/
u ~ʏ	uː	GOOSE	u	/u/
əi	aɪ	PRICE	aɪ	/ɑː/
ɔe	ɔɪ	CHOICE	ɔɪ	/oɪ/
u ~ʌu	aʊ	MOUTH	aʊ	/æə/
i	ɪə	NEAR	ɪr	/ɣ/, /ɻ/
e ~ę	ɛə	SQUARE	ɛr	/æə/
ę ~ạ	ɑː	START	ɑr	/ɑː/
o ~ɔ	ɔː	NORTH	ɔr	/ɔə/
o	ɔː	FORCE	or	/ɔə/
ju	ʊə	CURE	ʊr	/ʊə/
e~ë	[i]	HAPPY	[i]	i
ä	ɚ	LETTER	/ər, ɚ/	ə, ɚ
ä	ə	COMMA	ə	ə

TAB. A.4 – Voyelles de l'AAV et du scots, comparées aux classes d'isophones du GA et du RP

A.5 Fréquence des lettres en anglais

Cette illustration est dans le domaine public, et a été réalisée par Pavel Mička. Il s'inspire des travaux de cryptographie sur l'anglais pour tirer cette visualisation (LEWAND, 2000). En anglais, l'espace

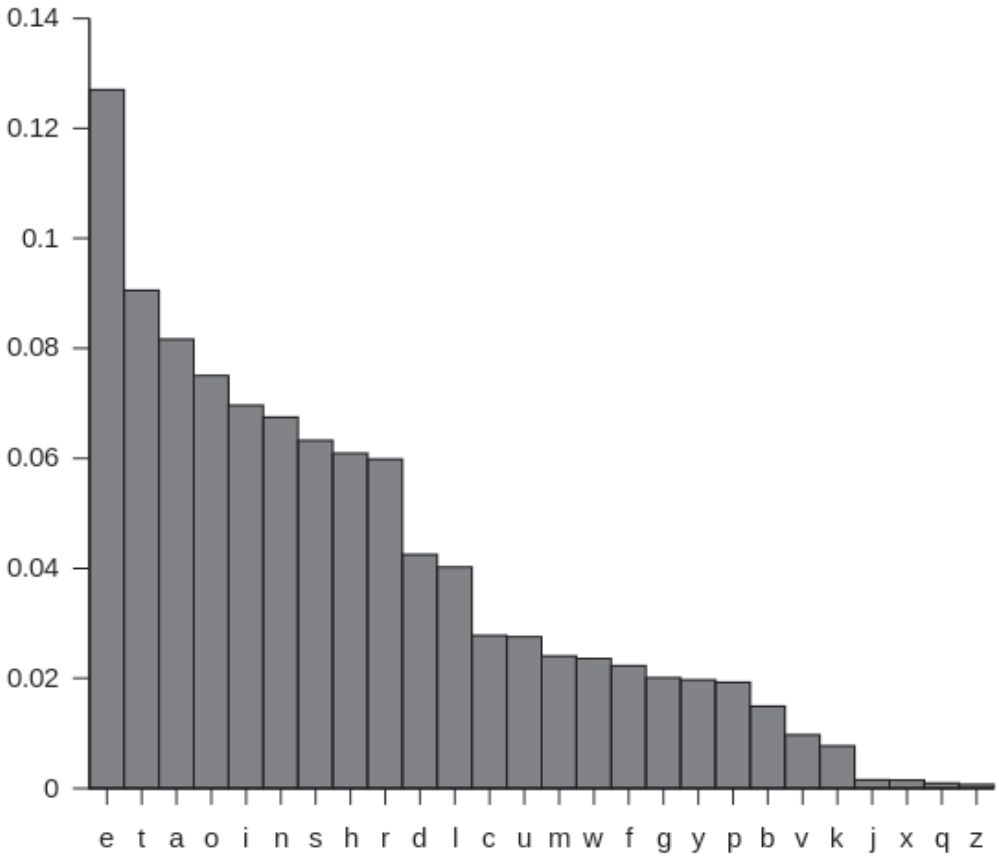


*Illustration A.2 – Fréquence des lettres en langue anglaise (*Lewand, *2000)*

est plus fréquent que la lettre <e> [1], et les caractères extra-alphabétiques, tous confondus, devraient occuper la quatrième place du tableau (Stewart, 1999).

A.6 *Scotster* et *Black Planet*

A.6.1 *Scotster*

A.6.1.1 Welcome to Scotster

Le texte qui suit est le premier message envoyé aux utilisateurs de SC dès leur inscription. C'est le manifeste de la communauté :

1. Source : <http ://www.data-compression.com/english.html> (Dernier accès le 20 juin 2014.

Hello !
Thank you for becoming a Scotster :)

You're now part of the world's leading social network for Scotland, dedicated to bringing people together online and through local events. Please do use our Tell A Friend feature to recommend your friends to help build Scotster in your local area. This is our website and the main way in which most of our Members keep up-to-date and stay in touch.

WHY IS SCOTSTER DIFFERENT ?

Our approach leads the way in social networking, successfully combining the web with the real world - whereas other sites may be about connecting you to your friends and family, we do that and more : we help you to expand your social circle and to meet new people on your wavelength. This is the new way to get to know people from the internet and you're far more likely to develop long-term friendships and great relationships.

Sending messages on Scotster is fast and efficient and we show you who is online right now so that you can start chatting.

Link up with your Scotster Friends and keep your contacts up-to-date with what you're doing and how you're feeling using the 'Today I…' box on the home page. You're about to discover a brand new way to connect with your friends and to meet people who share your interests and enjoy doing the same things you do.

SOME QUICK TIPS :

* Remember that our major difference is our approach to introducing you to other people - so move on from traditional sites and get up and do something different ! Explore the 'Events' area to get involved or start uploading your photos and sharing experiences.

* To update your profile and photos, or to manage anything that's specific to your Membership or the ways in which you use Scotster, go to 'My Stuff' on the main menu. Go to the 'Settings' menu and check that everything's set up correctly for you.

* Browse the 'People' area from the main menu or use our great search features that make finding stuff easy !

* Learn about tags - they're just key words or short phrases that help us find people, events, chat topics and even photos in which you may be interested.

* More features are coming every week. Scotster really is brand new and we've got lots more stuff to show you.

BUT RIGHT NOW…

Chat to more people in a relaxed environment and see the fantastic benefits of becoming a Scotster. It will feel great keeping in touch and making new friends in your area and you'll get to hear more about the stuff that everyone's doing so that you can join in !
Your next stop should be 'My Stuff' on the main menu then 'Edit Profile' - update your details, writings and start uploading your photographs :)

Very best wishes,
Jason, Chief Scotster.

A.6.1.2 Graphies alternatives - SC

Les graphies alternatives étudiées dans cette thèse et dans SC sont trop nombreuses et volumineuses pour être imprimées. Elles sont disponibles sur la clef USB d'annexes, avec les adresses des

fils de discussion de SC ainsi que des pages supplémentaires convoqués dans l'analyse, comme les fils hybrides du *Pulse*, à mi-chemin entre statut personnel à la FB et article à part entière. Tous les liens hypertextes sont actifs, au 20 juin 2014, et permettent de vérifier le CEC.

A.6.1.3 Architecture de SC

Le tableau récapitulant les indices alphanumériques utilisés pour identifier les sections et sous-sections du forum SC se trouve sur la clef USB d'annexes.

A.6.2 *Black Planet*

A.6.2.1 Architecture de BP

Le tableau récapitulant les indices alphanumériques utilisés pour identifier les sections et sous-sections du forum BP.

A.6.3 Graphies alternatives - BP

Les graphies alternatives étudiées dans cette thèse et dans BP sont trop nombreuses et volumineuses pour être imprimées. Elles sont disponibles sur la clef USB d'annexes, dans un document qui regroupe l'architecturation des fils de discussion de BP, les éléments supplémentaires convoqués dans cette analyse, comme les adresses des pages personnelles. Cette liste sert donc à vérifier le CEC virtuel, directement sur internet. Les liens ont été vérifiés au 20 juin 2014.

A.7 Liste de lieux virtuels évoqués dans cette thèse

Le but de cette liste de lieux virtuels est d'expliquer aux lecteurs les moins férus d'internet les différences entre les types de lieux virtuels, en complément du chapitre trois, ainsi que de donner une brève définition de chaque concept. La liste à proprement parler contient les noms des sites les plus proéminents ou les plus connus, par nombre d'utilisateurs, des lieux virtuels qui m'ont paru les plus importants dans la rédaction de cette thèse.

Tout d'abord, voici le bilan des fréquentations des médias sociaux sur internet, par Fred Cavazza : (Frédéric Cavazza, 2014)

Illustration A.3 – Panorama des réseaux sociaux

Wikis, ressources générales : Presque toutes ces ressources utilisent un modèle de Wiki, du nom de *Wikipedia* : ce sont des sites collaboratifs, dont les membres se corrigent les uns les autres. Ceci a pour but d'assurer la qualité des articles. L'*Encyclopaedia Dramatica* est un exemple de Wiki très subversif, à l'humour provocant (le trolling et le sarcasme y sont omniprésents), qui semble curaté par des membres appartenant à la culture internet des Anonymous et de *4Chan*. L'*Urban Dictionary* se caractérise par la création parallèle d'entrées pour une même définition, contrairement à un dictionnaire classique. Les utilisateurs votent pour ou contre une entrée, ce qui change son classement général. Les mots qui y sont définis appartiennent soit à l'argot anglo-saxon, soit à la culture internet.

Forums et *Message Boards* : Les forums et *message boards* principaux utilisés dans cette thèse font l'objet de parties à part entière, mais je caractérise ici *4Chan* et *Something Awful*, qui ont contribué à cette étude. *4Chan* est une plateforme créée en 2003, qui a commencé en tant qu'*image board*. C'est un lieu virtuel où le groupe de pirates et activistes (les « hacktivistes ») Anonymous a vu le jour. 4Chan est l'un des principaux créateurs de mèmes et de pratiques culturelles internet, comme le *rick-rolling*, qui consiste à placer un lien vers une vidéo du clip de « Never gonna give you up » de Rick Astley à la place d'un lien potentiellement intéressant. Certaines pratiques linguistiques issues de ce site ont été décrite dans cette thèse (les histoires en *green-text*…). *Something Awful* est légèrement moins important, mais ce site de nouvelles, de critiques assorti d'un forum gigantesque et à l'inscription payante est historiographiquement essentiel pour comprendre la culture internet. De nombreux mèmes et éléments culturels internet comme Slenderman (figure mythologique et inquiétante d'un homme très mince, aux bras tentaculaires) y ont été créés.

Réseaux Sociaux : J'utilise ici la définition large des réseaux sociaux. *Facebook* (création : 2004) est construit autour de l'individu qui crée son profil, et peut, ensuite, interagir avec ses « amis ». La communication passe par des messages, qui peuvent être des liens externes, des images, ou des messages textuels, postés sur son « mur » ou celui de ses connaissances. *Facebook* a des groupes d'utilisateurs qui s'agrègent autour d'intérêts communs, et surtout, un système de

valorisation des messages (bouton « Like »). C'est le réseau social le plus fréquenté, avec plus d'un milliard de profils actifs en septembre 2012. *MySpace* est un réseau social construit autour de la musique, et sert de tremplin aux jeunes artistes. Il est possible de répondre en commentaire aux profils. Sa spécificité l'a rendu fragile, et, depuis 2008, le nombre de ses profils ne cesse de diminuer au profit de *Facebook*. Il a 36 millions d'utilisateurs en septembre 2013. La plateforme *Skyblog* permet à une population plutôt jeune de créer un blog gratuitement. Les commentaires sur les messages étaient « monnayés » par les utilisateurs, avec des formules comme « lâche tes comm », et autres négociations (contre un commentaire posté, l'auteur du blog s'engageait à en envoyer dix sur celui de l'autre…). Enfin, *Twitter* est un réseau tourné vers la communication. Les messages y sont courts (140 caractères maximum) et consultables même pour les non-membres. Contrairement à Facebook, où les profils sont liés réciproquement, il est possible de « suivre » un membre sans qu'il ne suive l'autre en retour. Le réseau compte 274 millions d'utilisateurs actifs (juillet 2014).

Ressources linguistiques en langue française : Le Trésor de la Langue Française Informatisé, WikiLF et les publications gratuites et en ligne de la Délégation Générale de la Langue Française et des Langues de France (DGLFLF).

A.7.1 Données avec distances de Levenshtein

Les distances pondérées de Levenshtein sont présentées dans les annexes sur clef USB.

A.8 Résultats du test de Wilcoxon

Voici les résultats de la console de R, pour les deux corpus de graphies alternatives présentées dans le chapitre 5 (section 5.4.2) :

```
> wilcox.test(ScotCalc$ProportionSc, ScotCalc$ProportionNet, alternative='l', correct=FALSE,
exact=FALSE, paired=TRUE, conf.int=TRUE)
Wilcoxon signed rank test
data : ScotCalc$ProportionSc and ScotCalc$ProportionNet
V = 83, p-value = 3.349e-06
alternative hypothesis : true location shift is less than 0
95 percent confidence interval :
-Inf -3217.134
sample estimates :
(pseudo)median
-6349.581


  > wilcox.test(AAVCalc$ProportionBP, AAVCalc$ProportionNet, alternative='l',
paired=TRUE, conf.int=TRUE)
Wilcoxon signed rank test
data : AAVCalc$ProportionBP and AAVCalc$ProportionNet
V = 282, p-value = 0.003558
alternative hypothesis : true location shift is less than 0
95 percent confidence interval :
-Inf -43.39062
sample estimates :
(pseudo)median
-109.988
```

A.9 Codes Informatiques

Les codes informatiques qui ont servi à extraire des données ou à faire des remarques statistiques n'ont pas tous été écrits par l'auteur de ce travail. Certains ont été adaptés, et leurs sources sont mentionnées lorsque les macros ou les codes sont copiés. Les macros en VBA (Visual Basic Advanced) furent utilisées sous Excel 2013, et les codes sous R ont été écrites et/ou adaptées sous R Studio.

Le travail de dendrogrammes du chapitre 5 a été réalisé avec RExcel, qui implémente les menus de R Commander dans les documents Excel.

A.9.1 Macros en Visual Basic Advanced [Excel]

A.9.1.1 Macro pour le calcul pondéré de la distance de Levenshtein-Damerau sous Excel

Elle a été codée par l'utilisateur rgmatthes de Stack Overflow (Atwood & Spolsky, 2008) et est accessible à l'adresse <http ://stackoverflow.com/questions/13693149/weighted-damerau-levenshtein-in-vba> (dernier accès le 2 mai 2014) :

```
Function WeightedDL(source As String, target As String) As Double
   Dim deleteCost As Double
   Dim insertCost As Double
   Dim replaceCost As Double
   Dim swapCost As Double

   deleteCost = 1
   insertCost = 1.1
   replaceCost = 1.1
   swapCost = 1.2

   Dim i As Integer
   Dim j As Integer
   Dim k As Integer

   If Len(source) = 0 Then
```

```
    WeightedDL = Len(target) * insertCost
    Exit Function
End If

If Len(target) = 0 Then
    WeightedDL = Len(source) * deleteCost
    Exit Function
End If

Dim table() As Double
ReDim table(Len(source), Len(target))

Dim sourceIndexByCharacter() As Variant
ReDim sourceIndexByCharacter(0 To 1, 0 To Len(source) - 1) As Variant

If Left(source, 1) <> Left(target, 1) Then
    table(0, 0) = Application.Min(replaceCost, (deleteCost + insertCost))
End If

sourceIndexByCharacter(0, 0) = Left(source, 1)
sourceIndexByCharacter(1, 0) = 0

Dim deleteDistance As Double
Dim insertDistance As Double
Dim matchDistance As Double

For i = 1 To Len(source) - 1

    deleteDistance = table(i - 1, 0) + deleteCost
    insertDistance = ((i + 1) * deleteCost) + insertCost

    If Mid(source, i + 1, 1) = Left(target, 1) Then
        matchDistance = (i * deleteCost) + 0
    Else
        matchDistance = (i * deleteCost) + replaceCost
    End If

    table(i, 0) = Application.Min(Application.Min(deleteDistance, insertDistance), matchDistance)
Next

For j = 1 To Len(target) - 1

    deleteDistance = table(0, j - 1) + insertCost
    insertDistance = ((j + 1) * insertCost) + deleteCost

    If Left(source, 1) = Mid(target, j + 1, 1) Then
        matchDistance = (j * insertCost) + 0
```

```
    Else
        matchDistance = (j * insertCost) + replaceCost
    End If

    table(0, j) = Application.Min(Application.Min(deleteDistance, insertDistance), matchDistance)
Next

For i = 1 To Len(source) - 1

    Dim maxSourceLetterMatchIndex As Integer

    If Mid(source, i + 1, 1) = Left(target, 1) Then
        maxSourceLetterMatchIndex = 0
    Else
        maxSourceLetterMatchIndex = -1
    End If

    For j = 1 To Len(target) - 1

        Dim candidateSwapIndex As Integer
        candidateSwapIndex = -1

        For k = 0 To UBound(sourceIndexByCharacter, 2)
            If sourceIndexByCharacter(0, k) = Mid(target, j + 1, 1) Then candidateSwapIndex = sourceIndexByCharacter(1, k)
        Next

        Dim jSwap As Integer
        jSwap = maxSourceLetterMatchIndex

        deleteDistance = table(i - 1, j) + deleteCost
        insertDistance = table(i, j - 1) + insertCost
        matchDistance = table(i - 1, j - 1)

        If Mid(source, i + 1, 1) <> Mid(target, j + 1, 1) Then
            matchDistance = matchDistance + replaceCost
        Else
            maxSourceLetterMatchIndex = j
        End If

        Dim swapDistance As Double

        If candidateSwapIndex <> -1 And jSwap <> -1 Then

            Dim iSwap As Integer
            iSwap = candidateSwapIndex

            Dim preSwapCost
```

```
            If iSwap = 0 And jSwap = 0 Then
               preSwapCost = 0
            Else
               preSwapCost = table(Application.Max(0, iSwap - 1), Application.Max(0, jSwap - 1))
            End If

            swapDistance = preSwapCost + ((i - iSwap - 1) * deleteCost) + ((j - jSwap - 1) * insertCost) + swapCost

         Else
            swapDistance = 500
         End If

         table(i, j) = Application.Min(Application.Min(Application.Min(deleteDistance, insertDistance), matchDistance),
              swapDistance)

      Next

      sourceIndexByCharacter(0, i) = Mid(source, i + 1, 1)
      sourceIndexByCharacter(1, i) = i

   Next

   WeightedDL = table(Len(source) - 1, Len(target) - 1)

End Function
```

A.9.1.2 Macro pour le calcul de la proportion de similarité orthographique

Cette macro prend en charge une variante de la distance de Levenshtein et des correspondances partielles (*fuzzy match* ou *approximate match*) sous Excel. Elle a été codée par l'utilisateur Apostolos55 de Stack Overflow(Atwood & Spolsky, 2008) et est accessible à l'adresse <http ://stackoverflow.com/questions/424
distance-in-excel> (dernier accès le 2 mai 2014) :

```
' Levenshtein3 tweaked for UTLIMATE speed and CORRECT results
' Solution based on Longs
' Intermediate arrays holding Asc()make difference
' even Fixed length Arrays have impact on speed (small indeed)
' Levenshtein version 3 will return correct percentage
'
Function Levenshtein3(ByVal string1 As String, ByVal string2 As String) As Long

Dim i As Long, j As Long, string1_length As Long, string2_length As Long
```

```
Dim distance(0 To 60, 0 To 50) As Long, smStr1(1 To 60) As Long, smStr2(1 To 50) As Long
Dim min1 As Long, min2 As Long, min3 As Long, minmin As Long, MaxL As Long

string1_length = Len(string1) : string2_length = Len(string2)

distance(0, 0) = 0
For i = 1 To string1_length : distance(i, 0) = i : smStr1(i) = Asc(LCase(Mid$(string1, i, 1))) : Next
For j = 1 To string2_length : distance(0, j) = j : smStr2(j) = Asc(LCase(Mid$(string2, j, 1))) : Next
For i = 1 To string1_length
   For j = 1 To string2_length
      If smStr1(i) = smStr2(j) Then
         distance(i, j) = distance(i - 1, j - 1)
      Else
         min1 = distance(i - 1, j) + 1
         min2 = distance(i, j - 1) + 1
         min3 = distance(i - 1, j - 1) + 1
         If min2 < min1 Then
            If min2 < min3 Then minmin = min2 Else minmin = min3
         Else
            If min1 < min3 Then minmin = min1 Else minmin = min3
         End If
         distance(i, j) = minmin
      End If
   Next
Next

' Levenshtein3 will properly return a percent match (100%=exact) based on similarities and Lengths etc...
MaxL = string1_length : If string2_length > MaxL Then MaxL = string2_length
Levenshtein3 = 100 - CLng((distance(string1_length, string2_length) * 100) / MaxL)

End Function
```

A.10 Travaux universitaires réalisés au cours de cette thèse

Ces travaux sont classés suivant la nomenclature établie par l'AERES, dont les clefs seront données avant chaque partie.

A.10.1 Communications

— C-ACTI : Communications avec actes dans un congrès international

— C-ACTN : Communications avec actes dans un congrès national

— C-COM – Inter : Communications sans actes dans un congrès international

— C-COM – Nat : Communications sans actes dans un congrès national

C-COM – Nat *« Identity and self-description in African American communities online »* (22 avril 2010) Faculty and Graduate Student Colloquium, Université du Texas, Austin, États-Unis. Organisateurs : D. Bini, C. Blyth. Site internet : http ://www.utexas.edu/cola/depts/frenchitalian/events/14510

C-COM – Inter *"Grapho-phonemic representations of African American Vernacular English on the Internet"* (14 sep-tembre 2010), PAC Workshop "The phonology of English : usage, varieties and structure" Montpellier. Organisateurs : P. Carr et J. Durand. Site : http ://w3.pac.univ-tlse2.fr/

C-ACTN *« Recherche de propriétés linguistiques d'une communauté virtuelle : communaulectes et codifica-tions linguistiques »* (7 décembre 2010) Colloque des Doctorants de l'École Doctorale ERASME à l'occasion du $40me$ anniversaire de Paris 13-Nord. Villetaneuse

C-COM – Nat *« Communaulectes et codifications linguistiques »* (3 février 2011) : Doctoriales – Journée d'Études du Centre de Recherche Interculturelles sur les Domaines Anglophones et Francophones (CRIDAF), Uni-versité Paris 13-Nord, Villetaneuse. Organisatrice : C. Parfait.

C-ACTI *"Graphic conventions on niche websites : Sedimentation of alternative spelling"* (5 mai 2011), PhD conference, Kosiče, Slovaquie. Organisateurs : Prof. Dr. Pavel Štekauer et Dr. Slávka Tomaščíková.

C-ACTI *"Alternative spelling and censorship : the treatment of profanities in virtual communities"* (27 mai 2011), Linguistic Impoliteness & Rudeness in Communication and Society, Lyon. Organisateurs : D. Jamet et M. Jobert.

C-ACTI *"Internet representations of dialectal English"* (31 mars 2012), 16ème colloque sur l'anglais oral ALOES et CRIDAF "Multicultural Spoken English", Villetaneuse. Organisateurs : N. Ballier, Ch. Migette. Site : http ://www.univ-paris13.fr/cridaf/fr/activites-cridaf-2011/seminaires-colloques/activites-cridaf-2012/208-evenements-030112-cridaf.html

C-COM Nat – *« Les références culturelles à l'heure de la mémétique : Internet et les échanges culturels »* (6 novembre 2012) : Séminaire de l'Équipe de Recherche Interdisciplinaire sur les Aires Culturelles (ERIAC), Uni-versité de Rouen, Mont-Saint-Aignan. Organisatrices : F. Cabaret et O. Louiset. Site : http ://eriac.net/les-references-communes-a-lheure-de-la-memetique-internet-et-les-echanges-culturels/

C-COM Nat – *« Dialectologie grapho-phonématique de deux communautés virtuelles pour une approche discursive des communalectes »* (15 novembre 2012) : Séminaire de Master 2-Doctorat- Post Doctorat « Théorie du Texte et du Discours » de M.-A. Paveau. Université Paris 13-Nord, Villetaneuse. Organisatrice : M.-A. Paveau

C-COM Nat – *« Multimodalité du Corpus : Pistes de réflexion »* (8 février 2013) : Séminaire de Doctorat et Post-Doctorat « Discours et Doctorat » PRES Sorbonne Paris Cité (Universités Paris III, V et XIII). Université Paris V, Paris. Organisateurs : M.-A. Paveau, S. Moirand et P. Von Münchow et F. Mourlhon-Dallies

C-COM Nat – *« Aspects Méthodologiques de l'Exploitation d'un Corpus de Parole Dialectale Écrite sur Forums Internet »* dans la table ronde « Nouvelles Perspectives de Recherches à l'Heure du Web 2.0 » avec N. Marignier et A. Zineddine, dirigée par L. Goudet (18 Avril 2013). Colloque des doctorants de l'École Doctorale ERASME, Villetaneuse.

C-COM – Nat	*« Aller-Retour Conceptuel : Les Hip Hop Studies »* (26 avril 2013) : Séminaire de Doctorat et Post-Doctorat « Discours et Doctorat » PRES Sorbonne Paris Cité Cité (Universités Paris III, V et XIII). Université Paris V, Paris. Organisatrices : M.-A. Paveau, S. Moirand et P. Von Münchow et F. Mourlhon-Dallies
C-COM	*« Phonologie et Lexicologie « sauvages » :l'Accent Tag sur Youtube »* (17 mai 2013) Congrès 2013 de la Société des Anglicistes de l'Enseignement Supérieur. Dijon. Site : http ://saes2013.u-bourgogne.fr/
C-ACTI	*« « That foo just talks to much » : Légitimité et économie du pouvoir dans un forum de gang afro-américain »* (4 juillet 2013), Colloque International du Réseau Francophone de Sociolinguistique « Les locuteurs et les langues : pouvoirs, non-pouvoirs et contre-pouvoirs », Corte. Organisateurs : Romain Colonna. Site : http ://rfs2013.univ-corse.fr/
C-COM –Nat	« *"You niggaz is just like bitches" : Discours en ligne, genre et hip hop* » (2 décembre 2013), Séminaire de Master 2-Doctorat de Luca Greco « Langage(s), genre(s), sexualité(s) », ILPGA, Paris. Site : http ://www2.univ-paris8.fr/RING/spip.php ?article2884
C-COM	« Migrer vers LaTeX, utiliser Zotero et Biblatex/Biber » (17 décembre 2013), séminaire des doctorants de Nicolas Ballier, Université Paris-Diderot, Paris.
C-COM –Inter	« Alternative (poly)spellings on two English-speaking niche forums » (19 février 2014), dans le symposium « Web-native writings » dirigé par M.-A. Paveau, Congrès international « Writing Research Across Borders », Université Paris-Ouest Nanterre la Défense, Nanterre. Site : http ://conference2014.fr/wrab_accueil_en.php
C-COM	« LaTeXet le non-textuel » (18 juin 2014), séminaire des doctorants de Nicolas Ballier, Université Paris-Diderot, Paris.

A.10.2 Publications

Clef de l'AERES :

— ACL : Articles dans des revues internationales ou nationales avec comité de lecture répertoriées

— ACLN : Articles dans des revues avec comité de lecture non répertoriées

— ASCL : Articles dans des revues sans comité de lecture

— CHAP : Chapitres d'ouvrages scientifiques

ACLN	"Graphic conventions on niche websites : sedimentation of alternative spelling" in *Language, Literature and Culture in Present-Day Context : Contemporary Research Perspectives in Anglophone PhD Studies*. p. 25-37 (2011) Kosiče : Skase
CHAP	"Alternative spelling and censorship : the treatment of profanities in virtual communities" in *Aspects of Linguistic Impoliteness*. p. 209-222 (2013). Cambridge : Cambridge Scholars Publishing
ACLN	« Discours de gangs afro-américains sur internet : « calligraphie et stylistique dans le monologue de présentation personnelle ». *Itinéraires (Littérature, Textes, Cultures)* sous la direction de M.-A. Paveau [parution 2014]
ASCL	« Typologie des usages langagiers sur internet à partir de trois communautés virtuelles » *Actes du Col-loque des doctorants de l'École doctorale ERASME à l'occasion du 40me anniversaire de Paris 13-Nord*. Villetaneuse [parution 2015]
ASLC	"Internet Representations Of Dialectal English" *Actes du 16ème colloque sur l'anglais oral ALOES et CRIDAF "Multicultural Spoken English"* [en préparation, parution 2015]

A.10.3 Autres travaux

Clef de l'AERES :

— CORP : Corpus

— TRAD : Traductions

— AP : Autres productions : bases de données, logiciels, comptes rendus d'ouvrages, guides techniques, catalogues d'expositions, rapports intermédiaires de grands projets internationaux, etc.

TRAD	*The Reconstruction of Proto-Romance* (Robert A. Hall, 1950) avec S. Mougenot (2007) pour le Centre d'Études et de Recherche de l'Occitan, Université Paris IV-Sorbonne.
AP	Enregistrement de matériel pédagogique pour *Pause-Café*, un manuel de Français Langue Étrangère (2008) pour C. Blyth, N. Meghrabi, S. Pellet, et S. Foerster. McGraw-Hill Humanities Social eds
AP	*Mes Langues aux Chats*, Hypotheses.org (2012 ; messages remontant à 2006) http ://-lac.hypotheses.org
AP	Rédaction du manuel d'utilisation de *WinPitch*, un logiciel de phonétique (2013) avec Ph. Martin et V. Pouillon.
CORP	*CATSSAA : Corpus of the Accent Tag – Spoken Scots and African American* – (2013) [site web] Spoken Language Data Repository (SLDR), Université d'Aix-en-Provence. Adresse : http ://sldr.org/sldr000830/fr

j'ai également été dans le comité d'organisation des colloques suivants :

— « Spoken English and the Media », 16me colloque de l'ALOES à l'Université Paris 13-Sorbonne Paris Cité, Villetaneuse.

— « Données hybrides et contextualisation des corpus : nouvelles approches » (22-23 mai 2014), PLEIADE, Université Paris 13, Sorbonne Paris Cité ; EDA, Éducation, Discours, Apprentissage, EA 4071, Université Paris Descartes ; SYLED, Systèmes linguistiques, énonciation, discours, EA 2290, Université Sorbonne-Nouvelle, Paris 3 à l'Université Sorbonne-Nouvelle, Paris.

J'ai été encadrante d'un mémoire de Master 1 intitulé « A sociolinguistic analysis of African American Vernacular English : the expression of masculinity in rap music from 1993 to 2013 », par Audrey Tesson et soutenu le 16 septembre 2014 à l'Université de Rouen, Mont Saint-Aignan.

A.11 Utiliser la clef USB d'annexes

Certains éléments sont trop longs ou complexes pour être imprimés ; il s'agit de l'index général des fils de discussion de BP, ainsi que ceux de BP, par exemple, dont les clefs de lecture des citations sont expliquées dans le chapitre 3. Chaque forum est marqué par : son nom (SC ou BP), la section du forum (chiffre), la sous-section du forum (lettre ; après <Z>, l'alphabet reprend en <AA, AB> …), et enfin, par le numéro du fil de discussion. Les annexes sont données en clef USB et non sur CD-ROM parce que tous les ordinateurs ont des prises USB, mais les lecteurs de CD ne sont plus inclus sur tous

les ordinateurs, en particulier les ordinateurs portables.

Tous les documents présentés sur la clef USB sont lisibles sous n'importe quel ordinateur équipé d'*Acrobat Reader*. Tous les fichiers dans un autre format natif sont également disponibles en .pdf. Les données quantitatives (distances de Levenshtein…) sont données en .txt, séparées par des tabulations ou en .csv, séparées par des points-virgules. Elle contient évidemment la présente thèse, sous format .pdf, navigable grâce aux explications sur les hyperliens dans la partie . Les fichiers de travail en .tex et les illustrations sont également

Pour effectuer une recherche dans un des fichiers, pressez <Ctrl + F> et entrez le code que vous recherchez (attention, l'opération peut prendre quelques minutes). Les index ne contiennent pas d'hyperréférences.

La clef contient également des extraits sonores ou audio-visuels de *Precious* et *Trainspotting*, pour compléter l'étude de certains cas phonétiques dans le chapitre 1.

Index Nominum

50 Cent (artiste), 156, 363

Aitchison, J., 82
Aitken, A. J., 78, 374, 385
Allan, K., 281, 308
Anderson, C., 118
Anis, J., 174
Asante, M., 147
Auroux, S., 431

Bakhtine, M., 6
Ball, A., 118
Ballier, N., 220, 231, 232, 235
Baratz, J., 118
Baron, N., 257
Barratt, M., 219
Baugh, J., 115, 116
Bays, H., 189, 283
Bellama, D., 381
Benjamin, W., 286
Benveniste, E., 1
Big L (artiste), 152
Black Eyed Peas (artistes), 131
Black, I., 151
Blake, A., 160
Blake, N., 10
Blake, W., 128
Blevins, W., 350
Body Count (artiste), 149
Bolichowski, J., 252
Boyle, D., 54
Boèce, 94
Bradford, J., 124
Brower, K., 128
Brown, C. (artiste), 133
Brown, E., 80
Bulot, T., 249
Burns, R., 64, 161
Burov, I., 358

Calvet, L.-J., 120
Campbell, K., 4, 118, 124, 146
Carr, P., 350, 351
Catach, N., 75, 336, 349
Cavazza, F., 3, 219
Chomsky, N., 337
Cohen, G., 123
Connery, S. (artiste), 149
Corbett, J., 2, 5, 78, 101–103, 107
Costa, J., 4, 80, 159, 163, 340
Cotter, J., 125
Croft, W., 21
Crystal, D., 199, 283, 328, 357, 358, 431
Cukor-Avila, P., 133

Davis, M., 252
Dawkins, R., 251, 284
Dehaene, S., 25, 29
Deleuze, G., 126, 433
Deorowicz, S., 263, 267
Deschamps, A., 336, 350, 384
Descombes, V., 220
Deseilligny, O., 219
Dillard, J., 78, 137, 383
DiNucci, D., 175
Donath, J., 220
Durand, J., 2

E-40 (artiste), 355
Eagle, A., 68, 96, 99, 147, 151, 164, 255, 336
Eloy, J.-M., 101
Eminem (artiste), 301, 324

Fasold, R., 117
Frege, G., 1, 231
Freire, P., 14
Fry, E., 144

Gadet, F., 249
Gambier, I., 50
Giegerich, H., 336
Goffman, E., 193
Graphie démarcative, 14
Green, L., 77

Haraway, D., 7
Hemart, L., 26
Herring, S., 168
Hickey, R., 80
Holloway, J., 123
Hopper, P., 173
Hughes, L., 125
Hurston, Z., 147
Hyperexpressivité, 316
Hémart, L., 33

Ice T (artiste), 149

Jacewicz, E., 399
Jakobson, R., 187
Jay-Z (artiste), 231
Jones, C., 102
Jouxtel, P., 285

Kay, B., 78, 82, 83, 89, 111, 346
Keane, A., 119
Kearse, R., 142, 145, 146, 156
Keegan, K., 5
Kelis (artiste), 152
Kerbrat-Orecchioni, C., 303
Klinkenberg, J.-M., 349
Kortmann, B., 2, 112, 130, 137–139
Krahmer, B., 128
Kripke, S., 220, 239

Labov, W., 9, 120, 133, 242, 404
Ladefoged, P., 138
Lanehart, S., 137
Larson, K., 28
Lawson, R., 358
Lecercle, J.-J., 4, 147, 165
Levitt, S., 19
Lewand, R., 412
Liberman, M., 417
Lievrouw, L., 3
Lil Wayne (artiste), 132, 372

Macafee, C., 92
MacLeod, I., 144
Maingueneau, D., 3
Malcolm X, 119
Marcoccia, M., 3
Mauss, M., 172
Mazière, F., 96
McClure, D., 95, 99
McFedries, P., 282
McGregor, E. (artiste), 149
McPake, J., 88
Merzeau, L., 219
Miethaner, U., 2
Mountford, L., 64
Mufwene, S., 5, 121, 130
Murison, D., 80
Mème, 284

Nelly (artiste), 383
Neuman, Y., 370, 371
Nicki Minaj (artiste), 366
Nowson, S., 61
Nurse, L., 119

Ong, W., 95, 343

Papi, C., 252, 282
Paul, S. (artiste), 131, 147
Paveau, M.-A., 7, 43, 157, 161, 283, 294
Poulter, S., 85
Preston, D., 10
Preston, R., 14
Pullum, G., 302
Purves, D., 94, 100

Quirk, R., 25

Ramsay, A., 95
Rayner, K., 252
Richardson, E., 118, 146, 147, 235

Rickford, J., 115
Roach, P., 366
Robinson, C., 108
Rosenthal, R., 242
Rossari, C., 264
Rourke, L., 184, 250

Sapphire, 14, 16
Saussure, F., 20
Schneider, E., 2, 78, 115, 117
Scoobie, J., 108
Sellar, W., 163
Sherman, J., 124
Sidnell, J., 373
Skretkowicz, V., 148
Smitherman, G., 142, 145–147
Snoop Dogg (artiste), 131, 154, 155, 355, 356
Sorlin, S., 79
Soulja Boy (artiste), 356, 366
Spears, B. (artiste), 231
Stevenson, J., 144
Strauss, L., 218
Stuart-Smith, J., 78, 101, 108–110, 112, 113, 383
Suga Free (artiste), 413
Sullivan, C., 50
Switzer, K., 49
Sylvain, S., 121

Thomas, E., 387, 388
Too Phat (artiste), 356
Trovato, V., 14
Trudgill, P., 83, 343, 385

Walker, H., 150
Walker, P., 143
Warrack, A., 82, 144, 147
Wasserman, L., 39
Wells, J., 117, 138, 383
Welsh, I., 14, 54
Wenger, E., 172, 173, 180, 249
Will.i.am (artiste), 231
Williams, J., 72
Wise, C., 117, 119, 121
Wolfram, W., 78, 115, 137

Zimmer, B., 417
Zwicky, A., 338, 348, 354

Index Rerum

Bitch, 411
Ebonics, 78
Fuck, 417
Nigger, 405
Twitter, 172

Ambigramme, 171
Apostrophe d'excuse, 95, 424
Archigraphème, 349, 383

Bouma, 28, 29, 405, 407

Cockney, 150
Communauté de pratique, 173

Dialectologie perceptive, 149
Distance de Levenshtein, 403

E-bonics, 126
eWAVE (Electronic World Atlas of Varieties of English), 130

Gallo (langue), 89
Gallois (langue), 85
Gaélique écossais, 89
Grammar Nazi, 25, 74, 269
Graphie contournative, 430
Graphie discriminante, 99
Graphie démarcative, 425, 430
Graphies alternatives, 23
Graphies contournatives, 331
Graphies démarcatives, 331
Graphies erronées (fautes d'orthographe), 24
Graphies gabaritiques, 21
Grapho-phonématique, 13

Hyperexpressivité, 283, 300, 321, 332

Jamaïcain (langue), 16, 49

Leetspeak, 306, 350, 407
Leith (lieu), 64
Littéracie, 39, 252
Loi d'Aitken, 374, 385, 395
Loi de Skitt, 271

Majuscule, 42
Marque phonétique, 348
Mème, 253
Mémoire numérique, 173

Netspeak, 431
Négation, 134

Onomastique, 18

Paradoxe de l'observateur, 9
Parlécrit, 14
Point Godwin, 269
Principe consonantique, 23
Principe d'intelligibilité, 25
Principe de conservation, 28, 29
Prononciation, 26

Regraphie, 48

Schwa, 392
Score de Flesch-Kincaid, 39
Scouse, 150
Smiley, 171
Snowclone, 302, 354
Sous-titres, 50, 71
Standardisation graphique, 24

Test de Shapiro-Wilk, 400
Test de Wilcoxon, 400
Typobonics, 126, 362
Typoglycémie, 252

Résumé de la thèse

Cette thèse étudie la situation de deux langues minoritaires parlées dans des pays anglophones, le scots et l'afro-américain vernaculaire utilisés sur internet dans deux forums de niche, dont les scripteurs sont majoritairement afro-américains pour l'un, et s'intéressent à la culture écossaise pour le second.

Les deux disciplines linguistiques mobilisées sont la phonologie des langues minoritaires et de l'anglais, à travers l'étude de graphies alternatives, et l'analyse de discours. Ces discours asynchrones permettent des relevés inédits de l'apparition de nouvelles formes lexicales ou graphiques, qui sont propres à ces communautés. À l'aide d'un corpus contrastif composé d'œuvres de fiction produits par des locuteurs natifs, les caractéristiques grapho-phonémiques des deux langues qui s'y trouvent permettent de prédire les usages graphiques présents sur les forums de discussion. L'utilisation de forums est l'occasion de se pencher sur plusieurs phénomènes discursifs propres à leur écologie. Les deux langues ne sont pas pratiquées dans les mêmes conditions : les utilisateurs du scots l'apprennent aux autres membres, les locuteurs de l'AAV s'en servent comme marqueur socio-culturel.

Cette thèse défend l'idée que les pratiques discursives sur un forum créent un lecte qui lui est spécifique, le communaulecte. Ce dernier est en partie observable dans les graphies alternatives. Ces graphies naissent du désir d'adéquation entre la graphie et la phonie, et sont d'autant mieux observables que les membres de BP et SC utilisent une langue minoritaire à laquelle ils sont principalement exposés à l'oral. Cette langue provoque des changements graphiques plus marqués.

Abstract

This dissertation deals with two minority languages spoken in English-speaking countries, Scots and African American vernacular, used on two niche forums whose population is interested in Scottish culture for the former, Scotster, or mainly African American for the latter, Black Planet.

The two linguistic domains which will be summoned are phonology dealing with minority languages and English, because both are discernible within alternative spellings ; as well as discourse analysis, because these asynchronous discourses allow unique data mining and insights into the creation of new lexical or graphic forms, which are more common on online communities they appear on. Thanks to a contrastive corpus made of works of fiction produced by native speakers, the grapho-phonemic traits of the two minority languages can be predicted on internet forums. The use of billboards also allows to study discursive phenomena which are specific to their ecology. The two languages practiced there are not used the same way : the users of Scots tend to teach it to others, while speakers of African American use it as a socio-cultural marker.

This dissertation's intent is to show that discourse practices on a forum create a platform-specific lect, called a communaulect. It is partly noticeable through alternative spellings. These are harbored by a will to lessen the difference between spelling and sound, and are even more detectable because members of BP and SC use a minority language they are mostly exposed to orally, hence twisting words more dramatically.