HAL
open science

La hiérarchisation des constituants discursifs dans un corpus d'anglais oral spontané

Frédérique Passot

To cite this version:

Frédérique Passot. La hiérarchisation des constituants discursifs dans un corpus d'anglais oral spontané. Linguistique. Université de la Sorbonne nouvelle - Paris III, 2004. Français. NNT : . tel-00011240

HAL Id: tel-00011240

https://theses.hal.science/tel-00011240v1

Submitted on 19 Dec 2005

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.

Université Paris III – Sorbonne Nouvelle

Institut du Monde Anglophone

La hiérarchisation des constituants discursifs dans un corpus d'anglais oral spontané

Thèse de linguistique anglaise présentée par

Frédérique Passot

en vue de l'obtention du titre de Docteur de l'Université

Directeur : Monsieur le Professeur Claude Delmas

2004

Remerciements

Ces remerciements vont tout d'abord à Claude Delmas, qui m'a fait confiance en acceptant de diriger mes recherches, menées pour une bonne partie outre Altantique.

Je tiens également à remercier tout particulièrement Mary-Annick Morel, membre du jury, pour le cadre théorique que son enseignement et ses écrits m'ont donné. Qu'elle soit remerciée pour les nombreux ouvrages et le matériel d'enregistrement qu'elle a bien voulu me prêter.

Ma reconnaissance va également à Laurent Danon-Boileau, directeur de précédents travaux et membre du jury, qui a su me donner le goût de l'analyse linguistique, et en particulier de la linguistique de l'oral, à travers son enseignement et ses écrits.

Merci également à Aliyah Morgenstern, membre du jury, grâce à qui j'ai pu soumettre un certain nombre d'hypothèses à la sagacité du groupe TELOS. Qu'elle soit remerciée pour sa disponibilité et son soutien.

Je remercie enfin Jean-Louis Duchet d'avoir accepté de faire partie de mon jury.

Gérard Deléchelle doit être remercié pour m'avoir gentiment envoyé les microfiches de sa thèse.

Merci également à Anne Catherine Simon pour m'avoir fait parvenir le manuscrit de l'ouvrage à paraître issu de sa thèse.

Enfin, ces remerciements vont à Maria Candea pour l'envoi de bon nombre de ses publications, pour sa rigueur et sa disponibilité.

Un grand merci à Elodie Vialleton (qui a mis à ma disposition un long corpus) et Laure Geschwind pour leur relecture attentive de mon manuscrit, pour leurs critiques et leurs suggestions. Qu'elles soient remerciées, avec Jean Szlamowicz, pour les nombreuses heures de discussion que nous avons passées ensemble.

Ces remerciements vont également à Benoît Toïgo et Pierre Boragno pour leurs précisions concernant la notation musicale à l'époque grégorienne décrite au chapitre 1.

Alexis Michaud a procédé aux mesures acoustiques présentées au chapitre 5, qu'il en soit ici remercié, ainsi que pour ses suggestions et commentaires.

Merci à Martine de Cola-Sekali pour ses suggestions concernant mon approche de *because* et *cos* et pour sa disponibilité.

Merci enfin aux membres du groupe SeSyLiA, dirigé par Claure Delmas, Pierre Cotte et Geneviève Girard, qui m'ont permis d'exposer à plusieurs reprises les derniers avancements de mon travail.

Toute ma gratitude va également à Celia, James, Helen, Rebecca, Tony, Frazer, Ben, Reena et Jonathan, ainsi que Nick, Matthew, Mike, Hannah, Ian et Claire qui ont bien voulu faire abstraction des micros et continuer à discuter.

Merci également à tous ceux — famille, amis, étudiants ou collègues de Paris III — qui se sont soumis à mes tests de perception.

A special thank you goes to everyone at Axiomatic and at Intertrust (and especially to Talal Shamoon, Dave Maher, Mike MacKay, Bill Rainey, Jack Lacy, Knox Carey, Gilles Boccon-Gibod, Steve Mitchell, and William Bradley), for supporting me and providing me with the nicest work environment I'll probably ever know.

Merci également à Emmanuelle de Champs et Karine Bigand, mes compagnes de BNF, sans qui les journées à la bibliothèque auraient été bien plus austères.

Enfin, ce travail doit énormément à Julien Bœuf, tant sur le plan pratique que théorique, à qui j'adresse ici un immense merci. Son attention de tous les jours et son soutien indéfectible m'ont permis de mener cette thèse à bien.

Table des matières

Table des figures **15**

Liste des tableaux **17**

Introduction **18**

0.1 Considérations terminologiques 18

0.1.1 Discours 19

0.1.1.1 De quelques oppositions majeures en linguistique . . . 19

0.1.1.2 Une linguistique du discours 21

0.1.2 Segmental et suprasegmental 30

0.1.2.1 Segmental 30

0.1.2.2 Suprasegmental 31

0.1.3 Hiérarchisation et constituants 32

0.1.3.1 Hiérarchisation 32

0.1.3.2 Constituants 35

0.1.3.3 Bilan 42

0.2 Plan général de l'étude 44

I Cadre théorique et méthodologique **47**

1 Etudes antérieures **49**

1.1 Etudes sur l'intonation 49
1.1.1 Quelques points de repère chronologiques 49
1.1.1.1 Du Moyen-Age au XVIIe siècle : les précurseurs 50
1.1.1.2 XVIIIe et XIXe siècles : la naissance de deux écoles . . 54
1.1.1.3 XXe et XXIe siècles : la diversification des approches . 57
1.1.2 Un domaine théorique et descriptif 59
1.1.2.1 Le modèle de la théorie des contours 61
1.1.2.2 Le modèle de la théorie des niveaux 62
1.1.2.3 Le modèle de la théorie des configurations 63
1.1.2.4 Le modèle de la théorie autosegmentale-métrique . . . 65
1.1.2.5 Le modèle de M.-A. Morel et L. Danon-Boileau 68
1.1.3 Un domaine expérimental 71
1.1.3.1 Apport de l'acoustique et de l'informatique 71
1.1.3.2 Production et perception 74
1.2 Décrire la langue et l'échange verbal 89
1.2.1 Décrire la langue 89
1.2.1.1 Le modèle distributionnel 90
1.2.1.2 Le modèle génératif 93
1.2.1.3 Limites de l'influence de ces modèles dans notre travail 93
1.2.1.4 Autres modèles hiérarchiques pour la description de faits syntaxiques 95
1.2.2 Décrire l'échange verbal 98
1.2.2.1 Le modèle communicationnel et informationnel 98
1.2.2.2 Le modèle interactionnel 103
1.2.2.3 Le modèle énonciatif 106

2 Principes théoriques 108

2.1 Présentation du modèle général adopté 109
2.1.1 Le modèle de M.-A. Morel et L. Danon-Boileau 109
2.1.1.1 Les marqueurs suprasegmentaux 109
2.1.1.2 Un modèle de structuration du discours oral 111
2.1.2 Application de ce modèle à la description de l'anglais 114
2.1.2.1 Le paragraphe oral en anglais 115
2.1.2.2 Valeur conventionnelle des paramètres en anglais . . . 121
2.2 Présentation des hypothèses de départ 133
2.2.1 Segmental et suprasegmental : une syntaxe autre 133
2.2.2 Fonctions et significations de l'intonation 135
2.2.3 Système de l'oral et fractales 139

3 Présentation du corpus utilisé 145

3.1 Constitution d'un corpus d'oral spontané 145
3.1.1 Les enjeux 145
3.1.1.1 Eléments de définition 146
3.1.1.2 Corpus existants 148
3.1.1.3 Représentativité d'un corpus 151
3.1.2 Corpus originaux analysés et rejetés 154
3.1.2.1 Corpus primaire 155
3.1.2.2 Corpus secondaire 162
3.1.2.3 Corpus de comparaison 166
3.1.2.4 Corpus rejeté 167
3.1.3 Bilan qualitatif et quantitatif du corpus exploité 169
3.2 Traitement du corpus 171
3.2.1 Etablissement de la transcription 171
3.2.1.1 Considérations méthodologiques 172

3.2.1.2 Transcrire le verbal 174
3.2.1.3 Transcrire le paraverbal 176
3.2.1.4 Difficultés de transcription 177
3.2.2 Etablissement des courbes intonatives 179
3.2.2.1 Choix du logiciel d'analyse prosodique 179
3.2.2.2 Contraintes et choix techniques 181
3.2.2.3 Lecture des courbes 183
3.2.3 Traitement informatique et réflexion linguistique 184
3.2.3.1 Apport du traitement informatique 185
3.2.3.2 Limites du traitement informatique 186
3.2.3.3 Bilan 188

II De quelques marqueurs segmentaux de hiérarchisation 191

4 Éléments pour une analyse distributionnelle des marqueurs de subordination 193

4.1 De la subordination : définition de travail 194
4.1.1 Définition consensuelle 194
4.1.1.1 Caractéristiques syntaxiques générales 194
4.1.1.2 Sémantisme de la circonstance 197
4.1.2 Incidence des subordonnées circonstancielles 198
4.1.2.1 Incidence syntaxique de la subordonnée 198
4.1.2.2 Incidence sémantique de la subordonnée 200
4.2 Eléments de distribution 202
4.2.1 Les subordonnées exprimant une circonstance 203
4.2.2 Les marqueurs de subordination 204
4.2.2.1 Données prises en compte et calculs effectués 204

4.2.2.2 Résultats comparés 207
4.3 Bilan et hypothèses explicatives 214
4.3.1 Résumé des résultats 214
4.3.1.1 Points communs entre les corpus dépouillés 214
4.3.1.2 Points de divergence 215
4.3.2 Hypothèses explicatives 216
4.3.2.1 Deux hypothèses linguistiques concurrentes 216
4.3.2.2 Arguments, objections et conclusions 217

5 Le cas des circonstancielles de cause 220
5.1 Généralités sur l'expression de la cause 221
5.1.1 Une relation logique 221
5.1.2 La cause en langue et en discours 222
5.1.2.1 Outils disponibles et leurs caractéristiques 222
5.1.2.2 Outils effectivement utilisés : éléments de distribution . 224
5.2 Aspects de la subordination causale en *because* et *cos* 228
5.2.1 Deux facettes d'un même marqueur 228
5.2.1.1 *Cos* forme affaiblie de *because* 228
5.2.1.2 Emploi de *cos* pour *because* 229
5.2.2 Caractéristiques acoustiques 230
5.2.3 Caractéristiques distributionnelles communes 232
5.2.3.1 Ordre des propositions 233
5.2.3.2 Présence et place des pauses 237
5.3 *Because* et *cos* : deux marqueurs à distinguer ? 244
5.3.1 Emploi de *because* pour *cos* 244
5.3.2 Caractéristiques distributionnelles et valeurs intersubjectives divergentes 245

5.3.2.1 Sujet syntaxique des propositions 245
5.3.2.2 Position des pauses 249
5.3.2.3 Formulation et prise en charge 251
5.4 Sur la relation entre p et *because/cos* q 254
5.4.1 Restriction et opérations syntaxiques 254
5.4.1.1 Indicateurs d'une relation restrictive dans notre corpus 254
5.4.1.2 Un *because* plus restrictif ? 259
5.4.2 Le cas des énoncés métalinguistiques 260
5.4.2.1 Caractéristiques intonatives attendues 261
5.4.2.2 Caractéristiques intonatives observées 263
5.4.3 Vers une dissolution du lien causal ? 266
5.4.3.1 Traitement appositif de la subordination 266
5.4.3.2 Diversification du rôle de *because/cos* 268
5.4.3.3 Spécialisation des marqueurs et bilan 271

6 Subordination et organisation discursive 279
6.1 Hypotaxe ou parataxe ? 279
6.1.1 Un fonctionnement hypotaxique 280
6.1.2 Un fonctionnement parataxique 281
6.1.2.1 *Because/cos* + pause + proposition 282
6.1.2.2 *For* et *because/cos* 283
6.2 Pérennité d'une hiérarchisation discursive des constituants liés 285
6.2.1 Une hiérarchisation assurée par les marqueurs ? 285
6.2.1.1 Arrière plan et premier plan 285
6.2.1.2 *So* marqueur de « position » 288
6.2.1.3 *And* marqueur polyvalent 289
6.2.2 Sur la délimitation et l'organisation des unités du discours . . . 290

6.2.2.1 Délimitation et chaînage 290
6.2.2.2 Niveau hiérarchique des éléments reliés 291
6.2.2.3 Réseaux de marqueurs 294
6.3 Progression discursive en spirale 300
6.3.1 Spirale et circularité 301
6.3.1.1 Parataxe de l'enfant 301
6.3.1.2 Analyse d'un exemple : la question des clés 303
6.3.2 Gestion de la relation intersujets 310
6.3.2.1 Regarder vers l'arrière 310
6.3.2.2 Pour aller vers l'avant 312

III De quelques faits suprasegmentaux de hiérarchisation 315

7 La gestion de l'espace intersubjectif par la prosodie 317
7.1 Gestion tripartite de l'espace intersubjectif 317
7.1.1 Quelques travaux antérieurs 318
7.1.2 Division retenue dans ce travail 319
7.1.3 De la théorie à l'exemple 321
7.1.3.1 Consensus établi 322
7.1.3.2 Discordance et forçage du consensus 323
7.1.3.3 Egocentrage 324
7.2 Trois plages mélodiques concordantes ? 325
7.2.1 Calcul des bandes 326
7.2.1.1 Hypothèse de la loi normale 327
7.2.1.2 Bande neutre non modulée 331
7.2.2 Confrontation à l'intuition linguistique 332
7.2.2.1 De l'exemple à la théorie 332

7.2.2.2 Redondance et indépendance des marques 334
7.3 Limites et perspectives 337
7.3.1 Limites imputables aux logiciels 337
7.3.2 Limites imputables à la méthodologie adoptée 338
7.3.2.1 Phénomènes à considérer 338
7.3.2.2 Partis pris contestables 340
7.3.3 Perspectives 340

8 Les pauses silencieuses dans la construction du discours 343
8.1 Apport des études antérieures 343
8.1.1 La pause silencieuse : éléments de définition 344
8.1.1.1 Seuils adoptés par les études antérieures 345
8.1.1.2 Seuils adoptés pour cette étude et premières analyses . 346
8.1.2 La part du temps de pause dans l'activité langagière 348
8.1.3 Perceptibilité des pauses silencieuses 349
8.1.3.1 Durée, site d'occurrence et valeur perceptuelle 349
8.1.3.2 Quelle place pour le contexte ? 351
8.2 Mise à l'épreuve des hypothèses de travail 353
8.2.1 Test de perception 353
8.2.1.1 Matériel utilisé 353
8.2.1.2 Procédure appliquée 355
8.2.1.3 Sujets sélectionnés 355
8.2.1.4 Dépouillement des résultats 356
8.2.2 Vérification de nos hypothèses de travail 360
8.2.2.1 Hypothèse 1 : durée et emplacement des pauses 361
8.2.2.2 Hypothèse 2 : durée et perceptibilité des pauses 362
8.2.2.3 Hypothèse 3 : perceptibilité et emplacement des pauses 365

8.2.3 Limites de nos hypothèses de travail 368
8.2.3.1 Première hypothèse 368
8.2.3.2 Deuxième hypothèse 369
8.2.3.3 Troisième hypothèse 370
8.3 Les pauses dans la construction du discours et de l'interaction 371
8.3.1 Niveaux de pertinence des différentes pauses 371
8.3.1.1 Niveaux de construction du discours conversationnel . 372
8.3.1.2 Différents types de charnières 373
8.3.2 Un marqueur mixte dans un contexte prosodique complexe . . . 377
8.3.2.1 Contexte prosodique 377
8.3.2.2 Contexte pragmatique 381
8.3.2.3 Contexte syntaxique des pauses d'hésitation 383

9 Aspects de la gestion des tours de parole 387
9.1 Définitions préliminaires et problèmes 388
9.1.1 Tour de parole 388
9.1.1.1 Définition mécaniste 388
9.1.1.2 Définition empirique 391
9.1.2 Place transitionnelle 393
9.1.2.1 Caractéristiques syntaxiques 393
9.1.2.2 Caractéristiques prosodiques et mimo-gestuelles 394
9.1.2.3 Caractéristiques pragmatiques et interactionnelles . . . 395
9.2 Le système de l'alternance des tours à l'épreuve 396
9.2.1 Règles d'alternance 397
9.2.1.1 Anticipation des places transitionnelles 397
9.2.1.2 Anticipation de l'identité du locuteur suivant 401
9.2.2 Alternance, silence et simultanéité de parole 402

9.2.2.1 Assouplissement des règles 402
9.2.2.2 Accordage temporel des prises de parole 403
9.3 Les chevauchements de parole : des ratés de l'interaction ? 406
9.3.1 Caractéristiques générales du phénomène 406
9.3.1.1 Présentation de l'extrait de corpus analysé 406
9.3.1.2 Fréquence d'apparition 407
9.3.1.3 Deux configurations 408
9.3.2 Un phénomène perturbateur ? 409
9.3.2.1 Durée des segments recouverts 409
9.3.2.2 Caractéristiques syntaxiques 410
9.3.2.3 Caractéristiques sémantiques 411
9.3.3 Un ciment dans l'édifice conversationnel 413
9.3.3.1 Sur le plan de l'organisation discursive 413
9.3.3.2 Sur le plan intersubjectif 415

Conclusion **420**
Le discours conversationnel comme objet fractal 420
La syntaxe spécifique de l'oral 422
Réduction du champ des possibles 422
Un marqueur supplémentaire : *cos* 424
Réinvestissement des structures existantes 426
Signification et rôle des paramètres prosodiques 428
Pour une méthodologie de l'analyse prosodique 428
Niveaux de pertinence et délinéarisation de la chaîne 430

A Annexe au chapitre 3 **436**
A.1 Enregistrement du corpus 436
A.2 Numérisation du corpus et édition des fichiers sonores 437

A.3 Analyse prosodique 437

B Annexe au chapitre 4 **439**

B.1 Données issues de la LGSWE 440

B.2 Données issues du BNC 441

B.3 Données issues du corpus FP 444

B.4 Bilan : données comparables 444

C Annexe au chapitre 5 **446**

C.1 Référence du sujet de la principale 446

C.2 Position des pauses silencieuses 447

C.3 Présence et densité des marques du travail de formulation 448

C.3.1 Présence de marques du travail de formulation 448

C.3.2 Densité des marques du travail de formulation 448

C.4 Force argumentative : *so* et *but* 449

Références bibliographiques **450**

Table des figures

1 Le discours, objet composite 22

1.1 Manuscrit de Munster : formulaire pour la récitation liturgique 51

1.2 Représentation de l'intonation par D. Jones 62

1.3 Représentation de l'intonation par K. Pike 63

1.4 Représentation de l'intonation par A. Crompton 64

1.5 Représentation de l'intonation par J. Pierrehumbert & J. Hirschberg 66

1.6 Représentation de l'intonation de l'anglais selon ToBI 67

1.7 Courbe de la perception humaine d'après la loi de Weber-Fechner 85

1.8 Diagramme schématique du système de la communication d'après C. Shannon . . 100

2.1 Deux schémas d'analyse hiérarchique parallèles 123

2.2 Correspondance des niveaux d'analyse par rapport au temps 124

2.3 Points de contact entre suprasegmental et segmental 136

2.4 Le triangle de Sierpinski 139

2.5 Le flocon de neige de Von Koch 140

3.1 Courbes prosodiques produites par WinPitch Classic 184

3.2 Erreur du logiciel (en haut) corrigée après écoute (en bas) 187

4.1 Exemple de représentation hiérarchique de la subordination complétive en grammaire générative 196

4.2 Fréquence d'occurrence (pour 100 mots) de chaque conjonction dans chaque corpus 208
4.3 Part de chaque conjonction par rapport au nombre total de conjonctions dans chacun des trois corpus (pour 100 conjonctions) 209
4.4 Densité de conjonctions dans le corpus FP 211
4.5 Part de chaque conjonction par rapport au nombre total de conjonctions dans chacun des textes de notre corpus 213

5.1 Articulation comparée de *because* et de *cos* 232

6.1 Spirale discursive entremêlant les thèmes des emplois du temps (T) et des clés (K) 307

7.1 Distribution des valeurs de F_0 dans [Celia] en quarts de ton répartis autour de la moyenne 329
7.2 Distribution des valeurs de F_0 dans [James] en quarts de ton répartis autour de la moyenne 329
7.3 Courbe de l'exemple 7.8 avec matérialisation des trois plages mélodiques 336
7.4 Courbe de l'exemple 7.12 avec matérialisation des trois plages mélodiques 339

8.1 Distribution des pauses en fonction de leur durée dans le corpus Celia&James . . 347
8.2 Comparaison des paramètres du signal filtré (en bas) et non filtré (en haut) 386

9.1 Modulation mélodique sur le pronom interrogatif 399
9.2 Signes d'une stratégie de conservation de la parole 400
9.3 Signaux de fin de tour 405
9.4 Configurations des recouvrements de parole : L1> (à gauche) ou L2> (à droite) . . 408

B.1 Dénombrement des conjonctions simples dans le corpus de la LGSWE 440
B.2 Dénombrement des conjonctions simples dans le BNC 443
B.3 Dénombrement des conjonctions simples dans notre corpus 443

Liste des tableaux

1.1 Seuil différentiel des fréquences (SDF) pour F_0 entre 50 et 600 Hz d'après M. Kitantou 86

1.2 Seuil différentiel de glissando (SDG) pour F_0 entre 50 et 600 Hz d'après M. Rossi . 88

1.3 Modèle hiérarchique interactionnel 103

2.1 Exemple de découpage du français en paragraphes 114

2.2 Exemple de découpage de l'anglais en paragraphes 121

2.3 Niveaux de pertinence 132

3.1 Comparaison de la densité lexicale de notre corpus et du corpus LGSWE 153

3.2 Corpus utilisé 170

4.1 Eléments de comparaison entre les trois corpus concernant le nombre de conjonctions simples relevées par rapport au nombre de mots 207

4.2 Part de *if, because/cos* et *when* par rapport au nombre total de conjonctions . . . 210

5.1 Ordre d'apparition des propositions en *because/cos* 236

5.2 Place des pauses silencieuses dans les 122 énoncés en *because/cos* de notre corpus 243

5.3 Densité des marques du travail de formulation dans les énoncés en *because* et en *cos* 252

7.1 Division de l'espace intersubjectif 321
7.2 Moyenne, médiane et écart type de F_0 dans les deux fichiers 327
7.3 Test de conformité à la loi normale appliqué à [Celia] 328
7.4 Test de conformité à la loi normale appliqué à [James] 330
7.5 Bande neutre pour les deux fichiers 331

8.1 Distribution des pauses par tranches en fonction de leur durée dans le corpus Celia&James 347
8.2 Distribution des pauses en fonction de leur site d'occurrence dans notre extrait de corpus 359
8.3 Durée moyenne des pauses en fonction de leur site d'occurrence dans notre extrait de corpus 361
8.4 Corrélation durée/perceptibilité d'après le coefficient de Spearman 363
8.5 Débit, longueur des suites sonores et perceptibilité des pauses dans notre extrait de corpus 378

9.1 Segments produits en situation de chevauchement de parole dans l'épisode *Travel* 407

A.1 Comparaison de quatre logiciels d'analyse prosodique 438

C.1 Référence du sujet de la principale 447
C.2 Position des pauses silencieuses 447
C.3 Présence de marques du travail de formulation 448
C.4 Densité des marques du travail de formulation 449
C.5 Densité des marques du travail de formulation 449

Introduction

Le travail que nous présentons ici aborde la question de la hiérarchisation des constituants discursifs en anglais oral spontané tel qu'il se manifeste dans un corpus conversationnel. En nous appuyant sur la description de données originales, nous souhaitons analyser dans quelle mesure l'organisation du discours est le résultat de la combinaison de phénomènes issus de la double dimension « segmentale » et « suprasegmentale » de l'oral que nous définirons plus loin. Une attention portée à ces deux plans nous permettra d'appréhender divers modes de structuration de l'anglais oral, du point de vue syntaxique, mais également du point de vue thématique et interactionnel, sans jamais perdre de vue que le discours conversationnel est le lieu d'enjeux énonciatifs.

0.1 Considérations terminologiques

Toute réflexion linguistique s'inscrit dans une tradition, ou du moins dans un contexte. Les termes « discours », « hiérarchisation », « constituants », ainsi que « segmental » et « suprasegmental », sont porteurs de significations issues de ces traditions. Afin de définir dans quel sens nous les utilisons dans ce travail et d'affiner notre projet de recherche, il convient de rappeler les acceptions que les linguistes ont pu leur attribuer dans le cadre de problématiques souvent différentes des nôtres.

0.1.1 Discours

0.1.1.1 De quelques oppositions majeures en linguistique

La tradition platonicienne dissocie deux plans dans l'appréhension du monde, qui fonctionnent comme les deux faces d'une même réalité. Il s'agit d'une part du monde sensible (celui des choses qui nous entourent), et d'autre part du monde des Idées (celui des modèles dont relèvent ces choses). Cette dichotomie s'applique également à la linguistique, qui opère plusieurs distinctions fonctionnant sur cette opposition entre modèle et image. Quelle est la place du discours dans cette série d'oppositions ?

Langue et langage, langue et parole Les langues (le terme est le plus souvent employé au pluriel) sont conçues dans le premier cas comme la manifestation particulière d'un objet qui les subsume et les dépasse — le langage. L'objet ultime de la linguistique étant sans doute la connaissance profonde du langage, celle-ci doit cependant s'attacher à décrire et comprendre les formes sous lesquelles il se réalise. Les langues, vues sous l'angle de leur multiplicité, font donc signe collectivement en vertu de leur relation au langage dont elles participent.

Opposée à la parole, la langue n'est plus l'image qui éclaire un objet conceptuel impossible à observer de façon directe. Elle reçoit au contraire le statut d'objet, qui s'instancie dans la parole. Pour F. de Saussure, en effet, la langue est forme et non substance. Elle offre le cadre dans lequel s'exprime la matière, c'est-à-dire la parole. Cette conception est résumée par O. Ducrot et J.-M. Schaeffer :

> La langue se définit comme un code, en entendant par là la mise en correspondance d'« images auditives » et de « concepts ». La parole, c'est l'utilisation, la mise en oeuvre de ce code par les sujets parlants.[1]

La langue se livre ainsi dans les données empiriques de la parole, la forme ne se

[1] DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Le Seuil. (p.245).

comprenant qu'à partir de la substance.

Parole et discours Ce dernier couple ne regroupe pas des éléments opposables. Il s'agit de termes proches dont il faut cependant préciser la définition. Certains chercheurs récusent en effet le terme de *parole* utilisé par F. de Saussure par opposition à la langue. G. Guillaume reproche à ce terme d'être éminemment ambigu. Sans pour autant remettre en question l'opposition du couple langue/parole, il substitue à ce dernier terme celui de *discours*. C'est ce dont son élève R. Valin rend compte dans le passage suivant :

> La véritable dichotomie est celle de la *langue* et de son *emploi* : la langue est la chose qu'on emploie quand on parle. L'inconvénient d'appeler, avec Saussure, *parole* l'emploi de la langue est qu'il y a *de la parole* — sous la forme de signes virtuels — dans la langue. Aussi G. Guillaume a-t-il, lui, de tout temps préféré le mot *discours* [...].[2]

D'où la définition donnée par R. Valin, qui, comme son maître, fait du discours « un emploi de la langue ».[3] On retrouve la même distinction chez A. Reboul et J. Moeschler qui parlent de « phrases en usage ».[4] C'est une **première définition** que nous retiendrons pour *discours*, sur laquelle nous serons néanmoins amenée à revenir.

Les distinctions que nous venons d'évoquer sous-tendent nécessairement toute réflexion linguistique. Pour G. Guillaume, le mouvement de pensée du concret vers l'abstrait et de l'abstrait vers le concret constitue le fondement de la recherche linguistique. Il permet même d'en définir la méthode :

[2] VALIN, Roch (1964). La méthode comparative en linguistique historique et en psychomécanique du langage. *Cahiers de psychomécanique du langage* 6 : pp.1-57. Québec : Presses de l'Université Laval. (p.23n).

[3] *ibid.* (p.56).

[4] REBOUL, Anne & MOESCHLER, Jacques (1998b). *Pragmatique du discours : de l'interprétation de l'énoncé à l'interprétation du discours*. Paris : Armand Colin. (p.41).

> Les discours sont des suites, des séquences, de phrases en usage [...].

> La méthode en question [...] est le mouvement incessamment inversé d'une observation qui ne s'approche, et le plus possible, du concret que pour pouvoir mieux s'en évader en direction de l'abstrait, et ne prend pied dans l'abstrait qu'afin de pouvoir mieux ensuite, par une nouvelle et plus puissante attaque du concret, le pénétrer et s'introduire à une connaissance de son intime structure.[5]

Ce mouvement est un des principes méthodologiques que nous retenons dans notre travail.

0.1.1.2 Une linguistique du discours

Si ces distinctions sont bien essentielles, il est nécessaire ici de situer notre objet d'étude par rapport à elles. N'ayant jamais accès qu'à la face concrète du langage, le linguiste ne devrait pouvoir se donner pour domaine que celui du discours. Cependant, d'autres formes d'« usage de la langue » s'offrent à lui. Ces formes ne relevant pas toutes selon nous du discours, nous serons amenée ici à réfuter notre première définition du discours.[6] Comment donc définir ce domaine et ses caractéristiques ? En quel sens notre démarche s'inscrit-elle dans une linguistique de discours ?

Un domaine polymorphe Nous considérons le discours comme un domaine polymorphe et vivant se définissant à la fois par des données qui lui sont propres et d'autres, extérieures à lui-même.

Définition de travail En nous inspirant de la théorie informatique, nous proposons de considérer ici le discours comme une classe abstraite composite dont les manifestations peuvent aussi bien être temporelles qu'atemporelles, comme l'illustre la Figure 1.[7] Discours juridique et discours religieux constituent des exemples canoniques d'un discours atemporel ; tandis que récit et dialogue définissent les deux branches des discours de type temporel. La caractéristique composite de cette catégorie permet une

[5] GUILLAUME, Gustave (1952). La langue est-elle ou n'est-elle pas un système ? *Cahiers de*

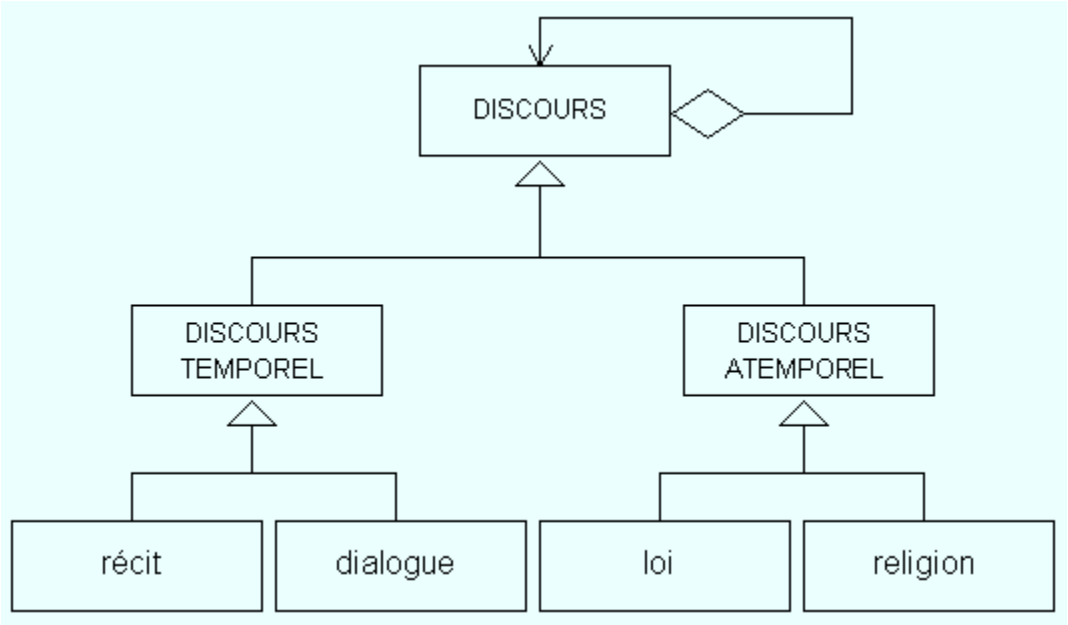

FIG. 1 – Le discours, objet composite

perméabilité des frontières entre types de discours. Ainsi, la loi, discours atemporel, s'amende et agrège des discours temporels. De même, un récit peut contenir un dialogue et inversement. Un discours, quel qu'il soit, peut agréger un ou plusieurs autres types de discours. Ainsi conçu, le discours ne constitue pas une unité supérieure à la phrase mais un domaine d'opération de la langue qui prend la forme d'énoncés.

Cette conception contrevient à l'opposition généralement admise entre récit et discours, où le récit relève du domaine de la narration, et le discours de celui de l'argumentation. Or cette opposition ne nous semble pas satisfaisante dans la mesure où le récit relève précisément de ce que nous entendons par discours et ne s'y oppose pas. Nous préférons en revanche introduire la catégorie du dialogue (réel ou fictif), qui, relevant également du discours, nous semble plus à même de figurer en contrepoint

Linguistique Structurale 1 : pp.3-30. Québec : Presses de l'Université Laval. (p.5).

[6]Voir p.20.

[7]Ce schéma est conforme aux arbres d'héritage UML (unified modeling language) décrits par exemple dans GAMMA, Erich *et al.* (1995). *Design Patterns : Elements of Reusable Object-Oriented Software). Boston : Addison-Wesley.*

du récit dans notre schéma.

Les attributs du discours Nous avons jusqu'ici trouvé quasiment indifféremment dans les définitions proposées plus haut les termes de « phrase » et d'« énoncé ». Il convient ici de les distinguer. Contrairement à la syntaxe, par exemple, qui travaille sur la phrase, une linguistique du discours s'intéresse à l'énoncé. Certes, phrases et énoncés constituent tous deux une forme d'usage de la langue. Leurs caractéristiques pragmatiques et énonciatives les opposent cependant.

Qu'ont en commun loi, religion, récit et dialogue qui justifie pour nous leur parenté et les distingue des phrases isolées ? Il nous semble que ces manifestations du discours ont en commun de relever de ce que A. Reboul et J. Moeschler appellent, avec D. Sperber et D. Wilson,[8] une « double intentionnalité ».[9] Pour ces auteurs, le locuteur ou l'auteur d'un énoncé est animé de deux intentions complémentaires :

> I. **L'intention informative**, c'est-à-dire l'intention qu'a le locuteur d'amener son interlocuteur à la connaissance d'une information donnée.
>
> II. **L'intention communicative**, c'est-à-dire l'intention qu'a le locuteur de faire connaître à l'interlocuteur son intention informative.[10]

Une telle perspective nous oblige à revenir sur le terme d'« emploi » utilisé par R. Valin, ou celui d'« usage » chez A. Reboul et J. Moeschler,[11]. Comparons le fonctionnement d'une phrase et celui d'une gamme musicale. La phrase isolée constitue bien un usage de la langue, tout comme la gamme est un usage des notes de musique. Mais dans les deux cas, il ne s'agit que d'une application de principes théoriques et non d'une forme plus inventive relevant du discours. La gamme est un exercice qui n'a pas

[8] SPERBER, Dan & WILSON, Deirdre (1986, 2e éd. 1995). *Relevance. Communication and Cognition.* Oxford : Blackwell.

[9] REBOUL, Anne & MOESCHLER, Jacques (1998a). *La Pragmatique aujourd'hui : une nouvelle science de la communication.* Paris : Le Seuil.

[10] *ibid.* (p.71).

[11] Voir p.20

pour vocation de transmettre une émotion musicale. Au contraire, de même qu'une pièce de musique, l'usage de la langue appelé *énoncé* par A. Reboul et J. Moeschler fait sens par lui-même et au-delà de lui-même. Isolée, la gamme n'est rien d'autre qu'une gamme. Elle n'a de sens que pour et par elle-même. Au sein d'une pièce de musique, elle participe de cette pièce et contribue à l'émotion musicale suscitée par l'ensemble.

La distinction entre intentionnalité informationnelle et intentionnalité communicationnelle nous fait ainsi revenir à la dichotomie entre forme et substance, *modus* et *dictum*. Si l'énoncé combine les deux, la phrase reste du côté de la forme, du *dictum*, sans pouvoir offrir de substance susceptible d'une interprétation. Le travail que nous nous proposons de mener ici se place du point de vue de l'interprétation (par opposition à la production). Il ne s'agit pas de juger de la recevabilité grammaticale d'une construction, mais d'expliquer les conditions de possibilité de son interprétation.[12] Ainsi, puisque la définition du discours comme un usage ou un emploi de la langue nous conduit à considérer comme relevant du discours ce qui n'en relève pas, elle est insuffisante.

L'analogie avec la gamme et la pièce de musique souligne la nécessité de prendre en compte une deuxième notion, après celle de la double intentionnalité, qui permet d'affiner ce que nous entendons par *discours*. Par opposition aux phrases isolées, l'énoncé fonctionne en effet au sein de ce que nous appellerons pour l'instant un ensemble. C'est le sens des termes de « suites » et « séquences » dans la définition déjà mentionnée plus haut :

> Les discours sont des suites, des séquences, de phrases en usage [...].[13]

La notion d'ensemble se définit à son tour par la cohésion de ses membres et par une organisation interne — organisation de la structure de chaque membre, et orga-

[12]Nous justifierons plus amplement notre choix de l'énoncé comme unité opératoire dans le cadre de notre étude.

[13]REBOUL, Anne & MOESCHLER, Jacques (1998b). *op. cit.* (p.41).

nisation de la structure globale — dont ce travail se donne pour objectif de détailler les caractéristiques. La cohésion d'un tel ensemble est par essence dynamique. Elle n'est en aucun cas donnée une fois pour toutes : il s'agit d'une construction et d'une reconstruction permanente au cours de l'échange.

Enfin, l'analyse du discours prend en compte des données extérieures au discours, mais qui influencent son organisation et sa matière. Nous avons mentionné plus haut la caractéristique composite du discours qui consiste pour une partie A d'un tout à se redire en entier ou en partie dans une partie B du même tout. Il s'agit d'une caractéristique « fermée » du discours (au sens d'A. Reboul & J. Moeschler),[14] dont les parties s'auto-définissent et s'auto-décrivent. Cette vision fractale, que nous allons développer au cours de notre étude, ne permet pas d'épuiser les propriétés du discours. Le concept de double intentionnalité induit une ouverture du discours vers le méta-, le para- ou même le non-verbal.[15] Pour être interprété, le discours nécessite la prise en compte de données interactionnelles et encyclopédiques, aussi bien que syntaxiques et sémantiques. Ceci explique en partie la place réservée à l'étude du « contexte » sous toutes ses formes dans les analyses linguistiques dépassant le niveau de la phrase. On note une large polysémie de ce terme, qui fait tout à la fois référence, selon les auteurs, au verbal ou au situationnel, c'est-à-dire à l'intérieur du texte et à l'extérieur du texte. Intervient donc ici la caractéristique composite ouverte du discours, c'est-à-dire sa propriété à intégrer des données extérieures dont il faut rendre compte pour l'expliquer.

Bilan De ce domaine polymorphe, nous retiendrons trois implications qui le placent au centre de notre problématique.

[14] *ibid.*

[15] Nous entendons par métadiscursif toute forme de discours sur le discours, par para-verbal les productions vocales accompagnant la production du discours (telles que la toux ou les éclats de rire). Le non-verbal fait référence aux données extérieures à la production verbale (telles que la situation).

Le discours en tant que reflet d'un système transcendant Tout d'abord, c'est en effet le *discours* tel que défini plus haut qui nous intéresse, pour lui-même, mais également en tant que manifestation tangible de la langue et de son système. Replaçons notre première définition[16] dans cette perspective et revenons à nouveau sur ce qu'implique le terme d'« emploi » utilisé par R. Valin, ou celui d'« usage » chez A. Reboul et J. Moeschler.

Se pose ici crucialement le choix des données de départ sélectionnées par le linguiste parmi ce que lui offre son expérience concrète du langage. Si certains chercheurs construisent pour l'occasion les segments sur lesquels ils conçoivent et mettent à l'épreuve leurs théories, nous préférons travailler sur des productions avérées, contemporaines de leur élaboration. « Usage » et « emploi » nous semblent définir une ligne méthodologique forte pour le linguiste, dont l'objectif consistera à expliquer des faits observés en contexte afin de proposer une théorie unifiée. De ce fait, comme nous l'avons vu plus haut, on parlera plus difficilement de discours dans le cas de phrases isolées créées dans le but d'illustrer une particularité grammaticale observée par ailleurs. Bien que ces phrases constituent un usage de la langue (par le linguiste lui-même), elles ne s'inscrivent pas à proprement parler dans le domaine du discours. Elles ne sont pas de même nature que le discours et n'en ont pas les attributs. Ainsi, nous estimons qu'elles ne constituent pas une image fidèle du discours, dont elles ne relèvent pas, et ne peuvent donc pas en éclairer le système spécifique.

Notre travail sur un corpus d'oral spontané correspond donc d'abord à un choix théorique et méthodologique. Certes, toute production langagière participe du langage et est à ce titre digne d'intérêt. Mais si l'objet de la linguistique est la connaissance de la langue et de son système, et si le discours en est la manifestation tangible, alors il convient de travailler sur la matière linguistique la plus à même de faire progresser cette connaissance. C'est pourquoi notre recherche porte, au sein du domaine du

[16] Voir p.20.

discours, sur un corpus d'exemples oraux authentiques en contexte. Nous reviendrons en détail sur ses caractéristiques[17] mais nous pouvons ici préciser que notre corpus relève de la catégorie « dialogue » de notre diagramme, et qu'il s'agit, au sein de cette catégorie, de conversations naturelles mettant chacune en présence deux locuteurs.

Bien entendu, un corpus, si long soit-il, ne peut prétendre à la représentativité totale de l'ensemble dont il n'est qu'un exemple. Cet ensemble est de toute façon en perpétuelle expansion, et il est impossible de le tenir tout entier sans recourir à l'abstraction. Nous montrerons néanmoins dans quelle mesure notre corpus est comparable aux corpus de référence en ce qui concerne l'anglais oral spontané (le British National Corpus, le corpus de la *Longman Grammar of Spoken and Written English*),[18] et comment il donne à ce titre une image aussi fiable que ceux-ci du système de la langue dans lequel ils s'inscrivent.

Le discours en tant que lieu de prise en charge des opérations énonciatives Le concept de double intentionnalité exposé plus haut comme un attribut du discours confère à ce domaine une place essentielle au regard de la problématique énonciativiste dont nous nous réclamons. Au sein du discours, l'oral spontané est le lieu où se jouent crucialement les enjeux soulignés par la théorie initiée par E. Benveniste et continuée par A. Culioli. Reprenant les termes d'E. Benveniste,[19] J. Dubois propose cette **deuxième définition** :

> Le discours est le langage mis en action, la langue assumée par le sujet parlant.
>
> (Syn. : PAROLE.)[20]

[17] Voir Chapitre 3, p.145.

[18] BIBER, Douglas *et al.* (1998). *Longman Grammar of Spoken and Written English.* Londres : Longman.

[19] Voir la définition non pas du « discours » mais de l'« énonciation » dans BENVENISTE, Emile (1974). *Problèmes de linguistique générale II.* Paris : Gallimard. (p.80) :

> L'énonciation est cette mise en fonctionnement de la langue par un acte individuel d'utilisation.

[20] DUBOIS, Jean *et al.* (1973). *Dictionnaire de linguistique.* Paris : Larousse. (p.150).

La théorie développée ici place le sujet parlant au centre de son discours à travers les notions de mise en jeu des représentations mentales et de prise en charge du discours. Le sujet énonciateur laisse son empreinte dans le discours à travers des marqueurs spécifiques ou plus généralement à travers son organisation. une influence tangible sur son organisation. Dès lors, de même que le contexte se crée (au moins en partie) à mesure qu'avance le discours, l'interprétation, et notamment le jugement de cohésion, doivent se faire au fil du discours.

La mise en jeu des représentations mentales intervient à la fois globalement et localement. Décrite par A. Reboul et J. Moeschler en termes d'intentionnalité globale et locale,[21] elle correspond à l'effort que les locuteurs doivent fournir pour inférer à chaque instant le sens de la prise de parole de l'autre, ou simplement de l'énoncé qui précède. Le sens, qu'il soit local ou global, ne provient pas de la somme des signifiés. Il s'élabore en contexte, à partir de présuppositions et de préconstructions.

C'est une intuition semblable qui fonde la théorie des opérations énonciatives : toute « énonciation » est le produit d'une subjectivité pétrie de représentations et se trouve insérée dans un contexte linguistique et une situation extra-linguistique nourrissant ces représentations mentales. La formule d'A. Culioli est la suivante :

> L'idée fondamentale est qu'il n'existe pas d'occurrence textuelle isolée. Ainsi, lorsqu'un sujet produit un énoncé (c'est-à-dire une occurrence textuelle, elle-même composée d'occurrences de termes constitutifs), cet énoncé, et chacune de ses parties constitutives, est situé dans un espace énonciatif muni d'un système de coordonnées subjectives et spatio-temporelles, pris dans un champ de relations inter-sujets, et toute occurrence fait partie d'un agrégat structuré d'occurrences qui forment un domaine.[22]

[21] REBOUL, Anne & MOESCHLER, Jacques (1998b). *op. cit.* (chap.VII).

[22] CULIOLI, Antoine (1990). La linguistique : de l'empirique au formel. *Sens et place des connaissances dans la société.* Centre de Meudon-Bellevue, CNRS. 1987. Reprinted in *Pour une linguistique de l'énonciation.* Gap : Ophrys. (Tome 1, débat p.29).

Le discours tel que nous souhaitons l'appréhender dans ce travail correspond donc à cette deuxième définition, qui intègre les notions de contexte et d'intersubjectivité.

Le discours oral en tant que lieu de l'interaction verbale Enfin, le domaine du discours, à l'oral en particulier, est celui où se jouent les relations interpersonnelles. Opposée à la conception « fondamentalement *unilatérale* et *linéaire* ».[23] traditionnellement adoptée jusque vers le milieu du XXe siècle, la prise en compte de la dimension interactive de la communication se fonde sur les faits suivants :

1. Les phases d'émission et de réception sont en relation de *détermination mutuelle* [...].
2. Ces déterminations mutuelles s'exercent de façon aussi bien *successive* que *simultanée* [...]. D'autre part, en ce qui concerne l'axe des successivités, le déroulement des événements communicatifs n'est pas exclusivement linéaire, mais il incorpore des mécanismes d'anticipation et de rétroaction. [...]
3. Le récepteur doit, au même titre que l'émetteur (quoique dans une moindre mesure), être considéré comme *actif* : il produit une activité non seulement cognitive — c'est le « travail interprétatif » —, mais aussi somatique — c'est l'activité régulatrice [...].
4. Il faut admettre enfin que la « clé » (le code) qui permet d'effectuer les opérations d'encodage et de décodage, c'est-à-dire de mettre en correspondance signifiants et signifiés, est en partie *construite* au cours du déroulement de l'interaction.[24]

L'analyse des conversations de notre corpus devra donc rendre compte de ces quatre dimensions, qui en font

[23] KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales.* Paris : Armand Colin. (Tome I, p.25).
[24] *ibid.* (Tome I, pp.25-28).

> [...] le lieu d'une activité collective de production du sens.[25]

Nous verrons notamment la nécessité d'adopter une perspective interactive de la communication pour expliquer la progression thématique ainsi que les changements de tours de parole et les nombreux chevauchements. Ici encore, le discours conversationnel réunit des sujets parlants dont les représentations mentales se construisent et se confrontent en contexte.

En somme, le travail entrepris ici se situe résolument dans le domaine du discours en tant que manifestation tangible d'un système conceptuel transcendant, en tant que lieu de prise en charge des opérations énonciatives, et en tant que lieu de l'interaction verbale par excellence.

0.1.2 Segmental et suprasegmental

Ce travail se donne pour objectif de contribuer à la description des plans « segmental » et « suprasegmental » du discours oral. Qu'entendons-nous derrière ces deux termes ? Comment sont-ils généralement utilisés ?

0.1.2.1 Segmental

Le terme désigne pour nous les unités qui se donnent dans l'ordre des segments de la chaîne linéaire, identifiables par la morphosyntaxe. En ce sens, notre définition du terme se distingue de celle de la phonologie, pour laquelle le phonème relève du domaine du segmental. Selon le niveau d'analyse, le segmental correspond pour nous à des unités allant du mot ou même du morphème jusqu'à des ensembles plus larges, par exemple, l'énoncé. D'une manière générale, le segmental est consigné dans les transcriptions que nous donnons des conversations de notre corpus.

[25] *ibid.* (Tome I, p.28).

0.1.2.2 Suprasegmental

La définition du terme « suprasegmental » est plus complexe. Nous reprenons ici la recherche terminologique de M. Rossi *et al.*[26] Le terme apparaît en 1942 sous la plume de C. Hockett,[27] par opposition au segmental :

> [...] features which clearly follow each other in the stream of speech are segmental. Those which clearly extend over a series of several segmental groupings are suprasegmental[28]

D'après M. Rossi *et al.*, cependant, l'émergence de cette nouvelle terminologie n'a pas résussi à en harmoniser les usages. Les auteurs relèvent au moins cinq emplois différents du terme, qu'il préfèrent abandonner au profit d'« intonation » ou de « prosodie » — deux termes que nous utiliserons comme des synonymes de « suprasegmental ».

Les auteurs reviennent également sur la notion de continuité qui émerge de la plupart des définitions :

> La mélodie est continue, comme toute substance ; mais nous considérons l'intonation comme une forme discontinue, constituée d'unités discrètes, coextensives au phonème et commutables en un même point de la chaîne, ces unités, toujours significatives s'organisant pour former des contours ou syntagmes intonatifs dont le domaine est la phrase ou ses constituants.
>
> Dans cette optique, l'intonation est un signe linguistique qui associe une forme de l'expression à une forme ou une substance du contenu.[29]

Notre démarche s'inscrit dans cette problématique qui cherche à déterminer différents niveaux de contribution des phénomènes intonatifs, compris comme des signes linguistiques, dans la hiérarchisation des constituants du discours (et non de la phrase).

[26] ROSSI, Mario *et al.* (1981). *L'intonation, de l'acoustique à la sémantique.* Paris : Klincksieck.
[27] HOCKETT, Charles F. (1942). A System of Descriptive Phonology. *Language* 18 : 3-21.
[28] *ibid.* (p.100).
[29] ROSSI, Mario *et al.* (1981). *op. cit.* (p.14).

La description de l'articulation de double dimension morphosyntaxique (ou segmentale) et intonative (ou suprasegmentale) du discours est au cœur de notre projet de recherche.

0.1.3 Hiérarchisation et constituants

Ainsi défini dans sa diversité et dans sa complexité, le domaine du discours se prête de façon particulièrement intéressante à une étude portant sur la hiérarchisation de ses éléments constitutifs. Que peut-on cependant entendre par « hiérarchisation » dans le cadre de notre travail ? De quels éléments s'agit-il ?

0.1.3.1 Hiérarchisation

Une notion centrale Sans entrer dans des considérations relevant de la philologie, il est remarquable que les développements majeurs de la pensée tant en linguistique qu'en biologie ou en anthropologie correspondent à l'émergence de conceptions hiérarchisées de leur objet d'étude, ainsi qu'à la classification et à l'analyse de leur organisation sous forme arborescente. On peut ainsi, avec J. Lekomcev, affirmer son importance primordiale en linguistique :

> Hierarchy is, apparently, one of the fundamental concepts of linguistics.[30]

Il semble en effet raisonnable de dire que tout ce qui a un sens peut s'appréhender de façon hiérarchisée, et que c'est en cela qu'il peut faire sens au-delà de lui-même. Une vue synthétique surplombante d'éléments d'abord isolés de façon analytique favorise leur compréhension : elle permet littéralement de tout prendre ensemble, afin d'en reconnaître l'inscription dans un système.

L'agencement des différentes parties respecte un ordre qu'il convient d'étudier. Pour reprendre les termes du « théorème de Platon » formulé dans le *Sophiste* et

[30] LEKOMCEV, Ju. K. (1974). On certain formal characteristics of linguistic hierarchies. *Linguistics. An international review* 131 : 39-48. (p.39).

rappelé par S. Auroux[31] :

> Des noms tous seuls énoncés bout à bout ne font donc jamais un discours, pas plus que des verbes énoncés sans l'accompagnement d'aucun nom.[32]

Platon met ici le doigt sur la nécessité de mener de front une étude paradigmatique mais aussi syntagmatique du discours. Celui-ci acquiert dès lors une double dimension — une dimension horizontale (celle des syntagmes) et une dimension verticale (celle des paradigmes), au sein desquelles les locuteurs opèrent des choix et se plient à des contraintes. La réflexion sur la linéarité première du langage, la surface, consiste à déterminer, en profondeur, ces choix et contraintes.

Cependant, pour C. Delmas, une langue se donne forcément dans une linéarité qui tend à obscurcir les principes de structuration de ses énoncés :

> La structuration, bien que présente dans la constitution de l'énoncé, peut être occultée dans le formel de surface : le linéaire.[33]

La prise en compte des deux dimensions paradigmatique et syntagmatique permet de percer la linéarité et constitue une approche réellement hiérarchisée de la construction du discours.

Une notion vaste Le terme de « hiérarchisation » désigne communément à la fois « l'action d'organiser selon une hiérarchie [et] cette organisation ».[34] Cette définition, confondant l'effet et sa cause, confère à la notion une portée extrêmement vaste et témoigne d'un large domaine d'application. Si l'on conçoit l'objectif de la linguistique comme S. Auroux, la notion de hiérarchisation promet d'être très féconde :

[31] AUROUX, Sylvain *et al.* (1996). *La philosophie du langage.* Paris : PUF. (p.25).
[32] PLATON. *Sophiste* : 362a.
[33] DELMAS, Claude (1987). *Structuration abstraite et chaîne linéaire en anglais contemporain.* Paris : CEDEL. (p.4).
[34] REY, Alain & REY-DEBOVE, Josette (dir.) (1994). *Le Petit Robert 1 : dictionnaire alphabétique et analogique de la langue française.* Paris : Dictionnaires Le Robert. (p.1091).

> [...] le but du linguiste est de rendre compte des relations qu'ont les signes linguistiques entre eux, en tant qu'ils donnent lieu aux énoncés des langues naturelles.[35]

La fécondité du concept de hiérarchisation en linguistique lui permet d'être pertinent aussi bien au niveau de la description de l'organisation d'unités de même ordre et de divers ordres qu'au niveau de la réflexion générale sur les faits auxquels le linguiste cherche à donner sens. Elle peut correspondre, d'un côté, au constat d'une structure, et, de l'autre, à l'imposition d'une grille de lecture sur un tout à interpréter. La notion entre ainsi dans le projet initial de nombreux modèles dont nous allons décrire la contribution à notre réflexion au cours de ce travail. Etablir et rendre compte de hiérarchies syntaxiques est ainsi au cœur du modèle distributionnel et du modèle génératif. Sur le plan pragmatique, le modèle communicationnel, appliquant cette notion de façon plus large, propose une vision hiérarchique de l'interaction verbale découpée en unités de diverses tailles.

Notre recherche a pour ambition d'explorer la notion de hiérarchisation des constituants du discours aussi bien au sens strict des relations syntaxiques qu'au sens plus large des relations sémantiques, thématiques, pragmatiques, ou encore interactionnelles. Ainsi, la notion trouve pour nous sa pertinence au niveau de la mise en place du sens, de l'organisation thématique du discours, du positionnement énonciatif des parties prenantes du dialogue, ainsi que dans la gestion des tours de parole.

Une fois admis le principe selon lequel le discours s'organise en un système compositionnel hiérarchisé, il nous faut définir plus précisément les éléments qui entrent dans cette organisation. Que désigne le terme de constituants en linguistique ? Dans quelle mesure peut-on parler de constituants propres au discours, c'est-à-dire de constituants discursifs ?

[35] AUROUX, Sylvain *et al.* (1996). *op. cit.* (p.114).

0.1.3.2 Constituants

Les différentes écoles de linguistique s'accordent pour donner au terme de constituant une même définition de base, fidèle au sens commun — celle d'être une partie constitutive d'un tout. Cette première définition générale pose le problème essentiel des critères permettant d'identifier, donc de définir, des constituants spécifiques, les uns par rapport aux autres. Quel rapport les différents constituants entretiennent-ils au sein du tout dont ils font partie ? Quel rapport entretiennent-ils avec ce tout ?

La notion d'unité constitutive

Une notion centrale De même que la hiérarchisation, la notion d'unité constitutive est centrale en linguistique :

> La linguistique, depuis ses débuts, s'est constituée comme une tentative pour dégager les unités élémentaires en question et pour décrire leurs relations.[36]

Se développe ainsi une méthode permettant de définir une unité et ses relations avec les autres éléments du système. Deux opérations de délimitation sont menées parallèlement, sur le plan syntagmatique et sur le plan paradigmatique : segmentation et substitution servent de caution l'une à l'autre. En effet, une unité est atteinte lorsque l'on ne peut plus décomposer l'ensemble en question en segments plus petits. D'autre part, l'unité dégagée doit pouvoir faire sens en elle-même dans le système, et, par exemple, accepter d'être remplacée par une autre.

Une notion relative Un travail sur la notion de constituant qui s'attache à identifier ce à quoi le terme d'« unité » linguistique réfère se heurte d'emblée à une multitude de données de nature très différente. La détermination des unités constitutives d'un corpus admet des constantes indéniables. Cependant, nombre de constituants

[36] REBOUL, Anne & MOESCHLER, Jacques (1998b). *Pragmatique du discours : de l'interprétation de l'énoncé à l'interprétation du discours*. Paris : Armand Colin. (p.22).

ainsi définis sont spécifiques à chaque champ d'investigation, voire à chaque école théorique. En effet, la liste des unités opératoires d'un corpus donné dans un champ donné se conçoit non pas comme un fait objectif ou une donnée brute, mais comme un outil de description, manipulé derrière un prisme particulier. Ainsi, découper en unités, c'est déjà interpréter.[37]

La question de l'adéquation des outils de description à l'objet décrit et à la perspective adoptée pour sa description est présente dans le travail de L. Bouquiaux. La démarche de cet auteur consiste à déterminer des outils pour la description des langues à tradition orale, à partir de ceux utilisés pour les langues à tradition écrite. Son hypothèse de départ est la suivante :

> Loin d'être *a priori* identifiables d'une langue à l'autre, nous pensons [...] que les catégories grammaticales, qui constituent le pivot de toute description syntaxique, doivent être redéfinies dans un cadre donné, pour chaque langue, suivant des méthodes adaptées de celles utilisées pour la définition des phonèmes.[38]

Cette hypothèse trouve un écho certain dans la démarche adoptée dans cette thèse. La langue anglaise est bien entendu une langue à tradition écrite. Pourtant, nous souhaitons montrer ici que la description du fonctionnement de son utilisation spontanée à l'oral ne s'accommode pas d'une simple application des catégories définies pour l'écrit. En outre, si l'hypothèse de L. Bouquiaux ne concerne que la description syntaxique de

[37]Ce fait est par exemple mis en évidence dans ERPICUM, D. & PAGE, M. (éds.) (1986). Dimensions de l'interaction dans une conversation. *Cahiers du Département de Psychologie de l'Université de Montréal* 5, dont C. Kerbrat-Orecchioni rappelle l'expérience dans KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales.* Paris : Armand Colin. (Tome 1, p.220n) :

> [...] travaillant sur *le même* corpus, les différents auteurs en proposent un découpage chaque fois différent... Il est vrai que la perspective, et les objectifs de la description, varient également d'une étude à l'autre. Il est donc tout à fait normal que les découpages divergent en conséquence (selon que l'on s'intéresse plutôt à l'organisation thématique, ou pragmatique, ou bien encore, au problème bien différent de l'évolution du rapport des « places »).

[38]BOUQUIAUX, Luc (1986). Syntaxématique. Définition et classement des unités syntaxiques ou catégories grammaticales ou parties du discours. *Modèles Linguistiques* 8 (1) : 63-68. (p.63).

différentes langues, il nous semble raisonnable de l'étendre aux différentes approches d'une même langue, ainsi qu'aux différents types d'unités descriptives, sans se limiter à la seule syntaxe.

De la sorte, il semble impossible d'adopter sans réflexion préalable l'arsenal descriptif créé pour la description d'un champ de recherche différent, ou dans une optique différente de la nôtre. Le corollaire de cette hypothèse est bien entendu la nécessité de définir notre domaine d'investigation et l'approche théorique que nous souhaitons adopter. C'est l'objet de ce développement introductif. Le rejet en bloc des unités existantes est néanmoins impossible. Il est nécessaire d'envisager dans quelle mesure les unités qui s'offrent à nous sont susceptibles d'être opératoires pour la description de la langue orale dans le cadre théorique que nous nous fixons. Ainsi, bien que relative à chaque domaine d'étude, la notion d'unité lui est à chaque fois centrale car elle en fait la spécificité.

Notre objectif est de décrire une « grammaire » de l'oral qui mêle la syntaxe et la phonologie à la pragmatique, l'analyse de discours, et à l'énonciation. Nous proposons donc de passer en revue les unités pertinentes à ces différents domaines d'étude et de déterminer s'il est possible, et dans quelle mesure, au prix de quel aménagement, de les transposer à ce domaine complexe.

Une notion théorique à l'épreuve du corpus La détermination des constituants opératoires dans un domaine se heurte au passage de la théorie à la pratique où des irrégularités apparaissent nécessairement. La confrontation constante avec le corpus permet de s'assurer de la pertinence de cette liste et des éventuels aménagements à opérer.

L'une des difficultés principales de la détermination des constituants du tout qui s'offre à nous tient tout d'abord au caractère continu de la chaîne sonore. Comment discrétiser des phénomènes qui nous parviennent en bloc ? Formulé plus haut

par M. Rossi *et al.*,[39] c'est le pari de toute les théories de l'intonation (dont nous proposons une description sommaire au chapitre 1) que de décrire le fonctionnement des phénomènes suprasegmentaux — essentiellement la mélodie — en terme de phénomènes distincts et opposables.

En outre, le domaine du discours tel que défini plus haut nous met en présence d'éléments hétérogènes qu'il n'est pas possible de faire figurer dans le même arbre hiérarchique. Le caractère spontané des échanges de notre corpus rend la segmentation en groupes syntaxiques ou en groupes intonatifs difficile. On pense par exemple aux traces du travail de formulation que constituent la pause pleine ou silencieuse, ou l'allongement, ou encore les faux départs, réinitialisations et auto-corrections qui viennent contrarier la formation des constituants traditionnels. L'hétérogénéité des constituants considérés provient également du fait que notre attention se porte à différents niveaux de structuration de l'échange verbal.

Nous nous préparons donc à être confrontée à des problèmes relatifs à la délimitation des unités ainsi qu'à la détermination de leurs relations. L'analyse de la conversation spontanée se situe au confluent de plusieurs domaines. Dans quelle mesure les unités opératoires de ces domaines s'appliquent-elles au nôtre ?

Les constituants du discours Parler de constituants suppose tout d'abord la possibilité de rendre discrets les éléments qui nous parviennent comme formant un continuum. Quelles unités opératoires retenons-nous pour décrire le tout dont elles participent ?

[39]ROSSI, Mario *et al.* (1981). *L'intonation, de l'acoustique à la sémantique.* Paris : Klincksieck. (p.100). Voir plus haut, p.31.

Du phonème au mot Pour R. Jakobson & L. Waugh,[40] et pour P. Delattre[41] avant eux, les **phonèmes** constituent la première unité de sens. Ils les considèrent comme les « constituants ultimes du langage ».[42] D'un point de vue sémantique, le phonème nous intéresse lorsque, isolé, il fait percevoir un mot tronqué dont il fait partie. C'est le cas, par exemple, de ce que nous notons *b-* et que nous comprenons en divers endroits comme un *but*. Il s'agit donc ici d'une place marginale du phonème qui ne fait sens qu'en tant qu'avatar d'un mot. Plus essentielle est la pertinence de cette unité de l'oral sur le plan énonciatif et interactionnel. Un phonème isolé, faisant signe vers un mot ou non, exprime déjà la présence du locuteur, sa capacité à prendre la parole, ou encore son intérêt pour ce que dit l'autre. On pense aux régulateurs et aux marques d'écoute qui sont souvent réduits à un phonème. Notre approche de la langue orale n'est pas phonologique. Un phonème donné n'est pas seulement signifiant par opposition à tel autre (sur la base de traits pertinents qui les distinguent), mais aussi par opposition au silence qui l'entoure. De ce fait, toute manifestation vocale fait sens en soi, indépendamment de son inscription dans le système de la phonologie.

Si notre approche n'est pas non plus morphologique, elle emprunte à ce domaine son unité minimale.

De même que le phonème, le **morphème** nous intéresse de façon marginale en tant qu'avatar du mot. D'un point de vue structurel, en revanche, il détermine les frontières des segments susceptibles d'être répétés en cas d'hésitation de la part du locuteur. Ainsi, dans les deux exemples suivants :

C hm {0,45} <h> I can {0,75} <h> I can **introd- duce** you to the alarm system and::

(Celia&James : Keys, *introd- duce*)

J *well I suppose they're **pro- promised*** they might meet other Westerners

[40] JAKOBSON, Roman & WAUGH, Linda (1979, trad. fr. 1980). *La charpente phonique du langage.* Paris : Editions de Minuit.

[41] DELATTRE, Pierre (1972). The Distinctive Function of Intonation. in Bolinger, Dwight (éd.), *Intonation, selected reading* : 159-174. Harmondsworth : Penguin Books.

[42] JAKOBSON, Roman & WAUGH, Linda (1979, trad. fr. 1980). *op. cit.* (p.101).

(Celia&James : English corner, *pro- promised*)

En ce sens, le morphème contraint la structure du linéaire telle qu'elle nous parvient.

Nous empruntons également à la phonologie l'unité de la syllabe. De façon marginale, cette unité nous permet d'étiqueter des unités supérieures au phonème mais inférieures au morphème.[43] Plus essentiellement, la syllabe nous intéresse pour la description de certains phénomènes prosodiques. L'alternance des syllabes accentuées et inaccentuées permet la perception du rythme. D'autre part, les syllabes accentuées reçoivent typiquement un surcroît d'intensité et constituent le site d'occurrence privilégié (mais non exclusif) des variations mélodiques.

Emprunté à la syntaxe et à la sémantique, le **mot** constitue bien entendu une unité privilégiée de notre analyse dans la mesure où il permet la mise en place des sèmes. En outre, d'un point de vue phonologique, une syllabe est accentuée parce qu'elle fait partie d'un mot accentué. L'accentuation est d'une part fonction de la catégorie grammaticale à laquelle le mot appartient (mot lexical, par opposition à mot grammatical), mais peut affecter n'importe quel mot pourvu que sa mise en relief fasse partie d'un projet discursif ou intersubjectif. C'est sur le mot et le syntagme nominal dont il est la tête que se porte l'essentiel du travail de la construction discursive. C'est donc sur cette unité que se fondent les hiérarchies les plus pertinentes à notre étude.

Du syntagme à l'énoncé Unité de rang intermédiaire, le **syntagme** nous intéresse dans la mesure où il regarde à la fois vers le rang inférieur où se situe le mot, et vers le rang supérieur que constitue l'énoncé. Sur le plan sémantique, il permet l'approfondissement des sèmes contenus dans les mots isolés (un nom s'y en-

[43] Voir l'exemple suivant :

C is tha::t*
J *yeah **vo-***
C just Chinese *or is that everybody*
J *yes they have er* they have much

(Celia&James : Facilities, *vo-*)

toure par exemple d'un adjectif). Syntaxiquement, c'est le lieu où s'expriment les déterminations, et donc un certain type de relations entre les sèmes. Sur le plan discursif, le syntagme est une unité centrale puisqu'elle peut être déplacée, manipulée selon les principes de la gestion de l'information.

La **proposition**, unité subsumant le syntagme, est construite autour d'un verbe, conjugué ou non. Cette unité est utile à notre étude pour des raisons proches de celles décrites pour le syntagme. En outre, sa capacité à réunir un sujet et un prédicat la rend propice à une utilisation indépendante. Nous pensons notamment aux relatives rhématiques, forme courante à l'oral.

L'**énoncé**, que nous préférons à la phrase pour les raisons évoquées plus haut, est l'unité au sein de laquelle les relations prédicatives trouvent leur validation, en fonction des coordonnées de l'énonciation. C'est une unité fondamentale dans notre étude. Sur le plan syntaxique s'y exprime un niveau supérieur de déterminations, entre sujet et prédicat. D'un point de vue discursif, les unités de rang inférieur y sont agencées conformément aux principes du dynamisme communicatif qui conçoit une division en deux pôles : thème/rhème, ou topique/focus.

Du paragraphe oral au tour de parole Le **paragraphe oral** est l'unité au sein de laquelle s'organisent éléments du cadre thématique et contenu rhématique. Le paragraphe oral occupe une place centrale dans la théorie de M.-A. Morel & L. Danon-Boileau[44] élaborée pour le français. Nous verrons dans quelle mesure cette unité reste opératoire pour la description de l'anglais.

Enfin, la dernière unité que nous considérerons, en-deçà de celle, conjoncturelle, formée par chacune des conversations enregistrées, est le **tour de parole**. Intuitivement aisée à définir, cette unité interactionnelle échappe aux définitions qui cherchent à lui donner des frontières strictes.

[44]MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation. L'exemple du français.* Gap : Ophrys.

0.1.3.3 Bilan

Unités constitutives et hiérarchisation sont indissolublement liées, la hiérarchie ne pouvant se comprendre sans constituants, et les unités prenant sens uniquement les unes par rapport aux autres. Un travail sur la hiérarchisation des constituants discursifs doit donc prendre en compte la complémentarité de ces niveaux d'analyse au sein du système qu'ils contribuent à édifier.

Système et énonciation Les linguistes s'entendent depuis plusieurs siècles pour dire que la langue est un système, voire, selon G. Guillaume « un système de systèmes ».[45] La description de ce système est au cœur d'un projet préoccupé par l'organisation hiérarchisée, la structuration des constituants du discours. En ce sens, notre recherche trouve certaines de ses racines dans le structuralisme défini ici par E. Benveniste :

> Le principe fondamental est que la langue constitue un système, dont toutes les parties sont unies par un rapport de solidarité et de dépendance. Ce système organise des unités, qui sont les signes articulés, se différenciant et se délimitant mutuellement. La doctrine structuraliste enseigne la prédominance du système sur les éléments, vise à dégager la structure du système à travers les relations des éléments, aussi bien dans la chaîne parlée que dans les paradigmes formels, et montre le caractère organique des changements auxquels la langue est soumise.[46]

Dans quelle mesure la langue orale spontanée fait-elle également système ? En quoi diffère-t-elle de la langue écrite ? Une linguistique du discours tel que défini plus haut se conçoit en ce sens comme une étude des liens entre les éléments qui le constituent. Au sein de ce système, elle s'intéresse également à la place du sujet

[45] GUILLAUME, Gustave (1952). La langue est-elle ou n'est-elle pas un système ? *Cahiers de Linguistique Structurale* 1 : 3-30. (p.7).

[46] BENVENISTE, Emile (1966, 1974). *Problèmes de linguistique générale, I & II.* Paris : Gallimard. (p.98).

parlant et à la latitude dont celui-ci dispose. Quels types de choix se présentent à l'énonciateur ? Si l'on considère que la modification d'un élément ou de la relation entre deux éléments entraîne un rééquilibrage au sein du système, les choix des locuteurs sont-ils susceptibles de transformer le système ?

La chimie de la langue orale Notre corpus d'oral spontané suscite des questions supplémentaires. Comment le système s'accommode-t-il de la double dimension de la langue orale ? En quoi segmental et suprasegmental contribuent-ils au système ? Ce n'est pas tant un système bien réglé que nous nous donnons à décrire que certains aspects de la chimie de la langue orale. La prosodie, qui se superpose à la syntaxe, tend à délimiter des constituants de l'oral et à signaler leur statut au sein de l'ensemble. Les phénomènes intonatifs et les phénomènes syntaxiques, de natures différents, se combinent non pas tant à la façon de constituants que comme des ingrédients dont le mélange est bien plus que la somme de ses composants. C'est le sens de la remarque de L. Tesnière à propos de la relation prédicative :

> [...] une phrase du type *Alfred* parle n'est pas composée de **deux** éléments 1° *Alfred*, 2° *parle*, mais bien de **trois** éléments, 1° *Alfred*, 2° *parle* et 3° la connexion qui les unit et sans laquelle il n'y aurait pas de phrase. Dire qu'une phrase du type Alfred parle ne comporte que deux éléments, c'est l'analyser de façon superficielle, purement morphologique, et en négliger l'essentiel, qui est le lien syntaxique. [...] Il en va de même en chimie, où la combinaison du chlore Cl et du sodium Na fournit un composé, le sel de cuisine ou chlorure de sodium NaCl, qui est un tout autre corps et présente de tout autres caractères que le chlore d'une part et le sodium Na d'autre part.[47]

Le lien essentiel, dans notre étude, se situe non seulement au niveau syntaxique, mais également sémantique, thématique et intersubjectif. Nous montrerons en ef-

[47] TESNIERE, Lucien (1959,1988). *Eléments de syntaxe structurale.* Paris : Klincksieck. (pp.11-12).

fet que ce lien change de nature en fonction du niveau auquel nous nous plaçons, conformément à l'économie particulière de l'oral.

Notre travail s'intéresse donc à la façon dont des phénomènes de divers ordres se complètent pour contribuer à l'effet d'un tout régulier. Il cherche également à mettre en lumière les irrégularités et les redondances de ce système.

0.2 Plan général de l'étude

Notre **première partie** cherche à établir le cadre théorique et méthodologique dans lequel ce travail se situe.

Il s'agit tout d'abord, dans le chapitre initial, de mesurer l'apport des études antérieures en retraçant rapidement les étapes de la constitution de l'étude de l'intonation de la langue anglaise en tant que domaine de recherche à part entière. Nous évoquons ensuite les modèles existants pour la description et l'interprétation des phénomènes intonatifs, ainsi que ceux relatifs à la langue et au discours.

Dans un deuxième temps, nous présentons les principes théoriques adoptés dans notre recherche. Nous nous attardons sur le modèle de M.-A. Morel & L. Danon-Boileau,[48] établi pour le français, dont nous envisageons les possibilités d'adaptation à la description de l'anglais oral spontané. Ce chapitre est également l'occasion d'énoncer nos hypothèses de départ. La première concerne la syntaxe de l'oral, dont nous faisons l'hypothèse qu'elle est sensiblement différente de celle de l'écrit. Sa moindre densité et sa moindre variété permettrait ainsi aux paramètres intonatifs de jouer un rôle compensateur et structurant. C'est ainsi qu'il nous semble possible de postuler des fonctionnements prosodiques interprétables à différents niveaux pour chacun des paramètres. Notre dernière hypothèse s'inspire de la théorie mathématique des objets fractals et voit dans l'oral spontané un système dont la cohésion tient au

[48] MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *op. cit.*

moins en partie à la répétition du même phénomène à différentes profondeurs de structuration.

La présentation des données sur lesquelles nous nous proposons de vérifier nos hypothèses occupe le dernier chapitre de cette première partie. Celui-ci explique les enjeux et les modalités de la constitution d'un corpus d'oral spontané, et décrit le traitement que nous avons appliqué à notre corpus de conversations pour établir nos données de départ.

Notre **deuxième partie** s'attache à décrire le fonctionnement de quelques phénomènes de structuration hiérarchisée observables sur le plan segmental.

Dans un souci distributionnel, nous nous intéressons en premier lieu aux marqueurs de subordination tels qu'ils sont représentés dans notre corpus, pour nous concentrer ensuite sur le cas des circonstancielles de cause et en particulier sur les marqueurs *because* et *cos* auxquels nous attribuons un fonctionnement distinct.

Enfin, le chapitre 6 propose d'étudier la façon dont la subordination syntaxique et thématique sert l'organisation du discours en assurant notamment la délimitation d'unités argumentatives ou thématiques de statuts différents et en favorisant la progression discursive à la façon d'une spirale, par un constant retour en arrière.

Enfin, notre **troisième et dernière partie** s'intéresse à la façon dont la structuration de la conversation spontanée est liée à la prosodie, et notamment à certains phénomènes en particulier.

Au nombre de ces phénomènes, nous comptons le fonctionnement du paramètre de hauteur dans la gestion de l'espace intersubjectif. S'appuyant sur la statistique et sur l'analyse linguistique, le chapitre 7 montre une relation entre état de la coénonciation et plage mélodique utilisée.

Le chapitre suivant rend compte du rôle des pauses silencieuses dans la construction du discours, en se fondant notamment sur une expérience visant à évaluer l'impact

de la durée et de la fonction des pauses sur leur perceptibilité.

Enfin, c'est vers la gestion des tours de parole que se tourne notre attention, et plus spécifiquement sur une des particularités de l'oral spontané : les chevauchements de parole. En nous attachant à définir les caractéristiques syntaxiques, sémantiques et prosodiques des segments prononcés en situation de recouvrement de parole, nous chercherons à montrer que les chevauchements de parole ne témoignent pas d'un problème de gestion de l'interaction mais constituent une richesse de la langue conversationnelle.

Première partie

Cadre théorique et méthodologique

L'objectif de ce premier développement est d'exposer les bases théoriques et méthodologiques sur lesquelles se fonde la recherche que nous présentons ici.

Sur le plan théorique, toute réflexion linguistique se situe nécessairement par rapport à une tradition. L'objet d'étude que nous nous donnons nécessite de considérer à la fois la tradition des études des « éléments musicaux »[49] de la langue, et celle des études linguistiques proposant une description de la langue ou du discours. Un tour d'horizon des études antérieures doit permettre de déterminer quels modèles existants prendre en compte. Dans quelle mesure peut-on conserver leurs hypothèses, leurs méthodes et leurs résultats ?

Nous décrivons ensuite le modèle théorique que nous retenons, et anticipons un certain nombre d'aménagements susceptibles d'être nécessaires pour appliquer ce modèle à notre objet d'étude. Notre deuxième chapitre présente également les trois hypothèses de départ qui définissent notre projet.

Enfin, il s'agit de convoquer les outils et méthodes que nous nous donnons pour vérifier ces trois hypothèses. La présentation du corpus conversationnel qui constitue notre réservoir d'exemples s'accompagne d'une réflexion sur la méthodologie adoptée pour son enregistrement et pour l'établissement de la transcription et des courbes d'analyse prosodique figurant dans le second volume de ce travail.

[49] FAURE, Georges (1962). *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise.* Paris : Didier.

Chapitre 1

Etudes antérieures

L'analyse de la hiérarchisation des constituants discursifs que nous proposons dans ce travail se situe au confluent d'une double tradition. Elle s'inscrit en effet dans la longue tradition des études portant sur l'intonation de la langue. D'abord liée à la musique, l'intonation devient un domaine de recherche à part entière, et développe des outils et méthodes qui lui sont propres, ou qu'elle emprunte à l'acoustique, la psychoacoustique ou à la phonétique. Dans un second temps, nous décrirons un certain nombre de modèles relevant de la tradition linguistique de description de la langue et du discours. Il s'agira d'évaluer dans quelle mesure ils sont pertinents pour la description de l'objet d'étude que nous nous donnons.

1.1 L'intonation dans la tradition des études linguistiques

1.1.1 Quelques points de repère chronologiques

Les premiers travaux témoignant d'un intérêt pour l'intonation des langues s'attachent à décrire la langue orale en termes musicaux. L'apparition du mot en français,

à la fin du Moyen-Age, le lie directement à la musique. L'intonation désigne alors les formules mélodiques de quelques signes servant à rappeler aux chanteurs le début des psalmodies du chant grégorien. Le terme a conservé son rapport avec la mélodie musicale, dans la mesure où il fait aujourd'hui plus généralement référence à la justesse des notes jouées ou chantées, et à leur accord avec l'harmonie. En tant que domaine de recherche, néanmoins, il s'est progressivement détaché de son origine musicale.

1.1.1.1 Du Moyen-Age au XVIIe siècle : les précurseurs

Chant liturgique et notation musicale au Moyen-Age Dès avant le IXe siècle, la musique médiévale européenne se dote d'un système de notation plaçant le verbe et sa signification au cœur de son esthétique. Proche de l'écriture, ce système emprunte à celle-ci accents, lettres de l'alphabet et marques de ponctuation.

Le chant religieux byzantin fait usage d'une notation ekphonétique mimant les mouvements de la voix chantée, que l'on retrouve, en occident, dans la notation neumatique *a campo aperto* (ou à champ ouvert) du chant grégorien. Les neumes (du grec *pneuma*, souffle) sont eux aussi des signes graphiques qui, sous la forme de points et d'accents placés au-dessus des syllabes du mot, indiquent au chanteur la direction de la ligne mélodique. Avec un certain nombre de variations selon les régions, on distingue deux signes de base correspondant respectivement à un ton montant, plus aigu, et à un ton descendant ou plus grave : l'*accentus acutus* (ou *virga*), et le *accentus gravis* (ou *punctum*). Ces deux tons se combinent pour former des tons complexes de deux, trois ou quatre notes. Mentionnons simplement, en plus des deux tons simples, le *clivis* (ton montant-descendant) ou le *podatus* (ton descendant-montant) pour leur proximité avec les inflexions de la langue anglaise décrites au XXe siècle par la théorie des contours intonatifs.[1]

Les neumes sont au départ de simples moyens mnémotechniques destinés à faciliter

[1] Voir p.57 et p.61.

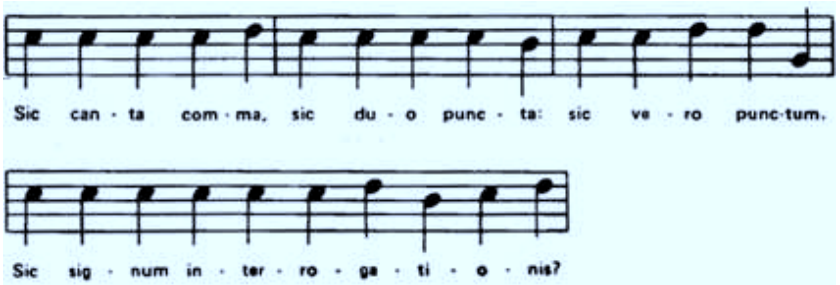

FIG. 1.1 – Manuscrit de Munster : formulaire pour la récitation liturgique

l'exécution d'une mélodie déjà connue du chanteur. Nombre de formulaires rappellent les schémas mélodiques de base auxquels est associée la ponctuation neumatique, comme celui du manuscrit de Munster dont la Figure 1.1 présente une version moderne citée par K. Hadding-Koch[2] et C. Goodwin[3] :

Les indications données par les neumes permettent ainsi de se remémorer la direction générale de la ligne mélodique, sans renseigner sur l'intervalle entre les notes, ni véritablement sur le rythme de celles-ci. De plus en plus riche, la notation médiévale, qui cherche en fait à indiquer plus en détail les paramètres du son, développe les moyens de son ambition. Ainsi, un certain nombre de signes supplémentaires (*épisèmes* et « lettres significatives ») permettent de noter des nuances accentuelles, rythmiques ou dynamiques :

> Dès le IXe s., on éprouve le besoin de préciser ce qui était vague dans l'intonation, l'accentuation, le rythme des neumes. D'où l'addition de traits — épisèmes — à certains neumes, pour indiquer l'appui de certaines notes, la durée de certaines autres, des liaisons entre les groupes, etc. De là aussi l'addition de « lettres significatives » : *a* voulant dire augmentez ou avec ampleur ; *c*

[2]HADDING-KOCH, Kerstin (1961). *Acoustico-Phonetic Studies in the Intonation of Southern Swedish.* Lund : C.W.K. Gleerup. (p.9).

[3]La figure provient de GOODWIN, Charles (1981). *Conversational Organization : Interaction between Speakers and Hearers.* New York : Academic Press. (p.48). On y lit : :

> Sic canta comma, sic duo puncta : sic vero punctum. Sic signum interrogationis ?

= *celeriter*, brièvement ; *e* = *equaliter*, unisson ; *f* = *frange*, avec renforcement ; *i* = *iusum*, plus bas ; *m* = modérément, etc.[4]

Faisant ainsi moins appel à la mémoire du chanteur, les neumes à champ ouvert tendent à s'inscrire sur la page à des hauteurs différentes au-dessus du texte, mimant véritablement une ligne mélodique : les notes de même hauteur sont positionnées sur une même ligne imaginaire. A partir du XIe siècle, cette ligne se matérialise pour faciliter la lecture. L'apparition de la portée (d'abord à une ou deux lignes) permet de noter avec davantage de précision la hauteur des sons et les intervalles. Les marques stylisées que sont les neumes se déforment pour s'inscrire dans la portée et devenir des notes de forme carrée.

D'abord pauvre en notations rythmiques, le chant grégorien épouse néanmoins la métrique du mot. Le musicologue P. Wagner montre en effet que l'opposition entre longues et brèves présente dans l'écriture musicale aux XIe et XIIe siècles correspond à celle de la langue latine, utilisée dans la liturgie.[5] Cette opposition débouche sur la distinction rythmique consignée dans la notation carrée entre notes de valeurs différentes.

Voix parlée et voix chantée à la Renaissance L'humanisme renaissant approfondit la réflexion sur la relation entre les inflexions du chant et celles de la parole, qui aboutit en musique à la naissance de l'opéra au tournant des XVIe et XVIIe siècles en Italie. Parallèlement, se développe en grammaire un intérêt pour la ponctuation et son rapport avec la prosodie de la voix parlée.

D'après K. Pike,[6] repris par G. Faure dans son ouvrage sur les éléments musicaux

[4]BRENET, Michel (1926). *Dictionnaire pratique et historique de la musique.* Paris : Armand Colin. (p.296).

[5]WAGNER, Peter (1895, 1905, 1921). *Einführung in die gregorianischen Melodien.* 3 vols. Fribourg, Leipzig.

[6]PIKE, Kenneth L. (1945). *The Intonation of American English.* Ann Arbor, MI : University of Michigan Press.

de la langue anglaise,[7] la première allusion connue à la mélodie de l'anglais dans un ouvrage de grammaire date de la moitié du XVIe siècle, dans le travail de J. Hart sur la ponctuation.[8] Celui-ci se donne comme problématique celle de l'imitation par l'orthographe et la ponctuation de la prosodie naturelle de la voix humaine. J. Hart s'intéresse ainsi plus particulièrement à la mélodie de la voix et aux pauses silencieuses.

En 1634, C. Butler observe dans sa grammaire de l'anglais[9] une différence intonative entre les énoncés déclaratifs, les énoncés interrogatifs et les énoncés incomplets. L'énoncé déclaratif se termine sur une chute de la voix. La montée finale est réservée à certaines phrases interrogatives et aux phrases inachevées. Il note une distinction importante entre les questions ouvertes, marquées par une montée mélodique finale, et les questions fermées (en *wh-*), pour lesquelles cette montée est absente. Il propose également une révision de l'orthographe dont le but est d'établir une meilleure adéquation entre l'image graphique et l'image phonique des mots. Comme J. Hart, C. Butler rapproche intonation et ponctuation. Il montre l'existence de quatre tons (*level, fall, extra-low* et *fall*) et de trois degrés de pauses qu'il relie aux différences existant entre les marques de ponctuation. Dans son chapitre sur l'histoire des études sur l'intonation jusqu'à la fin du XXe siècle, A. Cruttenden rend compte de son attention pour l'intonation spécifique des énoncés parenthétiques ou incidents, sur une note basse.[10]

Ainsi, la réflexion sur la notation de la monodie grégorienne et celle des grammairiens de la Renaissance contiennent en germe tous les éléments de l'étude de l'intonation de la voix parlée qui se développe aux siècles suivants.

[7]FAURE, Georges (1962). *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise.* Paris : Didier.

[8]HART, John (1569). *An orthographie : Conteyning the Due Order and Reason Howe to Write and Paint Thimages of Manes Voices, most like to Life or Nature.* Londres : William Seres.

[9]BUTLER, Charles (1634, 1910). *English Grammar.* Eitchler, A. (éd.). Niemeyer : Halle A. S.

[10]CRUTTENDEN, Alan (1986, 1997). *Intonation.* Cambridge : Cambridge University Press.

1.1.1.2 XVIIIe et XIXe siècles : la naissance de deux écoles influentes

Place grandissante de l'intonation dans les grammaires Au XVIIIe siècle apparaissent de nombreux manuels d'élocution, fondés sur la similitude entre voix parlée et voix chantée. G. Faure[11] nous rapporte les deux essais suivants, dont l'apport concernant l'étude des phénomènes intonatifs est intéressant.

Dans sa *Prosodia Rationalis*,[12] J. Steele distingue cinq paramètres prosodiques indépendants : l'accent (impliquant la notion de hauteur et de variation de la hauteur mélodique), la quantité (qui recouvre la durée et l'allongement), la pause, la force (c'est-à-dire l'intensité), ainsi qu'un dernier paramètre, l'emphase, qui correspond probablement à une conjonction des paramètres mélodiques et intensifs. La notation de la mélodie et de ses variations se fait sur une portée musicale avec un grand souci pour le détail.

G. Faure mentionne ensuite l'important travail de J. Walker, dont la préoccupation pédagogique pour l'établissement de principes d'élocution a permis de poser les bases des études suprasegmentales. On retient notamment de lui deux ouvrages majeurs, les deux volumes de ses *Elements of elocution* (1781), et *The Melody of Speaking, Delineated ; or Elocution taught like music, by visible signs* (1787).[13] Sa contribution principale est d'avoir développé une « théorie des inflexions » et un système de notation adapté. Il distingue cinq tons différents, appelés *slides*, en fonction de l'évolution de la hauteur de la voix. Il met au point un système de représentation visuelle à l'aide de signes diacritiques pour noter ces tons — un ton statique : *monotone*, et quatre tons mélodiques : *rise, fall, fall-rise* et *rise-fall*. En outre, on voit apparaître la notion

[11] FAURE, Georges (1962). *op. cit.*

[12] STEELE, Joshua (1775, 2e éd. 1971). *Prosodia Rationalis, or An Essay towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols.* Olms : Georg Publishers

[13] WALKER, John (1781, 1974). *Elements of elocution.* Londres : Scolar Press.
WALKER, John (1787, 1974). *The Melody of Speaking, Delineated ; or Elocution taught like music, by visible signs.* Londres : Scolar Press.

de registre, qui rend les intervalles et la hauteur relative de la voix plus significatifs que sa hauteur absolue. Les variations mélodiques sont envisagées de façon indépendante pour chaque individu. A. Cruttenden fait de ce chercheur l'un des précurseurs de la notion d'information donnée par le contexte ou par la situation, dont il étudie l'impact en termes intonatifs.[14]

Jusqu'au XVIIIe siècle, les approches de l'intonation des langues, et en particulier de l'anglais, sont principalement descriptives. Elles témoignent d'une très grande influence des théories musicales, tant pour la description des phénomènes prosodiques que pour leur notation. La problématique est celle de l'imitation, sans se préoccuper des fonctions de l'intonation. Cependant, l'intonation de la langue orale confirme ainsi sa place dans les grammaires de l'anglais. En outre, le travail sur les outils descriptifs et la précision des descriptions qui en découlent contribuent à l'intérêt que présentent les contributions des chercheurs du XVIIIe siècle en particulier. C'est le sens de la remarque de G. Faure :

> [...] il semble bien que dès la fin du XVIIIème siècle, l'étude de l'intonation anglaise ait non seulement retenu l'attention d'authentiques chercheurs, mais les ait aussi amenés à poser, souvent en termes excellents, tout un ensemble de problèmes qui font encore l'objet de travaux et de discussions, et même à proposer à leur sujet des essais de solution qui ne seraient pas déplacés dans des ouvrages beaucoup plus récents.[15]

L'intonation comme domaine de recherche Le XIXe siècle voit la naissance, de part et d'autre de l'Atlantique, des fondements des deux écoles théoriques les plus influentes jusqu'à nos jours. L'étude des éléments suprasegmentaux accède au rang de domaine de recherche spécifique, au sein duquel s'ajoute une préoccupation

[14]CRUTTENDEN, Alan (1986, 1997). *Intonation.* Cambridge : Cambridge University Press.

[15]FAURE, Georges (1962). *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise.* Paris : Didier. (p.19).

interprétative au souci descriptif des siècles précédents.

Aux Etats-Unis, les études intonatives sont essentiellement dominées par J. Rush, dont s'inspirent la plupart des traités d'élocution américains.[16] Il considère que les signes vocaux (*vocal signs*) constituent une représentation fidèle de la pensée et des émotions du locuteur. Sa problématique principale est donc de tenter d'établir de façon précise et rigoureuse le lien existant entre ces signes et les innombrables nuances de la pensée humaine. Il assigne à l'intonation une valeur de témoin de l'état psychologique du locuteur. S'il reprend la distinction entre tons statiques et ton mélodiques qu'il affine, sa contribution principale réside, pour A. Cruttenden,[17] dans le fait qu'il soit le premier à noter l'importance de la variation mélodique affectant la finale des énoncés. Il annonce ainsi les *terminals* du siècle suivant, utilisés par les chercheurs américains. En outre, il établit une distinction importante entre les énoncés « expressifs » et les énoncés « inexpressifs », préfigurant ainsi la distinction moderne entre énoncés « marqués » et « non marqués ».

Les avancées de la recherche sur l'intonation en Grande-Bretagne sont dues, à la même époque, aux travaux de la famille Bell. On retient en particulier ceux d'Alexander Melville Bell, dont le père et le frère se sont intéressés à la phonétique, et dont le fils a inventé le téléphone. Ses recherches sur l'élocution[18] l'amènent à travailler sur l'accent, et à montrer qu'il coïncide avec une inflexion de la ligne mélodique. D'un point de vue mélodique, il définit ces inflexions selon des termes similaires à ceux des chercheurs du siècle précédent. Simples (*fall, rise, monotone*) ou composées (*rise-fall, fall-rise* et même *rise-fall-rise*), les inflexions peuvent être prononcées sur un mode aigu (*high*) ou un mode grave (*low*). Ces précisions lui permettent de proposer une

[16]RUSH, James (2e éd. 1833). *The Philosophy of the Human voice, Embracing its Psychological history, together with a system of principles by which criticism in the Art of Elocution may be Rendered intelligible.* Philadelphie : Grigg and Elliott.

[17]CRUTTENDEN, Alan (1986, 1997). *op. cit.*

[18]BELL, Alexander Melville (1859). *The elocutionary manual.* Londres : Hamilton, Adams and Co.

théorie détaillée concernant le sens associé à l'inflexion finale d'un énoncé. Dans son ouvrage de 1867 intitulé *Visible Speech*,[19] il contribue également à la réflexion sur la transcription des phénomènes suprasegmentaux entamée par J. Walker[20] au siècle précédent (dont il retient un certain nombre de principes).

Les XVIIIe et XIXe siècles contiennent ainsi en germe la division entre les adeptes américains de la théorie des niveaux et les tenants anglais de la théorie des contours.

1.1.1.3 XXe et XXIe siècles : la diversification des approches

Les tendances divergentes s'affirment au XXe siècle quant à la perspective adoptée par les chercheurs américains et celle des britanniques.

Parmi les chercheurs britanniques les plus influents, on compte tout d'abord D. Jones,[21] qui, au début du XXe siècle, grâce à l'utilisation du phonographe récemment apparu, propose un modèle de transcription de l'intonation. Celui-ci se fonde sur deux schémas intonatifs de base : le *fall*, et le *rise*, auxquels il ajoute le *fall-rise* non marqué et le *fall-rise* emphatique. D. Jones mentionne également un ton statique (*level intonation*). Ces contours intonatifs sont repris et explorés plus avant par H.E. Palmer,[22] ainsi que L.E. Armstrong & I.C. Ward.[23] Les principaux continuateurs de ces fondateurs de l'approche britannique (sur les traces de J. Walker et plus tard de A.M. Bell) sont R. Kingdon,[24] J.D. O'Connor & G.F. Arnold,[25] D. Crystal,[26] et A. Cruttenden.[27]

[19]BELL, Alexander Melville (1867). *Visible Speech, the science of universal alphabetics, or self-interpreting physiological letters.* Londres : Simpkin, Marshall and Co.

[20]WALKER, John (1787, 1974). *op. cit.*

[21]JONES, Daniel (1909, 1956). *The Pronunciation of English.* Cambridge : Cambridge University Press.

[22]PALMER, Harold E. (1924). *A Grammar of Spoken English.* Cambridge : Heffer.

[23]ARMSTRONG, Lilias Eveline & WARD, Ida Caroline (1926). *A Handbook of English Intonation.* London : Heffer.

[24]KINGDON, Roger (1958). *The groundwork of English intonation.* London : Longman.

[25]O'CONNOR, Joseph Desmond & ARNOLD, Gordon Frederick (1961, 1973). *Intonation of Colloquial English.* London : Longman.

[26]CRYSTAL, David (1969). *Prododic Systems and Intonation in English.* Cambridge : Cambridge University Press.

[27]CRUTTENDEN, Alan (1986, 1997). *op. cit.*

Tous ces chercheurs ont en commun d'attribuer une signification à la direction de la pente mélodique, indépendamment de son amplitude. A. Cruttenden justifie l'inscription de ses recherches au sein de la théorie des contours dans les termes suivants :

> [...] the decision to allow an analysis based on contours to predominate in this book remains to some extent an act of faith. But there is at least one factor which tips the balance towards a contour analysis: this is my conviction that there is some basic similarity of meaning in all falling contours as opposed to a basic similarity of meaning in all rising contours [...].[28]

Ce modèle considère l'intonation comme un système de syllabes accentuées, chacune d'entre elles étant affectée d'un ton (montant, descendant, ou statique). La mélodie des syllabes inaccentuées importe peu et découle de celle des syllabes accentuées. La représentation graphique se fait entre deux lignes horizontales à l'aide de points ou de tirets symbolisant la hauteur des syllabes situées dans la tessiture du locuteur.

L'école américaine développe dans la deuxième moitié du XXe siècle une théorie des niveaux mélodiques. Ses représentants sont notamment K. Pike,[29] et G.L. Trager & H.L. Smith.[30] Le modèle de représentation de la théorie des niveaux utilise une échelle de niveaux phonémiques (la plupart du temps au nombre de quatre). Chaque niveau est représenté par un chiffre, le plus grand correspondant en général à la hauteur mélodique la plus élevée. Une intonation descendante, par exemple, est notée comme suit : /21/, /31/, /41/, /32/... Un certain nombre d'éléments de notation viennent ponctuer les groupes intonatifs, notamment les pauses, ou l'allongement de la syllabe précédente, représentés à l'aide de signes diacritiques.

Certains commentateurs voient derrière cette opposition entre chercheurs britanniques et chercheurs américains une différence de conception de l'intonation en tant

[28] *ibid.*

[29] PIKE, Kenneth, L. (1945). *The Intonation of American English.* Ann Arbor, MI : University of Michigan Press.

[30] TRAGER, George Leonard & SMITH, Henry Lee (1951). *An outline of English Structure.* Norman, OK : Battenberg Press.

qu'objet d'étude. L'approche britannique est souvent qualifiée de descriptive et empirique, alors que l'approche américaine serait davantage fonctionnelle et théorique. Il est vrai que D. Jones[31] lui-même présente son travail comme celui d'un phonéticien, et non pas celui de quelqu'un se préoccupant spécifiquement de l'intonation de la langue anglaise. Son chapitre consacré à l'intonation de l'anglais s'inscrit dans le cadre d'un travail dont l'objectif est d'établir plus généralement une description de la prononciation de l'anglais. Cependant, les deux écoles nous semblent aussi bien concernées par les questions formelles que fonctionnelles.

Ainsi, les écoles britanniques et américaines établissent les bases descriptives et théoriques de l'enseignement et de la recherche sur l'intonation de la langue anglaise. Très présent au milieu du XXe siècle, le clivage tend à s'estomper dans le domaine de la recherche au bénéfice d'une diversification des approches. En effet, le siècle dernier voit la popularisation d'outils techniques donnant accès automatiquement à l'image du signal non stylisé. Grâce à cette image brute, non encore interprétée par l'homme, il est désormais possible de prêter attention aux différents paramètres intonatifs (et non plus quasiment exclusivement à la seule mélodie) individuellement et séparément. Nous allons à présent nous concentrer sur ces deux aspects du domaine de recherche que constitue l'intonation : d'une part la contribution des modèles théoriques proposés pour la description des phénomènes intonatifs, et d'autre part l'apport du recours à la machine, qui fait de ce domaine un domaine expérimental.

1.1.2 Un domaine théorique et descriptif

Avant d'évoquer le type d'approche adopté dans ce travail, il nous paraît utile de revenir sur les différents modes d'analyse et de représentation phonologique issus de la tradition linguistique anglo-saxonne. La description des faits prosodiques et l'attention portée sur certains au détriment d'autres est en effet en partie fonction de la théorie

[31] JONES, Daniel (1909, 1956). *op. cit.*

adoptée. Nous reprenons ici le travail de A. Crompton qui, dans un article de 1980, établit deux critères d'évaluation d'une grammaire de l'intonation, qu'il applique aux modèles existants.[32]

Le premier critère de A. Crompton est d'ordre phonétique. Pour être pertinente, une représentation phonétique doit contenir des indications précises concernant les paramètres physiques et perceptifs de l'intonation, c'est-à-dire en particulier la fréquence fondamentale et son corrélat, la hauteur mélodique.[33] Il distingue deux types de contours. Le *pitch contour*, qui contient des informations relatives aux trois paramètres intonatifs, est affecté par leurs variations les plus fines et présente un aspect irrégulier. Le contour intonatif (*intonation contour*) constitue la stylisation du *pitch contour* en une ligne lisse et continue.

Le second critère est de nature phonologique. Une grammaire de l'intonation doit exprimer toutes les oppositions pertinentes d'une langue et seulement celles-ci. Elle doit également être en mesure de rendre compte à la fois des tons simples et des tons complexes (qui ne sont autre chose que la combinaison de tons simples). Enfin, elle doit être capable d'énoncer des règles.

A. Crompton évalue trois modèles : celui de la théorie des contours (*tonetic stress marks*) de l'école britannique, celui de la théorie des niveaux de la tradition américaine, et celui de la théorie des configurations. Si l'évaluation des trois modèles selon les mêmes critères préalablement établis semble être un gage d'objectivité, il faut garder à l'esprit que le but de A. Crompton dans cet article est de promouvoir son propre modèle, celui de la théorie des configurations.

[32]Notons que ces modèles font la part belle au seul paramètre mélodique. Intensité et durée, parfois mentionnées comme importantes, sont la plupart du temps ignorées, ou encore simplement assimilées à la fréquence fondamentale.

[33]Bien que mentionnées, l'intensité et la durée sont écartées.

1.1.2.1 Le modèle de la théorie des contours

A. Crompton relève les insuffisances suivantes, dues à un certain nombre d'idéalisations inhérentes au mode de représentation utilisé. Pour lui, ce système manque de précision concernant à la fois la définition de la tessiture de chaque locuteur et la hauteur relative de chaque syllabe. En outre, l'échelle temporelle figure rarement sur les graphiques. Enfin et surtout, les contours intonatifs sont traités essentiellement comme des séquences de hauteurs mélodiques affectant des séquences de syllabes. A. Crompton conclut que la théorie des contours propose un modèle insuffisant d'une part phonétiquement, parce que le paramètre de hauteur n'est pas assez précis, d'autre part phonologiquement, parce qu'il ignore certaines des oppositions pertinentes du français, langue particulièrement étudiée par A. Crompton.

Ce premier modèle pose tout de même avec R. Kingdon[34] les bases d'une théorie moderne de l'intonation telle qu'on l'enseigne dans les universités britanniques et dans la grande majorité des départements d'anglais en France. En effet, contrairement à la plupart des analyses antérieures au XXe siècle ou inspirées de celles-ci, c'est la hauteur relative des tons, dans les limites d'une voix donnée, et non pas par leur hauteur intrinsèque précise, qui assure les fonctions de la mélodie dans l'acte de communication. Le développement de la théorie des contours en Grande-Bretagne coïncide avec la disparition de la notation sur portée musicale, difficilement compatible avec un traitement de la hauteur de la voix selon une échelle relative. Si A. Crompton reproche au modèle de la théorie des contours son manque de précision quant à la hauteur relative des syllabes, la version qu'en propose D. Crystal[35] tente précisément de situer chaque syllabe par rapport à la précédente en spécifiant si celle-ci est de même hauteur, un peu plus haute ou un peu plus basse, plus haute ou plus basse, ou encore beaucoup plus haute ou beaucoup plus basse. Les hauteurs absolues des différentes syllabes et

[34] KINGDON, Roger (1958). *op. cit.*
[35] CRYSTAL, David (1969). *op. cit.*

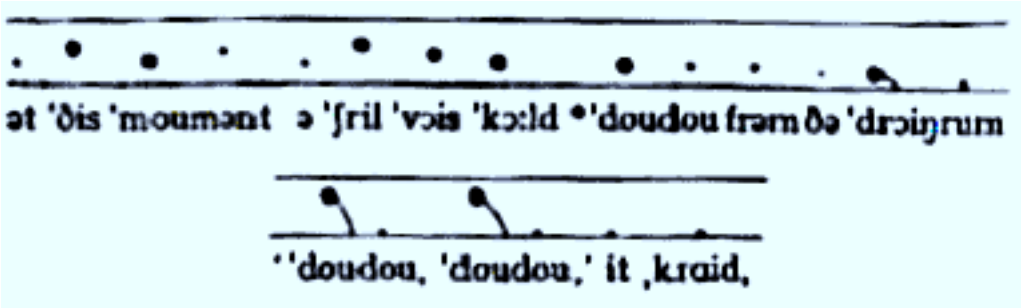

FIG. 1.2 – Représentation de l'intonation par D. Jones

de la tessiture des locuteurs n'entrent pas en considération.

La théorie des contours propose autant de versions de son modèle que de chercheurs la pratiquant, son évolution est donc constante. La Figure 1.2 donne un exemple de représentation graphique de l'intonation de l'anglais présenté par D. Jones dans son ouvrage de 1909.[36]

1.1.2.2 Le modèle de la théorie des niveaux

Ce type de théorisation et de description de l'intonation subit très tôt la critique d'un certain nombre de linguistes, parmi lesquels D. Bolinger.[37] Selon le critère phonétique de A. Crompton, la représentation des phénomènes intonatifs par la théorie des niveaux est insuffisante car elle équivaut à une représentation de la mélodie en escalier. Il reproche à G.L. Trager & H.L. Smith[38] non seulement de considérer la mélodie comme s'incarnant dans un nombre de points dans l'énoncé et non pas comme un continuum, mais encore d'avoir arbitrairement établi ce nombre de points de hauteurs mélodiques perçues en anglais à seize (ce qui n'a selon lui aucun fondement). Cependant, A. Crompton ne récuse pas totalement le mode de représentation de la théorie des niveaux, qui offre une représentation relativement fidèle de schémas mélodiques

[36]JONES, Daniel (1909, 1956). *The Pronunciation of English.* Cambridge : CUP. (p.190).

[37]A. Crompton fait notamment référence à un article de 1951 : BOLINGER, Dwight (1951). Intonation : levels versus configurations. *Word* 7 : 199-210.

[38]TRAGER, George Leonard & SMITH, Henry Lee (1951). *op. cit.*

Every man on the force is working on it sir
°3--2 °3--2- °3- -2- °4- -1-2/

FIG. 1.3 – Représentation de l'intonation par K. Pike

observés. K. Pike[39] considère que ce n'est pas tant le niveau de toutes les syllabes qui importe que celui atteint aux points de contour (*contour points*). Il s'agit en définitive de prendre en compte les contours intonatifs et de les intégrer à une représentation par niveaux. Il définit la tessiture de chaque locuteur, tessiture dans laquelle il établit de façon absolue les différents niveaux. Si A. Crompton admet que la version de la théorie développée par K. Pike est assez satisfaisante d'un point de vue phonétique, il déplore que, par manque de clarté et de précision, les problèmes relatifs à la hauteur des niveaux ne soient pas résolus.

Du point de vue phonologique, A. Crompton regrette la restriction la plus fréquente des niveaux mélodiques à quatre (telle qu'adoptée par K. Pike). Si elle est souhaitable méthodologiquement, elle se révèle erronée en pratique. En outre, cette conception statique de l'intonation comme une série de niveaux atteints ou à atteindre lui semble davantage convenir à une description du chant que de la parole. Néanmoins, de façon intéressante et contrairement à la théorie de l'école britannique, la théorie des niveaux postule implicitement l'indépendance des phénomènes mélodiques et de l'accentuation lexicale.

1.1.2.3 Le modèle de la théorie des configurations

Ce modèle, développé par D. Bolinger,[40] et par A. Crompton lui-même,[41] est né d'une remise en cause des modèles issus des théories que l'on vient d'évoquer. Il

[39] PIKE, Kenneth, L. (1945). *op. cit.*
[40] BOLINGER, Dwight (1951). *art. cit.*
[41] CROMPTON, Andrew (1980). Intonation : Tonetic stress marks versus levels versus configurations. *Word* 31 : 151-198.

```
[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]
 |   +R    | -R | -R  | -R  | +R |
 |   -F    | +F | -F  | +F  | -F |
```

FIG. 1.4 – Représentation de l'intonation par A. Crompton

considère l'intonation comme un phénomène suprasegmental, ou plus précisément, d'après A. Crompton, suprasyllabique. Il est caractérisé par l'importance donnée aux configurations, c'est-à-dire à la direction des variations mélodiques. Il s'agit de déterminer la structure suprasyllabique des énoncés en sortes de groupes intonatifs afin de préciser la direction de la variation mélodique qui les affecte. On retrouve ici des termes appartenant à la théorie des contours. Chez A. Crompton, une montée (*rise*) se représente par les symboles suivants : + R, – F. Une chute (*fall*) est symbolisée comme suit : – R, + F. Un état statique (*static*) est noté – R, – F. D'après A. Crompton, ce type de représentation permet de considérer de façon séparée le segmental et la prosodie, qui peuvent varier indépendamment l'un de l'autre.

Dans l'exemple typique présenté par la Figure 1.4, les crochets représentent les différentes syllabes, les traits verticaux indiquent l'alignement avec celles-ci et déterminent des groupes suprasyllabiques dont la direction des variations mélodiques est spécifiée par les lettres.[42] D'après A. Crompton, ce type de représentation permet de considérer de façon séparée le segmental et la prosodie, qui peuvent varier indépendamment l'un de l'autre. S'il est difficile (mais possible) de rendre compte de tons complexes au sein de la même syllabe, les notations de direction admettent des degrés : un chiffre de 1 à 4 par exemple peut se substituer à + ou – pour indiquer la pente relative du mouvement mélodique.

A. Crompton émet cependant un certain nombre de réserves quant à la pertinence d'un tel mode de représentation, qui implique lui aussi une stylisation des phénomènes

[42] *ibid.* (p.185).

rapportés. Il suggère en effet que l'éventail de degrés possibles dans les montées et les chutes est susceptible d'être plus vaste au niveau phonétique qu'au niveau phonologique. S'il admet que ce modèle est incapable de spécifier les oppositions pertinentes d'une langue en ce qui concerne la hauteur de syllabes individuelles, il considère que ce modèle est apte à faire apparaître toutes celles concernant la direction des variations mélodiques.

Pour sa part, D. Bolinger[43] utilise un mode de représentation très évocateur pour l'œil qui consiste à disposer le texte (orthographique ou phonétique) sur la page de façon à reproduire intuitivement la hauteur des différentes syllabes les unes par rapport aux autres. Le nombre de niveaux utilisés peut être infini. La mise en page devient néanmoins très complexe.

1.1.2.4 Le modèle de la théorie autosegmentale-métrique

Avec l'apparition de cette théorie à la fin des années 1970 s'opère un retour à une approche de la représentation des phénomènes de l'intonation par des niveaux. Mais cette démarche s'effectue dans un souci de rendre compte des phénomènes mélodiques au niveau phonétique et au niveau phonologique. Elle entend résoudre le conflit opposant les tenants de la théorie des niveaux et ceux de la théorie des configurations. Les représentants de cette théorie appliquée à la langue anglaise sont principalement J. Pierrehumbert,[44] C. Gussenhoven[45] et D.R. Ladd.[46]

La théorie autosegmentale-métrique conçoit plusieurs « tons ordinaires » déterminés par la hauteur du son. C. Gussenhoven se sert des contours intonatifs de l'école britannique (trois tons : le *rise* : L*H, le *fall* : H*L, et le *fall rise* : H*LH). D.R. Ladd

[43]BOLINGER, Dwight (1951). *art. cit.*

[44]PIERREHUMBERT, Janet B. (1980). The Phonology and Phonetics of English Intonation. Thèse de doctorat, MIT, Département de Linguistique.

[45]Voir par exemple GUSSENHOVEN, Carlos (1984). *On the grammar and semantics of sentence accents.* Dordrecht : Foris.

[46]Voir notamment LADD, D. Robert (1996). *Intonational Phonology.* Cambridge : Cambridge University Press.

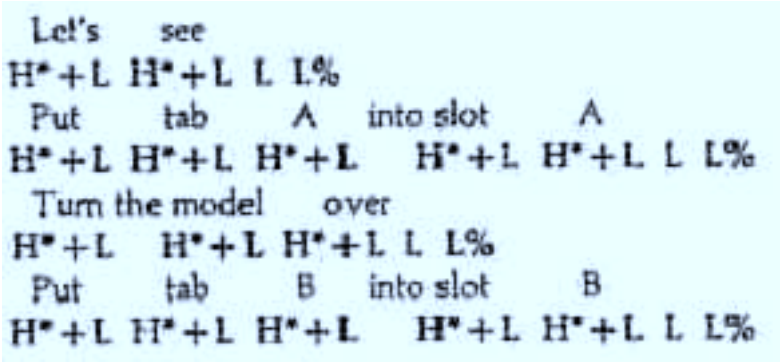

FIG. 1.5 – Représentation de l'intonation par J. Pierrehumbert & J. Hirschberg

reprend les quatre niveaux de l'école américaine (H : ton très haut, L : ton très bas, HL : ton haut, LH : ton bas). J. Pierrehumbert réduit quant à elle le nombre de niveaux à deux (H : ton haut, L : ton bas). La notion de *downstep*, centrale chez J. Pierrehumbert, permet de concevoir la succession des tons comme des séquences linéaires. Si la mélodie de la phrase décline naturellement, seuls deux niveaux suffisent pour la représenter. La hauteur de chaque ton se lit relativement à celle du ton précédent : la fonction d'un ton L, qui n'a aucune manifestation directe dans le contour de F_0, est d'indiquer que la valeur de F_0 du ton H suivant est inférieure à celle du ton H précédent. A ces tons ordinaires s'ajoutent des « tons de frontière » signalés par le signe %.

Ce mode de description de l'intonation est illustré par la Figure 1.5.[47] Il s'insère dans un système plus complexe de représentation des phénomènes mélodiques nommé ToBI (*Tones and Break Indices*) :

> A ToBI transcription for an utterance consists minimally of a recording of the speech, an associated electronic or paper record of the fundamental frequency contour, and (the transcription proper) symbolic labels for events arranged in four parallel tiers. [...]:

[47]PIERREHUMBERT, Janet & HIRSCHBERG, Julia (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. in Cohen, Philip R., Morgan, Jerry, *et al.* (éds.), *Inten-*

(1) a tone tier (2) an orthographic tier
(3) a break-index tier (4) a miscellaneous tier

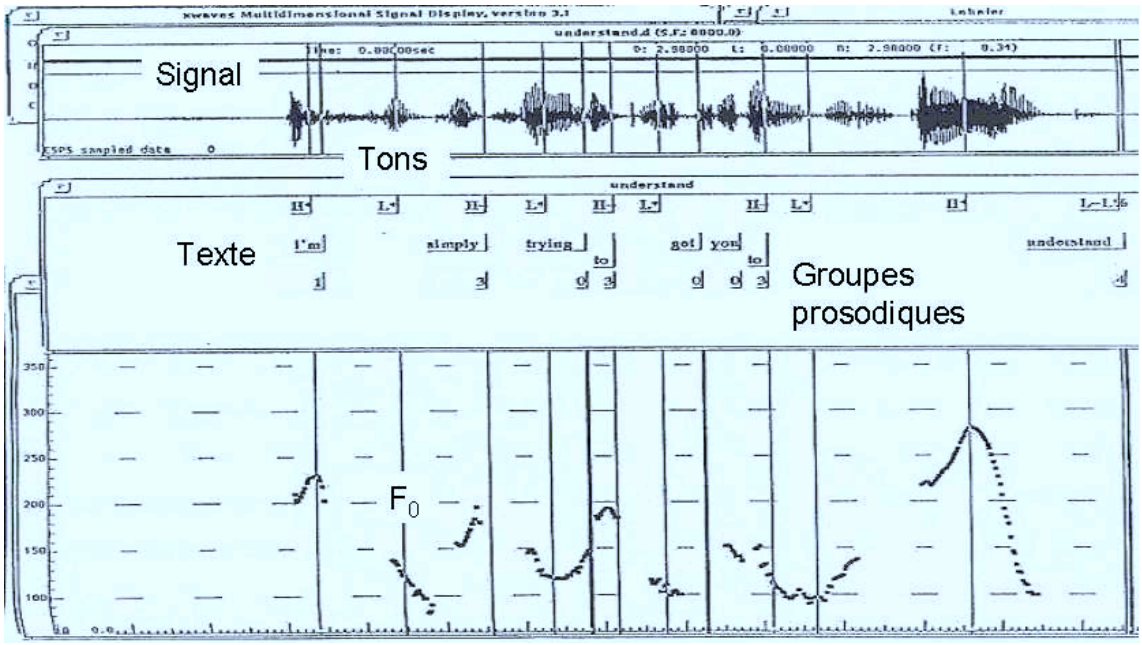

FIG. 1.6 – Représentation de l'intonation de l'anglais selon ToBI

Ce système de représentation se sert de quatre lignes horizontales différentes. La ligne supérieure (1), proche de l'analyse phonologique, est celle où figurent les tons. Se succèdent des tons hauts et des tons bas (H et L), accompagnés de signes diacritiques indiquant les limites des diverses unités et l'importance des tons au sein de ces unités : *, –, et %. Les lignes réservées à la transcription orthographique (2) et aux remarques diverses (4) contiennent des éléments, qui, bien que ne relevant pas de la prosodie, sont utiles à son interprétation. Elles permettent de noter des événements comme le rire ou la toux des locuteurs enregistrés qui viennent perturber le déroulement des énoncés et leur courbe mélodique. La troisième ligne (*the break-index tier*) sert à marquer les groupes prosodiques formés par les mots d'un énoncé. Il s'agit de mesurer subjectivement le degré d'association d'un mot avec le suivant. Ce degré est figuré

tions in Communication : 271-311. Cambridge, MA : MIT Press.

sur la ligne immédiatement inférieure à celle de la transcription orthographique à la hauteur de la fin de chaque mot par un chiffre sur une échelle de 0 (pour les mots les plus conjoints) à 4 (pour les plus disjoints). La Figure 1.6 présente l'analyse proposée par M. Beckman & G. Ayers d'un énoncé de leur corpus : *I'm simply trying to get you to understand.*[48]

Cette théorie propose donc un mode de représentation visuelle complexe et détaillée. Elle se fonde sur des données obtenues par des mesures instrumentales vérifiées (si l'on ne tient pas compte des éléments figurant sur la ligne appelée *break-index tier*). Elle a recours à la mesure et au calcul de la hauteur mélodique par rapport au temps. Si ce mode de description constitue pour nous une avancée considérable par rapport aux stylisations des modèles présentés plus haut, on remarque cependant qu'elle privilégie la mélodie aux dépens des autres paramètres suprasegmentaux.

1.1.2.5 Quelques remarques concernant le modèle de M.-A. Morel et L. Danon-Boileau

Si l'on a longuement évoqué ces problèmes de représentation des phénomènes intonatifs, c'est qu'ils ont non seulement beaucoup occupé les linguistes, mais qu'ils témoignent de points de vue différents sur ce qui fait l'intonation d'une langue. Dans la mesure où la représentation est une stylisation, elle constitue déjà une interprétation. Chaque école a ses détracteurs, mais chaque théorie est en constante évolution et propose des apports intéressants qu'il faut savoir mettre à profit. L'idée de P. Delattre[49] de combiner niveaux et contours est séduisante, c'est d'ailleurs une démarche relativement proche de celle adoptée à l'Université Paris III pour le français par M.-A. Morel

[48]BECKMAN, Mary & AYERS Gayle (1994). *Guidelines for ToBI Labelling*, version 2.0. Linguistics Department, Ohio State University.

[49]DELATTRE, Pierre (1972). The Distinctive Function of Intonation. in Bolinger, Dwight (éd.), *Intonation, selected reading* : 159-174. Harmondsworth : Penguin Books.

et L. Danon-Boileau.[50] Celle-ci fait l'objet d'une description détaillée au chapitre suivant.[51] Il nous semble utile néanmoins de donner ici quelques précisions concernant le mode de représentation adoptée dans leur *Grammaire de l'intonation.*

Pour la représentation de la mélodie d'énoncés cités à titre d'exemple dans le corps d'un texte, M.-A. Morel & L. Danon-Boileau adoptent un mode de représentation qui ressemble à celui de D. Bolinger. Il permet de citer un extrait de corpus et de montrer certains rapports de hauteur mélodique. On voit pourtant dans l'exemple suivant qu'il est difficile de les faire apparaître de façon précise sans recours aux chiffres :

$^{\text{just}}{}_{\text{ement}}$ $^{\text{j'voulais savoir avec c't(te) histoire de}}$ $_{\text{math}}$ $^{\text{est-ce qu'on fait}}$ $^{\text{d'l'informa}}{}_{\text{tique}}$[52]

Si n'apparaissent ici que trois niveaux, ce mode de représentation est construit autour de quatre niveaux. Un tel découpage ne correspond à aucune réalité acoustique : la mélodie consiste en un continuum de fréquences. Cependant, comme nous le verrons plus loin,[53] l'oreille n'est pas sensible à une infinité de niveaux sonores relatifs. De la sorte, une réduction des variations pertinentes au sein du continuum à un nombre restreint n'est pas aberrant. On retiendra ici les réserves d'A.C. Simon, qui regrette que ce modèle s'attache à l'établissement de niveaux fixes, ne tenant pas compte des variations locales du registre utilisé par les locuteurs.[54] Après avoir défini la tessiture du locuteur, il s'agit d'étudier les rapports des niveaux entre eux de façon dynamique, la direction des mouvements et la pente de ces mouvements. Une fois établi un repère moyen, on situera chaque point en fonction de ce repère. L'analyse fonctionnelle consistera ensuite à justifier la hauteur de tous les points par rapport au repère et les uns par rapport aux autres. Il est tentant d'établir ce repère une fois pour

[50]Voir par exemple MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation, L'exemple du français.* Gap : Ophrys.

[51]Voir Chapitre 2, p.108.

[52]*ibid.* (p.128).

[53]Voir Section 1.1.3, p.71.

[54]SIMON, Anne Catherine (à paraître). *Segmentation et structuration prosodiques du discours : Une approche multidimensionnelle et expérientielle.* v.2.0 2002 destinée à la publication. Thèse de Doctorat. Louvain-la-Neuve : Université Catholique de Louvain. (p.28).

toutes, ce que fustige précisément A.C. Simon. Mais il est bien entendu souhaitable qu'il soit constamment redéfini, d'où l'importance d'une prise en compte des positions relatives des différents points. Bien qu'imparfait, ce découpage arbitraire nous semble utile sur le plan méthodologique et constitue un préalable à l'analyse fonctionnelle du discours : c'est plus un outil qu'une fin en soi. Appliqué de façon intéressante pour le français, il s'agira pour nous de déterminer dans quelle mesure il est pertinent pour l'anglais.

La réduction de la représentation des phénomènes intonatifs à la stylisation du seul paramètre de la mélodie nous semble problématique. C'est pourtant ce que pratiquent les théories évoquées plus haut. Une étude de l'intonation d'une langue telle que nous la concevons doit tenir compte de ses paramètres les plus pertinents, c'est-à-dire la mélodie, l'intensité et la durée. En vue d'un couplage de ces paramètres, il est indispensable de les faire figurer dans les représentations. M.-A. Morel & L. Danon-Boileau proposent le codage suivant :

> Syllabe en exposant : montée de F_0/ Syllabe en indice : chute de F_0
>
> Chiffres donnés sous les syllabes (ou à la suite de la syllabe finale) : ils correspondent au niveau de variation de Fo, déterminé en fonction de la plage de variation du locuteur :
>
> > H4 : niveau le plus haut / H1 : niveau le plus bas
> >
> > H2 et H3 : niveaux intermédiaires / H2,5 : entre H2 et H3
> >
> > H2+ : niveau légèrement supérieur à H2
> >
> > H2– : niveau légèrement inférieur à H2
> >
> > Plage haute : de H2,5 à H4 / Plage basse : de H2,5 à H1
>
> Intensité : les signes + / – / = ou I+ / I– / I= indiquent les variations d'intensité évaluées de façon empirique à partir des données fournies par les tracés mélodiques.[55]

[55]MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *op. cit.* (p.5).

Dans l'exemple suivant, {xx} indique la durée de la pause en centisecondes, et l'allongement est noté par (::). Ce type de notation permet de coupler les indices ou de les décorréler et de focaliser l'attention sur un seul paramètre :

> moi je trouve (H3 I+) e j/je trouve e: (H2 I=) fantastique (H3 I+) de bon (H1) d'en/d'entendre (H2– I=) un musicien (H2 I=) qui qui joue très vite (H3– I+) et:: {50} mais (H2+ I+) je je crois qu'ça (H3 I=) c'est /c'est l'travail (H2 I–) à la maison::: (H3 I–)[56]

Ce mode de représentation nous semble être le plus complet de ceux que nous avons évoqués. Cependant, ce n'est pas celui que nous retenons dans le cadre de ce travail. Si nous conservons le même mode de notation pour les phénomènes de durée (pauses et allongements), nous ne reportons généralement pas la hauteur mélodique ou l'intensité atteintes sur un segment dans la transcription. Il est parfois nécessaire d'y faire référence, ce que nous faisons, pour la mélodie, à l'aide des simples symboles F_0+, F_0= et F_0–, en fonction du contexte local (l'énoncé) mais également du contexte élargi (l'intégralité de l'échange). Nous procédons de la même façon pour l'intensité. On se reportera le plus souvent possible aux tracés prosodiques de WinPitch qui nous semblent constituer la plus fidèle représentation des phénomènes intonatifs. Le modèle de M.-A. Morel & L. Danon-Boileau reste le cadre théorique principal dans lequel s'inscrit notre recherche.

1.1.3 Un domaine expérimental

1.1.3.1 Apport de l'acoustique et de l'informatique

Les études sur l'intonation se sont développées de façon de plus en plus scientifique à la faveur de la création de passerelles entre la linguistique et un domaine relativement éloigné : la physique. La physique décrit les paramètres acoustiques de

[56] *ibid.* (p.73).

la parole, auxquels la linguistique tente d'attribuer un sens au sein d'un système de significations. L'informatique permet au linguiste de procéder à l'acquisition du signal acoustique, à sa numérisation et à son stockage avant le traitement proprement dit et l'établissement de la représentation graphique qui en découle. Il est cependant nécessaire, pour pouvoir en donner une interprétation linguistique, de se familiariser avec la nature de ces phénomènes physiques.

Le son est l'effet d'un phénomène de vibration qui se traduit par une variation locale de la pression atmosphérique. Se propagent ainsi des ondes qui, portées par l'air, franchissent l'espace qui sépare la source (corde d'un violoncelle, corde vocale) du récepteur (micro, oreille). Les paramètres suivants permettent d'en donner une première description, et constituent les outils du linguiste pour l'analyse suprasegmentale.[57]

Le spectre fréquentiel

F_0, sons voisés, sons non voisés La fréquence fondamentale (F_0) d'un son correspond au nombre de vibrations par seconde de l'onde émise par la source sonore. Elle représente la fréquence dominante de ce son. Exprimée en Hertz (Hz), elle se traduit, d'un point de vue perceptif, par la sensation de hauteur mélodique. En termes de production, phonéticiens et acousticiens s'accordent sur les affirmations suivantes :

> La fréquence de vibration dépend de la pression de l'air dans les poumons et des propriétés mécaniques des cordes qui sont commandées par de nombreux muscles laryngés. En général, plus la pression de l'air dans les poumons est forte, plus les cordes vocales sont amincies et tendues et plus haute est la fréquence de vibration de ces cordes, qui produisent les pulsations de l'air. Le train de pulsations engendre des variations rapides de pression dans le conduit vocal,

[57] Nous nous intéresserons principalement aux sons produits par la parole.

ou, en d'autres termes, un son. La hauteur du son correspond à une certaine fréquence de vibration. [58]

Les sons de la parole se séparent théoriquement en deux catégories, selon leur mode de production. Les sons voisés sont produits par la vibration des cordes vocales, tandis que les sons non voisés ne font pas intervenir celles-ci. En conséquence, il est impossible de déterminer une fréquence fondamentale pour les sons non voisés. Seuls les sons voisés donnent une impression de hauteur (et apparaissent sur les tracés mélodiques des logiciels d'analyse prosodique, d'où l'aspect discontinu du tracé de la courbe).

La hauteur des voix masculines peut varier entre 80 et 200 Hz. Les femmes utilisent un ambitus plus large, en moyenne entre 150 à 450 Hz, en voix parlée.

Harmoniques et formants Les autres pics d'énergie des sons voisés pour une fréquence donnée sont les harmoniques et les formants. Produits en même temps que la fréquence fondamentale, ils sont de fréquence supérieure à celle-ci et en constituent des compléments, d'un point de vue physique et perceptif. Les harmoniques correspondent à des multiples entiers de F_0. Perceptivement, ils permettent de reconstituer F_0 dans le cas où cette dernière serait filtrée. Les formants et leur place par rapport à F_0 correspondent à la signature des voyelles. Les harmoniques et les formants permettent plus généralement de définir le timbre d'une voix.

L'intensité globale (I) correspond à l'énergie contenue dans le signal de parole. D'un point de vue physique, elle se manifeste comme l'amplitude des vibrations des cordes vocales dont elle est proportionnelle au carré. Elle s'exprime ici en décibels (dB). D'un point de vue perceptif, l'intensité correspond au volume sonore dégagé.

[58] MILGRAM, Maurice & ZARADER, J.L. (1998). *Traitement du signal II : analyse spectrale, modélisation et filtrage adaptatif.* Brochure Maîtrise EEA et Maîtrise MPA. Université Paris 6, Jussieu : Service des publications. (p.23).

Ainsi, chuchoter ou crier, c'est faire varier l'intensité.

La durée (D) est un paramètre dont l'appréhension précise est facilitée par le traitement informatique des données. Toute parole s'inscrivant dans le temps, il est indispensable de se préoccuper de cette dimension au sein de laquelle se joue la perception des phénomènes d'allongement, de pause, de rythme, et grâce à laquelle la notion d'ordre et de succession prend sens. L'unité de mesure de la durée est la seconde (s).

1.1.3.2 Production et perception

Considérons la réflexion suivante, qui invite à compléter notre bref propos sur la dimension acoustique de la parole par quelques remarques concernant la perception auditive humaine.

> Il existe différentes manières de définir les paramètres prosodiques, selon qu'on les considère sur le plan de la production, sur le plan acoustique, ou perceptif. [...] Lorsque l'on se situe d'un point de vue acoustique, l'étude des paramètres se base sur ce qui est mesurable. Les paramètres peuvent être alors considérés comme parfaitement indépendants. Cependant, sur le plan de la production ou de la perception, il existe une forte interaction entre les différents paramètres.[59]

Traditionnellement, les études scientifiques sur le langage s'intéressent au phénomène acoustique dans la mesure où il est le produit d'une combinaison de facteurs physiques. Cependant, toujours plus nombreuses sont les recherches dont l'objectif est la conception et l'amélioration de programmes informatiques de synthèse et de reconnaissance de la parole, et qui doivent donc s'intéresser à la fois à la production et à la perception du phénomène acoustique. Si les linguistes font dans l'ensemble

[59]BEAUGENDRE, Frédéric (1995). *Modèle de l'intonation pour la synthèse.* Ecole thématique « Fondements et perspectives en Traitement Automatique de la Parole ». CNRS, Marseille-Luminy, 17-25 juillet 1995. Version électronique (p.2). http://www.bibliotheque.refer.org/html/parole/beaugend/beaugend.htm (site consulté le 11 août 2003.)

davantage cas de la perception, les raisons en sont tout autres. En effet, l'étude de la langue est ancrée dans le réel et s'écarte parfois des canons théoriques. Il s'agit de la langue en contexte, de la langue de l'échange. Chaque locuteur fait l'expérience immédiate du discours de l'autre, et chaque locuteur a connaissance de la façon dont il parle à travers l'image que lui renvoie son interlocuteur de son propre discours. C'est d'ailleurs l'intérêt d'un corpus comme le nôtre que de permettre de saisir l'impact du discours de chacun des interlocuteurs sur l'autre.

Les notions d'acoustique présentées plus haut ne suffisent donc pas à épuiser la complexité des vibrations sonores. Si elles nous permettent de mieux saisir le phénomène physique, elles doivent être complétées par une réflexion sur l'effet qu'elles produisent sur l'être humain, c'est-à-dire une réflexion sur leur perception. L'acception même du mot « son », rappellent P. Bourcet & P. Liénard, témoigne de cette double dimension :

> Les fluctuations rapides (plusieurs dizaines de fois, jusqu'à des milliers par secondes) de la pression de l'air au niveau de nos oreilles engendrent une sensation auditive, et le mot SON désigne à la fois la **vibration physique** capable d'éveiller cette sensation et la **sensation** elle-même.[60]

Ainsi, continuent-ils, l'attention doit être portée également sur la réception de l'objet sonore produit, donc sur le sujet auditeur et/ou interlocuteur :

> Le son possède un aspect **objectif** et peut ainsi être considéré comme une **cause**, **objet** naturel des sciences et techniques. Sous son aspect **subjectif** le son est un effet étroitement dépendant du **sujet** qui le ressent.[61]

Cependant, il existe des constantes perceptives chez l'homme qui permettent de mettre en rapport objet et sujet, production et perception.

[60] BOURCET, Patrice & LIENARD, Pierre (1987, 2e éd. 1994). Acoustique fondamentale. in Mercier, Denis (éd.), *Le Livre des techniques du son* 1 : pp.13-43. Paris : Eyrolles. (p.13).
[61] *ibid.* (p.13).

Hiérarchie de perceptibilité des paramètres Une hiérarchie des indices suprasegmentaux selon leur degré de perceptibilité ressort notamment des travaux de D. Jones[62] et de G. Faure.[63]

Pour ces deux chercheurs, l'intensité est un phénomène essentiellement subjectif. Le locuteur le ressent en fonction des efforts qu'il exige de lui. Ces efforts se gravent dans sa mémoire et leur souvenir lui permet d'interpréter, en tant qu'auditeur, ceux déployés par son interlocuteur dans une situation similaire. G. Faure récuse ici toute perception intensive en l'assimilant à une erreur de perception de phénomènes non pas intensifs mais mélodiques, accompagnés de signes mimo-gestuels.

> C'est en fonction de ces signes, et aussi sous l'effet des souvenirs évoqués chez l'auditeur par le message qu'il perçoit et qui lui fait revivre ses propres efforts en tant que locuteur, dans des situations analogues, que la manifestation acoustique, essentiellement mélodique, de cette « tension » accentuelle est finalement imaginée comme une variation d'intensité.[64]

Il apparaît néanmoins que certains phénomènes intensifs sont perceptibles objectivement et en tant que tels, et qu'il est possible de leur donner une interprétation linguistique.

La durée des syllabes est perçue avec plus de netteté que leur intensité. Une oreille naïve est capable de distinguer clairement et spontanément au moins trois degrés de durée (correspondant chez D. Jones à l'accent principal, l'accent secondaire, et l'inaccentuation). Mais ici encore se retrouve l'écart entre perception et production :

> La durée perçue n'a que des rapports assez lointains avec la durée mesurée, [...] et elle est toujours, d'autre part, une sorte d'effet second, c'est-à-dire la simple conséquence de la nécessité physique dans laquelle se trouve le locuteur

[62] Voir par exemple, JONES, Daniel (1950). *The Phoneme, its nature and use.* Cambridge : Cambridge University Press.

[63] FAURE, Georges (1962). *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise (essai de description phonologique).* Paris : Didier.

[64] *ibid.* (p.140).

> de disposer d'assez de temps pour affecter la syllabe dite accentuée de l'intensité et surtout du ton et des variations de ton qui la détachent, pour l'oreille, des syllabes voisines.[65]

La relativité de la perception de durée est confirmée par l'acousticien M. Kitantou.

> La perception subjective du temps est une grandeur très difficile à évaluer. Contrairement à celle de la fréquence (oreille absolue) ou à celle du niveau, l'[évaluation] de la durée d'un phénomène est sujette à une grande variabilité. Elle dépend de nombreux facteurs psychologiques (état d'éveil, conditions d'écoute, habitude...).[66]

Cependant, nous montrerons que certains faits de durée ont une pertinence sur un corpus comme le nôtre et sont interprétables en termes linguistiques.

La hauteur mélodique (F_0) semble être le paramètre le plus perceptible. On remarque d'ailleurs que c'est le paramètre auquel les linguistes, travaillant souvent uniquement à l'oreille (sans ordinateur), ont traditionnellement accordé le plus d'importance dans leurs travaux sur l'intonation. Dans son article, A. Gribenski[67] soutient que l'oreille, peu sensible aux variations de durée et d'intensité, est d'une finesse extrême lorsqu'il s'agit de distinguer des nuances de tons, c'est-à-dire des variations mélodiques subtiles. Grâce à un seuil de perception différentielle élevé, elle peut détecter une différence de hauteur entre deux sons donnés de façon assez fine.

Il apparaît que l'oreille humaine peut plus communément différencier un petit nombre de niveaux de hauteur mélodique. Ainsi, M. Rossi & M. Chafcouloff[68] affirment qu'il est difficile de distinguer plus de six niveaux fonctionnels de hauteur différents.

[65] *ibid.* (p.143).

[66] KITANTOU, Mpaya (1987, 2e éd. 1994). La perception auditive. in MERCIER, Denis (éd.). *Le Livre des techniques du son* 1 : 155-181. Paris : Eyrolles. (p.162).

[67] GRIBENSKI, André (1957). La hauteur des sons : sa mesure, sa perception. *La Nature* 3269, septembre 1957 : 359-363.

[68] ROSSI, Mario & CHAFCOULOFF, M. (1972a). Les niveaux intonatifs. *Travaux de l'Institut de Phonétique d'Aix-en-Provence* 1 : 167-176.
ROSSI, Mario & CHAFCOULOFF, M. (1972b). Recherche sur le seuil différentiel de fréquence fondamentale de la parole. *Travaux de l'Institut de Phonétique d'Aix-en-Provence* 1 : 179-185.

Les expériences d'A. Di Cristo & M. Chafcouloff[69] établissent que distinguer entre trois niveaux (haut, moyen, et bas) est chose bien plus naturelle pour leurs informateurs. C'est une distinction entre trois niveaux que nous conserverons dans le cadre de ce travail.

Outre les différents niveaux de hauteur mélodique, il convient de prendre en compte la façon dont les mouvements mélodiques sont interprétés par l'oreille humaine. D'après les mêmes sources, il est généralement difficile de distinguer plus de quatre niveaux de pente mélodique (petite, moyenne, grande, très grande).

Indépendance et interdépendance des paramètres La question de l'indépendance des paramètres se pose d'un point de vue physiologique pour la production et d'un point de vue psychoacoustique pour la perception. Elle est d'autant plus importante que, si nous cherchons à attribuer un sens linguistique aux variations de ces paramètres, il convient de savoir dans quelle mesure leur perception et leur production sont décorrélées.

D'un point de vue physiologique P. Lieberman, dans son article de 1965 sur la perception de l'intonation qu'ont les linguistes, fait remarquer que la structure spectrale des voyelles influe sur la perception de leur intensité, de sorte qu'un locuteur ne produit pas le même effort pour donner à deux voyelles différentes la même intensité :

> Owing to the spectral structure of the vowels in question, in particular the higher first formant of [a] vs. the lower first formant of [i], the vowel [a] will involve higher-energy acoustic output that [i] for the same level of energy input from the larynx-source.[70]

[69] Cette publication de 1975 est mentionnée par A. Di Cristo mais n'est pas référencée dans la bibliographie de son ouvrage de 1985 *De la microprosodie à l'intonosyntaxe*. Aix-en-Provence : Publications de l'Université de Provence.

[70] LIEBERMAN, Philip (1965). On the acoustic basis of the perception of intonation by linguists.

D'autre part, G. Faure affirme qu'il est quasiment impossible de faire varier l'intensité d'une syllabe, qui est fonction de l'amplitude des vibrations des cordes vocales, sans modifier avec elle la hauteur de cette syllabe, qui est liée à la fréquence de ces mêmes vibrations. Cependant, cette remarque nous semble fort contestable. Certes, il s'agit sans doute d'une tendance physiologique selon laquelle une plus grande tension des cordes vocales s'accompagne d'une plus grande pression sous-glottique. Néanmoins, dissocier hauteur et intensité, loin d'être impossible, est chose commune chez tous les chanteurs.

D'un point de vue psychoacoustique Les études sont plus nombreuses qu'en physiologie de la parole et nous apportent davantage d'enseignements. Ainsi, certains chercheurs établissent qu'il est difficile, perceptivement, d'isoler la mélodie des variations d'intensité et de durée qui lui sont associées.

La sensation de hauteur varie principalement avec la fréquence. Mais elle dépend également partiellement d'autres phénomènes acoustiques, ce que rappelle M. Kitantou. Ainsi, la sensation de hauteur est également sensible aux variations d'intensité.

> Quand l'intensité d'un son pur croît, la sensation de hauteur varie : elle croît pour les fréquences élevées et décroît pour les fréquences basses.[71]

Cela dit, cette différence, inférieure à 8% de la fréquence, n'est établie que pour les sons purs — ce que le discours n'est pas.

La durée d'un son influe également sur la perception de sa hauteur, comme le rappelle P. Fraisse dans son ouvrage de 1974 sur le rythme.

> Le facteur temps joue un rôle primordial dans sa relation avec les autres variables : fréquence, niveau, espace qu'il affecte de façon significative.[72]

Word 2 : 40-54.

[71]KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.* (p.160).

[72]FRAISSE, Paul (1974). *Psychologie du rythme.* Paris : PUF. (p.162).

Du fait de la constante de temps de l'oreille, un son de moins d'une seconde sera perçu comme plus grave qu'il ne l'est en réalité. A. Di Cristo[73] rend également compte de ce phénomène et affirme que les tons sont perçus moins hauts lorsque leur durée est inférieure à 0,025 s. Il rappelle également les recherches de C. Witting,[74] confirmées par M. Rossi,[75] établissant que des variations mélodiques sur moins de 0,05 s sont perçues comme des tons statiques. Inversement, du point de vue de la perception de la durée en fonction de F_0, I. Lehiste[76] montre en 1976 que les voyelles isolées recevant un ton circonflexe ou creusé sont perçues comme légèrement plus longues que celles affectées d'un ton statique, montant ou descendant.

Enfin, sur des segments très brefs, la perception de hauteur peut être influencée par le timbre des sons prononcés.

> La présence d'un formant dans un spectre peut modifier la sensation de hauteur d'une note. [77]

Certains traits relatifs à la sensation d'intensité dépendent également d'autres paramètres. Il existe en effet une corrélation entre perception du volume sonore et de la fréquence fondamentale. La pression, donc l'intensité produite, doit être supérieure pour percevoir un son de 50 Hz à celle nécessaire à un son de 100 Hz au même volume. Ainsi, plus la fréquence fondamentale est élevée, plus les différences de volume sonore sont finement perçues.

---

[73]DI CRISTO, Albert (1985). *De la microprosodie à l'intonosyntaxe.* Aix-en-Provence : Publications de l'Université de Provence. (p.551).

[74]WITTING, C. (1962). A method of evaluating listeners' transcriptions of intonation on the basis of instrumental data. *Language and Speech* 15 (3) : 138-150.

[75]ROSSI, Mario (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 24 : 129-161.
ROSSI, Mario (1978). The perception of non repetitive intensity glides on vowels. *Journal of Phonetics* 6 : 9-18.

[76]LEHISTE, Ilse (1976). Influence of Fundamental Frequency Pattern on the Perception of Duration. *Journal of Phonetics*, n°4 : 113-117.

[77]KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.*(p.161).

Discussion L'interdépendance des paramètres intonatifs, qui ressort des études mentionnées plus haut, nous semble devoir être relativisée dans le cadre qui nous intéresse. Nous conserverons la notion d'effort du locuteur, d'un point de vue physiologique, pour produire un son très aigu (F_0 élevée) ou très fort (intensité élevée), que nous relierons, d'un point de vue interprétatif, à la perception d'un phénomène marqué.

Il est en revanche possible de mettre en question l'incidence des phénomènes relevés par les études psychoacoustiques sur le fonctionnement des paramètres suprasegmentaux au sein d'interactions verbales. Il nous semble important de revenir sur deux points en particulier. Le premier a trait au niveau d'abstraction caractérisant le corpus adopté par les psychoacousticiens. La plupart des études mentionnées sont issues d'un travail sur des sons purs. Or la parole humaine n'est pas faite de sons purs, et l'on peut se demander dans quelle mesure les résultats obtenus sur ceux-ci lui sont applicables. En outre, comme le souligne A. Di Cristo,[78] il convient de relativiser la pertinence de résultats obtenus en laboratoire, dans des conditions expérimentales très éloignées de celles de la perception de la parole spontanée en milieu naturel. Il n'en demeure pas moins que certaines études ont été effectuées à partir d'un corpus de parole naturelle, comme celles de M. Rossi.[79]

Le deuxième point qui nous pousse à relativiser l'interdépendance des paramètres intonatifs au sein des interactions verbales concerne le niveau d'analyse auquel les expériences psychoacoustiques ont été effectuées. Les durées mentionnées (0,025 s, 0,050 s) témoignent d'une microanalyse des phénomènes, en-deçà des unités pertinentes à notre étude de la parole spontanée. Les études de A.W.F. Huggins,[80] puis celles de F. Goldman Eisler,[81] ont montré que les 150 à 200 mots prononcés par minute

[78]DI CRISTO, Albert (1985). *op. cit.* (p.541).
[79]Notamment ROSSI, Mario (1971). *art. cit.*
[80]HUGGINS, A.W.F (1964). Distortions of the temporal pattern of speech. *Journal of the Acoustical Society of America* 36 : 1055-1064.
[81]GOLDMAN-EISLER, Frieda (1968). *Psycholinguistics : Experiments in Spontaneous Speech.*

au cours d'une conversation en anglais correspondent à 4 à 7 syllabes par seconde. Ainsi, on peut calculer que la durée d'une syllabe oscille en moyenne entre 0,15 s et 0,25 s, c'est-à-dire une durée 3 à 10 fois supérieure au seuil en-dessous duquel s'applique le principe d'interdépendance des paramètres. En outre, la perception du rythme, et donc des relations entre éléments dans le temps concerne des durées bien supérieures au seuil fixé par A. Di Cristo[82] et M. Rossi.[83] P. Fraisse[84] établit en effet deux types de durée pour la rythmisation spontanée. Les temps courts, de 0,15 à 0,30 s permettent de percevoir une collection d'éléments. Les temps longs, qui sont ceux de la perception spontanée de la durée, sont caractérisés par une durée supérieure à 0,40 s et peuvent s'étendre jusqu'à 1 s.

Nous faisons donc l'hypothèse de la décorrélation des paramètres, au niveau d'analyse où nous nous plaçons. Nous étudierons donc les variations de chaque paramètre afin d'évaluer sa participation à un effet de sens ou à un effet énonciatif. Cependant, il ne s'agit pas d'ignorer la nature « pluriparamétrique[85] » de l'intonation d'une langue et d'éclater son organisation. Un paramètre n'apparaît jamais seul (il n'y a pas de F_0 sans durée, par exemple), d'où l'importance de la prise en compte du contexte intonatif de chaque phénomène propre à un seul paramètre. La valeur spécifique donnée à une montée de F_0 ne sera pas la même si elle s'accompagne d'une baisse ou d'une hausse de I. Le croisement des paramètres apparaît donc non pas comme une nécessité psychoacoustique mais comme une nécessité théorique.

Échelles et unités : la loi de Weber-Fechner Les analyses présentées dans cette thèse se fondent sur l'examen des courbes et des tableaux de valeurs fournies indépendamment pour chaque paramètre. Afin de donner une interprétation linguis-

New-York : Academic Press.

[82]DI CRISTO, Albert (1985). *op. cit.*

[83]ROSSI, Mario (1971). *art. cit.* et ROSSI, Mario (1978). *art. cit.*

[84]FRAISSE, Paul (1974). *Psychologie du rythme.* Paris : PUF.

[85]DI CRISTO, Albert (1985). *op. cit.* (p.551), rapportant les conclusions de ROSSI, Mario (1977). L'intonation et la troisième articulation. *Bulletin de la Société de Linguistique de Paris* LXII : 55-68.

tique valable de leurs variations, il nous faut encore pouvoir lire ces données et comprendre ce qu'elles représentent.

Les limites de la perception sonore chez l'homme, en terme de hauteur mélodique, s'étendent de 20 Hz à 15-20 kHz. En-deçà et au-delà de ces valeurs (qui définissent la bande passante de l'oreille humaine) se situent respectivement les infra-sons et les ultra-sons, inaudibles à l'homme. La fréquence fondamentale de la voix humaine s'étend dans une bande beaucoup plus restreinte : 80 à 500 Hz environ. L'oreille est beaucoup plus sensible aux phénomènes intensifs. Dès quelques décibels, un son est détecté. Le seuil de douleur s'élève à 130 dB.

Pour les valeurs de F_0 et de I qui nous intéressent pour l'étude de la parole, la perception répond à la loi de Weber-Fechner, dont M. Kitantou rappelle ici la définition :

> La loi de Fechner s'applique à l'ensemble des organes des sens et affirme que l'augmentation relative de la sensation est proportionnelle au logarithme de la grandeur excitatrice.[86]

De la sorte, pour qu'un son de forte intensité soit perçu encore plus fort, il doit subir une augmentation plus importante que si son intensité de départ était faible. En d'autres termes, pour obtenir une augmentation arithmétique de l'intensité d'un phénomène perceptif, il faut faire subir une augmentation géométrique à son stimulus.

Cette non-linéarité de la perception humaine est reflétée, pour la mesure de l'intensité, par l'échelle des décibels. Son fonctionnement logarithmique est particulièrement bien adapté au phénomène perceptif pour lequel elle fournit une mesure. M. Kitantou montre ici que le son d'un violon et celui de dix violons, 10 fois plus fort, seront marqués par une différence de 10 dB :

> La variation de sensation serait, d'après Fechner, la même lorsque l'on passe d'un violon à dix violons ou de 10 à 100 violons ou encore de 100 à 1000 violons

[86] KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.* (p.156).

> puisque le rapport est toujours de 10. L'estimation de ce gain peut alors être exprimée sous forme d'un rapport logarithmique :
> soit S = k log ΔI/I. Il vient pour notre exemple : ΔI/I = 100/1000 = 10/100 = 1/10
> et S = k log 1/10 exprimé en Bels. Nous choisirons k égal à 10 afin d'exprimer S en décibels, soit S = 10 log 1/10 = 10 ∗ (– 1) = – 10 dB
> Nous aurons augmenté de 10 dB la sensation en passant de 1 à 10 violons jouant tous à la même intensité.[87]

En revanche, de même que l'échelle temporelle, l'échelle des Hertz est linéaire. Mais la perception de la hauteur mélodique et de la durée n'en reste pas moins affectée par la loi de Weber-Fechner. L'échelle musicale au contraire reflète bien notre perception logarithmique des hauteurs mélodiques. En effet, si l'intervalle de la quinte correspond toujours à trois tons et demi, sa valeur en Hz varie en fonction de la hauteur. Si l'on considère les cordes d'un violon (de la plus aiguë à la plus grave) accordées selon un tempérament égal au la = 415 Hz, l'intervalle sera d'environ 206 Hz entre mi et la, 138 Hz entre la et ré, et de 92 Hz entre ré et sol. Les précautions inverses de celles prises plus haut doivent donc être les nôtres pour l'interprétation des variations des valeurs de F_0. Comme nous le voyons sur la courbe de la Figure 1.7, il nous faudra tenir compte du fait qu'une variation de x Hz située en haut du registre d'un locuteur sera perçue comme moins importante qu'une variation dans le registre bas. Pour influente qu'elle soit et qu'elle ait pu être, cette loi du XIXe siècle est remise en question par certains chercheurs pour son manque de flexibilité. Ceux-ci lui en substituent d'autres plus complexes qui retiennent cependant le principe de la perception logarithmique de l'oreille humaine.[88]

Dans le cadre de notre étude, toutes les précautions envisagées sont à relativiser

[87] KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.* (p.157).

[88] C'est notamment le cas des chercheurs de l'Uppsala Universitet, en Suède, à qui nous avons emprunté la Figure 1.7. Voir http://www.neuro.uu.se/fysiologi/gu/nbb/lectures/WebFech.html (site consulté le 11 août 2003).

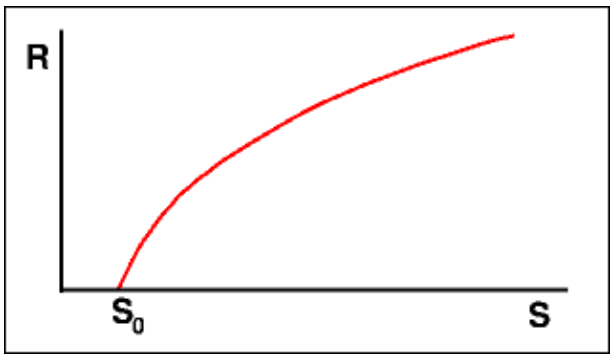

FIG. 1.7 – Courbe de la perception humaine d'après la loi de Weber-Fechner

du fait de la variation en fonction de la fréquence du seuil de perception différentielle des fréquences et de glissando.

Les seuils de perception différentielle On appelle seuil de perception différentielle le seuil à partir duquel l'oreille détecte un changement de niveau (de durée, d'intensité ou de hauteur mélodique). Ainsi, d'un point de vue perceptif, une variation (ou une comparaison) est bien plus pertinente que la détermination de valeurs absolues. Pour prendre un exemple ressortissant au domaine de la perception visuelle, il est plus facile de situer deux nuances de gris sur une échelle de gris allant de blanc à noir qu'une seule.

> Dans la vie quotidienne, il est plus utile de détecter une information qualitative ou relative qu'une information quantitative ou absolue. L'oreille, à l'instar des autres organes des sens, est plus performante pour détecter une différence entre deux grandeurs (analyse différentielle) que pour en évaluer les niveaux absolus (mesure). [89]

Quel est ce seuil pour les paramètres qui nous intéressent ? A. Di Cristo[90] rapporte que les différentes recherches sur le **seuil différentiel pour la perception des**

[89]KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.* (p.156).
[90]DI CRISTO, Albert (1985). *op. cit.*

hauteurs mélodiques ne concordent pas. Nous nous servirons des données fournies par l'acousticien M. Kitantou :

> L'oreille est sensible aux variations en fréquence d'un phénomène acoustique. Il a été possible de déterminer la plus petite différence de fréquence Δf qui permette de distinguer la hauteur de deux sons : l'un étant plus grave ou plus aigu que l'autre. Le plus petit intervalle perceptible Δf dépend de la fréquence. Ce n'est pas l'écart que l'on juge mais le rapport entre cet intervalle et la fréquence de référence f, autrement dit Δf/f. [91]

D'après la courbe des variations de Δf/f avec la fréquence présentée par M. Kitantou, nous proposons le Tableau 1.1 qui détermine le seuil de perception différentielle et sa valeur en Hz pour les hauteurs de F_0 susceptibles de nous intéresser (50 à 600 Hz).

HAUTEUR	SEUIL DE PERCEPTION DES FRÉQUENCES	
	rapport	valeur en Hz
50 Hz	10/1000e (0,010)	0,5
100 Hz	5/1000^e(0,005)	0,5
125 Hz	4/1000e (0,004)	0,5
150 Hz	3,8/1000e (0,0038)	0,57
175 Hz	3,6/1000e (0,0036)	0,63
200 Hz	3,5/1000e (0,0035)	0,7
300 Hz	3/1000e (0,003)	0,9
325 Hz	(rapport inchangé	0,975
350 Hz	de 300 à 2000 Hz)	1,05
375 Hz		1,125
400 Hz		1,2
425 Hz		1,275
450 Hz		1,35
475 Hz		1,425
500 Hz		1,5
525 Hz		1,575
550 Hz		1,65
575 Hz		1,725
600 Hz		1,8

TAB. 1.1 – Seuil différentiel des fréquences (SDF) pour F_0 entre 50 et 600 Hz d'après M. Kitantou

[91] KITANTOU, Mpaya (1987, 2e éd. 1994). *art. cit.* (pp.159-161).

Les résultats de ce rapide calcul permettent de relativiser l'importance du caractère logarithmique de la perception de la mélodie rapportée aux valeurs absolues en Hz. Dans la bande de fréquence qui nous intéresse pour le calcul de F_0, nous constatons en effet que les écarts ne dépassent pas 2 Hz. Néanmoins, une telle expérience est fondée sur la perception de sons discrets à classer sur une échelle du plus grave au plus aigu. Elle ne permet donc pas de rendre compte de la perception différentielle des hauteurs mélodiques en parole continue. Ainsi, il faut également prendre en compte le seuil de glissando, c'est-à-dire le seuil qui détermine le moment où l'on commence à percevoir un mouvement mélodique dans un signal continu. M. Rossi[92] donne le seuil de 14% en plus ou en moins à partir de la hauteur de départ. A partir de ses calculs et sur le modèle du Tableau 1.1 précédent, nous proposons le Tableau 1.2 où figurent le seuil de glissando pour des fréquences de 50 à 600 Hz et les valeurs correspondant à la perception d'une chute ou d'une montée à partir de la hauteur spécifiée dans la première colonne. Le Tableau 1.2 présenté ici, montre ainsi qu'une différence de 7 Hz suffit à permettre à l'oreille de percevoir une chute ou une montée mélodique à 50 Hz. A 150 Hz, la différence devra atteindre 20 Hz. A 500 Hz, un mouvement mélodique sera perçu à partir de 70 Hz de différence. Ce tableau nous renseigne donc davantage que le précédent sur la perception des mouvements mélodiques de la parole continue. Les résultats qui en émanent tendent à confirmer la perception logarithmique des phénomènes mélodiques.

En ce qui concerne **seuil différentiel de l'intensité**, nous proposons de nous référer, comme A. Di Cristo, au travail de M. Rossi,[93] qui rend compte d'un seuil de 3 dB entre deux voyelles identiques. Les chercheurs de l'INRP[94] calculent le seuil pour

[92]ROSSI, Mario (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 24 : 129-161.
ROSSI, Mario (1978). The perception of non repetitive intensity glides on vowels. *Journal of Phonetics* 6 : 9-18.

[93]ROSSI, Mario (1971). *art. cit.*

[94]Ces données sont celles présentées sur le site de l'INRP, consulté le 28 avril 2003 :
http ://www.inrp.fr/Acces/JIPSP/phymus/m_lexiq/lexbc2.htm.

Hauteur	Seuil de glissando (SDG) en Hz		
	SDG	F_0 – SDG	F_0 + SDG
50 Hz	7	43	57
100 Hz	14	86	114
125 Hz	17,5	107,5	142,5
150 Hz	21	129	171
175 Hz	24,5	150,5	199,5
200 Hz	28	172	228
300 Hz	42	258	342
325 Hz	45,5	279,5	370,5
350 Hz	49	301	399
375 Hz	52,5	322,5	427,5
400 Hz	56	344	456
425 Hz	59,5	365,5	484,5
450 Hz	63	387	513
475 Hz	66,5	408,5	541,5
500 Hz	70	430	570
525 Hz	73,5	451,5	598,5
550 Hz	77	473	627
575 Hz	80,5	494,5	655,5
600 Hz	84	516	684

TAB. 1.2 – Seuil différentiel de glissando (SDG) pour F_0 entre 50 et 600 Hz d'après M. Rossi

l'intensité de la façon suivante. Soit deux intensités I1 et I2, où I2 est supérieure à I1. Le seuil différentiel est issu du rapport (I2 − I1)/I1. Ainsi l'oreille nous donne-t-elle accès, d'après ce calcul, à plus de 320 sensations intensives différentes entre 0 dB et le seuil de douleur (130 dB), ce qui correspond approximativement au seuil de M. Rossi.[95]

La **perception de la durée** a été quantifiée par différents auteurs. A. Di Cristo rapporte que l'examen de nombre de ces travaux par I. Lehiste[96] l'amène à conclure que les segments phonétiques faisant partie d'une phrase sont caractérisés par un seuil égal à 0,040 s pour une durée de 0,030 à 0,300 s. Dans son article de 1972, M. Rossi[97] teste ce seuil pour la voyelle naturelle isolée [a]. Ses données sont conformes à la loi de Weber-Fechner entre 0,130 et 0,290 s. Dans cette zone qui l'intéresse pour l'étude

[95]ROSSI, Mario (1971). *art. cit.*

[96]LEHISTE, Ilse (1970). *Suprasegmentals.* Cambridge, MA : MIT Press

[97]ROSSI, Mario (1972). Le seuil différentiel de durée. in Valdman, Albert (éd.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre.* The Hague : Mouton. 435-450.

de la durée vocalique et où l'oreille est la plus fine pour détecter des différences de durée, le seuil différentiel de durée est égal à 22,5% en moyenne. En somme, entre 0,060 et 0,140 s, le seuil est constant et proche de 0,030 s. Il augmente au-delà et jusqu'à 0,290 s où il atteint 0,065 s.

Nous conclurons cette partie technique en réaffirmant la nécessité de considérer les différents paramètres comme bénéficiant d'une indépendance certaine, bien que relative, et méritant ainsi une analyse séparée. Nous retiendrons ainsi la notion de geste vocal délibéré et non pas physiologiquement conditionné. Nous étudierons également les paramètres suprasegmentaux en combinaison. On verra que le couplage des indices est nécessaire à une bonne compréhension des phénomènes de mise en place discursive ainsi que des rapports co-énonciatifs. En ce qui concerne la perception, nous avons pu mesurer l'importance de la notion de seuils à franchir, et relativiser l'importance de l'écart entre échelle linéaire et échelle logarithmique lorsqu'il est question de perception différentielle à l'échelle où nous nous plaçons.

1.2 Décrire la langue et l'échange verbal

1.2.1 Décrire la langue

Les modèles présentés ici nous intéressent à plusieurs titres. Tout d'abord, ce sont des modèles extrêmement influents : ils ont constitué et constituent encore le cadre de réflexion d'un grand nombre de chercheurs en linguistique. En outre, avec des moyens et des objectifs différents, ils proposent tous les deux un modèle hiérarchisé de la structuration du discours. Enfin, notons que les instigateurs de ces deux modèles se sont également intéressés à l'intonation. Quelles avancées ces modèles constituent-ils ? Quelle est leur contribution à notre domaine ? Quels enseignements pouvons-nous en retirer, dans la mesure où, en écartant la subjectivité, la perspective adoptée nous

semble être davantage celle de l'élucidation de faits de langue et que de discours ?

1.2.1.1 Le modèle distributionnel

Issu de la réflexion structuraliste américaine, ce modèle est proposé par L. Bloomfield dès les années 1920, avant d'être repris par Z. Harris et par C.F. Hockett. Il a pour objectif d'expliquer la parole de façon externe, mécanique. Ce type de description s'oppose à la conception mentaliste de la parole, qui en fait essentiellement l'expression de la seule subjectivité du locuteur.

Distribution Ce modèle cherche avant tout à déterminer la « distribution » des unités qui l'intéressent. L'étude du contexte linéaire d'occurrence de ces unités, appelé « environnement », constitue une étape importante du travail. Elle consiste à répertorier les unités précédant et suivant l'unité étudiée. A la description de l'environnement linéaire succède une analyse faisant intervenir la notion de complexité, de relation hiérarchique. O. Ducrot & J.-M. Schaeffer rappellent ici la définition de la notion d'« expansion » :

> Soit *b* un segment (unité ou suite d'unités) de l'énoncé *E*. Soit *c* un segment d'un autre énoncé *E'* du corpus. On dira que *b* est un expansion de *c*, si : 1° *c* n'est pas plus complexe que *b* (en ce sens que *c* ne comporte pas plus d'unités que *b*), 2° la substitution de *c* à *b* dans *E* produit un énoncé *E"* du corpus (*b* et *c* ont donc un environnement commun).[98]

La distribution, qui se définit par l'ensemble des environnements où l'on rencontre une unité donnée, aboutit à la création des « classes distributionnelles », c'est-à-dire de classes regroupant toutes les unités ayant la même distribution. Ainsi, l'un des apports du distributionnalisme est de proposer une méthode permettant de restreindre le champ des possibles pour l'apparition d'une unité donnée, et ainsi de noter les

[98] DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Le Seuil. (p.50).

déviances par rapport au schéma le plus couramment rencontré. C'est en ce sens que notre approche se veut également distributionnelle : par exemple en décrivant la réalisation dans notre corpus des possibles offerts par la grammaire.

Segmentation Ce n'est qu'une fois la segmentation opérée et les unités déterminées que l'étude distributionnelle proprement dite peut s'effectuer. Ce modèle part du principe selon lequel le continuum du discours est segmentable de façon non arbitraire :

> The first distributional fact is that it is possible to divide (to segment) any flow of speech into parts, in such a way that we can find some regularities in the occurrence of one part relative to others in the flow of speech. These parts are the discrete elements which have a certain distribution (set of relative locations) in the flow of speech; and each bit of speech is a particular combination of elements.[99]

Pour être efficace, cependant, c'est-à-dire permettre d'isoler des classes d'éléments, la segmentation doit s'opérer à plusieurs niveaux :

> If we wish to be able [...] to obtain elements (or classes of elements) whose distributions will have maximum regularity, we have to divide not only the time flow into successive portions, but also any single time segment (or succession of time segments) into simultaneous components (of one segment length, e.g. a tone, or longer, e.g. a pitch-stress contour).[100]

Les auteurs de l'*Encyclopædia Universalis* nous rappellent l'apport considérable du distributionnalisme avec la notion de « constituants immédiats » (C.I.), qui permettent la décomposition de l'énoncé :

> Chaque phrase se voit attribuer une structure hiérarchisée d'éléments emboîtés les uns dans les autres. On détermine d'abord des éléments assez vastes, les

[99] HARRIS, Zellig S. (1954). Distributional structure. *Word* 10 (2) : 146-162. (p.158).
[100] *ibid.* (p.158).

> C.I. de la phrase, puis on décompose ceux-ci à leur tour en C.I. de C.I., et ainsi de suite jusqu'à ce qu'on obtienne les unités minimales, les morphèmes. Les procédures utilisées diffèrent selon les linguistes, mais peuvent toujours se ramener à une combinaison d'opérations de segmentation et de substitution.

De ce point de vue encore, la recherche des constituants immédiats rapproche ce modèle de nos intérêts. En effet, notre recherche sur les modes de structuration de l'anglais oral passe nécessairement par la détermination des éléments de base qui s'organisent en un ensemble toujours plus complexe. Un modèle proposant une méthode permettant d'établir les principes hiérarchiques de structuration de la langue revêt un intérêt tout particulier pour nous.

Sur quelle base la segmentation en constituants immédiats s'effectue-t-elle donc ? Puisque la méthode de segmentation est non arbitraire, sur quels principes se fonde-t-elle ? La question, essentielle à nos yeux, est d'autant plus problématique que le modèle distributionnaliste tient certes compte des relations entre éléments linguistiques, mais à l'exclusion de leurs relations sémantiques et fonctionnelles. C'est ce que souligne Z.S. Harris au début de son article déjà cité plus haut :

> Here we will discuss how each language can be described in terms of a distributional structure, i.e. in terms of the occurrence of parts (ultimately sounds) relative to other parts, and how this description is complete *without intrusion of other features such as history or meaning.*[101]

O. Ducrot & J.-M. Schaeffer [102] reprochent aux distributionnalistes de ne pas se poser la question des critères de délimitation des unités de façon satisfaisante. En fait, la méthode distributionnelle se fonde sur des opérations de base permettant de délimiter des segments : la substitution et la division.

Pour découper un segment b sans se fonder sur son sens, et d'une façon qui

[101] *ibid.* (p.146). Nous soulignons.
[102] DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *op. cit.*

> ne soit pas pour autant arbitraire, on le compare à un segment *c* dont *b* est un expansion, et dont l'analyse s'impose par le fait de deux unités minimales *c'* et *c"*. On découpe alors *b* en deux segments *b'* et *b"*, choisis pour être, respectivement, des expansions de *c'* et *c"*.[103]

1.2.1.2 Le modèle génératif

La structure en constituants et sa représentation arborescente, issues du distributionnalisme, se retrouvent dans les règles syntagmatiques développées par N. Chomsky. Elles sont intégrées à un modèle fondé sur une opposition entre surface et profondeur, la structure de surface étant le résultat de transformations appliquées à la structure profonde selon des règles précises. Ces règles sont construites à partir des relations syntagmatiques inscrites dans la langue et permettent de prédire toutes les phrases grammaticales possibles d'une langue. Les règles transformationnelles permettent ensuite d'expliquer l'ordre des mots dans le linéaire de surface, transformé à partir de la structure profonde. Ainsi, par exemple, une forme passive est conçue comme la transformation d'une phrase originelle à la voie active. De même, l'interrogation procède de l'assertion.

La représentation arborescente de la structure de ces phrases fait apparaître plusieurs niveaux de profondeur hiérarchique et illustre la caractéristique récursive de la langue. Elle permet la visualisation de la position de chaque élément dans la hiérarchie en fonction des nœuds où ils apparaissent. Ainsi, ce modèle formel assigne à tout constituant un cas, une place identifiée de façon non ambiguë.

1.2.1.3 Limites de l'influence de ces modèles dans notre travail

Ces modèles hiérarchiques de description de la langue nous intéressent pour la méthode qu'ils ont mise au point et pour leur rigueur dans son application. Cependant,

[103] *ibid.* (pp.50-51).

leur influence sur notre travail reste relativement limitée.

La raison essentielle réside dans le fait que notre objet d'étude n'est pas la langue mais bien le discours. Celui-ci se donne comme un tout dont les multiples unités ne sont pas données d'avance. Or O. Ducrot & J.-M. Schaeffer affirment que la méthode distributionnelle américaine est une démarche diamétralement opposée à celle du structuralisme saussurien en ce qu'elle part des unités pour définir le système qu'elles contribuent à édifier.

> Du point de vue de la linguistique saussurienne, le distributionalisme soulève certaines difficultés, dont la plus souvent signalée concerne la détermination des unités. Pour Saussure, les éléments ne sont jamais donnés, et leur découverte ne fait qu'un avec celle du système [...]. Or une étude distributionnelle semble impliquer, par définition, la connaissance préalable des éléments ; pour établir la distribution d'une unité, il faut avoir déterminé cette unité (i.e. l'avoir délimitée [...] dans la chaîne parlée, et être capable de l'identifier à travers ses diverses occurrences), et avoir déterminé aussi les unités qui constituent ses environnements.[104]

En outre, une analyse du discours telle que nous la concevons doit prendre en compte la dimension subjective de l'usage de la langue. Les transformations de la grammaire générative éliminent cette dimension en faisant d'un énoncé le résultat de l'application de règles. Cette théorie, de même que la théorie distributionnelle, ne prête pas suffisamment attention, d'une part, au contexte discursif et intersubjectif dans lequel il est produit et, d'autre part, aux choix conscients ou inconscients qui émanent du contexte et dont il est le fruit.

Enfin, les « imperfections » de la langue spontanée sont considérées par la grammaire générative comme des erreurs de performance, tandis qu'elles occupent pour nous une place importante dans la structuration du discours. La méthode hypothético-

[104]DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *op. cit.* (p.54).

déductive de N. Chomsky cherche à écrire les règles capables de produire toutes les phrases possibles d'une langue donnée. Ce n'est pas notre objectif. Nous avons l'ambition plus empirique de proposer quelques hypothèses concernant l'utilisation de l'anglais à l'oral spontané en nous fondant sur l'observation d'énoncés avérés formant la texture des conversations spontanées de notre corpus.

1.2.1.4 Autres modèles hiérarchiques pour la description de faits syntaxiques

D'autres hiérarchies ont été proposées à partir de l'observation des faits syntaxiques. Moins ambitieuses que les deux modèles précédents, elles s'attachent à décrire certains phénomènes précis. Leur portée dépasse pour certaines la seule description de la langue pour intéresser également le discours.

Hiérarchie d'accessibilité On pense en premier lieu à la hiérarchie d'accessibilité des groupes nominaux (*Accessibility Hierarchy*) de E. Keenan et B. Comrie.[105] Ces deux auteurs établissent l'existence d'une échelle des fonctions grammaticales endossées par les groupes nominaux et placent le sujet au sommet de celle -ci devant l'objet direct et l'objet indirect.

sujet > objet direct > objet indirect [106]

Issue d'une réflexion sur la probabilité pour un groupe nominal d'accepter une subordonnée relative en fonction de la position de celui-ci dans la phrase, la hiérarchie d'accessibilité s'applique également à d'autres phénomènes syntaxiques.

[105] KEENAN, Edward & COMRIE, Bernard (1977). Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8 : 62-100.

[106] COLLINGE, N.E. (éd.) (1990). *An Encyclopædia of Language.* Londres, New-York : Routledge. (p.310).

Hiérarchie de personne De même, la hiérarchie de personne (*Animacy Hierarchy*), formulée par M. Silverstein[107] en 1976, trouve un écho dans de nombreux travaux, notamment ceux de G. Lazard[108] qui lui ajoute les notions défini/indéfini et générique/spécifique. Nous reproduisons ici la formulation d'origine telle que mentionnée dans l'encyclopédie de N.E. Collinge :

> 1st & 2nd person non-singular pronouns > 1st & 2nd person singular pronouns > 3rd person pronouns > proper nouns > human proper nouns > animate common nouns > inanimate common nouns.[109]

Ce modèle a d'abord été conçu comme un moyen de déterminer la probabilité pour un élément de la hiérarchie de devenir sujet grammatical dans une construction ergative. Plus un élément se situe à gauche (en haut) de cette échelle, plus il est susceptible d'être à l'origine d'une action, et inversement. De ce fait, lorsqu'un élément situé à droite (en bas) de l'échelle est, fait exceptionnel, actant, il est marqué comme ergatif. Cette remarque a un corollaire important : le cas accusatif sera d'autant plus marqué que l'élément concerné se trouvera en haut de l'échelle.

C'est en ce sens que ce genre de réflexion hiérarchique sur la langue intéresse notre travail. Si les hiérarchies d'ordre syntaxique font partie de notre réflexion, c'est en tant qu'elles reflètent un usage de la langue — usage qui s'illustre à travers les choix opérés par les locuteurs au sein des possibilités ouvertes par la grammaire.

Hiérarchies pragmatiques De telles hiérarchies sont essentiellement syntaxiques, comme le précise G. Lazard.

> On voit ainsi s'esquisser une hiérarchie de l'ensemble des termes nominaux

[107]SILVERSTEIN, Michael (1976). Hierarchy of Features and Ergativity. in Dixon, R. M. W. (éd.), *Grammatical Categories in Australian Languages* : 112-171. Canberra : Australian Institute of Aboriginal Studies.

[108]LAZARD, Gilbert (1985). La notion de distance actancielle. in Bouscaren, J., Franckel, J.-J. & Robert, S. (éds.). *Langues et langage. Problèmes et raisonnement en linguistique — Mélanges offerts à Antoine Culioli* : 135-146. Paris : PUF.

[109]COLLINGE, N.E. (éd.) (1990). *op. cit.* (p.312).

de la proposition, depuis les termes incorporés jusqu'aux actants les plus périphériques et les circonstants. (16) la figure à l'aide de quelques jalons : terme incorporé (NV ; l'ordre des sigles N et V est arbitraire), objet proche (N_{-1}), objet plus distant (N_0), terme « datif » (N_i), actant périphérique (N_j), circonstant (N_k).

(16) NV N_{-1} N_0 N_i N_j N_k

Deux remarques s'imposent ici pour éviter tout malentendu. La première est que l'échelle figurée par (16) n'a qu'une valeur d'exemple. Il va de soi que la hiérarchie des actants classés selon leur « distance » au verbe doit être établie dans chaque langue et que les positions pertinentes varient d'une langue à l'autre. [...]

Deuxième remarque. Cette hiérarchie est grammaticale, non sémantique (ni pragmatique). Elle se fonde exclusivement sur des critères morphosyntaxiques : distance dans la chaîne et autres propriétés distributionnelles ou transformationnelles, telles que l'incorporation, le comportement dans une mise au passif (comme en hébreu, en anglais) et d'autres, qu'il convient de découvrir dans chaque langue. C'est justement sa nature proprement grammaticale qui en fait l'intérêt linguistique.[110]

Cependant, pour une langue donnée, le domaine d'application d'une hiérarchie syntaxique dépasse largement le cadre purement syntaxique dès que l'on s'intéresse à sa pertinence pour l'explication du discours. Si l'établissement de l'échelle de G. Lazard s'est faite à partir de critères syntaxiques et non pragmatiques, son incidence est mesurable en termes pragmatiques ou plus précisément énonciatifs. En effet, la langue en usage est produite et reçue. A ce titre, elle met en jeu l'anticipation et la représentation de la pensée de l'autre. Or, comme le suggère R. Artstein,[111] l'application des hiérarchies d'accessibilité et de personne en discours permet de déterminer si

[110]LAZARD, Gilbert (1985). *art. cit.* (pp.141-142).
[111]ARTSTEIN, Ron (1998). Hierarchies. Manuscrit disponible en ligne :
http ://www.cs.technion.ac.il/∼artstein/hierarchies.pdf, consulté le 22 décembre 2002.

une construction est marquée ou non marquée, ce qui revient précisément à déterminer dans quelle mesure le locuteur joue sur les pensées et représentations qu'il prête à son interlocuteur.

> [...] the co-occurrence of an element from the high end of one scale with an element from the low end of another scale is marked. It is not that themes are more marked than agents or that agents are more marked than themes, but it is more marked to have a theme as subject than to have an agent as subject; it is also more marked to have a theme as a subject than to have a theme as object. The markedness relations derived from the abstract hierarchies enable us to incorporate the effects of the hierarchies into the grammar, by excluding marked configurations in favor of less marked ones.[112]

1.2.2 Décrire l'échange verbal

Différents modèles proposent une description non plus de la structure du contenu de l'échange mais de l'organisation générale de sa forme. Nous pensons ici au modèle communicationnel et informationnel, au modèle interactionnel, et enfin au modèle énonciatif.

1.2.2.1 Le modèle communicationnel et informationnel

Le modèle communicationnel prend son origine dans les travaux de C. Shannon[113] pour sa conception télégraphique sémioticienne, et de H.P. Grice[114] pour son versant inférentiel. Le modèle de la communication télégraphique est établi par C. Shannon & E. Weaver à la fin des années 1940.[115] Les travaux de cet ingénieur des télécom-

[112] *ibid.* (p.2).
[113] Voir notamment SHANNON, Claude E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27 : pp.379-423 et 623-655.
[114] GRICE, Herbert Paul (1957). Meaning. *Philosophical Review* 66 : 377-388.
[115] SHANNON, Claude E. & WEAVER, Warren (1949, trad. 1975). *Théorie mathématique de la communication.* Paris : Retz.

munications et de ce philosophe ont notamment consisté à définir un schéma de la communication impliquant une source d'information, un message, un canal, une machine réceptrice et un destinataire :

> By a communication system we will mean a system of the type indicated schematically in Fig. 1.[116] It consists of essentially five parts:
>
> 1. An information source which produces a message or sequence of messages to be communicated to the receiving terminal. [...]
> 2. A transmitter which operates on the message in some way to produce a signal suitable for transmission over the channel. In telephony this operation consists merely of changing sound pressure into a proportional electrical current. [...]
> 3. The channel is merely the medium used to transmit the signal from transmitter to receiver. It may be a pair of wires, a coaxial cable, a band of radio frequency, a beam of light, etc.
> 4. The receiver ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
> 5. The destination is the person (or thing) for whom the message is intended.[117]

Ce schéma, qui est à la base de tous les systèmes de communication actuels, rend essentiellement compte de la dimension technique d'un échange d'information unilatéral consistant en une simple opération d'encodage-décodage. Pour influent qu'il ait pu être et soit encore parmi les ingénieurs des télécommunications et les sémioticiens, il ne nous paraît pas satisfaisant pour décrire l'échange linguistique tel qu'il se déroule dans notre corpus.[118]

[116] Voir Figure 1.8, tirée de SHANNON, Claude E. (1948). *art. cit.* (p.381).
[117] *ibid.* (pp.380-381).
[118] Notons que la théorie cybernétique de N. Wiener introduit au même moment la notion de rétroaction, ou *feedback*, du récepteur vers l'émetteur. Voir WIENER, Norbert (1948). *Cybernetics.*

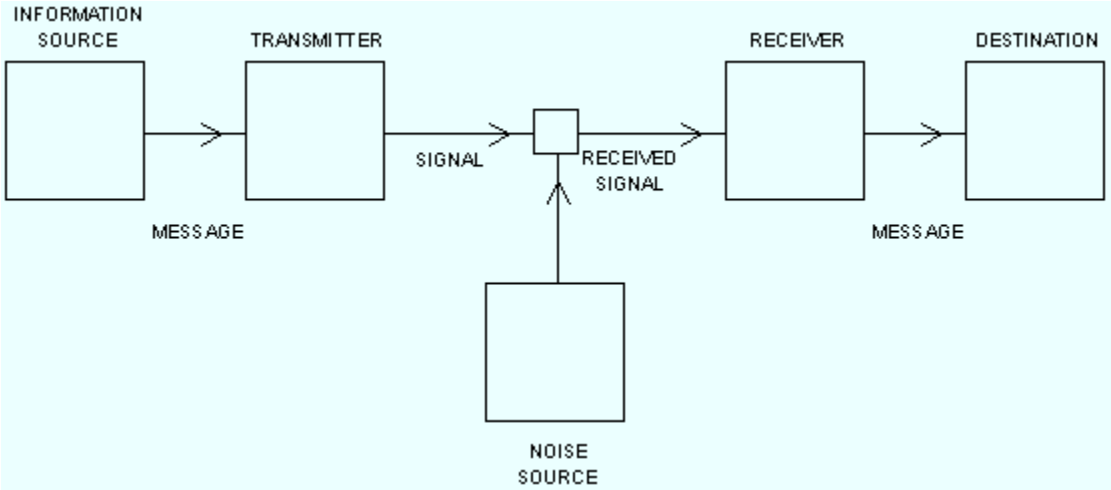

FIG. 1.8 – Diagramme schématique du système de la communication d'après C. Shannon

Au sein du domaine de la linguistique, il s'oppose au modèle de la communication proposé par H.P. Grice dès la fin des années 1950.[119] Celui-ci estime que le décodage littéral du message ne suffit pas à épuiser le contenu de ce qui est communiqué. L'extralinguistique a sa place dans ce qui est non seulement signifié mais également induit par le message : le moment et le lieu de l'énonciation, l'identité des interlocuteurs, et, chose importante, les intentions du locuteur et les attentes de l'interlocuteur. Il propose ainsi un modèle de communication inférentiel fondée sur un rapport de coopération bi-latérale entre les participants. Il formule neuf maximes classées en quatre catégories. Quantité, qualité, rapport et manière doivent respecter certains principes pour un bon fonctionnement de l'échange verbal.[120]

Maxims of quantity

1. Make your contribution as informative as is required (for the current purposes of the exchange).
2. Do not make your contribution more informative than is required.

J. Wiley : New York.

[119] GRICE, H.P. (1957). *art. cit.*

[120] Ces maximes sont rappelées dans SPERBER, Dan & WILSON, Deirdre (1986, 2e éd. 1995). *Relevance — Communication and Cognition.* Oxford, Cambridge, MA : Blackwell. (pp.33-34).

Maxims of quality.

Supermaxim: Try to make your contribution one that is true.

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Maxim of relation

Be relevant.

Maxims of manner

Supermaxim: Be perpiscuous.

1. Avoid obscurity of expression.
2. Avoid ambiguity.
3. Be brief (avoid unnecessary prolixity).
4. Be orderly.

Pour D. Sperber et D. Wilson, ces deux modes de communication co-existent sans être réductibles l'un à l'autre. Dans le cadre de leur théorie de la pertinence, ils reprennent de P. Grice l'idée de double intentionnalité déjà évoquée dans notre introduction générale.[121] Ils définissent la notion de contexte à la fois comme un construit psychologique, qui constitue pour l'interlocuteur un ensemble de prémisses qu'il est susceptible d'utiliser dans son activité non seulement de décodage, mais aussi d'inférence, et comme un construit textuel, qui s'élabore au fur et à mesure que l'échange avance. Le contexte conçu comme construit psychologique est propre à chaque locuteur, si bien qu'il est impossible de postuler une notion de connaissance partagée.[122] La théorie communicationnelle de C. Shannon est aussi une théorie de

[121] Voir p.23.

[122] Voir SPERBER, Dan & WILSON, Deirdre (1986, 2e éd. 1995). *op. cit.* (p.16) :

> A context in this sense is not limited to information about the immediate physical environment or the immediately preceding utterances: expectations about the future, scientific hypotheses or religious beliefs, anecdotal memories, gen-

l'information, dont ils retiennent la notion essentielle d'information et de quantité d'information.

> In information theory, one of the basic notions is that of the amount of information associated with a given situation. "Information" here, although related to the everyday meaning of the word, should not be confused with it. In everyday usage, information usually implies something about the semantic content of a message. For the purposes of communication theory, the "meaning" of a message is generally irrelevant; what is significant is the difficulty in transmitting the message from one point to another.
>
> From this point of view, information exists only when there is a choice of possible messages. If there were only one possible message there would be no information; no transmission system would be required in such a case, for this message could be on a record at the receiving point. Information is closely associated with uncertainty. The information I obtain when you say something to me corresponds to the amount of uncertainty I had, previous to your speaking, of what you were going to say. If I was certain of what you were going to say, I obtain no information by your saying it.[123]

A partir de la probabilité associée au contenu d'un message se conçoit une théorie de la pertinence, c'est-à-dire de l'effet produit en fonction de l'énergie dépensée. Plus le contenu d'un message est prévisible et plus il est redondant avec d'autres signes, par exemple, moins la quantité d'information qui lui est associée est importante, moins il est efficace. De même, plus le contenu d'un message est difficile à interpréter, moins il est efficace. Lorsque le contenu d'un message est peu prévisible, la quantité d'information qui lui est associée est élevée. Il revêt de ce fait un haut degré de

eral cultural assumptions, beliefs about the mental state of the speaker, may all play a role in interpretation.

[123]SHANNON, Claude E. (1950). Communication Theory — Exposition of Fundamentals. *IRE Transactions on Information Theory* 1 : pp.44-47. 1950. (p.44).

pertinence. D'où la possibilité d'une théorie sur les notions d'information donnée et information nouvelle, sur l'organisation générale des énoncés autour du thème et du rhème, et sur les notions de topique et focus.

1.2.2.2 Le modèle interactionnel

La théorie de la pertinence est un modèle pragmatique d'analyse de l'échange verbal, à laquelle une dimension interactionnelle peut s'ajouter. Nous nous intéressons à présent au modèle proposé notamment par l'école de Genève, où l'influence de D. Sperber et D. Wilson est perceptible. Autour d'E. Roulet, et à la suite des chercheurs de Birmingham (J. Sinclair et M. Coulthard), l'école de Genève se départit d'une conception de l'échange verbal comme une succession d'unités informationnelles pour proposer une analyse en actes de langage. En outre, les actes de langage sont organisés de façon rigoureuse en diverses unités d'ordre supérieur. Dans son compte rendu sur l'organisation structurale des conversations, C. Kerbrat-Orecchioni retient cinq niveaux principaux, deux relevant du monologue (c'est-à-dire dont la production revient à un seul locuteur) et trois du dialogue (produites par plusieurs locuteurs).[124]

unités dialogales	Interaction Séquence Echange
unités monologales	Intervention Acte de langage

TAB. 1.3 – Modèle hiérarchique interactionnel

C. Kerbrat-Orecchioni décrit ensuite les unités des cinq rangs relevés. L'*interaction*, parfois appelée *incursion*, est théoriquement définie par quatre paramètres qui

[124]Le tableau 1.3 est celui de KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales*. Paris : Armand Colin. (Tome 1, p.213).

ont font l'unité de rang maximal à laquelle la pragmatique s'intéresse actuellement, et dont l'auteur remet la pertinence en question. Ainsi, elle reformule la définition d'une interaction :

> *Pour qu'on ait affaire à une seule et même interaction, il faut et il suffit que l'on ait un groupe de participants modifiable mais sans rupture, qui dans un cadre spatio-temporel modifiable mais sans rupture, parlent d'un objet modifiable mais sans rupture.*[125]

La *séquence* revêt une unité thématique et pragmatique : elle est le lieu d'accomplissement d'un seul « but ». D'après l'auteur, les séquences les plus clairement perceptibles sont celles qui ouvrent ou closent une conversation, dans la mesure où elles sont caractérisées par un certain rituel. L'*échange*, considéré comme l'unité minimale qui puisse être produite par plusieurs locuteurs, est néanmoins décrit de façon peu satisfaisante indépendamment des unités qui le composent. Celles-ci relèvent du domaine monologal. L'*intervention*

> [...] ne se confond pas avec le tour de parole : unité fonctionnelle, elle ne se définit que par rapport à l'échange, et plus précisément, comme *la contribution d'un locuteur particulier à un échange particulier.*[126]

Cependant, sa délimitation est également souvent problématique du fait de la multiplicité des types d'interventions et de leurs réalisations effectives en conversation naturelle. Un échange est typiquement constitué de deux interventions, l'une dite initiative et l'autre réactive, et rarement de plus de trois ; la dernière, constituant en général une ratification de la deuxième, est appelée évaluative. L'organisation des interventions, et de l'unité supérieure qu'est l'échange, n'est pas forcément linéaire en profondeur. C. Kerbrat-Orecchioni propose ainsi de distinguer des échanges suivis, croisés et des échanges embrassés, impliquant donc un certain nombre d'interventions.

[125] *ibid.* (Tome 1, p.216).
[126] *ibid.* (Tome 1, p.225).

L'*acte* textuel se distingue de l'acte de langage tel que défini par Austin et par Searle. Unité minimale, elle ne correspond pas forcément aux unités syntaxiques, mais elle est la plus petite à pouvoir entrer en relation de dominance ou de subordination avec un autre acte. Elle s'entend en outre comme

> la plus petite unité qui provoque un changement dans l'état du savoir partagé (mémoire discursive) des interlocuteurs[127]

Le modèle interactionnel nous semble d'autant plus intéressant qu'il propose une analyse de l'échange verbal en termes hiérarchiques. S'il est fondé sur une définition du discours qui prend en compte l'« autre » ainsi que le contexte dans lequel l'échange a lieu,[128] il ne constitue cependant pas le modèle que nous souhaitons adopter pour notre recherche. La raison principale en est son orientation essentiellement pragmatique, c'est-à-dire orientée vers l'action :

> La dimension hiérarchique contient également l'une des hypothèses constitutives du modèle, selon laquelle le discours est conçu comme la face émergente d'un processus sous-jacent de **négociation** qui se traduit en une structure textuelle hiérarchisée.[129]

Notre intérêt pour l'échange verbal ne va pas à sa matérialité. Nous ne concevons en effet pas toute prise de parole comme le lieu d'une négociation concrète. Nous préférons négliger la dimension qu'A.C. Simon appelle « praxéologique » de l'échange verbal, pour nous tourner vers sa dimension intersubjective sur le plan symbolique — dimension prise en charge par une approche énonciative.

[127]SIMON, Anne Catherine (à paraître). *Segmentation et structuration prosodiques du discours : Une approche multidimensionnelle et expérientielle.* v.2.0 2002 destinée à la publication. Thèse de Doctorat. Louvain-la-Neuve : Université Catholique de Louvain. (p.100).

[128]Nous avons vu ce que la linguistique pragmatique apporte en conséquence à la définition du discours utilisées dans ce travail. Voir p.19.

[129]*ibid.* (p.69).

1.2.2.3 Le modèle énonciatif

Le modèle issu des travaux d'E. Benveniste puis d'A. Culioli propose un autre schéma de la communication verbale, et un autre modèle de structuration de l'échange. L'échange verbal est conçu par A. Culioli comme une rencontre de subjectivités.

> [...] l'activité de langage ne consiste pas à véhiculer du sens, mais à produire et à reconnaître des formes en tant que traces d'opérations (de représentation, référenciation et régulation). La signification n'est donc pas véhiculée, mais (re)-construite. La relation entre production et reconnaissance suppose la capacité d'ajustement entre les sujets. Cette capacité ne permet que rarement un ajustement strict. C'est parce qu'il y a un jeu inter-sujets qu'il y a *du jeu* dans l'ajustement.[130]

A partir de cette conception se développe une analyse des textes, écrits ou oraux, dont l'attention est portée à la détection des marques de l'inscription dans la langue de la subjectivité des parties prenantes de l'échange. L'énonciateur, origine des opérations de repérage, est une entité abstraite représentant le locuteur dans sa capacité à attribuer un contenu de pensée à son interlocuteur, porté au rang de co-énonciateur. Enonciateur et co-énonciateur construisent au cours de leur échange un ensemble de représentations mentales constituant leur connaissance partagée. Cet ensemble toujours en expansion s'appuie sur la construction d'un consensus.

La sémantique formelle des énonciativistes propose un modèle syntaxique fondé sur le principe de la dépendance. La relation prédicative potentielle contenue dans le schéma de lexis n'est validée qu'en fonction des coordonnées énonciatives de l'énoncé. Dans les limites de ce que propose la grammaire, l'organisation des unités syntaxiques suit un ordre qui fait l'objet d'un choix de la part de l'énonciateur. On retrouve donc ici des préoccupations très proches de celles des tenants de certaines hiérarchies

[130]CULIOLI, Antoine (1990). La linguistique : de l'empirique au formel. *Sens et place des connaissances dans la société.* Centre de Meudon-Bellevue, CNRS. 1987. Reprinted in *Pour une linguistique de l'énonciation.* Gap : Ophrys. (Tome 1, débat p.25).

syntaxiques à portée pragmatique.

Le modèle énonciatif trouve dans l'analyse de l'oral un terrain privilégié dans la mesure où, d'une part, le temps de l'énonciation coïncide avec celui de la locution, et, d'autre part, les entités du couple énonciateur et co-énonciateur sont davantage que des construits théoriques. La conversation spontanée voit s'inscrire dans sa langue les traces des enjeux de la relation intersubjective qui se noue à mesure qu'elle se construit.

Au terme de ce premier chapitre s'est dégagé l'apport des études antérieures à notre projet de recherche. Issus de traditions différentes, plusieurs modèles de description des phénomènes intonatifs s'offrent à nous. Chacun d'entre eux propose une conception intéressante de l'intonation, fondée sur un certain nombre d'hypothèses et de partis-pris plus ou moins différents des nôtres. Nous avons également pu mesurer l'apport non négligeable de l'acoustique et de la psychoacoustique à l'analyse de l'oral spontané, qui permet notamment de déterminer dans quelle mesure une variation intonative peut être perçu comme un signe linguistique. Enfin, les modèles de description de la langue, caractérisés par une grande rigueur méthodologique, nous sont néanmoins apparus inaptes à saisir la particularité du discours oral. Une linguistique du discours nécessite en effet des partis-pris et hypothèses propres, ainsi que des outils spécifiques. Si certains des principes sur lesquels s'appuie notre projet de recherche sont apparus au cours de ce chapitre, il nous faut les définir de façon plus approfondie.

Chapitre 2

Principes théoriques

Ce chapitre rappelle les bases du cadre théorique général dans lequel s'inscrit notre recherche sur l'anglais oral. Celle-ci s'appuie sur les travaux de M.-A. Morel et L. Danon-Boileau tels qu'ils nous parviennent et ont pu nous parvenir à travers leur enseignement et leur ouvrage intitulé *Grammaire de l'intonation*,[1] ainsi que sur ceux du Centre de Recherche en Morphosyntaxe du Français contemporain de l'Université Paris III, dirigé par M.-A. Morel.

Nous présentons en premier lieu ce modèle d'analyse de la langue orale, établi tout d'abord pour le français, et enrichi par une réflexion sur des langues telles que le turc, le persan, l'arabe, le coréen, le thaï et l'anglais. Fondé en partie sur la distinction entre valeurs iconiques et valeurs conventionnelles, il n'est pas possible de l'adopter tel quel. Nous verrons ainsi ce que nous pouvons en conserver pour l'anglais, et ce qui nécessite une adaptation pour rendre compte de la spécificité de cette langue.

La deuxième moitié du chapitre est consacrée à la présentation de nos hypothèses de départ. Formulées au sein du cadre théorique général que nous aurons circonscrit, elles concernent notre objet d'étude, c'est-à-dire le fonctionnement de la hiérarchi-

[1] MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation, L'exemple du français*. Gap : Ophrys.

sation des constituants du discours.

2.1 Présentation du modèle général adopté

2.1.1 Le modèle de M.-A. Morel et L. Danon-Boileau

Notre recherche s'inscrit dans le sillage du travail effectué par M.-A. Morel et L. Danon-Boileau sur le français, consigné notamment dans leur *Grammaire de l'intonation*.[2] Les deux auteurs y développent non seulement une théorie de l'intonation, mais aussi de la construction du français oral. Leur modèle, issu de recherches pluridisciplinaires, prend en compte les dimensions morphosyntaxiques, sémantiques, et interactionnelles du discours. Il offre ainsi une vision globale de la construction du discours à l'oral tout en reconnaissant le caractère hétérogène du matériau. C'est à ce titre que nous nous en réclamons.

A partir de l'observation d'un important corpus obtenu dans des situations de parole extrêmement variées, les deux auteurs formulent des hypothèses concernant aussi bien l'organisation du segmental que du suprasegmental, et de la dimension mimo-gestuelle (que nous n'exploitons pas ici).

2.1.1.1 Les marqueurs suprasegmentaux

M.-A. Morel et L. Danon-Boileau proposent une interprétation co-énonciative des phénomènes intonatifs en termes de marqueurs. Leur travail prend pour point de départ la valeur iconique de base de chacun des paramètres intonatifs (F_0, intensité, durée et pauses), c'est-à-dire leur valeur universelle. Il consiste ensuite, en observant les effets des diverses combinaisons possibles en contexte, à isoler leur valeur conventionnelle pour le français. Nous nous intéressons ici à la valeur iconique de ces paramètres.

[2] *ibid.*

Les **variations mélodiques** indiquent de façon iconique l'état de la relation inter-subjective. Elles témoignent des représentations que l'on se fait de la pensée de celui auquel on s'adresse. La montée mélodique constitue pour M.-A. Morel et L. Danon-Boileau une *deixis* vocale : il s'agit de faire appel à l'attention de l'autre, et de susciter l'ouverture d'un champ de co-énonciation (situé en plage mélodique haute). La chute mélodique indique en revanche la fin de la prise en compte de la pensée de l'autre. Elle marque un égocentrage, un repli sur soi de la part de l'énonciateur qui refuse de discuter avec son interlocuteur de ce qu'il avance, ou simplement qui indique à celui-ci que ce qu'il dit coule de source et n'a pas lieu d'être remis en question. La chute mélodique et le maintien en plage basse sont caractéristiques des incises qui, retirées du champ de la co-énonciation, n'appellent pas de commentaire.

L'**intensité** et ses variations sont les témoins de la gestion du droit à la parole. Ainsi, le maintien d'une intensité haute permet en cas de conflit de conserver la parole. Inversement, une chute d'intensité indique à l'interlocuteur qu'il est invité à prendre la parole.

> Les variations d'intensité indiquent si celui qui parle veut défendre son tour de parole ou y renoncer, ce qui peut prendre parfois une valeur iconique d'expressivité. De manière conventionnelle, toute montée d'intensité permet au locuteur de gérer les tours de parole, de prévenir une interruption du co-locuteur, ce qui souligne l'importance qu'il accorde à ce qu'il est en train de dire.[3]

L'intensité se situe donc iconiquement dans la dimension matérielle et interactionnelle de l'échange, appelée co-locution par les deux auteurs.

Il existe plusieurs types de **pauses silencieuses**. Le son étant créé par des différences de pression d'air et propagé par l'air, le locuteur doit constamment procéder à de courtes pauses pour respirer. Ces pauses, de 0,20 à 0,30 s, reviennent une vingtaine de fois par minute. Nécessité biologique, elles n'ont pas de valeur particulière. Il est

[3] *ibid.* (p.16).

cependant possible d'attribuer une valeur iconique aux pauses silencieuses de durée supérieure.

> La *pause silencieuse* indique un tournant au sein d'un cadre déjà constitué.
>
> Sur la base d'une attention accordée, elle met en relief le discours qui va suivre, et permet d'homogénéiser ce qui a précédé ou d'annuler une opération qui a pu y être ébauchée.[4]

On désigne par le terme générique de **durée** tous les phénomènes d'allongement des syllabes. Dans le modèle proposé par M.-A. Morel et L. Danon-Boileau, la durée est fonction de la représentation que l'énonciateur a du discours qu'il est en train de construire. En ce sens, elle est révélatrice du travail de formulation.

Ainsi, de même que le segmental, le suprasegmental possède des marqueurs dont la valeur est formulable en termes co-énonciatifs. Ceux-ci nous intéressent dans la mesure où il participent à la structuration du discours à l'oral.

2.1.1.2 Un modèle de structuration du discours oral

Les variations conjointes des paramètres intonatifs fonctionnent pour les deux auteurs comme autant d'indices de coupe et de cohésion au sein du continuum sonore. Ils proposent ainsi une segmentation de ce continuum en paragraphes dont ils décrivent les trois constituants immédiats.

L'organisation générale d'un paragraphe oral est issue de la distinction entre thème et rhème, opposition courante en linguistique. Elle est cependant plus complexe. La composition du paragraphe oral tel qu'il existe en français se fait autour d'un ou plusieurs constituants. Un paragraphe consiste au moins d'un rhème, précédé d'un ou plusieurs préambules. Le dernier rhème d'un paragraphe est parfois suivi d'un postrhème (également appelé *afterthought*, *postjet*, *postfix*, *mnémème* ou *antitopic*

[4] *ibid.* (p.13).

par d'autres auteurs). Un paragraphe oral type en français est donc construit sur le modèle suivant :

préambule + rhème + (postrhème)

Cette unité, nous signalent les deux auteurs, n'est cependant pas parfaitement stable. Elle obéit en effet au principe de recatégorisation qui permet à une séquence préambule-rhème de servir de préambule aux séquences suivantes. L'apparition d'un postrhème, en revanche, permet de clore la séquence préambule-rhème de façon non ambiguë. Au sein des paragraphes, ce sont les indices segmentaux uniquement qui indiquent la nature des constituants discursifs.

Généralement très développé, le **préambule** présente typiquement au moins l'un des éléments suivants. Rarement tous présents, ils confèrent au préambule différentes valeurs — une valeur relationnelle et connective avec ce qui précède, une valeur thématique, et/ou une valeur modale. Les deux auteurs notent que, en l'absence de marques de fonctions syntaxiques, l'apparition de ces éléments respecte un ordre fixe.

ligateur + point de vue + modus dissocié + cadre + support lexical disjoint[5]

D'un point de vue intonatif, la fin du préambule français est marquée par une montée mélodique.

Le **rhème** désigne le court segment qui suit le préambule. Utilisé dans diverses acceptions en linguistique, le rhème ne fournit pas forcément une information nouvelle. En revanche, d'après les deux auteurs, son contenu contrevient aux attentes que l'on prête à l'interlocuteur concernant l'objet de discours établi dans le préambule. D'un point de vue syntaxique, il est construit autour d'un verbe conjugué, précédé par un pronom (personnel, existentiel ou présentatif), et suivi d'une séquence de nature variable (groupe nominal, prépositionnel ou adverbial). D'un point de vue intonatif, le rhème français est affecté d'un motif mélodique en cloche (bas-haut-bas).

[5] *ibid.* (p.38).

Le **postrhème** a des caractéristiques syntaxiques, prosodiques et sémantiques très strictes. D'un point de vue intonatif, le postrhème est nécessairement caractérisé en français par une absence de pause après le rhème, un maintien en plage basse mélodique (H1 ou H2), une chute de l'intensité, une absence de variations mélodiques et une absence de remontée sur la dernière syllabe. Les syllabes formant le postrhème n'excèdent pas le nombre de sept à huit.[6] Du point de vue de la syntaxe et de la sémantique, deux types de fonctions suffisent à caractériser les postrhèmes. Soit ils expriment une modalité épistémique ou un point de vue. Soit ils contiennent un argument nominal en relation de co-référence et de cataphore avec un pronom du rhème (ce qui s'apparente au support lexical disjoint).

Nous reprenons ici le découpage en paragraphes d'un extrait du corpus présenté dans la *Grammaire de l'intonation*.[7] Dans cet exemple, les trois constituants du paragraphe oral sont représentés, comme on le voit dans le Tableau 2.1. La segmentation provient de la seule prise en compte des paramètres intonatifs, indépendamment du plan morpho-syntaxique.

tu vois moi j'crois qu'c'est pas comme ça qu'ça doit marcher la société

Le modèle de structuration du français oral spontané proposé dans la *Grammaire de l'intonation* admet des configurations plus complexes que celle dont nous rendons brièvement compte ici. Nous serons amenée à y revenir au moment d'évaluer l'adéquation du modèle à la description de la langue qui nous intéresse.

[6]Cette valeur (communication personnelle de M.-A. Morel) infirme les premiers résultats des recherches de M.-A. Morel et L. Danon-Boileau qui donnaient pour valeur maximale le nombre de douze. Ce résultat corrigé correspond du reste au nombre maximal de syllabes par groupe constaté en français. Il correspond également, d'un point de vue cognitif, à la capacité de la mémoire de travail.

[7]*ibid.* (p.22).

EXEMPLES	STRUCTURE ÉTROITE	STRUCTURE LARGE
tu vois moi j'crois	*ligateur* *point de vue* *modus dissocié*	PRÉAMBULE
qu'c'est pas comme ça qu'ça doit marcher		RHÈME
la société		POSTRHÈME

TAB. 2.1 – Exemple de découpage du français en paragraphes

2.1.2 Application de ce modèle à la description de l'anglais

Le modèle présenté ci-dessus a constitué le point de départ de nos recherches et de nos hypothèses. Cependant, il a été établi pour le français, et son adaptation à l'anglais oral spontané nous oblige à nous positionner parfois en légère rupture par rapport à lui. Une telle adaptation suppose la prise en compte d'un fonctionnement morphosyntaxique différent. La langue anglaise possède une morphologie beaucoup moins développée que celle du français. A l'oral, elle reste plus compacte, c'est-à-dire plus proche de l'ordre canonique des constituants : SVO. En outre, si les valeurs iconiques des paramètres intonatifs ont vocation universelle, ce n'est pas le cas des valeurs conventionnelles qu'il faut déterminer pour chaque langue, en fonction des caractéristiques morphosyntaxiques de celle-ci. La principale nouveauté apportée par nos recherches sur l'application du modèle de M.-A. Morel et L. Danon-Boileau à la description de l'anglais réside dans la démonstration de l'existence de différents niveaux de pertinence pour les paramètres intonatifs, et d'interprétations différentes en fonction de ces niveaux au sein de la coénonciation et de la colocution, deux notions dont nous proposons de redéfinir l'extension. Deux questions se posent donc à présent. Tout d'abord, dans quelle mesure le modèle de structuration de l'oral en paragraphes convient-il à l'anglais ? Ensuite, quelle est la valeur conventionnelle des paramètres

intonatifs en anglais ?

2.1.2.1 Le paragraphe oral en anglais

Si le français oral fonctionne selon le schéma prototypique exposé ci-dessus, l'anglais oral présente certains écarts mis en lumière par de récents travaux — notamment ceux de G. Jouët-Pastré[8] et de J. Szlamowicz,[9] ainsi que nos propres recherches.[10] Détaillons ici les similitudes et les différences de fonctionnement du préambule, du rhème et du postrhème en anglais, en faisant l'inventaire de leurs éléments constitutifs.

Le **préambule**, sur lequel Jean Szlamowicz s'est particulièrement penché, existe en anglais mais ne présente pas les mêmes caractéristiques qu'en français. Il est plus court, et se réduit souvent à un ligateur. Dans les deux langues, les ligateurs fonctionnent comme des éléments de cohésion et d'homogénéisation du discours en précisant le lien existant entre ce qui suit et ce qui précède. Contrairement au français, cependant, on les trouve en anglais aussi bien en fin qu'en début de paragraphe. Selon la typologie établie par M.-A. Morel et L. Danon-Boileau, deux sortes de ligateurs coexistent : les ligateurs énonciatifs (*yes, yeah, ok, right, no, well, oh*) et les ligateurs discursifs (*and, but, so, because*). On trouve dans notre corpus des exemples de ligateurs énonciatifs en début de paragraphe :

J **well** I mean *the second subject* second subject I had yesterday I forgot to turn the tape recorder on so <rires>

(Celia&James : Experiments)

[8]JOUËT-PASTRÉ, Gabrielle (1998). Le paragraphe anglais (britannique). in Morel, Mary-Annick (resp.) (1998). Intonation, oral spontané. Comparaison de langues. in CARON, B. (éd.) (1998). *CIL 16*. Actes du Congrès International des Linguistes. Paris, 1997.

[9]SZLAMOWICZ, Jean (1997). Etude des indices segmentaux et suprasegmentaux de structuration de l'anglais oral spontané. Mémoire de DEA, Université Paris III – Sorbonne Nouvelle.
SZLAMOWICZ, Jean (2001). Contribution à une approche intonative et énonciative du rôle des ligateurs dans la construction du discours en anglais oral spontané. Thèse de doctorat, Université Paris III – Sorbonne Nouvelle.

[10]PASSOT, Frédérique (1999). Accent de mot, accent de phrase : Aspects de l'accentuation et de l'intonation de l'anglais oral spontané. Mémoire de DEA, Université Paris III – Sorbonne Nouvelle.

de ligateurs discursifs en début de paragraphe :

C *<h> **but** we had that* years ago

(Celia&James : Photo)

ou de ligateurs discursifs en fin de paragraphe :

J *well I:: <rires> <h> {} it is <h> {} **because***

(Celia&James : Languages)

La notation du point de vue est beaucoup moins fréquente en anglais qu'en français et se limite la plupart du temps à des expressions du type *I mean*. Elle sert à préciser l'identité de l'énonciateur, lequel endosse la responsabilité de ce qui suit. Il peut s'agir du point de vue du locuteur-énonciateur :

H *I can't cope with going out* *tonight at all*
R *no no no no* that's fine
H {} really <rires>
R well I don't particularly want to
I mean I want to meet them cos they're really *nice people*
H *yeah I'd like to meet them*
but just not tonight

(Helen&Rebecca : Messages)

On trouve également des expressions telles que *according to X* ou du discours rapporté. Les deux exemples donnent le point de vue d'un énonciateur rapporté par le locuteur :

J **they** {} **they** {} **they say** it was a dictatorship

(Jonathan : Chile)

H *no::* my and **my mum said** on the phone yesterday in- that she'd being thinking about it and **she was gonna say** to me <h> don't just apply for any odd job

(Helen&Rebecca : Jobs)

Dans le modèle de M.-A. Morel et L. Danon-Boileau, le modus dissocié est le lieu de l'expression de la modalité épistémique ou appréciative. En anglais, l'existence des auxiliaires modaux cantonne le plus souvent le modus à l'intérieur du rhème. On

trouve cependant dans le préambule des adverbes de modalité influant sur la portée du rhème et opérant une prise en charge énonciative.

R and he said that they were only given a comp- really limited budget to actually help these people because it's although <h> they're **supposed** to spend so much money <h> to on customers but <h> hm **obviously** they want to all to go to profit

(Helen&Rebecca : Franck)

Le cadre établit le topos dans lequel se situe l'échange au cours du paragraphe. Il est réalisé en anglais par le biais de mises en relief intonatives : il n'est donc pas forcément lié à une place déterminée dans la syntaxe. Du point de vue de l'ordre des constituants, le cadre peut prendre la forme d'un circonstant initial, et plus rarement se situer en position de postrhème. Dans l'exemple suivant, le cadre se trouve au sein du préambule :

J there's **in Tsing Tao** there's {0,70} a place called English corner or something

(Celia&James : English Corner)

Il est également assez rare de trouver un support lexical disjoint, c'est-à-dire une structure, fréquente en français, dans laquelle se trouve répétée par coréférence la référence de l'argument qui sert de support à la prédication du rhème. M.-A. Morel et L. Danon-Boileau distinguent d'une part les structures à présentatif existentiel du type *il y a X* suivi d'un pronom relatif, que l'on trouve en anglais aussi sous la forme *you have X* suivi d'un pronom relatif, et d'autre part les structures avec reprise pronominale beaucoup plus rares en anglais qu'en français, mais moins qu'on ne le dit souvent. Certains supports lexicaux disjoints sont introduits par un présentatif existentiel :

J there's in Tsing Tao there's {0,70} **a place called English corner or something where** <h> {1,50} well when my sister and her colleagues first arrived they were encouraged to go **there** but it {0,50} appears just to be a place and Chinese people sort of congregate to see Westerners

(Celia&James : English Corner)

d'autres par un présentatif en *have* ouvrant une structure existentielle :

N yeah and he's the guitarist Matthew's {0,8} a guitarist and keyboard player {1,2} **we have** another chap **who** plays the bass another historian {0,4} hm <h> and I do a little bit on the drums

(Nick&Matthew1 : Group)

Le support lexical disjoint peut se situer en début de paragraphe :

M e: **all my f all my friends in North we** sound like we're from Leeds

(Nick&Matthew2 : North)

comme en fin de paragraphe :

M he's still he's not going out with that bird anymore e is he she gave him the boot Kath

(Nick&Matthew : Wedding)

En somme, le préambule anglais est capable d'accueillir la plupart des constituants du préambule français. Signalons ici cependant qu'il se limite le plus souvent, comme l'a montré Jean Szlamowicz,[11] au seul ligateur *and*.

Le poids du **postrhème**, bien plus étoffé en anglais qu'en français, constitue une importante différence entre les deux langues. Il semble qu'une logique de rééquilibrage dans la distribution des constituants discursifs soit à l'œuvre au sein du paragraphe oral anglais par rapport au français. En dehors des nombreux ponctuants de type *you know* ou *I mean*, le postrhème anglais, qu'il faudra donc redéfinir précisément pour cette langue, peut accueillir des propositions subordonnées et coordonnées. Constituant par là un prolongement du rhème, il acquiert une nature quasi rhématique et un statut ambigu qui soulève encore une fois des questions relatives au découpage du matériau oral. M.-A. Morel et L. Danon-Boileau font du postrhème une marque infaillible de clôture de paragraphe oral. Ceci semble ici remis en cause pour l'anglais dans la mesure où le postrhème présente encore des éléments rhématiques. L'orientation des deux langues semble donc contraire. Le français ne parvient à la formulation du rhème qu'au terme d'un ou plusieurs préambules. L'anglais, au contraire, donne immédiatement le rhème, quitte à y apporter des modifications ultérieures dans un

[11] SZLAMOWICZ, Jean (2001). *op. cit.*

ou plusieurs constituants supplémentaires à valeur rhématique.[12] Cette configuration est envisagée par les deux auteurs sous le nom de « configuration noyau-catalyse »[13] : elle est pour eux l'effet d'une stratégie intonative particulière en français. Il nous faut déterminer dans quelle mesure elle peut constituer le modèle de fonctionnement du paragraphe oral en anglais.

Par son orientation postrhématique, l'anglais oral spontané se rapproche du thaï, langue étudiée par R. Gsell et K. Tinothai :

> A series of rhemes, put together in the same segment by prosodic cues [...] constitutes a larger unit of paragraph, an Overall Rheme. The series of rhemes can be predicates sharing the same subject in the first constituent, preamble, as well as a series of successive added informations about the topic mentioned in the preamble. [...] Every rheme can be interpreted as a topic for the following rheme: e.g. rheme 1 = topic of rheme 2, rheme 2 = topic of rheme 3.[14]

Il est possible d'expliquer ces phénomènes par l'intégrité de la structure sujet-verbe-objet en anglais (d'où la quasi absence de support lexical disjoint) et par l'existence d'auxiliaires de modalité intégrés au rhème (d'où la quasi absence de modus dissocié en préambule).

Dans un travail préalable, nous avons montré qu'une telle hyperorganisation est perceptible en anglais. Celle-ci fonctionne, comme l'« overall rheme » du thaï, sur le principe de recatégorisation de ses éléments constitutifs : un rhème est recatégorisé comme préambule du rhème suivant qui a lui-même un rôle de préambule vis-à-vis du rhème suivant.[15] Une articulation hiérarchique en macro- et hyperparagraphes a ainsi été démontrée sur des bases intonatives (relatives à la continuité ou à la rupture

[12] Ceci remet donc en question la définition même du postrhème énoncée par les deux auteurs pour le français.

[13] MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *op. cit.* (p.71).

[14] GSELL, René & TINOTHAI, Kittipol (1999). On Intonation in Thai Spontaneous Discourse: Broadcasting Interview. *Actes du Congrès des Phonéticiens*, San Francisco. (p.3).

[15] PASSOT, Frédérique (1999). *op. cit* (pp.71-74).

de la ligne de déclinaison), et vérifiée sur des bases sémantiques. Ceci a néanmoins été établi à partir d'un corpus d'oral spontané qui ne relève pas de la catégorie des conversations à bâtons rompus.[16] Il nous appartient de vérifier la pertinence de cette analyse sur d'autres corpus, et en particulier sur des corpus proprement conversationnels. Inversement, il est envisageable de tenter de retrouver le modèle de structuration proposé pour le français sur des bases uniquement intonatives en le confrontant avec un découpage antérieur effectué sur des bases morpho-syntaxiques et sémantiques. Le Tableau 2.2 applique les étiquettes proposées dans la *Grammaire de l'intonation* sans tenir compte de l'intonation.

Considérons rapidement l'exemple suivant, dont la syntaxe le rend analysable dans des termes proches du français[17] :

C *and of course India {0,25} I mean* India English is there's no language problem at all {1,00}
<h> although they speak Portuguese in Goa everybody speaks English

(Celia&James : Languages)

Dans la mesure où les notions de préambule et de rhème sont fonctionnelles, il est tentant de chercher à les retrouver indépendamment de leurs traits intonatifs. C'est ce que nous proposons dans le Tableau 2.2, qui fait apparaître un niveau large de structuration.

Si cette première analyse fonctionnelle doit se doubler d'une analyse intonative, elle permet néanmoins d'affirmer qu'il semble possible conserver les trois constituants du paragraphe oral pour décrire l'anglais oral. Deux modèles d'orientation du paragraphe sont envisageables : il est hautement probable que le choix de l'orientation préambule-rhème ou noyau-catalyse dépende de la densité de la syntaxe. Il convient donc de quantifier l'importance des phénomènes de décondensation ou au contraire de densification de la syntaxe et leur influence sur une analyse en paragraphes.

[16] Il s'agissait du corpus Jonathan, présenté plus bas (voir Section 3.1.2.3, p.166).
[17] Voir le Tableau 2.1, p.114.

EXEMPLES	STRUCTURE ÉTROITE		STRUCTURE LARGE
and	*ligateur*	préambule	PRÉAMBULE
of course	*point de vue*		
India	*cadre*		
I mean	*point de vue*	préambule	
India	*cadre*		
English is	*rhème avorté (cadre)*		
there's no language problem at all		rhème	RHÈME
although	*ligateur*	préambule	PRÉAMBULE
they speak Portuguese		rhème	
in Goa	*cadre*	posthrème	
everybody speaks English		rhème	RHÈME

TAB. 2.2 – Exemple de découpage de l'anglais en paragraphes

2.1.2.2 Valeur conventionnelle des paramètres intonatifs en anglais

L'analyse de la valeur conventionnelle des paramètres intonatifs est pour nous indissociable d'une réflexion sur la portée des phénomènes concernés. En effet, les courbes font apparaître une multitude de variations intonatives d'amplitude et de durée différentes. L'examen des courbes, rendu difficile par cet enchevêtrement de signes, pose immédiatement la question de l'importance relative de la variation des valeurs. Comment interpréter un pic ponctuel d'intensité ? Que dire d'une modulation mélodique s'étendant sur un paragraphe oral entier ? Notre analyse de la valeur conventionnelle des paramètres intonatifs en l'anglais cherche à déterminer quelles variations sont linguistiquement interprétables. Il nous semble nécessaire d'y faire entrer la notion de « niveaux de pertinence », qui permet de dégager plusieurs échelles d'analyse.[18]

[18]Nous rendons partiellement compte ici du travail effectué en collaboration avec Elodie Vialleton et Jean Szlamowicz depuis plusieurs années sur la valeur conventionnelle des paramètres intonatifs en anglais. Voir par exemple : PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (2001). Les interrogatives en anglais oral spontané : intonation et énonciation. in *Anglophonia-Sigma*

Les niveaux de pertinence Nous avons démontré l'intérêt de considérer les indices en fonction de leur zone d'influence, que nous appelons « niveaux de pertinence ». Les niveaux de pertinence correspondent à des différences relatives à l'étendue des phénomènes concernés mais surtout à des différences relatives à leur portée interprétative.

Ce type d'analyse nivelée n'est pas réservée aux paramètres suprasegmentaux. Elle s'applique par exemple aux phénomènes liés à l'accord en genre et en nombre. Celui-ci se lit localement dans l'orthographe des mots de l'énoncé, chaque marque morphologique étant rendue nécessaire par la première. Au niveau de l'énoncé, les marques tissent ainsi un réseau qui rend l'incidence du genre ou du nombre générale et non plus locale. De la sorte, l'interprétation de l'énoncé s'en trouve modifiée selon le genre ou le nombre régissant l'accord de ses termes. Les phénomènes suprasegmentaux ont pour nous non seulement une étendue horizontale mais aussi une étendue verticale. Notre réflexion provient de l'intuition d'une possible mise en parallèle de deux systèmes hiérarchisés d'analyse du même matériau oral. Celle-ci permet d'établir les points de jonction existant entre segmental et suprasegmental.

Quels sont les deux systèmes en question ? Le premier provient de l'observation du segmental, et de son analyse en termes syntaxiques. Celle-ci, comme nous l'avons dans notre introduction générale, nous intéresse à partir du mot et jusqu'au niveau de l'énoncé (au moins). Dans ce système d'analyse syntaxique, on considère comme valable l'assomption selon laquelle les énoncés sont constitués de syntagmes qui sont eux-mêmes constitués de mots. C'est ce qu'illustre le schéma de gauche présenté sur la Figure 2.1. Le second système, représenté à droite sur la même figure, provient quant à lui de l'observation du suprasegmental, et du type d'interprétation que l'on peut en donner. Les données simplement phonologiques se combinent pour devenir interprétables au niveau discursif — de même du niveau discursif au niveau interac-

8 : 125-152. Toulouse : Presses Universitaires du Mirail.

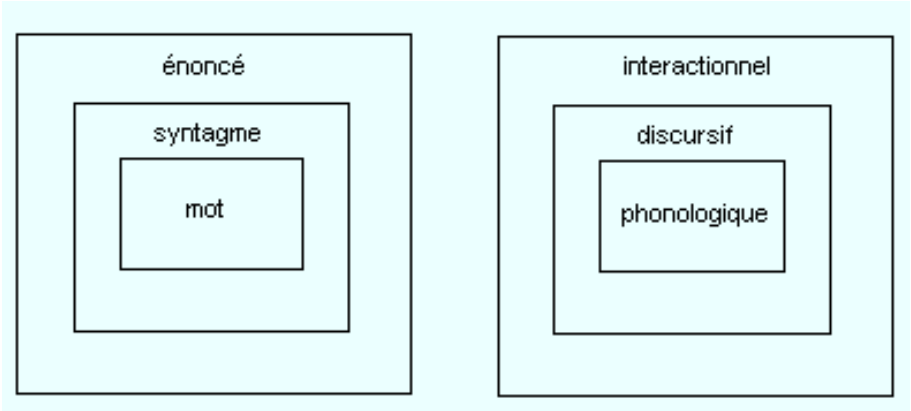

FIG. 2.1 – Deux schémas d'analyse hiérarchique parallèles

tionnel.

Or, les marques suprasegmentales affectent des éléments segmentaux situés à différents niveaux dans notre schéma selon qu'on peut en donner une interprétation phonologique, discursive ou interactionnelle. D'où un rapprochement possible des deux schémas d'analyse. La portée interprétative et la portée syntaxique des phénomènes intonatifs ont partie liée.

Ce rapprochement des niveaux segmentaux et suprasegmentaux, pour être complet, doit également se lire par rapport au temps. On admettra qu'il faut en règle générale plus de temps pour prononcer un énoncé qu'il n'en faut pour prononcer un syntagme ou un simple mot.[19] De même, une marque suprasegmentale ayant une incidence au seul niveau phonologique sera *a priori* plus ponctuelle qu'une marque susceptible d'une interprétation au niveau interactionnel. De la sorte, les trois niveaux des deux schémas d'analyse retrouvent certains parallélismes quand on les projète par rapport à l'axe du temps. La Figure 2.2 de la page 124 en propose une représentation schématique.

Sur l'axe horizontal, figurent les unités syntaxiques de l'analyse segmentale, de la plus petite à la plus grande, du contenu au contenant, du dominé au dominant.

[19]Un énoncé peut cependant n'être constitué que d'un seul mot.

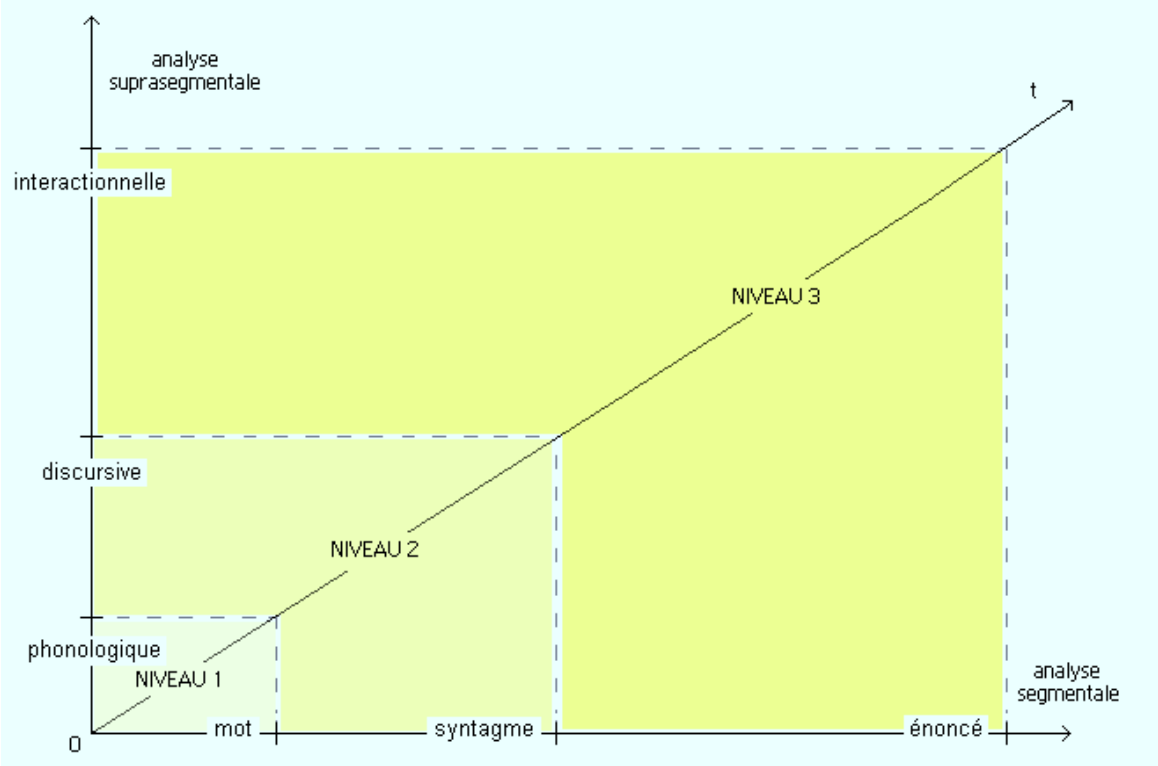

FIG. 2.2 – Correspondance des niveaux d'analyse par rapport au temps

Projetés sur un axe temporel, les éléments de l'axe horizontal sont de durée croissante. Sur l'axe vertical, celui de l'analyse suprasegmentale, sont reportés les domaines où s'interprètent les variations intonatives, du plus ponctuel au plus large, du contenu au contenant, du dominé au dominant. Projetés sur un axe temporel, les éléments de l'axe vertical correspondent à des phénomènes intonatifs de durée croissante.

Notre schéma fait apparaître ici encore une corrélation entre le niveau de pertinence des variations intonatives sur le plan interprétatif et sur le plan syntaxique. La valeur de la projection des éléments des deux axes sur l'axe temporel n'est qu'un construit théorique : la durée relative des phénomènes et leur niveau de pertinence n'est pas démontrée. Il est cependant raisonnable de postuler une corrélation entre les deux niveaux du fait de la probabilité qu'un phénomène intonatif affectant un mot et seulement un mot ne soit interprétable qu'en termes phonologiques, et qu'un phénomène intonatif interprétable au niveau interactionnel se retrouve non pas simplement au niveau des syntagmes, mais au niveau de l'énoncé.

En somme, des analyses que l'on peut formuler concernant la valeur des données suprasegmentales se dégage le même type de structure hiérarchique que des analyses syntaxiques du segmental. Ce constat nous permet de déterminer les points de contact entre segmental et suprasegmental. Cette théorie, pour séduisante qu'elle soit, doit encore être mise à l'épreuve du corpus. Il convient également d'examiner les deux remarques suivantes, qui constituent à nos yeux deux réserves importantes. Tout d'abord, il est sans doute arbitraire de rapprocher deux grilles de lecture qui n'ont pour seul point commun que de proposer une vision hiérarchisée de la constitution du même matériau. Deuxièmement, il est possible que la concordance de forme des deux grilles de lecture ne provienne pas de l'objet analysé, mais de la propension de l'esprit classique à adopter une approche tripartite quel que soit son objet. En outre, la correspondance entre niveaux d'analyse empruntés aux deux systèmes n'est pas stricte. Un aménagement est nécessaire pour passer de l'intuition à son ap-

plication. En effet, la syllabe semble davantage être le domaine de pertinence des phénomènes phonologiques que le mot, qui est quant à lui autant susceptible que le syntagme de bénéficier d'une interprétation discursive. Au dernier niveau, on parlera non seulement d'énoncés, mais également de paragraphes oraux et de tours de parole. Il reste néanmoins qu'une telle conception des fonctions de l'intonation par niveaux de pertinence nous semble éclairante dans bien des cas. Elle nous a notamment permis de dégager un invariant pour chacun des paramètres intonatifs en anglais, après analyse du fonctionnement de chacun d'entre eux aux trois niveaux de pertinence. En nous fondant sur la terminologie et les préoccupations de la théorie des opérations énonciatives, nous associons ainsi intensité et identification, mélodie et interprétation. Durée et pauses, pour lesquelles notre hypothèse est encore balbutiante, participent quant à elles des problématiques de formulation et segmentation.

Intensité : identification Les variations d'intensité nous semblent interprétables à trois niveaux différents, qui correspondent à des échelles d'analyse différentes.[20]

Notre premier niveau est celui du mot, voire de la syllabe. Pour des raisons physiologiques, les indices intonatifs s'entremêlent souvent. Ce phénomène se retrouve dans la réalisation de l'accent lexical en discours, de sorte que l'on peut observer des variations de la qualité de la voyelle, de la fréquence fondamentale et de la durée des phonèmes conjointement à une variation d'intensité. Au terme d'études préalables,[21] il est apparu qu'un mot donné se reconnaît à sa charpente intensive, articulée autour de sa syllabe accentuée, marquée par un surcroît d'intensité. C'est en ce sens que nous attribuons une fonction morphologique à l'intensité. Ainsi, phonologiquement,

[20]Remarquons que notre prise en compte de l'intensité n'est pas incompatible avec l'utilisation du micro-cravate par rapport au micro-casque toujours situé à égale distance de la bouche. En effet, notre approche se place du côté de la perception, de la réception du signal et non pas de sa production. Lorsque la tête du locuteur bouge, le volume sonore émis ne change pas, mais le volume perçu varie, ce que permet de capter un micro-cravate.

[21]Voir PASSOT, Frédérique (1999). Accent de mot, accent de phrase : Aspects de l'accentuation et de l'intonation de l'anglais oral spontané. Mémoire de DEA, Université Paris III – Sorbonne Nouvelle.

l'accent lexical permet de désigner le cœur du mot qu'il affecte. L'intensité a ainsi une fonction d'identification à son premier niveau de pertinence.

Si nous élargissons à présent notre champ de vision, nous pouvons observer des variations d'intensité de nature et d'ampleur différentes. Une double hiérarchisation s'opère par l'intensité entre les mots lexicaux et les mots grammaticaux d'une part (ces derniers, ne recevant normalement pas d'accent lexical, sont marqués par une intensité moins forte), et, au sein des mots lexicaux d'autre part, entre ceux qui sont particulièrement saillants et les autres. Nous avons montré, à l'aide de quelques exemples tirés de notre corpus, que l'intensité participe à la construction des objets de discours. Portant de façon privilégiée sur des éléments à valeur référentielle (et non sur des mots grammaticaux), l'intensité établit des relations entre les mots lexicaux, et contribue au soulignement des termes qui vont faire l'objet d'un développement thématique.

Au niveau interactionnel, les variations d'intensité concernent l'énoncé ou le paragraphe oral. Leur valeur est établie à partir de la constatation de deux phénomènes. Bon nombre de pauses au sein d'un même tour de parole sont suivies d'une plage intensive plus forte que ce les segments précédents et suivants. Nous avons donc émis l'hypothèse d'une intensité compensatrice qui unifie le tour de parole. L'intensité acquiert ici une fonction structurante en participant à la gestion des tours de parole. Ce surcroît d'intensité se vérifie également souvent en début d'intervention, dans la mesure où ce paramètre est le témoin de la revendication du droit à la parole. En cas de lutte pour la parole, c'est celui qui maintient une intensité haute qui garde la parole. Une baisse d'intensité sans marques d'hésitation équivaut à une cession de parole.

En somme, à ses trois niveaux de pertinence, l'intensité stabilise des éléments linguistiques et interactionnels sans préjuger de leur valeur intersubjective. Elle contribue à l'identification du cœur du mot, du thème de l'énoncé (ou du paragraphe) et du

locuteur principale.

Mélodie : interprétation Le même type d'analyse en trois niveaux s'applique selon nous aux variations mélodiques. Au niveau de l'analyse phonologique, on constate que les mouvements mélodiques s'effectuent de façon privilégiée sur les syllabes portant une accentuation lexicale, conjointement à un pic d'intensité. Ce phénomène est d'autant plus sensible dans les énoncés de laboratoire que la lecture (où l'on met le ton) a tendance à exagérer le paramètre mélodique au détriment des autres. D'un point de vue discursif, il semble que les segments marqués par une modulation mélodique bénéficient d'une mise en relief polémique, sont investis d'une prise de position. Au niveau supérieur, les mouvements du fondamental intervenant sur la fin d'un énoncé ou d'un paragraphe oral renseignent sur son statut intersubjectif. Ainsi, si l'on considère que la ligne mélodique en anglais est globalement descendante (en vertu de la déclinaison naturelle ou *downdrift*), une remontée mélodique en finale permet d'interpréter ce qui précède non comme un tout, mais comme une partie dont l'énonciateur annonce qu'il appelle une suite. S'il dispose lui-même des éléments pour compléter ce qu'il a avancé, la valeur de la remontée mélodique est continuative. S'il ne dispose pas (ou feint de ne pas disposer de ces éléments), la valeur de la remontée mélodique est une valeur d'appel, dont une réalisation courante est l'interrogation.

Durée : formulation et segmentation L'application à ce paramètres de notre analyse en niveaux de pertinence doit encore être approfondie. En effet, la durée concerne des phénomènes aussi différents que les allongements et les pauses.

Comme nous l'avons vu plus haut, les pauses peuvent être de deux types. Les pauses vides sont des moments de silence, de longueurs diverses, pendant lesquels le canal sonore est inoccupé. Les pauses pleines, en revanche, correspondent à des moments où le canal sonore est rempli par des particules dites d'« hésitation », le plus

souvent allongées. Les pauses vides ont bénéficié de davantage d'attention que les pauses pleines. Parmi celles dont la durée (supérieure à 0,20 s) les rend interprétables linguistiquement, Jean Slamowicz distingue ainsi trois fonctions différentes attachées aux pauses vides : la segmentation, la focalisation et la formulation.[22] Cette typologie se fonde sur la constatation de différences de longueur, de site d'occurrence, et d'indices supplémentaires auxquels les pauses sont associées. Après confrontation des données de ce chercheur avec notre corpus dans la suite de ce travail, nous pourrons envisager un rattachement à notre hypothèse théorique des niveaux de pertinence. Nous pouvons néanmoins avancer quelques hypothèses.

Au premier niveau se situe la pause respiratoire, dont la durée se situe autour de 0,20 s. Elle n'est généralement pas perçue et est à ce titre écartée des analyses linguistiques. Cependant, son site d'occurrence n'est pas anodin : il garantit l'intégrité de l'unité mot, ce qui n'est pas le cas du syntagme. Cette caractéristique est propre à l'anglais, quelles que soient la longueur et la fonction des pauses. Les cas de pauses entre les syllabes d'un même mot sont extrêmement rares, et méritent la plupart du temps d'être réévaluées en fonction du débit du locuteur dans le segment concerné.[23] Notre corpus donne plusieurs exemples d'auto-corrections engendrant une courte interruption du signal sonore inférieure au palier que nous avons choisi pour le repérage des pauses.[24] C'est le cas de l'exemple suivant où l'interruption de 0,10 s n'est pas l'occasion pour Celia de respirer, mais est probablement due au fait que la locutrice hésite à continuer de parler sans corriger la syllabe *-tro-* qu'elle n'a pas prononcée très clairement :

C hm {0,45} <h> I can {0,75} I can **introd-** {**0,10**} **duce** you to the alarm system

[22]SZLAMOWICZ, Jean (2002). Les pauses en anglais : de la faillite du silence à la structuration linguistique, ou de l'iconique au conventionnel. in Delmas, Claude & Roux, Louis (éds.), *Correct, incorrect en linguistique anglaise. CIEREC Travaux* 113 : 157-174. Saint-Etienne : Publications de l'Université de Saint-Etienne.

[23]Voir discussion sur le recours aux outils informatiques, à l'intuition linguistique et à la perception, p.184.

[24]En conséquence, elles n'apparaissent pas dans les transcriptions du Volume 2.

(Celia&James : Keys)

En revanche, au deuxième niveau, nombreuses sont les pauses intervenant en milieu de syntagme. Associées ou non à des marques d'hésitation, elles en prennent cependant la valeur et sont le témoin d'un travail de formulation. Lorsque le travail de formulation est feint, il correspond à une stratégie de focalisation, permettant à ce qui suit de se détacher avec plus de force. Enfin, au dernier niveau de pertinence, la valeur interactionnelle de la pause vide en anglais est liée au fait qu'un silence autonomise les segments qu'elle sépare. Elle peut donc constituer une charnière entre des constituants de grande taille, ou signifier une cession du tour de parole.

L'allongement relève, nous semble-t-il, d'une problématique très proche de celle des pauses vides. Au niveau phonologique, on remarque que le calcul du débit de parole se fonde sur le nombre de syllabes prononcé en un temps donné. La syllabe, et en particulier la voyelle qu'elle contient, est le lieu privilégié de la variation de durée. La corrélation des paramètres lors de la réalisation de l'accentuation lexicale touche également la durée : les syllabes accentuées ont tendance à être plus longues que les syllabes non accentuées. Cela dit, les variations de durée peuvent apparaître indépendamment de toute considération accentuelle. Elles ne sont donc plus interprétables en termes phonologiques, mais discursifs ou interactionnels. Au deuxième niveau, l'allongement participe soit d'une mise en relief, soit d'un problème de formulation (en particulier lorsqu'il s'accompagne de pauses vides ou pleines ou d'autres marques d'hésitation). Témoin de l'état de formulation de ce que le locuteur a à dire, la durée intervient également au niveau interactionnel où un allongement d'hésitation permet d'occuper le canal sonore pour prévenir toute prise de parole de l'interlocuteur. Il peut également, sans pour autant relever de l'hésitation, accompagner une mélodie continuative en vue de la conservation du tour de parole.

Il est difficile de dégager un invariant de ces diverses caractéristiques fonctionnelles. Sans doute faut-il préférer une analyse permettant d'embrasser cette variété

(notamment en analysant les conjonctions de paramètres) à une analyse cherchant à réduire cette richesse à une seule valeur. Les paramètres apparaissent rarement seuls en discours. Il faut tenir compte dans nos analyses de l'apport propre de chacun, dans la mesure où il est dissociable du tout. Nous proposons néanmoins le Tableau 2.3, qui résume nos hypothèses relatives à l'invariant linguistique de chacun des trois niveaux de pertinence, et les rapproche des éléments du segmental concernés et des paramètres suprasegmentaux.

	ANALYSE SEGMENTALE	ANALYSE SUPRASEGMENTALE			
		I	**F_0**	**D**	**P**
Niveau 1	syllabe	accentuation	accentuation	accentuation	respiration
Niveau 2	mot, syntagme	mise en relief (thématique)	mise en relief (polémique)	formulation, mise en relief	formulation, mise en relief
Niveau 3	énoncé, paragraphe, tour de parole	gestion du tour de parole	gestion du tour de parole	gestion du tour de parole	gestion du tour de parole, segmentation

TAB. 2.3 – Niveaux de pertinence

Il nous semble que le Niveau 1 de l'analyse suprasegmentale est largement celui de l'accentuation lexicale. Au Niveau 2, les indices permettent de mettre en relief un élément lexical en vue d'un développement thématique ou polémique, ou témoignent de difficultés de formulation, feintes (et donc relevant de la problématique de la mise en relief) ou réelles. Au Niveau 3, les indices sont interprétés en fonction de leur participation à la gestion des tours de parole (cession ou conservation).

2.2 Présentation des hypothèses de départ

L'objectif de notre recherche est de montrer que la langue orale spontanée possède des fonctionnements propres, qui s'illustrent de façon privilégiée lorsque l'on s'intéresse aux opérations de repérage d'un constituant discursif par rapport à un autre. Nous formulons ici quelques hypothèses concernant ce fonctionnement qu'il nous faudra confronter aux données de notre corpus.

2.2.1 Segmental et suprasegmental : une syntaxe autre

L'une des hypothèses que nous chercherons à vérifier est celle de la désyntactisation de l'oral par rapport à l'écrit, et de la compensation de ce phénomène par des marques suprasegmentales. Le terme de désyntactisation a pour nous une double signification : nous pensons ici à la fois à une syntaxe décondensée à la française, et à une moindre variété des formes syntaxiques à l'oral par rapport à l'écrit. Il semble que si un même message peut prendre des formes très différentes en ce qui concerne l'ordre des mots et les marques segmentales de hiérarchisation dans la chaîne linéaire, celui-ci soit néanmoins structuré, en profondeur, par les mêmes opérations. La théorie de l'énonciation, qui postule l'existence d'un niveau abstrait d'opérations sous-jacentes à l'activité de langage, permet précisément de considérer une distribution de marqueurs différents pour une même opération. Si l'on peut parler de désyntactisation, celle-ci

n'a évidemment pas la même ampleur en anglais qu'en français. Formuler l'hypothèse d'une syntaxe de l'oral ne peut se comprendre qu'à un niveau global. Il ne s'agit pas de tenter de montrer que la syntaxe de l'anglais écrit et la syntaxe de l'anglais oral sont totalement étrangères l'une à l'autre.

On l'a dit, cependant, les cas de syntaxe décondensée sur le modèle français existent en anglais. En outre, si l'on se départit beaucoup plus rarement du schéma canonique, une forme de désyntactisation intervient dans la mesure où les liens logiques font moins appel à la syntaxe qu'à l'écrit. La logique prend donc d'autres relais, tels que la sémantique et l'intonation. Nous serons amenée à évaluer la proportion de mots lexicaux dans notre corpus et à constater que si les mots outils sont une catégorie particulièrement bien représentée, ils sont relativement peu variés. Les marques segmentales de structuration sont donc polysémiques et moins spécifiantes du point de vue du lien qu'elles établissent entre les éléments qu'elles relient. Une syntaxe autre est à l'œuvre, tant au niveau segmental qu'au niveau suprasegmental.

Quelles sont donc les marques des opérations de hiérarchisation à l'oral ? Notre attention se portera bien évidemment de façon privilégiée sur ce qui est spécifique à l'oral, c'est-à-dire le suprasegmental. Il s'agit en effet d'évaluer le rôle de l'intonation dans les opérations de structuration et d'organisation de la chaîne parlée. Dans le cadre restreint de ce travail, il sera nécessaire de ne sélectionner qu'un petit nombre de faits de langue relatifs à ces questions de structuration.

Si l'on choisit par exemple le cas de la subordination et de la coordination, il nous faudra répertorier les types de relation qu'instaurent la subordination et la coordination selon la grammaire traditionnelle (établie pour l'écrit et à partir de l'écrit). Une étude distributionnelle confirmera qu'un type de relation n'est pas réservé à une forme, et que le vide formel ne correspond pas forcément à un vide sémantique et syntaxique. La preuve en est donnée par exemple par le relatif zéro qui, à l'oral comme à l'écrit, ne remet pas en cause le repérage de la subordonnée et de son antécédent. A

l'oral, quatre types de configurations devront être étudiés parmi les énoncés que l'on aura retenus :

1. L'énoncé présente des marques syntaxiques répertoriées par la grammaire traditionnelle. L'intonation est-elle marquée également ?
2. L'énoncé présente peu de marques syntaxiques. L'ordre canonique des mots est-il modifié ? Quelles marques intonatives le caractérisent ?
3. L'énoncé présente des marques intonatives particulières. Sur quels éléments portent-elles ? Quel effet ont-elles ?
4. L'énoncé ne présente pas de marques intonatives particulières. Que comprend-on de la relation entre les éléments ?

En somme, on cherchera toujours à déterminer si les indications données par la syntaxe et l'intonation sont contradictoires, redondantes ou encore complémentaires.

Plus généralement, il semble que l'intonation soit l'exact contrepoint de la syntaxe visible. Elle constitue une autre syntaxe — la syntaxe du suprasegmental. Ces deux systèmes ont bien évidemment des fonctionnements différents, mais ils n'en concernent pas moins le même matériau : le discours. Nous les concevons comme deux couches superposées, orientées selon la flèche du temps. Si deux droites parallèles ne se rencontrent jamais, il est raisonnable de postuler, comme l'illustre la Figure 2.3, que syntaxe et intonation admettent des points de contact. L'un des objectifs de ce travail est de déterminer ces points afin d'analyser l'articulation entre les deux syntaxes.

2.2.2 Fonctions et significations de l'intonation

La forme même des relations syntaxiques leur confère un invariant, un sens rémanent. *Because* introduit une subordonnée circonstancielle de cause, tandis que *if* indique une condition ou une concession. N'étant pas relayée par du segmental, et donc du sémantisme, l'intonation n'est cependant pas exempte de sens. Indépendamment

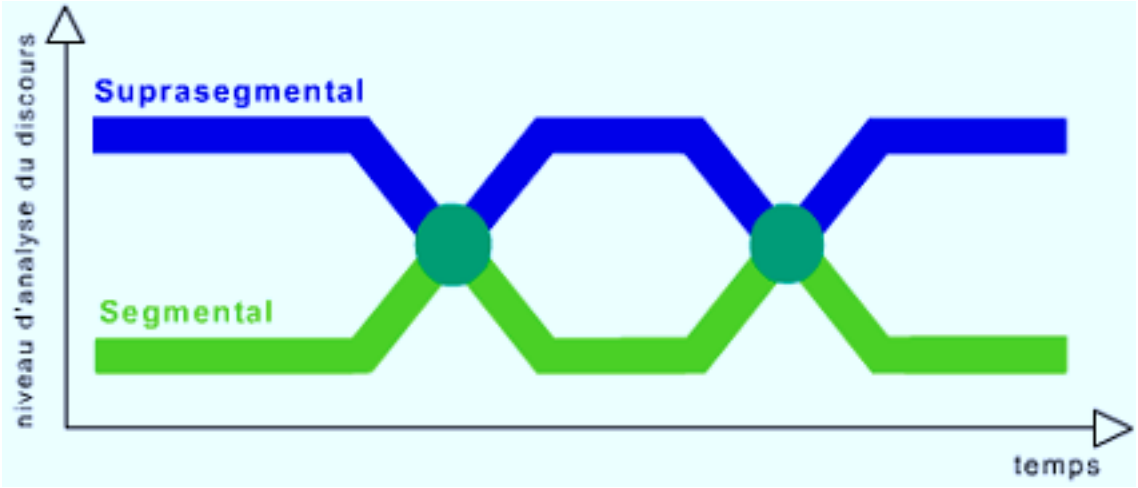

FIG. 2.3 – Points de contact entre suprasegmental et segmental

de l'identification des paramètres suprasegmentaux, les variations intonatives ont une réalité physique qui les inscrit, le long d'une ligne temporelle, sur un axe haut/bas. Intuitivement, il ne semble pas déraisonnable de penser que cet axe est lui-même signifiant de façon universelle, et cela en dehors de tout système linguistique. Les recherches de J. Ohala sur la hauteur sonore vont d'ailleurs dans ce sens. Il montre en effet que la hauteur répond à un code que l'on retrouve aussi bien dans une société animale comme celle des singes que dans les sociétés humaines. A travers ce code non linguistique, attitude et rôle social ont partie liée : les hauteurs basses correspondent à un rôle d'autorité, tandis que les hauteurs élevées sont davantage interprétées comme un signe de soumission.

> [...] the frequency code underlies the shape of human vocalizations [...]: low F_0 to signal threat or selfconfidence, high F_0 to signal non threat or deference.[25]

C. Gussenhoven approfondit cette hypothèse des codes biologiques en spécifiant, en plus du code fréquentiel (relatif à la hauteur absolue), un code de l'effort (*effort*

[25] OHALA, John J. (1994). The frequency code underlies the sound symbolic use of voice pitch. In Hinton, L., Nichols, J., & Ohala, John J. (éds.). *Sound symbolism.* Cambridge : Cambridge University Press. 325-347.

code) et un code de la production (*production code*). Le premier est lié au surcroît d'énergie nécessaire pour produire un pic mélodique, c'est-à-dire à l'amplitude des mouvements intonatifs. Deux types d'interprétations linguistiques sont possibles : sur le plan informationnel, l'emphase, et sur le plan des affects, la surprise :

> The Effort Code is based on the fact that greater articulatory effort tends to create more elaborate and more explicit phonetic realisations. In the case of fundamental frequency variation, greater explicitness leads to more canonical, wider pitch movements, less explicitness to slurred, narrow-range movement. As a result, a wider pitch span is associated with meanings that can be derived from speakers' motivations for the expenditure of articulatory effort. [...] The informational interpretation [...] of the Effort Code is 'emphatic', derived from the perception that the speaker regards his or her message as important and thus spends more energy on its production. An affective interpretation is 'suprised', derived from the perception that the speaker shows agitation thereby spending more effort on speech production.[26]

Le code de la production rend compte du phénomène de la diminution de la pression sous-glottique avec le temps, ce qui se traduit par la ligne de déclinaison. D. House décrit l'interprétation linguistique du phénomène dans les termes suivants :

> [...] the production code [...] associates high pitch with phrase beginnings (new topics) and low pitch with phrase endings.[27]

On note que les mêmes codes s'appliquent la plupart du temps aux variations d'intensité. En valeur absolue, une intensité haute est signe de volonté de domination, contrairement à une intensité basse. Conformément au code de l'effort, un surcroît d'énergie est également nécessaire pour produire un pic d'intensité. En ce qui concerne

[26]CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002 conference* : 215-218. (p.215).

[27]HOUSE, David (2003). Hesitation and Interrogative Swedish Intonation. *PHONUM* 9 : 185-188. (p.185).

la production, on constate une même tendance de la voix à produire un volume plus faible à mesure que l'énoncé se déroule.

Au sein d'un système linguistique, comme nous l'avons vu plus haut, les paramètres intonatifs se partagent des valeurs iconiques universelles et des valeurs conventionnelles propres à une langue. L'intonation ouvre des possibles parmi lesquels le système d'une langue particulière ou, simplement, le locuteur de cette langue, sélectionnent ce qui est susceptible de servir ses besoins langagiers. Dans le cadre de cette étude sur la structuration du discours à l'oral, nous chercherons à éclairer ce que la langue anglaise fait des possibles ouverts par l'intonation et ses codes universels.

A cet effet, nous approfondirons notre hypothèse théorique déjà évoquée plus haut des fonctions des paramètres intonatifs réparties selon leur niveau de pertinence.[28] Nous nous intéresserons en particulier aux Niveaux 2 et 3, qui concernent la construction du discours et de l'interaction. Sur le plan mélodique, une modulation correspond localement à une mise en relief polémique ou un soulignement des mots ou segments que le locuteur souhaite placer au cœur du débat, et qu'il investit d'une prise de position par rapport à la représentation qu'il prête à son interlocuteur. La mélodie, envisagée au Niveau 2, renseigne donc sur le *modus*. Plus globalement, une modulation mélodique constatée au Niveau 3 indique le statut intersubjectif du tour de parole. L'intensité intervient au Niveau 2 pour la sélection des cadres thématiques. Elle renseigne donc de façon privilégiée sur le *dictum*. Au Niveau 3, elle joue un rôle important dans la conservation ou la cession des tours de parole. Nous nous attacherons également à déterminer la fonction des allongements de phonèmes et plus particulièrement des pauses silencieuses.

Il s'agit donc d'évaluer dans quelle mesure les différents paramètres intonatifs contribuent à l'édification d'une syntaxe de l'oral dans le cadre de l'appauvrissement syntaxique ou de la re-syntactisation du tissu segmental.

[28] Voir p.121, et en particulier le Tableau 2.3, p.132.

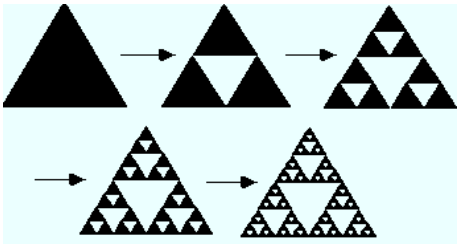

FIG. 2.4 – Le triangle de Sierpinski

2.2.3 Système de l'oral et fractales

Notre troisième hypothèse de départ concerne la forme de notre objet d'étude en tant que tout fait de parties, et la façon dont nous souhaitons l'analyser. Elle prend pour modèle une théorie mathématique relativement récente.[29] La théorie des objets fractals, initiée par J. Perrin, a été formulée par le mathématicien B. Mandelbrot[30] pour décrire la forme d'objets aussi complexes que la planète Terre, l'océan et le ciel.

> Bien que leur étude appartienne à des sciences différentes, entre autres la géomorphologie, l'astronomie et la théorie de la turbulence, les objets en question ont en commun d'être de forme extrêmement irrégulière ou interrompue ; la notion qui servira à unifier leur étude sera désignée par l'un ou l'autre des deux néologismes (synonymes), « objet fractal » et, quelquefois, « fractum », tous deux formés, pour les besoins de ce livre, à partir de l'adjectif latin « fractus », un dérivé de « frangere », briser, qui a précisément, entre autres significations, celle d'« interrompu ou irrégulier », et qui d'ailleurs est déjà utilisé pour désigner des nuages contenant des petits éléments de forme accidentée.[31]

L'hypothèse est séduisante de considérer l'oral spontané, éminemment irrégulier et interrompu, comme un objet fractal. Dans un tel objet, toute partie est identique

[29] Les considérations présentées ici ne font que s'inspirer de la théorie mathématique et ne prétendent pas l'appliquer de façon rigoureuse.

[30] MANDELBROT, Benoît (1975). *Les objets fractals : forme, hasard et dimension.* Paris : Flammarion.

[31] *ibid.* (p.1).

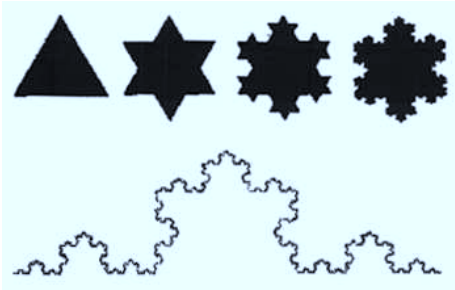

FIG. 2.5 – Le flocon de neige de Von Koch

au tout, comme l'illustrent les Figures 2.4 et 2.5,[32] où le même motif (le triangle) se répète pour former un motif plus grand.

C'est en quelque sorte le pari de la grammaire générative, qui figure sous forme d'arbre ou de parenthèses gigognes les relations syntaxiques existant entre les constituants. On remarquera à ce propos que l'arbre est régulièrement cité comme exemple dans les ouvrages mathématiques traitant des objets fractals, tant théoriques que naturels.[33] Citons ici l'introduction de l'ouvrage de L. Haegeman et J. Guéron :

> We assume that the sentence and all the units which constitute the sentence, i.e. its constituents, have the same internal organization. We adopt the X-bar theory of phrase structure which stipulates that all syntactic constituents are organized around a head, X. The head, X, is a simple syntactic constituent drawn from the lexicon and consisting of any word or morpheme category. X is expanded by the addition of a complement to form a larger unit, X' (X-bar), which in turn combines with a specifier to form XP, the maximal projection of X. The complement ZP and the specifier YP [...] have an internal structure identical to that of XP itself [...].[34]

[32]Les images reproduites ici sont empruntées à R.L. Devaney (Figure 2.4) et à L.D. Bradley (Figure 2.5). Elles proviennent de leurs sites internet, consultés les 11 et 12 janvier 2003 : http://math.bu.edu/DYSYS/chaos-game/node2.html pour le premier et http://www.pha.jhu.edu/~ldb/seminar/fractals.html pour le second.

[33]Voir notamment BARNSLEY, Michael F. (1988). *Fractals everywhere.* San Diego, CA : Academic Press. (pp.41, 94).

[34]HAEGEMAN, Liliane & GUERON, Jacqueline (1999). *English Grammar: a Generative Per-*

Ce travail ne s'inscrit pas dans le cadre de la théorie générative et transformationnelle, qui ne prend en compte que le fait syntaxique. Nous avons l'ambition d'apporter des éléments permettant d'avancer vers l'explication du système de l'anglais oral dans son entier. Pour ce faire, il nous semble essentiel de prendre en compte toutes les dimensions de construction et d'organisation à l'œuvre dans ce système. Dans son ouvrage fondateur, B. Mandelbrot fait référence à la lentille du microscope, qui permet d'appréhender le même objet à des distances différentes. Les dimensions phonologiques, thématiques et interactionnelles qui nous intéressent constituent autant d'échelles d'analyse pour un même phénomène. Comme nous l'avons montré plus haut, notre approche est fondée par exemple sur la reconnaissance de niveaux de pertinence des phénomènes intonatifs.

Penser l'oral en tant qu'objet fractal rejoint également la conception guillaumienne de la langue déjà évoquée plus haut. Le « système de systèmes » postule en effet que la forme générale de la langue se redit plusieurs fois en elle-même. Cette caractéristique itérative de la langue en fait un objet fractal :

> [...] la langue est un système de systèmes — un assemblage systématisé de systèmes contenants (ayant un contenu propre de positions intérieures) s'emboîtant les uns dans les autres et qui, inscrits chacun dans un plus étendu, le plus étendu de tous étant celui de l'assemblage qu'en fait la langue, diffèrent entre eux sous toutes sortes de rapports, sauf celui de leur forme commune de contenant, laquelle se répète identique à elle-même, et en réalité invariante, du plus étendu au moins étendu, de sorte que celle du plus étendu, la langue, assemblage de tous, serait connue au cas où l'on réussirait à voir en traits nets la forme de l'un de ceux, riche ou pauvre de substance, qu'elle contient.[35]

spective. Oxford : Blackwell. (p.2).

[35]GUILLAUME, Gustave (1952). La Langue est-elle ou n'est-elle pas un Système ? *Cahiers de Linguistique Structurale* 1 : 3-30. Québec : Presses Universitaires Laval. (p.7).

Il s'agira donc pour nous de déterminer dans quelle mesure cette caractéristique itérative de la langue se vérifie. Ici encore, on voit la nécessité de combiner les niveaux de recherche dans une perspective à la fois de micro- et de macro-analyse.

Ce que B. Mandelbrot appelle le degré de résolution est un point essentiel de sa théorie. Il pose notamment la question de l'objectif du chercheur et de la constitution de ses outils d'analyse. Au sein d'un ensemble circonscrit se dessine une infinie complexité à mesure que l'on agrandit le degré de résolution, c'est-à-dire que l'on resserre le champ.

> [...] montrons qu'un objet plus complexe [qu'un fil, un voile ou une petite boule], à savoir une pelote de 10 cm de diamètre, faite de fil de 1 mm de diamètre, possède, de façon en quelque sorte latente, plusieurs dimensions physiques distinctes. Au degré de résolution de 10 mètres, elle apparaît comme un point, donc comme une figure zéro-dimensionnelle ; au degré de résolution de 10 cm, c'est une boule, donc une figure tridimensionnelle ; au degré de résolution de 10 mm, c'est un ensemble de fils, donc une figure unidimensionnelle ; au degré de résolution de 0,1 mm, chaque fil devient une sorte de colonne, et le tout redevient une figure tridimensionnelle ; au degré de résolution de 0,01 mm, chaque colonne se résoud [sic] en fibres filiformes, et le tout redevient une figure unidimensionnelle, et ainsi de suite, la valeur de la dimension ne cessant de sautiller. A un certain niveau d'analyse, la pelote se représente par un nombre fini d'atomes ponctuels, et le tout redevient zéro-dimensionnel. Qu'un résultat numérique dépende ainsi des rapports entre l'objet et l'observateur, n'a rien de nouveau ; cela est bien dans l'esprit de la physique de ce siècle, et c'en est même une illustration particulièrement exemplaire. Il en résulte même, inévitablement — selon le critère utilisé — divers observateurs risquent d'évaluer le nombre de dimensions distinctes d'un même objet de façons différentes [...].[36]

[36] MANDELBROT, Benoît (1975). *op. cit* (pp.13-14).

En somme, si la partie est semblable au tout, l'objet n'est jamais épuisé, comme l'ont éprouvé les travaux sur la longueur de la côte britannique — exemple d'objet fractal naturel. Sans préjuger du caractère fractal ou non de l'anglais oral spontané, nous retrouvons cette problématique dans la discussion concernant les unités pertinentes à notre description de l'anglais oral spontané, et ceci au moins pour deux raisons essentielles.

Fragmenter en unités constitutives nous semble en effet participer de la même démarche de mesure que celle à laquelle se livrent les mathématiciens que nous mentionnons. S'il est évident que l'on n'utilisera pas le même outil pour mesurer des valeurs de l'ordre de 10 cm ou de 0,01 mm, B. Mandelbrot précise dans l'extrait cité plus haut que le choix de l'unité de mesure conditionne le résultat. Dès lors, il nous faut déterminer le « degré de résolution » que nous recherchons afin de pouvoir choisir les unités que nous retiendrons pour l'analyse de l'anglais oral spontané, en ne perdant pas de vue le fait l'outil d'analyse prédétermine l'objet. On rejoint donc ici les préoccupations d'A. Einstein dont nous avons parlé plus haut. Le résultat d'une étude étant fonction de l'échelle à laquelle elle a été menée, on peut penser que le caractère fractal d'un objet tel que le nôtre n'apparaît pas forcément de façon immédiatement perceptible, mais qu'il s'agit d'une construction mentale qui provient *a posteriori* du regard porté sur lui par de multiples paires d'yeux. Notre objectif est de retrouver une vision unifiée de l'objet à travers ses multiples dimensions et à plusieurs échelles.

Deuxièmement, la micro-analyse du domaine de l'oral a traditionnellement suscité davantage d'intérêt de la part des chercheurs que la macro-analyse — on pense notamment à la recherche sur les parties du discours et sur le système phonologique des langues. Encore une fois, si la partie est semblable au tout, il est tentant de décrire le tout dans les termes mêmes que ceux utilisés par la syntaxe et la phonologie pour en décrire les parties. D'un point de vue méthodologique, il semble recevable de vouloir appliquer à petite échelle (grande réalité) ce qui fonctionne à grande échelle (petite

réalité). Si l'on peut envisager cette perspective, on se gardera néanmoins de tomber dans le même piège que G. Gougenheim, signalé par O. Ducrot & J.-M. Schaeffer, qui a éprouvé les limites de l'application sans aménagement des principes du fonctionalisme phonologique à la grammaire.[37]

Quelle que soit la conclusion concernant la possible application d'une telle analyse à l'anglais oral spontané tel que représenté dans notre corpus, le recours à la théorie des objets fractals ne peut être qu'éclairante dans notre perspective. Nous n'ambitionnons évidemment pas de pouvoir appliquer à la lettre la théorie mathématique à notre objet d'étude. Il nous semble cependant intéressant d'en emprunter l'esprit, dans la mesure où cette théorie permet à la fois de penser le rapport de la partie au tout, et la place de l'analyste dans ce rapport.

Ces trois hypothèses, sans avoir l'ambition d'épuiser les questions posées par l'oral spontané, sous-tendent toute notre réflexion sur son organisation. Il s'agit en effet de proposer une description détaillée du fonctionnement de la syntaxe et de l'intonation au sein du système complexe de l'anglais conversationnel. C'est en particulier à la complexité de la structure que se porte notre attention, et à l'analyse de ses différentes couches. Afin de mener notre projet à bien, il nous faut rassembler des données propices à l'analyse. Sur quelles données notre travail s'appuie-t-il ? Nous souhaitons à présent décrire les outils qui vont nous permettre de mettre ces hypothèses à l'épreuve.

[37]Voir DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Le Seuil. (pp.44-45).

Chapitre 3

Présentation du corpus utilisé

Ce dernier chapitre de notre partie concernant le cadre théorique et méthodologique de notre recherche est pour nous l'occasion de présenter les outils que nous nous donnons afin de mettre à l'épreuve nos hypothèses de départ : notre corpus constitue une base de données originale et essentielle à notre travail. Sa constitution et son traitement ont fait l'objet de toute notre attention. Ils sont le fruit d'une réflexion méthodologique sur ce que signifie travailler sur un échantillon de discours conversationnel authentique, dont une partie de l'analyse est confiée à un logiciel.

3.1 Constitution d'un corpus d'oral spontané

3.1.1 Les enjeux

La constitution d'un « corpus » revêt des enjeux importants que nous souhaitons aborder à présent. Quels critères doivent présider à son élaboration ? Si la linguistique de corpus propose une méthodologie éclairante à ce sujet, sa conception de la définition d'un corpus et de sa raison d'être ne nous semble pas toujours adéquate pour le travail que nous nous proposons d'accomplir.

3.1.1.1 Eléments de définition

Tout texte est susceptible de faire corpus. Le choix n'est pourtant pas anodin, et différentes définitions apportent une vue contrastée de ce qu'un « bon » corpus doit être. Pour les chercheurs se réclamant de la linguistique de corpus, un corpus doit constituer un échantillon de discours le plus large donc le plus représentatif possible. C'est ce qui ressort de la définition suivante :

> Corpus: A collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language.[1]

Comme le précise le site internet de l'Université de Lancaster, cet échantillon recherche donc la diversité en ce qui concerne le sexe des locuteurs, leur âge, la classe sociale à laquelle ils appartiennent, et le type de texte collecté.[2] Une fois les textes compilés, ils subissent une analyse manuelle puis un traitement informatique. Nombreuses sont les tentatives d'étiquetage de corpus longs afin de les rendre automatiquement lisibles par la machine pour un traitement quantitatif rapide et efficace des données. Le but de la constitution d'un tel corpus est alors de déterminer, principalement pour des faits de syntaxe, des tendances générales et des variations, et d'effectuer des calculs statistiques à grande échelle.

Notre objectif est également de détecter des tendances et des variations, au sein d'un corpus nous éclairant sur le fonctionnement de l'oral spontané. Cependant, comme nous espérons l'avoir montré au chapitre précédent,[3] notre approche se veut

[1] Cette définition, tirée du rapport EAGLES (1996e) (*Preliminary recommendations on corpus typology.* EAG-TCWG-CTYP/P. Pisa : Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale.), est disponible sur le site officiel du groupe d'experts (site consulté le 15 août 2003) : http://www.ilc.cnr.it/EAGLES96/corpintr/node13.html#SECTION00040000000000000000

[2] Voir le site internet de l'Université de Lancaster, conçu comme un complément de l'ouvrage suivant : McENERY, Tony & WILSON, Andrew (1996). *Corpus Linguistics.* Edimbourg : Edinburgh University Press : http://www.ling.lancs.ac.uk/monkey/ihe/linguistics/corpus2/2fra1.htm (site consulté le 25 octobre 2001).

[3] Voir Chapitre 2, p.108.

davantage qualitative, et nécessite un autre type de corpus. Comme le précise C. Leroy, le choix du corpus étant susceptible d'orienter les résultats d'une recherche, il doit être déterminé en fonction des objectifs de cette recherche.

> [...] il s'agit donc de savoir ce que l'on fait, c'est-à-dire d'éviter le mélange des genres (s'en tenir à la sacro-sainte homogénéité !) et de savoir que le choix fait dans les multiples possibilités de corpus déterminera, ou pour le moins influera sur les résultats obtenus ; il s'agit également, *a contrario*, de savoir adapter son corpus au(x) but(s) que l'on fixe à sa description.[4]

Nous avons donc choisi de compiler un corpus original, afin qu'il ne porte que le biais que nous choisissons de lui donner (ou que nous lui donnons inconsciemment), dans l'objectif de nos recherches. Notre corpus, augmenté de celui mis à notre disposition par Elodie Vialleton, constitue un tout homogène à l'échelle d'une thèse, compte tenu du temps nécessaire pour l'établissement de la transcription orthographique et du tirage des courbes. Nécessité matérielle, le resserrement de notre corpus sur un nombre restreint de conversations longues nous permet d'en avoir une connaissance profonde, qui tient compte non seulement de leur forme mais aussi de leur contenu, explicite ou allusif (notamment en référence à la situation d'enregistrement dans laquelle nous avons placé les locuteurs), et des relations intersubjectives qui s'y nouent. En outre, nous disposons des enregistrements et avons accès à une transcription de première main la plus fidèle possible que nous avons effectuée nous-même.

Outre ce corpus original, nous aurons également recours, à titre de comparaison, à des corpus existants, pour lesquels nous n'avons malheureusement pas accès aux enregistrements. Récemment compilés, certains d'entre eux font référence dans le domaine de la langue parlée.

[4]LEROY, Christine (1985). La notation de l'oral. *Langue française* 65 : 6-16. (p.7).

3.1.1.2 Corpus existants

Différents corpus sont mis à la disposition des chercheurs. Leurs caractéristiques propres les rendent diversement intéressants pour notre étude. Nous n'en retiendrons que deux.

Corpus existants utilisés Nous pensons ici tout d'abord au British National Corpus.[5] Cette compilation, que nous devons aux chercheurs de l'Université d'Oxford, est constituée de 100 millions de mots, dont 10% (c'est-à-dire 10 millions) ont été produits à l'oral dans des situations diverses. Grâce au logiciel SARA, il est possible d'effectuer des requêtes précises sur le corpus entier, ou uniquement sur la partie produite à l'oral. Ainsi pouvons-nous connaître le nombre total d'occurrences d'un même mot. Si celui-ci se retrouve dans plusieurs catégories de parties du discours, il est possible de resserrer les critères de recherche afin d'obtenir uniquement les chiffres concernant la partie du discours qui nous intéresse. Enfin, nous aurons recours à des éléments du contexte immédiat en élargissant la requête aux segments voisins du mot concerné.

Outre le BNC, nous ferons usage des données chiffrées présentées par les auteurs de la *Longman Grammar of Spoken and Written English* (LGSWE).[6] Ce corpus, collecté par D. Biber et S. Conrad et dépouillé à l'Université de Northern Arizona, comprend de 40 millions de mots, dont 25% produits à l'oral spontané, et constitue également une source de données très intéressante. Ce corpus rassemble des productions langagières relevant de quatre catégories différentes : la fiction, le discours scientifique et académique, la presse écrite, et la conversation. Le corpus conversationnel se divise lui-même en deux parties selon que les locuteurs parlent un anglais britannique ou un anglais américain. Sauf indication contraire, les données relatives à la conversation concernent donc les deux variétés d'anglais. Nous privilégierons les données issues du

[5]Le BNC est disponible en ligne à l'adresse suivante : http://www.natcorp.ox.ac.uk.

[6]BIBER, Douglas (éd.) (1999). *Longman Grammar of Spoken and Written English.* Longman : Harlow, Essex.

corpus d'anglais britannique lorsque cela est possible. Remarquons que, contrairement au BNC, le corpus LGSWE n'est pas consultable ou interrogeable directement. Nous n'avons accès qu'aux résultats des requêtes faites par les auteurs sur leur corpus. Ces résultats étant la plupart du temps présentés sous la forme de tableaux de valeurs ou de graphiques de statistiques, il est néanmoins possible de les utiliser à titre de comparaison. Ils s'accompagnent également d'une conclusion détaillée de chaque expérience statistique, mettant en lumière la spécificité de chaque type de corpus, et notamment celle du corpus conversationnel qui fait l'originalité de cette grammaire.

Corpus existants écartés Nous avons délibérément écarté le corpus Brown et son jumeau britannique le corpus LOB (Lancaster-Oslo/Bergen), compilé dans les années 1960 par G. Leech, S. Johannson et K. Hofland. Ces deux corpus, largement utilisés par les linguistes anglicistes, comportent un million de mots chacun, provenant de 500 textes en prose d'environ 2 000 mots. Les textes ont été sélectionnés parmi les publications de l'année 1961 et relèvent pour les deux corpus de 15 catégories différentes (par exemple, reportages, documents officiels, fiction...). Le LOB et le Brown sont une mine pour un travail sur l'écrit. Mais, contrairement au BNC, ils sont exclusivement constitués de textes écrits. En outre, si le LOB rassemble des textes dont les auteurs sont, en principe, natifs des îles britanniques, le Brown est un corpus américain.

Deux entreprises conjointes de compilation de textes, *the Survey of English Usage* (de University College à Londres) et *the Survey of Spoken English* (de l'Université de Lund, en Suède) ont abouti à la fin des années 1970 à la publication du corpus London-Lund. Il s'agit d'un corpus d'un million de mots comportant 200 textes de 5 000 mots produits à l'oral par des anglophones britanniques issus d'une classe moyenne cultivée. Ces textes ont néanmoins des caractéristiques différentes : la moitié d'entre eux a pour origine une source écrite (publiée : lecture de textes de fiction, documents administratifs, pièces de théâtre, journaux télévisés ; ou non publiée : lettres, journaux intimes...).

Les textes sans antécédent écrit se divisent encore en deux groupes selon qu'ils ressortissent au monologue ou au dialogue. Parmi les textes monologaux se situent par exemple des commentaires sportifs ou des discours (préparés à l'avance mais non rédigés). Le dialogue regroupe des conversations téléphoniques ainsi que des conversations en face-à-face entre deux à quatre locuteurs naïfs ou non naïfs. Ainsi, le corpus proprement oral est constitué de 500 000 mots, tandis que le corpus de conversation spontanée en face-à-face correspond à 30 000 mots. Le principal inconvénient de ce corpus est qu'il n'est pas disponible en ligne, à notre connaissance, en libre accès. Il est en effet difficile de mener à bien des recherches automatiques lorsque l'on n'a accès qu'à la version imprimée du corpus.[7] Cependant, certaines des observations présentées dans la grammaire de R. Quirk *et al.* sont fondées sur l'analyse de ce corpus.[8]

Il existe encore bien d'autres corpus de la langue anglaise parlée. Mentionnons-en seulement quelques-uns. Par exemple, le corpus Lancaster/IBM se compose 52 000 mots issus des archives de la BBC. Il a rapidement subi un étiquetage rigoureux afin de le rendre exploitable automatiquement. On le désigne ainsi par l'acronyme MARSEC (Machine-Readable Corpus of Spoken English). Il est notamment exploité par les chercheurs aixois.[9] Après l'avoir envisagé, nous n'utiliserons pas le corpus Christine, établi par G. Sampson, chercheur à l'Université du Sussex (Brighton), dans la mesure où il est constitué pour partie d'extraits du corpus oral du BNC (sont également inclus des extraits du corpus de l'Université de Reading sur les émotions,[10] et du corpus de London-Lund).[11] Enfin, remarquons que des corpus plus spécialisés sont également disponibles. Le corpus COLT (Corpus of London Teenage Language),

[7]Voir SVARTVIK, Jan & QUIRK, Randolph (1980). *A Corpus of English Conversation.* Lund : C.W.K. Gleerup.

[8]QUIRK, Randolph *et al.* (1985). *A Comprehensive grammar of the English language.* London : Longman.

[9]Ceux-ci sont sur le point de rendre public leur propre corpus étiqueté : Aix-MARSEC.

[10]Des renseignements sur ce corpus d'accès protégé sont disponibles à l'adresse suivante (site consulté le 11 août 2003) : http://www.rdg.ac.uk/AcaDepts/ll/speechlab/emotion/.

[11]Le corpus Christine est disponible en ligne à l'adresse suivante (site consulté le 20 décembre 2002) : http://www.cogs.susx.ac.uk/users/geoffs/ChrisDoc.html.

compilé par des chercheurs de l'Université de Bergen (Suède) est constitué de 500 000 mots prononcés par des adolescents de Londres au début des années 1990. L'intégralité de ce corpus figure dans le BNC. Les variétés linguistiques font également l'objet de compilations, hors des îles britanniques (en Nouvelle-Zélande, par exemple), ou sur le territoire britannique (nous pensons ici à l'immense travail mené à bien à l'Université de Cambridge et à l'Université d'Oxford par l'équipe d'E. Grabe et de F. Nolan sur l'intonation dans les îles britanniques[12]).

A la différence du nôtre, ces différents corpus sont interrogeables automatiquement grâce notamment à l'étiquetage syntaxique qu'ils ont subi. Nous avons envisagé un tel étiquetage car c'est une garantie qui rend un corpus propre à l'utilisation par d'autres chercheurs. Toutefois, la formation technique supplémentaire et le temps de réalisation nécessaires à une telle entreprise la rendaient impossible dans le cadre de la présente étude.

3.1.1.3 Remarques concernant la représentativité d'un corpus

Sur la constitution d'un corpus s'exercent des contraintes multiples, parfois extérieures à l'objectif linguistique initial — temps et financement sont au nombre de celles-ci. Un corpus n'a pas la prétention d'épuiser les richesses d'une langue et d'en donner une image totalement fidèle. Une compilation, si grande soit-elle, reste un réservoir d'exemples avérés, fruits de la créativité potentiellement infinie de locuteurs dans le cadre limitatif d'une langue. La taille optimale d'un corpus est atteinte lorsque la redondance des phénomènes les plus fréquents n'est plus contrebalancée par la découverte d'une proportion plus large de faits jugés jusque-là rares. Elle est par conséquent également fonction de ce que le chercheur souhaite y trouver. Se pose donc inévitablement la question de la représentativité d'un corpus : dans quelle mesure des

[12]Leur corpus (IViE) est disponible à l'adresse suivante (site consulté le 11 août 2003) : http://www.phon.ox.ac.uk/∼esther/ivyweb/index.html.

données issues d'un ensemble réduit de textes peuvent-elles renseigner sur la langue dans son ensemble ?

Si elle constitue sans doute un objectif de départ lors de la constitution d'un corpus, il semble que la représentativité n'est qu'un leurre. C'est ce qui ressort de l'introduction de l'ouvrage de présentation du corpus LOB :

> It follows then that the LOB Corpus is not representative in a strict statistical sense. It is, however, an illusion to think that a million-word corpus of English texts selected randomly from the texts printed during a certain year can be an ideal corpus.[13]

Quel crédit accorder à notre corpus ? Faute de pouvoir prendre la langue orale dans son ensemble comme point de référence, nous nous sommes attachée à vérifier sur plusieurs points la ressemblance de notre corpus aux deux corpus de référence présentés plus haut.[14] Il s'agit ici en fait de tester l'appartenance de notre corpus au même échantillon que celui dont est extrait le corpus de la LGSWE. Les auteurs de cette grammaire présentent une expérience intéressante permettant de donner une caractéristique distinctive de la conversation par rapport aux trois autres types de corpus. A partir d'un extrait de 70 mots de leur corpus conversationnel qu'ils déclarent être représentatif du reste, ils calculent la densité lexicale de l'extrait.

Nous avons procédé à la sélection de six échantillons d'environ 70 mots dans le corpus Celia&James (422 mots au total), issus de quatre épisodes différents (*Summer, Beer, Keys* et *John's permission*). Après avoir dénombré les mots lexicaux, grammaticaux et les *inserts* (pauses pleines, marques d'acquiescement, phatiques...) dans les six textes et avoir calculé leur proportion respective par rapport à 70 et à 100 mots, nous avons comparé les chiffres obtenus avec ceux donnés par les auteurs de la LGSWE.[15]

[13] JOHANNSON, Stig & HOFLAND, Knut (1989). *Frequency Analysis of English Vocabulary and Grammar* 1. Oxford : Clarendon Press. (p.3).

[14] Voir p.148.

[15] BIBER, Douglas (éd.) (1999). *Longman Grammar of Spoken and Written English.* Longman : Harlow, Essex. (p.61).

Comme le montre le Tableau 3.1, les résultats sont très proches : notre corpus contient 40% de mots lexicaux contre 41 pour le corpus LGSWE, 49% de mots grammaticaux contre 44, et 11% d'*inserts* contre 14%. Cette expérience tend à prouver que la densité lexicale est comparable dans notre corpus et dans le corpus LGSWE.[16]

		Mots lex.	Mots gramm.	Inserts
Echantillons Celia&James	Total	169	208	45
(422 mots)	Pour 70 mots	28	35	7
	Pour 100 mots	**40%**	**49%**	**11%**
Echantillon LGSWE	Total	29	31	10
(70 mots)	Pour 100 mots	**41%**	**44%**	**14%**

TAB. 3.1 – Comparaison de la densité lexicale de notre corpus et du corpus LGSWE

Faute de pouvoir prouver la représentativité de notre corpus ou même celle des grands corpus, nous ne pourrons qu'avoir recours à de tels tests statistiques afin d'évaluer le jeu existant entre les données issues de notre corpus par rapport à celles des corpus de référence, qui constituent un échantillon sensiblement plus large de la réalité. Ainsi, nous garderons à l'esprit l'affirmation suivante selon laquelle aucun corpus ne peut épuiser les richesses de complexité et de créativité d'une langue :

> [...] le corpus reste sélectif, il ne peut rendre compte de toute la réalité : nombres de possibilités linguistiques ne se retrouvent jamais dans un corpus, tout simplement parce que le langage humain est trop vaste, et l'on n'est jamais sûr de la raison pour laquelle telle ou telle donnée est absente d'un certain (ensemble de) corpus : cela peut représenter un accident, une impossibilité dans le contexte, dans tout contexte, ou au contraire un choix. [17]

Cependant, le recours au corpus nous semble être une nécessité. L'observation de productions avérées et la théorisation constituent deux pans indissociables de notre

[16] D'après le test du χ^2 (khi2), la différence entre les deux distributions n'est pas significative (χ^2 = 0,789, 2 ddl, p $\leqslant$ 1).

[17] COPPIETERS, René (1997). Quelques réflexions sur la question des données : corpus et intuitions. *Recherches sur le français parlé* 14 : 21-41. Aix-en-Provence : Publications de l'Université de Provence. (p.36).

étude, l'un se nourrissant de l'autre et vice-versa. Certes, une partie de notre travail s'apparente au dépouillement d'une masse d'exemples que l'on cherche à expliquer, classer, et à partir desquels on souhaite aboutir à une généralisation de certains principes. En cela, nous nous rapprochons d'une logique de type semblable à la *data-driven approach*, décrite par C. Kerbrat-Orecchioni dans les termes suivants :

> [...] c'est au contraire de l'observation scrupuleuse des faits qu'il faut partir pour élaborer petit à petit la théorie. Dans cette approche « data driven », les constructions théoriques doivent entièrement être mises au service des données empiriques, et non l'inverse. [18]

Pourtant, comme nous venons de le voir, la constitution même d'un corpus doit être guidée par des considérations théoriques préalables. L'apparente dualité entre observation et théorisation se résout donc par de constants allers-retours entre l'une et l'autre, et par une évolution dynamique des construits théoriques et de l'attention de l'observateur.

3.1.2 Corpus originaux analysés et rejetés

Notre corpus se compose de plusieurs enregistrements originaux, effectués entre juillet 1999 et juin 2001. Le corpus primaire est celui que nous avons dépouillé depuis le départ. C'est celui sur lequel nous avons commencé à travailler et qui est à ce titre notre réservoir d'exemples privilégié. Nous utiliserons le corpus qu'Elodie Vialleton met à notre disposition comme corpus d'appoint, pour appuyer nos analyses sur des exemples supplémentaires. Nous nous servirons également d'un corpus de comparaison auquel nous n'aurons recours que pour vérifier des tendances présentes dans les autres corpus. L'utilisation de différents corpus nous permet de nous prémunir quelque peu contre l'écueil que constitue l'« effet de corpus » qui tend à généraliser des ten-

[18] KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales.* Paris : Armand Colin. (Tome 1, p.47)

dances qui ne sont caractéristiques que d'un corpus et à en masquer d'autres qui sont davantage représentatives d'une règle générale. Nous n'en sommes pas néanmoins totalement protégée dans la mesure où, comme nous allons le voir à présent, les corpus ont des similitudes non négligeables (ce qui confère par ailleurs une homogénéité au tout).

3.1.2.1 Corpus primaire

Ce corpus, composé de Celia&James et de Helen&Rebecca, est notre principale source d'exemples. Nous en avons établi la transcription et édité les courbes intonatives.

Corpus Celia&James Ce corpus met en présence deux personnes de statut et d'âge très différents, dont l'anglais est représentatif de ce que l'on a coutume de désigner comme la « *Received Pronunciation* ».[19] Celia est secrétaire du laboratoire de phonétique de l'Université d'Oxford. Elle a 54 ans. Originaire d'une banlieue huppée de Londres (Hampstead), elle s'est installée à Oxford il y a 30 ans. Agé de 23 ans, James vient de Londres et habite Oxford depuis 5 ans. Il est étudiant en phonétique et prépare un MPhil. Etant amené à faire ses recherches dans le laboratoire, il s'y est rendu dans le but d'obtenir une clef d'accès auprès de Celia. L'enregistrement a eu lieu l'après-midi dans le laboratoire de phonétique et dure 27 minutes. Les deux locuteurs

[19]Remarquons que si ce terme reste largement utilisé pour désigner l'accent le plus répandu dans le sud de l'Angleterre, l'existence d'une telle référence est contestée. D. Jones, désireux de ne pas adopter une attitude prescriptive dans son manuel de prononciation, décrit la RP comme l'accent le plus largement compris dans le monde anglophone : (JONES, Daniel (1909, 1956). *The Pronunciation of English.* Cambridge : CUP. (pp.3-4).)

> In [this book] is given a fairly detailed description of one form of English pronunciation which, though not a standard, can at least be said to be easily understood throughout the English-speaking world, and attention is called to some of the more outstanding divergences commonly heard in various localities and to differences of style employed by individual speakers. The 'widely understood pronunciation' here described may be termed 'Received Pronunciation' (abbreviation RP). This term is not a particularly good term, but it is doubtful whether a better one can be found.

se sont d'abord sentis peu à l'aise. D'une part, ils ne s'étaient jamais rencontrés auparavant. D'autre part, ils ont été troublés par le manque de directives reçues de notre part concernant le contenu et la forme de leur échange.

Cependant, celui-ci présente des caractéristiques qui le rendent exploitable dans le cadre de ce travail. Tout d'abord, il relève de l'oral spontané. Si l'enregistrement a eu lieu au laboratoire de phonétique d'Oxford, il ne s'agit pas de ce que l'on appelle un enregistrement de laboratoire. Ensuite, effectué dans un lieu insonorisé, il a bénéficié de bonnes conditions d'enregistrement, même si les locuteurs n'étaient pas en cabine ou en chambre anéchoïque comme c'est parfois le cas. Libres de leurs mouvements et de se placer comme ils le souhaitaient, les deux locuteurs ont pris place sur des chaises éloignées d'environ deux ou trois mètres.

Sans doute la présence des microphones a-t-elle légèrement perturbé l'échange, de même que l'incongruité de la situation. Cependant, James est linguiste et pratique lui-même ce genre d'expériences. Celia travaille quant à elle en contact avec des linguistes au sein d'un important laboratoire de phonétique. Il est donc à penser que la gêne a dû être minime.

Il est intéressant de noter que la matérialité de l'enregistrement devient objet de discours à plusieurs moments dans la conversation. Dans cet extrait, James s'amuse de ce que nous écouterons ce que Celia et lui disent à propos de nous au moment de dépouiller l'enregistrement :

J oh I it's *quite funny*
C *but*
J we can talk about them on the tape <rires>
C <rires>
J they'll get back to France <rires> see what <bruit> things we had to say

(Celia&James : Esther)

Ici, Celia est en train de perdre son micro-cravate. James lui demande de vérifier que le magnétophone enregistre toujours en regardant le vu-mètre. Il enchaîne ensuite

sur une expérience malheureuse liée à des problèmes d'enregistrement. L'enregistrement présent devient point de départ pour un long échange.

C *and* the combination of for both front {1,10} <problème de micro> doo::r <problème de micro> that looks as if that microphone's <h> {0,70} s::lightly *at a funny angle* <h>

J *w-ye* {0,50} probably then you can see your recorder there is there a

C yes it is moving <rires>

(Celia&James : Recording)

Enfin, à la fin de l'enregistrement, les locuteurs font un bilan de leur conversation : bilan quantitatif autant que qualitatif.

C so they've g- 've got twenty {0,25} twenty m- {0;35} twenty minutes or twenty five minutes on tape <bruit>

J er twenty f-

C something like *that*

J *twenty* five

C twenty five minutes on tape <rires> <h> (0,80)

J o::kay::

C *that will probably do*

(Celia&James : John's permission)

Le caractère naturel et spontané de la conversation ne nous semble pas non plus remis en cause par l'incongruité de la situation. Enregistrés dans un lieu familier, les interlocuteurs se laissent rapidement prendre dans un échange animé marqué par un grand nombre d'interruptions et de chevauchements de parole. D'autre part, l'échange ne relève pas de l'interaction type, qui intéresserait les pragmaticiens, entre secrétaire et étudiant en quête d'un renseignement ou d'une faveur. Il est intéressant de noter que le thème de la clef, s'il est relativement récurrent, n'est absolument pas central dans l'échange. En outre, l'incidence du statut et de l'âge des locuteurs sur la forme de l'échange nous semble négligeable par rapport à celle de leur personnalité.

Nous faisons donc l'hypothèse que ce corpus relève de la catégorie des conversations naturelles. C'est d'ailleurs une revendication formulée par les deux locuteurs à la fin de l'enregistrement.

C *this is* genuine RP speech
guaranteed genuine R- P speech
J *genuine RP speech except that* {1,00} I should open *my mouth more
C *free conversation*

(Celia&James : GRPS)

La conversation tourne autour de quelques sujets principaux. Ceux-ci ont trait tout d'abord aux études de linguistique à l'université d'Oxford en général et à celles de James en particulier. Il relate notamment ses difficultés à établir un corpus sur lequel appuyer ses recherches et décrit une récente série d'expériences malheureuses. Ainsi, James explique la raison de sa venue au laboratoire et réitère sa requête, à laquelle Celia accède. Enfin, l'enregistrement se clôt sur une longue séquence au cours de laquelle les locuteurs parlent de voyages en Asie, et en particulier en Chine.

La totalité de ce corpus figure dans le seconde volume de ce travail sous forme de transcription orthographique et de tracés intonatifs.

Corpus Helen&Rebecca Ce corpus réunit deux lectrices d'anglais à l'Université Paris III. L'enregistrement s'est déroulé par une chaude journée du mois de juin 2001 sur leur lieu de travail, dans un bureau du département d'anglais du centre Censier. Assises autour d'une petite table ronde, les deux locutrices ont parlé pendant 25 minutes environ.

Agée de 25 ans, Helen vient de Londres où elle a vécu depuis son enfance jusqu'à son entrée à l'Université d'Oxford, et où elle est ensuite revenue pour passer un PGCE. C'est à l'Université de Londres qu'elle a rencontré Rebecca, inscrite dans le même cursus. Rebecca, 27 ans, a passé son enfance dans les Cornouailles, au sud ouest de l'Angleterre. Elle a ensuite étudié les langues à l'Université de Leeds avant de s'installer à Londres. Toutes les deux parlent un anglais standard.

Helen et Rebecca se connaissent très bien puisqu'elles sont co-locataires et ont l'habitude de bavarder ensemble. Dans ce corpus, l'échange est très vif. Les tours de

parole sont pour la plupart très brefs et enchevêtrés. Ce corpus présente de nombreuses marques de connivence. On y trouve fréquemment des références totalement propres à la sphère des deux amies : il est longuement question de leur vie non seulement professionnelle mais également privée, ainsi que de leurs amis. Ce fonctionnement par sous-entendus rend parfois notre interprétation difficile. Il nous a néanmoins été possible de contacter les locutrices après l'enregistrement pour éclaircir certaines allusions. Cela a notamment été le cas pour « Pigs in Paris », surnom donné à l'entreprise de restauration pour laquelle Rebecca a travaillé, et qu'elle ne tient visiblement pas en haute estime.[20]

Placées dans une situation certes un peu particulière, Helen et Rebecca semblent rapidement à l'aise. Il est vrai qu'en tant que lectrices de langue étrangère dans une université, elles sont régulièrement amenées à parler en présence de micros et à être enregistrées. Bien entendu, leur rapport avec un magnétophone lors d'une conversation amicale est différent de celui qu'elles peuvent avoir au laboratoire de langues, par exemple. Ainsi, les deux locutrices font état d'une certaine gêne à désigner certaines personnes par leur nom.

Dans cet extrait, Helen raconte qu'une de ses collègues lui rappelle constamment qu'elle est anglaise. Même si tout cela est fait sur le ton de la plaisanterie, Helen se sent chaque fois offensée. On voit ici s'illustrer le fonctionnement par allusion puisque Helen ne veut pas nommer sa collègue, dont Rebecca connaît déjà l'identité. Faisant appel à la connivence, marquée par « you know », elle la désigne par « thingy ».

H well I've just spent all day being {} harassed about being English again
R what <en imitant un accent français> talking like this to you or
H no just
oh bah elle est anglaise so oh she's English and
R why
H you know:: thingy {} mentioning no names
R yeah quite good <rires> yeah

[20]Voir Helen&Rebecca : English.

H always hm {} always kind of {} acts yeah like {} it's really hilarious to {} point out I'm English every five minutes <h>

(Helen&Rebecca : English)

Plus loin, Rebecca dit regretter d'avoir identifié par son prénom quelqu'un sur qui elles ont donné des informations qui les mettent mal à l'aise. Helen tente de la rassurer.

R I don't think I like because of names
H *<rires>*
R *copyright* <rires>
H what is that uh
R no no no XX it's the only reason I won't
H no no but we didn't
R *no*
H *I mean* there are lots of *Myriams <rires>*
R *yeah that's right* yeah exactly
H so don't try and *work out who she is*

(Helen&Rebecca : Weird families)

Néanmoins, leurs scrupules ne les empêchent pas à plusieurs reprises de mentionner nommément plusieurs collègues de Paris III. On pense par exemple à l'épisode New Red Shoes qui clôt l'enregistrement. Les éclats de rire sont provoqués par un qui-pro-quo entre deux collègues de Helen portant le même prénom, l'un étant un homme, l'autre une femme. L'homme, qu'elles savent toutes deux que nous connaissons, est identifié par son nom de famille sans qu'aucune marque de gêne ne soit perceptible. Il est donc à penser que leur gêne à désigner les personnes dont elles parlent par leur nom ne provient pas spécifiquement du fait d'être enregistrées ou du fait que nous puissions les connaître. En revanche, il semble qu'elle soit davantage liée au contenu de leurs propos. Dans le cas de Myriam, par exemple, les informations données concernant la folie de sa mère sont particulièrement tragiques. Dans le premier cas, il s'agit d'un problème relationnel qui blesse Helen dans sa personne et sur lequel elle veut probablement rester discrète sans dénoncer qui que ce soit.

Habituées à travailler avec du matériel audio, les deux locutrices prennent la liberté d'interrompre l'enregistrement environ 15 minutes après le début. Pendant l'interlude, elles ouvrent l'une des fenêtres du bureau afin d'en rafraîchir l'atmosphère (le changement d'ambiance sonore est audible). Au moment de remettre les magnétophones en route, elles s'adressent directement à nous, comme elles ont déjà fait plus tôt.

R hello here we are again
H yeah plans for Nîmes *[inaudible]*
R *you missed all* the gossip *though*
H *<rires>*
R we felt that you didn't *deserve to hear it all*

(Helen&Rebecca : Nîmes1)

Le lieu d'enregistrement lui-même n'a sans doute pas beaucoup troublé les deux locutrices. Dans le cadre familier de Censier, néanmoins, le bureau où elles étaient installées leur était inconnu. La description des objets qui s'y trouvent (notamment un cendrier) sont le point de départ d'un récit anecdotique concernant les cours de poterie suivis par Rebecca dans son enfance.

Nous avons présenté ici, comme pour le corpus Celia&James, les traces évidentes laissées par la présence de magnétophones et de micros dans la conversation. Celles-ci ne nous semblent cependant pas en mesure de remettre en cause le caractère naturel et spontané de l'échange. Comme pour le corpus précédent, nous faisons l'hypothèse que Helen et Rebecca ont enregistré une conversation naturelle susceptible de servir de support à notre étude. Il constitue notre deuxième corpus de référence, dont la transcription et une partie des courbes prosodiques se trouvent dans le volume 2.

La conversation s'ouvre sur un bilan concernant les perspectives d'emploi des deux locutrices et des divers rendez-vous qu'elles ont pris avec de potentiels employeurs. Une grande partie de la conversation concerne la vie professionnelle de Helen et Rebecca au sein de l'université Paris III : elles parlent de leurs collègues, de leurs horaires, de leurs cours, de leurs étudiants, et de leur lieu de travail en général (et en parti-

culier du lieu dans lequel s'effectue l'enregistrement). Les deux locutrices évoquent également longuement leur vie sociale chargée et les difficultés qu'elles ont à organiser leurs nombreux rendez-vous avec leurs amis — amis dont elles sont donc amenées à parler. Immédiatement après l'interruption momentanée de l'enregistrement, une longue séquence est consacrée à l'évocation par Helen de ses projets de vacances à Nîmes à l'occasion de la feria.

3.1.2.2 Corpus secondaire

Les trois corpus suivants constituent pour nous une précieuse réserve d'exemples supplémentaires. Nous les considérons comme secondaires dans la mesure où nous ne les avons pas dépouillés nous-même. Nous les devons tous les trois, pour la transcription et pour les courbes, à Elodie Vialleton, que nous remercions ici.[21]

Corpus Mike&Hannah L'enregistrement de ce corpus a été réalisé en juillet 1999 dans la maison de Mike, dans la banlieue d'Oxford. Mike est guide touristique pour une agence de voyage dont il emmène les clients visiter la Grèce. Né en Grèce il y a trente ans de parents originaires du Hertfordshire, il vit dans le sud de la Grande-Bretagne depuis son plus jeune âge. Hannah, âgée de 28 ans, est née à Londres et vit à Oxford depuis l'âge de 10 ans. Elle travaille pour les services sociaux. Voisins et amis, Mike et Hannah se retrouvent régulièrement le soir chez l'un ou l'autre pour bavarder. Au moment de l'enregistrement, ils ne s'étaient pas vus depuis deux semaines, pendant lesquelles Mike était en vacances.

La situation de parole est donc familière aux deux locuteurs, en dehors de la présence de notre matériel d'enregistrement. Tout d'abord étonnés par notre démarche, ils perdent rapidement la retenue qu'ils se croient obligés d'avoir, surveillés par le « chaperon high-tech »[22] que constitue le magnétophone.

[21] Pour une version complète de ce corpus, voir la thèse d'Elodie Vialleton.
[22] Voir la remarque de Mike :

Ce corpus est marqué par la domination de Mike en termes de temps de parole. Ce déséquilibre a plusieurs explications. Tout d'abord, du fait de l'heure tardive, Hannah est fatiguée. En outre, Mike, qui avoue être plus bavard que son amie, rentre de vacances : il a donc beaucoup d'aventures à lui raconter. Ce déséquilibre se note dans la forme que prend la conversation : certains passages relèvent du récit, mais le dialogue ne devient pas monologal puisque Hannah continue à intervenir pour demander des précisions à Mike ou pour s'exclamer. Nous faisons donc l'hypothèse que ce corpus est la trace d'une conversation naturelle et est donc à ce titre susceptible d'éclairer le fonctionnement de l'anglais oral spontané. C'est pourquoi nous en ferons usage en tant que corpus secondaire en complément de notre corpus primaire.

Les thèmes abordés par les deux locuteurs sont en nombre réduit. Tout d'abord, devant la surprise de Hannah, Mike explique à son amie comment nous avons pris contact avec lui pour l'enregistrer. Bien vite, celle-ci enchaîne sur le récit de plusieurs anecdotes concernant ses parents, ses amis ou encore elle-même. Ensuite commence la narration des vacances de Mike, qui occupe quasiment tout le reste de l'enregistrement. Celui-ci est interrompu par notre retour dans la pièce environ 30 minutes après avoir laissé les locuteurs.

Corpus Nick&Matthew L'enregistrement de ce corpus a été fait en juillet 1999 dans la chambre d'étudiant de l'un des deux locuteurs à Balliol College. Les deux amis, étudiants à l'Université d'Oxford, se retrouvaient comme à leur habitude autour d'une tasse de thé en début d'après-midi. Nick, âgé de 26 ans, est originaire du Surrey, dans le sud de l'Angleterre. Le sujet de la thèse de doctorat qu'il est sur le point d'achever porte sur l'histoire de l'euthanasie. Matthew vient quant à lui de Leeds, au nord de

M this is like a chaperone {0,6} it's a high tech chaperone *and they're just going to* and the tape will be sent to your parents afterwards to prove that *we only discussed*

(Mike&Hannah : Recording)

l'Angleterre. Il a 27 ans et prépare également une thèse, en chimie.

La durée de l'enregistrement est d'un peu moins d'une heure, mais la totalité ne nous semble pas exploitable pour plusieurs raisons. A la demande des deux locuteurs, nous avons assisté à l'enregistrement de la première demi-heure de ce corpus (Corpus Nick&Matthew1). Les quelques passages où nous participons à la conversation ont donc dû être écartés de nos analyses prosodiques. Dans la deuxième demi-heure (Corpus Nick&Matthew2), une troisième locutrice fait une brève apparition : Nick et Matthew reçoivent la visite de leur amie Libby. Ce passage a également dû être écarté dans la mesure où l'on entend mal la locutrice qui n'est pas équipée de micro.[23] Au total, un peu plus de 40 minutes de ce corpus sont exploitables dans le cadre de notre étude.

Le caractère spontané de ce corpus est assuré par la familiarité des locuteurs avec le cadre et la situation dans lesquels l'enregistrement s'est déroulé. Cependant, l'influence de la présence des micros est perceptible au début de l'enregistrement. Matthew se montre particulièrement timide et attend la plupart du temps d'être sollicité par Nick ou par les deux linguistes françaises à qui il avait demandé de rester. Très rapidement, les locuteurs se montrent très à l'aise, ce dont témoigne le style relâché de leur conversation, qui alterne entre des passages où Nick occupe le canal sonore, et des passages plus équilibrés où Matthew, moins bavard que son ami, prend également en charge la conduite du dialogue.

Les thèmes abordés par les deux locuteurs dans ce corpus touchent d'abord à l'enregistrement auquel ils prennent part, leur façon de parler, et les accents des membres de leurs familles respectives. La conversation s'établit ensuite sur leur travail de recherche et leurs perspectives d'emploi. Nick et Matthew discutent également de leurs loisirs, et parlent longuement de voyages, de cuisine et de musique.

[23]En outre, Libby a un accent distinctement américain, qui est susceptible d'influencer son utilisation des paramètres intonatifs.

Corpus Ian&Claire Contrairement aux autres enregistrements, celui-ci s'est déroulé sur deux jours. Constitué de deux conversations distinctes (Ian&Claire1 et Ian&Claire2), il est notablement plus long que les autres (près d'une heure et demie). L'enregistrement a eu lieu à Leicester en avril 2001, chez les locuteurs.

Agée de 50 ans, Claire a grandi à Leicester où elle est maintenant enseignante. Ian, son compagnon, a quitté le nord de l'Angleterre il y a de nombreuses années pour venir vivre à Leicester et y travailler en tant qu'ingénieur. Deux autres personnes sont présentes lors des enregistrements : leur fille Eleanor, 11 ans, qui ne souhaitait pas prendre part à la conversation mais que l'on entend finalement dans le premier enregistrement ; et Anna, leur locataire, dans le second.

Ce corpus présente la particularité de saisir une heure et demie de la vie quotidienne d'une famille d'anglophones au moment du dîner. Une grande partie de la conversation est directement liée à la thématique du repas, comme dans l'exemple suivant :

I *Eleanor could you go and get me another* glass of water as I'm chained to the microphone please *thanks*

(Ian&Claire1 : Truffles)

D'autres thèmes sont abordés. Dans Ian&Claire1, en présence d'Eleanor, la conversation tourne autour de la vie quotidienne de la famille et de ses membres. Dans Ian&Claire2, en présence d'Anna, les trois locuteurs qui se connaissent depuis peu parlent de leurs goûts personnels en matière de politique, cinéma, restaurants, décoration intérieure.

Du fait de la situation d'enregistrement, la qualité sonore des deux bandes est parfois perturbée par des bruits parasites. Ainsi, on perçoit distinctement le tintement de la vaisselle ou encore des bruits de mastication. Il arrive également que les locuteurs parlent la bouche pleine ou doivent détacher leur micro pour aller chercher un plat à la cuisine. La présence pour chaque enregistrement d'un troisième locuteur nuit par

moments légèrement à la compréhension dans la mesure où la locutrice supplémentaire ne portait pas de micro. Si leurs paroles ont cependant pu être transcrites, il n'a pas été possible de leur associer des courbes intonatives.

Malgré tout, nous avons retenu ce corpus pour les mêmes raisons qui nous ont fait prendre en compte les données présentées plus haut. Tout d'abord, les locuteurs parlent un anglais standard, représentatif de la *Received Pronunciation*. Ensuite, la qualité sonore est bonne, et la longueur des enregistrements permet éventuellement d'écarter les passages les moins clairs. Notons que ce corpus ne fait cependant pas l'objet d'une analyse prosodique de notre part. Avant tout, le caractère spontané de ces deux conversations est indubitable. Certes, il est fréquemment fait référence à la matérialité de l'enregistrement et aux fils des micros qui obligent les deux locuteurs à réduire leurs allées et venues entre la table et la cuisine. Cependant, Ian et Claire, enthousiastes à l'idée de fournir le matériau d'une analyse linguistique, ont eux-mêmes décidé du moment de l'enregistrement et de sa durée en mettant les magnétophones en route à l'heure du dîner. Ces deux conversations capturent donc un moment privilégié de la vie familiale ou sociale, non sollicité par nos besoins de linguiste.

3.1.2.3 Corpus de comparaison

Outre le corpus présenté jusqu'ici, il nous a semblé utile de pouvoir recourir à un ensemble plus restreint de données ne se conformant pas exactement à notre projet. Ainsi, nous pourrons faire référence à un corpus d'oral spontané supplémentaire. Celui-ci présente la particularité de relever davantage de l'interview que de la conversation à bâtons rompus. Ce corpus, différent de ceux retenus au titre de corpus primaire et de corpus secondaire, a pour vocation d'éclairer ceux-ci par la comparaison.

Corpus Jonathan Agé de 22 ans, Jonathan est originaire de la banlieue sud de Londres (Croydon). Il vient de terminer brillamment ses études de français et d'espa-

gnol à l'Université d'Oxford. L'enregistrement, long de 30 minutes, s'est déroulé dans une salle de cours de St Hugh's College à Oxford, dont Jonathan est membre. Mal réveillé ce matin-là après avoir fêté la fin de ses études la veille, il montre relativement peu d'enthousiasme à être enregistré.

Les thèmes abordés dans les trois extraits que nous exploitons ici concernent tous la vie de Jonathan, ses ambitions professionnelles et ses voyages. Tout d'abord, il se présente et décrit sa ville natale. Il explique ensuite qu'il a fini ses études et qu'il aimerait entrer dans le corps diplomatique. Il décrit la formation de diplomate telle qu'il se l'imagine. Il parle également de son séjour au Chili et d'événements politiques survenus dans ce pays.

Le format de ce corpus est davantage celui d'une interview que d'une véritable conversation à bâtons rompus. Deux francophones demandent au locuteur anglophone de les renseigner sur des points le concernant. Il se trouve donc dans une position d'informateur et est peu soumis à des contradictions. Son débit est assez lent, les pauses sont longues, et la syntaxe proche de celle de l'écrit. Néanmoins, il s'agit bien ici d'anglais oral spontané, dont l'intérêt principal réside dans le fait que Jonathan parle librement, sans avoir préparé l'entretien, et sans que son tour de parole ne soit menacé. Il peut donc dérouler naturellement son discours. Selon notre schéma du discours de la Figure 1,[24] ce corpus est plus proche de la catégorie du récit que de celle du dialogue.

Nous donnons ici la transcription, à titre de comparaison, des moments les plus parlants sur la logique d'une interaction et l'organisation du discours.

3.1.2.4 Corpus rejeté

La constitution d'un corpus doit se faire dans un souci d'exigence. Pour s'assurer de la cohérence et de la qualité du corpus retenu, il est important, nous semble-t-il, de

[24] Voir p.22.

pouvoir rejeter des corpus potentiels au vu de critères de sélection. Nous avons ainsi été amenée à écarter les deux corpus suivants pour des raisons différentes.

Corpus Tony&Frazer Ce corpus a été enregistré dans un pub d'Oxford, le Three Goats' Heads, en juillet 1999. Il rassemble deux collègues enseignant les langues à l'Université d'Oxford. Tony, âgé d'environ 55 ans, est professeur de français. Originaire de Liverpool, il a perdu son accent régional au profit d'un accent représentatif de la *Received Pronunciation*. Frazer est âgé de 27 ans. Il enseigne l'allemand tout en travaillant sur sa thèse de doctorat. Il vient de Chichester dans le Sussex.

L'enregistrement place les locuteurs dans une situation qui leur est très familière puisqu'ils se rencontrent chaque semaine dans ce même pub pour bavarder pendant des heures. Nous n'avons enregistré qu'une heure de leur conversation. Celle-ci est très animée, la plupart du temps conduite par Tony. Son caractère spontané est d'autant moins contestable que les locuteurs ont vidé un certain nombre de pintes de bière et de cidre au cours de leur échange. Cependant, nous avons choisi de ne pas exploiter ce corpus, et ce pour plusieurs raisons relatives à la situation dans laquelle l'enregistrement a été effectué. En effet, d'autres locuteurs traditionnellement invités à cette rencontre hebdomadaire, dont certains étaient francophones, participent à l'interaction verbale. Ceux-ci sont difficilement audibles puisqu'ils n'étaient pas munis de micros. Ce corpus ne relève donc pas à proprement parler de la conversation à deux. Deuxièmement, l'atmosphère du pub est trop bruyante pour rendre ce corpus exploitable du point de vue de l'intonation. A la musique de fond s'ajoute le bruit des pintes qui s'entrechoquent.

En dépit du fait que la syntaxe est sans doute intéressante, nous avons décidé de ne pas prendre en compte ce corpus pour nos analyses. Nous n'en donnons ni les courbes ni la transcription.

Corpus Ben&Reena Comme le corpus Helen&Rebecca, ce corpus réunit deux lecteurs anglophones de l'Université Paris III. L'enregistrement, long d'environ 30 minutes, s'est déroulé une semaine plus tard dans le même bureau du centre Censier, dans les mêmes conditions.

Ben, âgé d'une vingtaine d'années, est originaire de Bath, tandis que Reena, de dix ans son aînée, vient de Londres. Ils sont tous les deux habitués à travailler avec du matériel audio-visuel et ont l'habitude d'être enregistrés. L'influence potentiellement inhibante des micros et magnétophones a donc dû être minime. Collègues depuis deux ans, les deux locuteurs se connaissent bien, mais n'ont pas le même degré d'intimité que Helen et Rebecca. Cette différence est perceptible dans la forme de leur conversation : les tours de parole sont plus longs, on y trouve beaucoup moins de chevauchements, d'où un échange moins vif. Les sujets de conversation sont abordés de façon moins personnelle. Il est principalement question de l'immobilier à Londres et à Paris, et de la vie en banlieue et à Paris.

Ce corpus, par sa nature spontanée et sa bonne qualité sonore, est amené à l'avenir à figurer parmi nos corpus principaux. N'étant pas en mesure d'en proposer une transcription satisfaisante par manque de temps, il nous faut cependant le rejeter ici avant de pouvoir l'exploiter pleinement à l'occasion de prochains travaux.

3.1.3 Bilan qualitatif et quantitatif du corpus exploité

Le corpus retenu nous semble présenter des caractéristiques qui lui confèrent une certaine unité. D'un point de vue technique, toutes les conversations ont été enregistrées avec le même matériel, ce qui tend à garantir une qualité sonore uniforme.[25] Les différences de qualité proviennent du lieu dans lequel les enregistrements ont eu lieu. Ainsi, le corpus Tony&Frazer a été écarté non pas parce qu'il a été mal enregistré, mais parce qu'il a été enregistré dans un endroit particulièrement bruyant. Nous avons

[25]Les caractéristiques de ce matériel, de qualité professionnelle, sont détaillées en Annexe A.

favorisé des lieux d'enregistrement qui soient familiers aux informateurs, qui, la plupart du temps se connaissaient bien voire très bien avant l'enregistrement. Ceux-ci, qui ignoraient pour la plupart notre intérêt pour l'intonation, n'ont reçu aucune autre directive de notre part que de bien vouloir se parler. Le but étant de faire entrer les paires de locuteurs en conversation libre, il nous a la plupart du temps semblé préférable de les laisser seuls en tête à tête dans la mesure du possible. L'innocence des informateurs est pour nous une garantie de l'authenticité de l'échange : ne sont appelées à figurer dans notre corpus primaire ou secondaire que des conversations dites « naturelles ». Enfin, l'unité du corpus a également des racines sociologiques. Tous les locuteurs enregistrés parlent un anglais britannique standard, représentatif de la « *Received Pronunciation* », et tous sont issus d'une classe moyenne cultivée.

Le Tableau 3.2 présenté ici résume de façon quantitative le corpus exploité dans le cadre de ce travail. Nous fondons nos recherches sur l'observation des productions orales de 11 locuteurs, qui ont prononcé près de 50 000 mots sur un total de plus de 7 heures d'enregistrement.

Corpus		**Durée**	**Nbre de mots**	**Script et courbes**
Primaire	Celia&James	30 mn x 2	6 400 mots	total en annexe
	Helen&Rebecca	25 mn x 2	5 600 mots	total en annexe
Secondaire	Mike&Hannah	30 mn x 2	8 000 mots	extraits
	Nick&Matthew1	20 mn x 2	4 500 mots	extraits
	Nick&Matthew2	20 mn x 2	3 700 mots	extraits
	Ian&Claire1	25 mn x 2	4 500 mots	extraits
	Ian&Claire2	60 mn x 2	12 000 mots	extraits
Comparaison	Jonathan	30 mn	4 500 mots	extraits
TOTAL	11 locuteurs	+ de 7 h	$\sim$ 50 000 mots	

TAB. 3.2 – Corpus utilisé

En somme, nous faisons l'hypothèse que le corpus retenu ici peut nous permettre, par sa qualité et sa quantité, de comprendre le fonctionnement de certains principes hiérarchiques d'organisation des constituants discursifs en anglais oral spontané bri-

tannique standard.

3.2 Traitement du corpus

Le travail présenté ici est fondé en grande partie sur l'observation de notre corpus. Toute notre attention a été portée, dans un premier temps, à transformer ces conversations enregistrées sur une période d'environ deux années en matériau exploitable pour étudier les phénomènes de structuration de l'anglais oral spontané. La question de C. Leroy est donc toujours pertinente :

> [...] que faire de nos enregistrements ? Comment travailler sur une réalité orale essentiellement fugace [...] ?[26]

Paradoxalement, travailler sur l'oral nécessite de se poser la question de son passage à l'écrit. Il ne s'agit pas de rendre l'oral semblable à l'écrit, mais simplement de le saisir et d'en rendre compte, pour des questions de commodité, de façon pérenne. Le changement de support permet en effet de fixer la fugacité de l'échange et de faciliter son exploitation linguistique. Dans un deuxième temps, notre travail a consisté à coucher sur le papier non plus la dimension verbale mais plus généralement la dimension sonore de ces échanges enregistrés. Nous avons donc choisi un logiciel d'analyse prosodique auquel confier le tracé de courbes intonatives. En dernier ressort, nous avons fait correspondre la transcription et les données suprasegmentales.

3.2.1 Etablissement de la transcription

L'établissement de la transcription orthographique constitue une étape fondamentale du traitement du corpus. Si rien ne peut se substituer à l'enregistrement, il s'agit de donner une version écrite la plus fidèle possible de l'échange qui s'est déroulé à

[26] LEROY, Christine (1985). La notation de l'oral. *Langue française* 65 : 6-16. (p.7).

l'oral. La question du mode de transcription et des conventions à adopter se pose donc d'emblée.

3.2.1.1 Considérations méthodologiques

De nombreuses études sur l'oral font état d'un débat sur la méthodologie à adopter pour l'établissement de la transcription de documents sonores. Cette réflexion, relayée par des groupes d'experts, a pour objectif d'établir des conventions de transcription standardisées, afin de permettre l'utilisation d'un même corpus par de nombreux chercheurs. Ce débat aboutit en 1992 avec les conventions NERC,[27] qui proposent de considérer quatre niveaux de précision dans l'établissement d'une transcription. C. Blanche-Benveniste,[28] dont les travaux témoignent d'un vif intérêt pour l'établissement d'une méthodologie rigoureuse, en rend compte ici :

- Le premier donnerait une représentation orthographique avec un minimum de ponctuation (en 1996, Sinclair[29] suggère qu'on omette la ponctuation). A ce niveau, aucune marque des interactions, ni même des changements de locuteurs.
- Le niveau deux introduirait les tours de parole des locuteurs et certains éléments non verbaux.
- Le niveau trois aurait des indications d'intonation et d'interactions. Les frontières d'unités tonales et syllabiques seraient indiquées. Seuls des phonéticiens pourraient le faire, et sur de bonnes qualités d'enregistrement.
- Le niveau quatre additionnerait les indications acoustiques et phonétiques, les schémas mélodiques, un tracé du fondamental (avec d'excellents enregis-

[27] Network of European Reference Corpora.

[28] BLANCHE-BENVENISTE, Claire (1997). Transcriptions et technologies. *Recherches sur le français parlé* 14 : 87-99. Aix-en-Provence : Publications de l'Université de Provence.

[29] C. Blanche-Benveniste fait ici allusion à l'intervention de J. Sinclair lors du colloque de l'EAGLES (Expert Advisory Group in Language Engineering Standards) tenu à Madrid en janvier 1996. Les recommandations de ce groupe sont disponibles en ligne à l'adresse suivante : http://www.ilc.cnr.it/EAGLES/browse.html (site consulté le 15 août 2003).

trements).[30]

Au quatrième niveau, la complexité croissante des recommandations du NERC dépasse nos compétences, et nous semble proposer un texte trop touffu. Ces recommandations présentent néanmoins l'intérêt de hiérarchiser les phénomènes dont il faut rendre compte dans une transcription, et d'affirmer la nécessité de faire des choix. Il est en effet impossible d'être totalement fidèle à l'enregistrement. Seul l'enregistrement peut être fidèle à lui-même, sinon à l'échange dont il est la trace. Le passage d'un support à l'autre entraîne nécessairement une certaine déperdition d'information. Or,

> La qualité de la transcription revêt malgré tout une importance considérable puisque ce sera en dernier ressort le matériau qui servira à l'analyse [...].[31]

Il s'agit donc de déterminer les points précis qui devront faire l'objet d'une attention spécifique, en fonction de nos objets de recherche.

Nous suivons ici les grandes lignes posées par C. Leroy[32] et reprises par M.-A. Morel au sein du Centre de Recherche en Morphosyntaxe du Français Contemporain de l'Université Paris III, ainsi que par M.-A. Morel et L. Danon-Boileau dans leur *Grammaire de l'intonation.*[33] Nous retenons deux principes fondamentaux : une transcription doit rendre compte des mots prononcés, mais également d'un certain nombre de phénomènes perceptibles sur le canal sonore.

Nous nous éloignons cependant de ces trois auteurs sur différents points ayant trait aux phénomènes prosodiques. Contrairement à M.-A. Morel et L. Danon-Boileau, dont nous nous inspirons, nous ne notons pas les variations de hauteur et d'intensité dans la transcription. Pour faciliter la lecture des transcriptions, C. Leroy préconise l'usage de la ponctuation en tant que marqueurs prosodiques. Elle utilise ainsi la virgule pour marquer une pause, un point d'interrogation pour signifier qu'elle in-

[30] BLANCHE-BENVENISTE, Claire (1997). *art. cit.* (p.91).
[31] LEROY, Christine (1985). *art. cit.* (pp.8-9).
[32] *ibid.*
[33] MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *op. cit.*

terprète la séquence qui précède comme une interrogative. Nous ne suivons pas ces conseils concernant l'utilisation de la ponctuation. Mis à part les deux points (:), communément admis pour marquer un allongement, nous ne faisons aucun usage de la ponctuation. Nous ne faisons entrer dans la transcription que la marque des allongements et la mesure des pauses. Pour les variations d'intensité et de la fréquence fondamentale, mais renvoyons le lecteur aux courbes d'analyse prosodique présentées dans notre second volume et aux enregistrements consignés sur le CD-ROM joint.

En somme, nous proposons une transcription de niveau deux, pour reprendre la hiérarchie établie par le NERC, sans ponctuation. Nous y faisons également intervenir quelques données relevant du niveau quatre, dans la mesure où ceux-ci sont audibles, ont pu perturber ou influencer les locuteurs, ou sont visibles sur les courbes.

3.2.1.2 Transcrire le verbal

La première question est celle du choix de l'alphabet à utiliser. C. Leroy fait l'inventaire des avantages et inconvénients de l'utilisation de l'alphabet phonétique pour transcrire les productions orales. Elle y voit deux avantages. Le premier consisterait à soumettre au lecteur un texte visiblement différent de l'écrit, afin de permettre un accès plus direct aux faits oraux, sans passage par l'orthographe traditionnelle et les conventions de l'écrit. L'absence de ces conventions constitue également le second avantage, en permettant au transcripteur de ne plus se soucier du sens et de l'orthographe de ce qu'il entend, mais simplement des sons qu'il entend. Les inconvénients d'une telle pratique sont pourtant nombreux. Ceux-ci tiennent tout d'abord à la lecture d'une telle transcription, impossible au néophyte. De même, la maîtrise de l'écriture de l'alphabet phonétique international est le fait de phonéticiens, et non pas de n'importe quel linguiste intéressé par l'oral. C. Leroy mentionne en dernier lieu la difficulté de dactylographier de longues transcriptions phonétiques à l'aide de claviers standards. Ces inconvénients majeurs lui font préférer la transcription orthographique.

La question d'un recours ponctuel à une notation phonétique peut cependant se poser pour rendre compte des séquences qui ne peuvent pas figurer dans un dictionnaire parce que les segments prononcés sont incomplets ou non reconnaissables. Nous avons choisi de conserver une transcription orthographique dans tous les cas. Lorsque les mots prononcés sont incomplets, nous cherchons toujours à inférer le mot complet, et à donner les lettres des phonèmes perçus. Lorsque le mot demeure non identifié, nous donnons les lettres qui correspondent communément aux phonèmes perçus.

Notons que nous avons recours à la transcription phonémique dans deux épisodes du corpus Celia&James : Diphthongs1 et Diphthongs2, afin de transcrire les sons prononcés par les deux locuteurs au cours du récit par James de ses récentes expériences de phonétique. La transcription phonémique, indispensable ici, est signalée comme telle, entre deux barres obliques.

Une autre solution est envisageable : celle d'approcher la réalisation phonétique tout en conservant une orthographe non phonétique peu conventionnelle. Cette pratique, appelée par C. Blanche-Benveniste « trucage orthographique », est largement contestée, dans la mesure où elle obscurcit le contenu du texte sans pour autant en éclaircir de façon significative la forme. Nous y avons néanmoins recours dans un cas notoire pour transcrire les diverses réalisations de *because*. Il nous semble ici important de distinguer les cas où la conjonction est réalisée sur deux syllabes ou sur une seule. Dans ce dernier cas, nous transcrivons donc par *cos*. On notera cependant que cette forme orthographique est attestée par les dictionnaires,[34] si bien que la notation de cette variante ne relève pas à proprement parler de la pratique condamnée par C. Blanche-Benveniste. En outre, nous utilisons les contractions lexicalisées de *going to* et *want to* en *gonna* et *wanna*, lorsqu'il est approprié de le faire. En tout autre lieu, nous avons opté pour une transcription orthographique conventionnelle.

[34]notamment par le dictionnaire de référence pour la langue anglaise : SIMPSON, John & WEINER, Edmund (1989, 2e éd.). *Oxford English Dictionary*. Oxford : Clarendon Press.

3.2.1.3 Transcrire le paraverbal

Transcrire ne consiste pas seulement à rendre compte des propos échangés. Avec les mots doivent figurer ce que l'on considère parfois comme les scories de l'oral,[35] mais qui en constituent aussi une des grandes originalités. Toutes les répétitions de mots ou de parties de mots sont notées avec précision. Nos transcriptions signalent également les pauses silencieuses et leur durée (inscrite entre accolades en secondes), ainsi que les allongements vocaliques et consonantiques (par le signe (:), seul ou répété selon la longueur approximée de l'allongement, placé après le phonème concerné). Nous faisons mention des pauses pleines sous trois formes différentes en fonction des phonèmes prononcés : *hm*, *mm*, et plus rarement *er*. Toute perturbation importante du signal sonore a également été notée. Nous mentionnons et identifions, entre crochets pointus <>, les soupirs, toux, rires, et autres bruits. Une attention toute spécifique a été portée à la notation des chevauchements de parole. Les segments produits en situation de chevauchement de parole, c'est-à-dire prononcés alors que les deux locuteurs parlent en même temps, sont notés entre étoiles : *segment recouvert*.

Dans l'exemple suivant, c'est tout cela que nous avons choisi de transcrire.

H <**rires**>
R you can get wild on the **e-** on the internet
that's exciting
H ***yeah*** I know I need to do my emails and I need to phone that man
R yes *yeah*
H *but how* can I phone him because **th-** that staff-room phone won't let me phone out to England {**0,80**} will it
R no but what about the man with this with **the:: hm** France radio one {**0,80**} French ***[inaudible]***
H *****<**h**> **oh yeah*** but I haven't got his number with me

(Helen&Rebecca : Staying awake till seven)

[35]Voir KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales*. Paris : Armand Colin. (Tome 1, p.41).

3.2.1.4 Difficultés de transcription

Certains passages de notre corpus, soumis à des anglophones, restent ambigus, voire incompréhensibles. Nous proposons parfois plusieurs transcriptions possibles, séparées par des barres obliques (la première étant la plus plausible). Il nous arrive de proposer une transcription unique, suivie d'un point d'interrogation. D'autres fois, comme dans l'exemple que nous venons de citer, nous notons qu'un segment inaudible n'a pu être transcrit.

Ces difficultés ont fait l'objet d'un grand intérêt par nombre de linguistes soucieux d'établir une méthodologie de la transcription. L'examen de ces erreurs par les chercheurs du GEDA (Groupe d'Etude sur les Données Orales) a permis de dégager des régularités parmi les erreurs commises par les transcripteurs. D'après P. Cappeau, la recherche du sens oriente tout d'abord le transcripteur vers l'attendu, laissant peu de place à l'inattendu.

> Le relevé des corrections apportées aux travaux de transcripteurs novices a permis de vérifier à quel point le locuteur est impliqué dans le travail de transcription et de quelle façon il tend à reconstruire les données. Il occulte souvent ce qu'il ignore ou maîtrise imparfaitement (les modes de production de l'oral), substitue le lexique dont il dispose, et adapte la syntaxe à certaines de ses habitudes.[36]

Bien évidemment, une transcription nécessite de multiples écoutes, de préférence effectuées par des personnes différentes. Cela permet d'éviter la plupart de ces erreurs, sans pour autant le garantir totalement. La transcription reste un difficile exercice d'écoute minutieuse et attentive, au cours duquel l'oreille est sollicitée par des sons que nous restituons en les interprétant malgré nous. Comme le rappellent les membres

[36]CAPPEAU, Paul (1997). Données erronées : quelles erreurs commettent les transcripteurs ? *Recherches sur le français parlé* 14 : 117-126. Aix-en-Provence : Publications de l'Université de Provence. (p.126).

du GEDA autour de M. Bilger, à la suite de C. Blanche-Benveniste et C. Jeanjean,[37]

> [...] l'écoute semble être avant tout une activité interprétative qui provoque plus d'erreurs perceptives que ne le font les difficultés d'ordre purement perceptif. L'oreille n'est pas un traître, elle est surtout asservie à la recherche de signification.[38]

L'activité langagière est également une activité interprétative, si bien que les interlocuteurs dont l'échange fait l'objet d'une transcription par une tierce personne sont eux-mêmes pris dans le même type d'activité, au sein de leur échange. Ceci ne doit pas être de nature à nous faire considérer la transcription comme un miroir de l'échange. Si l'on peut parler de miroir, il s'agira toujours d'un miroir déformant, permettant à un observateur d'étudier un échange de l'extérieur et *a posteriori*. C'est le sens de la remarque de R. Coppieters, pour qui toute transcription est une interprétation de la réalité :

> La difficulté fondamentale posée par le corpus reste bien sûr celle de son interprétation. D'une part, tout corpus donné en termes linguistiques, qu'ils soient phonétiques ou lexicaux, représente déjà une interprétation de la réalité, ce qui n'est pas un reflet direct de la réalité à l'état « brut », et en particulier ce n'est pas la réalité telle que les locuteurs la comprenaient au fur et à mesure de leur interaction [...]. C'est toujours de la réalité telle qu'elle a été organisée et comprise par après, presque toujours par d'autres locuteurs.[39]

Les transcriptions que nous proposons dans le volume 2 sont le fruit d'un minutieux travail d'écoute. Elles représentent les échanges enregistrés de la façon la plus

[37]BLANCHE-BENVENISTE, Claire & JEANJEAN, Colette (1987). *Le français parlé : transcription et édition.* Paris : Didier.

[38]BILGER, Mireille *et al.* (Groupe d'Etude sur les Données Orales) (1997). Transcription de l'oral et interprétation. Illustration de quelques difficultés.*Recherches sur le français parlé* 14 : 57-86. Aix-en-Provence : Publications de l'Université de Provence. (p.58).

[39]COPPIETERS, René (1997). Quelques réflexions sur la question des données : corpus et intuitions. *Recherches sur le français parlé* 14 : 21-41. Aix-en-Provence : Publications de l'Université de Provence. (p.35).

fidèle possible. Cependant, elles ne prétendent absolument pas être parfaites. Nous signalons les éléments inaudibles et douteux, et renvoyons le lecteur aux enregistrements disponibles sur le CD-ROM joint.

3.2.2 Etablissement des courbes intonatives

Complément nécessaire de la transcription, les courbes intonatives de notre corpus primaire figurent dans le second volume de ce travail. L'établissement de courbes utilisables passe par le choix d'un logiciel adapté. Il s'agit pour nous d'obtenir le tracé de courbes des valeurs de la fréquence fondamentale et de l'intensité en fonction du temps. A partir de ces tracés, il nous faut ensuite faire correspondre ce qui apparaît dans notre transcription avec les phénomènes intonatifs dont les courbes rendent compte.

3.2.2.1 Choix du logiciel d'analyse prosodique

La recherche en prosodie se développant, le choix de logiciels adaptés est relativement large. Nous avons pu comparer les avantages et inconvénients de quatre logiciels utilisés dans différents laboratoires de recherche, et avons arrêté notre choix sur l'un d'entre eux pour plusieurs raisons. Les logiciels que nous avons comparés sont Anaproz, WinSnoori, Praat et WinPitch.[40]

Nous aurions pu choisir de travailler avec le logiciel Anaproz, conçu par F. Colombo (ingénieur en automatique, spécialiste du dialogue Homme-Machine). Ce logiciel non commercialisé a été mis au point spécifiquement pour la recherche en prosodie. Il a été utilisé par plusieurs chercheurs ayant récemment soutenu leur thèse.[41] C'est également

[40] Les avantages et inconvénients comparés des quatre logiciels sont résumés dans le Tableau A.1 figurant en Annexe A (p.438).

[41] On pense par exemple, à l'Université Paris III, à M. Candea et à J. Szlamowicz.
CANDEA, Maria (2000). Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané. Thèse de doctorat. Université Paris III – Sorbonne Nouvelle.

le logiciel qui a servi à tracer les courbes intonatives apparaissant dans la *Grammaire de l'intonation* de M.-A. Morel et L. Danon-Boileau.[42]

Développé par les ingénieurs de BaBel technologies et de l'équipe CNRS Loria, WinSnoori est destiné à la recherche en phonétique et en intonation. Il est utilisé par certains doctorants de l'Université Paris V et de l'ILPGA à l'Université Paris III, et par des chercheurs de l'IRIT à l'Université Toulouse III.[43]

Praat a été mis au point par P. Boersma pour ses propres recherches à l'institut de phonétique de l'Université d'Amsterdam. Distribué gratuitement en ligne,[44] il est utilisé par exemple au Laboratoire Parole et Langage à l'Université d'Aix-Marseille et par une grande communauté de phonéticiens et de linguistes à travers le monde. Nous avons eu ponctuellement recours à ce logiciel notamment pour l'établissement des spectrogrammes présentés dans ce travail.

Enfin, WinPitch est le logiciel avec lequel nous avons commencé à travailler dans le cadre de notre DEA et sur lequel notre choix s'est établi pour le présent travail.[45] Conçu par P. Martin[46] pour ses propres recherches sur la prosodie du français, il s'est révélé le plus adapté à nos besoins.[47]

Sont entrés en considération plusieurs critères importants à différentes étapes de notre travail. Outre la facilité d'installation, l'ergonomie et la maniabilité du logiciel et des données qui en sont issues ont été examinées. Etant donné que notre souhait de

SZLAMOWICZ, Jean (2001). Contribution à une approche intonative et énonciative du rôle des ligateurs dans la construction du discours en anglais oral spontané. Thèse de doctorat Université Paris III – Sorbonne Nouvelle.

[42]MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *op. cit.*

[43]Une version gratuite de démonstration de ce logiciel payant est disponible en ligne à l'adresse suivante :
http://www.loria.fr/∼laprie/WinSnoori/
C'est la seule version à laquelle nous ayons eu accès.

[44]Pour télécharger le logiciel Praat : http://www.fon.hum.uva.nl/praat.

[45]C'est également le logiciel utilisé par Elodie Vialleton. Il nous a paru important d'utiliser le même logiciel afin de travailler sur des documents compatibles.

[46]Université Paris VII

[47]La version professionnelle WinPitch Pro, récemment sortie, offre des fonctionnalités très intéressantes. Une version gratuite de démonstration de WinPitch Classic est disponible à l'adresse suivante : http://www.winpitch.com.

montrer que certaines variations de temps, d'intensité et de hauteur mélodique sont linguistiquement signifiantes, il nous fallait opter pour un logiciel traitant ces trois dimensions (la plus rare étant l'intensité). L'affichage du spectrogramme ne constituait pour nous qu'un critère secondaire. Afin de nous permettre de traiter des épisodes conversationnels longs, le logiciel devait également supporter des fichiers de taille importante tout en gardant une précision d'analyse suffisante. La souplesse des paramètres de calcul, de mesure et d'affichage des résultats est un critère important, dans la mesure où elle permet de corriger des erreurs éventuelles dues au caractère inadapté des paramètres établis par défaut. Enfin, la possibilité d'imprimer les tracés intonatifs, nécessaire pour présenter notre corpus sur un support visuel, a constitué un dernier critère indispensable.

3.2.2.2 Contraintes et choix techniques

Pour faciliter le traitement informatique des corpus analysés avec WinPitch Classic, nous avons procédé à un découpage du continuum sonore en séquences de 30 secondes à 2 minutes 30. Les titres donnés à la cinquantaine d'épisodes du corpus principal ainsi délimités et à la centaine d'autres appartenant au corpus secondaire reflètent pour la plupart le thème abordé par les locuteurs ou simplement un trait saillant du passage. Certains choix techniques dans la manipulation du logiciel ont dû être faits afin que celui-ci nous fournisse des données analysables et visualisables conformément à nos objectifs de recherche. Ne faisant pas l'objet d'un calcul, la mesure de l'intensité est moins dépendante de tels choix que la fréquence fondamentale. Ces choix se fondent sur une connaissance des caractéristiques acoustiques du matériau sonore[48] ainsi que sur une évaluation des moyens nécessaires aux objectifs que nous nous sommes fixés dans le cadre de ce travail.

P. Martin propose plusieurs algorithmes pour le calcul de F_0 par WinPitch : Fast

[48] Voir p. 71.

Comb, Classic Comb, et AMDF. Après avoir comparé l'aspect visuel de la courbe en fonction de l'algorithme adopté, nous nous sommes fixée sur Fast Comb, qui allie rapidité et précision, sans pour autant être trop sensible aux plosives et aux sifflantes, par exemple, qui font subitement monter la valeur de F_0.

S'est également rapidement posé le problème de l'échelle des valeurs de F_0 sur lesquelles doit porter le calcul. Les paramètres par défaut n'étant pas pleinement satisfaisants pour embrasser toute la tessiture des femmes (en particulier pour des valeurs élevées) comme celle des hommes (pour des valeurs très basses), nous avons étendu le champ de l'analyse afin d'obtenir des valeurs de F_0 de 70 à 600 Hz.

La représentation visuelle des données prosodiques concernant F_0 est le dernier point sur lequel nous souhaitons nous arrêter ici. Trois modes de représentation visuelle de la fréquence fondamentale sont disponibles dans WinPitch, correspondant à des échelles différentes. Nous avons éliminé rapidement l'échelle de représentation musicale sur deux portées. C'est bien le mode de représentation, peu pratique et trop influencé par la tradition solfégique, qui est en cause ici, et non pas l'échelle musicale en soi. En effet, celle-ci présente l'intérêt de calquer la perception humaine de F_0, les demis-tons et quarts de tons étant répartis sur une échelle logarithmique. Après avoir envisagé d'utiliser précisément l'échelle de représentation logarithmique non musicale offerte par WinPitch (ce que nous avions fait dans notre DEA), nous avons finalement opté pour une représentation selon une échelle linéaire. La représentation selon une échelle logarithmique facilite les comparaisons entre variations de l'intensité et variations de F_0 puisque l'échelle des décibels est une échelle logarithmique. En outre, elle donne une vision plus fidèle de ce qui, dans les variations de F_0, est effectivement perçu par l'oreille humaine (les valeurs basses sont favorisées, par rapport aux valeurs élevées). Néanmoins, les courbes utilisant l'échelle logarithmiques sont assez peu lisibles. C'est donc dans un souci de facilité de lecture que nous avons choisi l'échelle de représentation linéaire. Nous garderons toujours à l'esprit que la percep-

tion des faits prosodiques reste logarithmique et que notre lecture des courbes de F_0 doit en tenir compte, ainsi que nous l'avons souligné plus haut.[49] Nous avons montré que les différences entre les deux échelles sont minimes pour les valeurs qui nous intéressent. Il reste cependant possible et souhaitable, sur certains points précis, proches de préoccupations psycholinguistiques, de travailler sur les données chiffrées issues du logiciel en leur appliquant un coefficient particulier de façon à rendre leurs variations logarithmiques.

3.2.2.3 Lecture des courbes

Les courbes d'analyse prosodique présentées dans le second volume de ce travail suivent toutes le même modèle. Chaque vignette représente quatre secondes de parole. Chaque page propose deux graphiques par locuteur, soit quatre en tout pour les conversations : sont ainsi mis en regard les courbes de chaque locuteur pour les 8 mêmes secondes de parole. Quelques décalages sont malheureusement à noter, du fait d'une légère différence de vitesse de déroulement des cassettes analogiques sur lesquelles les enregistrements ont été faits. Il est donc conseillé de se reporter au texte de la transcription également présenté dans le volume 2 pour suivre avec précision le déroulement de l'interaction. Comme le montre la Figure 3.1, en haut de chaque vignette se situe le texte de la transcription, aligné manuellement sur la ligne du temps. Au milieu, la courbe de F_0 est présentée selon une échelle linéaire de 0 à 500 ou 600 Hz en fonction des locuteurs. La courbe d'intensité est inscrite dans la partie inférieure du graphique, sur une échelle logarithmique de 0 à 40 dB. Nous avons délibérément omis les tracés du spectrogramme et du signal sonore qui n'apportent pas, il nous semble, de renseignements supplémentaires utiles à notre étude.

Quelques précautions doivent être prises au moment de la lecture des courbes. Il convient de ne pas accorder trop d'importance à des phénomènes qui, d'un point de

[49] Voir p. 71.

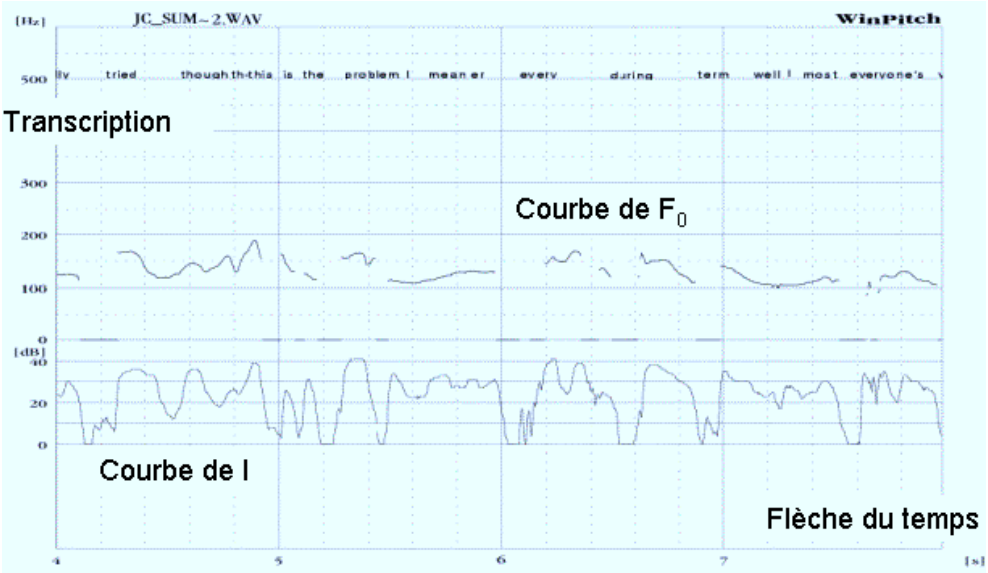

FIG. 3.1 – Courbes prosodiques produites par WinPitch Classic

vue perceptif et donc interprétatif au niveau qui nous intéresse, n'en ont pas. C'est le cas des variations micromélodiques dont la pertinence est pour nous négligeable. Une variation d'une durée inférieure à celle d'une syllabe est généralement à ne pas prendre en compte. C'est davantage la forme générale de la courbe qui nous intéresse. Cela permet également de se prémunir contre certaines erreurs de calcul du logiciel. Si l'algorithme Fast Comb y est moins sensible, il peut apparaître sur la courbe de F_0 de grands décrochements ponctuels causés par la prononciation d'une plosive ou d'une sifflante. Il convient donc de procéder à une sorte de lissage mental des courbes afin de ne prêter attention qu'aux phénomènes linguistiquement signifiants, correspondant aux niveaux de pertinence retenus dans le cadre de notre recherche.[50]

3.2.3 Traitement informatique et réflexion linguistique

Le débat sur la subjectivité des sources touche l'étude de la langue orale pour plusieurs raisons. Nous avons évoqué plus haut la question de la représentativité d'un cor-

[50]Voir p.121.

pus, qui n'est jamais qu'un échantillon du réel qui n'a pas pour vocation de l'épuiser. Nous avons également traité de la nécessaire interprétation de l'échange par le transcripteur. Venons-en à présent au traitement informatique, aux mesures et aux calculs des phénomènes intonatifs effectués par la machine. Si cette partie du travail semble la plus objective, elle a néanmoins ses limites, que la subjectivité du linguiste doit être en mesure de rectifier. Quelle place accorder aux données issues d'un traitement informatique objectif par rapport à celles issues de notre écoute subjective et de notre expérience de la langue ?

3.2.3.1 Apport du traitement informatique

L'utilisation d'outils spécifiques pour l'étude des phénomènes suprasegmentaux continus s'est développée de façon plus tardive que pour celle des phonèmes. Les outils numériques ont succédé aux outils analogiques et sont devenus incontournables. L'analyse expérimentale, effectuée à l'aide d'outils informatiques, constitue pour nous une grande avancée dans la méthode d'établissement des données sur lesquelles les analyses linguistiques se fondent. Elle assure une plus grande fiabilité et permet une plus grande souplesse dans la manipulation des données.

Le désavantage majeur du seul recours à l'oreille dans l'établissement des tracés intonatifs provient du fait que l'oreille est un instrument facilement influencé. Soumise au besoin de sens, elle est d'autant plus susceptible de chercher à percevoir des schémas attendus. En quelque sorte, on n'entend que ce que l'on veut bien entendre. Source potentielle d'erreurs au moment de la transcription orthographique, l'oreille l'est également dans la perception des phénomènes intonatifs lorsqu'il s'agit de détecter la part de chaque paramètre et son évolution en fonction du temps. Mentionnons par exemple le cas des pauses subjectives : l'oreille humaine a tendance à percevoir une pause dans l'émission de la parole à certaines charnières mélodiques (généralement en fin de segment, lorsque la ligne mélodique baisse avant de remon-

ter sur le segment suivant). Une confrontation avec la représentation du seul signal acoustique permet d'infirmer cette impression.

Il est probable que la grande attention accordée aux phénomènes mélodiques par les études les moins récentes ait en partie pour origine le fait que l'oreille ait été le premier outil de mesure de l'intonation. En effet, la mélodie étant sans doute le paramètre le plus perceptible, c'est à elle que l'on s'intéresse en premier. C'est au bénéfice de la mélodie, dans un souci de quantification, qu'ont été utilisés les premiers instruments dédiés à l'analyse prosodique. Les autres paramètres, s'ils ont été identifiés, ont longtemps été ignorés.

Le recours à une analyse paramétrique expérimentale dans le cadre d'un traitement informatique des données acoustiques permet, en les dissociant, d'accorder une attention égale à davantage de paramètres suprasegmentaux.

3.2.3.2 Limites du traitement informatique

Le traitement informatique des données ne nous met cependant pas totalement à l'abri d'erreurs. Les limites d'une telle démarche rejoignent celles qui caractérisent toute entreprise d'analyse aveugle, automatisée par la machine. La fiabilité des données fournies par le logiciel n'étant pas à toute épreuve, l'écoute attentive et le jugement linguistique resteront toujours nos premiers guides.

Outre les erreurs de calcul dues à l'explosion ou au sifflement d'une consonne, certains phénomènes nécessitent une attention particulière. Une erreur courante est celle du saut d'octave : ne disposant pas de suffisamment de place sur la grille du graphique pour tracer la courbe mélodique sur toute la tessiture du locuteur (ou plus généralement de la locutrice), le logiciel traite les segments situés au-delà de la valeur maximale comme s'ils étaient prononcés une octave plus bas. Dans l'exemple suivant, une simple écoute a permis de détecter l'erreur commise par le logiciel et de rectifier la bande sur laquelle est représentée F_0. Comme le montre la Figure 3.2, le changement

d'échelle de 500 à 600 Hz fait du segment de gauche la partie la plus aiguë de l'énoncé et non pas la plus grave.

H how did you learn to do all this

(Mike&Hannah : River)

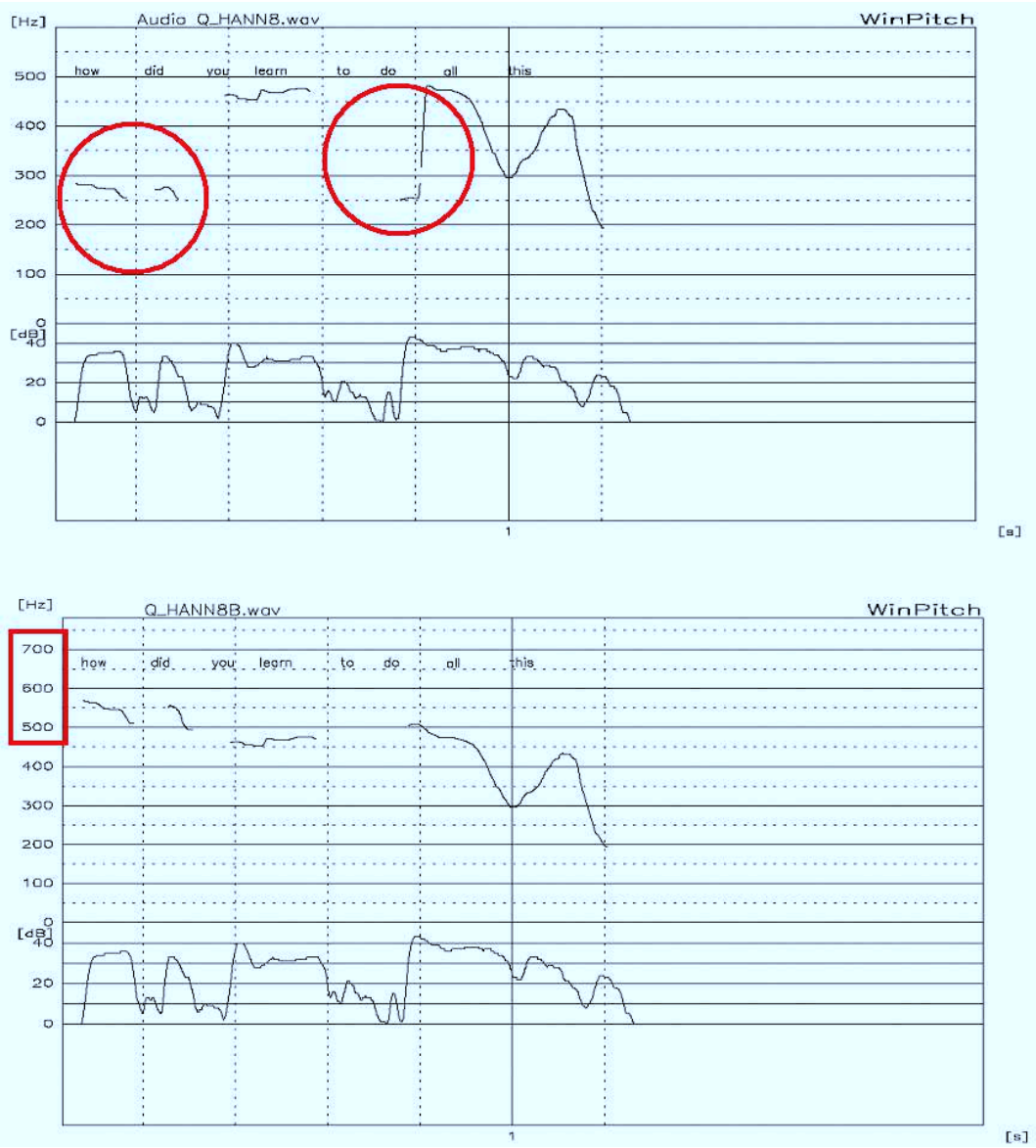

FIG. 3.2 – Erreur du logiciel (en haut) corrigée après écoute (en bas)

Comme nous le verrons pour une éventuelle détection automatique des pauses, la caution de la perception humaine et du jugement linguistique est partout souhaitable.

3.2.3.3 Bilan concernant le traitement informatique des données prosodiques

Achevons ce chapitre en réaffirmant que les données issues d'une analyse expérimentale si poussée soit elle ne sont pas exploitables telles quelles : elles doivent nécessairement être analysées et interprétées d'un point de vue linguistique pour devenir signifiantes pour nous. Pourtant, les limites au traitement automatique des données ne nous semblent pas constituer une raison suffisante pour ne pas y recourir. Nous ne partageons pas l'opinion exprimée par P. Garvin & M. Mathiot[51] dans leur article de 1958 qui préfèrent se passer de tout instrument de mesure et de calcul, justifiant leur position par une référence à un écrit de F. Daneš paru l'année précédente que nous reproduisons ici :

> It is true that instruments are more accurate and sensitive than the human ear [...], but instrumental records and their interpretation in terms of physical acoustics do not give us a true picture of the way in which the speakers hear and understand (evaluate) their own language [...].
>
> [...]
>
> The most complete instrumental record becomes useful to linguists only after analysis and evaluation from the standpoint of the function and significance of the various components in terms of the system of the language. The instrumental record, however, does not by itself allow us to ascertain the significance and function of the various waves, formants, etc.—these must first be discovered, albeit only in outline, by an auditory analysis of spoken language [...].[52]

La position de F. Daneš, loin d'autoriser l'abandon d'une analyse instrumentale, nous semble simplement signaler l'importance du travail interprétatif de linguiste. L'oreille

[51] GARVIN, Paul L. & MATHIOT, Madeleine (1958). Fused units in prosodic analysis. *Word* 14 : 178-186. La citation de F. Daneš se trouve dans une note de bas de page (pp.179-180).

[52] DANEŠ, František (1957). *Intonace a věta ve spisovné češtine.* Prague. (pp.39-41).

et l'expérience de la langue et des échanges verbaux doivent rester nos principaux outils d'analyse, épaulés par la machine.

Si le traitement informatique permet la quantification précise de certains phénomènes, il nous place devant une masse d'informations qu'il nous faut apprendre à trier. D. Crystal fait remarquer ici la nécessité du filtre de l'oreille et de la subjectivité humaine afin de rendre utilisables les données de l'analyse instrumentale :

> But even with simpler, speedier, cheaper, more sensitive and reliable instruments to analyse a corpus of speech than at present, the results would still be of limited value in trying to reach any understanding of the meaning of such vocal effects as intonation and other prosodic features when perceived by the listener, for the obvious reason that the instrumental analyses produce pictures of speech which are too sensitive to detail to provide any clear pattern.[53]

Ces données sont donc sans valeur en dehors de ce que le linguiste peut leur attribuer en tant qu'animal parlant et entendant.

Ainsi semble être réhabilitée la subjectivité du chercheur dans un domaine qui, bien que s'appuyant sur des analyses instrumentales, reste fortement ancré dans la tradition des études linguistiques. La subjectivité liée à l'expérience de la langue est un critère positif, car elle donne accès au système de signification et de représentation symbolique dans lequel la conversation spontanée s'inscrit.

Au terme de cette première partie où se sont dégagés les principes généraux adoptés dans ce travail, nous souhaitons revenir sur les rapports entre un corpus et la réalité dont il n'est qu'un échantillon. D'après les mots de H. Frei,

> [...] le discours, qui correspond aux rapports syntagmatiques ou *in praesentia* de Saussure, et qui ne peut, ni fonctionner ni être interprété par le linguiste sans les rapports associatifs ou *in absentia* correspondants, est une manifestation

[53] CRYSTAL, D. (1969). *Prosodic Systems and Intonation in English.* Cambridge : CUP. (p.13).

> du système linguistique, mais il n'en est jamais qu'une manifestation partielle. Il est donc possible que deux *corpus* différents, c'est-à-dire deux collections de matériaux prélevés dans le discours, reflètent le même système. Or, le but essentiel du linguiste n'est pas l'étude exclusive du *corpus*, qui ne pourrait aboutir qu'à un inventaire, mais celle du système à travers le *corpus*.[54]

De même, nos hypothèses de départ n'ont pas l'ambition de nous permettre d'épuiser notre objet d'étude. Ce ne sont que des pistes ouvrant des perspectives, que nous espérons prometteuses, pour l'analyse de quelques phénomènes de structuration de l'anglais oral spontané tel qu'il se manifeste dans notre corpus. Nous proposons d'explorer deux domaines où s'illustre la hiérarchisation des constituants discursifs : le domaine du suprasegmental, et, en premier lieu, celui du segmental.

[54]FREI, Henri (1954). Critères de délimitation. *Word* 10 (2) : 136-145. (p.138)

Deuxième partie

De quelques marqueurs segmentaux de hiérarchisation

La hiérarchisation des constituants discursifs à l'oral s'appuie largement sur l'ordre des mots tel qu'il est déterminé par les structures syntaxiques de la langue. Celles-ci organisent en effet le discours en unités dépendantes et unités indépendantes et les disposent dans le linéaire de surface. C'est ainsi que les unités propositionnelles sont organisées dans le cadre de la relation de subordination syntaxique.

Cette deuxième partie de notre travail s'intéresse à certains faits de hiérarchisation induits par la subordination syntaxique. L'une de nos hypothèses de départ postule que l'oral spontané est caractérisé par une syntaxe différente de celle de l'écrit. En quel sens les éléments subordonnés syntaxiquement sont-ils hiérarchisés ? En quoi l'oral spontané illustre-t-il un fonctionnement spécifique de la subordination ?

Afin d'apporter des réponses à ces questions, nous proposons de commencer par établir quelques éléments pour une analyse distributionnelle des marqueurs de subordination dans notre corpus. De cette première étude émerge la catégorie sémantique de la subordination causale, catégorie la plus répandue, dont nous étudions ensuite en détail le fonctionnement des marqueurs privilégiés : *because* et *cos*. L'analyse d'un certain nombre d'exemples de notre corpus suggère enfin une remise en question de la pertinence d'une description syntaxique traditionnelle de la subordination dite causale, au profit d'une analyse davantage sémantique. S'illustre, selon nous, le caractère fractal de l'anglais oral, conformément à notre hypothèse de départ, qui voit un principe d'organisation syntaxique s'appliquer au domaine de l'organisation discursive.

Chapitre 4

Éléments pour une analyse distributionnelle des marqueurs de subordination

Le procédé syntaxique de subordination d'une proposition à une autre et les marqueurs de subordination (en particulier les conjonctions) ont fait l'objet d'études nombreuses et variées dans l'histoire de la linguistique. L'intérêt pour la subordination provient à la fois de la préoccupation pour classer les mots, dont les conjonctions de subordination, au sein des parties du discours et de celle, plus tardive, d'étudier la façon dont se combinent les propositions au sein de la phrase. Ce chapitre cherche tout d'abord à établir une définition de travail concernant l'opération syntaxique et sémantique de subordination (et en particulier la subordination circonstancielle), à partir de la tradition linguistique. Il s'intéresse ensuite aux caractéristiques distributionnelles relatives aux propositions subordonnées et aux marqueurs de subordination à l'oral spontané.

Notre objectif est multiple : il s'agit tout d'abord de délimiter quelles possibilités sont offertes aux locuteurs, puis d'évaluer dans quelle mesure elles sont mises à pro-

fit. Nous souhaitons montrer à cette occasion que des tendances générales fortes se dégagent (de notre corpus comme des corpus de référence) et formuler deux hypothèses permettant de les expliquer.

4.1 De la subordination : définition de travail

Les constituants phrastiques peuvent s'organiser de deux façons différentes. Ils peuvent être mis sur un pied d'égalité : on parle alors de parataxe. Au contraire, ils peuvent entretenir une relation asymétrique, hypotaxique. Dans ce dernier cas, qui est celui dont la subordination relève, l'un des deux constituants est syntaxiquement dépendant de l'autre. La subordination se définit donc en anglais contemporain par opposition à la coordination, forme de mise en relation parataxique. Néanmoins, la distinction n'est pas toujours aisée à reconnaître, notamment pour les marqueurs de cause (*for* et *because*, par exemple). La perméabilité de la frontière entre coordination et subordination s'explique probablement par des origines communes que révèle la tendance parataxique du vieil anglais. Nous y reviendrons plus loin.[1]

4.1.1 Définition consensuelle

Si chaque étude sur la subordination adopte une perspective particulière, une définition minimale en émane, fondée sur les critères syntaxiques et sémantiques permettant de parler de subordination.

4.1.1.1 Caractéristiques syntaxiques générales

Les constituants phrastiques subordonnés peuvent s'insérer dans l'arbre hiérarchique de l'énoncé qui les accueille à toutes les places qu'un groupe nominal y est susceptible d'occuper :

[1] Voir p.285.

> Subordinate clauses may function as subject, object, complement, or adverbial in a superordinate clause. [...] In addition, subordinate clauses may function within these elements: postmodifier in noun phrase [...], prepositional complement [...], adjectival complementation [...].[2]

Ainsi, en fonction de la place occupée par la subordonnée dans la matrice, R. Quirk *et al.* distinguent quatre types de propositions subordonnées : nominalisées, circonstancielles, relatives et comparatives.

Dans le cadre de ce travail, nous ne nous intéresserons qu'aux propositions subordonnées circonstancielles, c'est-à-dire à celles qui occupent la place d'un groupe à fonction adverbiale dans la matrice. Plusieurs critères syntaxiques témoignent de leur dépendance vis-à-vis de cette matrice :

> Subordination is generally marked by a signal in the subordinate clause. The signal may be of various kinds:
>
> (i) The clause is initiated by a subordinating conjunction [...].
>
> (ii) The clause is initiated by a wh-element [...].
>
> (iii) Initial elements in the clause are inverted [...].
>
> (iv) The presence of certain verb forms in finite clauses is determined by the type of subordinate clause [...].
>
> (v) The verb element of the clause is either nonfinite or absent [...].
>
> More than one subordination signal may cooccur in the same subordinate clause.[3]

Ces marques de subordination ont donc trait à l'ordre des mots, à la forme (finie ou non finie) du verbe (et à son absence éventuelle), ou à la présence d'un mot introducteur en tête de subordonnée (conjonction ou mot en *wh*-). Au sein des propositions

[2] QUIRK, Randolph *et al.* (1985). *A Comprehensive Grammar of the English Language.* Londres : Longman. (p.1047).

[3] *ibid.* (p.997).

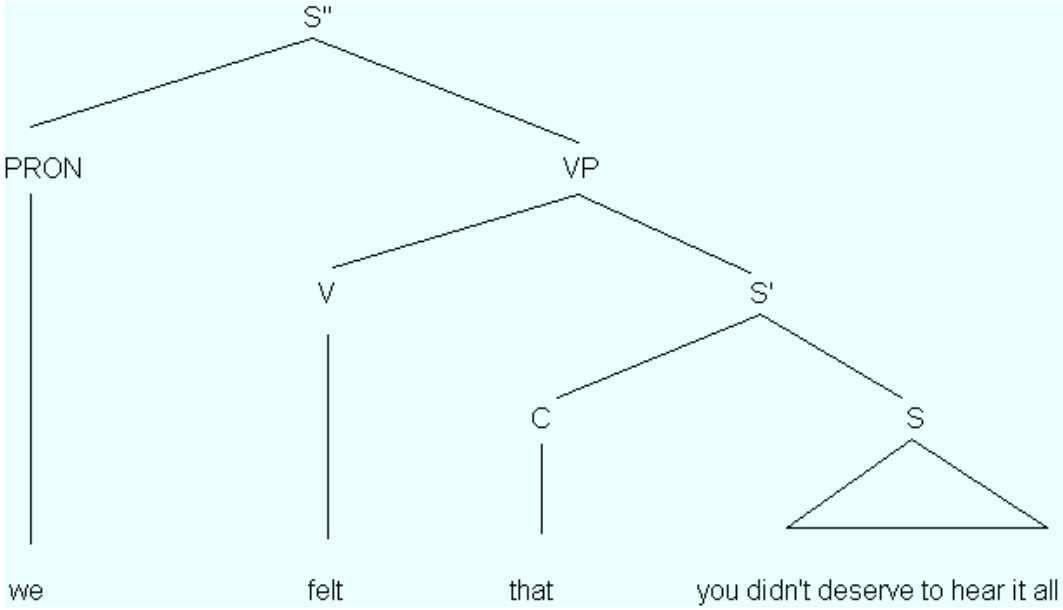

FIG. 4.1 – Exemple de représentation hiérarchique de la subordination complétive en grammaire générative

subordonnées circonstancielles c'est à celles introduites par une conjonction que nous nous souhaitons porter notre attention ici.

En ce qui concerne les relations syntaxiques entre les deux propositions, nous avons vu que les théories s'accordent sur la dimension hypotaxique, donc hiérarchique, du procédé de subordination.

> Du point de vue du *statut* syntaxique des deux structures de phrase en jeu, il résulte donc de ce processus que la proposition q (l'enchâssée) est « rétrogra -dée » à un « rang inférieur », cf. notions de *rank shift* (Halliday), ou encore de « translation » au sens de Tesnière.[4]

Cette hiérarchisation transparaît graphiquement dans l'arbre syntaxique de l'exemple suivant :

R we felt that you didn't *deserve to hear it all*

(Helen&Rebecca : Nîmes1)

[4]WYLD, Henry (2001). *Subordination et énonciation.* Gap : Ophrys. (p.9n).

La Figure 4.1, conservant la terminologie anglosaxonne de l'ouvrage d'introduction à la grammaire générative de L. Haegeman et J. Guéron,[5] montre que la complétive occupe la place de l'objet dans la phrase minimale, et se situe donc à une place hiérarchiquement dépendante par rapport à celle-ci.

4.1.1.2 Sémantisme de la circonstance

Les propositions subordonnées circonstancielles font traditionnellement l'objet d'un classement de type sémantique en fonction de la circonstance que leur contenu indique. Ce classement n'offrant pas de stricte correspondance avec une tentative similaire opérée sur les adverbes, on distingue ainsi les circonstancielles de temps, de cause, de raison, de concession, de condition, de manière, de but, et de conséquence.[6] Cependant, ce type de classification, longue et toujours incomplète, ne fait pas l'unanimité parmi les chercheurs : certains en viennent à mettre en question la validité d'une perspective typologique. C'est le cas de S. Rémi-Giraud qui relève un si grand nombre de configurations qu'elle aboutit au constat du manque d'efficacité d'une telle typologie.[7]

En dépit de la multitude des configurations, J. Feuillet propose au contraire une typologie réduite à trois catégories : les circonstancielles spatiales, temporelles, et notionnelles.[8] Les circonstancielles notionnelles, qui nous intéressent particulièrement ici, se subdivisent en trois sous-types (propositions dépendantes préparatoires, dépendantes complémentaires et dépendantes préparées), qui recouvrent nombre de types sémantiques différents :

> Quand on parle de dépendantes préparatoires, complémentaires et préparées, on prend pour référence les « principales ». Les premières sont constituées des

[5]HAEGEMAN, Liliane & GUERON, Jacqueline (1999). *English Grammar : a generative perspective*. Oxford : Blackwell. (p.100).

[6]L'analyse traditionnelle de l'anglais ne reconnaît pas de subordonnées circonstancielles de lieu.

[7]REMI-GIRAUD, Sylvianne (1988). Le complément circonstanciel : problèmes de définition. in Rémi-Giraud, S. & Roman, A. *Autour du circonstant* : 65-113. Lyon : Presses Universitaires de Lyon.

[8]FEUILLET, Jack (1992). Typologie de la subordination in Chuquet, Jean & Roulland, Daniel (éds.). *Travaux linguistiques du CERLICO* 5 : 7-28. Rennes : Presses Universitaires de Rennes.

causales (la raison explique l'existence de la principale), des conditionnelles (la raison conditionne l'existence de la principale) et des concessives (la raison n'empêche pas l'existence de la principale, car elle est la négation d'une implication). Les complémentaires comprennent les instrumentales, les sociatives, les adversatives, les comparatives et les substitutives. Les préparées regroupent les finales, les consécutives et les négatives (forme de neutralisation).[9]

Pour contesté qu'il soit, un classement de type sémantique ne manque pas de répertorier quelques grandes catégories de circonstancielles dont les différents auteurs s'accordent à reconnaître l'existence.

4.1.2 Incidence des subordonnées circonstancielles

Plus théorisé est le débat concernant le rapport entre matrice et subordonnée circonstancielle analysé du point de vue de l'incidence de cette dernière. Il est en effet tributaire de la perspective théorique adoptée.

4.1.2.1 Incidence syntaxique de la subordonnée

L'exemple reporté plus haut à la Figure 4.1 illustre l'enchâssement d'une proposition dans le cadre de la complémentation verbale analysée par la théorie X-barre. Où s'insère la proposition circonstancielle enchâssée dans l'arbre syntaxique ? Quelle incidence celle-ci a-t-elle sur la matrice ?

R. Quirk *et al.* répertorient quatre fonctions syntaxiques pour les groupes adverbiaux (dont les circonstancielles instancient la place) : les adjoints (*adjuncts*), subjoints (*subjuncts*), disjoints (*disjuncts*) et conjoints (*conjuncts*).[10] Seules deux catégories sont en fait pertinentes pour les circonstancielles : les adjoints et les disjoints.

[9] *ibid.* (p.16).
[10] QUIRK, Randolph *et al.* (1985). *op. cit.* (p.1068).

Les subordonnées disjointes se distinguent des adjointes par leur statut périphérique, non restrictif, vis-à-vis de la principale. Cette différence se manifeste dans ma possibilité d'appliquer des transformations syntaxiques aux adjointes, que les disjointes n'acceptent pas :

(i) Only the adjunct clause can be the focus of a cleft sentence

(ii) Only the adjunct clause can be the focus of a variant of the pseudo-cleft sentence

(iii) Only the adjunct clause can be the focus of a question [...]

(iv) Only the adjunct clause can be the focus of negation [...]

(v) Only the adjunct clause can be focused by focusing subjuncts [...] such as only, just, simply, and mainly.

(vi) Only the adjunct clause can be the response to a wh-question formed from the matrix clause.[11]

propositions restrictives acceptent donc la mise en relief contrastive.

La réalisation intonative des propositions est liée à leur incidence par rapport à la matrice. En position finale, les propositions adjointes tendent ainsi à être intégrées intonativement à l'énoncé complexe, contrairement aux propositions disjointes :

> In final position, content disjunct clauses tend more commonly than adjunct clauses to be separated from their matrix clauses by intonation or punctuation, though punctuation usage varies. [...] In initial position, all adverbial clauses, regardless of form or function, are separated by intonation and (usually) by punctuation from their matrix clause. [12]

[11] *ibid.* (p.1071).
[12] *ibid.* (p.1072).

4.1.2.2 Incidence sémantique de la subordonnée

Analyse traditionnelle La question de l'incidence de la subordonnée sur la principale sur la plan sémantique est envisagée par R. Quirk *et al.*. Les adjointes se rapportent aux circonstances internes du procès décrit dans p, tandis que les disjointes fournissent un commentaire sur la forme de p ou sur son contenu :

> Adjuncts and disjuncts tend to differ semantically in that adjuncts denote circumstances of the situation in the matrix clause, whereas disjuncts comment on the style or form of what is said in the matrix clause (style disjuncts) or on its content (content or attitudinal disjuncts).[13]

Le recours à la sémantique permet encore de distinguer, au sein des propositions disjointes, des propositions plus ou moins périphériques par rapport à la principale. Les propositions circonstancielles apportant une justification de l'énonciation en en décrivant les circonstances (*style disjuncts*) sont ainsi plus détachées de la principale que les propositions commentant le contenu de celle-ci (*content disjuncts*). D'un point du vue intonatif, d'après les auteurs de la grammaire, les propositions métadiscursives sont toujours séparées de la principale.[14]

Analyse énonciativiste Dans son ouvrage de 2001, H. Wyld[15] remet en cause la vision traditionnelle de l'enchâssement des propositions circonstancielles, qu'il considère tributaire d'une certaine analyse de la structure de base de la phrase. Il préfère à celle-ci l'énoncé canonique d'A. Culioli, soit le produit de l'opération $< \ \lambda \underline{\in} Sit \ >$, ou encore $< \ arb \underline{\in} Sit \ (S, T) \ >$, qui permet de considérer à la fois le niveau prédicatif, mais aussi énonciatif de l'énoncé.[16]

[13] *ibid.* (p.1070).
[14] Voir *ibid.* (p.1073).
[15] WYLD, Henry (2001). *op. cit.*
[16] Dans ce schéma, λ représente le contenu propositionnel, $\underline{\in}$ un opérateur de localisation, Sit la situation d'énonciation. Le contenu propositionnel se compose de a et b (premier et deuxième arguments), reliés par le relateur r. La situation d'énonciation comprend quant à elle une coordonnée T

Dans le cadre de la sémantique formelle de la théorie des opérations énonciatives, il propose une analyse de la subordination conçue

> [...] comme phénomène d'enchâssement opérant de telle sorte que la proposition enchâssée (q) soit envisageable comme étant incidente sur une des places constitutives du schéma abstrait associé à la structure matrice (p) ainsi conçue, et donc, le cas échéant, sur Sit (ou sur une de ses composantes).[17]

En posant le paramètre S comme site d'incidence possible, cette théorie permet d'expliquer avec précision la valeur modalisante de certaines subordonnées.

L'analyse sémantique se préoccupe également du statut respectif des segments reliés par le marqueur de subordination dans l'organisation thématique de l'énoncé complexe.[18] Evoquant le lien possible entre antéposition et thématisation, H. Wyld suggère que le statut thématique des propositions ne découle pas forcément de l'ordre dans lequel elles apparaissent. L'intonation permet en effet au repérage interpropositionnel de fonctionner indépendamment de l'ordre linéaire :

> [...] si au plus on peut dire qu'il y a affinité entre antépositionnement et statut de repère d'énoncé, il n'est pas rare qu'une circonstancielle jouant le rôle de repère d'énoncé puisse se trouver postposée (moyennant à l'oral un « réajustement » intonatif).[19]

Pour M. De Cola-Sekali, la relation de repérage provient simplement de la connexion des deux propositions. La chercheuse relève deux conséquences de cette relation sur la prise en charge de p qui nous intéressent particulièrement dans le cadre d'une étude sur la subordination causale. Tout d'abord, elle souligne le rôle sémantique d'agent ou instrument qu'assume la proposition repère vis-à-vis de la proposition

spacio-temporelle et une coordonnée subjective S. Voir CULIOLI, Antoine (1999). *Pour une linguistique de l'énonciation*. Gap : Ophrys. (Tome 2).

[17] WYLD, Henry (2001). *op. cit.* (p.11).

[18] Le terme d'organisation thématique fait ici référence à la polarisation de l'énoncé entre ce dont on parle et ce que l'on en dit.

[19] *ibid.* (p.71).

repérée (rappelant que la morphologie d'une conjonction de subordination comme *because* porte en elle la trace du cas instrumental du vieil anglais (*by cause*)) : par glissement ou « transfert », la prise en charge de q permet celle de p.[20] En corollaire à ce premier point, elle voit dans le statut de segment repéré de la principale le ferment de son inscription dans un projet argumentatif :

> L'origine de [la] prise en charge [de l'énoncé p] est **extériorisée** dans un autre énoncé. Or extérioriser la prise en charge d'un énoncé, [...] c'est [...] faire de sa prise en charge une prise en charge **contestée**, ou du moins a priori contestable.[21]

La dépendance syntaxique marquée par la conjonction de subordination induit une dépendance de prise en charge. La proposition principale, repérée, devient ainsi intersubjectivement dépendante de la proposition subordonnée causale qui est syntaxiquement dépendante d'elle. La subordonnée lui sert de repère pour la prise en charge de sa relation prédicative. Cette dépendance réciproque place la subordination causale au cœur de la problématique énonciativiste qui est la nôtre.

4.2 Eléments de distribution

Les possibilités offertes par la grammaire et le lexique d'une langue ne sont pas toutes exploitées à égalité par les locuteurs de cette langue. Une étude distributionnelle a pour vocation d'établir le contexte d'occurrence immédiat de certaines formes, mais également de les dénombrer afin de fournir un point de départ pour une analyse approfondie du discours. En effet, l'observation est un des fondements essentiels sur lesquels la théorie peut s'appuyer. C'est le sens de la réflexion des auteurs d'une étude descriptive du corpus LOB en introduction à leur ouvrage :

[20] Voir DE COLA-SEKALI, Martine (1991a). Connexion inter-énoncés et structuration des relations temporelles et argumentatives en anglais contemporain. Une étude énonciative des connecteurs polyvalents SINCE et FOR. Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (p.235).

[21] *ibid.* (p.235).

> To describe a language fully, we need to study not only the choices which are theoretically available but also the extent to which these options are actually utilized [...].[22]

D'un point de vue distributionnel, la grammaire de R. Quirk donne relativement peu d'indications sur la fréquence d'emploi des différents types de subordonnées. En revanche, la *Longman Grammar of Spoken and Written English* (LGSWE) de D. Biber *et al.*[23] offre de nombreuses données chiffrées exploitables dans le cadre d'une comparaison de tendances entre notre corpus et un corpus de grande taille. Le BNC nous servira également ici de point de comparaison.

4.2.1 Les subordonnées exprimant une circonstance

La LGSWE décrit les tendances rencontrées dans quatre types de textes différents : les textes de fiction, les bulletins d'information, les textes scientifiques ou académiques et la conversation. Les propositions circonstancielles, de même que les adverbes, se rencontrent de façon privilégiée dans le corpus de fiction. Cependant, elles ne sont bien entendu pas exclues des autres corpus, dont l'analyse révèle des tendances intéressantes. Le classement proposé par les auteurs de cette grammaire tient compte, au sein des circonstancielles, du sémantisme des propositions subordonnées. Ainsi, il apparaît que les conditionnelles et causales se retrouvent surtout en conversation, tandis que les subordonnées de but et de temps caractérisent davantage le corpus de fiction et d'information. Deux types de données nous intéressent ici : les données relatives comparant chaque type de corpus sur la base du nombre de circonstancielles relevé, et les données absolues établies pour chaque type sémantique de circonstancielle au sein de chaque corpus.

[22] JOHANNSON, Stig & HOFLAND, Knut (1989). *Frequency Analysis of English Vocabulary and Grammar.* Oxford : Clarendon Press.

[23] BIBER, Douglas *et al.* (1998). *Longman Grammar of Spoken and Written English.* Londres : Longman.

Nous retenons ainsi deux points importants. Tout d'abord, les propositions causales et conditionnelles sont les circonstancielles les plus répandues au sein du corpus de conversation de la LGSWE. Ensuite, sur le total des circonstancielles causales et conditionnelles relevées dans les quatre corpus étudiés, celles provenant du corpus de conversation sont les plus nombreuses (pour le même nombre de mots).

4.2.2 Les marqueurs de subordination

La langue offre une grande variété de choix en ce qui concerne les conjonctions de subordination. Si les grammaires traditionnelles les placent généralement sur le même plan, elles ne s'utilisent pas toutes aussi fréquemment les unes que les autres. Nous proposons ici une étude comparée de la fréquence d'emploi des principales conjonctions simples dans trois corpus différents : le corpus de la LGSWE, le BNC, et une grande partie de notre corpus (appelé ici corpus FP). Quels sont les marqueurs les plus utilisés dans chacun des trois corpus ? Quelles différences les trois corpus présentent-ils néanmoins ?

4.2.2.1 Données prises en compte et calculs effectués

Nous décrivons ici les données retenues pour l'étude comparative de la distribution des marqueurs de subordination dans les trois corpus.

Données prises en compte

Corpus dépouillés Nous analysons ici des données issues de trois corpus différents par leur taille et leur maniabilité. Le plus important en taille est le BNC dont nous ne considérons que la partie produite à l'oral, soit plus de 10 millions de mots que des requêtes ciblées permettent de consulter avec précision. Dans le corpus de la LGSWE ne nous intéressent que les données recueillies à partir de l'analyse des conversations

en anglais britannique, telles qu'elles sont présentées dans l'ouvrage de D. Biber *et al.* (soit en proportion, par million de mots). Un certain nombre de précautions ont été prises pour l'établissement des données que nous retenons ici. Elles sont rassemblées en annexe.[24]

Pour cette expérience, nous n'avons pas eu recours à l'intégralité de notre corpus. Seul notre corpus primaire et une partie de notre corpus secondaire ont été dépouillés. Sont désignés par le terme de corpus FP les textes suivants : Celia&James, Helen&Rebecca, ainsi que Mike&Hannah et les deux corpus Nick&Matthew, soit au total environ 30 000 mots.

Répertoire des différents marqueurs L'éventail des conjonctions répertoriées dans les grammaires est très large. Elles sont la plupart du temps séparées en plusieurs groupes, selon leur forme. R. Quirk *et al.* proposent de distinguer les conjonctions simples, les conjonctions complexes et les conjonctions corrélatives. Nous reprenons ici largement leurs travaux présentés dans la *Comprehensive Grammar of the English Language.*[25]

Parmi les conjonctions simples figurent *after, (al)though, as, because, before, if, once, since, that, unless, until, when, where*, et *while.* Les conjonctions complexes sont classées en fonction de la conjonction simple qui s'y retrouve, combinée avec au moins un autre élément adverbial, prépositionnel, verbal ou nominal. Le subordonnant *that*, que nous avons choisi d'écarter, permet également de former des conjonctions complexes. Par exemple, parmi les conjonctions complexes en *that*, on retrouve *in that, so that, such that*, ou encore *except that.* Dans d'autres conjonctions complexes en *that*, la conjonction simple est facultative, comme dans *now (that), provided (that), supposing (that), considering (that), seeing (that).* Sur le même modèle, on

[24]Voir p.439.

[25]QUIRK, Randolph *et al.* (1985). *A Comprehensive Grammar of the English Language.* Londres : Longman. (§14.12).

trouve des conjonctions combinées avec *as*, *than*, ou *if*. J.-R. Lapaire et W. Rotgé mentionnent également les expressions de temps *every time* et *the minute/the moment* parmi les conjonctions complexes qui ne se combinent pas, cette fois, avec une conjonction simple.[26] Les conjonctions corrélatives sont également des combinaisons de conjonctions simples pour la plupart, dont la particularité est d'être réparties à l'initiale de chacune des propositions, subordonnée et matrice. On cite notamment *if... then, (al)though... yet/nevertheless, as... so, more/-er/less... than, as... as, no sooner... than, whether... or* et *the... the*.

De cette liste non exhaustive, nous ne retiendrons que les conjonctions les plus simples dont le dénombrement a été fait sur une même base dans les trois corpus décrits ci-dessus. Les marqueurs dont la distribution nous intéresse sont donc au nombre de neuf : *if, because/cos, when, before, while, although, until, since* et *after*. Compte tenu de la forte représentation des propositions circonstancielles causales et conditionnelles relevée dans la LGSWE, il est à prévoir que les marqueurs *if, because/cos* et *since* sont les plus employés.

Calculs effectués Le dénombrement des conjonctions simples dans les trois corpus doit être complété par certains calculs pour rendre les données comparables. Trois valeurs nous intéressent ici :

1. le nombre de mots du corpus (N_m) ;
2. le nombre de conjonctions du corpus (N_c) ;
3. le nombre d'occurrences de chacune des neuf conjonctions étudiées (N_x).

Nous pouvons ainsi calculer le nombre moyen de conjonctions pour 100 mots dans chacun des trois corpus. $N_c/N_m * 100$ donne une idée de la richesse, de la densité d'un corpus en conjonctions. Il est également possible de calculer la part occupée par

[26] LAPAIRE, Jean-Rémi & ROTGE, Wilfrid (1992). *Réussir le commentaire grammatical de textes.* Paris : Ellipses. (p.170).

chacune des conjonctions au sein des conjonctions du corpus. $N_x/N_c * 100$ indique la probabilité (en %) pour que la conjonction x soit choisie lorsque l'une des neuf conjonctions simples est utilisée. Enfin, le dernier calcul, qui est le produit des deux précédents, révèle combien de fois la conjonction x sera utilisée pour 100 mots du corpus : $N_x/N_m * 100$.

4.2.2.2 Résultats comparés

Comparaison des trois corpus Le Tableau 4.1 résume brièvement le nombre d'échantillons de chaque corpus à partir desquels il est possible d'effectuer une comparaison. Celui-ci se monte à 128 360 au total. Il indique également la densité de conjonctions simples par corpus.[27]

	CORPUS FP	LGSWE	BNC	TOTAL
Nbre de conjonctions (N_c)	360	11 550	116 447	128 360
Nbre de mots (N_m)	30 000	1 000 000	10 365 464	11 395 464
Rapport en % ($N_c/N_m * 100$)	1,2000	1,1550	1,1234	1,1264

TAB. 4.1 – Eléments de comparaison entre les trois corpus concernant le nombre de conjonctions simples relevées par rapport au nombre de mots

Densité de conjonctions simples au sein des trois corpus Notre premier point de comparaison concerne la densité de conjonctions simples au sein de chaque corpus. Comme en témoigne le Tableau 4.1, les résultats sont très proches.[28]

Dans notre corpus, les neuf conjonctions totalisent 360 occurrences, soit une proportion de 1,2%, légèrement supérieure aux proportions constatées dans les deux autres corpus et dans la totalité des trois corpus. La plus grande densité de conjonc-

[27]Remarquons que les résultats portant sur la totalité des trois corpus sont fortement influencés par les résultats du BNC, qui contient à lui tout seul plus de 90% des mots des trois corpus réunis.

[28]Valeur calculée : $N_c/N_m * 100$.

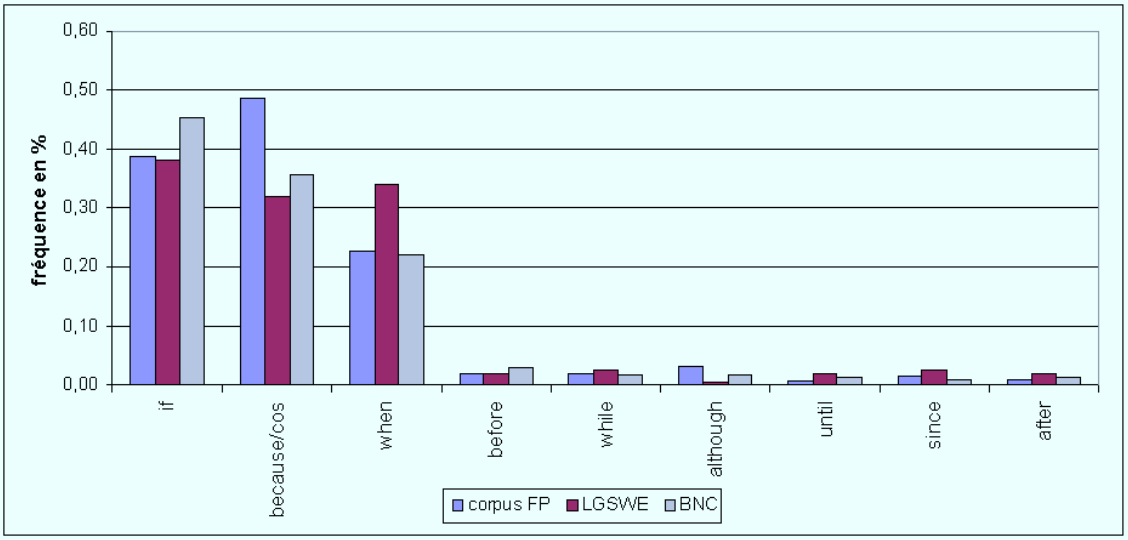

FIG. 4.2 – Fréquence d'occurrence (pour 100 mots) de chaque conjonction dans chaque corpus

tions simples dans notre corpus s'explique en partie par le grand nombre d'occurrences de *because/cos* relativement au nombre de mots (0,49% dans notre corpus contre 0,35% pour le BNC). La forte proportion de la conjonction *although* (0,03%) par rapport aux autres corpus et notamment au corpus de la LGSWE (0,005%) nous semble être une anomalie due au faible nombre d'échantillons étudiés. Néanmoins, elle contribue au même effet dans notre expérience.

Fréquence de chaque conjonction par corpus Si l'on considère à présent en détail la fréquence d'occurrence de chacune des neuf conjonctions, il apparaît que celle-ci est la plupart du temps très proche dans les trois corpus. On retrouve invariablement *if, when* et *because/cos* parmi les conjonctions les plus représentées. Certaines différences existent néanmoins, ce dont témoigne la Figure 4.2. Cet histogramme présente la fréquence comparée de chaque conjonction par rapport au nombre total de mots dans chacun des trois corpus.[29]

Nous voyons ici que chaque corpus se distingue des deux autres sur une conjonc-

[29]Valeur calculée : $N_x/N_m * 100$, où N_x représente le nombre d'occurrences de la conjonction x.

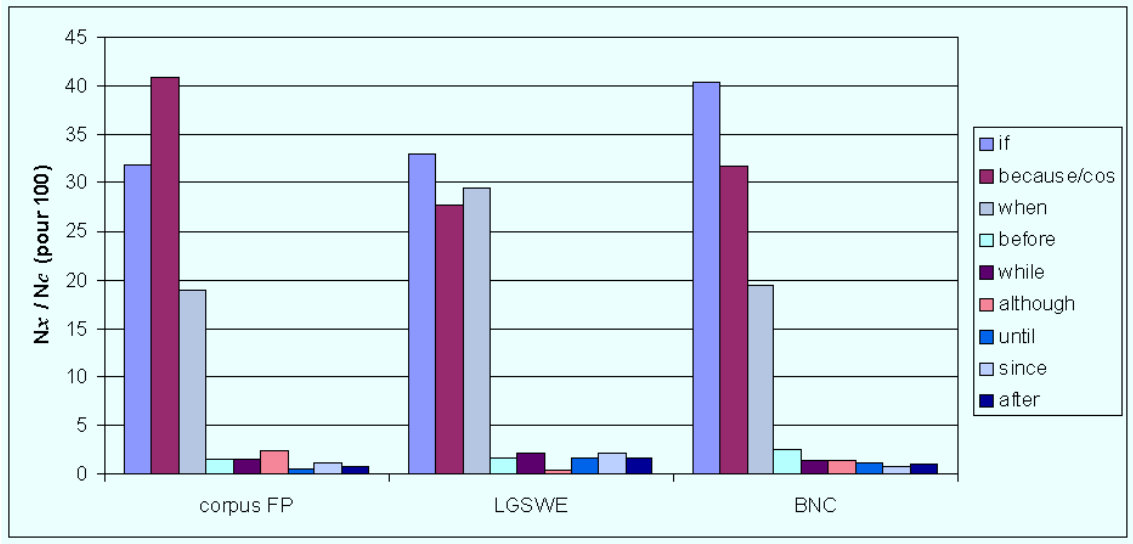

FIG. 4.3 – Part de chaque conjonction par rapport au nombre total de conjonctions dans chacun des trois corpus (pour 100 conjonctions)

tion particulière. En effet, *if* est proportionnellement plus représenté dans le BNC qu'ailleurs. De même pour *when* dans le corpus conversationnel de la LGSWE. Notre corpus se distingue quant à lui par une plus forte proportion de *because/cos* que les deux autres. Pour les six autres conjonctions, les variations sont infimes.

Part de chaque marqueur parmi les conjonctions simples employées Ces tendances se retrouvent partiellement dans la part prise par chacune des conjonctions par rapport à l'ensemble des conjonctions simples relevées pour un corpus, comme l'illustre la Figure 4.3.[30] Deux tiers des marqueurs considérés ayant une part très faible, concentrons-nous sur les trois marqueurs les plus fréquemment choisis parmi les conjonctions simples.

Certes, pour les trois corpus confondus, lorsque l'une des neuf conjonctions de subordination simples est employée, il s'agit dans près de 40% des cas de la conjonctions *if*, dans plus de 31% des cas de *because* ou *cos*, et dans environ 20% des cas de *when*.

[30] Valeur calculée : $N_x/N_c * 100$.

Cependant, comme le montre le Tableau 4.2, cette fréquence par rapport au total des conjonctions relevées dans les trois corpus masque certaines différences entre les corpus.

	CORPUS FP	LGSWE	BNC	TOTAL
if	31,9 %	32,9 %	40,4 %	39,7 %
because/cos	40,8 %	27,7 %	31,8 %	31,4 %
when	19 %	29,4 %	19,4 %	20,3 %
TOTAL	91,7%	90 %	91,6 %	91,4 %

TAB. 4.2 – Part de *if, because/cos* et *when* par rapport au nombre total de conjonctions

En effet, la conjonction *if*, qui représente 40% des conjonctions simples relevées dans le BNC, ne se retrouve que dans environ 32% des cas dans les deux autres corpus. De même, la conjonction *when* n'est utilisée que dans 19% des cas dans notre corpus et dans le BNC, contre 30% dans la LGSWE. La forte présence des conjonctions *because* et *cos* semble être une caractéristique importante de notre corpus. Il apparaît en effet que ces deux conjonctions réunies constituent 40% des conjonctions relevées dans notre corpus. Cette proportion est sensiblement plus faible dans les deux autres : 27% pour la LGSWE et 31% pour le BNC.

Spécificités de notre corpus A l'intérieur de notre corpus, les différents textes offrent-ils une image homogène de l'emploi des conjonctions de subordination les plus simples ? Les calculs appliqués pour la comparaison des trois corpus peuvent être utilisés pour comparer à présent les textes constituant notre corpus.

Densité de conjonctions simples au sein des textes de notre corpus La densité de l'ensemble du corpus FP en terme de conjonctions simples est, nous l'avons vu plus haut, de 1,2%. Comme en témoigne la Figure 4.4, la proportion de conjonctions

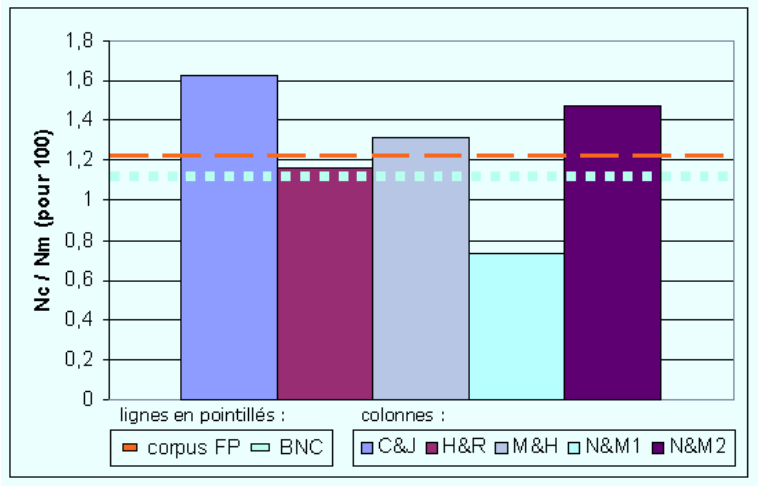

FIG. 4.4 – Densité de conjonctions dans le corpus FP

par rapport au nombre de mots des différents textes de notre corpus pris individuellement peut varier du simple au double. Les conjonctions de subordination simples sont, dans leur ensemble, sur-représentées en particulier dans Celia&James (1,62%) et dans Nick&Matthew2 (1,44%) par rapport au BNC mais également à l'ensemble du corpus FP. En revanche, Helen&Rebecca propose des valeurs extrêmement proches de la norme fixée par le BNC. Le corpus Nick&Matthew1 est particulièrement pauvre en conjonctions : la densité de conjonctions dans cette partie de notre corpus secondaire y est deux fois moindre que dans Nick&Matthew2, par exemple. Elle y est également largement inférieure à la densité relevée dans le BNC.

Cette différence mérite notre attention. Tout d'abord, il convient de vérifier si cet écart entre les textes de notre corpus touche toutes les conjonctions ou seulement certaines. Si l'on regarde les données conjonction par conjonction, on constate que les différences entre les corpus s'expliquent largement par les différences relevées pour les trois principales conjonctions. Dans le corpus Celia&James, la plus forte densité est due à une sur-représentation des conjonctions *if* et surtout *because/cos* par rapport au reste de notre corpus. *When* y est quant à lui légèrement sous-représenté. A l'inverse,

Nick&Matthew1 est caractérisé par une faible proportion de *if* et de *because/cos* que la plus forte présence de *when* ne suffit pas à contre-balancer. Dans le corpus Nick&Matthew2, c'est la plus forte représentation de *when* qui explique la densité relativement élevée de conjonctions. L'examen des données individuelles relatives à la représentation de chacune des conjonctions révèle des équilibres différents au sein des deux textes de notre corpus dont la densité globale est la plus proche de la norme. Dans Helen&Rebecca, un surplus de *if* vient contre-balancer une moindre représentation de *when*. La densité globale légèrement supérieure présentée par le corpus Mike&Hannah s'explique essentiellement par la plus forte présence de *because/cos* relativement au nombre de mots du corpus.

En somme, l'écart de densité des différents textes de notre corpus en ce qui concerne les conjonctions de subordination simples n'est jamais l'effet d'une sur-représentation collective des conjonctions relevées dans un corpus. Il provient toujours de l'équilibre des différentes composantes que constituent les neuf conjonctions simples étudiées, et plus particulièrement des trois principales. En terme de densité, les trois conjonctions les plus courantes permettent de distinguer les différents textes de notre corpus de la façon suivante : dans Celia&James et Mike&Hannah, *because* et *cos* sont sur-représentés par rapport à l'ensemble du corpus FP. *If* est sur-représenté dans Celia&James et dans Helen&Rebecca. *When* est plus présent dans les deux textes Nick&Matthew que dans le reste de notre corpus.

Part de chaque marqueur parmi les conjonctions simples employées dans notre corpus Ici encore, l'examen des seules conjonctions *if*, *because/cos* et *when* suffit à épuiser la question, dans la mesure où celles-ci constituent plus de 90% des conjonctions. Seul Nick&Matthew2 ne dépasse pas les 87,5%. Deux groupes se dessinent clairement lorsque l'on élimine *when* pour calculer la part occupée par les seuls

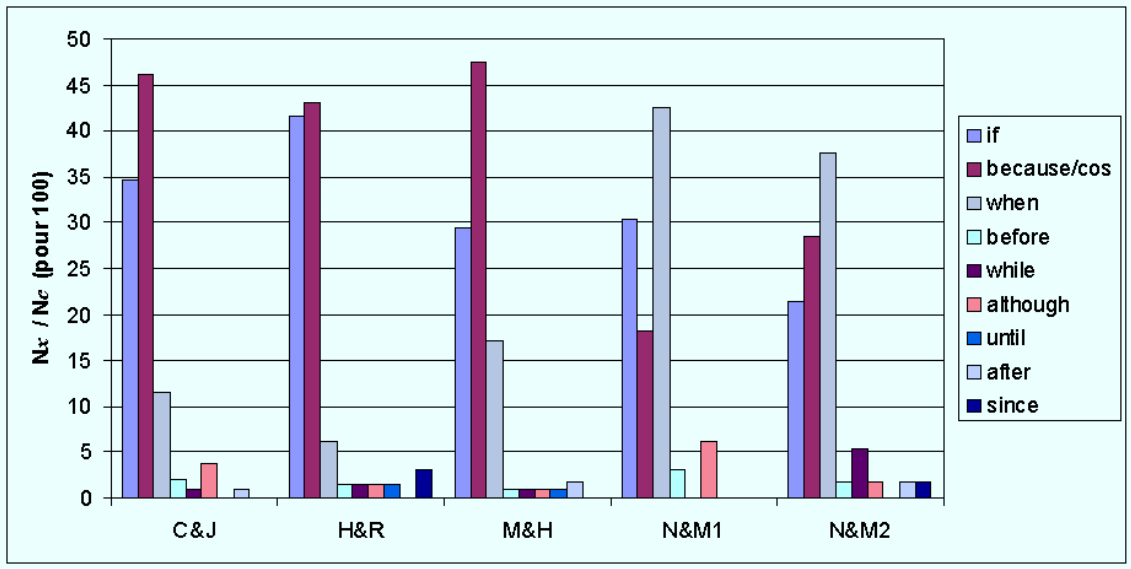

FIG. 4.5 – Part de chaque conjonction par rapport au nombre total de conjonctions dans chacun des textes de notre corpus

because/cos et *if* dans les différents textes.[31] La proportion atteint 75 à 85% pour Celia&James, Helen&Rebecca et Mike&Hannah, contre seulement 50% pour les deux textes Nick&Matthew. Si l'on resserre le champ de cette étude à *because/cos* et *when* (en éliminant donc *if*), l'opposition entre ces deux groupes est encore plus parlante. Comme le montre la Figure 4.5, la proportion des deux conjonctions fonctionne de façon inversée dans les textes Nick&Matthew par rapport au reste de notre corpus.

En somme, parmi les conjonctions simples étudiées, la conjonction privilégiée dans Celia&James, Mike&Hannah, et, dans une moindre mesure Helen&Rebecca est *because/cos*. *When* constitue la conjonction privilégiée dans les textes Nick&Matthew.[32]

[31] Valeur calculée : $N_x/N_c * 100$.

[32] Le test du khi2 révèle une différence significative entre les différents textes de notre corpus en ce qui concerne la distribution des conjonctions *if, when* et *because/cos* ($\chi^2 = 43{,}594$ pour 8 ddl, $p \leqslant 0{,}001$). Cependant, les tests détaillés montrent que les corpus Celia&James, Helen&Rebecca et Mike&Hannah ne sont pas significativement différents. L'écart entre les cinq textes de notre corpus provient essentiellement de la distribution observée dans les textes Nick&Matthew. (Pour une description du test du khi2, voir MULLER, Charles (1992). *Initiation aux méthodes de la statistique linguistique*. Paris : Honoré Champion.)

4.3 Bilan et hypothèses explicatives

Les calculs effectués plus haut révèlent des points communs et des différences entre les corpus sur lesquels nous souhaitons revenir afin de tenter de les expliquer.

4.3.1 Résumé des résultats

Le dénombrement des marqueurs simples de subordination à travers plusieurs corpus a mis en lumière des tendances fortes. Celles-ci se vérifient à différentes échelles à travers les points communs relevés entre les corpus, mais également à travers leurs divergences.

4.3.1.1 Points communs entre les corpus dépouillés

Les corpus dépouillés convergent sur de nombreux points. Les résultats relatifs à la densité de conjonctions simples par rapport au nombre total de mots dans un corpus montrent une relative similarité entre les corpus ($N_c/N_m * 100$) : un marqueur simple de subordination est utilisé à l'oral spontané tous les 80 à 90 mots environ.[33]

Un constat général s'impose à la lecture des données concernant la probabilité d'apparition d'une conjonction par rapport aux autres conjonctions simples ($N_x/N_c * 100$). Les données issues des trois corpus se rejoignent et font clairement apparaître une tendance forte, si ce n'est relativement à l'anglais oral spontané en général, du moins concernant l'utilisation des conjonctions de subordination les plus simples : la diversité des marqueurs effectivement utilisés est extrêmement faible. Dans les trois corpus, *if, because/cos* et *when* représentent à eux seuls plus de 90% des occurrences des neuf conjonctions simples prises en compte.

[33]Voir le Tableau 4.1, p.207.

4.3.1.2 Points de divergence

Certaines divergences apparaissent également entre les corpus, révélant des caractéristiques propres.

Entre les trois corpus dépouillés Les divergences constatées entre les trois corpus dépouillés concernent la prééminence d'une conjonctions par rapport à toutes les conjonctions simples envisagées ($N_x/N_c * 100$). Ainsi, l'une d'entre elle émerge au sein de chaque corpus comme marqueur de subordination privilégié. C'est le cas de *if* dans le BNC, qui est choisi dans 40% des cas, et pour *because/cos* dans le corpus FP, tandis que *when* est davantage utilisé dans la LGSWE que dans les deux autres corpus.[34]

Au sein de notre corpus Des divergences de divers ordres apparaissent entre les textes de notre corpus. Le calcul de la densité (N_c/N_m*100) a montré que Celia&James est plus riche en conjonctions de subordination simples que le reste des textes et que l'ensemble du corpus FP.[35]

Si l'on considère à présent la probabilité d'apparition d'une conjonction par rapport aux autres conjonctions simples ($N_x/N_c * 100$), il est apparu que *because/cos* est particulièrement présent dans Celia&James, Helen&Rebecca et Mike&Hannah (40% des occurrences), alors que *when* est privilégié par les textes Nick&Matthew — textes qui sont apparus sur ce point significativement différents des autres.[36]

Au sein de notre corpus se vérifie la double tendance relevée dans les corpus de référence : la faible diversité en marqueurs et la préférence pour un marqueur en particulier. Quels éléments pouvons-nous apporter pour expliquer ce phénomène ?

[34]Voir le Tableau 4.2, p.210.
[35]Voir la Figure 4.4, p.211.
[36]Voir la Figure 4.5, p.213.

4.3.2 Hypothèses explicatives

A travers ces divergences relevées entre les trois corpus dépouillés ou entre les textes de notre corpus se confirme encore la tendance suggérée par les points communs énoncés plus haut : celle de l'utilisation quasi exclusive d'une ou plusieurs conjonctions en dépit de la variété accessible aux locuteurs. Deux hypothèses concurrentes peuvent être formulées pour expliquer cette tendance.

4.3.2.1 Deux hypothèses linguistiques concurrentes

Hypothèse structurelle La première hypothèse que nous souhaitons proposer est relative à la notion linguistique de choix et dégage une tendance structurelle de l'anglais oral spontané. Si la grammaire place le locuteur devant une multitude de possibilités pour exprimer le contenu de sa pensée, il semble que celui-ci restreigne la plupart du temps son choix à un nombre très faible de solutions. En conséquence de cette réduction systématique du champ des possibles (qui s'illustre dans l'emploi privilégié des circonstancielles causales et conditionnelles), les trois conjonctions *if*, *when* et *because/cos* suffisent à épuiser presque entièrement l'éventail des conjonctions utilisées.

Hypothèse conjoncturelle La seconde hypothèse concerne le type de propos échangés dans le cadre d'une conversation spontanée. Elle s'attache donc davantage à dégager l'aspect conjoncturel de ces tendances. Il est en effet possible de supposer que la plus grande présence de telle ou telle conjonction dans un corpus ou dans une partie de ce corpus s'explique par le type de relation intersubjective qui s'y joue, ou par le contenu des propos échangés. Au-delà des tics et habitudes langagières de tel ou tel locuteur, il s'agit de détecter une possible corrélation entre la forte présence de la conjonction x et le discours dans lequel celle-ci apparaît. Selon cette hypothèse, la forte représentation des trois conjonctions provient du fait que les interlocuteurs dans une conversation spontanée s'occupent principalement d'apporter une cause ou une

explication, de formuler une condition ou de replacer un événement dans le temps — ce à quoi les subordonnées introduites par les trois conjonctions sont employées.

4.3.2.2 Arguments, objections et conclusions

N'ayant pas accès aux textes constituant le corpus de la LGSWE ou le BNC, nous ne pouvons pas vérifier cette deuxième hypothèse pour ces deux corpus. En revanche, notre propre corpus peut être analysé dans cette direction.

Peut-on attribuer ces différences, en vertu de notre seconde hypothèse, au contenu thématique[37] et à la nature de l'échange à l'œuvre dans chaque corpus, et dans chaque texte au sein de notre corpus ? Dans Celia&James, la forte densité en conjonctions peut s'expliquer par le fait que, les deux locuteurs ne se connaissant pas avant l'enregistrement, ils ont davantage besoin de baliser leur échange en situant les faits qu'ils relatent dans le temps et en justifiant leurs propos. Néanmoins, il est frappant de constater la non-corrélation entre le propos général de la conversation entre Helen et Rebecca et entre Mike et Hannah avec l'emploi de la conjonction la plus attendue. Dans ces deux textes, la dimension temporelle est très importante. Helen et Rebecca parlent en effet de leur emploi du temps de la journée et de la semaine ainsi que de leurs projets de sorties et de leurs perspectives d'emploi dans les mois à venir. Mike raconte longuement ses vacances mouvementées en Grèce. Or, dans ces deux textes, la proportion de *when* est faible (17% pour Mike&Hannah) voire très faible (6% pour Helen&Rebecca).

Certes, la dimension sémantique et thématique ne doit pas être évacuée. Elle explique au moins une partie des faits constatés. Dans Helen&Rebecca, l'importance de la projection dans l'avenir, qui constitue un thème de la conversation, se retrouve dans la part accordée à la conjonction *if* (40% des cas). Cependant, la forte proportion de *when* et sa relative densité constatée dans les textes Nick&Matthew n'en font pas

[37] Le « contenu thématique » est ici utilisé au sens large de « sujet de la conversation ».

un récit événementiel. D'un point de vue thématique et intersubjectif, il est difficile d'expliquer pourquoi les textes Nick&Matthew se singularisent par rapport au reste de notre corpus. Il est également difficile d'expliquer la différence qui existe entre les deux textes en ce qui concerne la densité des conjonctions utilisées.[38] Le seul examen du contenu thématique et de la nature de l'échange dans son ensemble ne nous donne pas les moyens de comprendre de façon pleinement satisfaisante de tels écarts entre les textes. Sans doute nous faudrait-il considérer avec davantage d'attention le détail de l'évolution des positionnements intersubjectifs au cours d'une conversation et la possible corrélation avec une plus grande ou plus faible utilisation de conjonctions en général ou de telle conjonction en particulier. Il nous semble en effet plus prometteur de nous concentrer sur le type de relation intersubjective que sur la dimension thématique de l'échange. C'est d'ailleurs le sens de la réflexion de M. De Cola dont nous nous sommes fait l'écho plus haut : en repérant la prise en charge d'un énoncé par rapport à celle d'un autre, la subordination contribue à la gestion des relations intersujets, qui, en tant que moteur de la dynamique conversationnelle, favorise la progression discursive.[39]

Une objection supplémentaire à notre seconde hypothèse provient de l'observation suivante. Le sémantisme des propositions introduites par les six autres conjonctions (en particulier *before, while, until, after* et *since* pour les subordonnées temporelles et causales) n'est pas foncièrement différent de celui des trois conjonctions les plus utilisées. Pourquoi sont-elles si peu employées ? Faisons dès lors l'hypothèse que, les conjonctions les plus simples introduisant précisément des subordonnées au sémantisme très proche, le locuteur porte naturellement son choix vers la conjonction la plus générique. Par exemple, la conjonction *when* occupe le centre organisateur du domaine notionnel de l'expression d'une relation temporelle. En périphérie se trouvent

[38] Voir test du χ^2, p.213n.
[39] Voir p.202.

des conjonctions moins polyvalentes dont le sémantisme les spécialise pour l'expression d'un certain type de relation temporelle.

Il semble donc bien, conformément à notre première hypothèse, que la faible diversité des conjonctions simples effectivement employées provienne d'une limitation structurelle du champ des possibles à l'oral. D'après nos résultats, les interlocuteurs d'une conversation spontanée ont fortement tendance, quel que soit le sujet abordé, à n'utiliser qu'une portion réduite de la gamme que la grammaire leur offre. Valable pour les conjonctions les plus simples introduisant une proposition subordonnée circonstancielle, cette hypothèse mériterait d'être vérifiée dans d'autres domaines pour lesquels s'offre un choix concernant la syntaxe et la sémantique.

Au terme de cette étude quantitative des conjonctions de subordination les plus simples, il est apparu que les corpus d'oral spontané, bien qu'inégalement riches en conjonctions, présentent la caractéristique générale de n'utiliser pour toutes conjonctions qu'un tiers de celles répertoriées dans les grammaires. Cette tendance forte s'explique semble-t-il par notre première hypothèse concernant principalement la notion de choix et se retrouvant dans l'attraction du centre organisateur de la notion dans laquelle le contenu de pensée du locuteur s'inscrit. Notre deuxième hypothèse cherche à expliquer la forte présence des trois principales conjonctions par le contenu thématique et le type de relation intersubjective à l'œuvre. Moins satisfaisante, cette hypothèse, dont il faudra vérifier la possible adéquation avec le détail des échanges, ne réussit pas à expliquer les différences existant entre les trois corpus, ni même entre les cinq textes constituant notre corpus. Enfin, notre étude a montré que notre corpus est particulièrement riche en *because* et *cos* et que la part de ces deux conjonctions y est très importante. Il est donc propice à une analyse plus précise de ces deux conjonctions que nous allons à présent étudier séparément.

Chapitre 5

Le cas des circonstancielles de cause

De notre étude distributionnelle des marqueurs de subordination émerge tout d'abord la catégorie sémantique de la subordination causale. Il apparaît que son marqueur privilégié, *because/cos*, est également particulièrement représenté dans notre corpus. Nous proposons donc dans ce chapitre d'approfondir notre analyse de la subordination à l'oral en nous concentrant sur le fonctionnement de ce marqueur complexe dans notre corpus. Que révèle-t-il des spécificités de la subordination syntaxique à l'oral ? Dans quelle mesure peut-on distinguer *because* et *cos* et les considérer comme des marqueurs apparaissant chacun de façon privilégiée dans un contexte prosodique, argumentatif et énonciatif différent ?

5.1 Généralités sur l'expression de la cause

5.1.1 Une relation logique

D'un point de vue logique, la relation de cause implique au moins deux propositions, l'une rapportant un effet constaté, l'autre une cause observée ou induite liée à cet effet. D'un point de vue conceptuel, une telle relation établit donc une hiérarchisation des deux propositions, l'une découlant de l'autre qui lui est préalable et la contient potentiellement.

Dans l'échange conversationnel, les locuteurs s'appuient sur des relations logiques telles que la cause pour assurer la progression de l'argumentation et agrandir l'espace de consensus qu'ils construisent. Cependant, la relation causale n'y est qu'une lointaine parente de celle de la logique, tant elle s'appuie sur les représentations mentales engagées dans l'échange. C'est le sens de la remarque de D. Schiffrin :

> [...] the multiple realization of 'cause' and 'result' relations is not a product of *because* or *so*. Rather, it is due to our understandings of causality. Although full consideration of how causal relations are actually worked out (either in scientific or everyday reasoning) is not important for discussion of causal connectives, it does seem that most conclusions which we draw about cause and effect are not based on observations which empirically warrant them: events are not usually observed both before and after a stimulus, as well as compared to a control event to which the stimulus has not applied. Rather, most conclusions are based on our interpretations of causal relations according to our own (culturally relative) schemata for making such interpretations. Since we assume that others are following parallel procedures — basing their inferences on their own schemata — we interpret many statements from others as conclusions warranted by the speaker's own perspective.[1]

[1] SCHIFFRIN, Deborah (1987, 2001). *Discourse markers.* Cambridge : CUP. (pp.211-212).

Nous ne nous attarderons pas sur la dimension formelle de la relation logique de cause, dont il apparaît que le mécanisme est largement déformé en discours au profit d'une logique plus personnelle.

5.1.2 La cause en langue et en discours

Comment la relation de cause s'exprime-t-elle ? Qu'advient-il en discours de ce que la langue offre aux locuteurs d'un corpus d'anglais oral spontané ?

5.1.2.1 Outils disponibles et leurs caractéristiques

La relation causale entre deux propositions enchâssées peut s'exprimer de plusieurs façons différentes, par l'intermédiaire d'une conjonction ou sans conjonction. Dans le cadre de ce chapitre, nous nous intéresserons essentiellement à la première.

Les différentes conjonctions utilisées en anglais pour exprimer la cause induisent chacune une relation particulière entre p et q. Dans tous les cas, le contenu de la proposition subordonnée constitue le repère permettant la validation de celui de la principale. Le degré de prise en charge de la validité de ce repère et de celle du lien entre repère et repéré varie en fonction de la conjonction utilisée.

Since Les propositions en *since* expriment une présupposition. Le contenu de q (Q) ne fait pas l'objet d'une prise en charge au moment de l'énonciation, mais la conjonction signale que la question de la prise en charge est dépassée, retirée du débat, celle-ci ayant déjà eu lieu. La préconstruction de Q découle naturellement de la préconstruction du lien entre les deux propositions. M. De Cola-Sekali conçoit l'opération effectuée avec *since* comme une double relation de repérage :

> [...] q est le repère d'une classe de p pour tout S [énonciateur] [...].
>
> Ce double repérage (q repère constitutif d'une classe d'énoncés incluant p, et S* [tout énonciateur] repère subjectif de ce lien) est utilisé par S_0 comme

> confirmation de sa propre modalisation de l'énoncé cible p. Il autorise ainsi l'énonciateur S_0 à modaliser p sans risque d'être contré.[2]

Lorsqu'il utilise la conjonction *since*, l'énonciateur prend en charge la validation de la relation prédicative dans p et ne fait que rappeler un lien entre les deux propositions dont il souligne la nature évidente et non problématique. Bien entendu, une proposition en *since* peut être utilisée par l'énonciateur sans prendre appui sur aucun construit préalable. En faisant passer automatiquement dans l'acquis Q et le lien entre celui-ci et le contenu de p, une telle construction peut relever d'une stratégie argumentative visant à retirer certains éléments du champ du discutable.

As La conjonction *as* semble fonctionner de la même façon que *since*. Sa parenté avec le *as* comparatif la rend en effet propice à introduire une proposition dont la validation est considérée comme acquise et qu'elle institue en point de repère.

Because En revanche, avec *because*, la question de la prise en charge n'est pas dépassée. Elle est même centrale. En T_0 se joue la prise en charge de q, de p et du lien entre les deux propositions, grâce à la présentation par l'énonciateur en q des conditions de réalisation de la relation prédicative dans p. Nous souscrivons ici à l'hypothèse de M. De Cola-Sekali selon laquelle

> [...] dans les énoncés du type *p because q*, l'opérateur *because* permet à l'énonciateur S_0 de poser q comme contenant les conditions de validation de la relation prédicative dans p, et de réimposer sa prise en charge.
>
> *Because* ouvre donc sur une nouvelle prise en charge par l'énonciateur S_0 de l'énoncé p dont la prise en charge est mise en débat par *why* (actualisé dans le jeu interlocutif ou anticipé dans le jeu co-énonciatif).[3]

[2] DE COLA-SEKALI, Martine (1991b). Connexions inter-énoncés et relations intersubjectives : l'exemple de *because, since*, et *for* en anglais. *Langages* 104 : 62-78. (p.69).

[3] *ibid.* (p.66).

Mettant en jeu la question de la prise en charge au moment de l'énonciation, la connexion par *because* est plus neutre que *since* ou *as* sur les plans argumentatif et intersubjectif. Cependant, la part occupée par le travail de l'énonciateur dans une telle construction la rend également propice à une utilisation polémique. En effet, le contenu de q ne fait pas l'objet d'un consensus préalable : il n'est pas préconstruit et ne reçoit sa validation en T_0 que par l'énonciateur. Or, il sert de justification ou d'explication à la relation prédicative dans p dont la validation est en question.

5.1.2.2 Outils effectivement utilisés : éléments de distribution

Des outils relevant d'une problématique causale disponibles dans la grammaire, quels sont ceux que les locuteurs de notre corpus emploient effectivement ?

Why ? La première remarque que nous souhaitons faire concerne la très faible utilisation de *why*, traditionnel catalyseur de proposition causale, en regard de la variété et du nombre de structures exprimant la cause et introduites par un marqueur spécialisé dans notre corpus. Les cinq textes que nous avons examinés contiennent en effet au total 24 occurrences de *why*, dans des emplois différents. Douze d'entre elles constituent des interrogatives au discours direct, comme dans l'exemple *why_010* où Helen et Rebecca résument leurs engagements du week-end :

H but that's not a problem though still
why is that a problem
R **because** we're supposed to be going to Hélène's on {} *Saturday evening*

(Helen&Rebecca : Week-end, *why_010*, *because_027*)

ou au discours direct rapporté, dans l'exemple suivant où Nick relate son récent entretien d'embauche dans une école (*why_018*) :

N and then they asked they asked ludicrous questions as well like er {0,8}
they asked y- you know **why** do you think you're gonna be a good teacher <h> (0,5)
I said well I I know I know it might sound rather cliché but actually I f- I feel it's it's something of a vocation it's what I've

M *<rires>*

N *always wanted* to do <h> and it's what I think I will be best at

(Nick&Matthew1 : Interview, *why_018*)

L'exemple *why_010* est exceptionnel dans la mesure où c'est le seul énoncé de notre corpus dans lequel *because* fait suite à une interrogative en *why*. Moins d'1% des 145 *because/cos* sont donc liés à une question en *why*.

Dans trois autres cas, *why* est utilisé comme pronom interrogatif indirect, comme dans l'exemple *why_007*. Celia et James s'interrogent sur les raisons qui peuvent pousser des Occidentaux expatriés en Chine à se réunir sous les yeux des curieux dans les *English corners* :

J it {0,50} appears just to be a place and Chinese people sort of congregate to see Westerners
I'm not sure <h> {0,70}

C oh the

J what {0,20} point {0,25}

C *to see Westerners*

J *there is for the Westerners* to go there
why they particularly congregate there
perhaps just <rires> they're asked to <rires>

(Celia&James : English corner, *why_007*)

The reason why Enfin, *why* se retrouve de façon récurrente dans notre corpus (9 occurrences sur 24) dans des structures qui le placent en position conclusive du type *(because) q is the reason why p*. Dans l'exemple suivant (*why_008*), *why* vient clore une séquence explicative sur la nécessité admise par Celia et James de pouvoir organiser des expériences perceptives. *Why* n'appelle pas une explication mais permet de désigner un contenu propositionnel énoncé dans son contexte gauche proche comme cause du contenu de la proposition qu'il introduit. Notons ici la présence d'une proposition causale en *cos* dans la réplique précédente de Celia :

C ***cos** you* can't get the people

J well that's that's **why** I want to <rires> do it out of hours *yes*

C *yes* {0,85} you just simply can't get the people

(Celia&James : John's permission, *why_008*)

Dans cet exemple, représentatif des huit autres faisant apparaître *why*, l'antécédent *the reason* est absent. Lorsqu'il est présent, c'est le pronom qui disparaît. Notre corpus ne contient que trois occurrences de ce phénomène. Dans l'exemple *reason_003*, l'orientation des propositions est conclusive, la cause figurant à gauche de l'énoncé sous la forme anaphorique du pronom démonstratif *that* :

N that's er **one of the key reasons** I don't go to church really

(Nick&Matthew1 : Vegetarian, *reason_003*)

L'ordre des propositions logiques est inversé dans l'exemple suivant (*reason_001*), où l'effet est énoncé avant la cause. Le schéma est donc *the reason (why) p is (because) q* :

J I think **the reason** we are lectured ?through? the afternoon is that we share the facilities with the Chinese people <h> {0,80} who: {0,20} *obviously have nicked* the morning slots <rires>

(Celia&James : Facilities, *reason_001*)

As et since Ces deux marqueurs sont extrêmement peu représentés dans notre corpus, que ce soit dans leur emploi argumentatif ou dans leur emploi temporel. Nous ne répertorions aucune occurrence du *as* causal. Parmi les quatre *since* que nous relevons sur l'ensemble des cinq textes de notre corpus, trois sont utilisés dans leur acception causale. Les deux premiers proviennent du corpus Celia&James :

J which is er {0,55} fairly stressful actually <h> (0,40)

C *mm* <bruit>

J *especially* **since** hm {0,70} I thought it'd be fun to do:: {0,55} sort of an experimental thing ?especially? John suggested it {0,70}

(Celia&James : Facilities, *since_001*)

J Sally will be going out to join her and <h> {0,55}
s- **since** she actually speaks Chinese that's really quite helpful

(Celia&James : Travel, *since_002*)

et le dernier de Mike&Hannah :

M I just said look let's just keep swimming **since** we're swimming <h> (0,25) but I'll give it a shot first

(Mike&Hannah : Little near-death experience, *since_003*)

Prépondérance de *because* et *cos* Le présent chapitre n'a pas pour objectif d'épuiser la question de l'expression de la cause ou de ses marqueurs. Il entend seulement apporter des éléments supplémentaires pour confirmer l'observation formulée au chapitre précédent à propos de l'éventail restreint des formes effectivement utilisées parmi les nombreuses formes possibles.[4] En outre, il permet de vérifier, conformément à notre hypothèse de départ concernant les fractales, que des phénomènes observés à un niveau dans l'analyse se retrouvent à un niveau différent.

Nous avons montré au Chapitre 4 comment trois conjonctions de subordination simples (*when, if* et *because* ou *cos*) suffisaient à couvrir 90% des marqueurs utilisés, dans notre corpus comme dans le BNC ou dans le corpus de la LGSWE. Or, la même tendance s'observe lorsque l'on s'intéresse uniquement à un type de subordination — la subordination causale. *Because* et *cos*, qui constituent 40% des conjonctions simples utilisées dans notre corpus, sont employés dans plus de 90% des cas lorsque apparaît une conjonction de subordination causale (telle que celles que nous avons relevées plus haut). Se trouve donc reproduit à l'échelle des conjonctions causales un phénomène observé à un niveau de précision inférieur, celui de toutes les conjonctions simples confondues.

S'opère semble-t-il ici encore une réduction du champ des possibles par les locuteurs. Ceux-ci tendent massivement à diriger leur choix vers le centre organisateur de la notion de cause, et, au sein de la subordination, à s'orienter vers les marqueurs les plus évidents, c'est-à-dire *because* et *cos*. Nous avons évoqué plus haut l'attirance des locuteurs pour des marqueurs polyvalents. Ainsi, *when* est largement préféré à *after* ou *before*. Notons que polyvalent s'entend non pas au sens de polysémique, mais

[4] Voir p.216.

plutôt au sens de non spécifiant ou générique. Il est remarquable en effet que les deux conjonctions polysémiques *as* et *since*, bien que susceptibles d'être employées dans de nombreux cas — relation temporelle, comparative ou argumentative —, soient largement ignorées par les locuteurs de notre corpus.

Nous souhaitons à présent nous concentrer sur l'étude de *because* et *cos*. Que nous disent-ils de la subordination causale à l'oral ? En quoi peuvent-ils être considérés comme un seul et même marqueur ? Dans quelle mesure peut-on au contraire les distinguer ?

5.2 Aspects de la subordination causale en *because* et *cos*

Commençons par considérer les deux formes comme deux facettes d'un seul et même marqueur, *because*. Quelles caractéristiques communes relève-t-on dans notre corpus ?

5.2.1 Deux facettes d'un même marqueur

Cos est communément considéré comme un avatar de *because* dont il constitue, par exemple pour l'*Oxford English Dictionary* (OED), une variante familière ou vulgaire, voire dialectale.[5]

5.2.1.1 *Cos* forme affaiblie de *because*

Selon le même dictionnaire, *cos*, sous ses diverses orthographes (*(')cause, cuz, case...*) n'apparaît que tardivement, pas avant le XVIe siècle. Le deuxième exemple

[5] Voir les entrées *cause, 'cause*, conj. et *cos, 'cos*, adv. and conj. dans SIMPSON, John & WEINER, Edmund (1989, 2e éd.). *Oxford English Dictionary*. Oxford : Clarendon Press.

donné par le dictionnaire sous l'entrée de la conjonction est tiré de *The Jew of Malta* de C. Marlowe, datant de 1592 :

> 1592 MARLOWE *Jew Malta* IV. ii. 1535 Do you mean to strangle me? Yes, **cause** you use to confess.[6]

En revanche, l'OED atteste que *because* tire ses origines du vieil anglais et se rencontre en moyen anglais sous une forme proche de l'actuelle dès la fin du XIVe siècle, par exemple dans les *Canterbury Tales* de G. Chaucer :

> c1386 CHAUCER *Frankl. Prol.* 8 **By cause** I am a burel man..Haue me excused of my rude speche.[7]

Etymologiquement *because* constitue une synthèse du syntagme prépositionnel *by cause*. *Cos* fait l'aphérèse de *because* en ne conservant que la tête lexicale du syntagme d'origine (*cause*), porteuse de l'accent lexical du dissyllabe.

5.2.1.2 Emploi de *cos* pour *because*

Sans souscrire à l'idée selon laquelle les deux formes relèvent forcément d'un niveau de langue très différent, nous observons dans certains cas que le monosyllabe se substitue bien à *because* et en constitue une version courte. Ce phénomène est notamment attendu dans des cas de reprise, une occurrence de *because* précédant une reprise du marqueur par sa forme monosyllabique. Puisque aucune référence n'est en jeu, nous ne pouvons parler de reprise anaphorique qu'au sens stylistique : il nous semble possible de voir ici un simple phénomène de répétition, qui permet l'affaiblissement des mentions suivantes du marqueur répété.

Dans l'exemple *because_045* ou *cos_053*, Mike raconte comment Ann, l'amie avec qui il a passé une semaine de vacances en Grèce, prend conscience du danger qu'il y a à descendre une rivière : pendant un instant, elle croit que Mike s'est noyé. Le

[6] *ibid.* (Entrée *cause, 'cause,* conj.).
[7] *ibid.* (Entrée *because,* conj.).

locuteur pose le sujet (*she*) et le verbe (*thought*) d'une proposition causale en *because* avant de l'interrompre pour ajouter un détail important concernant le temps qu'il a passé sous l'eau. Après une respiration audible, il entame une nouvelle proposition causale, modalisée par l'adverbe *apparently*, dont le sujet est cette fois le pronom personnel première personne. Notons que Mike revient ensuite à sa proposition initiale débutant par *she thought*. Ce faux départ s'accompagne d'une reprise du marqueur de subordination. *Because* est ici repris sous la forme de *cos* :

M but actually that freaked her out **because** she thought <h> **cos** apparently I went under for quite a long time and she thought I was <h> (0,35) a goner and she suddenly thought God <h> (0,5) what happens if {0,7} I disappear

(Mike&Hannah : Freaked out, *because_045, cos_053*)

5.2.2 Caractéristiques acoustiques

La phonologie, par souci de systématisme, oppose de façon binaire une forme pleine à une forme faible (les deux admettant une ou plusieurs réalisations). L'acoustique tend à montrer que ces variations sont en fait à lire sur un continuum entre réalisations canoniques, soignées ou en forme de citation d'un côté, et réalisation coarticulée, réduction articulatoire en parole continue ou réalisation dans une position prosodique faible de l'autre. Ce type d'opposition caractérise par exemple *I'm* par rapport à *I am*, ou *I'll* par rapport à *I will* qui, comme *cos* et *because* sont des couples dissyllabe/monosyllabe. Dans quelle mesure ce phénomène de réduction s'applique-t-il à *cos* par rapport à *because* ? Nous présentons ici une expérience pilote menée sur un petit nombre d'exemples issus de notre seul corpus Celia&James (15 exemples).[8]

Certains résultats suggèrent un affaiblissement de *cos* par rapport à *because*. C'est le cas tout d'abord de la tendance de *cos* à être plus court que la deuxième voyelle de *because*. En outre, l'analyse de la position du deuxième formant indique sur certains

[8]Nous tenons à remercier Alexis Michaud à qui nous devons les mesures acoustiques présentées ici.

exemples une voyelle légèrement plus antérieure pour *cos* que pour *because*, ce qui peut être interprété comme le signe d'une tendance à la centralisation, témoin d'une prononciation affaiblie. D'autre part, la voyelle de *cos* est parfois entièrement dévoisée.

D'autres résultats indiquent la permanence de certains traits entre les deux marqueurs. Ainsi, le *voice onset time* de l'occlusive est très proche dans les deux cas, soit environ 60 ms. Ceci suggère que l'occlusive reste aspirée dans tous les cas observés, et n'est donc pas affectée par une réduction. D'autre part, la remarque formulée à l'instant concernant la postériorité de la deuxième voyelle de *because* n'est valable que pour James. Chez Celia, nos exemples ne montrent pas de différence régulière entre les deux voyelles. L'analyse de la durée des quinze marqueurs envisagés montre que *cos* est en moyenne plus court que la moyenne des deux syllabes de *because*. Cependant, le débit des énoncés dans lesquels *cos* apparaît est également plus rapide, la durée d'une syllabe étant d'environ 0,18 s en moyenne (contre 0,21 pour les énoncés en *because*). Ces chiffres correspondent presque exactement aux durées moyennes constatées pour le marqueur *cos* et pour la moyenne des syllabes de *because* : dans son contexte, *cos* est donc aussi long que *because* dans le sien.

D'un point de vue général, le rapide examen de nos quinze exemples suggère que la réalisation de chacun des deux marqueurs n'est pas monolithique. *Cos* montre certains signes d'affaiblissement par rapport à *because*. Cependant, de même que *because* peut être prononcé de façon plus ou moins emphatique, *cos* admet un vaste éventail de prononciations réparties le long de l'axe du continuum entre réalisations canoniques et réalisations co-articulées. Cet axe est sans doute un peu moins long pour *cos* que pour *because*, qui ne connaît pas, dans notre corpus, de réalisation véritablement réduite. La Figure 5.1 propose une représentation de cette répartition des formes le long de l'axe de l'articulation.

Cette expérience n'a pas l'ambition de donner des résultats fermes mais simplement de suggérer des tendances à confirmer sur davantage de données. Au vu des

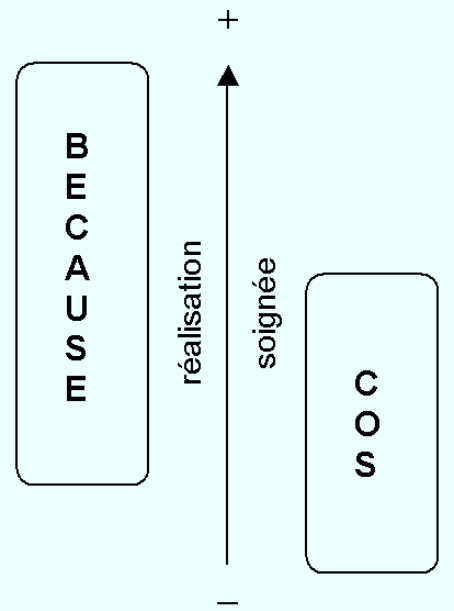

FIG. 5.1 – Articulation comparée de *because* et de *cos*

premières indications qu'elle fournit, nous sommes confortée dans l'idée que la grammaire de l'oral offre aux locuteurs deux marqueurs de causalité aux potentialités expressives relativement proches. Si *cos* peut être ici considéré comme une forme réduite de *because*, c'est parce que ce dernier ne présente pas, dans notre corpus, de forme totalement affaiblie.

5.2.3 Caractéristiques distributionnelles communes

Notre corpus contient au total 144 occurrences de *because* et *cos*. Après un premier dépouillement, nous ne retenons que 85% des énoncés concernés, soit 122, pour leur forme syntaxique suffisamment aboutie pour les rendre propices à l'analyse : au minimum, les places du sujet et du verbe des deux propositions y sont instanciées (à de rares exceptions près). Cette proportion est la même pour les énoncés en *because* et les énoncés en *cos*. Certaines caractéristiques distributionnelles communes nous éclairent sur le fonctionnement de la subordination causale, sinon à l'oral, du moins dans notre

corpus. C'est le cas en particulier de l'ordre d'apparition des propositions, ainsi que de la présence et de la position des pauses.

5.2.3.1 Ordre des propositions

Subordonnée initiale, finale ou médiane R. Quirk *et al.* répertorient trois positions possibles pour la subordonnée causale par rapport à sa principale : une position initiale (q *because/cos* p), une position finale (p *because/cos* q), et une position médiane (où la proposition *because/cos* q s'insère entre deux constituants de p). Leur dépouillement d'un extrait de 100 000 mots du corpus London-Lund indique une très nette préférence des locuteurs pour la position finale, choisie dans 97% des cas où *because/cos* est utilisé.[9]

Seules les deux premières positions sont représentées dans notre corpus. De façon écrasante, comme dans le corpus London-Lund, la place typique de la proposition subordonnée y est finale. Cet ordre se retrouve dans plus de 96% des cas, indiquant ici encore une forte restriction du champ des possibles. Seuls quatre exemples présentent l'ordre inverse. Nous n'en donnerons qu'un seul. Dans l'exemple suivant, Helen explique qu'en l'absence de Sarah, sa collègue lectrice, elle n'aura personne à qui parler pendant les heures précédant son dernier cours de la journée. La subordonnée causale précède la proposition principale :

H *hm* and then **because** Sarah {} didn't have to teach now she's not gonna come in

(Helen&Rebecca : Staying awake till 7, *because_028*)

Collaboration : subordonnée produite par l'interlocuteur La subordination causale à l'oral présente une caractéristique supplémentaire par rapport à l'écrit quant à l'apparition des deux propositions. Celles-ci peuvent en effet être prononcées par

[9]QUIRK, Randolph *et al.* (1985). *A Comprehensive grammar of the English language.* London : Longman. (§15.47, pp.1106-1107).

deux locuteurs différents. D. Biber[10] rend compte de ce phénomène en prévoyant une autre distribution possible des propositions. Il décompte ainsi les subordonnées en positions initiale ou finale ainsi que celles dont la principale est prononcée par un locuteur différent. De tels cas de collaboration entre les locuteurs pour la production d'une forme que la grammaire traditionnelle traite comme un tout se retrouvent dans notre corpus. L'ordre d'apparition des propositions y semble être contraint, la subordonnée venant compléter une principale préalablement énoncée.

Notre typologie prévoit, outre les subordonnées finales et initiales, les subordonnées apparaissant seules. Parmi celles-ci, toutes sauf une, soit en tout quatre exemples, relèvent de ce phénomène de collaboration. Citons l'exemple *because_027* que nous avons convoqué plus haut pour illustrer l'utilisation du pronom interrogatif *why* en tant qu'introducteur d'une interrogative directe.[11] A la question posée par Helen, Rebecca répond par une proposition en *because* seule, dont la proposition principale est à reconstruire à partir du présupposé contenu dans l'interrogative.

Dans l'exemple *cos_012*, James explique que, bien que linguiste, il désigne le bâtiment partagé par les étudiants de linguistique et de chinois comme l'Institut de chinois. Son interlocuteur tardant à trouver ses mots pour en décrire les raisons, Celia suggère une explication. La proposition qui nous intéresse ici est celle de Celia (en caractères gras dans l'extrait reporté plus bas), qui se rattache à une principale prononcée par James. La subordonnée de Celia entre en concurrence avec une construction similaire produite presque simultanément par James, les deux locuteurs étant engagés dans le même effort d'explicitation au même moment. Remarquons que le contenu des deux propositions subordonnées qui dépendent de la même principale concorde : James désigne le bâtiment comme l'Institut de chinois parce que c'est d'abord en tant qu'Institut de chinois qu'il l'a connu lorsque sa sœur le fréquentait.

[10]BIBER, Douglas (éd.) (1999). *Longman Grammar of Spoken and Written English.* Harlow, Essex : Longman. (tableau 10.9).

[11]Voir p.224.

J *so I'm quite strange* I'm a *linguist* who insists on referring to the building as the Chinese institute <rires>

C *mm*

J cos that's how <rires>

C ***cos you::***

J *how I used* to *think of it*

C ***that's how you* kn:: *met it fir::st***

(Celia&James : Facilities, *cos_012*)

Au sein des subordonnées finales, nous considérons comme participant du même phénomène de collaboration celles dont les principales se résument à une marque d'acquiescement ou de dénégation et ne contiennent pas de verbe. Dans 11 d'entres elles, *yes, yeah* ou *no* permettent de reprendre le contenu de l'énoncé précédent en lui ajoutant une modalité assertive et servent de support à une proposition causale. Dans certains exemples, cet énoncé sollicite parfois par lui-même une explicitation ou du moins une réponse. Ici (*because_020*), James cherche à savoir si Celia, la secrétaire du laboratoire, doit obtenir l'accord de John, le directeur, avant de délivrer la clef des locaux à un étudiant. Celia donne une réponse négative catégorique avant même que James ait terminé sa question. Etonné, il incite son interlocutrice à développer les raisons d'une telle certitude. Comme souvent dans le cas des *yes/no questions*, la réponse est circonstanciée et s'accompagne d'une explication :

J *do you technically* need John's sort of permission *then*

C *no*

J to get

C hm no

C no

J <h> (0,35) ok <h> (0,55

C **no because::** I mean {0,50} I know he wants you to have keys so that's fine

(Celia&James : John's permission, *because_020*)

Dans l'exemple suivant, James fait part à Celia de son manque d'expérience lorsqu'il s'agit d'organiser des tests perceptifs. Il n'est particulièrement pas satisfait de sa séance de la veille. Il raconte qu'il a oublié de mettre en route l'enregistrement de

l'un de ses sujets, dont l'accent ne correspondait pas à ses attentes. Celia, cherchant à relativiser la portée cette maladresse, suggère que le hasard a bien fait les choses : c'est aussi bien comme ça, dit-elle. James la détrompe en disant que ses problèmes n'étaient pas terminés dans la mesure où il ne peut pas non plus exploiter les résultats de la personne suivante qui n'a pas compris la tâche qu'il lui demandait. Il implique ainsi que le hasard aurait également pu faire qu'il n'enregistre pas non plus la suivante.

C *well just* as well that was the one you didn't turn the tape *on*

J ***yeah* well not really cos** the *next girl was* <h> {0,60}

C *wasn't ideal*

J I have to bin as well because she <h> {0,55} *she got this <rires>*

(Celia&James : Diphthongs 1, *cos_023*)

La réplique de James est complexe. Tout d'abord, celui-ci accuse réception de la remarque de Celia avec *yeah*. Ensuite, il en relativise la validité en utilisant le ligateur *well*, qui introduit le rejet de la validation du lien prédicatif. La proposition causale vient justifier ce pourquoi la remarque de Celia n'est pas totalement (*not really*) recevable et invoque des circonstances connues de James seulement. Il semble que l'utilisation de *really* permette de suggérer que la remarque de Celia a quelque chose de valable, mais que le champ d'application de la relation prédicative n'est pas défini de façon satisfaisante et doit être envisagé comme plus vaste.

	CORPUS FP	LGSWE
Subordonnée initiale	4%	5%
Subordonnée finale	84%	90%
Collaboration	12%	5%

TAB. 5.1 – Ordre d'apparition des propositions en *because/cos*

L'ordre d'apparition des propositions en *because/cos* de notre corpus, comme dans celui de la LGSWE, voit très majoritairement s'imposer la subordonnée finale.[12]

[12]Nous distinguons à présent les subordonnées finales prononcées par le même locuteur et par un

Au Tableau 5.1 figurent, aux côtés des chiffres relatifs à notre corpus, ceux donnés par D. Biber.[13] Nous remarquons que les cas de collaboration ont une proportion légèrement supérieure dans notre corpus. Ce chiffre est sans doute lié au fait que nous considérions les séquences du type *yes because...* ou *no cos...* comme des énoncés dont la subordonnée et la principale sont prononcées par des locuteurs différents. Une autre explication provient de la petite taille de notre corpus qui peut tendre à voir une sur-représentation de certains phénomènes. D'une manière générale, cependant, les deux distributions sont très semblables et tendent à confirmer qu'une réduction du champ des possibles s'opère à l'oral pour orienter les locuteurs vers des choix largement prévisibles.

5.2.3.2 Présence et place des pauses

L'une des fonctions de l'intonation est de segmenter le continuum sonore en unités de sens, ces unités prenant généralement la forme d'unités syntaxiques. Ainsi, comme on le trouve dans les grammaires traditionnelles, dans le cas de la subordination, l'intonation souligne les charnières entre unités et les transitions entre niveaux hiérarchiques. La pause silencieuse est largement invoquée dans son rôle de segmentation entre principale et subordonnée circonstancielle. Qu'en est-il dans notre corpus de la présence et de la place des pauses dans les énoncés en *because/cos* ?[14]

Présence des pauses Commençons tout d'abord par évaluer la présence de pauses dans nos exemples, indépendamment de toute considération fonctionnelle : les pauses que nous relevons donc ici peuvent aussi bien servir à la respiration, l'hésitation ou à la focalisation qu'à la segmentation. Nous remarquons que seuls la moitié des 122 énoncés aboutis en *because/cos* (60 en tout, dont 16 contiennent également au moins

locuteur différent.

[13]BIBER, Douglas (éd.) (1999). *op. cit.* (tableau 10.9).

[14]Les pauses silencieuses dont nous tenons compte sont au minimum longues de 0,20 s.

une pause pleine) contiennent au moins une pause silencieuse. Les observations que nous faisons ici n'ont donc pas la prétention d'être généralisables. Elles ne peuvent que décrire certaines tendances au sein de notre corpus.

Enoncés sans pause silencieuse Citons quelques énoncés dont aucune des deux propositions entrant en jeu dans la structure causale ne contient de pause. James revient ici sur le fait que l'anglais ne soit pas une langue très couramment parlée en Chine, du fait, pense-t-il, du faible développement du tourisme international :

J I suppose that probably is right **cos** they obviously don't see that many

(Celia&James : Beer, *cos_018*)

Dans l'exemple *because_026*, Helen et Rebecca réfléchissent à l'organisation de leur emploi du temps très chargé pour les jours suivants. L'énoncé ne fait pas apparaître de pause :

R *we'll do* that **because** I think *I don't feel like it*

(Helen&Rebecca : Messages, *because_026*)

Enoncés avec pause silencieuse Puisque c'est la présence d'une pause silencieuse qui nous intéresse ici, tournons-nous à présent vers les énoncés qui en contiennent. Dans l'exemple suivant, Rebecca explique à Helen que le père de son ami Franck a toujours beaucoup poussé son fils :

R *and he's* sort of living his dreams through his son **because** <**h**> **(0,35)** Franck went to:: Westminster {**0,25**} at the age of seven

(Helen&Rebecca : Westminster, *because_033*)

Rebecca fait une pause à deux reprises : une première fois juste après le marqueur de subordination, sous la forme d'une respiration audible, et la seconde, plus loin dans la subordonnée, avant d'ajouter un détail supplémentaire permettant de compléter le portrait de son ami d'enfance.

Enoncés avec pause pleine Les pauses pleines, dont la fonction ne relève pas de la segmentation, apparaissent très minoritairement seules (dans 4 énoncés), comme dans l'exemple suivant. Mike, tout juste rentré de vacances, regrette de ne pas pouvoir montrer ses photos à Hannah, d'autant plus qu'elles doivent déjà avoir été développées. Cet énoncé présente une pause pleine (notée *er*) accompagnant la répétition du sujet grammatical de la proposition subordonnée en *because* :

M it's a pity I haven't got the photos **because** they **er** they're supposed to be back [inaudible] anyway

(Mike&Hannah : Trip, *because_036*)

Plus souvent, la pause pleine apparaît en conjonction avec au moins une pause silencieuse. Une partie des 60 énoncés évoqués plus haut contiennent également au moins une pause pleine. Les deux exemples suivants font partie des 16 énoncés concernés, déjà comptabilisés parmi les 60 contenant au moins une pause silencieuse. Dans l'énoncé *because_013*, Celia récapitule quelles clés elle doit donner à James pour qu'il puisse accéder au laboratoire de phonétique en dehors des heures de bureau. Sa longue énumération occasionne de nombreuses pauses pleines et silencieuses dans la principale et entre les deux propositions. Notons également la présence d'allongements syllabiques dans la subordonnée et d'autres nombreuses marques du travail de formulation :

C well what I'll have to give you will be the **mm** {**0,40**} key to the door of forty-one:: {**0,25**} **hm** {**0,50**} <h> either the back or the front door I don't know what I've got *spares of* <h> {0,50}

J *okay*

C **hm** <**h**> {**0,65**} and:: the key for the room for the doors in the lab {**0,70**} <**h**> **because** they **all:: we::** lock them up when we go

(Celia&James : Keys, *because_013*)

Aucune différence significative n'est à observer entre *because* et *cos* quant à la présence ou à l'absence de pauses dans un énoncé, ni quant au type de pause relevé : la moitié des énoncés contiennent au moins une pause, celle-ci étant très largement une pause silencieuse. Cependant, nos observations ne permettent pas de distinguer

les pauses de segmentation des autres pauses. Le critère du site d'occurrence est sans doute un bon indicateur du rôle grammatical des pauses silencieuses accompagnant les propositions en *because/cos* de notre corpus.

Place des pauses : à la charnière syntaxique des propositions ? Parmi les 44 énoncés contenant au moins une pause silencieuse mais aucune pause pleine, quels sont ceux dont au moins une des pauses occupe la place grammaticale située entre les deux propositions ? Quels sont ceux, donc, dans lesquels la pause silencieuse joue un rôle de segmentation ? Sont-ils les plus nombreux ?

Enoncés avec une seule pause Notre relevé prend en compte des données complexes. Il envisage plusieurs sites d'occurrence pour les pauses silencieuses :

- entre la principale et la subordonnée (site 1) ;
- dans la subordonnée (site 2) ;
- ou dans la principale (site 3).

Parmi les 37 énoncés dont la pause silencieuse occupe un seul site, citons quelques exemples illustrant chacun de ces trois sites.

Dans l'exemple suivant, la pause segmente l'énoncé à la charnière entre les deux propositions, conformément à nos attentes. Helen raconte avec humour qu'elle a pris des cours de poterie dans son enfance et qu'elle a notamment fait un lion que sa mère a conservé. La courte pause silencieuse située entre les deux propositions donne l'occasion à Rebecca de rire :

H but my Mum's still got it
R <rires>
H **cos** it's such a great work of art

(Helen&Rebecca : Art, *cos_037*)

Ici, la pause silencieuse n'occupe pas la charnière des propositions mais se situe dans la subordonnée. James explique qu'il va tout de même pouvoir se servir de son expérience

phonétique assez désastreuse de la veille sous la forme d'une étude pilote :

J yeah I turned it on I thought well I'll record this and it might be useful anyway **because** it <**h**> {**0,50**} probably going to have to convert the whole thing into a pilot anyway

(Celia&James : Experiment, *because_004*)

Enfin, aucune pause n'occupe la place grammaticale attendue ici non plus. Mike explique que, malgré les apparences spartiates de la randonnée, Ann et lui ne se levaient pas aux aurores. C'est la principale qui accueille la pause silencieuse relevée dans cet énoncé :

M we would start really leisurely I mean we'd {**0,25**} wake up when we woke up which was usually **cos** in a gorge it's not till nine o'clock that the sun hits you or ten even sometimes but we'd usually wake up at nine

(Mike&Hannah : Storm, *cos_059*)

Enoncés avec plusieurs pauses Pour les quelques énoncés contenant des pauses à plusieurs endroits, les sites combinés sont au nombre de quatre :

- entre la principale et la subordonnée et dans la subordonnée (sites 1 et 2) ;
- entre les deux propositions et dans la principale (sites 1 et 3) ;
- entre les deux propositions et dans les deux propositions (sites 1, 2 et 3) ;
- ou seulement dans la subordonnée et la principale (sites 2 et 3).

Avec 4 énoncés, la première configuration est la plus courante. Ici, une première pause segmente l'énoncé à la charnière entre les deux propositions, et d'autres émaillent la subordonnée. Mike raconte que l'orage provoque l'apparition d'un précipité rouge dans la rivière le long de laquelle Ann et lui marchent :

M and then the river just turned red {**0,55**} **because** it brings down this sort of {**0,3**} bauxite {0,7} hm I don't know what it is some mineral {**0,4**} and <**h**> **(0,4)** {**0,3**} it dissolves in the water and turns the water dark red *and it looks like*

H *you're joking*

M blood

(Mike&Hannah : Storm, *because_050*)

Les trois autres sites combinés correspondent chacun à un énoncé.

Bilan provisoire concernant la place et la présence des pauses La distribution des pauses révèle des résultats intéressants. Seuls 15 énoncés sur les 122 (soit 12%) de notre corpus contiennent une pause située entre la principale et la subordonnée, c'est-à-dire à l'endroit le plus attendu. Dans 9 cas, il s'agit d'un énoncé à pause unique. Si l'on considère que notre corpus propose plus d'une centaine d'exemples, la segmentation des propositions par la pause est loin d'être répandue. Une rapide consultation du BNC confirme cette tendance : sur près de 37 000 énoncés produits à l'oral et comportant *because* ou *cos*, seuls 1 900 sont répertoriés comme contenant une pause silencieuse juste avant le marqueur, soit une très faible proportion (5%).[15] En revanche, nos résultats ne sont pas conformes avec ceux présentés par F. Grosjean & A. Deschamps. Ces derniers établissent que 23% des phrases du corpus utilisé pour leur article de 1975 (interviews radiophoniques) sont suivies d'une proposition subordonnée (sauf relative) et que, dans 60% des cas, la subordonnée est précédée d'une pause.[16] Comme nous avons pu le voir, notre corpus révèle des tendances très différentes.[17] Le Tableau 5.2 résume le nombre d'énoncés contenant au moins une pause silencieuse dans chacun des sites évoqués, et l'importance de chaque site par rapport au total d'énoncés en *because/cos*.

Si les pauses silencieuses relevées dans les énoncés où apparaissent *because* ou *cos* ne servent que très minoritairement à la segmentation des propositions, quelle est leur fonction ? Où les retrouve-t-on ?

Le Tableau 5.2 montre que, dans notre corpus, le site privilégié d'apparition d'une

[15] L'ordre des propositions selon lequel la principale précède la subordonnée n'épuise certainement pas les exemples du BNC. Cependant, compte tenu de la très faible proportion d'énoncés dont les propositions apparaissent dans un autre ordre dans notre corpus, le résultat de la requête à laquelle nous avons procédé donne une bonne évaluation de la part des énoncés en *because/cos* dont les propositions sont segmentées par une pause silencieuse.

[16] GROSJEAN, François & DESCHAMPS, Alain (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 31 : 144-184. (p.169).

[17] Notons que cette différence ne provient pas de notre seuil minimal de 0,20 s qui nous fait éliminer toute pause de durée inférieure. En effet, le seuil adopté par les deux chercheurs est encore supérieur au nôtre (0,25 s.).

SITE D'OCCURRENCE	NOMBRE	PROPORTION
Pas de pause :	78	63,9 %
Pause unique :		
Entre principale et subordonnée (site 1)	9	7,4 %
Dans subordonnée (site 2)	25	20,5 %
Dans principale (site 3)	3	2,5 %
Pauses multiples :		
Sites 1 et 2	4	3,3 %
Sites 1 et 3	1	0,8 %
Sites 1, 2 et 3	1	0,8 %
Sites 2 et 3	1	0,8 %

TAB. 5.2 – Place des pauses silencieuses dans les 122 énoncés en *because/cos* de notre corpus

pause silencieuse n'est pas le site grammatical de la charnière des propositions. La présence d'une pause de segmentation au sein d'un énoncé en *because/cos* est relativement peu fréquente : on la retrouve seulement dans 12% des exemples. En revanche, nous remarquons que l'intérieur de la subordonnée (site 2) est propice à l'apparition d'une pause silencieuse : dans un quart des énoncés en *because/cos*, la subordonnée contient une pause silencieuse. Ces pauses sont pour la plupart des pauses d'hésitation, sur lesquelles nous reviendrons plus loin.

Ces faits distributionnels communs aux marqueurs les plus courants que sont *because* et *cos* tendent à caractériser la subordination causale à l'oral : peu de variation dans l'ordre d'apparition des propositions (la principale précédant la subordonnée), et peu de segmentation des propositions par la pause (le site de l'intérieur de la subordonnée étant largement préféré à celui de la charnière des propositions). Cette uniformité peut néanmoins être relativisée au vu de certaines caractéristiques permettant de considérer *because* et *cos* comme deux marqueurs différents.

5.3 *Because* et *cos* : deux marqueurs à distinguer ?

Nous souhaitons à présent apporter des éléments permettant d'appuyer l'hypothèse selon laquelle *because* et *cos* sont véritablement distincts. Notre analyse part de l'observation d'un faisceau d'indices syntaxiques et prosodiques qui favorisent l'apparition de l'un des marqueurs plutôt que l'autre. Selon nous, et comme en témoignent les contextes distributionnels typiques dégagés pour chacun des marqueurs, *because* et *cos* mettent en jeu différemment la relation intersujets.

5.3.1 Emploi de *because* pour *cos*

Nous avons vu plus haut que, dans certains cas, *cos* apparaît dans un contexte de reprise du marqueur *because* énoncé précédemment par l'un ou l'autre des locuteurs. Nous en avons conclu que, la relation causale ayant déjà été annoncée par *because*, une reformulation peut se contenter de la forme dite « réduite » du marqueur, c'est-à-dire *cos*. Cependant, notre corpus montre que, lorsque les deux marqueurs apparaissent conjointement, *because* peut également servir à reprendre *cos*. Ces faits tendent à remettre en question l'hypothèse selon laquelle *cos* constitue une aphérèse et une reprise affaiblie de *because*.

Dans l'exemple suivant, James raconte ses difficultés à organiser des activités sportives avec ses camarades linguistes. Il dit penser donner l'impression de suivre un autre cursus tant son emploi du temps semble incompatible avec le leur. Il ouvre une subordonnée causale avec *cos* qu'il abandonne rapidement. Son projet explicatif est repris par Celia, qui propose une raison introduite par *because* :

J I mean I was to arrange er rowing last term and <h> {0,80} a- er {0,60} everyone would say we have an outing at six and I'd say I still have have lectures till then <rires> so they think I'm on some kind of evening class or something **cos** er <rires> or

C oh **because** they always got one late

(Celia&James : China, *because_005*, *cos_009*)

Peut-être ce phénomène est-il lié au fait que l'exemple présenté correspond à un « cas de collaboration », dans lequel la reformulation du propos ébauché par James est énoncée par Celia. Il tend néanmoins à montrer que la reprise d'un marqueur par l'autre n'est pas unidirectionnelle. Il pointe également vers une possible distinction des deux marqueurs quant aux valeurs qui leur sont associées. Selon nous, cette distinction est reflétée par une distribution qui différencie *because* et *cos* sur certains points.

5.3.2 Caractéristiques distributionnelles et valeurs intersubjectives divergentes

Au sein des tendances générales relevées plus haut, certains contextes favorisent l'apparition de l'un ou l'autre des marqueurs. Ces contextes qu'il faudra déterminer nous semblent témoigner de valeurs intersubjectives différentes associées aux marqueurs. Nous faisons ainsi l'hypothèse que *because* se situe davantage en rupture que *cos*, qui reste plus typiquement consensuel. Sur quels éléments cette hypothèse se fonde-t-elle ?

5.3.2.1 Sujet syntaxique des propositions

L'analyse des sujets syntaxiques de la proposition introduite par *because* ou *cos* et de la proposition principale à laquelle elle se rattache révèle des différences intéressantes.

Proche et lointain Dans les cinq textes de notre corpus, nous nous sommes concentrée sur certains types de sujets ayant comme référent des personnes (pronoms personnels et noms propres), par opposition à tous les autres sujets possibles (entrent dans la catégorie « autre » les groupes nominaux faisant ou non référence à des personnes, et les pronoms indéfinis). Nous avons ainsi pu établir trois catégories de sujets, dont les deux premières nous intéressent plus particulièrement. La première catégorie est

constituée de sujets relevant de la sphère du « nous », qui correspondent par exemple aux pronoms personnels sujets première et deuxième personne (au singulier et au pluriel). La deuxième correspond aux sujets relevant de la sphère du « eux », c'est-à-dire, par exemple, les pronoms sujets à la troisième personne. Enfin, nous avons rassemblé tous les autres sujets dans une troisième catégorie que nous appellerons « autre ».

Because et *cos* sont significativement différents dans notre corpus quant à la catégorie à laquelle appartient le sujet grammatical de la proposition principale ($\chi^2 =$ 7,691 pour 2 ddl, p $\leqslant$ 0,025).[18] Aucune différence n'est à remarquer concernant la catégorie « autre ». En revanche, nous constatons que *because* apparaît de façon significativement plus fréquente avec des propositions principales commandées par un sujet relevant de la sphère du « eux », tandis que les sujets relevant de la sphère du « nous » se retrouvent liés de façon privilégiée à l'occurrence de *cos*. Ainsi, en fonction de la proximité de la personne désignée par le sujet de la principale d'une subordonnée de cause apparaîtra plus volontiers l'un ou l'autre des deux marqueurs.

Typiquement, un énoncé en *because* et un énoncé en *cos* se présentent de la façon suivante en ce qui concerne le sujet de la proposition principale :

He/she/they + V ... + **because** + subordonnée

I/you/we + V ... + **cos** + subordonnée

Cos* et la sphère du *nous Cette tendance va dans le sens de notre hypothèse évoquée plus haut. S'inscrivant au sein de la sphère du « nous », *cos* apparaît comme un marqueur apte à exprimer une relation causale ou un contenu propositionnel dont la validité est supposée acquise ou n'est pas susceptible d'être mise en question. Dans l'exemple *cos_001*, James explique à Celia qu'une partie de ses recherches consiste en des expériences. Il se souvient que ses difficultés en travaux pratiques l'avaient

[18] Les effectifs pris en compte pour le test du χ^2 sont consignés dans le Tableau C.1, en annexe au présent chapitre (p.447).

justement fait arrêter les sciences lorsqu'il était plus jeune.

J I just remember that **I** actually gave up science at sixteen **cos** I was all so rubbish at the practicals <h> {0,75} so <rires> and I'm

C but this can't be exactly like practicals at school

J no but er you know if if you're sort of incompetent at one kind of practical it's <rires>

C *what you're you're bound to be in-*

J *and <rires> it spreads over*

(Celia&James : Experiments, *cos_001*)

Dans cet énoncé à la première personne, James évoque un événement de sa vie que Celia ne peut pas connaître. Les représentations mentales du co-énonciateur ne sont pas anticipées comme *a priori* conflictuelles : James n'a donc pas à construire d'argumentation détaillée autour de ce qu'il avance. En outre, son propos ne laisse place à aucune divergence de point de vue possible. La relation prédicative de p est posée comme validée : elle est inscrite par James dans le paysage consensuel à l'aide d'une subordonnée circonstancielle qui fait l'objet d'une prise en charge minimale de la part de l'énonciateur tant sa validation et le lien qu'elle entretient avec la principale sont posés comme évidents et incontestables. Dans de tels énoncés, la parole du sujet énonciateur et sujet dans l'énoncé suffit à garantir la validité de p et le lien entre p et q. Le marqueur *cos* contribue ainsi à mettre en relation des faits déjà acquis. La suite de notre exemple montre que Celia rebondit sur le paradoxe évoqué par James. Elle conteste la pertinence de l'exemple pris par James, mais ne revient pas sur la validité de la relation causale.

Because* et la sphère du *eux Au contraire, un contexte impliquant par exemple des personnes extérieures au couple formé par les interlocuteurs est construit sur des bases potentiellement moins consensuelles. C'est un contexte favorable à l'apparition d'un marqueur tel que *because*, qui permet d'anticiper et de travailler à combler une divergence dans les représentations mentales. Dans le passage suivant, Rebecca décrit la famille de son ami Franck. Contrairement à lui et à sa sœur, son frère n'a pas eu

une scolarité très réussie.

R yeah so he he's got one GCSE and all he's ever done is work in things like McDonald's and things but **presumably** <h> **he probably** feels like the complete black sheep of the family

H *yeah*

R ***because** everyone* else is <h> *really*

H *failure*

(Helen&Rebecca : Westminster, *because_034*)

Si Rebecca est la seule détentrice des informations qu'elle donne à Helen, sa position n'est pas celle d'un égocentrage qui place le contenu de p hors débat. Elle ne le considère pas non plus comme d'emblée consensuel mais le présente de façon à gagner l'assentiment de son interlocutrice. Comme en témoignent les deux adverbes utilisés successivement (*presumably* et *probably*), elle n'expose plus des faits mais formule une hypothèse au sujet de la psychologie du frère de Franck. Engagée sur un terrain qu'elle sait incertain, elle s'en remet à son interlocutrice pour juger de la validité de sa remarque. Sa proposition en *because* apporte une circonstance qu'elle soumet à Helen permettant de valider la relation prédicative dans p. Se construit ici une argumentation visant à obtenir de Helen qu'elle admette au sein du domaine consensuel un élément qui n'en fait pas *a priori* partie et que Rebecca reconstruit sous ses yeux. Remarquons que cette construction d'un consensus avec *because* se retrouve sur le plan interactionnel avec l'intervention de l'interlocutrice. Helen reprend en effet à son compte les propos de Rebecca et leurs implications. Dans sa bouche, le frère de Franck ne se considère plus comme une brebis galeuse (*black sheep*) mais comme un raté (*failure*). Avant même que Rebecca ne soit arrivée au bout de sa subordonnée, Helen reformule la proposition principale et accepte ainsi son contenu au sein du paysage partagé par les deux interlocutrices.

5.3.2.2 Position des pauses

Comme nous avons pu le voir plus haut, les énoncés en *because* ou *cos* ne présentent pas de différence quant à la présence d'une pause, qu'elle soit silencieuse ou pleine. En revanche, une différence significative existe entre les deux marqueurs en ce qui concerne la position des pauses qui les accompagnent.

Deux sites privilégiés Les pauses relevées dans les énoncés en *because* et *cos* sont caractérisées par une distribution significativement différente en fonction du marqueur (χ^2 = 8,969 pour 3 ddl, p $\leqslant$ 0,05).[19] Elles permettent ainsi de définir des contextes d'apparition distincts pour les deux marqueurs. Les écarts concernent essentiellement les pauses apparaissant entre les deux propositions et celles situées à l'intérieur de la proposition subordonnée. Dans les énoncés en *because*, les pauses se situent de façon privilégiée dans la subordonnée. Certes l'intérieur de la subordonnée constitue aussi un site d'occurrence fréquent pour les énoncés en *cos*, mais de façon moins saillante. En revanche, *cos* apparaît plus volontiers que *because* dans un énoncé complexe dont les deux propositions sont séparées par une pause.

Le schéma typique relatif à la position des pauses dans les énoncés en *because* et en *cos* est donc le suivant :

principale + **because** + **{pause}** + S + V ...

principale + **{pause}** + **cos** + S + V ...

Remarquons que les pauses des énoncés en *because* se retrouvent à n'importe quelle place à l'intérieur de la subordonnée. La notation dans notre schéma indique que la pause s'intercale entre le marqueur et le reste de la proposition, conformément à une tendance relevée dans notre corpus : le marqueur tend à être ainsi isolé du reste de la subordonnée.[20]

[19] Les effectifs pris en compte pour le test du χ^2 sont consignés dans le Tableau C.2, en annexe au présent chapitre (p.447).

[20] Ce phénomène est également mentionné dans COUPER-KUHLEN, Elizabeth (1996). Intonation

Pause, consensus et élaboration linguistique Cette distribution est pour nous l'indice d'une plus grande complexité et d'une consensualité moindre dans les énoncés en *because* par rapport à ceux en *cos*. En effet, les pauses accompagnant *because* sont davantage liées au travail de formulation. En dehors du site grammatical de la charnière entre les deux propositions, la pause silencieuse, ou, à plus forte raison, pleine, est principalement une pause d'hésitation. Elle permet à l'énonciateur de prendre davantage de temps pour l'élaboration d'un propos moins attendu par le co-énonciateur ou plus complexe à décrire, et qui nécessite donc davantage de précaution dans la formulation.

C'est le cas de l'exemple suivant où Mike explique qu'un matin, Ann et lui se réveillent sur un banc de sable au milieu de la rivière et se retrouvent nez à nez avec des villageois à dos de mule.

M there was this stream of people going past on their mules **because** {**0,4**} that was the only way to get to one of the villages it's this tiny village

H aha

M which has no road to it and so people just take their mules along the river bed <h> (0,8) and just as we were doing really

(Mike&Hannah : Mules, *because_054*)

L'utilisation de *because* permet à Mike selon nous de ne pas présenter le contenu de q et la relation qu'il entretient avec celui de p comme évidents mais de les construire en T_0 pour son interlocutrice : il vient au devant de représentations mentales divergentes qu'il attribue à Hannah. Si la subordonnée n'est pas complexe en soi, on constate que Mike choisit d'étendre son explication au-delà de la structure causale même. En effet, le village est conservé comme point de départ d'un développement plus long. Il nous semble que la pause silencieuse qui suit immédiatement *because* met en valeur le lien causal avant même de fournir les éléments qui l'étayent et permet de préparer une explication relativement élaborée qui déborde ici du cadre de la proposition subordonnée.

and Clause Combining in Discourse : the Case of Because. *Pragmatics* 6 (3) : 389-426. (p.422).

Cette intuition d'une plus grande complexité des énoncés en *because* est confirmée par une attention portée aux autres marques du travail de formulation et aux marques de prise en charge.

5.3.2.3 Formulation et prise en charge

A la position des pauses pleines et silencieuses s'ajoute un faisceau d'indices supplémentaires permettant de distinguer *because* et *cos*. Ceux-ci concernent principalement la place occupée par le travail de formulation dans les énoncés en *because*. Nous avons en effet procédé à un relevé systématique de marques qui nous semblent témoigner d'un important travail de mise en mots, en nous intéressant notamment à la densité de telles marques au sein d'un même énoncé. Parmi celles-ci, nous comptons tout d'abord les répétitions, corrections et faux départs et les allongements d'hésitation. A cette liste, nous avons ajouté les pauses pleines et les pauses silencieuses situées en dehors du site grammatical représenté par la suture entre les deux propositions. Enfin, nous avons également dénombré les ponctuants phatiques, tels *you know* ou *I mean*. Ceux-ci permettent non seulement de retarder l'apparition d'un contenu sémantique en cours d'élaboration, mais aussi, dans une problématique argumentative, de signaler une rectification ou une explicitation de faits pris en charge par l'énonciateur (*enfin, je veux dire...*) ou, sinon de s'assurer que celui-ci fait l'objet d'un consensus, du moins de tenter de le rendre consensuel (*mais si, tu sais bien...*).

Présence et densité de ces marques Dans notre corpus, *because* et *cos* sont significativement différents quant à la présence ou l'absence de marques du travail de formulation ($\chi^2 = 15{,}220$ pour 1 ddl, $p \leqslant 0{,}001$).[21] Ainsi, les énoncés en *because* s'accompagnent plus souvent de telles marques que les énoncés en *cos* : plus de 75% des énoncés en *because* comportent au moins une des marques décrites plus haut,

[21] Les effectifs pris en compte pour le test du χ^2 sont consignés dans le Tableau C.3, en annexe au présent chapitre (p.448).

contre 40% pour *cos*.

Considérons à présent la densité de telles marques au sein de chaque énoncé, c'est-à-dire le nombre d'énoncés comportant 0, 1, 2, 3, et 4 marques ou plus. Les données brutes de notre relevé font apparaître encore une différence significative entre les deux marqueurs ($\chi^2 = 19{,}352$ pour 4 ddl, $p \leqslant 0{,}001$).[22] Le Tableau 5.3 présente sous forme de pourcentage la distribution des marques du travail de formulation dans les énoncés en *because* et *cos* dans notre corpus. Les énoncés en *because* sont non seulement plus nombreux à présenter des marques du travail de formulation, mais celles-ci y apparaissent en plus grand nombre que dans les énoncés en *cos* : l'accumulation des marques est particulièrement flagrante lorsque l'on compare la proportion d'énoncés comportant au moins quatre marques (26,5% pour les énoncés en *because*, contre 9,5%).

	BECAUSE	COS
Enoncés sans marque	22,5 %	58,1 %
Enoncés avec 1 marque	14,3 %	16,2 %
Enoncés avec 2 marques	20,4 %	10,8 %
Enoncés avec 3 marques	16,3 %	5,4 %
Enoncés avec 4+ marques	26,5 %	9,5 %

TAB. 5.3 – Densité des marques du travail de formulation dans les énoncés en *because* et en *cos*

***Because* et la complexité du propos** Ainsi, d'après nos données, *because* apparaît typiquement dans un contexte où l'élaboration du propos pose problème, rendant le travail de formulation particulièrement visible. Comme nous avons pu le voir plus haut, le débit des énoncés en *because* est en outre plus lent.[23] Dans le long exemple suivant, nous relevons successivement un ponctuant (*I mean*) introduisant

[22] Voir Tableau C.4 (p.449).
[23] Voir 317, 231.

une incise dont la construction n'aboutit pas (*there was* constitue un faux départ par rapport à *when you have*), un allongement vocalique d'hésitation (*a::*) suivi de la forme corrigée du déterminant indéfini (*an*), une pause pleine (*er*) et une pause silencieuse d'hésitation longue de 0,70 s, et enfin une répétition de *to see*, qui permet d'unifier une principale complexe et décousue et de lui subordonner une proposition en *because* :

J so I still listened to see what she did there ***I mean there was** <rires>*

C *<claquement de langue> God*

J when you have **a::**

C maddening

J **er** {**0,70**} **an** onset cluster **to see** whether

C *mm*

J *she would* split it **because** that's what we were testing in the original experiment

(Celia&James : Experiments, *because_019*)

Outre la préférence pour des sujets aux référents étrangers au couple des interlocuteurs, et pour une pause dans la subordonnée et non pas à la charnière des propositions, *because* se distingue donc encore par son apparition privilégiée dans un contexte où les marques de prise en charge et les marques du travail de formulation abondent. Ces différences de distribution correspondent à nos yeux à une mise en jeu différente des représentations mentales, sur laquelle nous reviendrons plus loin, entre un *because* en rupture et un *cos* consensuel. Elles témoignent potentiellement d'une différence dans la relation entre les propositions liées en fonction du marqueur utilisé.

5.4 Sur la relation entre p et *because/cos* q

5.4.1 Restriction et opérations syntaxiques

5.4.1.1 Indicateurs d'une relation restrictive dans notre corpus

Certains tests syntaxiques permettent de déterminer la portée de la subordonnée.[24] Les opérations en question concernent par exemple l'antéposition de la subordonnée, sa mise en valeur polémique, la transformation de la principale en une interrogative partielle en *why*, la prémodification du marqueur, ou encore son utilisation dans une locution prépositionnelle. Examinons les quelques énoncés de notre corpus qui indiquent d'emblée par leur forme syntaxique la nature de la relation existant entre les deux propositions.

Ordre des propositions Une première indication de la portée de la subordonnée provient de l'ordre des propositions : seule la subordonnée explicative peut être antéposée. Comme nous l'avons dit plus haut, notre corpus contient très peu d'exemples d'énoncés en *because* ou *cos* dans lesquels la subordonnée précède la principale. Dans quatre cas seulement, on peut affirmer, au vu de l'ordre des propositions, que la portée de la subordonnée est restrictive.[25] Citons ici deux exemples. Dans le premier, Mike explique pourquoi Ann et lui se sont réveillés un matin sur un banc de sable au milieu de la rivière. Dans cet énoncé tortueux, la subordonnée (*because it was almost dark by the time we'd got to that camp site*) précède la principale à laquelle elle se rapporte (*we thought we'd found a nice little beach*) et dont elle fournit une circonstance restrictive permettant la validation. C'est parce qu'il faisait noir qu'ils pensaient avoir trouvé une jolie petite plage :

M and but it was okay because then we woke up and it was very strange cos we woke up and we were lying <h> (0,5) on what we thought {0,2} **because** it was almost dark by the time we'd got to that

[24] Voir p.199.
[25] Voir les exemples cités plus haut (p.233).

{0,2} camp site <h> (0,5) and er we thought we'd found a nice little beach <h> (0,4) and in fact it was like a sort of little island almost cos it was a little river there we were <h> (0,4) sleeping on this little sand bank in the middle of this river

(Mike&Hannah : Mules, *because_053*)

Dans le deuxième exemple, c'est une proposition subordonnée en *cos* qui précède la principale :

R I think what happened was she tried to phone you on your mobile telephone but if i::t rang twice
H right
R and **cos** it is as easy on a computer *I could find out she had left a message*

(Helen&Rebecca : RFI, *cos_026*)

La nature et la taille de nos données ne nous permet pas de discerner une différence de fonctionnement entre les deux marqueurs quant à l'ordre des propositions.

Mise en relief polémique de la subordonnée La mise en relief polémique d'une proposition subordonnée, propre aux propositions adjointes (selon la terminologie de R. Quirk *et al.*[26]) peut prendre plusieurs formes.

On pense tout d'abord au clivage en *it*. Cette forme syntaxique, proche de *the reason is...*, est rare dans notre corpus en général, et donc parmi les propositions causales. Nous relevons un seul exemple, construit avec *because* :

M that was like day four or something we {1} hm {1,5} we were getting a bit tired or something I just ? felt that we w- we were losing that sense of {0,3} magic and beauty <h> (0,3) I think **it was *because** the weather was bad*

(Mike&Hannah : Storm, *because_051*)

La clivée permet d'identifier la cause de la lassitude gagnant les deux randonneurs. Le présupposé qui, dans ce type de structure, est exprimé dans la subordonnée clivée est construit en discours dans le contexte immédiat de la clivée. Cette proximité permet l'ellipse de *that we were losing that sense of magic and beauty*. Cette structure est une réélaboration d'un énoncé tel que : *we were losing that sense of magic and*

[26] Voir p.199.

beauty because the weather was bad, I think. La valeur contrastive du focus donne une coloration résolument restrictive à la relation entre les propositions de l'énoncé d'origine. Construite sur l'opposition de deux valeurs, la clivée élimine d'emblée toute autre cause possible de la lassitude.

La négation ou l'interrogation contrastive permettent également la mise en relief d'une subordonnée et témoignent de sa relation restrictive avec la principale dont elle dépend. Ici encore, notre corpus donne peu d'exemples. Citons l'exemple *cos_032*, où Helen, qui cherche un emploi, rapporte au discours direct les propos que lui a tenus sa mère au téléphone. Celle-ci lui recommande de ne pas accepter n'importe quel travail :

H my and my mum said on the phone yesterday in- that she'd being thinking about it and she was gonna say to me <h> don't just apply for any odd job

R *no*

H *just cos* you think oh I can do it

(Helen&Rebecca : Jobs, *cos_032*)

Ici encore, deux valeurs sont opposées : la valeur rejetée (se croire apte à la tâche requise) pointe en creux vers une valeur non exprimée (par exemple, être intéressée par la tâche requise).

L'exemple *because_001* relève d'un même phénomène de mise en relief contrastive. Ce passage contient les premières répliques du corpus Celia&James. Peu convaincue que l'enregistrement auquel elle prend part puisse convenir à des recherches, Celia vient de nous inviter à profiter de la présence d'Esther Grabe dans le laboratoire pour lui demander de nous prêter des corpus de conversation. Après notre départ, l'échange entre les deux locuteurs révèle un malentendu :

J did you say go and talk to Esther if they want some conversation {0,25} *is that <rires>*

C *<h> (0,25) no* {0,70} no **because** Esther's hm {0,35} working in the same sort of field as they are

J right

C and she's got material for them {1,15} so she's got recordings that are going to help them so::

(Celia&James : Esther, *because_001*)

Le malentendu réside sur le terme *conversation*. James suggère par sa question inaboutie (*is that*) et par ses rires gênés que Celia a peut-être été désobligeante en nous recommandant d'aller passer le temps à bavarder avec Esther pendant l'enregistrement. La réponse de Celia rejette l'allégation de James (non, je n'ai pas été désobligeante) et lui oppose la seule raison justifiant la recommandation qu'elle nous a adressées. Remarquons que Celia éclaircit par deux fois ce qu'elle entendait par le terme de conversation (*material* et *recordings*).

Réponse à une question en *why* Comme nous avons pu le voir au début de ce chapitre, très peu de *because* ou *cos* sont directement liés à *why* dans notre corpus. Nous ne relevons qu'un seul exemple, construit avec *because*.[27] La réponse à la question partielle entretient une relation restrictive avec la principale dont Rebecca fait l'économie.

Prémodification du marqueur Phénomène rare dans notre corpus, la prémodification du marqueur par un adverbe est également le témoin d'une relation restrictive unissant les propositions. Dans l'exemple *cos_032* que nous venons d'évoquer, le marqueur est modifié par l'adverbe *just* qui renforce la relation entre les propositions. La portée locale de la subordonnée est également suggérée par *often* dans l'exemple *because_011*. James y évoque les difficultés rencontrées par sa sœur et ses collègues étrangers. Il en identifie pour raison principale leur maîtrise insuffisante de la langue chinoise. Notons la répétition de l'adverbe qui tend à resserrer le lien entre les propositions.

J they're always being taken off to places
and they often don't *really* know what's going to happen
C *mm*
J **often because** there's a slight language problem

(Celia&James : Beer, *because_011*)

[27] Voir l'exemple *because_027*, p.224.

Le même phénomène se lit dans un exemple proche, bien que ne présentant pas de prémodification du marqueur :

J actually so I know I **only** know about the Chinese people **cos** my sister actually does Chinese

(Celia&James : Facilities, *cos_089*)

Only opère une restriction sur le champ d'application de la subordonnée causale qu'il appelle et qui seule permet de valider la relation prédicative dans la principale.

Locution prépositionnelle La possibilité pour le marqueur d'apparaître dans une locution prépositionnelle signe la portée locale de la subordonnée qu'il introduit. Nous en relevons quatre exemples dans notre corpus pour les marqueurs qui nous intéressent : trois exemples en *because*, et un seul en *cos*.

Dans notre premier exemple, James explique quelle est la différence entre un accent rhotique et un accent non rhotique. Pendant ce temps, Celia s'essaie à prononcer quelques mots dont le « r » orthographique disparaît à l'oral chez les locuteurs du sud de l'Angleterre afin de vérifier le phénomène par elle-même. *Because of* permet à James de revenir à son propos initial qu'il pose comme complément circonstanciel conditionnant sa conclusion sur la plus grande variété vocalique des accents non rhotiques :

J *we just* {0,45} *don't ever pronounce them*
C *yeah here* {0,20} here
J an::d *that means* {0,85} <h>
C *where* {0,70} there
J so it k- **because *of** that*
C *mm*
J that means that we:: as Southern English speakers actually have more different vowels

(Celia&James : Recording, *because-of_001*)

Dans les deux autres exemples en *because*, la locution prépositionnelle suit la relation prédicative qu'elle commente. Ici, Rebecca avoue ne pas se sentir très à l'aise de parler de choses personnelles concernant ses amis tout en étant enregistrée. Elle ne souhaite pas dévoiler leur identité :

R I don't think I like **because of** names

(Helen&Rebecca : Weird families, *because-of_002*)

De la même façon, Mike identifie de façon restrictive la raison expliquant le contenu prédicatif exprimé en début d'énoncé :

M there'd been a a sense of humour failure brewing for several months <h> (0,5) **because of** all these clients I've had to be on my best behaviour with

(Mike&Hannah : Sense of humour failure, *because-of_003*

Reportons à présent le seul exemple de notre corpus où *cos* entre dans la composition d'une locution prépositionnelle, dans lequel James explique la spécificité du cursus de chinois par rapport aux autres langues vivantes :

J *er yeah* {0,20} and also it's mm {0,30} it's quite strange {0,60} for an undergraduate degree course in Oxford in {0,20} in:: tha::t you have so much classroom time especially in the first two years <h> {0,55} **cos of** people come and learn Chinese from scratch

(Celia&James : China, *cos-of_001*)

5.4.1.2 Un *because* plus restrictif ?

Au terme de ce répertoire des formes syntaxiques indiquant une relation restrictive entre principale et subordonnée de cause, il apparaît que les deux marqueurs peuvent entrer en jeu dans de telles formes. Cependant, il nous semble que *because* soit plus propice que *cos* à l'expression de la restriction et qu'il se prête donc plus facilement aux opérations syntaxiques détaillées plus haut.

La syntaxe de l'anglais oral spontané, et la dimension prosodique qui l'accompagne, permet difficilement de procéder à des manipulations sur les énoncés de notre corpus. Il est en outre délicat d'en tester l'acceptabilité dans la mesure où ces jugements normatifs sont largement conditionnés par la grammaire de l'écrit. La remarque de G. Deléchelle évoque encore une difficulté supplémentaire, touchant à la définition de la causalité :

> Les critères syntaxiques ne sont pas une preuve irréfutable parce qu'ils reposent sur des jugements d'acceptabilité qui impliquent aussi une certaine conception de la causalité.[28]

Faute de pouvoir procéder à des manipulations et de disposer de suffisamment de données convergentes, nous nous tournons donc vers des corpus plus importants, et notamment le BNC pour vérifier certaines tendances observées dans notre corpus. Le premier indice sur lequel se fonde notre intuition concerne la possibilité de transformer une subordonnée causale restrictive en un groupe prépositionnel introduit par une locution formée à partir du marqueur de causalité. Notre corpus contient un exemple de locution construite avec *cos*, contre trois avec *because*. Cette tendance à la prédominance de *because* dans les locutions prépositionnelle est largement confirmée par le nombre négligeable de *cos of* dans le BNC par rapport à *because of*.[29] Le contexte interrogatif permet encore de mettre en lumière un fonctionnement distinct des deux marqueurs : *because* peut répondre à lui tout seul à une question en *why*, sans même nécessiter l'expression du contenu de la subordonnée. Il ne semble pas que ce soit le cas pour *cos*.

5.4.2 Le cas des énoncés métalinguistiques

La subordonnée restrictive entretient un rapport direct avec la principale. L'autre type de rapport traditionnellement relevé pour les subordonnées causales est un rapport indirect, ou métalinguistique. Nous nous concentrerons ici essentiellement sur le rapport entre la pause silencieuse et le lien interpropositionnel dans les énoncés dits « métalinguistiques ».

[28]DELECHELLE, Gérard (1980). Because. *Etudes anglaises* 76 : 19-35. Paris : Dider Erudition. (p.33).

[29]Le BNC oral contient environ 10 fois moins d'occurrences de *cos of* que de *because of*.

5.4.2.1 Caractéristiques intonatives attendues

Nombre de chercheurs rendent compte du fait que la pause, parmi les paramètres intonatifs, est un indice du lien entre les propositions. Citons par exemple F. Grosjean & A. Deschamps, qui analysent le pourcentage de phrases suivies par une pause en fonction du contexte droit. Les phrases en question sont soit conclusives, c'est-à-dire suivies d'aucune proposition subordonnée ou coordonnée, soit suivies, selon les cas, d'une subordonnée (relative ou non) ou d'une coordonnée.

> [...] les pauses remplissent à des degrés variables les divers emplacements selon la liaison plus ou moins serrée de la phrase avec la proposition qui la suit.[30]

Ainsi, les subordonnées relatives, qui sont fortement intégrées dans la phrase, constituent la catégorie envisagée la moins souvent précédée d'une pause (24%) dans leur corpus. Au contraire, les propositions indépendantes, donc non suivies d'une subordonnée ou d'une coordonnée, sont à plus de 80% suivies d'une pause.

Au sein des subordonnées de cause, des types se dégagent en fonction notamment des caractéristiques intonatives des propositions. D'après certains auteurs, la pause en particulier (et son équivalent graphique, la virgule), par la distance physique qu'elle crée entre l'émission des deux propositions, est l'indice d'une distance conceptuelle existant entre elles. G. Deléchelle remarque que la présence d'une telle rupture entre matrice et subordonnée causale caractérise les énoncés où la proposition en *because/cos* a une valeur justificative et non explicative.

> La ponctuation de ces phrases, très variable, indique toujours une forte pause avant because q, qui est nettement détaché de p sans pour autant constituer une proposition indépendante.[31]

[30] GROSJEAN, François & DESCHAMPS, Alain (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 31 : 144-184. (p.168).

[31] DELECHELLE, Gérard (1989). L'expression de la cause en anglais contemporain. Etude de quelques connecteurs et opérations. Thèse de doctorat d'Etat. Université Paris III – Sorbonne Nouvelle. (p.23).

Ainsi, la proposition causale n'exprime pas une circonstance interne au procès permettant de valider la relation prédicative dans p. Elle constitue au contraire un commentaire métalinguistique portant sur la principale. La même observation est faite par W. Rutherford, qui distingue des énoncés *a* explicatifs et des énoncés *b* justificatifs. Ces derniers semblent davantage autoriser une intonation qu'il nomme *comma intonation* (ou virgule intonative), et dont la pause silencieuse constitue une composante importante :

> [...] comma intonation between main clause and subordinate clause generally seems more natural in *b* than in *a*, both in speech and in writing [...].[32]

D'après l'hypothèse performative proposée par W. Rutherford dans son article, ces énoncés sont en fait à considérer comme la subordonnée complétive d'une phrase dont le sujet est l'énonciateur, l'objet indirect le co-énonciateur et le verbe principal un verbe de dire. La construction de départ est donc du type : *I am saying p to you because q*, et se donne comme suit : *p because q* après élision de la matrice. Ainsi, dans les énoncés métalinguistiques en *because/cos*, la proposition causale ne se rapporte pas à la relation prédicative dans la principale apparente, mais à son énonciation, c'est-à-dire à la relation prédicative dans la proposition effacée. Dans l'exemple *cos_063*, la principale dont dépend la subordonnée causale est exprimée, faisant de la relation entre les deux propositions une relation restrictive. Mike explique qu'un matin, Ann et lui se sont réveillés sur un banc de sable au milieu d'une rivière, sous les yeux des passants qui empruntaient le chemin pour rejoindre un village voisin :

> **M** and in fact it was like a sort of little island almost cos it was a little river there we were <h> (0,4) sleeping on this little sand bank in the middle of this river which must have looked quite odd and **I say** it must have looked odd **cos** <h> (0,4) there was this stream of people going past on their mules because {0,4} that was the only way to get to one of the villages it's this tiny village
>
> (Mike&Hannah : Mules, *cos_063*)

[32]RUTHERFORD, William E. (1970). Some observations concerning subordinate clauses in English. *Language* 46 (1) : 97-115. (p.98).

La justification du dire anticipe une potentielle rupture du consensus entre deux interlocuteurs sur le plan des représentations mentales ou en tout cas cherche à établir un nouveau consensus en comblant un déficit d'information chez Hannah. L'énoncé entier fonctionne comme un commentaire visant à justifier l'énonciation de la proposition subordonnée relative précédente (*which must have looked odd*), dont le contenu lexical et modal est repris textuellement dans la complétive commandée par *I say*.

5.4.2.2 Caractéristiques intonatives observées

Peut-on conclure de nos données concernant la présence des pauses à la suture grammaticale représentée par le site interpropositionnel que notre corpus contient 12% de propositions en *because/cos* dont la portée est extérieure à la relation prédicative de la proposition principale à laquelle elles se rattachent ?[33] La proximité des proportions est frappante, puisque notre corpus contient 13% d'énoncés métalinguistiques. Cependant, il apparaît que les 16 propositions que nous analysons comme apportant un commentaire métalinguistique ne correspondent pas strictement aux énoncés de notre corpus comportant une pause silencieuse entre la principale et subordonnée. Si 8 d'entre eux contiennent au moins une pause silencieuse, celle-ci n'apparaît à la charnière des propositions que 6 fois, soit de façon minoritaire.

Six énoncés métalinguistiques ne comportent aucune pause, comme dans l'exemple *cos_003* où Celia justifie dans une subordonnée causale son exclamation préalable :

C *oh how frustrating* though
J *but er*
C ***cos** that wastes* one of your subjects

(Celia&James : Experiments, *cos_003*)

Contrairement à ce que notre mise en page suggère, les deux propositions de Celia ne sont pas séparées par une pause (la réplique de James se superposant au début

[33] Voir Tableau 5.2, p.243. Le chiffre de 12% est obtenu en additionnant les données des lignes 3, 7, 8 et 9.

de la subordonnée de Celia). La même solidarité apparente entre les propositions se retrouve dans l'exemple suivant *cos_066*. Mike raconte à Hannah que son séjour en Grèce avec Ann s'est bien passé : non seulement il a fait beau temps, mais en plus ils se sont très bien entendus. Il rappelle au passage que partir en vacances avec quelqu'un qu'on ne connaît pas bien relève un peu du pari. Ici encore, la subordonnée n'apporte pas une circonstance de la validation de la relation prédicative dans p mais commente l'énonciation même de cette proposition. En cela, elle lui confère le statut de remarque appropriée dans le contexte.

M and the weather was good and we were you know we got on so well **cos** you have no idea how you are gonna get on with someone in that sort of *situation*

(Mike&Hannah : Ann, *cos_066*)

Dans d'autres exemples, il existe bien une pause silencieuse, mais elle se situe au sein de la subordonnée, comme dans l'exemple *because_033* cité plus haut.[34]

Dans notre corpus, les propositions séparées par une pause silencieuse sont d'autant plus détachées l'une de l'autre que le silence est généralement l'occasion pour l'interlocuteur d'intervenir brièvement entre la matrice et la subordonnée. C'est le cas dans l'exemple *cos_020* dont nous reportons ici un contexte assez large. Au cours de la conversation, Celia doit rattacher son micro-cravate. James lui demande de s'assurer que le problème de micro n'a pas arrêté l'enregistrement. Après avoir reçu une réponse négative de Celia, il justifie son inquiétude et donc l'énonciation même de sa question en revenant à sa mésaventure de la veille. Il s'écoule plus de 3 secondes entre la fin de son interrogative et le début de sa proposition justificative en *cos*. Entre les deux propositions, Celia le rassure, rit, émet une marque d'hésitation et respire de façon audible, tandis que James rit et respire également, et pause pendant une demi-seconde.

C that looks as if that microphone's <h> {0,70} s::lightly *at a funny angle* <h>

J *w-ye* {0,50} probably then you can see your recorder there is there a

[34]Voir p.238.

C yes it is moving <rires>

J <rires>

C er <h> (0,55)

J <h> {0,50} **cos** when I forgot to turn the tape recorder on yesterday I'd still been adjusting recording levels all the time it's just the tape <rires> wasn't running

(Celia&James : Recording, *cos_020*)

On est fondé à se demander si l'on peut encore parler de subordination syntaxique si la matrice et la subordonnée sont séparées non seulement par des pauses silencieuses extrêmement longues mais surtout par des énoncés indépendants. Citons encore l'exemple *cos_039*, issu de la conversation entre Helen et Rebecca :

H *I think I kn-* I should have asked him if he knew some people I he probably hated them {} would have hated them

R I think they're pretty quite I don't know really

H no I'm just thinking like the people I know who went to Westminster

R yeah

H who I know from Oxford are really {} nasty

R okay

H wawa *public school*

R ***cos** I've met* some of his friends and they're really:: {}

H *nice*

R *laid-* yeah *quite nice people*

(Helen&Rebecca : Westminster, *cos_039*)

Dans cet exemple, les deux amies parlent de Franck, ancien élève de Westminster school. Helen se rend compte que Franck et elle ont sans doute des connaissances communes. Elle explique que celui-ci devait forcément détester les quelques personnes issues de cette école qu'elle a rencontrées à Oxford. Rebecca tente de lui démontrer que tous les anciens élèves de Westminster ne sont pas haïssables et que les amis de Franck sont plutôt sympathiques. Cette démonstration se fait en plusieurs étapes. Elle essaie tout d'abord d'avancer sa propre opinion sur les anciens de Westminster (*I think*) qu'elle veut plus nuancée que celle de Helen (avec l'utilisation des deux modifieurs *pretty* et *quite*). Ne trouvant pas de qualificatif adéquat, elle abandonne sa construction sur une note contestant l'opinion exprimée par son amie (*I don't know really*, la

principale que nous considérons). Helen en profite pour développer son point de vue, Rebecca se contentant alors du rôle d'énonciateur second qui produit des marques d'écoute (*yeah, okay*). Cette dernière reprend ensuite le fil de sa démonstration avec une proposition en *cos* en deux parties dans laquelle elle complète l'énoncé laissé en suspens quelques secondes auparavant et assortit le jugement qu'elle exprime d'une prise en charge qui prend la forme d'une garantie personnelle.

5.4.3 Vers une dissolution du lien causal ?

Notre relevé des exemples métalinguistiques et des exemples dont la syntaxe indique une relation restrictive entre p et q est loin d'épuiser notre corpus. Celui-ci contient bien entendu d'autres exemples à interprétation restrictive. Cependant, dans de nombreux cas, il est difficile de déterminer de quel type de relation il s'agit, si bien qu'il semble exister une relation intermédiaire entre les deux propositions, la subordonnée constituant une apposition plus ou moins directement liée à la validation de la relation prédicative dans p ou à ses conditions d'énonciation.

5.4.3.1 Traitement appositif de la subordination

E. Couper-Kuhlen distingue deux configurations dans lesquelles apparaissent *because* ou *cos* à l'oral (qu'elle considère la plupart du temps conjointement).[35] La configuration CSI se distingue de la configuration CSII, plus courante, en plusieurs points. La chercheuse montre que la première configuration se caractérise par une relation directe entre principale et subordonnée, proche de la relation restrictive que nous avons décrite plus haut. Les deux propositions forment une seule unité de prédication dont le noyau est la principale.

Dans la seconde configuration, les deux propositions sont présentées comme moins

[35]COUPER-KUHLEN, Elizabeth (1996). Intonation and Clause Combining in Discourse : the Case of Because. *Pragmatics* 6 (3) : 389-426.

dépendantes : elles constituent deux unités d'information susceptibles de former chacune un tour de parole. Sur le plan informationnel et pragmatique, la subordonnée peut elle aussi occasionner chez l'interlocuteur des marques d'écoute ou des réactions plus longues. Dans l'exemple suivant, Helen et Rebecca parlent de leurs sorties entre amis et se rendent compte que leur emploi du temps de la semaine est trop chargé. Principale et subordonnée y apparaissent comme deux unités d'information indirectement reliées par le marqueur *cos*.

H *yeah I'd like to meet them*
but just not tonight
R *but I thought it'd quite beneficial* if you came along and *had a chat*
H ***cos** we've* got people coming around tomorrow night *as well and*
R *yeah that's* true

(Helen&Rebecca : Messages, *cos_030*)

Les répliques de Rebecca indiquent le statut de relative indépendance des deux propositions. Sa première réplique commence avant la fin de la principale énoncée par Helen. Le ligateur *but* permet de l'interpréter comme une réponse anticipée à celle-ci, dans un contexte où Helen vient de dire qu'elle se sentait trop fatiguée pour sortir le soir même. Ensuite, *yeah that's true* ne porte pas un jugement sur le contenu de la principale mais commente la subordonnée.

Sur le plan prosodique, une subordonnée causale relevant de la catégorie CSII sera typiquement marquée par une réinitialisation partielle de la ligne de déclinaison. Ainsi, dans l'exemple que nous venons de citer, la première syllabe portant un accent lexical dans la subordonnée (*got*) est prononcée à une hauteur légèrement supérieure à celle observée sur la dernière syllabe accentuée de la principale (*night*).

Cette tendance vers l'autonomisation de la subordonnée tend à remettre en question son statut de constituant de l'énoncé complexe. Par là se trouve interrogée la notion même de hiérarchie syntaxique et donc de subordination.

5.4.3.2 Diversification du rôle de *because/cos*

L'anglais oral spontané voit donc le rôle des marqueurs *because* et *cos* se diversifier. Ils peuvent être utilisés aussi bien de façon hypotaxique, dans le cadre de la subordination, que de façon parataxique. M. Schleppegrell décrit trois rôles susceptibles d'être remplis par un marqueur non-subordonnant :

> [It can be used] as a discourse-reflexive textual link which introduces a reason for or explanation of a prior statement [...] as an expressive, non-causal link which introduces elaboration of a prior proposition [...] [or] as a discourse marker which indicates continuation and response in conversational interaction.[36]

Dans le premier de ces emplois (qui correspond à ce que nous avons appelé méta-linguistique), le marqueur conserve un rôle causal dans la mesure où il introduit un élément justifiant l'énonciation précédente. Les deux autres, en revanche, participent pour nous de cette dissolution du lien causal spécifique à l'oral.

Marqueur d'élaboration Dans l'exemple suivant, Mike explique qu'il a entendu la voix d'une amie de sa mère à la radio alors qu'il était en Grèce :

M I was in {0,3} Greece I was doing one of these treks and <h> (0,35) {0,3} I we got to the hotel and I sat down having a beer <h> (0,35) and this voice boomed out and I hadn't actually realised that the radio was on <h> (0,25) b- I think they'd just switched it on at seven o'clock or whatever <h> (0,4) {0,25} and the v- the news came on in English and it was this friend of my mother's reading it *Pat who*

H *you're joking*

M I knew I mean I knew that she read the news for the Greek television *occasionally*

H *and there she was*

M **cos** she does it {0,25} in English you know and does the translations into English and then reads it {0,6} and suddenly I heard this voice I thought bloody hell that's Pat I mean gracious *me*

H *Pat*

M <rires> gracious me this is Pat

(Mike&Hannah : Pat, *cos_043*)

[36]SCHLEPPEGRELL, Mary (1991). Paratactic *because*. *Journal of Pragmatics* 16 : 31-58. (p.323).

La relation causale entre subordonnée et principale est assez floue. Mike mentionne précédemment que le journal parlé qui commençait était donné en langue anglaise. Cependant, le fait que Pat lise les nouvelles en anglais n'explique pas qu'elle présente parfois le journal à la radio grecque. Il ne permet pas non plus de justifier l'énonciation précédente. *Cos* a ici un rôle relevant de la deuxième catégorie décrite par M. Schleppegrell. Il s'agit pour Mike de donner des détails supplémentaires, bâtir un contexte ou un décor plus consistant, afin de renforcer l'effet de son anecdote sur son interlocutrice. Celui-ci est d'ailleurs très réussi si l'on considère la force de l'exclamation qu'il provoque : la dernière réplique de Hannah, *Pat*, atteint un niveau intensif très élevé et sa hauteur mélodique dépasse les 500 Hz.

Nombre d'exemples de notre corpus prennent la forme d'un long développement à partir d'une appréciation formulée dans la principale. Dans la plupart des cas, c'est le contenu de la subordonnée qui devient objet de discours, et non la caractérisation du sentiment. *Cos_027*, cité ci-dessous, fournit un bon exemple de ce phénomène.

Marqueur de ligation L'exemple *cos_027* illustre le troisième rôle non causal relevé par M. Schleppegrell, celui du ligateur.

H and she thinks it's funny and then she says I'm only joking {}
but it's actually
R not funny
H not that funny **cos**
and today she w- went off on one about {} how there were two many English in Fran::ce and how hm

(Helen&Rebecca : English, *cos_027*)

Helen raconte qu'elle est victime d'incessantes moqueries anglophobes proférées par l'une de ses collègues françaises. Elle exprime ici sa lassitude et entame un énoncé expliquant son sentiment. Celui-ci fait écho à la réplique compatissante de Rebecca et n'aboutit pas, Helen estimant sans doute qu'il n'est pas la peine de donner les circonstances permettant de valider le contenu de sa proposition principale tant celles-ci font l'objet d'un consensus entre les deux locutrices. *Cos* apparaît donc seul, en

fin de principale. Sa valeur causale lui est retirée par l'absence même d'un contenu propositionnel subordonné. Son fonctionnement se rapproche de celui d'un ligateur : il permet ici à Helen de conserver la parole. En outre, du fait de son origine causale, il permet d'annoncer que le contenu de ses propos à venir est pertinent vis-à-vis de ce qu'elle vient de dire. L'énoncé suivant, introduit par le ligateur *and*, constitue en effet un exemple du harcèlement subi par Helen.

Le rôle de ligateur est ici assumé par le marqueur *because*, dans un exemple où James raconte qu'un de ses sujets d'expérimentation a transformé le test prévu en changeant les séquences de sons à prononcer en cours d'expérience. Celia propose une explication :

C *I think* it's easier actually if you're putting the vowel at the beginning
J oh it is it is an easier *thing but*
C ***because** it's* m- much easier to say
J mm but the the experiment I've based this on was based in in fixes which were er vowel consonant
C mm

(Celia&James : Diphthongs2, *because_018*)

L'apport informationnel de la subordonnée est très faible. La proposition, détachée de la principale, ne fait que réaffirmer le contenu de celle-ci de manière tautologique (c'est plus facile à dire parce que c'est plus facile à dire). En tête d'une proposition qui tend à l'autonomie, le marqueur permet donc ici à Celia non plus de conserver la parole mais de s'arroger un tour de parole supplémentaire tout en affirmant la pertinence de son propos par rapport à ce qui précède.

Notons que la réinitialisation partielle de la ligne de déclinaison, typique de la catégorie CSII de E. Couper-Kuhlen, corrobore la thèse d'un marqueur ligateur. Une réinitialisation, même partielle, signale en effet le plus souvent un début de paragraphe oral. Le ligateur constitue en outre la pièce maîtresse du préambule.

5.4.3.3 Spécialisation des marqueurs et bilan

Parallèlement à la diversification du rôle des marqueurs, nous constatons une tendance à la spécialisation. C'est ainsi que nous expliquons la préférence de l'un des deux marqueurs par rapport à l'autre dans des contextes précis, relatifs notamment à la relation intersubjective et à la relation argumentative en jeu. Les différences significatives relevées plus haut dans la distribution des deux marqueurs sont autant d'indices de cette spécialisation.

Relation intersubjective Comme nous l'avons vu, la distribution des sujets grammaticaux de la principale tend à indiquer une répartition de l'espace intersubjectif entre les deux marqueurs. Les sujets première et deuxième personnes sont plus volontiers suivis de *cos* que de *because*. *Cos* est donc préféré dans des contextes où, avec la deuxième personne, est anticipé un consensus. En disant *you*, l'énonciateur place ce qu'il énonce sous le contrôle du co-énonciateur sans prévoir de décalage avec le contenu de pensée qu'il lui prête. Quand il dit *we*, ou à plus forte raison *I*, l'énonciateur peut seul se porter garant de ce qu'il énonce. D'un point de vue intersubjectif, le contenu de ce qu'il énonce peut déjà faire partie des connaissances qu'il partage avec le co-énonciateur ou n'appartenir qu'à la sphère privée de l'énonciateur. Dans ce dernier cas, on est fondé à parler d'égocentrage lorsque l'anticipation des représentations du co-énonciateur n'est pas envisagée comme fondatrices du propos tenu.

L'exemple *cos_052* illustre cette préférence pour *cos* dans les cas où l'énonciateur est le siège de sensations (*it just somehow felt as if...*) fondatrices de certitudes (*sure*) ou d'hypothèses (*even though, though, probably, as if, guess*). La subordonnée vient étayer une appréciation personnelle dont Mike est le seul garant et au sujet de laquelle il est fondé à ne concevoir aucune contestation de la part de son interlocutrice :

M **I was sure** my glasses had come off <h> (0,4) and I was trying to scrab scrabble around in the water for my glasses {0,35} **even though** I had this {0,25} thing you know {0,8} whatever

H the band thing

M a band to secure them on my head <h> (0,7) hm {0,6} but **I thought** it had **probably** not worked **cos it just somehow felt as if** they'd come off and **I guess** they'd just come like slightly off and so I couldn't feel the pressure of *them on my head {0,5}*

(Mike&Hannah : Little near-death experience,*cos_052*)

La préférence de *cos* pour l'expression d'une relation causale relative à la sphère du proche indique une propension du marqueur à être utilisé lorsque la relation entre p et q est congruente. Nombre d'exemples attestent de l'évidence du rapport entre les deux propositions, éliminant toute discussion possible sur le bien-fondé de cette relation. Dans l'exemple *cos_008*, James raconte que la plupart des étudiants ont oublié le code d'accès à la salle informatique du centre de linguistique et philologie, d'où un accès difficile.

J if you use it regularly <rires> I suppose you keep track a lot more if you just <rires> {0,25}
C *hm*
J *stop* using it and then you
C and you *find you can't get in* <rires>
J *never start again **cos** you can't <rires> get in* <rires>
C *it's ridiculous*

(Celia&James : Summer, *cos_008*)

Avec *cos*, cause et effet sont présentés comme naturellement liés : à force de ne pas utiliser la salle d'informatique, on se retrouve dans l'impossibilité de l'utiliser. Le caractère irrémédiable de la relation entre p et q confine à l'absurde, comme en témoigne l'exclamation de Celia dans sa dernière réplique.

Because est préféré lorsque l'énonciateur anticipe des représentations divergentes et cherche à exprimer des relations moins congruentes. Les énoncés en *because* sont davantage marqués par d'importantes modulations mélodiques sur les éléments cristallisant la divergence anticipée ou permettant de la dissiper. Dans l'exemple suivant, la proposition causale de James vient contredire les propos de Celia selon qui on peut très bien se rendre compte qu'un magnétophone n'est pas en train d'enregistrer parce rien ne bouge.

J cos when I forgot to turn the tape recorder on yesterday I'd still been adjusting recording levels all the time it's just the tape <rires> wasn't running

C *but you can see there's nothing to show*

J *<rires> yeah well you can still* yeah w- y- well no **because** there **was** something to show that's the thing
you still have the display of the recording level as the people spoke <h> {0,40} just that the tape wasn't <rires> going *round and round <rires>*

C *oh:: God*

(Celia&James : Recording, *because_014*)

La subordonnée introduite par *because* nie emphatiquement la validité du lien prédicatif de l'énoncé de Celia qu'il reprend mot à mot en soulignant mélodiquement la copule.

Contrairement à un énoncé en *cos* qui tend à présupposer un consensus ou à l'ignorer, un énoncé en *because* s'interprète davantage comme une demande d'assentiment ou comme un moyen d'accéder à un consensus. En ce sens, la force argumentative des deux marqueurs nous semble différente.

Force argumentative Notre dernier exemple, *because_014*, illustre la force argumentative du marqueur. L'énoncé de James est conçu comme le prélude d'une démonstration qui trouve ses arguments dans la suite de la réplique. L'exclamation de Celia témoigne du fait qu'un nouveau consensus est atteint, James ayant réussi à convaincre son interlocutrice. Nous relions au caractère plus fondamentalement argumentatif de *because* une caractéristique distributionnelle qui le distingue de *cos* : l'abondance des marques du travail de formulation caractérisant les énoncés en *because*. L'énonciateur qui veut convaincre a naturellement tendance à utiliser des phatiques, et à chercher le terme le plus apte à ébranler le contenu de pensée du co-énonciateur, quitte à devoir ménager des pauses pleines ou silencieuses plus ou moins longues. Ce phénomène témoigne selon nous d'une plus grande complexité des relations interpropositionnelles exprimées avec *because*. La pause silencieuse située immédiatement après le marqueur nous semble emblématique de cette tendance, d'une part, à conserver l'inscription de

la subordonnée comme constituant de la principale, et, d'autre part, en posant le lien avant d'en exprimer le contenu, à permettre la construction d'énoncés plus élaborés.

Cette tendance à utiliser *because* plutôt que *cos* dans des constructions à visée argumentative est confirmée par plusieurs faits de distribution. Considérons tout d'abord la distribution des relateurs *so* et *but* dans le contexte immédiat des deux marqueurs. Il existe en effet dans notre corpus une différence significative entre les deux marqueurs quant à la distribution de *but* et *so* dans le contexte droit de la subordonnée ($\chi^2 =$ 6,073 pour 2 ddl, p $\leqslant$ 0,05).[37] Il ressort de l'examen de notre corpus que *cos* a moins d'affinités que *because* avec le relateur argumentatif *so*, tandis qu'il s'accompagne plus volontiers de *but*, dont la force argumentative est bien moindre. En outre, une requête concernant l'occurrence de chacun des marqueurs après un *why* interrogatif montre que 7% des *because* du BNC sont directement liés à *why*, contre seulement 2% pour *cos*. Cette proximité plus naturelle de *why* et *because* atteste encore de la place privilégiée de *because* dans un tel contexte.

C'est ainsi que *because* tend à être préféré lorsque l'énonciateur cherche à renforcer le lien argumentatif entre la subordonnée et la principale. C'est le cas dans les exemples où les deux propositions ne sont pas émises par le même locuteur : utiliser *because* permet de resserrer le lien entre les propositions malgré le changement de locuteur. Dans l'exemple suivant que Celia reprend le *cos* de James avec *because* et assure ainsi la permanence du lien causal en le reconstruisant. Elle apporte en outre des arguments permettant d'expliquer pourquoi les étudiants de chinois disposent de locaux plus grands que les linguistes dont James fait partie :

J but I'm sure we use it more
C yeah
J **cos** {0,25} and I guess
C ***because** the library's there*
J *they have graduates as well* {1,25} *er*

[37] Les effectifs pris en compte pour le test du χ^2 sont consignés dans le Tableau C.5, en annexe au présent chapitre (p.449).

C *if* they've got a good library though they're *bound to aren't they*

(Celia&James : China, *because_007, cos_014*)

Comme nous avons pu le voir plus haut avec l'exemple *because_018*, *because* permet par extension de donner une force argumentative à des éléments qui par eux-mêmes en manquent.[38]

Il nous semble que la force argumentative de *cos* provient par endroits du caractère congruent (ou présenté comme tel) de la relation exprimée entre les deux propositions. Nous pensons au trait d'humour qui se plaît souvent à usurper l'apparence d'une logique argumentative. Cependant, à un *because* explicatif, l'humour préfère un *cos* qui pose les validations comme acquises et élimine toute discussion possible. C'est le cas du trait sans appel de Matthew qui suggère, en s'adressant à nous, que les Français ne peuvent pas comprendre qu'on puisse devenir végétarien :

M I mean I g- I gave up eating red meat when I was about {0,4} seventeen or something and then <h> (0,5) gradually decided to just {0,15} see if I could do with other things (?) it's more of an obsession than a {0,9} real justification {0,3} but I'm happy now I don't (?) so {2,6} you can't understand **cos** you're French I'm sorry <h> (0,6) <rires>

(Nick&Matthew1 : Vegetarian, *cos_074*)

Ainsi, le phénomène évoqué plus haut, qu'E. Couper-Kuhlen appelle *semantic bleaching* ou décoloration sémantique, touche selon nous davantage *cos* que *because*. Un emploi typique de *cos* consiste en effet à servir de point de départ d'une description, et non comme élément d'une argumentation. C'est le cas dans l'exemple suivant. Celia et James parlent de tourisme, en particulier en Asie. James doute qu'il existe de grandes différences entre les pays dont Celia compare les attraits. Les répliques de Celia ne constituent pas une réponse argumentée à la question de James. La locutrice ne convoque pas des raisons politiques, culturelles, religieuses, linguistiques, climatiques ou autres pour expliquer son point de vue, mais entame le récit d'une expérience personnelle peu susceptible d'étayer son propos. Ce récit ne nous semble

[38] Voir p.270.

pas non plus en mesure d'apporter un commentaire métalinguistique à l'énonciation de *I'm told it is*.

J *b- d'you think it's quite* different

C <h> I'm told it is::

J *yeah*

C ***cos*** I met my daughter there <h> {0,55}
and she wasn't away for a gap year she's older {0,65} she's gonna be your age {1,00}
and she'd gone for:: {0,50} hm <h> {0,80}
well she wasn't sure whether it was three months or a year {1,55}
and rang me from India

(Celia&James : Goa1, *cos_016*)

Cos introduit un long développement sur le séjour de la fille de Celia en Asie. Celui-ci ne participe pas d'un projet argumentatif de Celia qui chercherait à expliquer son point de vue ou justifier son dire par l'exemple. La problématique de la différence entre les pays d'Asie semble être abandonnée : elle n'est évoquée, très rapidement, que plusieurs minutes plus tard. Celia profite au contraire de son tour de parole pour partager son expérience personnelle de l'Asie, à travers le séjour de sa fille mais également à travers le voyage qu'elle a pu faire en lui rendant visite. Le long passage qui fait suite à notre exemple est entièrement consacré à l'évocation de ce voyage. *Cos* permet à Celia de maintenir une certaine cohérence dans ses propos tout en s'éloignant considérablement du propos initial. Le marqueur combine ici tous les emplois à faible coloration causale décrits par E. Couper-Kuhlen.

Si elle ne se lit pas dans tous les exemples du corpus, cette différence de force argumentative entre les deux marqueurs apporte un élément supplémentaire en faveur de notre intuition selon laquelle *because* reste naturellement plus propice à l'expression des relations causales directes. Avec la division de l'espace intersubjectif, elle caractérise pour nous la différence de valeur existant entre *because* et *cos*, dont témoigne encore notre dernier exemple :

M but {0,3} in fact it was good **because** I didn't know that much about her really <h> (0,35) and nor she about me **so** we started asking each other questions about <**h**> **(0,3) you know**

H <bâillement>

M families and whatever jobs and I had no idea what sort of thing she did {0,7} **so** it was nice **cos** we didn't have to sort of {1,3} I don't know yeah we had *plenty to talk about yeah*

(Mike&Hannah : Ann, *because_055, cos_068*)

Mike explique qu'il ne connaissait pas vraiment Ann avant de partir en vacances en Grèce avec elle, ce qui leur a permis de trouver facilement des sujets de conversation. En quoi les deux énoncés apparemment si proches, *it was good because...* et *it was nice cos...*, diffèrent-ils ? La première appréciation de Mike est suivie d'une véritable démonstration. Débordant les limites de la subordonnée causale (*because I didn't... about me*), celle-ci se décline en plusieurs propositions évoquant le point de vue de l'un et l'autre. Elle s'appuie sur des exemples concrets et voit apparaître le relateur *so* par deux fois. Les marques du travail de formulation attestent de la relative complexité des propos avancés : la longue pause respiratoire permet à Mike de trouver ses mots, de même que *you know*, qui en appelle aux connaissances partagées entre les parties prenantes du dialogue et tente de capter l'attention faiblissante de son interlocutrice sur le point de bâiller. La prise en compte d'un contenu de pensée potentiellement divergent apparaît donc dans cet exemple dans lequel Mike déploie son argumentation de façon à bâtir un nouveau consensus. C'est sur ce consensus établi que fonctionnent *cos* et la proposition qu'il introduit. Mike cherche ici encore ses mots, certes, mais le contenu qu'il énonce finalement découle si naturellement de ce qu'il a dit précédemment qu'il ponctue son propos de deux *yeah* témoignant de l'évidence du consensus.

Notre étude de la subordination causale dans notre corpus fait apparaître une double tendance. La première est une tendance vers l'appauvrissement ou vers la spécialisation de l'expression de la cause. Celle-ci a en effet pour moyens quasiment exclusifs les marqueurs *because* ou *cos*. Cette tendance se confirme encore, à certains égards, lorsque l'on cherche à caractériser le fonctionnement de la subordination en

because ou *cos*. Les rôles des marqueurs révèlent néanmoins une tendance à la diversification : le marqueur de subordination tend à perdre sa coloration strictement causale pour se teinter d'autres valeurs et occuper des rôles différents plus proches de la ligation. Au sein de ce paysage uniforme et pourtant contrasté, les deux marqueurs nous semblent acquérir des valeurs spécifiques, tant se dégagent de nos exemples des emplois privilégiés de l'un et de l'autre. Certains faits de distribution distinguent les deux marqueurs de façon significative et contribuent à leur conférer une valeur intersubjective et une force argumentative différentes.

Chapitre 6

Subordination et organisation discursive

Notre étude de la subordination causale en *because/cos* tend à mettre en question la pertinence de la notion de subordination syntaxique à l'oral spontané. Ce chapitre a pour ambition d'évaluer l'ampleur de ce phénomène et de montrer dans quelle mesure le modèle de hiérarchisation syntaxique, en assouplissant ses principes, étend son influence à d'autres domaines de l'organisation du discours. Se pose donc ici la question des unités : la proposition en *because/cos* entre-t-elle en combinaison syntaxique avec une autre proposition pour former un énoncé complexe ? La combinaison ne se joue-t-elle pas davantage au niveau sémantique au sein du paragraphe oral ?

6.1 Hypotaxe ou parataxe ?

Un certain nombre d'exemples traités dans le chapitre précédent témoignent d'une relation très lâche entre la proposition subordonnée causale et sa principale. Dans l'exemple *because_019* traité plus haut,[1] la distance séparant l'énonciation des deux

[1] Voir p.253.

propositions contribue à la dissolution du lien. Dans l'exemple *cos_043*,[2] nous attribuons cet effet au faible rapport causal unissant le contenu des deux propositions. Une telle liberté de la proposition dite subordonnée par rapport à sa principale pose la question de l'indépendance des deux propositions, tant sur le plan syntaxique que sur le plan sémantique. Certains auteurs voient encore ici un fonctionnement hypotaxique, là où d'autres préfèrent parler de parataxe. Où nous situons-nous dans ce débat ?

6.1.1 Un fonctionnement hypotaxique

E. Couper-Kuhlen[3] se pose la question de la pertinence de l'analyse des propositions en *because* comme des propositions hiérarchiquement subordonnées à une matrice. Dans la catégorie CSII qu'elle définit, elle constate en effet une forte tendance à l'autonomisation de la proposition causale par rapport à la principale à laquelle elle se rapporte. Cependant, ce phénomène ne suffit pas à ses yeux à disqualifier une analyse hypotaxique des deux propositions liées.

> CSII qualifies as a construction despite its augmented or two-stage character because the clause or reason clause is configured lexically and prosodically as being in construction with a prior clause: It has a lexical marker of cohesion and its declination is partially reset with respect to this prior clause. Were the two TCUs not in construction with one another, we would expect no lexical marker and full resetting.[4]

La présence du marqueur et la réinitialisation partielle (et non totale) de la ligne de déclinaison attestent donc selon elle du fait que les deux propositions sont bien dépendantes d'un point de vue structurel.

[2] Voir p.268.
[3] COUPER-KUHLEN, Elizabeth (1996). Intonation and Clause Combining in Discourse : the Case of Because. *Pragmatics* 6 (3) : 389-426.
[4] *ibid.* (p.410).

Deux objections peuvent être formulées à l'encontre de ce raisonnement. La première se fonde sur la faculté d'un même marqueur à instancier des places syntaxiques différentes. Nous pensons ici à l'opposition entre opérateur et verbe lexical caractérisant par exemple *have* et *do*. Nous avons également vu que *because* peut apparaître aussi bien seul sous la forme d'une conjonction qu'en combinaison au sein d'une locution prépositionnelle. Puisque le fonctionnement du marqueur concerné est propre à l'oral, rappelons également qu'un marqueur comme *and* se retrouve dans la position d'un coordonnant intra-énoncé aussi bien que dans celle d'un ligateur inter-énoncé. La présence du marqueur ne nous semble donc pas suffire pour affirmer, à l'oral spontané, la permanence de la relation hypotaxique que celui-ci assure à l'écrit.

Le fonctionnement de la ligne de déclinaison nous semble justifier une seconde objection. Une réinitialisation même partielle est une marque de segmentation. Certes, elle atteste d'une certaine pertinence d'un segment par rapport à celui qui précède, et peut ainsi servir de suture à deux propositions d'un énoncé complexe. Néanmoins, ce même type de segmentation peut être utilisé à un niveau de structuration supérieur à celui qu'on attend entre deux propositions liées. Si la réinitialisation totale atteste d'une frontière plus forte, une réinitialisation partielle de la ligne de déclinaison constitue une délimitation suffisante entre deux paragraphes oraux.

6.1.2 Un fonctionnement parataxique

Si l'hypothèse hypotaxique permet d'expliquer un certain nombre d'exemples, l'hypothèse d'un fonctionnement parataxique des propositions en *because/cos* est également recevable. Pour M. Schleppegrell,[5] la relation entretenue par les propositions que la grammaire traditionnelle appelle « principale » et « subordonnée en *because* » est en effet plus ambiguë, du fait de la nature même des segments liés. Elle relève davantage d'un phénomène de « *clause combining* » parataxique que de véritable subordination.

[5]SCHLEPPEGRELL, Mary (1991). Paratactic *because*. *Journal of Pragmatics* 16 : 323-337.

C'est le sens de la remarque que nous reportons ici :

> While an embedding relationship is clear with relative and complement clauses, which are constituents of NP or VP, the situation with adverbial clauses is different. Adverbial clauses can stand in a clause combining relationship rather than an embedding relationship with other clauses [...]. Because clauses [...] can be constituents of sentences as a whole, or even be paratactic clauses (not constituents at all).[6]

La relation unissant les deux propositions est plus proche de la coordination que de la subordination dans la mesure où la hiérarchisation syntaxique tend à disparaître au profit d'une structure horizontale dans laquelle *because/cos* conserve un rôle de relateur. Comme nous avons pu le voir au chapitre précédent, ce rôle s'apparente alors à celui d'un ligateur, dont la caractéristique est de relier non pas des propositions au sein d'un énoncé complexe mais des segments plus larges.

6.1.2.1 *Because/cos* + pause + proposition

La lecture parataxique de la relation entre p et *because/cos* q en anglais oral spontané contemporain trouve également un écho certain dans l'origine de la subordination de cause :

> [...] the strongly deictic character of the causal marker, and the optional absence of the particle þ*e*, suggest that in OE causal constructions were not as distinctly subordinate as in PDE [...]. Indeed, the original construction in PrOE [primitive OE] was probably a paratactic one. [...] the *for + DEM* in the 'because'-clause was either anaphoric [...] or cataphoric [...]. By the earliest OE period, constructions of this type co-existed with others in which the 'because'-clause was clearly subordinate because it was marked with the subordinator þ*e*, but we

[6]SCHLEPPEGRELL, Mary (1991). *art. cit.* (p.324).

> cannot tell whether *for + DEM* in the sense of 'because' had been reanalysed as a subordinator.[7]

Deux types de constructions causales co-existent donc en vieil anglais, l'une parataxique et l'autre, hypotaxique, qui l'a finalement supplantée à travers les siècles.

La tendance à la fois parataxique et déictique de l'expression de la cause en vieil anglais nous semble éclairer le fonctionnement de la pause dans bon nombre de propositions en *because/cos* de notre corpus. La configuration suivante est en effet très fréquente :

P + *because/cos* + {pause} + Q

glose P pour la raison suivante : Q

La glose que nous proposons met en relief les caractéristiques cataphoriques du marqueur causal : *because/cos* pose le lien pour ne l'expliciter que plus loin. Elle souligne également la relation parataxique entretenue par p et q, et explique par là la présence récurrente de la pause entre le marqueur et le contenu de la proposition qu'il introduit.

6.1.2.2 *For* et *because/cos*

Le fonctionnement syntaxique et sémantique d'une conjonction de coordination telle que *for* nous semble très proche de *because/cos*, et fait ainsi signe vers leur origine commune. *For* est très rarement utilisé à l'oral : notre corpus n'en présente qu'un seul exemple. Celui-ci rappelle très fortement les énoncés collaboratifs en *because/cos* dans lesquels les propositions liées sont prononcées par deux locuteurs différents, ou du moins, comme dans notre exemple, où la proposition causale se rapporte à une proposition constituée d'une seule marque d'acquiescement ou de dénégation :[8]

[7] HOGG, Richard, BLAKE, Norman, LASS, Roger, & ROMAINE, Suzanne (éds.) (1992-1999). *The Cambridge History of the English Language.* Cambridge : Cambridge University Press. (p.253).

[8] Voir p.233, et en particulier l'exemple *because_020.*

J *I think er Rosemary Cross* is er {0,45} supposed to be able to tell you but she's never there hm {0,65} in the afternoon so

C no **for** she only works <bruit> mornings so

J yeah {0,80}
<rires> oh that's why she's never *there in the afternoon* <rires>

C *that's why she's never* <rires> there in the afternoon *<rires>*

J *<rires> yeah*

C it's quite <h> (0,50) quite a simple explanation

(Celia&James : Summer, *for_001*)

L'énoncé de Celia accepte de voir *for* remplacé par *because* ou *cos*, marqueurs couramment utilisés dans une telle structure :

yes/no + *because/cos/for* + proposition

Dans le contexte de notre exemple, le fonctionnement intersubjectif de *for* est très proche de celui que nous avons attribué à *cos*. Cette proximité des deux marqueurs est également remarquée par M.A.K. Halliday & R. Hasan dans leur ouvrage de 1976, qui suggèrent que *for* disparaît à l'oral au profit de *cos* :

> [*For*] is hardly ever heard in spoken English, where its nearest equivalent is the word *because* in phonologically reduced form.[9]

Nos données ne permettent pas d'émettre de conclusion quant au statut relatif des marqueurs *because/cos* et *for*. D'un côté, elles suggèrent que les marqueurs se situent ici totalement sur un axe paradigmatique, la même structure acceptant légitimement des marqueurs différents. De l'autre côté, il se peut simplement que l'un soit comme comme un coucou dans le nid de l'autre — l'un des marqueurs se substituant à l'autre, plus légitime, dans une même structure, le temps et le nombre de cas similaires finissant par légitimer sa présence. Dans ce dernier cas, indépendamment de la plus grande légitimité de l'un ou l'autre des marqueurs, c'est semble-t-il le marqueur le

[9]HALLIDAY, Michael A.K. & HASAN, Ruqaiya (1976). *Cohesion in English.* Londres : Longman. (p.258).

moins usité généralement à l'oral qui est le plus susceptible de faire place à l'autre, plus courant.

Il nous est difficile de déterminer si une telle structure collaborative est véritablement hypotaxique, comme le suggère (dans une interprétation traditionnelle) la plus forte représentation de *because/cos*, ou parataxique, comme le suggère dans notre exemple la possibilité de voir apparaître *for*. Se trouve cependant illustrée la perméabilité des catégories syntaxiques des marqueurs, à travers la permanence des formes. Inversement, nous constatons l'utilisation des mêmes marqueurs au niveau interpropositionnel au sein d'une structure restrictive (de type CSI comme défini par E. Couper-Kuhlen), et au niveau inter-énoncés au sein d'une structure aux attaches plus lâches (de type CSII).[10]

6.2 Pérennité d'une hiérarchisation discursive des constituants liés

Si l'anglais oral voit à certains égards se dissoudre les relations syntaxiques hiérarchiques au profit de relations moins serrées, la hiérarchisation des segments sur le plan discursif ne nous semble pas remise en cause.

6.2.1 Une hiérarchisation assurée par les marqueurs ?

6.2.1.1 Arrière plan et premier plan

Cette hiérarchisation provient en partie de la dimension sémantique des marqueurs et de leur capacité à organiser le linéaire du discours. D. Schiffrin remarque la récurrence d'un schéma opposant *because* et *so* non pas d'un point de vue causal mais du point de vue du traitement de l'information apportée par les segments qu'ils

[10]COUPER-KUHLEN, Elizabeth (1996). *art. cit.* Voir plus haut, p.280.

introduisent :

> [...] the markers *because, and*, and *so* differentiate ideational segments of the discourse, with the subordinate conjunction *because* marking a subordinate part of the discourse (support) and the coordinators *and* and *so* marking a more dominant part of the discourse (the position).[11]

Ainsi, selon D. Schiffrin, les marqueurs fonctionnent comme des balises signalant la lecture à donner à chaque segment par rapport au propos développé. Ils permettent de mettre en perspective la ligne directrice du discours et de mesurer la contribution apportée par les segments propositionnels qui se succèdent. Dans une argumentation ou une narration, le locuteur émet des jugements ou présente des faits (la *position* de D. Schiffrin) et convoque des éléments divers (*support*), susceptibles de conforter sa position auprès de l'interlocuteur dont il cherche à emporter l'adhésion. Ces deux moments du discours sont soulignés par des marqueurs différents. Située au premier plan, la ligne directrice contenant la position du locuteur tend à être marquée par *and* ou *so*, tandis que les segments satellites, en arrière plan, sont introduits par *because/cos*.

Ce procédé discursif contribue au phénomène de dissolution du lien causal relevé dans notre corpus : ce n'est plus forcément une proposition causale que *because* ou *cos* introduisent, mais une proposition permettant d'étayer la ligne principale, et dont l'apport informationnel est traité comme secondaire par le locuteur. Il se retrouve dans l'exemple *because_044* : Mike raconte comment il s'est fait peur en évitant de peu un rocher, obstacle vers lequel la rivière qu'il parcourait à la nage le précipitait tandis qu'il se préoccupait de ses lunettes.

M 1. **so** I was going oh my shit my glasses my glasses and I wasn't actually thinking about <h> (0,4) {0,3} the river which actually wasn't that difficult <h> (0,35)
2. **and then** I suddenly looked up

[11] SCHIFFRIN, Deborah (1987). *Discourse Markers.* Cambridge : Cambridge University Press. (p.51).

3. **because** I at this time I was travelling quite fast you know <h> (0,3)

4. **and** I looked up

and I realised this was coming

and I was like uuuaaah

5. **and then** I realised I'd had my glasses in it (?)

6. **but** I had to n- avoid this rock that was now sort of you know this bit of cliff that was right in front of me

7. **and** I just got out of the way <h> (0,5)

8. **and** I'd said to Ann look I'll give this a shot first <h> (0,4) and and when I get to the far side I'll give you a sign either sort of no:: or yes {0,4}

9. **so** I got out and <gestes>

H *<rires>*

M *I don't think you want to try that* she went okay

(Mike&Hannah : Little near-death experience, *because_044*)

Mike ne construit pas ici une argumentation mais un récit, qui utilise les marqueurs de l'argumentation tels que décrits par D. Schiffrin. Ceux-ci favorisent le dessin de la ligne principale du récit qui se déroule en s'appuyant sur des éléments périphériques. En effet, les segments relevant de la ligne directrice du récit sont signalés par *and* ou *so* (1, 2, 4, 5, 6, 7 et 9), tandis que *because* introduit un segment situé en arrière plan. En forme de commentaire, la proposition en *because* (3) s'assure en fait de la pertinence des faits précédemment relatés en les plaçant dans un contexte où ils prennent sens. C'est parce qu'il progressait rapidement que le fait d'être distrait par ses lunettes est devenu problématique. Le contenu de la proposition en *because* offre ensuite un contexte tout à fait favorable à l'apparition d'un rebondissement dans le récit. Ce rebondissement, souligné par l'adverbe *suddenly*, est introduit une première fois en 3, avant la proposition en *because* qu'il semble appeler. Il apparaît une seconde fois après celle-ci, lorsque, précisément, suffisamment d'éléments du « décor » sont réunis pour pouvoir poursuivre le propos principal.

6.2.1.2 *So* marqueur de « position »

Certains phénomènes semblent corroborer l'hypothèse d'une forte charge sémantique des marqueurs. C'est le cas notamment de la faculté des propositions en *so* à être employées à l'initiale d'un échange. La valeur conclusive habituellement associée à *so* ne peut être invoquée. En revanche, le marqueur peut être analysé comme permettant au discours de commencer, en désignant un segment de premier plan. Ainsi, *so* apparaît dès la première réplique de la conversation entre Celia et James :

C ok **so** now it's {0,80} <h> (0,30) {0,40} twenty {0,25} twenty-three minutes past <en riant>
J <rires> mm yeah
C and we stop when we're fed up <en riant>
J fair enough {0,30} did you say go and talk to Esther if they want some conversation {0,25} *is that <rires>*

(Celia&James : Esther, *so_001*)

Dans cet énoncé, Celia déclare sa conversation avec James ouverte. Première réplique de leur échange en tête-à-tête, elle s'inscrit en fait dans le contexte de la mise en place du matériel d'enregistrement et des rapides tests auxquels nous avons procédé avant de laisser les deux locuteurs seuls pour environ trente minutes. Il existe donc bien un avant sur lequel s'appuie cet énoncé. Nous sommes cependant loin de la valeur logique ou chronologique généralement associée au marqueur. *So* établit ici un point de repère initial pour le déroulement de la conversation et semble en cela plus fondamentalement participer à la construction narrative, et donc à une logique interne au discours. Avec *so*, Celia institue en effet ce premier énoncé en acte fondateur de la conversation. Elle construit le point de départ de la ligne directrice, ligne qui sera abandonnée dès la première réplique circonstanciée de James, mais qui permet à la conversation de démarrer. Notons par ailleurs que son énoncé établit une sorte de contrat entre les deux locuteurs : Celia initialise l'horloge et rappelle qu'ils peuvent décider d'arrêter la conversation à tout moment, conformément à ce que nous leur avions dit.

6.2.1.3 *And* marqueur polyvalent

D. Schiffrin associe *and* avec les segments de premier plan dans l'organisation discursive. Elle reconnaît néanmoins que, dans un contexte argumentatif, le marqueur peut également précéder des segments plus en retrait :

> [...] *and* can connect ideas locally within both of the functionnally differentiated sections of an argument — the position and the support.[12]

Dans le récit de Mike présenté plus haut, *and* sert à introduire des éléments de premier plan (2, 4, 5 et 7) mais également un élément d'arrière plan (8).[13] En 8, Mike fournit en effet à Hannah un détail nécessaire pour que celle-ci comprenne la situation. Il opère un retour en arrière dans la chronologie des événements et construit un cadre dans lequel prennent sens ses remarques précédentes, ainsi que les suivantes : pendant que Mike se faisait chahuter par le courant, Ann était en amont et attendait un signe de son compagnon lui indiquant si le cours d'eau était praticable ou s'il était préférable qu'elle regagne la rive et rejoigne Mike à pied.

And se distingue donc de *so* qui, par sa charge sémantique propre, semble pouvoir donner d'emblée un statut de premier plan à un énoncé, au point de suffire à donner son impulsion initiale à une conversation longue de près de trente minutes. Cette différence de fonctionnement tend à montrer que le statut principal ou secondaire des contenus informationnels vis-à-vis de la ligne directrice du discours n'est pas forcément lié au marqueur. L'interprétation du discours par l'interlocuteur s'appuie plus largement sur une lecture du contexte et de l'intention communicative du locuteur.

[12]SCHIFFRIN, Deborah (1987). *op. cit.* (pp.133-134).
[13]Voir l'exemple *because_044*, p.287.

6.2.2 Sur la délimitation et l'organisation des unités du discours

Les marqueurs considérés participent au moins à deux niveaux à la conduite du discours. Comme nous avons pu le voir, ils aident à la mise en place de ses unités les unes par rapport aux autres. Plus largement, ils contribuent à la segmentation du discours en unités, c'est-à-dire à leur identification et délimitation, sans forcément préjuger de leur statut discursif.

6.2.2.1 Délimitation et chaînage

En quoi *and* contribue-t-il à l'organisation des constituants du discours ? Dans l'exemple précédent, en dépit de son apparition dans le contexte d'un récit, *and* n'est pas cantonné à l'énumération à caractère chronologique. Nous venons de voir qu'il peut entrer en jeu dans une problématique de hiérarchisation des segments quant à leur pertinence vis-à-vis du propos principal du discours. Il apparaît que la charge sémantique de ce ligateur polyvalent, le plus fréquent dans le corpus utilisé par D. Schiffrin,[14] ne permet pas de présager de la nature du lien qu'il contribue à signaler. C'est ainsi qu'on peut le retrouver à la tête de segments au statut hiérarchique totalement différent, contrairement à *because/cos* et *so* qui semblent associés de façon plus régulière à un statut particulier. Si donc notre exemple montre en effet la capacité de certains marqueurs à signaler sur quel plan se situe l'incidence des propositions qu'elles introduisent, il contribue aussi à souligner l'importance du contexte et de l'intention communicative dans l'interprétation à donner à celles-ci. La fonction essentielle de *and* dans l'organisation discursive nous semble être d'opérer un simple chaînage de segments successivement énoncés.

Cette opération de chaînage a comme corollaire la délimitation des maillons de

[14]SCHIFFRIN, Deborah (1987). *op. cit.* (p.128).

la chaîne. Marqueur au sémantisme relativement neutre, *and* constitue de façon privilégiée une charnière entre les éléments qu'il relie. Il sert en cela à les délimiter. Ainsi, à la finale, le ligateur a certes une valeur continuative, mais il nous semble que, plus fondamentalement, il unifie le segment qui précède et lui donne un statut d'unité à part entière :

R *so*
H *well I've* just spent all day being {} harassed about being English again
R what <en imitant un accent français> talking like this to you::
H no just
oh bah elle est anglaise so oh she's English **and**
R why
H you know:: thingy {} mentioning no names
R yeah right <rires> yeah
H always hm {} always kind of {} acts yeah like {} it's really hilarious to {} point out I'm English every five minutes

(Helen&Rebecca : English)

L'unité ainsi délimitée peut ensuite être annulée par la suivante ou lui servir de point de départ. Ici, Rebecca profite de l'apparition du marqueur pour interrompre son interlocutrice : une place transitionnelle est atteinte, c'est-à-dire un moment où il est possible pour l'interlocuteur de prendre la parole sans menacer la face de l'autre.[15]

6.2.2.2 Niveau hiérarchique des éléments reliés

La conjonction de coordination relie des éléments de même niveau hiérarchique dans l'arbre syntaxique. Qu'en est-il, d'un point de vue discursif, du ligateur ? Nous avons vu plus haut que *so* signale le passage d'une unité discursivement subordonnée à une unité principale. Comment la hiérarchisation s'effectue-t-elle avec un marqueur sémantiquement neutre tel que *and* ?

D. Schiffrin affirme que les unités reliées par le marqueur sont « équivalentes », à la manière des unités syntaxiques dans la relation de coordination :

[15]Pour une définition détaillée des places transitionnelles, voir Chapitre 9, p.393.

> [...] the presence of *and* signals the speakers' identification of an upcoming unit which is coordinate in structure to some prior unit. But because texts contain units which are both locally and globally related, through either functional or referential means, *and* marks different levels of discourse structure. Thus, wherever we find *and*, we know we have a unit that is connected to a structurally equivalent unit somewhere in the prior discourse — but the identification of those units depends on the use of textual information beyond *and* itself.[16]

Comment réconcilier cette position, à laquelle nous souscrivons au moins en partie, avec l'utilisation de *and* pour introduire des remarques métalinguistiques ou pour passer du coq à l'âne, et donc relier des segments qui ne se situent pas sur un même plan référentiel ni même fonctionnel ? Nous avons vu plus haut que le marqueur *and* reste neutre quant à l'importance discursive de chaque segment : sa charge sémantique n'affecte pas l'interprétation de l'intérieur de l'énoncé, ce qui le rend apte à introduire tout type d'énoncé. C'est cette indifférenciation des segments qui, nous semble-t-il, permet de parler d'équivalence des unités reliées. La hiérarchisation des constituants du discours est une notion relative. Si tous les éléments sont susceptibles, dans le linéaire, d'être introduits par les mêmes marqueurs, ils sont potentiellement équivalents. La lecture du contenu sémantique et de l'intention communicative de l'énonciateur permet à l'interlocuteur de donner une interprétation à ce qui est laissé à l'état de chaîne indifférenciée par la syntaxe. Des unités complexes se dégagent, qui entrent à leur tour en relation avec des unités de même niveau hiérarchique sur le plan discursif.

Comparons le fonctionnement des différentes occurrences du marqueur *and* dans deux exemples tirés de notre corpus :

J 1. I know quite a few people in Oxford
2. **and** they all work
3. **and** that's why I have to ask John if

(Celia&James : Summer, *and_008*)

[16] *ibid.* (p.141).

Dans ce premier exemple, les unités signalées et reliées par *and* suivent l'ordre linéaire sur deux niveaux hiérarchiques différents. Les deux premiers énoncés appartiennent à des niveaux équivalents. A l'ensemble qu'ils forment, unifié par le second *and*, se greffe l'énoncé 3. Notons que le dessin formé par les ensembles n'apparaît qu'*a posteriori* :

[[1] and [2]] and [3]

Le niveau discursif supérieur atteint avec le dernier énoncé est confirmé par la possibilité de gloser par *so*, tandis que le premier *and* correspondrait davantage à un *but* adversatif :

glose I know quite a few people in Oxford **but** they all work **so** that's why I have to ask John if (I can do my experiments out of hours)

L'agencement hiérarchique des ensembles se fait différemment dans l'exemple suivant (*and_027*), où James explique la relation particulière que sa famille entretient avec la Chine.

J 1. yes Sally the older of my younger sisters is studying Chinese
2. **and** then Elspeth {0,50} who's in China at the moment is on a gap year
3. **and** she just chose to go to China <h>
C 4. yeah
J 5. ***and** my parents*
C 6. *well she probably* heard a lot about it
J 7. yeah I ff- I {0,30} possibly I don't know if that had much to do with it
8. **and** my parents <h> {1,00} had the idea of going on holiday there partly to visit Elspeth while she was out there

(Celia&James : China, Travel, *and_027*)

Trois ensembles hiérarchiques se dessinent, articulés par *and*. A l'intérieur de celui formé par les énoncés 2 et 3, une charnière d'ordre inférieur est soulignée par le même marqueur.

[1] and then [[2] and [3]] and [5, 8]

L'orientation hiérarchique des éléments est différente de celle observée pour l'exemple précédent. La portée du marqueur, variable, n'est pas fonction du marqueur lui-même :

and reste sémantiquement neutre, que son incidence soit locale ou globale. Elle ne répond pas non plus à une logique de recatégorisation implacable selon laquelle les énoncés 1 et 2, de niveau hiérarchiques équivalents, se combineraient pour former nécessairement un ensemble situé à un niveau hiérarchique supérieur, équivalent à celui de l'énoncé 3.

Ici encore, *and* se borne à signaler une charnière qui, lors de la reconstitution des ensembles hiérarchiques par l'interlocuteur, autorise celui-ci à considérer un ensemble comme clos et potentiellement suivi de n'importe quel autre ensemble clos. Nous pensons donc que le marqueur *and*, qui permet ainsi de traiter des unités comme *a priori* indifférenciées, les rend aptes à être hiérarchisées. C'est ainsi que la chaîne, horizontale, peut se transformer en arbre.

L'« équivalence » des unités ainsi reliées n'est pas celle des unités coordonnées, dans la mesure où la hiérarchisation discursive est un phénomène dynamique qui s'étend au-delà du domaine de l'énoncé. L'arbre syntaxique décrit l'architecture figée d'un énoncé. En discours, la portée d'un segment est réévaluée à mesure que se déroule l'échange.

6.2.2.3 Réseaux de marqueurs

L'interprétation de la portée d'un segment, laissée ouverte par l'emploi d'un marqueur comme *and*, est notamment guidée de façon dynamique par le réseau dans lequel le marqueur s'inscrit.

And Un principe de relativité semble gouverner les marqueurs et leur portée. Particulièrement lorsque leur apport sémantique est peu spécifiant, les marqueurs prennent sens essentiellement les uns par rapport aux autres. C'est ce que remarque D. Schiffrin, à propos de l'interchangeabilité de *and* et du marqueur zéro :

> [...] the more general principle underlying this flexibility between *and* and 'zero'

> concerns the creation of syntagmatic contrasts: once a textual regularity has been developed — even if it is developed for only a short sequence — a new idea unit can be introduced by a change from that textual norm. The participation of *and* in this syntagmatic pattern suggests that *and* conveys no more referential information than 'zero'.[17]

Revenons à notre exemple *because_044* cité plus longuement p.287 :

M 2. **and then** I suddenly looked up
3. because I at this time I was travelling quite fast you know <h> (0,3)
4. **and** I looked up
and I realised this was coming
and I was like uuuaaah
5. **and then** I realised I'd had my glasses in it (?)

(Mike&Hannah : Little near-death experience, *because_044*)

C'est en effet à la lumière du contexte d'apparition du marqueur que nous pouvons différencier le fonctionnement de *and* en 4 et en 5. En 4, Mike relate une succession d'événements, présentés dans un ordre chronologique : l'énumération, relayée par *and*, inscrit cette succession dans l'ordre du discours. La neutralité sémantique du marqueur et de la relation qu'il souligne permet à chaque segment de se lire comme découlant naturellement du précédent. *And* contribue à l'intensité dramatique du passage en permettant au discours de mimer l'enchaînement des événements et la vitesse à laquelle le danger est arrivé. Le véritable rebondissement, en forme d'anti-climax, se situe en 5. Il est introduit par un ligateur complexe, *and then*, qui répond à celui utilisé en 2 et tend à décrocher l'énoncé 5 par rapport à l'ensemble précédent, tout en le rapprochant de l'énoncé 2. *And then* brise ainsi la « norme textuelle » ou du moins la régularité installée en 4 par la succession des *and*. Au réseau formé par les différentes occurrences du marqueur en 4 se superpose un autre réseau. Le parallélisme de la construction en 2 et en 5 permet ainsi à deux moments du récit d'émerger sur

[17]SCHIFFRIN, Deborah (1987). *op. cit.* (p.131).

le même plan : Mike voit le danger venir, puis il se rend compte qu'il n'a finalement pas égaré ses lunettes.

En évaluant la portée d'un marqueur relativement à son contexte, la « norme textuelle » établit donc des réseaux sur le plan local et global. Générée dynamiquement, elle s'appuie cependant sur des schémas relativement figés. La séquence ∅ + ∅ + *and* fait partie de ceux-ci, de même que *because/cos* + *so*.

Un schéma récurrent : *because/cos* + *so* Le chapitre précédent a établi la préférence de *so* pour *because* par rapport à *cos*, sans pour autant écarter totalement ce dernier.[18] Nous avions relié cette corrélation à une valeur plus argumentative de *because* qui semble en fait émaner du contexte dans lequel le marqueur apparaît plus volontiers que *cos*. La valeur discursive de *so*, que nous décrivions comme un marqueur argumentatif, s'éloigne en outre souvent d'une valeur logique en dehors d'un contexte argumentatif.

Ne nous situant plus dans une problématique de différenciation entre *because* et *cos*, nous traitons ici les deux marqueurs comme un seul. Comme nous avons pu le voir plus haut, la valeur discursive des segments introduits par l'un et l'autre des marqueurs est relativement stable. En contexte argumentatif, *because/cos* tend à introduire une prémisse, tandis que la conclusion est précédée de *so*. En contexte descriptif, l'arrière plan est distingué de la ligne principale du récit. En présence de *so*, cette hiérarchisation binaire entre segments principaux et segments secondaires est en fait établie sur trois segments : d'une part, *because* q introduit des éléments secondaires par rapport à p, d'autre part, l'énoncé en *so* introduit des éléments principaux par rapport à q. Se pose la question du rapport entre l'énoncé en *so* et p dans l'organisation du discours.

Nous souhaitons ici nous attarder sur la valeur anaphorique de *so*, décrite par

[18] Voir p.274.

exemple par J.-R. Lapaire & W. Rotgé.[19] Celle-ci se trouve largement vérifiée dans notre corpus dans des cas où *because/cos* et *so* conservent un sémantisme causal. C'est le cas dans l'exemple suivant, où James s'inquiète de savoir si Celia doit obtenir l'accord de John Coleman avant de lui donner les clefs du laboratoire que celui-ci dirige :

J *do you technically* need John's sort of permission *then*
C *no*
J to get
hm no
C no
J <h> (0,35) ok <h> (0,55)
C no **because::** I mean {0,50} I know he wants you to have keys **so** that's fine
J does he {0,80}
did he tell you that *<rires>*
C *well if you're doing it* after hours you have to

(Celia&James : John's permission, *because_020*)

La question de James trouve une réponse articulée en trois points qui contribuent tous trois à lui assurer que Celia n'a pas besoin de l'autorisation de John : la négation (plusieurs fois répétée), l'énoncé en *because* et l'énoncé en *so*. L'explication apportée par le contenu de l'énoncé en *because* semble se rapporter aussi bien à sa gauche à la négation qu'à sa droite à l'énoncé en *so*. Ces deux segments sont en effet fortement redondants d'un point du vue informationnel. En somme, cet exemple fait apparaître un schéma circulaire en trois parties : Celia répond à la question de James, elle explique ensuite sa réponse à l'aide d'une proposition causale en *because* qui lui sert finalement de point d'appui pour une proposition causale en *so* anaphorique de la réponse qu'elle a déjà donnée à James. Nous pouvons schématiser la boucle discursive ainsi formée de la façon suivante :

A *because/cos* B *so* A'

[19]LAPAIRE, Jean-Rémi & ROTGÉ, Wilfrid (1991). *Linguistique et grammaire de l'anglais.* Toulouse : Presses Universitaires du Mirail. (pp.287-288).

Ce phénomène de réitération du contenu de p dans un énoncé en *so* suivant un énoncé en *because* semble être une caractéristique de l'anglais oral spontané. Nous le retrouvons par exemple dans le corpus de D. Schiffrin de façon récurrente.[20] Il est encore visible dans des exemples où la séquence p *because/cos* q est suivie du seul *so* : tous les éléments étant donnés pour reconstruire p', il n'est pas besoin de l'énoncer. Dans l'exemple *because_024*, Helen apprend qu'elle a manqué un coup de téléphone de la part de ce que Rebecca pense être une employeuse potentielle :

R I think what happened was she tried to phone you on your mobile telephone but if i::t rang twice
H right
R and cos it is as easy as on a computer *I could find out she had left a message*
H *actually probably it wasn't her **because*** I didn't *give them my f- my phone number **so***
R *oh ok probably she left* a message on the
H yeah

(Helen&Rebecca : RFI, *because_024*)

So invite Rebecca à reconstruire pour elle-même la proposition p à partir des éléments donnés par Helen dans la proposition en *because*. La réplique suivante de Rebecca fait état du parcours argumentatif effectué par celle-ci : l'exclamatif, la marque d'assertion et l'adverbe épistémique témoignent de son appropriation du contenu de pensée de Helen. L'amorce de la boucle suffit à la suggérer et à inscrire les segments de cet exemple dans un système ternaire.

M. De Cola-Sekali voit une organisation tripartite similaire dans les énoncés où la proposition en *because/cos* est antéposée par rapport à la principale. L'analyse traditionnelle de ce cas de figure, rare à l'oral spontané, est ici remise en question. La chercheuse s'appuie sur la position relative des propositions cible (p) et repère (q) :

> *Because* n'est en fait jamais réellement antéposé : la cible du repérage qu'il effectue est toujours située à gauche du repère. La configuration dite « antéposée » n'est que l'indice de la présence d'un contexte avant pertinent dans la relation : elle signale que c'est le contexte avant qui est cible de la connexion. Le statut

[20] SCHIFFRIN, Deborah (1987). *op. cit.* Voir notamment les exemples 1 pp.49-50 ou 2 p.129.

de p dans *because q, p* est alors différent du statut de p dans *p because q* : dans le premier cas, p n'est pas la cible de la connexion, mais une reformulation d'un énoncé p1 préalable. L'énoncé complexe n'est alors plus envisageable comme binaire, il constitue une progression argumentative en trois temps de type :

1. cible d'un débat → 2. relation à un repère → 3. réassertion

où *because q* occupe en réalité une position médiane.[21]

Ces ressemblances avec le phénomène d'organisation discursive ternaire rencontrée dans notre corpus nous permettent-elles d'assimiler la proposition p postposée de M. De Cola-Sekali à notre proposition en *so* ? L'exemple suivant suggère les limites d'une telle assimilation. Y apparaissent en effet conjointement deux séquences incompatibles si elles se situent sur un même paradigme : une séquence *because* q est antéposée et une séquence où p est suivie d'une proposition en *so* :

R how much actually I've got hm about fifteen francs if you want m- money for drink or something
H actually I've got my water that's fine
R you can use that with *pleasure*
H *no it's alright* I got water
R *hm*
H *hm* and then **because** Sarah 0,20 didn't have to teach now she's not gonna come in **so** 0,70 I'll be all alone
R how come she doesn't have to teach and you do

(Helen&Rebecca : Staying awake till 7, *because_028*)

Helen énumère ses soucis de la journée : elle ne reprend les cours qu'à 19h, se sent très fatiguée, a oublié son porte monnaie chez elle, et va devoir passer plusieurs heures seule dans la salle des enseignants. C'est sa dernière réplique qui nous intéresse. Le motif ternaire est toujours présent, mais il semble difficile de trouver une cible située à gauche du repère que constitue la proposition *because* q. Il ne semble en effet pas possible de restituer, dans le contexte avant, une proposition p1 dont p soit

[21] DE COLA-SEKALI, Martine (1991a). Connexion inter-énoncés et structuration des relations temporelles et argumentatives en anglais contemporain. Une étude énonciative des connecteurs polyvalents SINCE et FOR. Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (pp.338-339).

la réélaboration. La proposition qui constitue une réélaboration de p (sous la forme d'une généralisation) se situe en aval : il s'agit de la proposition en *so*. Faut-il sinon considérer l'exclamation de Helen (*oh I don't know if I can stay awake here till seven that's awful*) située quelques secondes plus tôt comme la proposition p1 cible de la connexion ? Nous n'avons pas l'ambition d'apporter ici une réponse ferme. Il nous semble seulement que le débat intersubjectif concerne ici le contenu de la proposition en *so*. Suivi d'une longue pause, le marqueur unifie les unités discursives précédentes pour leur rapporter une unité supplémentaire pertinente. Ici encore, *because* s'éloigne de la ligne principale en donnant des éléments d'arrière-plan, tandis que *so* permet d'y retourner, en énonçant un contenu déjà présent ou en suggérant de le reconstruire.

La charge sémantique des marqueurs et leur organisation en réseaux signifiants permet de baliser l'espace discursif et d'orienter l'interprétation à donner aux différents segments énoncés. La hiérarchisation tripartite des constituants discursifs nous semble pertinente au-delà d'une étude des marqueurs dits de subordination ou de coordination : en quoi une organisation thématique en spirale contribue-t-elle à la construction intersubjective du discours ?

6.3 Progression discursive en spirale et gestion de la relation intersujets

Derrière celle de la linéarité discursive se pose la question plus générale de la construction de la relation intersubjective. En quoi le motif de spirale contribue-t-il à la fois à la progression discursive et à la construction de la relation intersujets ?

6.3.1 Spirale et circularité

L'apparente circularité des argumentations relevées précédemment est loin de constituer pour nous ou pour les interlocuteurs une faute de logique. La ligne du discours n'est pas statique. Son caractère récursif lui assure un dynamisme important. Il semble en outre inscrit à différents niveaux dans notre expérience de l'échange langagier.

6.3.1.1 Parataxe de l'enfant

Les travaux d'A. Morgenstern & M. De Cola-Sekali témoignent d'une utilisation semblable de l'ordre linéaire dans l'acquisition de l'argumentation par le jeune enfant. A l'âge d'un an et dix mois, Léonard (L) dessine avec sa mère :

M : Regarde, on va faire un rond rouge.

Elle le fait. Il crie et il secoue la tête.

L : **anonanon / anonanononanon //**

M : Non, tu veux pas ? Regarde, c'est joli.

L : **arouz / non : : : //**

Il tape sur le dessin.

M : C'est pas beau ? Alors on fait un autre.

L tape sur la feuille.

L : **aa : / nonarouzarouzepabo** //[22]

Deux séquences énoncées par l'enfant nous intéressent ici sont celles formées d'une part par la première et la deuxième réplique de Léonard (séquence 1 : *anonanon / anonanononanon // arouz / non : : : //*), et d'autre part par la deuxième et la troisième (séquence 2 : *arouz / non : : : // aa : / nonarouzarouzepabo //*).

Dans la première séquence, l'enfant construit une boucle semblable à celle formée par les énoncés combinant *because* et *so* de notre corpus. En l'absence de marqueurs

[22]MORGENSTERN, Aliyah & SEKALI, Martine (à paraître). De la parataxe aux premiers connecteurs : les prémices de l'argumentation chez le jeune enfant. Cahiers de l'ENS-LSH. (pp.6-7). Nous soulignons.

de connexion, c'est l'ordre des segments et la reprise du *non* qui, dans ce contexte argumentatif, confère à *arouz* (un rouge) une valeur explicative :

> Si l'on formalise par A l'explanandum (*non*) et par B l'explanans (*rouge*), on a ici la configuration : A (B-A). C'est la reprise du non après une nouvelle prédication (*rouge*) qui établit une relation explicative entre les deux séquences verbales.[23]

A cette boucle en A B A' succède une configuration appelée « encerclement » par les deux chercheuses, dans laquelle les deux membres de la structure argumentative sont répétés : A B B A' (séquence 2)

> [...] où A' est une re-formulation de A : *est pas beau*, repris de la mère s'apparente à une re-formulation conclusive de son refus (*non*). Entre A et A', il y a eu explication : B (*rouge*) a acquis un statut explicatif dans une sorte d'encerclement entre deux assertions de son opposition.[24]

La duplication d'au moins l'un des termes de l'explication et leur alternance favorise la compréhension par l'adulte de la démarche explicative de l'enfant :

> C'est un ordonnancement linéaire, de type assertion - explication - re-assertion, qui marque l'existence d'une argumentation.[25]

L'enfant cherche à signaler et combler une différence de représentations. Ici encore, la spirale explicative de l'enfant n'est pas stérile mais orientée vers la poursuite de l'échange.

Les locuteurs adultes de notre corpus ont acquis l'usage des marqueurs argumentatifs. Notre corpus montre cependant qu'ils superposent à leur utilisation la caractéristique de l'argumentation enfantine décrite par A. Morgenstern & M. De Cola-Sekali. Comme nous l'avons vu, l'utilisation des marqueurs autorise les locuteurs à

[23] *ibid.* (p.7).
[24] *ibid.* (p.7).
[25] *ibid.* (p.7).

ne pas redire A après *so* dans une séquence combinant *because/cos* et *so*. Mais la présence de *so* suffit à elle seule à suggérer la spirale. En combinant les deux modes d'organisation, la langue adulte montre ici un aspect redondant de son économie.

6.3.1.2 Analyse d'un exemple : la question des clés

Nous avons pu voir comment la combinaison des marqueurs *because/cos* et *so* contribue à souligner cette spirale qui nous semble être une caractéristique plus générale de l'organisation discursive à l'oral spontané. Nous retrouvons en effet le même motif dynamique récursif non plus au niveau local inter-propositionnel ou inter-énoncés immédiat mais au niveau plus large des épisodes thématiques distingués dans notre corpus. Concentrons-nous sur l'épisode que nous avons intitulé *Keys* dans la conversation entre Celia et James. Deux thèmes principaux s'entremêlent : d'une part, celui des emplois du temps et des allées et venues des membres du laboratoire de phonétique pendant les vacances d'été, et, d'autre part, celui de la clef que James cherche à obtenir de Celia pour accéder au laboratoire en dehors des heures ouvrables.[26]

Etude linéaire de l'épisode *Keys* Sept séquences thématiques apparaissent dans cet épisode, qui s'ouvre par la réplique de Celia :

> **C** yeah {1,20} we must yeah {0,95} <h> actually {0,60} you're around next week aren't you I was just thinking about this thing of:: key::s and {0,40} <h>
>
> (Celia&James : Keys, séquence_1)

Cette première séquence constitue une transition relativement abrupte entre l'épisode *Keys* et la fin de l'épisode *English corner*, qui clôt la discussion sur les pays asiatiques entamée dix minutes plus tôt. Plusieurs phénomènes concourent cependant à faciliter cette transition entre le long ensemble thématique auquel Celia met fin et celui qu'elle

[26]Nous utilisons ici le terme de « thème » dans son sens courant de sujet ou isotopie.

initie. Nous pensons tout d'abord aux hésitations qui émaillent sa prise de parole — répétition de *yeah*, faux départ avec *we must*, et nombreuses pauses silencieuses — autant de moyens pour signaler la transition et annoncer la nouvelle séquence thématique tout en retardant l'apparition de son contenu. De nombreux indices d'une construction du consensus apparaissent également. L'utilisation des pronoms deuxième personne du singulier et première personne du pluriel, par exemple, tend à associer l'interlocuteur au contenu de l'énonciation. L'interrogative, sous la forme d'une *tag question*, permet d'asserter un contenu tout en le soumettant à l'autre pour vérification : le premier thème principal de cet épisode est introduit. Enfin, la mention tardive (retardée entre autres par la locution *this thing of*) et allongée de ce qui constitue l'enjeu de toute la conversation met en valeur la question des clés qui peuvent devenir une isotopie principale de cet épisode.

Si cette première séquence tient ensemble les deux thèmes principaux de cet épisode, la deuxième les sépare pour ne s'intéresser qu'aux allées et venues des uns et des autres, en l'occurrence, celles de John :

C you know John's going on a holiday
J a::t the end of this month
C at the end of:: next week
J oh *really*
C *yes* <h> (0,20)
J oh

(Celia&James : Keys, séquence_2)

La transition, bien que plus courte, utilise ici les mêmes moyens que dans la première séquence : Celia cherche à construire un consensus en sollicitant James sur un contenu de pensée qu'elle lui attribue.

La troisième séquence développe le second thème principal de l'épisode : la question des clés du laboratoire.

C so we m- so it's slightly at the back of my mind we must {0,25} get
hm {0,45} <h> I can {0,75} <h> I can introd- duce you to the alarm system and:: {0,45} issue you

with some keys either <h> {0,45} today I'm not working tomorrow {0,90} <h> or:: one day next week

J okay

(Celia&James : Keys, séquence_3)

La cohésion discursive est assurée tout d'abord par le marqueur *so*, qui indique une prééminence hiérarchique du contenu qu'il introduit par rapport à ce qui précède. Par là, le marqueur relie les segments situés de part et d'autre. Nous retrouvons ici encore certaines caractéristiques de la première séquence, les nombreuses hésitations assurant une transition souple. Notons également le rappel d'une formulation abandonnée : *we* et *we must {0,25} get* constituant une nouvelle tentative de faire aboutir l'énoncé commencé dans la première séquence (*we must*). La présente séquence, dont le thème est bien celui des clés, prépare l'isotopie de la séquence suivante :

C and I'm {0,30} normally work Monday {1,05} <claquement de langue> Wednesday and Thursday {0,60} hm {0,85} that's normal

J okay

(Celia&James : Keys, séquence_4)

Après cette courte séquence décrivant l'emploi du temps de Celia, la cinquième séquence offre une synthèse des deux thèmes :

C but whenever's convenient {0,65}
but we ought to do it before he's away because then there may be more of a problem {0,85} <h>
hm <h> {1,40} I shall certainly be away the first week {1,50} that he is away {0,80} <h>
so you will need keys <rires> if you're guaranteed to get in when you want to

J okay

(Celia&James : Keys, séquence_5)

Ici, il nous semble que le marqueur *so* sert non pas à introduire une séquence mais à apporter une conclusion mêlant les deux thèmes abordés : puisque les membres du laboratoire seront plus ou moins tous en vacances en même temps, James a besoin des clés pour pouvoir entrer.

Cette synthèse est néanmoins suivie par une sixième séquence marquant un retour à l'isotopie de l'emploi du temps des uns et des autres. Il s'agit ici de celui de Andy :

C hm {0,40} <h> and I don't know what what other people are going t- I think Andy will be around most of the *time*

(Celia&James : Keys, séquence_6)

Suit une longue séquence initiée par la question de James (que nous reportons ici), qui fait revenir Celia à la thématique de l'organisation des clés :

J *okay* is that {0,20} I mean is that a key for {0,20}

C *but*

J *the* door of forty-one *or: is it*

C *<h> ya*

J just

(Celia&James : Keys, séquence_7)

Cette séquence, qui clôt l'épisode *Keys*, atteint une première tentative de conclusion marquée par *so* et *basically*. La question n'étant pas épuisée à ce moment, la véritable conclusion n'est atteinte que plus loin. Avec *so* et *just*, Celia indique que la question des clés se résume finalement à peu d'éléments :

C <h> (0,60) hm {0,95} so it's just a door to get into er key to get into the buildings {0,45} <h> and a key to:: <h> (0,40) get into the lab

(Celia&James : Keys, séquence_7)

La prise de conscience amorcée ici se poursuit dans l'épisode suivant, qui constitue pour nous la conclusion générale de l'épisode *Keys*, détachée des deux thèmes principaux évoqués :[27]

C that's quite simple really

(Celia&James : Recording)

Cette conclusion ne semble cependant pas convenir aux locuteurs, si bien que le débat sur les clés reprend brièvement (il est à présent question du code d'entrée du bâtiment), avant d'être définitivement abandonné.

[27] La frontière entre épisodes n'est pas liée à une frontière thématique. Elle a pour seule justification la contrainte pour un épisode d'éviter de dépasser deux minutes de longueur, l'épisode *Keys* étant long d'une minute et cinquante-huit secondes.

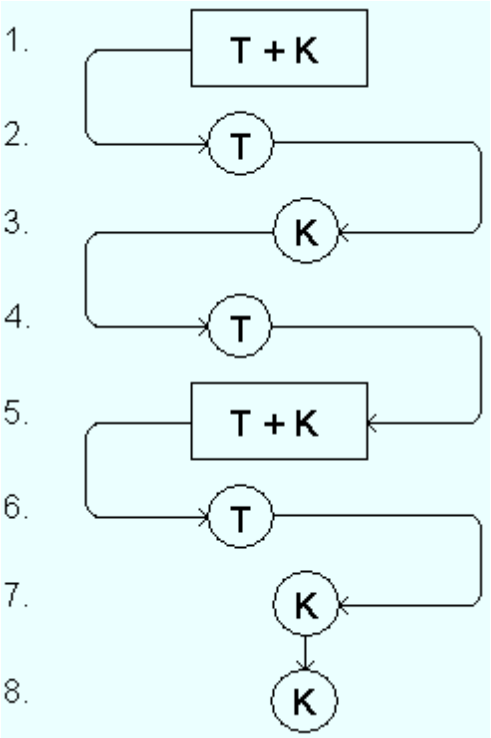

FIG. 6.1 – Spirale discursive entremêlant les thèmes des emplois du temps (T) et des clés (K)

L'analyse linéaire des thèmes abordés dans cet épisode de notre corpus suffit à montrer comment le motif de la spirale est un moteur de la construction thématique du discours. La Figure 6.1 schématise cette spirale discursive entremêlant les deux thèmes principaux de ce passage. L'alternance des thèmes leur permet de se nourrir l'un de l'autre et d'assurer la progression discursive de l'épisode.

Inscription de l'épisode dans l'ensemble de la conversation L'épisode *Keys*, dont l'organisation thématique interne est caractérisée par un fort dynamisme récursif, s'inscrit lui aussi au cœur d'une spirale à l'amplitude plus large, à l'échelle de la totalité de l'échange.

L'isotopie de la clef a une importance de premier plan dans l'échange. C'est la raison d'être de la conversation entre Celia et James. Elle a donc été abordée par les interlocuteurs avant que nous ne commencions à enregistrer leur échange (et donc à en influencer la teneur). Celia et James abordent à nouveau la question des clés dans l'épisode intitulé *Summer*. Le thème apparaît pour la première fois dans l'enregistrement au détour d'un énoncé, après que Celia a demandé à James s'il lui est difficile de trouver des sujets d'expérimentation. Suggéré par James dans sa première réplique explicative, le thème est identifié par Celia par *we must get keys organised* puis par *you need yeah {0,45} keys and the alarm* :

J I know quite a few people in Oxford and they all work and that's why I have to ask John if
C oh out of *hours you*
J *yeah*
C we must get *keys organised*
J *ye {0,25} yeah* {0,40} I mea::n
C cos that's fine but *hm*
J *and there's* the alarm as well *and stuff*
C *you need* yeah {0,45} *keys and the alarm*
J *cos I don't even know the::* er combination at the moment <h>
C for the door {0,70} well that's easier *though*
J *when I s-* bring people here they say don't you know the combination <rires> and I say well <rires> actually I've never asked <h> (0,25) don't know if I'm allowed to that or not that I just ring the bell

every time *and* <rires>

C *no* you're allowed to know it {0,35}

that won't be John's taste <rires>

J <rires> *yeah*

C *but yeah*

I'll te- tell it to you when we turn the tapes off

J right <rires>

C but hm {0,40} it changes anyway

J *yeah*

C *hm* {0,25} but it won't be *changing for a while*

(Celia&James : Summer, keys_1)

Les éléments thématiques de l'épisode analysé plus haut sont introduits ici : les clés et le code, ainsi que la question des emplois du temps. Il semble d'ailleurs que le problème qui a amené James dans le bureau de Celia soit déjà réglé : puisque James doit pouvoir accéder au laboratoire en dehors des heures ouvrables, Celia est d'accord pour donner des clés à James et lui promet de lui donner le code du bâtiment une fois l'enregistrement terminé. Comme nous avons pu le voir, la même isotopie est pourtant reprise et développée environ 15 minutes plus tard (soit 10 épisodes) dans l'épisode *Keys*, où les interlocuteurs identifient de quelles clés James doit disposer pour pouvoir accéder au laboratoire quand il en a besoin.

La même thématique est à nouveau abordée en fin d'enregistrement, cinq minutes plus tard, dans l'épisode intitulé *John's permission*. C'est le groupe prépositionnel de James, permettant le coq-à-l'âne, qui réoriente la conversation vers le sujet, avant de proposer un développement supplémentaire avec l'interrogative :

J *hm y- about getting in the room* I mean

C *mm*

J *do you technically* need John's sort of permission *then*

C *no*

J to get

(Celia&James : John's permission, keys_3)

Notons que l'isotopie de la clef est à nouveau fortement liée à celle des emplois du temps, qui constituent le thème de la suite de la séquence reportée ici. En outre,

la conclusion de ce nouvel épisode correspond à celle du premier, montrant que la question était bien déjà réglée en substance à la fin de la première mention du thème de la clef. Seuls des détails sont apportés par les nombreux développements observés.

Nous remarquons encore que, à la faveur d'un incident technique concernant le micro-cravate de Celia, la double thématique des clés et de l'emploi du temps des membres du laboratoire développée dans *Keys* fait place à un nouvel épisode thématique qui occupe les interlocuteurs pendant les cinq minutes suivantes (épisodes *Recording*, *Diphthongs1* et *2*). L'épisode thématique qui s'ouvre alors fait largement écho au second épisode de la conversation. Dans les deux cas, James explique qu'il n'est pas à l'aise lorsqu'il s'agit de mener des expériences pratiques.

Plusieurs spirales thématiques se dessinent donc, à différents niveaux, à travers le corpus Celia&James. Ce flux tourbillonnaire alimente le discours des mêmes thèmes sans pour autant le faire tourner en rond. Cette récursivité est un principe dynamique remarquable sur le plan de la progression discursive, mais également sur celui de la gestion de la relation inter-sujets.

6.3.2 Gestion de la relation intersujets

Inscrit dans l'économie de l'oral spontané, le motif de la spirale prend la forme d'une recatégorisation des constituants discursifs. Il contribue ainsi à la construction de l'attention partagée et à l'établissement des points consensuels grâce auxquels progressent le discours et l'échange en général.

6.3.2.1 Regarder vers l'arrière

La réutilisation du même pour créer du différent fait partie des principes essentiels décrits par M.A.K. Halliday & R. Hasan[28] assurant la cohésion du discours en lui conférant une « texture ».

[28]HALLIDAY, Michael A.K. & HASAN, Ruqaiya (1976). *Cohesion in English*. Londres : Longman.

Les deux auteurs distinguent plusieurs modes de relations qui unissent les constituants d'un texte, dont la substitution et l'ellipse d'une part, et, d'autre part, la relation de référence. Les premières unissent les éléments de façon formelle au niveau lexicogrammatical, expliquent-ils, tandis que la seconde est en fait une relation sémantique.[29] Ces deux types de cohésion ont en commun leur regard tourné vers l'arrière du discours, c'est-à-dire vers le contexte avant, qui permet de retrouver l'élément dont on a fait l'économie ou simplement d'en inférer la référence.

Si le discours regarde vers l'arrière pour pouvoir aller de l'avant, il est tout naturel de voir s'inscrire dans son économie le motif de la spirale à divers niveaux de sa construction. En termes informationnels, le nouveau dans l'énoncé *n* devient du donné sur lequel l'énoncé *n+1* peut s'appuyer. La zone d'information minimale (ou zone thématique)[30] située sur la gauche de l'énoncé accueille typiquement des éléments déjà connus ou plus consensuels. Elle laisse place en fin d'énoncé à un segment nouveau, ou traité comme nouveau, ou encore à un segment plus crucial sur le plan intersubjectif, et sur lequel porte le focus intonatif. Les choix de structures syntaxiques reflètent donc ces considérations. L'antéposition du complément indirect dans l'énoncé suivant permet, en plaçant *second subject I had yesterday* en zone thématique, de mettre en relief le segment qui fait l'enjeu de l'énoncé et de la séquence : James est si peu doué pour les travaux pratiques qu'il est capable, au cours d'une de ses expériences, d'oublier de mettre le magnétophone en route.

J well I mean *the second subject* [...] second subject I had yesterday I forgot to turn the tape recorder on

(Celia&James : Experiments)

Par contraste, l'ordre linéaire dans l'énoncé canonique que nous proposons (en ajoutant la préposition *with*) ne permet pas de faire ressortir le segment faisant l'objet d'une demande de consensus dans l'énoncé de James :

[29] *ibid.* (p.304).
[30] Thématique est ici utilisé par opposition à rhématique.

Enoncé canonique envisagé I forgot to turn the tape recorder on *with* the second subject I had yesterday.

Le même principe de retour sur soi est à l'œuvre, au sein du paragraphe oral, dans le phénomène de recatégorisation de la fonction des segments à mesure que le discours avance. C'est ainsi qu'un ensemble [préambule + rhème] peut servir de préambule dans le paragraphe suivant. Nous donnerons pour seul exemple la séquence suivante, dans laquelle le premier énoncé permet de restituer la référence du pronom personnel sujet du second. Le premier énoncé permet d'établir le cadre thématique sur lequel le second peut se construire. Il est recatégorisé en préambule :

M I went with a friend called Ann {0,35} and we walked down {0,4} a river in central Greece

(Mike&Hannah : Trip)

Ce n'est pas un souci de présentation chronologique des événements qui préside au choix de l'ordre des énoncés. La succession marquée par *and* a davantage à voir avec l'économie du discours : une fois les acteurs présentés, Mike peut commencer le récit dans lequel ils apparaissent.

Si l'ordre du linéaire ne peut établir que des relations binaires entre éléments voisins, la réutilisation du même permet aux différents segments d'étendre leur domaine relationnel. Le schéma de la spirale discursive, en regardant vers l'arrière et reformulant du déjà donné, crée des jeux d'écho entre maillons de la chaîne. Comme nous avons pu le voir plus haut, de l'intrication de deux thèmes émane typiquement un troisième. La progression du discours procède par contacts et glissements successifs inscrits dans sa linéarité. Ce travail sur la chaîne linéaire est essentiel à la gestion de la relation intersubjective.

6.3.2.2 Pour aller vers l'avant

Au-delà de la question du poids informationnel des éléments du discours, le motif de la spirale nous semble se justifier du point de vue de la construction de la relation

intersujets. Il constitue en effet un moyen de faire progresser le discours de consensus en consensus, ce qu'illustre de façon particulièrement claire son application au domaine causal.

Revenons sur les structures du types A *because* B *so* A'. Ces structures ont pour véritable enjeu la construction d'un consensus à partir de la reconnaissance d'une possible divergence de représentations mentales. Selon D. Schiffrin, *because q* octroie au contenu de *p* le statut d'information partagée, que vient confirmer le marqueur *so* :

> In sum, *because* can be used to preface information when the status of that information as shared background knowledge is uncertain and when that information is important for understanding adjacent talk. This use marks an information state transition because it is a shift from unshared to shared knowledge. And *so* can be used to preface information whose understanding is supplemented by information which has just become shared background.[31]

En termes énonciatifs, le contenu de *p* fait l'objet d'une demande de consensus. *Because q* offre des éléments permettant d'atteindre ce consensus. Dans l'énoncé en *so*, le contenu de *p* est considéré comme consensuel, compte tenu de ce qui en a été dit. Une fois ce point de consensus établit, il devient possible d'envisager la construction du suivant.

Avec ce chapitre se clôt notre partie concernant les modes d'organisation du discours à l'oral sur le plan segmental, c'est-à-dire essentiellement d'un point de vue syntaxique, discursif, et intersubjectif. A partir de l'exemple de la subordination et en particulier de la subordination causale, nous avons vu s'illustrer une caractéristique de l'oral spontané qui tend à voir les locuteurs ne profiter que d'une infime partie des possibilités offertes par la grammaire en terme de choix des marqueurs et des

[31]SCHIFFRIN, Deborah (1987). *Discourse Markers*. Cambridge : Cambridge University Press. (p.207).

structures. Parmi ces marqueurs, *because* et *cos* nous ont semblés particulièrement intéressants, à travers leurs points communs mais également leurs différences. Emerge donc une syntaxe autre avec ses marqueurs privilégiés et ses structures récurrentes. Un motif dominant, détaché du domaine simplement argumentatif, confirme notre hypothèse du discours comme un objet fractal. La spirale est à nos yeux un principe essentiel dans l'économie du discours conversationnel, à différents niveaux.

Deux de nos hypothèses de départ se trouvent donc confortées par ces résultats. Qu'en est-il de la troisième ? Comment les paramètres intonatifs entrent-ils en jeu dans la hiérarchisation des constituants discursifs à l'oral spontané tel qu'il s'illustre dans notre corpus ?

Troisième partie

De quelques faits suprasegmentaux de hiérarchisation

Une étude sur la hiérarchisation des constituants discursifs dans un corpus d'anglais oral spontané ne peut ignorer la dimension suprasegmentale de l'édifice conversationnel. L'une de nos hypothèses de départ suggère en effet que la syntaxe de l'oral, dont nous avons décrit certaines caractéristiques dans la partie précédente, trouve un contrepoint structurant dans l'intonation. En outre, nous faisons l'hypothèse que le fonctionnement de l'intonation est non seulement iconique mais également conventionnel, et que la gestion des paramètres fait au moins en partie l'objet de choix motivés de la part de l'énonciateur. Enfin, notre troisième hypothèse considère la conversation spontanée en anglais comme un objet fractal, ce qui nous invite à évaluer la pertinence de certains principes de structuration à différents niveaux.

Comment les paramètres prosodiques s'inscrivent-ils dans une logique de délimitation, de mise en place et d'organisation des éléments du discours ? Dans quelle mesure sont-ils pertinents aux niveaux discursif, intersubjectif et interactionnel ?

Il s'agit dans ce dernier développement d'évaluer la place des paramètres suprasegmentaux dans la construction de l'échange verbal. Nous examinerons tout d'abord de quelle manière la ligne mélodique renseigne sur l'état de la relation inter-sujets. Ensuite, nous nous intéresserons au rôle de la pause silencieuse dans la délimitation de différents types de charnières entre constituants de diverses tailles. Enfin, nous nous concentrerons sur le niveau interactionnel, qui permet de considérer la conjonction d'un faisceau de paramètres, parmi lesquels les paramètres suprasegmentaux, dans la gestion des tours de parole.

Chapitre 7

La gestion de l'espace intersubjectif par la prosodie

Nous avons évoqué dans un chapitre précédent une hypothèse dont la vérification fait partie du projet de départ de ce travail. Celle-ci concerne le fonctionnement décorrélé des paramètres prosodiques et leur faculté à participer à la gestion de l'espace thématique et intersubjectif. Le présent chapitre s'intéresse à l'hypothèse selon laquelle la conduite de la courbe mélodique est révélatrice de l'état de la relation inter-sujets. Nous souhaitons montrer que la gestion de l'espace intersubjectif relève d'une problématique de hiérarchisation à divers titres.

7.1 Gestion tripartite de l'espace intersubjectif

Les linguistes énonciativistes distinguent différentes postures permettant de rendre compte de l'évolution de la relation intersubjective au cours d'une conversation. Celles-ci concernent la place relative accordée par l'énonciateur au contenu de son discours, aux représentations mentales qu'il véhicule, et à celles qu'il prête à son co-énonciateur.

7.1.1 Quelques travaux antérieurs

Dans sa réflexion sur la deixis, L. Danon-Boileau mentionne deux types de *deixis* qui correspondent à différentes positions susceptibles d'être occupées par l'énonciateur vis-à-vis du co-énonciateur. Les notions de consensualité et d'anticipation de la pensée de l'autre y occupent une place prépondérante, comme en témoignent les remarques suivantes, formulées pour le français :

> [...] à côté de cette *deixis consensuelle*, dans laquelle énonciateur et coénonciateur « fusionnent » leurs regards, il en existe une autre dans laquelle au contraire l'énonciateur marque ses distances par rapport au coénonciateur — ou plutôt par rapport au « tout indistinct » qu'il formait avec l'autre dans la deixis consensuelle. C'est la *deixis de rupture*, laquelle constitue l'énonciateur en entité individuée.[1]

Le même type de division est développé et appliqué à l'anglais par M. de Cola-Sekali. Dans sa thèse, elle distingue les quatre valeurs suivantes pour la relation intersujets, qui correspondent aux quatre valeurs de repérage déterminées par A. Culioli :[2]

- So se situe en concordance avec S1 (identification) : So = S1.
- So ne se situe pas dans une altérité des sujets. Il est séparable de S1, mais non repéré par rapport à lui.
- So se situe en discordance avec S1 (différenciation) : So $\neq$ S1.
- So pose la prise en charge de façon décrochée : S*. Cette position servira par exemple à impliquer le co-énonciateur dans une prise en charge forcée des énoncés.[3]

[1] DANON-BOILEAU, Laurent (1992). Ce que « ça » veut dire : les enseignements de l'observation clinique. in Morel, Mary-Annick & Danon-Boileau, Laurent (éds.), *La deixis* : 415-427. Paris : PUF. (p.417).

[2] Voir par exemple CULIOLI, Antoine (1973). Sur quelques contradictions en linguistique. *Communications* 20 : 83-91.

[3] DE COLA-SEKALI, Martine (1991a). *Connexion inter-énoncés et structuration des relations temporelles et argumentatives en anglais contemporain. Une étude énonciative des connecteurs polyvalents SINCE et FOR.* Thèse de Doctorat de 3e cycle. Université Paris III – Sorbonne Nouvelle. (p.276).

Ces deux propositions se rejoignent dans le sens où elles considèrent que les relations intersubjectives sont essentiellement déterminées en fonction du repère que constituent les coordonnées énonciatives. Celles-ci permettent en effet le repérage des relations prédicatives par rapport à l'énonciateur ainsi que celui de l'énonciateur par rapport au co-énonciateur.

7.1.2 Division retenue dans ce travail

Notre travail se situe dans la continuité des deux précédents. Nous souscrivons à l'idée d'une division de l'espace intersubjectif au sein duquel nous proposons trois positions pour l'énonciateur.

Revenons au terme de « deixis de rupture » utilisé par L. Danon-Boileau dans la citation reportée plus haut. Dans le débat dont il est fait état à la suite de cet article, les remarques de M. Maillard font revenir l'auteur vers une autre dénomination : « deixis de différenciation ou d'individuation ».[4]

Il nous semble en effet que le terme de rupture soit plus approprié pour désigner non pas un état mais le mouvement qui fait sortir de l'**état neutre de consensus établi**. Pour nous, une telle rupture peut entraîner la relation entre énonciateur et co-énonciateur à l'un ou l'autre des pôles de l'axe dont le consensus établi constitue le centre. Dans l'état consensuel, énonciateur et co-énonciateur sont en concordance, pour reprendre les mots de M. de Cola-Sekali. Pour utiliser une métaphore corporelle rappelant la problématique de la deixis qui est celle de L. Danon-Boileau dans son article, nous dirons qu'ils sont côte à côte et qu'ils regardent dans la même direction.[5]

Lorsque l'énonciateur anticipe et prend en compte la pensée de l'autre, ce mouvement par lequel on sort de l'état consensuel aboutit à un **état de discordance ou de forçage du consensus**. L'énonciateur admet ou anticipe une divergence de

[4]DANON-BOILEAU, Laurent (1992). *art. cit.* (pp.424-425).

[5]Cette métaphore est utilisée dans MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation, L'exemple du français.* Gap : Ophrys. (p.12).

vues avec le co-énonciateur. Dans le cas du forçage du consensus, il essaie d'amener l'autre sur son terrain et d'établir ainsi un nouveau consensus. Dans le cas de la discordance, il décrète que son contenu de pensée et celui qu'il prête à l'autre sont incompatibles et irréconciliables. Enonciateur et co-énonciateur ne sont donc plus côte à côte mais face à face. Nous souhaitons conserver la catégorie unique déterminée par L. Danon-Boileau pour les deux manifestations de cet état non consensuel. Dissociées dans l'analyse de M. de Cola-Sekali, elles méritent selon nous pour l'instant d'être traitées ensemble.

Enfin, à l'autre pôle de l'axe évoqué plus haut se situe la deuxième valeur décrite par M. de Cola-Sekali. C'est en effet vers un **état d'égocentrage** que se fait la rupture lorsque anticipation et prise en compte de la pensée de l'autre sont absents : l'énonciateur est comme seul.

Différentes oppositions se dessinent entre les trois postures que nous venons de décrire brièvement. La métaphore corporelle de la position relative des parties prenantes du discours nous sert à distinguer trois attitudes énonciatives essentielles et permet d'illustrer la division ternaire de l'espace des relations inter-sujets. Cependant, celles-ci s'inscrivent également dans un jeu d'oppositions binaires. Les deux premières postures (consensus établi, et discordance/forçage du consensus), dans lesquelles la pensée de l'autre fait l'objet d'une anticipation et d'une prise en compte, correspondent à une attitude dialogale. La dernière, marquée par l'absence de prise en compte de la pensée de l'autre, relève davantage du monologue. Le Tableau 7.1 résume cette division tripartite de l'espace intersubjectif et les oppositions sur lesquelles elle est construite.

1. FACE À FACE	2. CÔTE À CÔTE	3. SOLITUDE, ISOLEMENT
forçage du consensus discordance	consensus établi	égocentrage
prise en compte de la pensée de l'autre		pas de prise en compte de la pensée de l'autre
attitude dialogale		attitude monologale

TAB. 7.1 – Division de l'espace intersubjectif

7.1.3 De la théorie à l'exemple

Afin de montrer comment les paramètres prosodiques contribuent à la gestion de l'espace tripartite que nous venons de décrire, nous souhaitons procéder à une expérience permettant de confronter notre modèle théorique à la réalité de l'anglais oral spontané tel que représenté dans notre corpus. Faute de pouvoir traiter l'intégralité de notre corpus, nous nous concentrerons sur un extrait d'un peu moins de deux minutes issu du corpus Celia&James, dans lequel James évoque la vie de sa sœur Elspeth, partie enseigner l'anglais en Chine. Dans cet extrait, que nous avons intitulé *Beer*, James est en position d'énonciateur principal. Il détient un savoir que Celia lui demande de partager, comme en témoigne l'énoncé interrogatif qui ouvre cet épisode :

C but Tsing Tao
so where about is that
is it *in*

(Celia&James : Beer)

A partir de la seule transcription orthographique, et de notre connaissance de l'extrait, nous proposons tout d'abord de chercher à identifier des segments (mots ou groupes de mots) susceptibles d'y être les témoins de l'une de ces trois postures co-énonciatives.

7.1.3.1 Consensus établi

La trace la plus évidente de la position consensuelle nous semble être la marque d'écoute *mm* (7.1). R. Gardner[6] analyse cette intervention typique de l'auditeur comme une réaction au discours du locuteur principal qui peut avoir différentes valeurs. Elle peut constituer par exemple une invitation à poursuivre adressée au locuteur, une marque d'assentiment, d'intérêt, ou d'évaluation de ce qui vient d'être dit, ou simplement, et plus fréquemment, de bonne réception. Pour ce chercheur, *mm* est une forme relativement faible manifestant de façon neutre l'activité d'écoute de l'auditeur.

En effet, le locuteur qui dit *mm*, loin de tenter de prendre la parole, indique à son interlocuteur que les représentations mentales mises en jeu à ce moment de l'échange sont consensuelles. Notre extrait compte sept occurrences, toutes prononcées par Celia. Dans l'exemple suivant, celle-ci invite James à développer sur les différents plats que sa sœur est amenée à goûter :

J I think the food is pretty:: interesting {0,60}
hm {0,50} actually I I'm sure my er {0,20} Elspeth is really into food and she's er very good at describing meals in <rires> great detail I haven't *heard any*

C [ex.7.1] ***mm***

J <h> {0,60} any of those details ye::t but {1,20}

James se montre ici incapable de fournir une longue description sur la gastronomie chinoise. Il ne répond que partiellement à la question de Celia. Celle-ci s'en contente pourtant : elle accuse réception de sa réponse. Se plaçant ainsi en auditrice, elle le conforte dans sa position d'énonciateur principal et l'invite ainsi à poursuivre. Elle manifeste la concordance des contenus de pensée véhiculés, et son adhésion au projet de James.

Dans notre extrait, une seule occurrence de *mm*, relayée par le ligateur *so*, mène

[6]GARDNER, Rod (2002). *When Listeners Talk : Response Tokens and Listener Stance*. Amsterdam/Philadelphia : John Benjamins.

à une véritable prise de parole. Nous analyserons cet exemple de façon plus détaillée dans la suite de ce chapitre. Les six autres fois, cette intervention forme un tout avec le discours de l'énonciateur, semblable au « tout indistinct » de L. Danon-Boileau. Celia se positionne de façon non conflictuelle dans le consensus établi entre les deux parties prenantes de l'échange. Notons que toutes les occurrences de *mm* sont prononcées en superposition par rapport au discours de James, c'est-à-dire pendant qu'il parle.

7.1.3.2 Discordance et forçage du consensus

L'état de la relation intersubjective que nous désignons par les termes de discordance et forçage du consensus est représenté par divers segments de notre extrait. Considérons l'exemple suivant où James, faute de pouvoir situer Tsing Tao, la ville où séjourne sa sœur, évoque son industrie agroalimentaire :

J it's sort of famous for beer well <h> {0,60} if ye if ye {0,25} <rires> *if ye {0,20}*
C <h> *Chinese beer*
J yeah Tsing Tao well it's one of the [ex.7.2] ***best known***
C *mm*
J {0,20} Chinese beers <h> {0,70} er *y- you get in (?)numerous(?) country*

Le caractère non consensuel qu'atteint la relation ici nous semble être concentré dans le segment *best known* (7.2). James s'inscrit en faux vis-à-vis des représentations mises en jeu. Celles-ci se manifestent, tout d'abord, dès sa première réplique, dans la modalisation répétée de la relation <Tsing Tao – be famous> qui tend à en affaiblir la validation. C'est ce que nous lisons dans la locution adverbiale *sort of*, l'utilisation de *well* (qui relativise la portée de ce qui vient d'être avancé), ainsi que dans la répétition ironique d'un début de proposition conditionnelle. Accompagnée de rires et de pauses (plus d'une seconde au total), la formulation de James n'aboutit pas mais se résout en quelque sorte dans l'intervention ironique de Celia. Avec sa réplique suivante, et notamment le groupe adjectival *best known*, James va donc au-devant de ces représentations qu'il estime erronées. Non seulement il existe de la bière chinoise,

rétablit-il, mais en plus elle est mondialement réputée et fait la célébrité de la ville de Tsing Tao.

Dans l'exemple suivant, Celia soulève le problème du code vestimentaire qu'elle croit occuper une place très importante en Chine. Or la sœur de James et ses collègues sont parfois amenés à participer à des banquets de façon impromptue :

C said you er you they'd have to dress:: {0,25} appropriately for

J <h> w- ba- *she hasn't she hasn't*

C [ex.7.3] ***how does she manage* tha::t**

J I d- well she hasn't mentioned anything like that so

L'énoncé *how does she manage that* témoigne pour nous d'une posture en rupture par rapport au consensus établi, relevant de la discordance. La forme interrogative permet d'interpréter la réplique de Celia comme un appel à l'autre, réel ou feint, pour qu'il comble le déficit sémantique représenté par l'adverbe interrogatif *how*. Sa dimension exclamative témoigne du fait que ce dont il est question dépasse les représentations que Celia en a. La sidération et la proximité de l'indicible caractérisent le passage dans l'état de discordance. L'écart entre les représentations étant irréductible, la construction d'un nouveau champ consensuel vers lequel retourner n'est pas envisageable.

7.1.3.3 Egocentrage

Dans notre extrait, l'échange entre les deux locuteurs est véritablement dialogal. Il s'agit bien d'un échange où Celia et James construisent un savoir partagé et confrontent des contenus de pensée. L'identification de segments marquant un égocentrage y est donc plus délicate.

Nous sommes en mesure de proposer un seul exemple nous semblant pouvoir relever de la posture égocentrée. Celui-ci fait immédiatement suite au passage précédemment cité :

J I d- well she hasn't mentioned anything like that so

C *cos it*

J *I guess* it's not a problem

C [ex.7.4] **it's not a problem**

Nous nous intéressons à la dernière réplique de Celia. La locutrice répète ici comme pour elle-même la réponse de James à son exclamation précédente. Après un mouvement vers un état de discordance, cet énoncé signe le retour vers le consensus ou même, proposons-nous, vers un stade de la relation intersubjective où ce problème qui n'existe pas est retiré du champ de l'échange. La sœur de James n'en ayant jamais parlé, James suppose que l'application d'un quelconque code vestimentaire ne pose pas problème. Celia reconnaît ici que la discordance née de l'écart de représentations évoqué plus haut n'avait pas lieu d'être. Puisqu'il n'y a rien à en dire, l'incident est clos.

Afin d'étayer ces intuitions concernant l'évolution de la relation intersubjective dans notre extrait, nous proposons d'examiner les modulations de la ligne mélodique. Dans quelle mesure celles-ci reflètent-elles la gestion de l'espace co-énonciatif ?

7.2 Trois plages mélodiques concordantes ?

L'hypothèse que nous souhaitons examiner à présent fait correspondre les trois postures de la co-énonciation à trois bandes de hauteur fréquentielle déterminées statistiquement pour chaque locuteur : nous proposons ainsi d'associer la bande moyenne avec le consensus, la bande haute avec le forçage de consensus et la discordance, et la bande basse avec l'égocentrage. Elle se fonde sur notre hypothèse de départ selon laquelle la mélodie est, dans une certaine mesure, dotée d'une valeur propre.[7] Elle s'appuie également sur le code de l'effort de C. Gussenhoven, qui associe un pic mélodique à une marque d'emphase ou de surprise, c'est-à-dire, pour nous, à une posture intersubjective particulière.[8] Le recours aux logiciels d'analyse prosodique nous

[7] Voir p.135.

[8] CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002*

est ici indispensable et nous permet de confronter notre première lecture de l'extrait avec des données chiffrées.

7.2.1 Calcul des bandes

Les valeurs de F_0 de chacun des deux locuteurs sont analysées à partir des données fournies par les fichiers .dat du logiciel WinPitch Classic. Nous pouvons ainsi obtenir les valeurs des paramètres intensifs et mélodiques à un temps donné et pour une séquence de longueur donnée (avec une précision possible à la milliseconde). Pour cette expérience, nous avons choisi un niveau de précision important afin d'obtenir une valeur de F_0 toutes les 20 ms : WinPitch fournit ainsi cinquante valeurs de F_0 par seconde. Au total, nous disposons donc potentiellement d'environ 5 500 valeurs de F_0 par locuteur pour un peu moins de deux minutes de parole.

Les deux locuteurs ne parlant pas aussi longtemps l'un que l'autre, le nombre d'échantillons différents de zéro sur lesquels s'effectuent nos calculs n'est pas du même ordre. Pour le locuteur principal, James, nous disposons de 2 382 valeurs de F_0, contre 482 pour l'auditrice, Celia. Dans le fichier de James (désormais [James]), les valeurs s'étalent de 77 à 455 Hz. Après examen du fichier, il apparaît que les valeurs les plus élevées (supérieures à 300 Hz), rares, correspondent à des sifflantes ou à des occlusives ou encore à des passages où les deux locuteurs parlent en même temps (il est donc probable que ces valeurs inhabituelles pour James soient en fait attribuables à Celia). Nous avons de ce fait décidé de ne pas prendre en compte, dans notre analyse du fichier de James, les valeurs situées au-delà de 300 Hz. Les valeurs de F_0 du fichier de Celia ([Celia]) s'échelonnent de 89 à 478 Hz. Les valeurs situées en-dessous de 100 Hz correspondent pour certaines à des moments où la voix de James se superpose à celle de Celia. D'autre part, elles sont en nombre extrêmement restreint (moins de 1%). Nous avons donc décidé de les écarter de nos calculs et de n'étudier que la bande

conference : 215-218. (p.215).

située pour Celia entre 100 et 500 Hz (donc 478 Hz).

7.2.1.1 Hypothèse de la loi normale

Afin de déterminer l'étendue des trois plages fréquentielles qui nous intéressent, nous procédons au calcul de la moyenne des valeurs de F_0, de la médiane et de l'écart type. Le Tableau 7.2 présente les valeurs obtenues pour les deux fichiers. Moyenne et médiane sont exprimées en Hertz, tandis que l'écart type a été converti en quarts de ton (la moyenne en constituant la valeur repère, c'est-à-dire correspondant à 0 écart type).

	[CELIA]	[JAMES]
nombre d'échantillons	479	2 340
moyenne (Hz)	198	135
médiane (Hz)	191	126
écart type (1/4 de ton)	10	6

TAB. 7.2 – Moyenne, médiane et écart type de F_0 dans les deux fichiers

Ce calcul est préalable à la vérification de la loi normale (ou loi de Laplace-Gauss), dont l'hypothèse est tentante ici. Une répartition de valeurs suivant la loi normale se traduit graphiquement par une courbe de Gauss (en cloche).[9] C. Muller précise :

> Tous les points de la courbe sont déterminés par deux paramètres : la moyenne et l'écart type ; la moyenne détermine son axe de symétrie, donc son point d'ordonnée maximum ; l'écart type, qui est la distance, sur l'axe des abscisses, de la moyenne au point d'inflexion, détermine la pente de la courbe à chacun de ses points.[10]

[9] Voir les caractéristiques de la courbe de Gauss, rappelées par MULLER, Charles (1992). *Initiation aux méthodes de la statistique linguistique.* Paris : Honoré Champion. (p.73) :

> Le tracé de cette courbe est symétrique par rapport à un axe vertical qui passe par son point le plus haut (son ordonnée maximale). L'abscisse de ce point est par conséquent le mode de la variable (sa valeur la plus fréquente ou la plus probable) ; elle marque aussi sa médiane et sa moyenne [...].
> De part et d'autre de ce maximum, les ordonnées décroissent d'abord lentement, puis plus rapidement [...] ; la décroissance ralentit, et la courbe se rapproche sensiblement (asymptotiquement) du 0 matérialisé par l'axe horizontal.

[10] *ibid.* (p.73).

Autour de la moyenne, l'écart type pondéré permet donc de définir une bande neutre, non marquée, au-delà et en-deçà de laquelle se situeraient les bandes marquées, positivement et négativement modulées.[11]

D'après la théorie statistique présentée par C. Muller,[12] trois calculs relatifs à l'aire sous la courbe permettent d'apprécier la conformité d'une répartition de valeurs avec la loi normale : sélectionner l'intervalle compris entre plus ou moins 1 écart type (σ) autour de la moyenne (m) isole 68% des individus d'une population normale, plus ou moins 2 σ correspondent à 95% des individus, et plus ou moins 3 σ à 99%.

Le Tableau 7.3 montre que les valeurs de F_0 dans [Celia], entre 100 et 500 Hz, ont une répartition relativement conforme à la loi normale. En revanche, comme l'indique le Tableau 7.4, elle ne s'applique pas de façon aussi claire à [James] (entre 70 et 300 Hz). Ces tendances se retrouvent dans l'allure de la courbe des valeurs des deux fichiers présentée aux Figures 7.1 et 7.2.[13] Moyennant quelques ajustements, la distribution des valeurs dans [Celia] semble bien s'approcher d'une courbe de Gauss : le mode de la variable (c'est-à-dire les valeurs les plus courantes) correspond approximativement à la moyenne ($x = 0$) et à la médiane. Par ailleurs, le nombre de valeurs décline lorsque x tend vers les extrêmes.[14] En revanche, celle de [James] offre moins de symétrie par rapport à la moyenne. Sa distribution est étalée à droite.

Pondération de l'écart type		Aire sous la courbe	
1 σ	$-10 < m < 10$	69% de distribution	($N = 334$)
2 σ	$-20 < m < 20$	95% de distribution	($N = 458$)
3 σ	$-30 < m < 30$	99% de distribution	($N = 478$)

Tab. 7.3 – Test de conformité à la loi normale appliqué à [Celia]

[11] Au sein de ces deux bandes s'applique également la loi normale.

[12] MULLER, Charles (1992). *op. cit.* (p.74).

[13] Dans l'histogramme concernant [Celia], chaque barre représente une tranche de 10 Hz. L'ambitus plus restreint de James nous a conduite à adopter un delta égal à 5 Hz.

[14] Les sommets relatifs atteints dans les deux bandes extrêmes suggèrent une possible distribution multimodale où la loi normale s'appliquerait également.

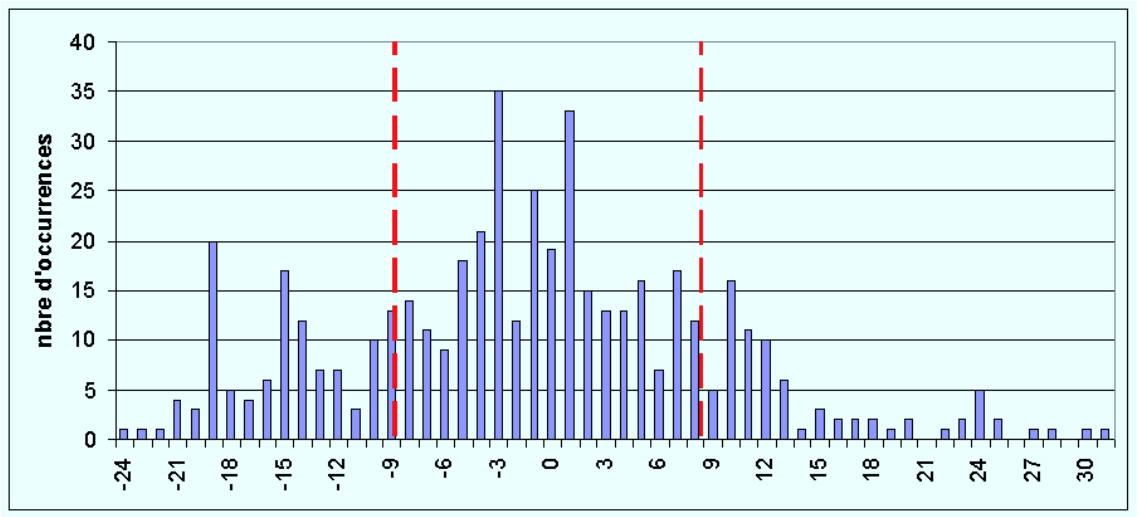

FIG. 7.1 – Distribution des valeurs de F_0 dans [Celia] en quarts de ton répartis autour de la moyenne

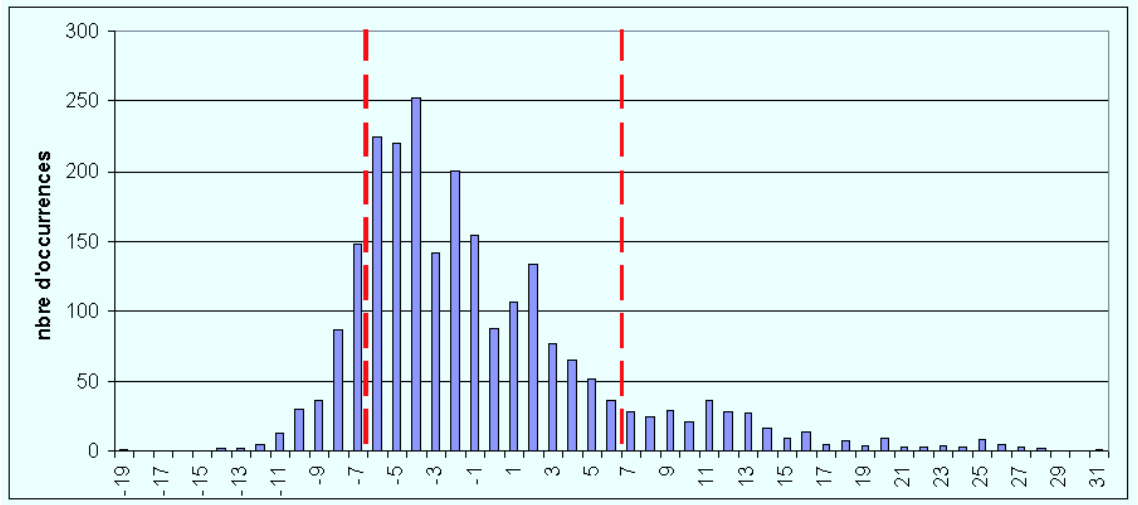

FIG. 7.2 – Distribution des valeurs de F_0 dans [James] en quarts de ton répartis autour de la moyenne

Pondération de l'écart type		Aire sous la courbe	
1 σ	$-6 < m < 6$	74% de distribution	($N = 1\ 751$)
2 σ	$-12 < m < 12$	94% de distribution	($N = 2\ 236$)
3 σ	$-18 < m < 18$	98% de distribution	($N = 2\ 318$)

TAB. 7.4 – Test de conformité à la loi normale appliqué à [James]

Doit-on pour autant exclure la pertinence de la loi normale pour expliquer la répartition des valeurs de F_0 dans notre extrait ? Il nous semble que l'hypothèse peut être maintenue pour plusieurs raisons. La raison essentielle que nous invoquons est que loi normale décrit une population idéale. Comme le rappelle J. Véronis, les réalisations se conforment plus ou moins imparfaitement au modèle.

> La DISTRIBUTION NORMALE est une DISTRIBUTION THÉORIQUE, en ce sens qu'elle est une idéalisation mathématique qui ne se rencontre jamais dans la nature.[15]

En outre, la petite taille des échantillons considérés fait s'élever la probabilité de s'éloigner de la loi normale. Un extrait de moins de deux minutes ne peut pas représenter la totalité de notre corpus, et encore moins la totalité de l'échange linguistique en général. Si c'est bien sur l'échantillon le plus petit que nos résultats sont les plus satisfaisants, on ne peut attendre d'un extrait comme celui que nous avons étudié qu'il présente toutes les caractéristiques d'une population normale. Certes, le test de l'écart type est insuffisant, mais il semble confirmé par l'écart entre la moyenne et la médiane : l'écart en quarts de ton est très proche dans les deux cas (1 pour [Celia], 2 pour [James]).

En dépit des objections formulées à son encontre, nous souhaitons conserver cette hypothèse et la mettre au service de notre intuition linguistique. Elle nous permet en

[15] VERONIS, Jean (2002). INF Z16 – Informatique et statistique I. Chapitre 6. Forme des distributions. Cours disponible en ligne, consulté le 12 janvier 2004 :
http://www.up.univ-mrs.fr/veronis/cours/INFZ16/index.html
http://www.up.univ-mrs.fr/veronis/cours/INFZ16/ch6.html

effet de diviser le registre utilisé dans notre extrait par les locuteurs en trois plages mélodiques correspondant aux trois postures co-énonciatives déterminées plus haut.

7.2.1.2 Bande neutre non modulée

Dans notre hypothèse, la bande neutre détermine la portion de discours que l'on considère comme témoignant d'un consensus établi. Son étendue, et donc le coefficient de pondération à appliquer à l'écart type, doit être le fruit d'un choix fondé sur l'appréciation du type d'interaction en jeu. Dans l'extrait analysé, James détient un savoir que Celia lui demande de partager. Aucune dimension agonale ou monologale n'est prévisible, mais elles ne sont pas pour autant à écarter totalement. Nous avons choisi ce coefficient de sorte qu'environ 60% des valeurs figurent dans cette bande neutre, c'est-à-dire que 60% de la conversation soit interprétable comme consensuelle.

	Pondération	Bande neutre (Hz)	Aire sous la courbe
[Celia]	0,8 σ	$155 < m < 250$	60%
[James]	0,65 σ	$115 < m < 155$	63%

Tab. 7.5 – Bande neutre pour les deux fichiers

Pour [Celia], 60% des valeurs sont circonscrites autour de la moyenne lorsque l'on sélectionne 0,8 écart type, soit une bande neutre située entre 155 et 250 Hz. Pour [James], 63% des valeurs se situent entre 115 et 155 Hz, grâce à une pondération de 0,65. Ces chiffres sont résumés dans le Tableau 7.5 et trouvent une représentation graphique entre les deux traits verticaux des Figures 7.1 et 7.2.

7.2.2 Confrontation de l'hypothèse statistique et de l'intuition linguistique

7.2.2.1 De l'exemple à la théorie

Revenons aux segments que nous avons isolés plus haut comme étant les témoins potentiels de chacune des trois postures co-énonciatives, et vérifions dans quelle mesure la plage mélodique où ils se situent correspond à notre intuition linguistique. Le résultat est largement positif.

Consensus établi Les sept *mm* sont tous prononcés à la même hauteur, dans la bande neutre que l'on a déterminée. Leur caractéristique consensuelle est présente encore ici, où apparaît la seule occurrence de *mm* suivie d'une conservation de la parole :

C [ex.7.5] *mm so what's the food like*

Deux phénomènes concourent ici, après la marque d'écoute, à la sortie du consensus établi et à la proposition d'un nouveau thème commun. Tout d'abord, la question s'interprète dans notre exemple comme un appel à l'autre, afin de l'amener sur un autre terrain. D'autre part, le ligateur *so* opère une clôture thématique et permet une réinitialisation par rapport aux repères précédents, relatifs à la bière. *So* constitue ici une charnière où s'articulent deux états de la relation intersubjective. Notre exemple est conforme à la vision de J. Szlamowicz :

> *So*, bien qu'il soit ligateur, introduit aussi une forme de rupture parce qu'il instaure une position énonciative qualitativement différente de l'énoncé précédent [...]. ***So* est donc au carrefour d'un apport discursif et d'une reprise de cet apport pour construire une suite.**[16]

[16]SZLAMOWICZ, Jean (2001). *Contribution à une approche intonative et énonciative du rôle des ligateurs dans la construction du discours en anglais oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (p.185).

Notons que Celia introduit un nouveau thème, *food*, dont la mention s'accompagne d'une légère incursion de la courbe dans la zone modulée supérieure qui en fait également le centre d'enjeux énonciatifs. L'énoncé suivant est néanmoins prononcé dans la bande médiane. Celia cherche en effet simplement confirmation d'un fait qu'elle tient pour acquis : la nourriture chinoise est excellente.

Nos exemples corroborent donc notre intuition linguistique, mais également notre hypothèse statistique du possible calcul d'une plage de fréquences non marquée correspondant à un état consensuel de la relation intersubjective.

Discordance et forçage du consensus Comme prévu, les segments (7.2) et (7.3) auxquels nous avons assigné des valeurs de forçage du consensus et de discordance sont prononcés en plage haute.[17] Se trouve ainsi confirmée la viabilité de notre intuition linguistique et de notre hypothèse d'une bande fréquentielle élevée correspondant à une rupture du consensus établi vers un état discordant dans lequel un retour vers un nouveau consensus est ou non envisagé.

Egocentrage Le segment (7.4), relevé comme potentiellement égocentré, offre cependant davantage de résistance :

C [ex.7.4] it's not a problem

La réplique de Celia se situe en effet en plage moyenne, et non pas en plage basse comme attendu. Trois possibilités sont à envisager pour expliquer cet échec. Premièrement, notre hypothèse des trois plages mélodiques est erronée (son principe, son mode de calcul, et/ou son résultat). Deuxièmement, notre interprétation de cet énoncé en termes co-énonciatifs est erronée. Troisièmement, notre hypothèse mélodique doit également prendre en compte d'autres paramètres intonatifs.

Cette dernière possibilité nous semble être à retenir. L'examen de la courbe d'intensité réhabilite en effet notre interprétation de l'état de la relation intersubjective.

[17] Voir le contexte restreint de ces deux exemples (p.323).

Le faible niveau d'intensité relevé ici permet en effet à Celia de retirer ce qu'elle dit du champ du discutable, sur le mode d'une incise ou d'un aparté. Elle se place au-delà de tout consensus, dans la mesure où la représentation de la pensée de l'autre n'est plus centrale voire plus envisagée.

Ainsi, à défaut de remettre en cause la validité de nos hypothèses de départ (linguistiques ou statistiques), cet exemple souligne certaines limites de l'hypothèse statistique. F_0 a bien un rôle prépondérant dans la gestion de la relation intersubjective, mais cette dernière n'est pas réductible à la seule conduite de la ligne mélodique.

7.2.2.2 Redondance et indépendance des marques

L'hypothèse des trois bandes confirme le fonctionnement de F_0 dans la gestion de la relation intersubjective. La sortie de la plage moyenne vers la plage haute correspond à des moments où l'investissement modal de l'énonciateur est palpable.

Redondance On constate une redondance des marques lorsque cet investissement est déjà marqué dans le segmental. Ainsi se retrouvent en plage haute certaines interrogatives et exclamatives :

C [ex.7.6] so where about is that

C [ex.7.7] but with no warning

C'est également le cas de structures dans lesquelles est opéré un travail sur le lien prédicatif, de nature emphatique ou épistémique :

J [ex.7.8] they do keep being taken to banquets[18]

C [ex.7.9] they're always being taken off to places

C [ex.7.10] they often don't really know what's going to happen

Le passage par la subjectivité de l'énonciateur est encore souligné par l'emploi des formes suivantes, affectées d'une modulation mélodique : *I think, I'm sure, I guess* où

[18]Les courbes de cet exemple et la matérialisation des trois bandes de fréquence sont présentées à la Figure 7.3.

le verbe de jugement est conjugué à la première personne, ou encore, *they seem* où le sujet du verbe de perception, obtenu par montée, voit la relation prédicative qu'il commande affectée de la modalité induite par le verbe *seem*. Dans tous ces exemples, James s'en remet à son interlocutrice en signalant que ce qu'il avance est sujet à caution. Il propose des points de consensus sans les imposer. Le même emploi de F_0 accompagnant des marques de point de vue se retrouve chez Celia :

C [ex.7.11] ***I suppose** they can just {0,60}*

J *sometimes they're*

C *talk* off the top of *their head*

L'accordage des représentations et la recherche d'un consensus sont ici rendus visibles à deux niveaux différents, d'où une redondance des marques : au niveau segmental par les formes utilisées et au niveau prosodique par la conduite de la ligne mélodique.

Non redondance Il apparaît cependant que segmental et suprasegmental sont la plupart du temps non redondants. Ainsi, l'observation des bandes de fréquence permet, par un mouvement inverse, de donner une interprétation co-énonciative à des éléments de discours moins marqués naturellement sur le plan intersubjectif. Le participe *taken off* apparaît trois fois littéralement, et une fois sous la forme tronquée *taken*. Forme répétée, il fait partie de l'information donnée et donc normalement de l'attention partagée, consensuelle. Or le traitement mélodique souligne qu'il est trois fois de suite investi par l'énonciateur (s'appuyant également deux fois, en (7.8) et (7.9), sur des marques segmentales) pour démontrer l'infantilisation de la sœur de James et de ses collègues étrangers. Le groupe verbal auquel il appartient apparaît clairement dans la zone modulée supérieure, comme nous pouvons le voir à la Figure 7.3. Seule sa dernière occurrence figure en plage neutre, le plaçant ainsi comme faisant partie de ce qui, dans l'échange, est considéré comme acquis par James comme pour Celia. On voit donc ici que le fonctionnement de F_0 est décorrélé de l'opposition

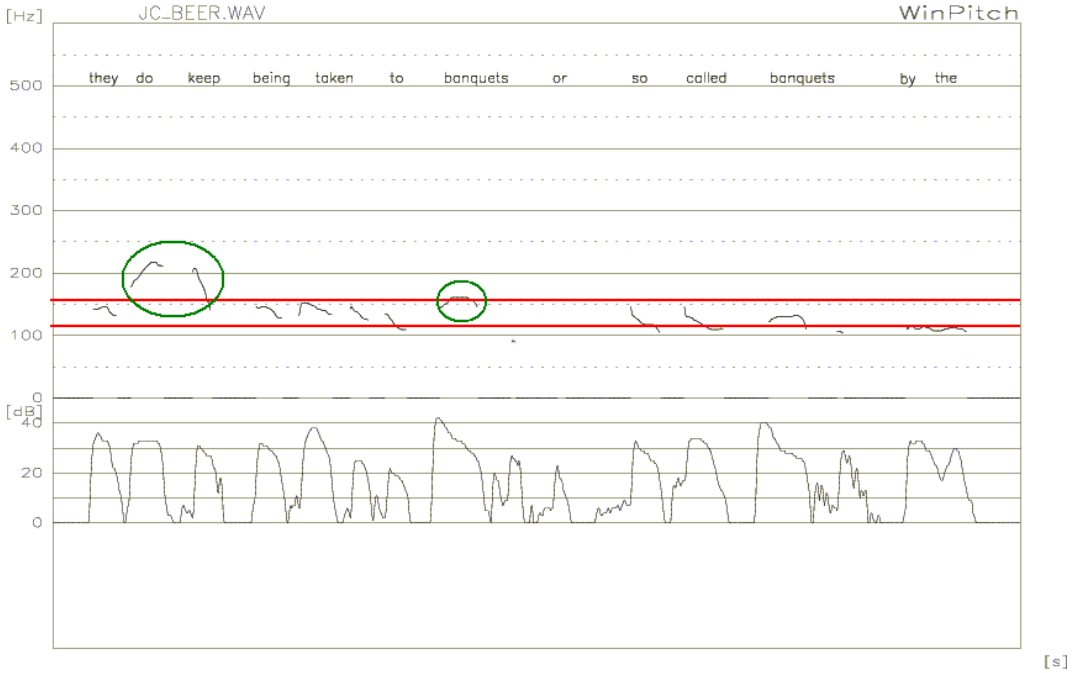

FIG. 7.3 – Courbe de l'exemple 7.8 avec matérialisation des trois plages mélodiques

entre information nouvelle et information donnée, mais renseigne sur l'investissement du sujet parlant dans le discours et de sa prise en compte du contenu de pensée qu'il prête à l'autre.

Ce fonctionnement non redondant de la prosodie par rapport aux marqueurs segmentaux permet précisément aux paramètres intonatifs de mettre en valeur certains segments par rapport à d'autres, indépendamment de leur place dans l'ordre linéaire de l'énoncé. C'est pourquoi nous parlons de « délinéarisation » de la syntaxe grâce à l'intonation, qui permet de tisser des réseaux par-dessus le linéaire de surface. Ce chapitre traite essentiellement des réseaux de segments dont l'investissement sur le plan intersubjectif est reflété par le traitement mélodique. L'intensité, quant à elle, tend à souligner des éléments importants sur le plan thématique et contribue elle aussi à ce phénomène de délinéarisation.

7.3 Limites et perspectives

L'application du modèle statistique des bandes correspondant aux trois positions de la théorie de la co-énonciation permet d'apporter un éclairage différent et complémentaire par rapport au seul mode d'analyse linguistique à l'oreille et à l'intuition. Cependant, elle n'est pas exempte d'importantes limites.

7.3.1 Limites imputables aux logiciels

Les erreurs des logiciels, outils indispensables à notre étude, sont susceptibles de fausser les chiffres à partir desquels nous fondons nos calculs statistiques. Nous ne nous attarderons pas longuement sur ces problèmes déjà évoqués plus haut.[19] Rappelons simplement que plosives et sifflantes font brutalement monter les valeurs de F_0. Réduire comme nous l'avons fait l'étendue des valeurs prises en compte permet de diminuer l'impact de ces erreurs mais ne l'annule pas. Les chevauchements de parole sont également source d'erreurs dans la mesure où les valeurs de F_0 de l'un des locuteurs tendent à venir s'inscrire dans le tableau de valeurs de l'autre. Les rires, toux et autres perturbations du signal sonore auxquelles le logiciel assigne une valeur de F_0 sont également susceptibles de fausser les données, que notre vigilance à l'examen des fichiers de valeurs ne peut pas toujours filtrer.

En outre, F_0 étant calculée à l'aide d'un algorithme, les résultats de nos calculs sont relatifs à chaque formule. Comme nous l'avons montré à l'occasion d'un récent colloque,[20] des variations sont donc possibles entre logiciels sur le même extrait sonore.

Les limites imputables aux logiciels nous semblent néanmoins négligeables par rapport à celles que l'on peut formuler à l'encontre de notre méthode.

[19] Voir 3.2.3, p.184.

[20] PASSOT, Frédérique (2003). Statistiques suprasegmentales et hypothèses linguistiques : problèmes et perspectives. *INTO-01, Intonation, Notation et Transcription de l'Oral.* Université de Rouen.

7.3.2 Limites imputables à la méthodologie adoptée

7.3.2.1 Phénomènes perceptifs à considérer

L'expérience menée ici ignore certains phénomènes importants relatifs au fonctionnement des paramètres intonatifs et de leur perception. Si nos calculs en quarts de ton tiennent compte de la perception logarithmique de F_0, ils ne reconnaissent que partiellement l'importance de la relativité des niveaux mélodiques perçus au sein d'un ambitus susceptible de varier. Nos hypothèses se fondent sur des bandes établies de façon absolue. Certes, l'échantillon de parole concerné est court : il dure moins de deux minutes. Il possède en outre une unité thématique et interactionnelle : Celia questionne James sur la vie de sa sœur en Chine. Nos hypothèses permettent donc de capturer les éventuelles particularités de cet extrait par rapport à un échange bien plus long dont il est issu. Pourtant, elles ne peuvent rendre compte des variations d'ambitus et du contexte mélodique immédiat dans lequel une hauteur prend son sens.

Notre modèle ne tient pas non plus compte du code de la production de C. Gussenhoven auquel nous avons fait allusion plus haut,[21] et en particulier de la ligne de déclinaison qui en découle.[22] Or certains passages de notre extrait tendent à montrer qu'une rupture de la ligne de déclinaison, au sein d'un même énoncé, investit celui-ci d'une résonance intersubjective particulière. La seule prise en compte de la plage mélodique ne permet qu'imparfaitement de comprendre le statut de l'énoncé (7.12) :

C [ex.7.12] said you er you they'd have to dress:: {0,25} appropriately for

Comme l'indique la courbe reproduite à la Figure 7.4, le départ de l'énoncé se situe en plage haute. Il constitue une demande de consensualité autour de quelque chose dont Celia a entendu parler mais dont elle n'endosse que partiellement la res-

[21] Voir Section 2.2.2, p.135.

[22] CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002 conference* : 215-218. (p.215).

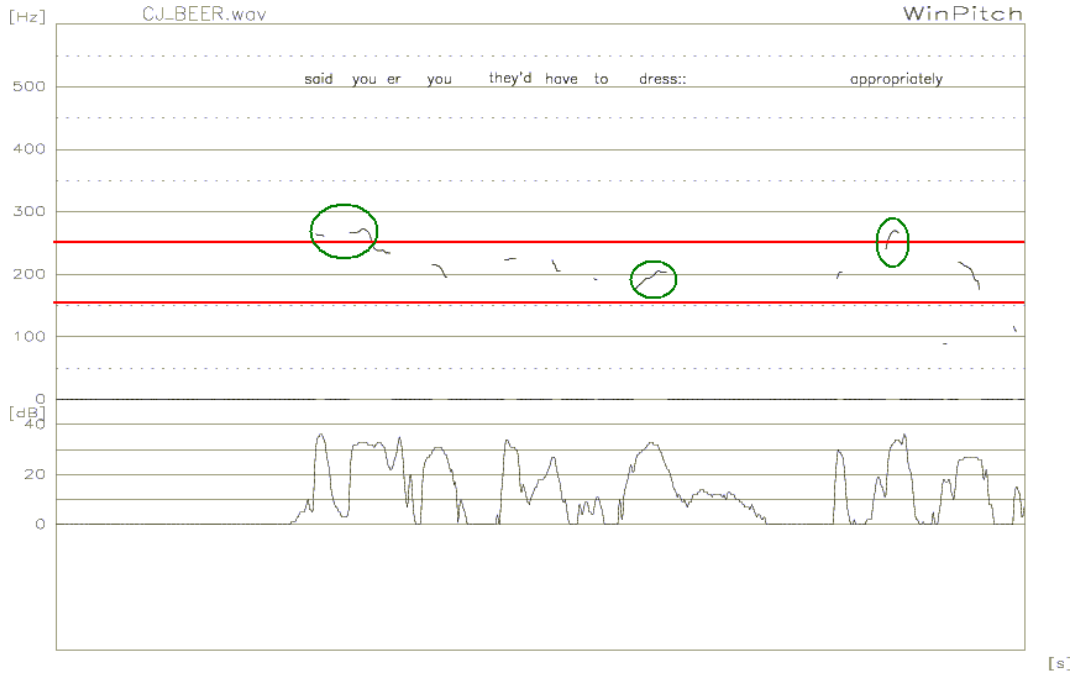

FIG. 7.4 – Courbe de l'exemple 7.12 avec matérialisation des trois plages mélodiques

ponsabilité (*said*). Ce qui fait question ici et qui est proposé comme nouveau thème commun, c'est l'existence d'un code vestimentaire strict en Chine. Or, le verbe *dress*, qui clôt temporairement l'énoncé, se situe en plage neutre. Notre hypothèse prédit ici une interprétation consensuelle. Souligné par l'allongement, le verbe est néanmoins prononcé avec une intonation montante, en forme d'appel à l'autre. De même que l'ensemble de l'énoncé où il figure, ce segment nous semble davantage être le témoin d'un état de la relation intersubjective où la locutrice cherche à établir un nouveau consensus. C'est le sens que nous donnons à la rupture de la ligne de déclinaison initiée par la montée intonative sur *dress*. Celle-ci est par ailleurs confirmée par l'aspect de la courbe sur l'adverbe *appropriately*, dont la syllabe accentuée franchit la frontière symbolique entre plage moyenne et plage haute.

7.3.2.2 Partis pris contestables

Les résultats de cette expérience sont également le fruit de plusieurs partis pris dont on peut contester la validité. Le premier est celui par lequel nous affirmons que la loi normale s'applique à la distribution des valeurs de F_0. Certes, toute réalisation admet certains ajustements par rapport au modèle. Cependant, l'application de cette loi statistique est moins convaincante pour [James] qu'elle ne l'est pour [Celia]. Un autre modèle statistique expliquant la distribution des valeurs de F_0 dans notre extrait aurait pu être adopté.

Le parti pris le plus contestable est sans doute celui de se limiter à la prise en compte de F_0. A bien des égards, les différents paramètres intonatifs ont un fonctionnement décorrélé et chacun est donc susceptible d'une interprétation propre — c'est ainsi que nous assignons une fonction essentiellement modale à F_0. Cependant, un paramètre ne se manifeste jamais seul. Il entre dans un jeu complexe où s'équilibrent les différentes composantes. Comme nous avons pu le montrer avec les exemples (7.4) et (7.12), l'intensité et la durée ont également leur rôle dans le positionnement coénonciatif.

7.3.3 Perspectives

L'hypothèse statistique présentée ici n'a de sens pour nous que dans la mesure où elle est mise au service de notre intuition linguistique. Notre recours à la loi décrivant une population normale n'a pas pour objet de proposer une description statistique précise de la distribution des valeurs de F_0 dans notre extrait. Nous cherchons en revanche à utiliser les outils de la statistique et les modèles qu'elle propose pour tenter d'expliquer des phénomènes qui nous semblent pertinents sur le plan linguistique. Comme nous avons pu le montrer, l'application d'un tel modèle nous permet non seulement d'expliquer et de préciser certaines interprétations que nous avons for-

mulées à propos de l'état de la relation intersubjective. C'est notamment le cas lorsque le plan morpho-syntaxique témoigne déjà d'un investissement intersubjectif particulier. Puisqu'elle place la courbe mélodique d'un segment dans le contexte général de l'extrait analysé, elle permet également d'anticiper la valeur intersubjective portée par le segment en question lorsque celui-ci n'offre pas d'indices particuliers quant à cette valeur.

En ce sens, nos résultats corroborent l'hypothèse du code de l'effort développée par C. Gussenhoven.[23] Produire un pic mélodique requiert un effort particulier de la part d'un locuteur. L'interprétation du segment concerné doit donc tenir compte de cette réalisation marquée. Comme nous avons pu le voir lors de l'exposé de nos hypothèses de départ,[24] le code de l'effort prévoit qu'une telle réalisation s'interprète sur le plan informationnel comme de l'emphase, ou sur le plan des affects comme de la surprise. Or, surprise et emphase nous semblent pouvoir décrire de façon assez satisfaisante les deux versants de la posture co-énonciative discordante et en forçage du consensus.

Au terme de cette expérience, il nous semble que la conjonction des deux hypothèses statistique et linguistique a apporté des résultats satisfaisants. Nous avons pu établir la concordance entre plage mélodique haute et forçage du consensus ou discordance, et entre plage basse et égocentrage. Cependant, cette relation n'est pas réciproque, et ne se vérifie pas de façon systématique entre plage moyenne et consensus. Si la mélodie est effectivement spécialisée dans la gestion de l'espace intersubjectif, elle ne fonctionne pas nécessairement de façon redondante par rapport à un segmental déjà porteur d'un investissement énonciatif. En outre, ses modulations ne suffisent pas à caractériser les trois postures de la co-énonciation. Sur le plan intonatif, d'autres paramètres doivent être pris en compte, tels que l'intensité, la durée, et la ligne de déclinaison. Le contexte et la situation ont également une part importante à jouer

[23] *ibid.* (p.215).
[24] Voir Section 2.2.2, p.135.

dans l'évolution de la relation intersubjective, qu'une analyse statistique aveugle telle que proposée ici ne peut quantifier. S'appuyant sur des outils, méthodes et théories toujours plus ambitieux, la tâche du linguiste demeure essentiellement guidée par la recherche du sens.

Chapitre 8

Sur la place des pauses silencieuses dans la construction du discours conversationnel

Ecouter et interpréter le silence constitue une partie non négligeable du travail cognitif engagé par les locuteurs d'une conversation spontanée. Ce chapitre s'intéresse aux pauses silencieuses et à la place qu'elles occupent dans la construction du discours et de l'interaction en anglais. Difficiles à appréhender du fait de leur nature volatile, les pauses silencieuses sont néanmoins signifiantes linguistiquement à différents niveaux que nous nous proposons de déterminer ici.

8.1 Apport des études antérieures

Comme nous avons pu nous en faire l'écho dans la deuxième partie de ce travail, dans l'attention que nous avons portée aux marques segmentales de subordination, la linguistique fait traditionnellement grand cas du rôle des mots dans la construction du discours. Si l'intérêt pour ce qui se passe entre les mots peut se lire dans les travaux sur

la ponctuation, celui-ci ne peut véritablement se développer qu'à partir du moment où une quantification de ce vide silencieux est possible : une étude sur les pauses nécessite l'utilisation d'outils de mesure suffisamment précis, auxquels les chercheurs n'ont pas forcément eu accès avant les années 1950. Les travaux plus récents sur les pauses silencieuses sont néanmoins nombreux et très documentés. Ils constituent des bases précieuses auxquelles nous nous référons.

8.1.1 La pause silencieuse : éléments de définition

Lorsque le silence fait sens, qu'il fait partie intégrale d'une prise de parole, on l'appelle pause silencieuse. Citons ici la définition donnée par M. Candea dans sa thèse :

> [...] une pause silencieuse est une interruption significative de toute émission sonore à l'intérieur d'une prise de parole ; dans la conversation, il existe parfois des pauses silencieuses également entre deux prises de parole de deux locuteurs différents. Une partie des pauses silencieuses qui surviennent dans la parole sont utilisées par les locuteurs pour respirer, mais la régularité morphosyntaxique de la distribution des pauses de respiration montre que les locuteurs choisissent les moments où ils vont respirer et ne subissent pas passivement cette contrainte physiologique [...].[1]

Ce chapitre ne s'intéresse qu'aux silences délimités par la parole d'un même locuteur. Nous ne traitons pas des *switching pauses*, c'est-à-dire des silences situés entre deux tours de parole attribués à deux locuteurs différents. Dans la définition précédente, la notion de contrainte physiologique à laquelle les locuteurs ne sont pas passivement soumis est bien évidemment fondamentale dans notre appréhension de ce phénomène.

[1] CANDEA, Maria (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (p.21).

8.1.1.1 Seuils adoptés par les études antérieures

A partir de quelle durée l'interruption est-elle « significative » ? E. Campione et J. Véronis[2] rappellent les différents seuils utilisés dans des études relativement récentes sur le français : les travaux de F. Grosjean & A. Deschamps[3] (minimum : 0,30 s, maximum : 2 s), ceux de D. Duez[4] (minimum : 0,18 à 0,25 s, pas de seuil maximal), et la thèse de M. Candea[5] (minimum : 0,20 s, maximum : 2 s). E. Campione et J. Véronis refusent d'adopter tout seuil, minimal ou maximal, et démontrent que leur application conduit à des résultats qu'ils jugent trompeurs, en particulier si l'on cherche à comparer la longueur moyenne des pauses dans des types de corpus différents (parole spontanée et lecture, par exemple).

Cependant, nous avons choisi d'adopter le seuil minimal de 0,20 s. Nous souscrivons en effet aux raisons invoquées par les chercheurs critiqués par E. Campione et J. Véronis. Celles-ci sont issues d'une attention portée à la fois aux faits de production et de perception. Le seuil adopté par D. Duez, par exemple, est le fruit d'un calcul opéré pour chaque locuteur tenant compte de la durée moyenne de ses occlusives intervocaliques à laquelle est ajouté quatre fois l'écart type.[6] Selon les locuteurs, elle aboutit donc à une valeur située entre 0,18 et 0,25 s — dont 0,20 s constitue, pensons-nous avec M. Candea, une moyenne satisfaisante. Cette dernière convoque ensuite les travaux de P. Fraisse[7] sur la perception et la production des groupes rythmiques et du tempo. Il ressort des recherches de ce dernier que 0,20 s correspond à la fois

[2]CAMPIONE, Estelle & VERONIS, Jean (2002). A large-scale multilingual study of silent pause duration. *Actes de Speech Prosody 2002* : 199-202. Aix-en-Provence.

[3]GROSJEAN, François & DESCHAMPS, Alain (1972). Analyse des variables temporelles du français spontané. *Phonetica* 26 : 129-156.
GROSJEAN, François & DESCHAMPS, Alain (1973). Analyse des variables temporelles du français spontané II : Comparaison du français oral dans la description avec l'anglais (description) et avec le français (interview radiophonique). *Phonetica* 28 : 191-226.

[4]Voir notamment DUEZ, Danielle (1982). Silent and non-silent pauses in three speech styles. *Language and Speech* 25 : 11-28.

[5]CANDEA, Maria (2000). *op. cit.*

[6]DUEZ, Danielle (1982). *art. cit.* (p.13).

[7]FRAISSE, Paul (1974). *Psychologie du rythme.* Paris : PUF.

approximativement au seuil en-dessous duquel il est difficile de produire mais aussi de percevoir des émissions sonores régulières, ce qui conforte la chercheuse dans son choix d'adopter un tel seuil.

8.1.1.2 Seuils adoptés pour cette étude et premières analyses

Dans le cas précis de notre étude, intéressée par la perception, il nous semble que la précision du seuil à choisir ne doive pas faire l'objet de calculs aussi précis que ceux de D. Duez. Si M. Candea a opté pour une mesure en centisecondes pour rester proche des unités pertinentes à la perception, nous conservons l'unité standard de la seconde mais la précision de nos mesures, à la centiseconde, bénéficie d'un arrondi à la demi-centiseconde la plus proche. Ainsi, une pause de 0,312 s et une pause de 0,284 s seront notées comme longues de 0,30 s, notre relevé admettant une marge certaine. En outre, nous n'avons pas pu procéder à une détection automatique des pauses, dans la mesure où, dans notre corpus conversationnel, les passages silencieux chez l'un des locuteurs peuvent correspondre non pas à une pause mais à une occupation du canal sonore par l'autre. Ainsi, du fait de la longueur d'un tel relevé manuel, seul le corpus Celia&James a fait l'objet d'un relevé systématique des pauses silencieuses d'une longueur minimale de 0,20 s, sans seuil maximal.

Signalons d'autre part que le résultat d'un calcul fondé sur la durée moyenne des occlusives doit également tenir compte, localement, de possibles variations importantes du débit. Dans l'exemple suivant, le signal s'interrompt pendant une durée supérieure au seuil de 0,20 s à l'intérieur du mot *about* :

C <h> {0,75} hm {0,55} she used to have I mean she was abou::t hm {1,15} eight or nine or something like that <h>

(Celia&James : Photo)

Cette interruption, localisée sur l'occlusive /t/, apparaît dans un contexte d'hésitation où le débit de la locutrice se ralentit subitement. Elle s'accompagne d'une marque de

travail de formulation (*hm*), suivie d'une longue pause (supérieure à 1 seconde). Ces données permettent, à notre sens, d'éliminer l'interprétation de cette interruption comme une pause silencieuse. Notre transcription fait donc état ici d'un allongement.

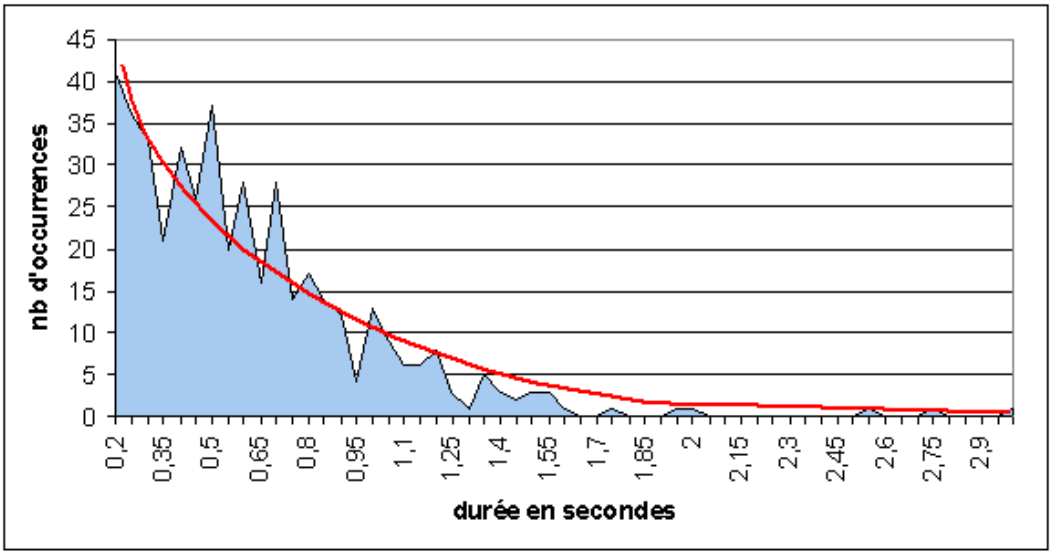

FIG. 8.1 – Distribution des pauses en fonction de leur durée dans le corpus Celia&James

Durée	Proportion
0,20 – 0,45 s	42%
0,50 – 0,95 s	42%
1 – 1,45 s	13%
1,50 – 1,95 s	2%
+ de 2 s	1%

TAB. 8.1 – Distribution des pauses par tranches en fonction de leur durée dans le corpus Celia&James

Comme nous l'avons dit plus haut, nous mesurons uniquement les pauses silencieuses dont les segments précédent et suivant ont été prononcés par le même locuteur. Le corpus Celia&James compte environ 450 pauses de ce type, pour une durée totale de 280 secondes, soit plus de 4 minutes. La durée moyenne des pauses y est de 0,61 s, la médiane de 0,55 s, et l'écart type de 0,38 s. La distribution générale des pauses dans ce corpus est représentée à la Figure 8.1. La courbe théorique superposée au gra-

phique mime la forte décroissance du nombre d'occurrences qui trouve son sommet à 0,20 s et tend vers zéro au-delà d'1,7 s environ. Le Tableau 8.1 analyse la distribution tranche par tranche (arbitrairement choisies pour correspondre à une demi-seconde). Il indique que les pauses supérieures à 2 secondes (et jusqu'à 3 secondes, pause la plus longue) constituent moins de 1% du nombre total des pauses relevées, tandis que plus de 80% d'entre elles se situent dans la tranche inférieure à 1 seconde. Ces chiffres correspondent à ceux avancés par F. Goldman-Eisler pour la conversation.[8]

8.1.2 La part du temps de pause dans l'activité langagière

Dans leurs travaux sur les variables temporelles du français et de l'anglais, F. Grosjean & A. Deschamps[9] suivent un protocole très strict emprunté à F. Goldman-Eisler et permettant notamment de calculer le temps passé par un locuteur à « pauser », au sein de son tour de parole.[10] Il s'agit en fait de mettre en rapport le temps d'articulation (temps où le locuteur prononce effectivement des sons) et le temps de locution (c'est-à-dire le temps de parole). Exprimée en pourcentage, la valeur issue de ce calcul du RTATL[11] indique, en creux, le temps de silence au sein du temps de parole.

F. Grosjean & A. Deschamps montrent ainsi que français et anglais ont des caractéristiques très proches dans deux types de corpus : un corpus d'interviews radiophoniques, et un corpus de description de dessins humoristiques. Dans le premier corpus, le RTATL de l'anglais est de 83,15%, contre 84,45% pour le français. Le second corpus propose des taux beaucoup plus faibles, du fait de la plus grande difficulté de la

[8]GOLDMAN-EISLER, Frieda (1968). *Psycholinguistics : Experiments in Spontaneous Speech.* New York : Academic Press. (pp.14-15).

[9]GROSJEAN, François & DESCHAMPS, Alain (1972). *art. cit.*
GROSJEAN, François & DESCHAMPS, Alain (1973). *art. cit.*
GROSJEAN, François & DESCHAMPS, Alain (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 31 : 144-184.

[10]GOLDMAN-EISLER, Frieda (1968). *op. cit.* (pp.18-26).

[11]Rapport Temps d'Articulation–Temps de Locution

tâche : 56,17% pour l'anglais et 58,71% pour le français. Après avoir relevé et mesuré toutes les pauses silencieuses de plus de 0,20 secondes dans notre corpus Celia&James, nous avons procédé au calcul global du rapport temps d'articulation–temps de locution pour ce corpus. Notre résultat est extrêmement proche de celui de F. Grosjean & A. Deschamps dans leur corpus d'interviews. Notre corpus présente en effet un RTATL de 83,32%. Du point de vue de l'effort cognitif, le corpus des deux auteurs, constitué d'entretiens portant sur les centres d'intérêts, les projets et les activités récentes des locuteurs, s'apparente au nôtre.[12] La différence essentielle réside dans le fait que, dans Celia&James, les rôles ne sont pas établis à l'avance et évoluent en cours de conversation. En somme, dans notre corpus comme dans celui de F. Grosjean et A. Deschamps, la part du silence dans le temps de parole d'un locuteur correspond à environ 17%. Cette place non négligeable occupée par la pause dans notre activité de parole invite à se pencher davantage sur ce phénomène.

8.1.3 Perceptibilité des pauses silencieuses

Nous l'avons vu plus haut, la pause silencieuse ne répond pas uniquement à une contrainte physiologique. La pause respiratoire ne constitue qu'un type de pause parmi d'autres. Site d'occurrence et durée sont donc le fruit d'un calcul signifiant de la part des locuteurs sur le plan linguistique. Les études antérieures fournissent des données relatives à la valeur perceptuelle des pauses et à leur rôle dans la construction du discours et de l'interaction.

8.1.3.1 Durée, site d'occurrence et valeur perceptuelle des pauses

Trois tendances complémentaires émanent des travaux de nos prédécesseurs. Tout d'abord, il ressort que les pauses les plus longues tendent à être situées aux frontières syntaxiques les plus importantes. F. Grosjean & A. Deschamps montrent par exemple

[12] Voir la description du corpus Celia&James, (p.155).

que, dans leur corpus, les pauses situées à l'intérieur de la phrase anglaise sont généralement plus courtes que celles situées à l'extérieur. La valeur médiane varie ainsi de 0,42 à 0,48 s.[13]

Ensuite, en corollaire de cette première remarque, il semble que la hiérarchie syntaxique existant entre propositions, syntagmes et mots soit reflétée dans la durée des pauses situées à leurs frontières. Ce phénomène est connu pour l'anglais depuis les travaux de F. Goldman-Eisler.[14] Celle-ci décrit en effet l'intégration des différents constituants au sein de la phrase, c'est-à-dire leur fonction grammaticale, en relation avec la longueur des pauses éventuelles les entourant. Par ordre croissant d'intégration fluide à la phrase, se trouvent placés sur cet axe mots, propositions subordonnées relatives, autres propositions subordonnées et propositions coordonnées :

> There is indeed a gradual decline in the degree of temporal integration in the sentence body from words within clauses, to relative subordinate, to other subordinate clauses, to co-ordinate clauses. [...]
>
> In other words, *the grammatical description of each of the clause types has its quantitative reflection in the length or absence of pauses which precede it, in natural, unprepared and spontaneous speech.*[15]

Ce phénomène se vérifie également en français. D. Duez[16] constate que F. Mitterrand organise la durée de ses silences en fonction du constituant qu'ils suivent. Lorsqu'une fin d'énoncé est suivie d'une pause, celle-ci dure typiquement 0,88 s. Une pause en fin de proposition est en moyenne longue de 0,64 s, et une pause suivant un syntagme de 0,55 s. Au sein d'un syntagme, les pauses de l'homme politique s'échelonnent de 0,35 à 0,40 s.

[13] GROSJEAN, François & DESCHAMPS, Alain (1975). *op. cit.* (p.71).

[14] GOLDMAN-EISLER, Frieda (1972). Pauses, Clauses, Sentences. *Language and Speech* 15 : 103-113.

[15] *ibid.* (pp.106-107).

[16] DUEZ, Danielle (1995). Silence et pouvoir : une comparaison de trois discours de François Mitterrand. *Travaux du Cercle Linguistique d'Aix-en-Provence* 13 : 153-164. (p.160).

Enfin, les pauses les plus longues tendent à être les plus audibles. C'est ce qui ressort de l'expérience de D. Duez sur la perception des pauses silencieuses en français.[17] Celle-ci utilise trois types de données sonores soumises à ses sujets d'expérimentation chargés de repérer les pauses silencieuses. Le premier échantillon est constitué d'un extrait de parole naturelle. Dans le deuxième échantillon, le signal du premier extrait a subi une inversion spectrale de sorte que la reconnaissance des mots est brouillée mais les paramètres mélodiques, intensifs et temporels sont préservés. Le troisième échantillon est un court passage de l'extrait original dont toutes les syllabes ont été remplacées par la voyelle de synthèse /a/ prononcée à une fréquence fondamentale uniforme. Elle constate que la plupart des pauses très brèves (c'est-à-dire de durée inférieure à 0,25 s) ont un taux d'identification faible, tandis que les pauses plus longues sont systématiquement bien mieux perçues.

> Duration seems to be the essential perceptual parameter of speech pauses, their identification rate being correlated with their duration.[18]

Si les pauses sont perçues indépendamment de leur durée dans les échantillons synthétiques, une importante corrélation est vérifiée, en revanche, entre durée et identification des pauses en parole normale et inversée. Notons que ces deux dernières configurations ne présentent pas de différence sensible, d'après D. Duez, quant à la perception des pauses silencieuses.

8.1.3.2 Quelle place pour le contexte dans l'identification des pauses ?

Nous pouvons formuler trois hypothèses de travail à partir des données fournies par les travaux précédemment cités.

1. Plus la pause est longue, plus la frontière qu'elle marque est importante.

[17] DUEZ, Danielle (1985). Perception of Silent Pauses in Continuous Speech. *Language and Speech* 28 (4) : 377-389.
[18] *ibid.* (p.388).

2. Plus la pause est longue, mieux elle est perçue.
3. Mieux la pause est perçue, plus la frontière est importante.

Ces assomptions, de forme fortement syllogistique, semblent écarter totalement le contexte autre que syntaxique dans notre appréhension des pauses silencieuses.

Il est néanmoins fait mention d'éléments contextuels relatifs à la pause (en dehors de son site d'occurrence syntaxique). Le propos de F. Grosjean et A. Deschamps est d'étudier les pauses et leurs conditions de production au sein des variables temporelles primaires du français et de l'anglais. A ce titre, ils s'intéressent notamment à la durée des suites sonores séparant chaque pause. Ils étudient également l'influence des variables secondaires, c'est-à-dire les pauses pleines, les allongements syllabiques, les répétitions et les faux départs. Plus intéressée par les phénomènes perceptifs, D. Duez analyse l'environnement acoustique immédiat des pauses silencieuses et son impact potentiel sur leur perceptibilité. Son intérêt pour le contexte prosodique des pauses reste limité à la dernière syllabe avant la pause dont elle relève la durée, l'intensité et la fréquence fondamentale, pour les comparer à la durée moyenne ou aux autres valeurs caractérisant l'avant-dernière syllabe. Elle en conclut que d'autres paramètres entrent également en jeu pour la perception d'une pause, notamment la durée de la voyelle précédant la pause, mais la durée de la pause reste un paramètre déterminant. Comme F. Grosjean et A. Deschamps, elle étudie la présence ou l'absence de pauses pleines et d'autres marques d'hésitation. Enfin, elle mentionne la possible importance du débit et dit souhaiter tester ce paramètre dans une étude à venir.

A nos yeux, la pause doit être considérée dans un contexte prosodique plus large, dans lequel elle est susceptible d'avoir un rôle dans la construction thématique du discours et dans la gestion de l'interaction. Sur un corpus de conversation naturelle dont nous avons pu identifier les différents enjeux, nous souhaitons vérifier dans quelle mesure s'appliquent les trois hypothèses dérivées des travaux antérieurs auxquels nous nous référons. Nous souhaitons enfin montrer qu'une attention portée à un contexte

plus large ne peut qu'enrichir notre compréhension du phénomène pausal.

8.2 Mise à l'épreuve des hypothèses de travail

La mise à l'épreuve des trois hypothèses formulées plus haut passe par une confrontation avec les données de notre corpus. Nous présentons ici une expérience à laquelle nous avons procédé visant à tester la perception des pauses silencieuses d'un extrait de notre corpus dans différentes conditions d'écoute.

8.2.1 Test de perception

8.2.1.1 Matériel utilisé

Le matériel de notre expérience est issu de notre corpus Celia&James. Nous avons sélectionné un passage répondant à plusieurs critères : l'unité thématique, le nombre de pauses, et la lisibilité de l'échange. Il s'agit de l'épisode intitulé Sri Lanka, long de 40 secondes, au cours duquel Celia dit à James combien elle aimerait partir en voyage en Asie. De forme relativement monologale, ce passage contient seize pauses silencieuses :

C mm I've been to Japan {pause 1 : 1,05}
but I've never been to:: {pause 2 : 0,45} China or Singapore any of these or {pause 3 : 1,20}
J yeah
C Malaysia at all {pause 4 : 0,70}
I'd love to {pause 5 : 0,65}
J yeah *I I would too*
C *absolutely love to* {pause 6 : 2,75}
J althou::gh I think
C my sister's gone just gone to Thailand {pause 7 : 1,35}
J yeah
C hm {pause 8 : 0,75} my daughter's been round {pause 9 : 0,85} all over the world
in fact she said that Sri Lanka was {pause 10 : 0,60} I don't know quite why but {pause 11 : 0,85} her
theory was that I would absolutely ado::re Sri Lanka {pause 12 : 1,10}
more than Thailand more than Bali {pause 13 : 0,80} more than India {pause 14 : 1,20}

but {pause 15 : 0,35} I haven't really been to {pause 16 : 0,45}
been to Goa but not to India proper

(Celia&James : SriLanka)

Quatre échantillons ont été créés à partir du seul fichier sonore où sont enregistrées les paroles de Celia. Dans les deux premiers, nous avons appliqué au fichier original un filtre passe-bas créé pour l'occasion,[19] qui, en supprimant les fréquences les plus aiguës, permet de brouiller la reconnaissance des phonèmes, en particulier consonantiques, et donc des mots prononcés par Celia. Les deux échantillons sont donc dépourvus de leur contenu informationnel sémantique et syntaxique. Seuls sont conservés les paramètres prosodiques.

La Figure 8.2 (p.386) présente l'analyse comparée des paramètres du signal du premier énoncé de Celia :

C mm I've been to Japan

La lecture des deux copies d'écran issues du logiciel Praat permet d'évaluer dans quelle mesure le spectre du signal a été modifié par le filtre : on voit clairement, en bas, que les attaques des consonnes ont disparu et que les formants sont généralement très affaiblis par rapport au fichier original, tandis qu'intensité, fréquence fondamentale et durée sont préservées.

Le fichier original, non filtré, est notre troisième échantillon, tandis que le dernier est constitué du troisième échantillon auquel nous avons ajouté la transcription orthographique sur support papier (les interventions de James n'étant pas mentionnées, et la mise en page étant réduite à sa plus simple expression). Pour rendre les analyses plus aisées, nous avons choisi de regrouper les résultats issus des deux premiers échantillons afin de comparer la perception des pauses de l'extrait de corpus par nos sujets dans trois configurations différentes : une configuration filtrée, une configuration

[19]Deux filtres passe-bas Kaiser FIR ont été créés par Julien Bœuf, caractérisés par les paramètres suivants : 250-500 Hz (60 dB) et 500-1000 Hz (60 dB), la valeur inférieure indiquant la fréquence à partir de laquelle s'applique le filtre.

normale non filtrée, et une configuration avec support visuel.

8.2.1.2 Procédure appliquée

L'expérience, individuelle, s'est chaque fois déroulée dans un lieu calme. A travers des écouteurs branchés sur un ordinateur portable, les échantillons ont été présentés dans l'ordre à nos sujets. La tâche demandée consistait à repérer les pauses silencieuses et à les signaler en appuyant sur une touche du clavier de l'ordinateur. La définition de ce que nous entendions par pause silencieuse a été laissée volontairement vague. Si nécessaire, nous avons simplement précisé que nous étions intéressée par les endroits où la locutrice interrompt son discours, s'arrête momentanément de parler. Les échantillons n'ont été diffusés qu'une seule fois. Les sujets ont été invités à appuyer sur la touche dès qu'ils avaient perçu que la locutrice faisait une pause ou avait fait une pause, si bien que le signal qu'ils donneraient pouvait correspondre à un endroit où la locutrice avait déjà recommencé à parler. Lorsque la transcription était disponible, il s'agissait d'insérer une marque dans le texte, par exemple un trait vertical, entre les mots séparés par une pause.

L'expérience n'a commencé pour chacun des sujets qu'après nous être assurée que le volume sonore était suffisamment fort et convenait au sujet, et que celui-ci ou celle-ci avait totalement compris les tâches requises.

8.2.1.3 Sujets sélectionnés

Sur les seize sujets ayant participé à l'expérience, seuls neuf ont été retenus, qui ont accompli les quatre tâches dans l'ordre. Nos sujets, quatre femmes et cinq hommes, viennent d'horizons variés. Un grand nombre de langues maternelles différentes sont représentées : anglais américain, français, allemand, italien, arabe et fon. Leur maîtrise de la langue anglaise est également variable, un sujet anglophone côtoyant aussi bien des sujets bilingues que d'autres vivant en pays anglophone depuis peu ou étudiant

l'anglais, ou encore d'autres ayant une connaissance plus limitée de cette langue.

La raison qui nous a fait sélectionner ce panel varié de sujets est simple. Il s'agissait de réduire la probable influence de leur connaissance de la syntaxe de l'anglais et des attentes créées par celle-ci, afin de favoriser une écoute plus naïve de tous les paramètres entrant en jeu — sémantiques, syntaxiques, mais aussi prosodiques. Ainsi seraient-ils sensibles aux pauses les plus perceptibles, et donc les plus signifiantes.

8.2.1.4 Dépouillement des résultats

Au total, 406 interruptions du signal ont été notées par nos sujets, dont 309 ont été correctement identifiées sur les 576 pauses effectives. Nous avons considéré une pause comme identifiée si le sujet l'a signalée en appuyant sur la touche au moins 0,40 s après le début de celle-ci ou au plus 0,70 s après la fin. La borne inférieure trouve sa justification dans les travaux de chercheurs en sciences cognitives sur la mémoire immédiate et le temps de réaction.[20] Elle correspond au retard minimal précédant l'émission d'une réponse issue d'une interprétation du signal perçu. C'est la valeur retenue par M.-A. Morel et L. Danon-Boileau comme intervalle nécessaire pour qu'un locuteur produise une marque d'écoute minimale telle que *mm*.[21] La valeur de la borne maximale a moins de justification théorique. Elle correspond néanmoins, d'après M.-A. Morel,[22] au temps minimal nécessaire pour qu'un locuteur émette une marque d'écoute élaborée. D'un point de vue cognitif, on peut donc considérer que la problématique de la simple identification des pauses silencieuses est dépassée lorsque plus de 0,70 s se sont écoulées après la reprise de la parole. Cette borne n'a été appliquée qu'à deux marques signalant une pause. La première, retenue comme iden-

[20] Citons par exemple le travail de BADDELEY, Alan (1990). *Human Memory. Theory and Practice.* Boston : Allyn and Bacon.

[21] Voir par exemple MOREL, Mary-Annick (2003). Parleur et écouteur. Formes intonatives couplées dans l'échange oral en français. in Mettouchi, Amina & Ferré, Gaëlle (éds.), *Actes de IP2003, Interfaces Prosodiques* : 293-300. Nantes : Université de Nantes. (p.294).

[22] *ibid.* (p.294).

tifiant une pause, est située juste en-dessous de la limite. L'autre, juste au-dessus de cette limite, n'a pas été considérée comme identifiant une pause. En outre, elle est tellement éloignée de la fin de la pause à laquelle elle fait potentiellement référence qu'elle peut presque correspondre à la pause suivante.

Durée des seize pauses Comme le montre la transcription du passage reportée plus haut,[23] les pauses de notre extrait ont des durées occupant presque tout le spectre des durées représentées dans la totalité du corpus Celia&James. La plus courte est de 0,35 s, la plus longue de 2,75 s. La distribution est cependant légèrement décalée vers la droite dans la mesure où les pauses de notre extrait sont globalement plus longues. La moyenne y est de 0,94 s (contre 0,61 s au total), la médiane de 0,83 s (contre 0,55), et l'écart type de 0,56 s (contre 0,38).

Le temps passé à faire une pause est ici supérieur à la moyenne constatée dans la totalité du corpus. Sur les 39 secondes de parole, plus de 15 sont occupées par du silence. Le RTATL n'est ici que d'un peu plus de 60%, ce qui signifie que les pauses de Celia occupent près de 40% de son temps de parole (contre 17% dans le corpus en général).

Les pauses de cet extrait ne sont donc pas représentatives des pauses de la totalité du corpus. Nous estimons néanmoins qu'elles sont éclairantes sur leur fonctionnement général.

Etablissement des scores Pour chaque sujet et chaque échantillon, nous avons établi une liste répertoriant le nombre de pauses signalées et le nombre de pauses effectivement identifiées. A partir de cette liste, nous avons ensuite attribué un score à chaque pause correctement identifiée. Le calcul du score prend en compte le rapport entre le nombre de tentatives et le nombre de succès dans l'identification des pauses.

[23] Voir p.353.

Toute pause correctement identifiée par un sujet dans un échantillon est affectée du même score, conformément à l'équation suivante (8.1) :

$$\text{score} = \frac{100}{nombre \quad de \quad pauses \quad signalées \quad dans \quad l'échantillon} \tag{8.1}$$

Par exemple, notre sujet identifié par l'initiale R a signalé 6 pauses dans l'échantillon non filtré. Sur celles-ci, 5 correspondent à des pauses existant effectivement dans l'échantillon. Seules celles-ci seront affectées d'un score, correspondant à l'équation (8.2), soit 16,67. Sur l'échantillon suivant, le même sujet a noté 9 pauses, toutes correctement identifiées. Le score de chaque pause pour ce locuteur et cet échantillon est donc égal au rapport résultant de l'équation (8.3), soit 11,11.

$$\text{score} = \frac{100}{6} = 16,67 \tag{8.2}$$

$$\text{score} = \frac{100}{9} = 11,11 \tag{8.3}$$

En somme, plus le sujet a signalé de pauses, moins les pauses correctement identifiées parmi celles-ci sont signifiantes. Moins les pauses signalées sont nombreuses, plus celles qui correspondent à une véritable pause prennent de la valeur.

L'objectif n'étant pas d'évaluer la perception de chaque sujet mais la perceptibilité de chaque pause, nous avons ensuite additionné, pour chaque pause, les scores témoignant de sa perception dans chaque échantillon par les neuf sujets. Ce dernier calcul nous permet enfin de classer les 16 pauses en fonction de leur taux de perceptibilité. Sur l'ensemble des quatre échantillons, la pause la moins reconnue a le score total le plus faible (55,53), tandis que la pause la mieux perçue est caractérisée par le score le plus élevé (305,13). Nous disposons ainsi de deux mesures différentes pour évaluer la perceptibilité d'une même pause. Le nombre ou la proportion de sujets l'ayant perçue constituent un indice brut, tandis que le score de cette même pause constitue un indice pondéré de sa perceptibilité.

Etiquetage syntaxique et fonctionnel Nous avons ensuite procédé à l'analyse du site syntaxique d'occurrence de chaque pause. Quatre places possibles ont été envisagées : milieu de constituant, début d'énoncé ou de proposition après ligateur, fin de proposition, et fin d'énoncé. La répartition des pauses en fonction de leur site d'occurrence dans notre extrait de corpus a été établie non sans difficulté. Le principal écueil provient de la subtile distinction entre des énoncés conclusifs et de simples propositions servant de préambule à la suivante. Ce classement a bénéficié de toute notre prudence. Les chiffres la décrivant, au Tableau 8.2, ne sont donnés qu'à titre indicatif et ne reflètent la répartition des pauses que dans notre seul court extrait de corpus. Celui-ci offre en effet une syntaxe beaucoup moins éloignée de celle de l'écrit que dans la majorité du corpus. Le monologue de Celia favorise l'apparition de propositions et d'énoncés complets qui ne sont pas représentatifs du reste du corpus où les tours de parole sont beaucoup moins longs.

Site	Nombre	Proportion
milieu de constituant	3	18,75%
début d'énoncé ou de proposition	4	25%
fin de proposition	2	12,5%
fin d'énoncé	7	43,75%

TAB. 8.2 – Distribution des pauses en fonction de leur site d'occurrence dans notre extrait de corpus

Les fonctions représentées dans cet extrait n'épuisent pas l'éventail des possibilités offertes par les pauses. Les fonctions généralement envisagées dans les études antérieures sont au nombre de quatre (respiration, hésitation, focalisation, et segmentation) et répondent aux définitions de base suivantes :

- La pause respiratoire correspond à une nécessité biologique. Elle n'a pas de fonction linguistique spécifique.
- La pause d'hésitation a pour fonction de faciliter la planification du discours en retardant l'énoncé des éléments qui suivent.

– La pause de focalisation permet de souligner l'élément suivant en le détachant de ce qui précède.
– La pause de segmentation délimite les unités linguistiques.

En gardant à l'esprit qu'une pause peut avoir plusieurs fonctions, et notamment que toute pause peut contenir une respiration, nous avons attribué une fonction à chacune des seize pauses de notre extrait. Nous avons tenu compte des contextes sémantique et syntaxique d'apparition de la pause, et de l'éventuelle présence de marques du travail de formulation. Il en ressort que notre extrait contient un nombre relativement équilibré de pauses d'hésitation (7) et de pauses de segmentation (9). Aucune des seize pauses n'a pour unique fonction de permettre à la locutrice de reprendre son souffle, si bien que nous n'avons comptabilisé aucune pause respiratoire. En outre, aucune fonction focalisatrice n'est décelable dans les pauses dans cet extrait. Les résultats de notre expérience s'appliquent donc uniquement aux pauses d'hésitation et aux pauses de segmentation. C'est pourquoi il nous faudra étudier le fonctionnement des pauses de focalisation dans un autre extrait de corpus.

En somme, notre extrait ne prétend pas concentrer les caractéristiques d'un corpus dont il n'est pas représentatif pour plusieurs raisons déjà évoquées. Néanmoins, nous pensons qu'il permet de tester de manière satisfaisante les hypothèses de travail dégagées plus haut à partir des travaux antérieurs auxquels nous nous référons. D'autre part, il permet de mettre en lumière le fonctionnement des pauses silencieuses dans le cadre de notre expérience sur la perception de ces pauses dans trois configurations différentes.

8.2.2 Vérification de nos hypothèses de travail

Les trois hypothèses formulées à la lecture des travaux de nos prédécesseurs mettent en relief le rôle de la durée et de la fonction dans l'identification des pauses

silencieuses. Dans quelle mesure les sujets de notre expérience se réfèrent-ils à ces deux paramètres pour détecter les silences dans les paroles de Celia ?

8.2.2.1 Hypothèse 1 : durée et emplacement des pauses

Notre première hypothèse concerne le rapport entre la durée de la pause et l'importance de la frontière signalée. Si l'on considère que les pauses de segmentation marquent des frontières plus conséquentes que les pauses d'hésitation, cette première hypothèse est globalement vérifiée dans notre extrait. En effet, aucune des pauses de segmentation ne dure moins de 0,60 s. Cependant, les pauses d'hésitation peuvent également être très longues : 0,85 ou même 1,20 s (pauses 11 et 3). Les pauses d'hésitation les plus longues sont situées en début d'énoncé. Les pauses de segmentation les plus courtes, en revanche, suivent les énoncés non conclusifs, c'est-à-dire les énoncés auxquels la locutrice ajoute une proposition sans l'avoir visiblement planifiée.

Nos données, bien que très limitées, confirment également la hiérarchie évoquée par F. Goldman-Eisler pour l'anglais et D. Duez pour le français.[24] Dans notre extrait, les constituants d'ordre supérieur sont suivis par des pauses plus longues que ceux d'ordres inférieurs. Le Tableau 8.3 montre que la durée moyenne des pauses augmente avec la taille des éléments où elles apparaissent.

Site	Durée moyenne	Ecart type
milieu de constituant	0,50 s	0,09
début d'énoncé ou de proposition	0,79 s	0,35
fin de proposition	1,08 s	0,04
fin d'énoncé	1,19 s	0,73

TAB. 8.3 – Durée moyenne des pauses en fonction de leur site d'occurrence dans notre extrait de corpus

Ces tendances, qui tendent à vérifier la première hypothèse, doivent encore être confirmées par davantage de données issues de notre corpus. En effet, l'écart type reste

[24] Voir p.350.

très important au sein des pauses apparaissant en début d'énoncé ou de proposition après un ligateur initial (0,35 s), et plus encore dans la catégorie des pauses suivant une fin d'énoncé (0,73 s).

8.2.2.2 Hypothèse 2 : durée et perceptibilité des pauses

La deuxième hypothèse, issue des travaux auxquels nous nous référons, suppose un rapport entre la durée des pauses et leur perceptibilité. Disposant de la durée des pauses et des scores témoignant de leur perceptibilité, nous avons pu procéder à un test statistique permettant de tester l'hypothèse d'une corrélation entre ces deux faits. Obtenu après calcul, l'indice ou coefficient de Spearman[25] (ρ, ou r_s) autorise ou non à rejeter l'hypothèse nulle que nous formulons, selon laquelle les scores des pauses et leur durée seraient décorrélés. Compte tenu de nos données expérimentales, il nous est en outre possible de comparer la perception des mêmes pauses dans trois conditions d'écoute différentes (filtrée, non filtrée, avec texte).

Test de corrélation Avant de procéder au test statistique, nous pouvons formuler plusieurs attentes. Tout d'abord, si la durée conditionne la perception, aucune différence ne devrait apparaître entre les scores des pauses dans les différentes configurations, et par conséquent le coefficient de corrélation, indépendamment de son orientation, devrait être stable. Deuxièmement, si nous constatons néanmoins des différences de coefficient, il est prévisible que la configuration filtrée favorise une sensibilité des sujets à la durée des pauses, dans la mesure où ils n'ont pas accès à l'information grammaticale et sémantique. De la sorte, nous prévoyons une corrélation plus élevée dans la configuration filtrée que dans les deux autres.

Appliqué aux résultats issus de l'ensemble des échantillons soumis à nos sujets d'expérimentation, le calcul révèle une forte corrélation entre durée et perceptibilité

[25] Voir par exemple MULLER, Charles (1992). *Initiation aux méthodes de la statistique linguistique.* Paris : Honoré Champion.

(r_s = 0,8125, p $\leqslant$ 0.001, 14 ddl). Il nous permet donc de rejeter l'hypothèse nulle niant l'existence d'une corrélation. Cependant, contrairement à notre première attente formulée plus haut, le coefficient de corrélation admet d'importantes différences en fonction de la configuration d'écoute. Celles-ci sont consignées dans le Tableau 8.4.

	r_s	p
toutes configurations confondues	0,8125	$\leqslant$ 0.001
configuration filtrée	0,7536	$\leqslant$ 0.001
configuration originale (non filtrée)	0,8860	$\leqslant$ 0.001
configuration originale + texte	0,5529	$\leqslant$ 0.03

TAB. 8.4 – Corrélation durée/perceptibilité d'après le coefficient de Spearman

De façon surprenante, la plus forte corrélation n'est pas constatée dans la configuration filtrée, où nous prévoyions que la durée aurait un rôle prépondérant, mais dans la configuration originale, non filtrée. Ce phénomène contrevient donc à notre deuxième attente. Compte tenu du résultat précédent, il semble encore surprenant que la valeur du coefficient chute autant alors que les sujets ont accès au texte en plus du document sonore original. La plus faible corrélation constatée dans cette dernière configuration provient en fait de la meilleure détection de toutes les pauses en général que dans les autres configurations où le nombre de pauses correctement identifiées est 30% inférieure. Le nombre de pauses signalées étant supérieure dans la dernière configuration, la différence des scores n'est que d'environ 10%.

Remarquons qu'il n'existe donc pas chez nos sujet de phénomène d'apprentissage induit par l'ordre d'écoute des échantillons entre les échantillons filtrés et le troisième échantillon (non filtré). Un tel phénomène favoriserait en effet une plus forte identification de toutes les pauses, indépendamment de leur durée, au fur et à mesure que se déroule l'expérience. En outre, sur le total des seize pauses, nous ne constatons de croissance significative de la proportion de sujets ayant correctement identifié les pauses qu'entre les deux derniers échantillons. Peut-être ce phénomène a-t-il donc une part dans les résultats de la configuration impliquant la lecture de la transcription. Il

nous semble néanmoins que la différence constatée entre les scores de perceptibilité entre les échantillons témoigne davantage d'un changement dans l'attention portée aux pauses et à leur durée par nos sujets selon qu'ils ont pour tâche d'écouter ou d'écouter et lire simultanément. En outre, la variation du coefficient de corrélation et la corrélation moyenne constatée dans la configuration filtrée nous semblent indiquer que la durée n'est pas le seul indice ou du moins pas le plus important permettant l'identification des pauses dans notre expérience, en particulier lorsque les sujets ont pour seul support les paramètres intonatifs.

Perception filtrée et perception non filtrée Notre expérience fait apparaître d'importantes différences entre la perception des pauses dans les différents échantillons selon que le signal de parole a été filtré ou non. Ces résultats sont en contradiction avec ceux que D. Duez expose dans son article de 1985, où le signal original était proposé en regard d'un signal inversé et d'un signal synthétique.[26] Comme nous l'avons dit plus haut, son étude révèle peu de différence entre la perception des pauses lorsque le signal est filtré ou qu'il n'est pas modifié. Elle constate en revanche un écart important entre la configuration synthétique et les deux autres.

Afin de comparer la perception des pauses dans les trois configurations de notre expérience, considérons à présent la contribution de chacune des conditions d'écoute au score total de chaque pause. Nous savons déjà que la perception des pauses est différente dans les trois cas. Dans quelle mesure la part de chaque configuration apportée au score de chaque pause s'éloigne-t-elle de l'équilibre des 33% ? Les pauses 12 et 14 sont moins détectées dans les échantillons filtrés et avec le texte que dans l'échantillon original, qui contribue à leur score à hauteur de plus de 45%. Ces deux pauses, longues de plus d'une seconde, ont une fonction de segmentation. Les pauses dont l'échantillon original est le moins responsable du score sont les pauses 1, 4, 10, et

[26] DUEZ, Danielle (1985). Perception of Silent Pauses in Continuous Speech. *Language and Speech* 28 (4) : 377-389.

en particulier 16. La pause 16 est la plus courte de l'extrait, et son faible taux d'identification est conforme avec la forte corrélation démontrée plus haut. En revanche, la pause 1 est l'une des plus longue (plus d'une seconde). Pour toutes les pauses, à l'exception des pauses 2 et 3, la part de chaque configuration est loin d'être constante. Elle est chaque fois différente du seuil des 33%.

D'autres différences intéressantes apparaissent à la comparaison de pauses de même durée. Les pauses 3 et 14 sont perçues aussi bien l'une que l'autre dans les conditions où le signal n'a pas été filtré. Le score de la première est cependant presque deux fois plus élevé lorsque l'information sémantique et syntaxique est brouillée. Dans une moindre mesure, le phénomène se retrouve entre les pauses 9 et 11 lorsque, cette fois, le signal n'est pas filtré.

Pour l'heure, notre objectif n'est pas de fournir des explications, mais simplement d'évaluer les différences perceptuelles concernant les mêmes pauses dans trois conditions d'écoute différentes. Celles-ci contribuent encore à remettre en question l'impact de la seule durée dans notre perception des pauses silencieuses. Les écarts constatés précédemment semblent en outre indiquer un rapport possible entre fonction et perceptibilité de la pause. Nos derniers exemples opposent en effet des pauses d'hésitation (3 et 11) et des pauses de segmentation. Dans quelle mesure ce rapport est-il vérifié ?

8.2.2.3 Hypothèse 3 : perceptibilité et emplacement des pauses

Notre troisième hypothèse établit un rapport entre la perceptibilité d'une pause et le type de frontière qu'elle signale. Il apparaît que les pauses de segmentation de notre extrait de corpus correspondent aux sites d'occurrence délimitant les constituants les plus larges (propositions et énoncés), tandis que les pauses d'hésitation se trouvent toutes en milieu de constituant ou en début de proposition ou d'énoncé. La perception des pauses reflète-t-elle ces différences fonctionnelles ?

Notre test montre que les pauses d'hésitation passent plus souvent inaperçues que

les pauses de segmentation. Au total, les pauses d'hésitation sont perçues par 43% des sujets, tandis que les pauses de segmentation sont détectées par plus de 62% d'entre eux. Six des sept pauses d'hésitation font partie des six pauses détectées par le moins de sujets dans la configuration filtrée et en général. Les pauses d'hésitation 15 et 16, par exemple, sont passées inaperçues pour 75% des sujets, tandis que les pauses de segmentation captent toujours l'attention d'au moins la moitié de nos sujets. La seule pause caractérisée par un taux d'identification de 100% des sujets dans l'une des configurations est une pause de segmentation (il s'agit de la pause 1, dans la configuration avec le texte).

Remarquons que ces résultats, qui vérifient notre troisième hypothèse, ne corroborent pas ceux de D. Boomer et A. Dittman, obtenus sur un test de nature différente du nôtre.[27] Ces deux chercheurs décrivent une expérience dans laquelle ils font varier la longueur de la pause d'hésitation et de la pause de segmentation (ou *juncture pause*) existant dans quatre longs énoncés issus de leur corpus de conversations radiodiffusées. Huit échantillons sont créés. Dans les trois premiers, la pause de segmentation est supprimée, et la pause d'hésitation manipulée de sorte qu'elle dure 100, 200 ou 500 ms. Dans les trois autres, c'est la pause d'hésitation qui est supprimée, et la durée de la pause de segmentation manipulée. Deux autres échantillons de référence sont créés, où toute pause est éliminée. Ces échantillons servent de point de comparaison avec les énoncés où la durée d'une des pauses est manipulée. La tâche demandée à leurs informateurs est de déterminer si les deux versions du même énoncé (sans pauses ou avec pause manipulée) sont identiques. Ils montrent que la détection d'une différence entre les énoncés soumis est meilleure lorsque la durée de la pause d'hésitation est manipulée. Les pauses grammaticales apparaissant à des endroits attendus, elles passent davantage inaperçues dans le test des deux auteurs. Ainsi, ils montrent que les pauses

[27] BOOMER, Donald & DITTMAN, Allen (1962). Hesitation pauses and juncture pauses in speech. *Language and Speech* 5 : 215-220.

d'hésitation ont un seuil de perceptibilité plus bas que les pauses de segmentation, les rendant ainsi plus faciles à détecter :

> [...] the thresholds would be about 200 msec. for hesitation pauses and somewhere between 500 and 1,000 msec. for juncture pauses.[28]

Si un test de différenciation fait apparaître de tels écarts, il ressort plus généralement des études antérieures que les pauses d'hésitation ont tendance à moins retenir l'attention des sujets lorsqu'on leur demande de repérer toutes les pauses dans un document sonore. Un autre phénomène bien connu tend à confirmer la plus grande perceptibilité des pauses de segmentation. Le site d'occurrence privilégié des pauses subjectives (phénomène de perception d'une pause silencieuse en l'absence de toute pause) indique que la détection des pauses est favorisée par la présence d'une frontière syntaxique. M. Candea formule la remarque suivante :

> Ce constat montre que les pauses silencieuses structurant et hiérarchisant les énoncés sont attendues par les auditeurs et, loin d'être gommées de la perception, elles sont au contraire rétablies subjectivement même là où elles étaient en réalité absentes.[29]

Dans notre expérience, considérons à présent les pauses dont les scores sont en-deçà ou au-delà des prévisions compte tenu de leur durée. Notre test de corrélation permet de comparer un classement des pauses en fonction de leur durée avec un classement issu des scores obtenus, et ainsi d'évaluer dans quelle mesure la perception de certaines pauses s'écarte du modèle suggéré par notre première hypothèse. Trois pauses se distinguent particulièrement. La pause 8 a un score bien moindre que sa durée ne laissait attendre, tandis que les pauses 4 et 5 se détachent comme exceptionnellement bien détectées compte tenu de leur durée. Il est frappant de constater

[28] *ibid.* (p.217).

[29] CANDEA, Maria (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (p.131).

que cette distinction recoupe une différence de fonction, la pause 8 étant une pause d'hésitation et les deux autres servant à la segmentation.

La fonction permet encore d'expliquer les différences de score entre des pauses de même longueur, là où l'hypothèse précédente a montré ses limites. Comparons les trois pauses suivantes : les pauses 4, 8 et 13. Les pauses 4 et 13 ont des caractéristiques très similaires, en termes de durée, de site d'occurrence et de fonction : toutes les deux sont des pauses de segmentation relativement longues (0,70 et 0,80 s) situées en fin d'énoncé. Le score des deux pauses est comparable dans toutes les configurations. En revanche, si la pause 4 et la pause 8 sont de longueurs presque semblables (0,70 et 0,75 s), les deux pauses ont des fonctions et des sites d'occurrence différents. La pause 8 est une pause d'hésitation située en début d'énoncé, alors que la pause 4 sépare deux énoncés. Cette différence se traduit par un important écart dans les scores de ces deux pauses dans toutes les configurations. Les pauses 5 et 10 confirment encore que, lorsque deux pauses de fonctions différentes ont la même durée, la pause de segmentation a tendance à être mieux détectée.

8.2.3 Limites de nos hypothèses de travail

D'une manière générale, notre test de perception valide les trois hypothèses formulées à partir des études antérieures auxquelles nous nous référons. Cependant, d'importantes zones d'ombre persistent, qui suggèrent les limites que nous souhaitons exposer à présent.

8.2.3.1 Première hypothèse

Selon la première hypothèse, les pauses les plus longues correspondent à des frontières importantes. En effet, la durée des pauses de segmentation (qui marquent les frontières les plus importantes dans notre extrait) est bien supérieure à celle des

pauses d'hésitation. Ces dernières durent en moyenne 0,66 s (écart type : 0,30), contre 1,17 s (écart type : 0,64) pour les autres. Ce résultat ne corrobore cependant pas celui de G. Ferré sur un corpus semblable au nôtre, qui ne relève pas d'écart significatif entre les durées des différents types de pauses.[30] Rappelons que nos données sont ici extrêmement limitées. Comme le montre l'important écart type caractérisant la durée des pauses de segmentation, seize pauses ne peuvent suffire à confirmer ou infirmer une règle. Elles ne peuvent qu'indiquer une tendance, conforme ici à notre première hypothèse. Ce constat ne constitue pas en soi une limite de l'hypothèse en question mais de nos résultats.

8.2.3.2 Deuxième hypothèse

La deuxième hypothèse prévoit que les pauses les plus longues seront les mieux perçues. Comme nous avons pu le constater, elle est en grande partie vérifiée par notre expérience. Cependant, cette hypothèse ne peut pas expliquer que le coefficient de corrélation que nous avons calculé soit différent selon les configurations d'écoute alors que la durée des pauses reste inchangée. En outre, la part de chaque configuration apportée au score de chaque pause n'est qu'en de rares cas équilibrée (c'est-à-dire proche du seuil des 33% où la pause est perçue avec autant de succès dans les trois configurations). Certes, l'ajout du texte permet d'augmenter l'acuité de la détection. Mais la contribution de chaque configuration ne présente pas suffisamment de régularités à travers toutes les pauses pour nous permettre de caractériser une écoute comme systématiquement plus réceptive aux pauses. Au plus, elle suggère un décalage de l'attention des auditeurs d'un type de phénomène vers un autre type de phénomène que nous souhaitons caractériser plus loin.

[30] FERRE, Gaëlle (2003). Discursive, Prosodic and Gestural Marking of Focalisation Pauses in British English. in Mettouchi, Amina & Ferré, Gaëlle (éds.), *Actes de IP2003, Interfaces Prosodiques* : 265-270. Nantes : Université de Nantes.

8.2.3.3 Troisième hypothèse

D'après la troisième hypothèse, à un fort taux de perception correspond une frontière importante. Nos résultats révèlent une tendance semblable dans la mesure où les pauses de segmentation, qui délimitent des frontières syntaxiques plus importantes que les pauses d'hésitation, sont généralement mieux perçues. En outre, à longueur égale, la pause de segmentation émerge plus facilement. Ces résultats plaident pour une prise en compte conjointe de la durée et de la fonction de la pause pour expliquer sa perceptibilité. Pourtant, notre expérience montre que les deux critères ne sont pas toujours suffisants. Elle présente en effet des contre-exemples criants de pauses dont le score diffère alors que leurs durée et fonction sont les mêmes. Citons par exemple les pauses d'hésitation 2 et 16, toutes deux longues de 0,45 s. Deux fois plus de sujets ont détecté la pause 2 dans la configuration filtrée, et quatre fois plus dans la configuration originale. Le score de la première est 2,8 fois supérieur à celui de la seconde lorsque le signal est filtré, et presque 6 fois supérieur dans le fichier original. Nos hypothèses sont incapables d'expliquer de telles différences de perceptibilité entre des pauses de même durée et de même fonction.

En outre, il nous semble intuitivement réducteur d'imputer à la seule pause la perception de la frontière à laquelle notre dernière hypothèse fait référence. Si notre définition de la frontière ne prend en compte que des données syntaxiques, comment concilier notre hypothèse et les résultats issus des échantillons filtrés qui privent nos sujets de l'information relevant de la syntaxe ? Or les pauses de segmentation sont perçues par davantage de sujets que les pauses d'hésitation dans toutes les configurations, y compris donc, dans la configuration filtrée. L'importance de la frontière ne peut pas selon nous se réduire à la taille de l'unité syntaxique délimitée.

Devant ces écueils, nous souhaitons proposer une hypothèse supplémentaire permettant de comprendre la perception des pauses en les inscrivant dans leur plus large

contexte d'apparition, incluant des paramètres relevant des plans pragmatiques et prosodiques.

8.3 La pause silencieuse comme indice de la construction du discours et de l'interaction

Le succès mitigé des trois hypothèses formulées à partir de notre lecture des travaux antérieurs nous amène donc à proposer une hypothèse supplémentaire permettant de prendre en compte des données extérieures à la pause même. Selon nous, ces données entrent en jeu dans la perceptibilité des pauses silencieuses et permettent d'approfondir notre appréhension de leur rôle à plusieurs niveaux dans la construction du discours et de l'interaction. La pause silencieuse nous semble en effet tout d'abord prendre sens au sein d'un contexte prosodique plus complexe qui contribue à son identification. Nous montrerons ensuite comment elle s'insère également dans la gestion pragmatique de l'échange.

8.3.1 Niveaux de pertinence des différentes pauses

Nous souhaitons ici reprendre l'une des hypothèses fondatrices de ce travail définie au chapitre 2,[31] selon laquelle les indices suprasegmentaux sont interprétables à différents niveaux selon le degré de résolution choisi pour l'analyse du matériau oral. Il s'agit pour nous de montrer ici comment le silence participe à la construction hiérarchisée et à l'organisation du discours et de l'interaction aux niveaux définis précédemment.

[31] Voir en particulier Section 2.1.2.2, p.121.

8.3.1.1 Les niveaux de construction du discours conversationnel : rappel

L'activité conversationnelle implique au moins trois couches élaborées simultanément auxquelles nous souhaitons porter notre attention. Au niveau le plus bas de la construction du discours, le locuteur doit choisir de quoi parler et trouver ses mots. Il doit ensuite se préoccuper de la façon dont il souhaite présenter ce dont il veut parler : il s'agit, au niveau discursif, de mettre en relation les mots sur lesquels son choix s'est arrêté, de façon marquée ou non marquée selon ce qu'il estime nécessaire compte tenu de son intention communicative et du contenu de pensée qu'il attribue à son interlocuteur. Enfin, parler signifie également envisager la prise de parole de l'autre : conserver la parole ou la donner constituent la couche supérieure de la construction du discours conversationnel que nous souhaitons envisager.

Comme nous avons pu le voir plus haut,[32] la tradition linguistique considère les phénomènes d'hésitation et d'auto-correction (au nombre desquels figurent certaines pauses silencieuses) comme des scories, des accidents, qui retardent l'émission du message et contribuent à en rendre la compréhension difficile. Pour W. Chafe, une telle conception de la langue orale spontanée, fondée sur la distinction entre « performance » et « compétence » est erronée. Dans le passage suivant de son article de 1980, il reproche notamment à N. Chomsky[33] de considérer que le locuteur est davantage préoccupé par la dimension syntaxique que par la dimension sémantique et pragmatique de son discours et, de ce fait, que la tâche du linguiste est de lire les règles de compétence à travers la performance :

> "A record of natural speech will show numerous false starts, deviations from rules, changes of plan in mid-course, and so on. The problem for the linguist [...] is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual

[32] Voir Section 3.2.1.3, p.176.
[33] CHOMSKY, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA : MIT Press.

> performance." (Chomsky, 1965 : 4). This view is, I think, misleading to the extent that it suggests that the speaker has some grammatical ideal toward which he is striving, and which he is often prevented from attaining because of various imperfections in the psyshological processes involved in "performance".[34]

Pour W. Chafe comme pour nous, au contraire, la dimension sémantique et pragmatique est centrale dans l'activité conversationnelle :

> [...] the speaker's chief goal is to get across what he has in mind, and that he is not likely to be interested in grammaticality unless there is some special reason to think of it, as there usually is not. The speaker is interested in the adequate verbalization of his thoughts. Pauses, false starts, afterthoughts, and repetitions do not hinder that goal, but are steps on the way to achieving it.[35]

Ainsi, si l'on considère que la transmission du contenu prime sur l'état du contenant, les pauses d'hésitation sont bien mieux analysées en termes de charnières qu'en termes de points d'achoppement, d'imperfections et d'échecs localisés de la formulation. Nous proposons d'adopter la même perspective pour toutes les pauses silencieuses.

8.3.1.2 Différents types de charnières

Les pauses silencieuses sont le témoin de charnières de divers ordres dans la construction du discours et de l'interaction. D'un point de vue cognitif, le silence procure une sensation de groupement, comme en témoigne la remarque du scientifique A. Baddeley :

> Even a brief pause is sufficient to produce the grouping effect, which probably arises as a result of the underlying auditory system being specifically evolved

[34]CHAFE, Wallace (1980). Some Reasons for Hesitating. in Dechert, Hans W. & Raupach, Manfred (éds.). *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler* : 169-180. The Hague : Mouton. (p.170).

[35]*ibid.* (p.170).

to detect and use the rhythmic and prosodic aspects of speech.[36]

Le silence se donne donc en discours comme le reflet d'une certaine complétude. En fait, les pauses silencieuses indiquent, chez le locuteur, l'état de la planification et de l'organisation des constituants au sein de son discours. Associées à des marques du travail de formulation, elles révèlent en particulier les moments où la poursuite du discours est problématique : en dépit de ce qu'il laisse dans le non-dit, le silence signale à l'interlocuteur que le locuteur est arrivé à un point où ce qui précède est complet, ne peut être défait que par des amendements *a posteriori*. Ce type de pause peut apparaître en relation avec des difficultés immédiates de mise en mots (pour trouver le bon mot, la bonne nuance). Bien que l'attention à la grammaticalité du dit soit accessoire, de telles pauses peuvent également correspondre au constat d'une trop grande incompatibilité entre la structure grammaticale prévue ou désirable et celle en cours de réalisation pour continuer à délivrer le message. Les pauses silencieuses permettent également de délimiter des groupes de sens et de baliser la charpente syntaxique du discours. Elles ont pour effet de faciliter le travail de décodage de l'interlocuteur et d'en diminuer le coût cognitif.

Dans tous les cas, le silence a pour effet de signaler que le locuteur est allé au bout de ce qu'il prévoyait de dire ou de ce qu'il pouvait dire compte tenu de la planification de son discours. On comprend donc pourquoi l'un des effets possibles de la pause silencieuse d'hésitation est d'autonomiser le segment qui précède pour éventuellement l'annuler en repartant sur des bases permettant d'aller plus avant dans le développement du discours. On comprend également pourquoi les pauses marquant une telle complétude sont attendues par notre oreille qui a tendance à mieux repérer les silences lorsqu'ils correspondent à des frontières syntaxiques, au point d'en percevoir en leur absence. Ces pauses sont pertinentes au niveau minimal de la constitution et

[36]BADDELEY, Alan (1990). *Human Memory. Theory and Practice.* Boston : Allyn and Bacon. (p.40).

de la mise en place des constituants discursifs, c'est-à-dire le niveau où le locuteur décide de quoi il va parler et cherche les mots adéquats.

Au niveau discursif, les pauses silencieuses correspondent à un moment où l'énonciateur choisit de souligner un élément au sein de son discours. Le silence permet ici de briser la linéarité du discours. En l'isolant de ce qui se trouve sur sa gauche et en mimant une difficulté de formulation, il met en relief l'élément qui suit et délimite un site dont le contenu doit recevoir une attention particulière de la part de l'interlocuteur. Ce type de pause correspond à une stratégie de focalisation, pertinente au niveau de la mise en relation modalisée des éléments du *dictum*.

Enfin, les pauses silencieuses sont également pertinentes au niveau supérieur de la construction du discours conversationnel, c'est-à-dire au niveau proprement interactionnel. Elles constituent en effet d'importantes charnières sur le plan pragmatique, liées à la problématique de la conservation et de la cession de la parole. En conjonction avec d'autres paramètres prosodiques (comme la baisse conjointe de la fréquence fondamentale et de l'intensité), elles indiquent une possible fin de tour de parole et donc un endroit où il est opportun pour l'interlocuteur de prendre la parole. Certes, le silence ne peut pas être considéré comme un corrélat acoustique suffisant ni même nécessaire des places transitionnelles. Néanmoins, au niveau interactionnel, le silence en conversation est toujours l'effet d'une co-construction. D. Tannen a la remarque suivante à propos du silence au cours d'une interaction verbale :

> [...] like other features of discourse, it is always a joint production.[37]

Absence de parole de part et d'autre du canal, le silence comprend toujours en lui la perspective d'une prise de parole de l'un ou l'autre des interlocuteurs. D. Tannen poursuit sa remarque, affirmant que si le silence constitue bien une absence de marqueur d'interaction, il ne marque aucunement l'absence d'interaction.

[37]TANNEN, Deborah (1985). Silence: Anything But. in Tannen, Deborah & Saville-Troike, Muriel (éds.), *Perspectives on Silence* : 93-111. Norwood, NJ : Ablex. (p.100).

> [...] when there are two or more participants in a conversation—in other words, when there is conversation—anything that happens or doesn't happen, is said or unsaid, is the result of interaction among the two [...]. At any point that one person is not talking and thereby produces a silence, no one else is talking either—or there wouldn't be a silence.[38]

Nous rejoignons ainsi le débat concernant absence de marqueur et marqueur de l'absence engagé notamment par M. Candea dans son article de 1997.[39]

D'un point de vue interactionnel, l'opposition entre absence et présence d'une pause est néanmoins signifiante. Elle l'est d'autant plus lorsque la pause silencieuse est accompagnée d'une pause pleine. Loin de signaler une place transitionnelle, la combinaison des deux marques (appelée « pause mixte » par M. Candea) entre dans une stratégie de conservation de la parole :

> Les pauses mixtes présentent un fonctionnement différent des pauses simples, elles permettent des durées nettement plus longues à l'intérieur des syntagmes, elles garantissent le maintien de la parole dans des situations interactives où les temps de parole sont négociés sans cesse [...].[40]

Puisque les niveaux de construction du discours conversationnel fonctionnent comme des couches superposées, on comprend qu'une même pause puisse revêtir plusieurs fonctions. Un même silence peut en effet être pertinent à différents niveaux et témoigner de charnières de plusieurs types à la fois. D'autres pauses peuvent en revanche n'être pertinentes qu'à un seul niveau. Nous proposons ici l'hypothèse selon laquelle notre perception des pauses reflète notre perception de leur rôle dans la construction du discours et de l'interaction. Celle-ci s'appuie amplement non seulement sur leur durée et leur fonction, mais également sur le contexte dans lequel les différents types

[38] *ibid.* (p.100).

[39] CANDEA, Maria (1997). Peut-on définir la pause dans le discours comme un lieu d'absence de toute marque ? *Travaux linguistiques du CerLiCo* 10 : 231-244. Rennes : Presses Universitaires de Rennes.

[40] *ibid.* (p.238).

de pauses silencieuses apparaissent — c'est pourquoi la pause silencieuse nous semble devoir être considérée comme un marqueur mixte.

8.3.2 Un marqueur mixte dans un contexte prosodique complexe

Clarifions tout d'abord notre emploi du terme « marqueur mixte » pour décrire les pauses silencieuses. M. Candea utilise le même qualificatif en référence à la combinaison de deux pauses simples, par exemple une pause silencieuse et une pause pleine. Elle considère ainsi que la pause mixte, combinant deux pauses, est différente de la pause simple. Nous pensons que si une telle différence peut être faite, c'est parce que la pause ne se définit qu'en fonction du contexte dans lequel elle apparaît. Elle forme ainsi pour nous une marque mixte avec son contexte. Hors contexte, tout silence ressemble à un autre silence. Ce qui le définit, c'est certes sa durée (caractéristique propre, indépendante d'autre chose), mais aussi son site d'occurrence et sa fonction (qui ne se dessinent qu'en contexte), et encore les paramètres prosodiques et pragmatiques des segments qu'il délimite. Dans quels contextes prosodiques, syntaxiques et pragmatiques les pauses silencieuses s'inscrivent-elles ? Que révèlent-elles de notre appréhension de la construction du discours et de l'interaction dans les trois configurations envisagées dans notre test de perception ?

8.3.2.1 Contexte prosodique

Quelles sont les combinaisons de paramètres prosodiques qui favorisent la perception d'une pause ? Pour répondre à cette question, revenons à notre extrait de corpus exploité dans notre test de perception (épisode *Sri Lanka* de Celia&James).

Le débit et la longueur des segments précédant la pause Nous avons ainsi testé le rôle du débit sur le segment précédant chacune des seize pauses de l'ex-

trait de corpus étudié. Il est possible qu'un débit ralenti ait tendance à masquer la présence d'une pause, fût-elle longue, et qu'inversement, un débit rapide permette la meilleure perception de pauses relativement brèves. Rappelons que, d'après les travaux de G. Ferré sur les pauses d'hésitation, le débit est plus lent avant qu'après la pause.[41]

Les données relatives au débit de Celia sur les seize suites sonores suivies des seize pauses de l'extrait sont consignées au Tableau 8.5. Le nombre de syllabes prononcées entre chaque pause est extrêmement variable : il s'étend de 1 à 15, avec une moyenne de 6,5, une médiane de 6 et un écart type de près de 4 syllabes. La durée des suites sonores varie quant à elle de 0,40 à 2,60 s. Le débit de Celia est relativement lent dans notre extrait puisqu'elle prononce en moyenne 4,55 syllabes par seconde, soit 273 par minute.

	Suites sonores		Débit	
	syllabes	**secondes**	**syll/sec.**	**syll/min.**
moyenne	6,5	1,36	4,55	273
médiane	6	1,21	4,48	269
écart type	3,9	0,71	1,07	64
perceptibilité corrélée ?	NS	NS	NS	NS

TAB. 8.5 – Débit, longueur des suites sonores et perceptibilité des pauses dans notre extrait de corpus

Comme l'indique la dernière ligne de notre Tableau 8.5, ce test n'a pas permis d'établir de corrélation entre le débit du locuteur sur la suite sonore précédant la pause et la perceptibilité de celle-ci. Dans tous les cas, le coefficient est très faible, mais le degré de confiance insuffisant. Nos calculs ne permettent pas non plus de faire apparaître de corrélation entre la longueur des suites sonores et la perceptibilité des pauses. En complément, à la suite de l'article de 1975 de F. Grosjean et A. Des-

[41] La tendance inverse se vérifie pour les pauses de segmentation. Voir FERRE, Gaëlle (2003). Discursive, Prosodic and Gestural Marking of Focalisation Pauses in British English. in Mettouchi, Amina & Ferré, Gaëlle (éds.), *Actes de IP2003, Interfaces Prosodiques* : 265-270. Nantes : Université de Nantes. (p.266).

champs,[42] nous avons également envisagé la corrélation entre longueur des suites sonores et durée de la pause. Le résultat de ce dernier test est conforme à ceux de deux auteurs, qui établissent la non-corrélation entre la longueur des suites sonores et celle des pauses silencieuses. Notre extrait étant court et offrant une variété de constructions syntaxiques limitée, nous n'avons pas pu vérifier l'application à l'anglais des résultats d'A. Butcher pour l'allemand. Celui-ci montre dans son article de 1980 que la longueur de la pause de segmentation attendue par les sujets d'expérimentation ou effectivement prononcée par les locuteurs augmente de façon logarithmique avec la complexité des segments précédents. Si complexité et longueur des suites sonores ont partie liée, il est néanmoins prévisible, au vu de nos résultats, qu'une telle augmentation de la durée de la pause ne soit pas vérifiée dans notre extrait de corpus.

L'intensité Dans notre extrait, il n'apparaît pas de corrélation entre la perceptibilité de la pause et l'intensité des syllabes accentuées de la suite sonore précédente.

D'une façon générale, notre étude montre que l'intensité n'augmente jamais sur le segment précédant la pause. Elle diminue la plupart du temps, et plus rarement reste inchangée. En revanche, le segment débutant juste après la pause silencieuse, indépendamment de la fonction de celle-ci, est dans la majorité des cas marqué par une intensité accrue. Il semble que la locutrice ressente le besoin de compenser pour le silence qu'elle vient de produire, et de réaffirmer son droit à la parole, même lorsque celui-ci n'a pas été mis en danger.[43] Les cas où l'intensité décroît ne présentent pas d'unité en ce qui concerne la fonction ou la longueur de la pause concernée. Leurs points communs sont à trouver dans le lien entre les suites sonores séparées par la

[42] GROSJEAN, François & DESCHAMPS, Alain (1975). *art. cit.*

[43] L'intensité compensatrice est lié au fait que la pression sous-glottique redevient normale après un relâchement dû à l'interruption de l'émission du signal sonore et notamment après une respiration. Ce phénomène correspond pour l'intensité à ce que C. Gussenhoven décrit sous le nom de code de la production. Voir Section 2.2.2, p.2.2.2, et CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002 conference* : 215-218. (p.215).

pause. Les pauses marquées dans notre extrait par une baisse d'intensité entre la dernière syllabe accentuée la précédant et la première syllabe accentuée la suivant sont les pauses 2, 5, 7, 9 et 15.[44] Elles correspondent à des charnières relativement faibles dans la construction du discours. La pause 2 se combine avec une marque d'hésitation (un allongement sur la voyelle de la préposition *to*) afin de laisser à Celia le temps de trouver ses mots. Bien que pause de segmentation, la pause 5 ne suit pas un segment conclusif, puisque la locutrice fait un ajout *a posteriori*. Il nous semble possible de donner la même interprétation pour la pause 7, dans la mesure où le segment qu'elle suit appelle un développement plus circonstancié. De même pour la pause 9. Enfin, la pause 15 entre en jeu dans une stratégie dilatoire où la locutrice pose le lien, même factice, entre ce qu'elle vient de dire et ce qu'elle cherche encore à formuler.

De manière générale, le fonctionnement de l'intensité sur les segments séparés par des pauses silencieuses ne permet pas d'expliquer les différences de perceptibilité relevées dans notre expérience.

Un paramètre corrélé à la perceptibilité des pauses silencieuses : la ligne de déclinaison Nos tests de corrélation ont davantage de succès lorsque nous nous tournons vers les caractéristiques mélodiques des segments séparés par les pauses de notre extrait. Nous nous sommes concentrée sur la hauteur moyenne des syllabes accentuées d'une même suite sonore. Nous avons ensuite comparé la valeur calculée pour chacune des seize séquences et le score de perception des pauses correspondantes. Il apparaît ainsi qu'un rehaussement général de la hauteur mélodique d'un segment par rapport au précédent favorise la perception de la pause qui les sépare.

Notre extrait donne plusieurs cas de réinitialisation de la ligne de déclinaison, qui correspondent tous à une meilleure perception de la pause précédente. Par exemple,

[44] Voir la transcription de l'extrait, p.353.

la pause 5 est généralement bien mieux perçue que la pause 10, bien que de durée comparable. Certes, cette différence peut s'expliquer comme nous l'avons fait plus haut par une différence de fonction, la pause 5 étant une pause de segmentation. Le contexte mélodique des deux pauses nous semble constituer un meilleur élément d'explication, et notamment un élément plus fiable en particulier lorsque l'accès à la syntaxe et à la sémantique est brouillé. Le segment qui suit la pause 5 est marqué par une réinitialisation de la ligne de déclinaison. La même réinitialisation se retrouve après la pause 12, qui, nous l'avons vu, jouit de scores de perceptibilité largement supérieurs à ceux de la pause 1, malgré une durée et une fonction comparables. En configuration filtrée, les pauses suivies d'une réinitialisation de la ligne de déclinaison sont mieux perçues que les pauses de même longueur et que toutes les autres pauses. Dans notre extrait, les pauses 5, 6 et 12 s'inscrivent dans un tel contexte prosodique. Elles font partie des quatre pauses les mieux perçues lorsque syntaxe et sémantique ne guident pas notre écoute.

Si la réinitialisation de la ligne de déclinaison semble constituer un élément fiable, le contexte prosodique accompagnant la pause n'est que partiellement responsable de la meilleure perceptibilité de certaines pauses par rapport à d'autres, même en configuration filtrée.

8.3.2.2 Contexte pragmatique

Le rôle de la ligne de déclinaison dans la perception des pauses nous conduit à nous intéresser plus généralement au rôle du contexte pragmatique dans lequel s'inscrivent les pauses de notre extrait. Selon notre hypothèse des niveaux de pertinence, l'organisation du discours conversationnel s'intéresse au niveau supérieur à la mécanique interactionnelle, et en particulier à la gestion des tours de parole. Dans quelle mesure les pauses sont-elles pertinentes à ce niveau ?

Se pose ici la question de la contribution des pauses à la perception des charnières

interactionnelles d'une conversation. De façon très modeste, compte tenu du peu de données dont nous disposons, nous proposons de mesurer celle-ci à la lumière de la perceptibilité des pauses correspondant à des places transitionnelles dans le discours de Celia, c'est-à-dire à des endroits où un changement de locuteur est envisageable. La transcription du passage reportée plus haut montre que James participe à la conversation de façon sonore à quatre reprises, toujours pendant une pause de Celia (pauses 3, 5, 6 et 7). Nous plaçant du côté de la réception, nous considérons pour l'instant qu'une place transitionnelle est atteinte lorsque l'un des interlocuteurs tente de prendre ou obtient la parole. On nous objectera ici à raison que les répliques de James prononcées pendant les pauses 3 et 7 ne constituent pas des menaces pour la parole de Celia mais simplement des marques d'écoute qui tendent au contraire à conforter Celia dans son rôle d'énonciateur. Le terme de place transitionnelle est sans doute ici inopportun : il ne devrait s'appliquer qu'aux pauses 5 et 6, par ailleurs suivies par une réinitialisation de la ligne de déclinaison. Il nous semble néanmoins pertinent de considérer ces deux pauses au même titre que les autres dans la mesure où James s'y trouve justifié à apporter sa pierre à l'édifice conversationnel. Lorsque l'information prosodique seule est disponible, ces quatre pauses jouissent en effet d'importants scores de perception qui les placent parmi les pauses les mieux perçues.

Ces résultats, à confirmer par des tests de plus grande ampleur, indiquent au moins une incidence possible du contexte pragmatique sur la perceptibilité des pauses situées à des charnières pragmatiques suffisamment importantes pour que l'interlocuteur se manifeste vocalement. Ils nous permettent également de revenir sur notre troisième hypothèse de travail, selon laquelle mieux la pause est perçue, plus la charnière où elle se situe est importante. Nous avons montré que la charnière en question peut se situer à différents niveaux : en plus des charnières syntaxiques importantes, les pauses les mieux perçues correspondent encore à des charnières prosodiques majeures (relatives à la conduite de la ligne de déclinaison), ou à des charnières pragmatiques.

8.3.2.3 Remarque sur le contexte syntaxique des pauses d'hésitation

Nous avons vu plus haut que les pauses d'hésitation ont tendance à ne pas être repérées aussi fidèlement que les pauses de segmentation par nos sujets. Au sein des pauses d'hésitation, cependant, certaines sont davantage perçues que d'autres, indépendamment de leur durée. Il semble que leur contexte syntaxique d'apparition permette d'apporter un élément d'explication supplémentaire.

Au cours de son étude perceptive, M. Candea établit la nécessité de prendre en considération l'apparition conjointe d'une pause silencieuse et d'une marque du travail de formulation (pause pleine, allongement sur la dernière syllabe d'un mot, faux départ...). C'est en ce sens que, pour elle, pause silencieuse et marque du travail de formulation forment une marque mixte. Cette marque est à différencier de la seule pause silencieuse dans la mesure où la marque du travail de formulation tend à masquer la présence de la pause silencieuse :

> [...] la présence d'une pause silencieuse subséquente à une marque de TdF [marque du travail de formulation] **inhibe en partie la perception de cette pause silencieuse**.[45]

Cette remarque est infirmée par nos résultats lorsque les sujets ont accès à la transcription de ce qui est dit. En effet, il semble que la marque du travail de formulation constitue une infraction à la grammaticalité, ou du moins un phénomène relatif au premier niveau de la mise en mots, auquel nos sujets d'expérimentation sont plus sensibles à l'écrit.[46] Ainsi, au lieu d'être détournée des pauses, l'attention des auditeurs est comme attirée par la présence d'une marque du travail de formulation, si bien que les pauses silencieuses les accompagnant bénéficient toutes d'une perceptibilité

[45] CANDEA, Maria (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle. (p.132).

[46] Notons que les marques du travail de formulation font l'objet d'une transcription et qu'elles ont été soumises à la lecture de nos sujets d'expérimentation au même titre que tous les mots prononcés par la locutrice.

équivalente, indépendamment de leur durée. D'après nos données, la présence d'une pause pleine ou d'un faux départ sert d'indice particulièrement fiable de la présence d'une pause silencieuse. Ce phénomène explique par exemple la perception de la pause 16, qui suit un faux départ et permet à Celia de reformuler son propos autrement. Cette pause passe presque totalement inaperçue dans la configuration filtrée et reçoit les deux-tiers de son score lorsque les sujets ont accès au texte.

Ainsi, si l'on peut parler de marqueur mixte, c'est bien que la perception des pauses silencieuses ne dépend pas seulement de propriétés intrinsèques comme la durée, mais plus largement de critères qui lui sont extérieurs comme leur contexte prosodique, pragmatique et syntaxique. La perceptibilité d'une pause tient en effet partiellement à sa durée, mais également à sa fonction, à son site syntaxique d'occurrence, ainsi qu'à la conduite de la ligne de déclinaison et à l'alternance potentielle des tours de parole. Si certaines pauses sont mieux perçues que d'autres, c'est souvent parce qu'elles correspondent à des charnières plus importantes dans la conversation.

Il est apparu que la contribution des pauses silencieuses à l'appréhension de la construction du discours et de l'interaction se joue à différents niveaux et notamment, pour les pauses étudiées, au niveau de la mise en mots et de la gestion des tours de parole. Il nous reste à montrer dans des travaux à venir comment la pause de focalisation, en isolant certains éléments dans la chaîne linéaire, intervient au niveau discursif et a une fonction modale. Notre test a permis de mettre en lumière une différence importante entre perception filtrée et perception non filtrée. L'accès à l'information prosodique seule concentre l'attention des sujets sur la dimension interactionnelle de l'échange, c'est-à-dire sur notre niveau supérieur de l'édifice conversationnel. L'accès aux dimensions syntaxiques et sémantiques quant à lui l'oriente davantage, au niveau inférieur, vers la perception des groupes de sens et de ce qui peut la perturber, notamment les phénomènes relatifs au travail de formulation.

Nous sommes consciente des nombreuses objections qui peuvent être formulées au sujet de cette étude perceptive, concernant notamment la procédure, le choix de l'extrait à tester, le choix des sujets auxquels nous l'avons soumis, le dépouillement et l'interprétation des résultats. Les pistes que notre expérience ouvre nous semblent néanmoins suffisamment valables et intéressantes pour figurer ici.

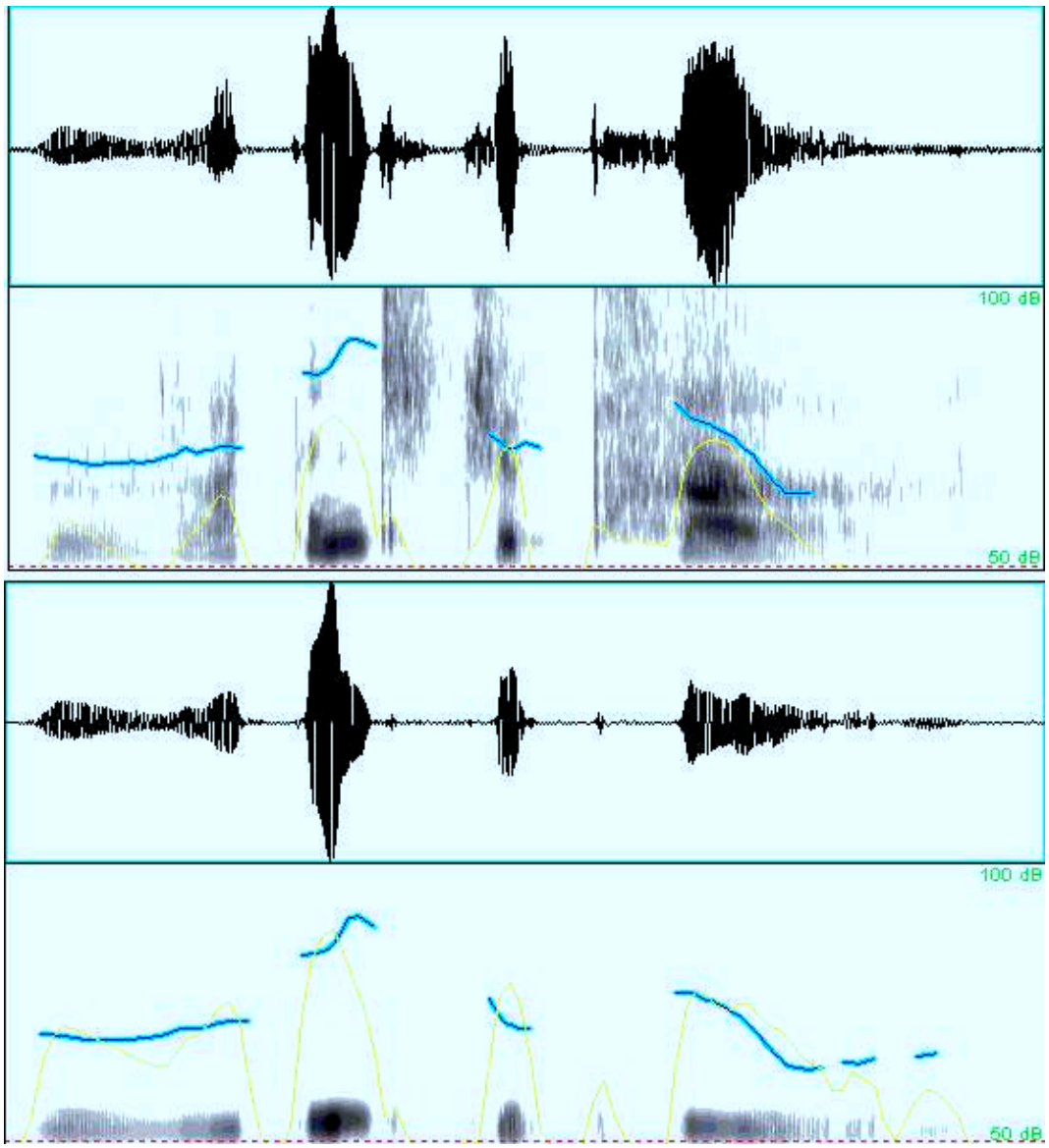

FIG. 8.2 – Comparaison des paramètres du signal filtré (en bas) et non filtré (en haut)

Chapitre 9

Aspects de la gestion des tours de parole

L'objectif de ce dernier chapitre est d'évoquer la gestion de la dimension interactionnelle du discours telle qu'elle s'illustre dans l'alternance des tours de parole. Nous indiquons au chapitre 3 que l'un des rôles de l'intensité est précisément de gérer cette dimension. Or, les places transitionnelles, qui constituent les charnières entre tours, sont généralement définies par la théorie pragmatique en termes syntaxiques. Il semble donc nécessaire d'adopter une perspective permettant d'embrasser un faisceau de paramètres.

Dans un premier temps, nous nous proposons de répertorier un certain nombre de définitions relatives à ce qu'il faut entendre par tour de parole et par place transitionnelle, et d'examiner leurs limites. Ensuite, nous nous concentrerons sur des passages de notre corpus où le relais entre deux tours semble problématique. Notre corpus présente en effet un grand nombre de chevauchements de parole et de longues pauses de segmentation. Ceux-ci constituent-ils des ratés de l'interaction ?

9.1 Définitions préliminaires et problèmes

9.1.1 Tour de parole

Comment définir un tour de parole ? Nombre d'études utilisent le tour de parole comme unité interactionnelle repère. La plupart d'entre elles utilisent des critères mécanistes, indépendants du contenu. D'autres préfèrent une définition faisant du tour de parole une unité thématique. Les deux approches, appliquées de façon stricte, posent des problèmes différents que nous souhaitons examiner ici.

9.1.1.1 Définition mécaniste

La définition du tour de parole que nous appelons mécaniste est fondée sur un principe d'alternance qui peut s'exercer à plusieurs niveaux selon les chercheurs.

Silence et parole Le premier type d'alternance envisagé concerne silence et parole. Pour Z. Harris, un tour de parole se définit par le silence qui précède et qui suit un segment produit par un locuteur donné :

> [It is] a stretch of talk, by one person, before and after which there is silence on the part of the person.[1]

Cette définition du tour de parole pose tout d'abord le problème de la définition du type de silence susceptible d'en constituer les bornes : tout silence ne semble pas apte à délimiter un tour de parole. Considérons l'exemple suivant, où Rebecca exprime sa compassion à Helen, qui se plaint de harcèlement moral :

R yeah I know but it it like weighs {0,60}
it's ev- it sov- <h> (0,30) builds up isn't it like a little sort of snowball effect *where you think oh it's harmless and*

H *yeah although I did have quite a lot* today actually so it's alright

[1]HARRIS, Zellig (1951). *Methods in Structural Linguistics*. Chicago : University of Chicago Press. (p.14).

(Helen&Rebecca : English)

Nous avons vu au chapitre précédent qu'une pause peut revêtir des fonctions différentes non exclusives les unes des autres, et que celles-ci sont partiellement corrélées avec la durée de la pause. Les pauses d'hésitation, situées de façon privilégiée en milieu de constituant ou en début de proposition ou d'énoncé, ne constituent pas des frontières de tour. C'est également le cas de la longue pause respiratoire de 0,30 s. Celle-ci entre dans une stratégie de recherche de formulation marquée par une intensité très faible et ne permet pas de déterminer, selon nous, la limite d'un nouveau tour de parole. Les pauses de segmentation, les plus longues en moyenne, sont celles qui intéressent Z. Harris. Celles-ci se situent exclusivement à une frontière syntaxique. La pause située à la fin de la première ligne dans notre transcription de la réplique de Rebecca semble ainsi apte à déterminer une frontière de tour.

La question qui se pose est ensuite celle de savoir qui est autorisé à parler après le silence. En effet, l'alternance entre silence et parole, comme nous l'avons vu dans l'exemple précédent, peut être le fait d'un seul locuteur. Les silences délimiteurs de tours de parole doivent-ils nécessairement laisser la parole à l'autre ? Les pauses de segmentation ne sont-elles que des *switching pauses* potentielles, qui le deviennent uniquement si le silence est suivi d'une véritable alternance des locuteurs ? En somme, le locuteur se tait-il toujours pour laisser parler l'autre ou peut-il enchaîner plusieurs tours de parole à la suite ?

Alternance des locuteurs Le second type d'alternance le plus couramment envisagé est celui qui voit les locuteurs parler à tour de rôle. Ainsi, J. Local *et al.* définissent un tour de parole comme toute production verbale d'un seul et même locuteur :

> The structural unit *turn* is identified in the first instance as a spate of talk by one speaker followed by a change of speaker in the clear (i.e., not in overlap).[2]

[2]LOCAL, John, WELLS, W.H.G, & SEBBA, M. (1985). Phonology for Conversation: Phonetic

Deux contraintes supplémentaires s'ajoutent à celle de l'unicité du locuteur. Tout d'abord, d'après cette définition, un tour non suivi d'une prise de parole ne semble pas concevable. Que dire donc de la dernière intervention lors d'une séquence conclusive ?

R *shall we stop there*
H *oh yeah* I think so
R <rires>
H bye
R <rires> bye

(Helen&Rebecca : New red shoes, *bye*)

Cette séquence n'est bien entendu que la mise en scène d'un rite de clôture à destination des linguistes qui transcriront et analyseront les propos tenus par les deux locutrices. Elle conserve et illustre cependant certaines caractéristiques importantes d'une véritable clôture. Elle met tout d'abord fin à la conversation entre les deux locutrices telle qu'elle nous parvient. Ensuite, elle est constituée de l'échange ritualisé de salutations réciproques prononcées à tour de rôle. D'après la définition de J. Local *et al.*, et de façon problématique nous semble-t-il, la dernière réplique de Rebecca ne saurait constituer un tour de parole puisqu'elle n'est pas suivie d'une prise de parole de Helen.

La deuxième contrainte mentionnée par J. Local *et al.* concerne le réglage temporel de l'alternance entre locuteurs. Celle-ci doit se faire, d'après les auteurs, sans chevauchement de parole. Il ne nous semble pas possible d'accepter une telle contrainte dans la définition d'un tour, dans la mesure où nombre de transitions entre prises de parole se font par tuilage.

Revenons à l'exemple précédent, issu du corpus Helen&Rebecca (épisode *English*). Que dire de la deuxième partie de la réplique de Rebecca, dont la fin est produite en même temps que les propos de Helen ? Le segment délimité à gauche par une pause de 0,60 s n'a pas pour borne de droite un silence : l'alternance entre parole et

Aspects of Turn Delimitation in London Jamaican. *Journal of Pragmatics* 9 : 309-330. (p.315).

silence n'est pas ici respectée. En outre, le passage de témoin entre les deux locutrices s'opère sur une longue séquence au cours de laquelle toutes deux parlent en même temps : l'alternance tranchée entre les locuteurs n'est pas respectée. Il ne semble pas raisonnable de dire pour autant qu'il ne s'agit pas d'une alternance de tours de parole.

9.1.1.2 Définition empirique

Si le principe d'alternance est une dimension indiscutable de la définition d'un tour de parole, il nous semble également important de pouvoir prendre en compte le contenu des propos échangés et le type de relation existant entre les interlocuteurs.

Une unité thématique Une définition possible considérerait alors qu'un tour est terminé lorsque le locuteur est arrivé au bout de son propos. L'utilité de cette définition est illustrée par l'exemple suivant :

H you know:: thingy {} mentioning no names
R yeah right <rires> yeah
H always hm {} always kind of {} acts yeah like {} it's really hilarious to {} point out I'm English every five minutes <h>
R *yeah but to:: everyone who comes*
H *so that's what she was doing*
R in there or::
H no just to me::

(Helen&Rebecca : English)

En vertu du principe d'alternance des locuteurs, et de l'alternance entre silence et parole, les deux premières répliques de Helen forment bien deux tours de parole. En revanche, une définition thématique du tour de parole ne voit qu'une seule et même unité, interrompue par les marques d'acquiescement de Rebecca. Sans impliquer que les propos de Rebecca ne constituent pas un tour de parole, nous notons que ces marques n'ont pas vocation à permettre à Rebecca de garder la parole. Elles confortent au contraire Helen dans son statut d'énonciateur.

Taille de l'unité Se pose bien entendu la question de la taille de cette unité thématique. Notre exemple précédent montre la correspondance possible entre cette unité et celle du paragraphe oral. Il illustre le découpage du paragraphe en un préambule et un rhème. Helen pose tout d'abord les éléments dont il est question dans le préambule et obtient l'assentiment de son amie, qui confirme qu'elle voit bien de qui Helen parle même si celle-ci ne veut pas donner de nom. L'unité thématique déterminée par le paragraphe n'est complète que lorsque Helen a énoncé le contenu du propos (rhème). Rebecca ne prend un statut d'énonciateur principal que lorsqu'elle demande à Helen si elle est la seule à être persécutée ainsi par sa collègue.

Dans le long extrait suivant, James n'interrompt à aucun moment la locutrice, dont les propos offrent en outre une grande cohérence thématique. Peut-il s'agir cependant d'un seul tour de parole ?

> **C** <h> {0,30} in the middle of she'd been away for about four months {0,30} five months {1,10} <h> hm {0,65} no actually longer than that six months <h> {1,20} hm {0,30}
> and she rang to say why don't I v- meet her in Goa {1,10} in ten days' time if you can imagine that <rires> {0,35}
> so <h> {0,85} I found a {0,90} hm flight {0,40} in the <h> {1,40} Sunday papers {0,65} and in fact went {0,60} met her in Goa she rang about two days later {0,65} said well I you know I'll ring you in two days and find out if you're coming <h> {1,55}
>
> (Celia&James : Goa 2)

A plusieurs reprises en effet, l'occasion est donnée à James de prendre la parole. Ce que reflètent ces points de transition possible entre locuteur, c'est bien une certaine complétude de l'unité constitutive du tour de parole, quelle que soit la nature de celle-ci.

Pour C. Goodwin, le tour de parole ne saurait se définir de façon statique dans la mesure où il s'inscrit dans le temps de l'échange verbal :

> [...] a definition of the turn as a static unit with fixed boundaries does not accurately describe its structure; rather, the turn has to be conceptualized as a

time-bound process.[3]

Cette remarque permet de concevoir le tour de parole comme une unité définie dynamiquement par la prise de parole d'un locuteur et l'arrivée à une place transitionnelle. Une définition empirique du tour de parole doit prendre en compte des données interactionnelles et thématiques aussi bien que des données syntaxiques, prosodiques et mimo-gestuelles.[4]

9.1.2 Place transitionnelle

Un tour de parole est jalonné de « places transitionnelles » qui sont autant de bornes potentielles. Ces places sont des moments où la prise de parole de L1 (le locuteur qui a la parole) peut être considérée comme achevée et donc laisser la place à la parole de L2 (l'interlocuteur). Comment définir une place transitionnelle ? Quels sont les indices permettant précisément de considérer qu'une telle place transitionnelle est atteinte ?

9.1.2.1 Caractéristiques syntaxiques

D'après H. Sacks *et al.*, une place transitionnelle correspond, en discours, au premier endroit où une fin d'unité est possible. Selon le contexte, cette unité peut être constituée d'une phrase, d'une proposition, d'un syntagme ou de toute autre construction lexicale.[5] Ainsi, cette unité n'est pas par essence prosodique mais syntaxique.

Cette conception pose le problème des segments inachevés, abandonnés, ou même interrompus par l'interlocuteur, très fréquents dans la conversation spontanée. Dans l'exemple suivant, un certain nombre de segments sont suivis d'une prise de parole

[3]GOODWIN, Charles (1981). *Conversational Organization: Interaction between Speakers and Hearers*. New York : Academic Press. (p.20).

[4]Nous ne sommes pas en mesure de tenir compte de ce dernier type de données pour l'analyse de notre corpus.

[5]SACKS, Harvey, SCHEGLOFF Emanuel A. & JEFFERSON Gail (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50 (4) : 696-735.

ou d'une cession de parole. Leur complétude syntaxique n'est pourtant pas toujours évidente.

C but this can't be exactly like practicals at school
J no but er you know if if you're sort of incompetent at one kind of practical it's <rires>
C *what you're you're bound to be in-*
J *and <rires> it spreads over*
well I mean *the second subject*
C *incompetent at others*

(Celia&James : Experiments)

La première réplique de James se termine sur un début de proposition dont une partie du prédicat manque. Cette proposition entamée avec *it's* correspond cependant à un constituant requis pour compléter la construction binaire ouverte par la proposition conditionnelle précédente. Ce constituant est interrompu par le rire de James, que Celia interprète sans doute comme un indice d'abandon de la construction. La locutrice commence en effet un nouveau tour de parole sans attendre que l'énoncé de James forme une unité syntaxique complète. La fin de cette réplique de Celia, prononcée en même temps que celle de James, fait apparaître le même phénomène. En s'achevant sur la première syllabe de l'adjectif *incompetent*, elle laisse la parole à James avant de l'interrompre à nouveau pour compléter l'adjectif, donné en entier à la fin de notre extrait.

9.1.2.2 Caractéristiques prosodiques et mimo-gestuelles

Pour S. Duncan la fin d'une unité grammaticale ne constitue qu'un indice parmi d'autres. La fin d'un tour de parole s'accompagne également de signaux de nature prosodique et mimo-gestuelle. Suivant en cela G.L. Trager & H.L. Smith, il considère la *phonemic clause* comme l'unité constitutive de base intéressant la gestion des tours de parole. La fin de cette unité s'interprète comme la fin d'un tour de parole si, par exemple, elle est marquée par une intonation montante ou descendante sur les dernières syllabes, si la syllabe finale est allongée ou le débit ralenti sur les dernières

syllabes, ou si le mouvement de la main entamé au cours du tour de parole prend fin.[6] Sur le plan mimo-gestuel, C. Kerbrat-Orecchioni mentionne également le regard soutenu sur les dernières syllabes et éventuellement un signe de tête (levée de menton ou des sourcils), ainsi qu'un « relâchement général de la tension musculaire ».[7] Sur le plan prosodique, une chute d'intensité sur les dernières syllabes ou un long silence peuvent également signaler qu'une place transitionnelle est atteinte. Nous renvoyons ici à nos remarques concernant l'effet de complétude procuré par le silence et les différents types de charnières qu'il signale.[8] *A posteriori*, la rupture de la ligne de déclinaison a également pour effet de souligner qu'un point de transition possible a été atteint, comme nous avons pu le montrer plus haut.[9]

De même que les critères syntaxiques, ces signaux prosodiques ne sont pas infaillibles. Ils ne sont pas non plus exclusifs les uns des autres mais tendent à apparaître en combinaison, favorisant ainsi le passage de relais entre locuteurs. Continuer la gesticulation des mains pendant l'émission du signal de fin de tour permet en revanche d'annuler l'effet de celui-ci.

9.1.2.3 Caractéristiques pragmatiques et interactionnelles

Une place transitionnelle peut également être suggérée par le statut illocutoire de l'énoncé. Ainsi, un énoncé interrogatif direct interprétable comme une demande d'information ponctuelle ou de développement plus large invite l'interlocuteur à prendre la parole. De même, parmi les signaux de fin de tour répertoriés par S. Duncan se trouve un ponctuant comme *you know*, pour sa valeur phatique d'appel à la participation de l'autre dans l'échange.

[6]DUNCAN, Starkey D., Jr. (1974). Some Signals and Rules for Taking Speaking Turns in Conversations. in Weitz, Shirley (éd.), *Non-verbal Communication* : 298-311. New York : Oxford University Press.

[7]KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales*. Paris : Armand Colin. (Tome 1, p.166).

[8]Voir p.373.

[9]Voir p.381.

Les paires adjacentes (*adjacency pairs*) constituent également des marques claires exigeant un changement de locuteur. On pense ainsi aux *tag questions* mais aussi aux énoncés qui font partie de rituels codés. Ainsi, dans l'exemple *bye* cité plus haut,[10] le *bye* de Rebecca fait sans ambiguïté suite au premier qui ouvre le rituel de salutation finale des locutrices et entraîne la fin de l'enregistrement. Si, comme nous l'avons fait remarquer plus haut, les interlocutrices s'adressent aussi bien l'une à l'autre qu'aux linguistes qui écouteront leur conversation, leur échange est conforme à un rituel où il est attendu que chacune des parties prenantes intervienne à tour de rôle et organise donc l'alternance des tours de parole. Notons que les deux premières répliques de notre extrait sont partiellement prononcées de façon simultanée, mais que le couple *bye/bye* se répond sans chevauchement.

En somme, le tour de parole est une unité aux frontières mouvantes gérées localement, dont une définition doit pouvoir comprendre des critères syntaxiques, mais aussi prosodiques et mimo-gestuels, et avant tout interactionnels.

9.2 Le système de l'alternance des tours à l'épreuve de la conversation spontanée

L'alternance des tours de parole admet des régularités qu'il est possible d'exprimer sous forme de règles et qui nous permettent de continuer d'affiner notre définition de ce qu'est l'unité tour de parole. Théoriques, ces règles sont fréquemment mises à mal par la conversation spontanée, et notamment par les nombreux chevauchements de parole qui la caractérisent.

[10] Voir p.390.

9.2.1 Règles d'alternance

Construite dans l'interaction, le tour de parole est une unité à géométrie variable. Certains phénomènes contribuent à assurer des transitions fluides entre deux tours de parole, en minimisant d'une part les silences et, d'autre part, en évitant les chevauchements de parole. Nous rapportons ici essentiellement les travaux de H. Sacks *et al.*, présentés dans leur article de 1974.[11]

9.2.1.1 Anticipation des places transitionnelles

La souplesse du passage de relais entre locuteurs est favorisée par l'anticipation du moment où une telle transition est possible.

> A speaker can talk in such a way as to permit projection of possible completion to be made from his talk, from its start, allowing others to use its transition places to start talk, to pass up talk, to affect directions of talk etc.; and [...] their starting to talk, if properly placed, can determine where he ought to stop talk. That is, the turn as a unit is interactively determined.[12]

D'après H. Sacks *et al.*, il est du ressort du locuteur en place (L1) de faciliter la « lecture » des limites potentielles de son tour de parole. D'un point de vue interactionnel, la synchronisation de la prise de parole par L2 dans la fenêtre temporelle ouverte par la place transitionnelle prévue par L1 détermine la fin du tour de parole de L1.

L'exemple suivant montre que la projection de la fin du tour de parole de L1 est facilitée notamment par des critères syntaxiques et par des critères intonatifs :

R it's only twenty minutes away as well isn't it
H what
R Montpellier

(Helen&Rebecca : Nîmes 2)

[11] SACKS, Harvey, SCHEGLOFF, Emanuel A., JEFFERSON, Gail (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50 (4) : 696-735. (p.706).
[12] *ibid.* (pp.726-727).

L'enchaînement des trois tours de parole est ici très souple : les transitions sont caractérisées par un temps de silence réduit (environ 0,15 et 0,40 s), et par une absence de chevauchements de parole. Les deux énoncés qui nous intéressent constituent des demandes d'information. Ils contiennent donc en eux le principe de l'alternance des tours.

La projection de la fin du premier tour de parole de Rebecca est facilitée par l'utilisation du *question tag*. Situé en position postrhématique, un *question tag* constitue un signe quasiment infaillible non seulement que le passage de témoin est souhaité par le locuteur en place mais surtout que la fin de son tour est imminente. Il permet en outre de confirmer l'intuition que L2 a pu former au cours du tour de L1 quant à sa borne de droite. Il est à penser en effet que la très courte durée du silence séparant les deux tours est liée au fait que d'autres places transitionnelles potentielles ont été perçues par Helen, notamment après *away* ou plus probablement après *as well* qui ponctue un énoncé syntaxiquement complet. Le temps de préparation de son interrogation, de l'ordre de 0,40 à 0,85 s, correspondrait alors davantage à des durées conformes aux contraintes de la cognition.[13]

La réplique finale de Rebecca suit un énoncé monosyllabique dont la Figure 9.1 présente les caractéristiques prosodiques. Il s'agit d'une interrogative prononcée en écho à l'énoncé précédent de Rebecca, dont l'objectif est d'obtenir une précision de sa part. D'un point de vue syntaxique, le pronom interrogatif est susceptible d'introduire un énoncé bien plus long, du type *what is only twenty minutes away* ou *what are you talking about*. Il apparaît cependant que Rebecca anticipe correctement le caractère elliptique de l'énoncé. Ce phénomène s'explique à nos yeux par la modulation mélodique particulièrement ample portée par le pronom, à laquelle s'ajoute la pause dont l'interprétation la plus probable devient donc celle d'une *switching pause*.

[13]Voir les seuils retenus dans notre expérience perceptive concernant les pauses silencieuses : Section 8.2.1.4, p.356.

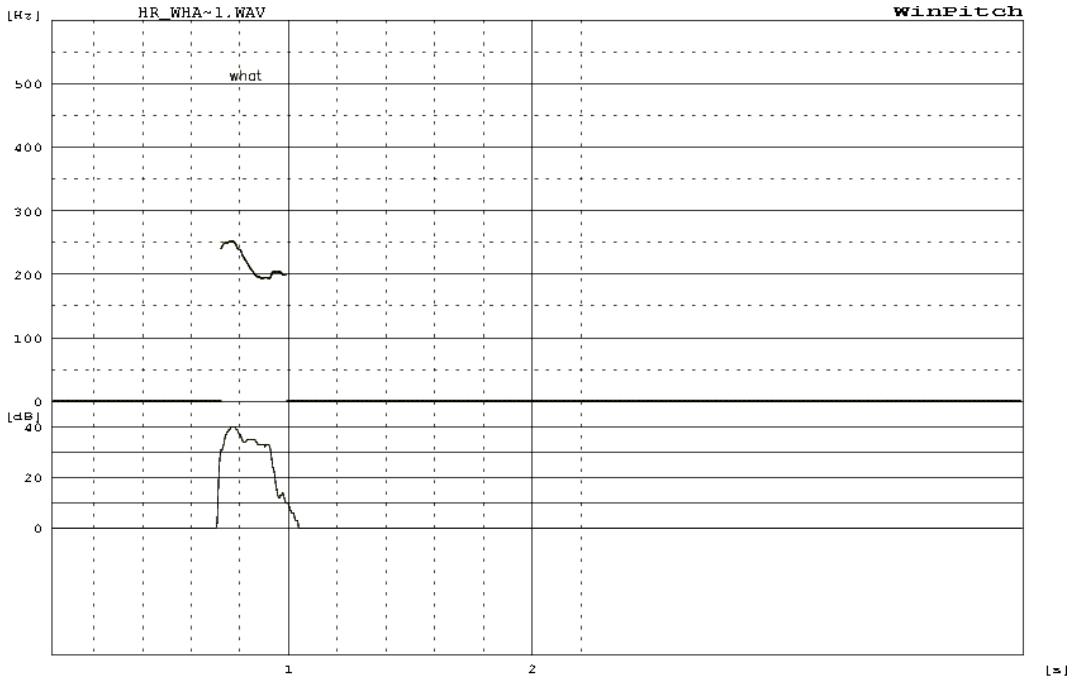

FIG. 9.1 – Modulation mélodique sur le pronom interrogatif

L'arrivée à une place transitionnelle engage L1 à céder la place à son interlocuteur si celui-ci profite de l'occasion qui lui est donnée pour prendre la parole. Si L2 ne se manifeste pas ou s'il ne le fait pas à l'endroit prévu, le droit de parole de L1 est renouvelé jusqu'à la prochaine place transitionnelle. Celle-ci reste, pour H. Sacks *et al.*, le seul moment où le passage de témoin entre deux locuteurs est possible :

> First, the system allocates single turns to single speakers; any speaker gets, with the turn, exclusive rights to talk to the first possible completion of an initial instance of a unit-type — rights which are renewable for single next instances of a unit-type [...]. Second, all turn-transfer is coördinated around transition-relevance places, which are themselves determined by possible completion points for instances of the unit-types.[14]

L'exemple suivant illustre le principe de reconduction du droit à la parole lorsque L2 ne profite pas de l'occasion que lui fournit l'apparition d'une place transitionnelle

[14]SACKS, Harvey, SCHEGLOFF, Emanuel A. & JEFFERSON, Gail (1974). *op. cit.* (p.706).

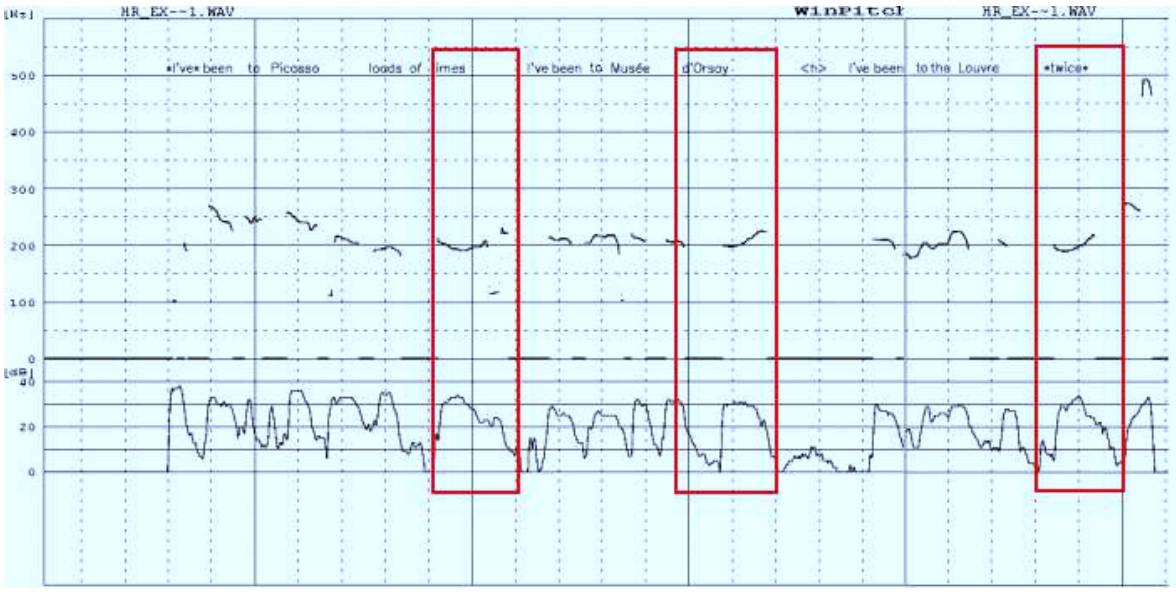

FIG. 9.2 – Signes d'une stratégie de conservation de la parole

pour prendre la parole à son tour.

H *I've* been to Picasso loads of times
I've been to Musée d'Orsay <h> {0,45}
I've been to the Louvre *twice*

R *(?)when* did you use the(?) Musée d'Orsay

(Helen&Rebecca : Museums)

Trois places transitionnelles émaillent le discours de Helen. Les deux premiers énoncés s'enchaînent sans aucune pause, mais les deux autres sont séparés d'une pause de près d'une demi-seconde. C'est seulement à la dernière occasion que Rebecca utilise son droit à la parole. L'énumération de Helen participe d'une stratégie de conservation de la parole qui s'accompagne, sur le plan suprasegmental, par des montées mélodiques et par le maintien de l'intensité en finale des énoncés. C'est ce que montre la Figure 9.2.

Le fonctionnement des paramètres intonatifs est ici pertinent au niveau interactionnel de la gestion des tours de parole. Il concerne le code de la production de C. Gussenhoven,[15] selon lequel, du fait de la moindre pression sous-glottique en fin

[15]CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002*

d'énoncé, fréquence et intensité diminuent. Le maintien d'une intensité forte en fin d'énoncé, de même qu'une mélodie montante, ont une valeur continuative.

9.2.1.2 Anticipation de l'identité du locuteur suivant

Dans le cas de la conversation à deux, la question de l'identité du prochain locuteur se pose de façon moins cruciale que lorsque de nombreux candidats à la prise de parole sont présents. H. Sacks *et al.* montrent que l'alternance des prises de parole répond néanmoins à une série de règles précises concernant la désignation de L2. Les règles que nous rappelons ici contribuent à la fluidité de la transition entre locuteurs, et sont valables quel que soit le nombre de participants engagés dans la conversation :

> (1) For any turn, at the initial transition-relevance place of an initial turn-constructional unit:
>
> (a) If the turn-so-far is so constructed as to involve the use of 'current speaker selects next' technique, then the party so selected has the right and is obliged to take next turn to speak; no others have such rights or obligations, and transfer occurs at the place.
>
> (b) If the turn-so-far is so constructed as not to involve the use of 'current speaker selects next' technique, then self-selection for next-speakership may, but need not, be instituted; first started acquires right to a turn, and transfer occurs at that place.
>
> (c) If the turn-so-far is so constructed as not to involve the use of 'current speaker selects next' technique, the current speaker may, but need not continue, unless another self-selects.
>
> (2) If, at the initial transition-relevance place of an initial turn-constructional unit, neither 1a or 1b has operated, and, following the provision of 1c, current speaker has continued, thus the rule-set a-c re-applies at the next transition-relevance place, and recursively at each next transition-relevance place, until

conference : 215-218.

transfer is affected.[16]

La plupart du temps, l'identité du locuteur suivant n'est pas inscrite dans le contenu ou la forme du propos tenu par L1. Le cas le plus courant dans la conversation spontanée est couvert par la règle 1c. Une interrogative directe invitant l'interlocuteur à prendre la parole constitue un exemple de désignation du locuteur suivant, conformément à la règle 1a. Cette règle s'illustre dans l'exemple suivant où la question de Helen engage Rebecca à prendre la parole :

H where have you been

(Helen&Rebecca : Museums)

Désignée par l'emphatique *you*, Rebecca énumère à son tour les musées parisiens qu'elle a visités.

Les règles développées par H. Sacks *et al.* décrivent ce qu'ils considèrent être le fonctionnement idéal d'une conversation, dans laquelle l'alternance des tours de parole n'est que rarement perturbée par des silences ou des chevauchements de parole. Ce système admet néanmoins quelques « imperfections ».

9.2.2 Alternance, silence et simultanéité de parole

Dans quelle mesure ces règles sont-elles mises à l'épreuve par la conversation spontanée telle qu'elle s'illustre dans notre corpus, et notamment par le phénomène récurrent des chevauchements de parole ?

9.2.2.1 Assouplissement des règles

L'alternance parfaite entre deux états ne devrait pas admettre de situation intermédiaire où aucun des deux états n'est vérifié ou encore où les deux états sont activés en même temps. Ce principe admet cependant des seuils de tolérance qui rend possible un assouplissement de cette règle :

[16]SACKS, Harvey, SCHEGLOFF, Emanuel A. & JEFFERSON, Gail (1974). *art. cit.* (p.704).

> Transitions (from one turn to a next) with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up the vast majority of transitions. The components and the rule-set, in organizing transfer exclusively around transition-relevance places, provide for the possibility of transition with no gap and no overlap.[17]

Pour H. Sacks *et al.*, les silences et chevauchements de parole, fréquents en conversation spontanée, sont donc admissibles s'ils restent de courte durée. Nous chercherons à déterminer plus loin à quelle fréquence nous rencontrons des chevauchements de parole de parole dans notre corpus, et pour quelle durée.

9.2.2.2 Accordage temporel des prises de parole

Alternances fluides, longs silences et chevauchements de parole relèvent de l'accordage temporel des prises de parole. Les circonstances de la prise de parole tardive ou précoce de L2 sont décrites dans les termes suivants par C. Kerbrat-Orecchioni :

> Le successeur peut prendre la parole :
>
> (1) trop tard,
>
> – soit que le signal de tour ait été mal perçu [...]
>
> – soit que le(s) successeur(s) potentiel(s) n'ai(en)t pas le désir ou les moyens, de produire l'enchaînement requis ;
>
> (2) trop tôt,
>
> – soit que L_2 ait cru percevoir une fin de tour qui n'était pas en fait programmée par L_1
>
> – soit qu'il s'empare de la parole en connaissance de cause (i.e. : tout en sachant pertinemment que L_1 n'a pas terminé son tour).[18]

[17] *ibid.* (p.708).

[18] KERBRAT-ORECCHIONI, Catherine (1990, 1998).*Les interactions verbales.* Paris : Armand Colin. (Tome 1, pp.172-173).

Notre attention se porte essentiellement sur les prises de parole précoces engendrant des recouvrements de parole. L'exemple suivant illustre une fin de tour dont l'anticipation de la fin est problématique :

R do you know what that tray reminds me *of*
H *what*
R my pottery lessons

(Helen&Rebecca : Art)

L'interrogation de Helen intervient après un signal de fin de tour prosodique. Comme nous pouvons le voir sur la courbe présentée à la Figure 9.3, l'allongement sur *me*, d'une durée légèrement supérieure à *tray*, la chute conjointe de l'intensité et de la fréquence fondamentale invitent à interpréter *me* comme une frontière d'énoncé. Rebecca, par souci de correction grammaticale, allonge son énoncé en y ajoutant la préposition *of* qu'elle n'avait pas prévue dans son projet initial. Après une pause inférieure à notre seuil des 0,20 s, celle-ci a pour effet d'étendre la durée de son tour de parole alors que la fin de celui-ci a déjà été annoncée. Elle se voit donc recouverte par l'interrogation de Helen.

L'interruption volontaire intervient en l'absence de signaux de fin de tour, comme dans l'exemple suivant où les deux locutrices s'interrogent sur les utilisateurs habituels du bureau dans lequel l'enregistrement se déroule :

R I reckon this is the <épelle en français> D I F S little study area
they have their own *room that they can come and use*
H *it is cos it's got a* the sign on the door

(Helen&Rebecca : Staying awake till 7, *cos_038*)

La réplique de Helen, qui recouvre la seconde moitié de l'énoncé de Rebecca, ne correspond pas à une place transitionnelle. De façon tardive, Helen apporte un élément permettant de confirmer l'affirmation de son amie selon qui elles se trouvent dans le bureau du DIFS. Cette intervention constitue une interruption volontaire du tour de Rebecca, qui, selon nous, ne menace pas la face de celle-ci.

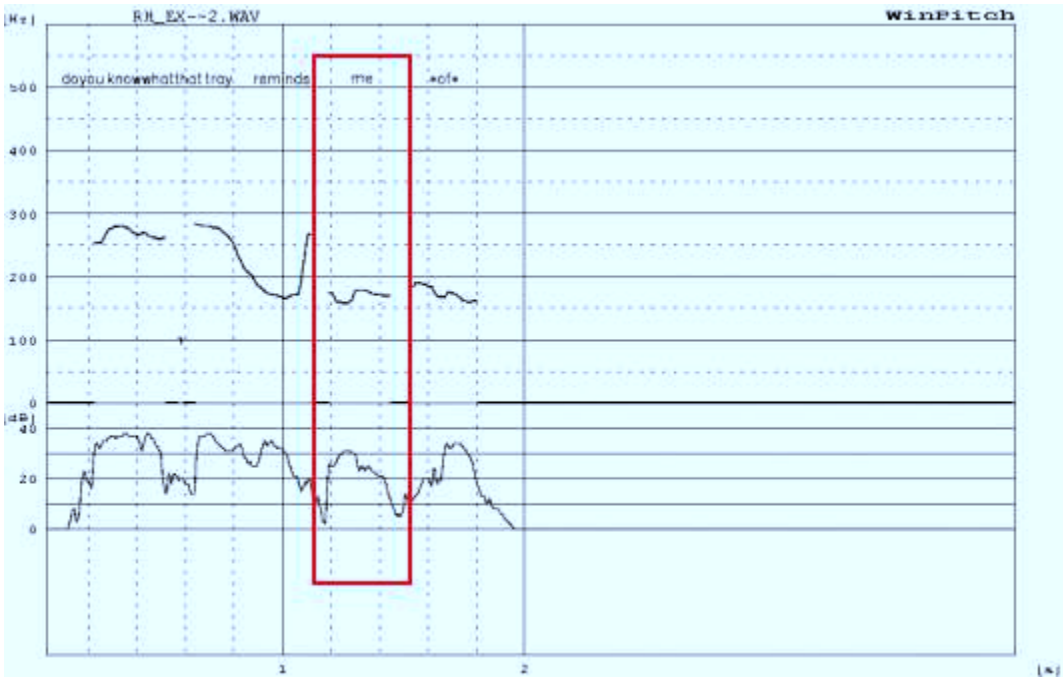

FIG. 9.3 – Signaux de fin de tour

C'est la caractéristique perturbatrice de cette occupation simultanée du canal qui nous intéresse ici. Admis par la théorie de H. Sacks *et al.*, les chevauchements de parole témoignent pour S. Duncan d'un échec de la gestion des tours de parole :

> [...] the turn-taking mechanism may be said to have broken down, or perhaps to have been discarded, for the duration of that state[19]

Après une évaluation de la fréquence du phénomène dans un extrait de notre corpus, l'analyse des segments concernés nous permettra de formuler des hypothèses concernant les facteurs favorisant un chevauchement de parole. Enfin, il s'agira de déterminer quelles sont les conséquences d'un tel phénomène sur les plans syntaxique, discursif et interactionnel, ainsi que sur celui de la relation inter-sujets. Les chevauchements de parole constituent-ils des ratés de l'interaction ?

[19]DUNCAN, Starkey D., Jr. (1974). Some Signals and Rules for Taking Speaking Turns in Conversations. in Weitz, Shirley (éd.), *Non-verbal Communication* : 298-311. New York : Oxford University Press. (p.302).

9.3 Les chevauchements de parole : des ratés de l'interaction ?

Afin de répondre à ces questions, nous proposons d'étudier en détail les chevauchements de parole dans un extrait de notre corpus, dont l'analyse permettra d'éclairer le fonctionnement de l'alternance des tours de parole. Il s'agit de caractériser les chevauchements de parole de notre corpus et d'évaluer leur impact sur la construction du discours et de l'interaction.

9.3.1 Caractéristiques générales du phénomène

9.3.1.1 Présentation de l'extrait de corpus analysé

L'extrait de corpus que nous avons choisi d'analyser est situé autour de la dixième minute de conversation entre Celia et James. Il s'agit de l'épisode que nous avons intitulé *Travel*. Long d'un peu plus de 2 minutes et 20 secondes, il met en lumière certaines caractéristiques de la totalité du corpus Celia&James quant à la gestion des tours de parole. Pris dans un débat souvent très animé, les locuteurs s'interrompent l'un l'autre fréquemment, ce qui occasionne de nombreux chevauchements, répertoriés dans le Tableau 9.1. L'extrait contient également des passages où les tours de parole sont beaucoup moins disputés, laissant apparaître de longues pauses aux points de transition. En dépit de la différence de statut entre la toute puissante secrétaire du laboratoire de phonétique et un étudiant venu lui demander un service, les règles d'alternance des tours de parole telles que définies par H. Sacks *et al.* n'y sont donc pas strictement respectées.

Après la discussion à propos des locaux partagés par les étudiants de phonétique et les étudiants de chinois (dans l'épisode *Facilities*), James commence à décrire le cursus suivi par ces derniers (dont sa sœur Sally fait partie), et le compare à celui des

	L1		L2	
[ex.9.1]	**J**	which is	**C**	mm
[ex.9.2]	**J**	and a lot of the	**C**	mm
[ex.9.3]	**C**	bear	**J**	er
[ex.9.4]	**C**	<rires>	**J**	and then my er
[ex.9.5]	**J**	why she's	**C**	mm
[ex.9.6]	**C**	-er wouldn't it	**J**	yeah
[ex.9.7]	**J**	no	**C**	but she hasn't
[ex.9.8]	**J**	you know like	**C**	oh ri-
[ex.9.9]	**J**	you know	**C**	mm
[ex.9.10]	**C**	-er one	**J**	yeah
[ex.9.11]	**J**	Chinese	**C**	but who's not
[ex.9.12]	**C**	Chinese	**J**	but
[ex.9.13]	**J**	sort of hard work	**C**	mm
[ex.9.14]	**J**	quite hard work	**C**	it's a different
[ex.9.15]	**C**	altogether	**J**	so there's three
[ex.9.16]	**J**	five months	**C**	mm

TAB. 9.1 – Segments produits en situation de chevauchement de parole dans l'épisode *Travel*

étudiants d'autres langues (épisode *China*). Juste avant l'épisode qui nous intéresse, James mentionne les voyages de sa famille en Chine. Dans le cadre de ses études, Sally y a passé une année pour notamment suivre des cours intensifs. Elle y retourne très prochainement pour faire du tourisme avec sa plus jeune sœur, Elspeth, qui vient également d'y passer une année à enseigner l'anglais. Ses parents, quant à eux, rentrent tout juste d'un voyage au cours duquel ils ont pu rendre visite à Elspeth. C'est donc principalement James qui occupe la parole. La position de Celia est celle de la personne qui écoute, acquiesce, s'étonne et demande des explications supplémentaires.

9.3.1.2 Fréquence d'apparition

Dans le corpus Celia&James Le corpus Celia&James dans son ensemble comporte près de 500 segments produits en situation de chevauchement de parole (250 par locuteur). Un locuteur voit l'intégralité ou une partie son discours recouvert par celui de l'autre approximativement toutes les 7 à 8 secondes. Si les segments concernés sont de durée et de nature différentes, il apparaît qu'environ 15% des mots prononcés sur l'ensemble de notre corpus le sont en recouvrement.

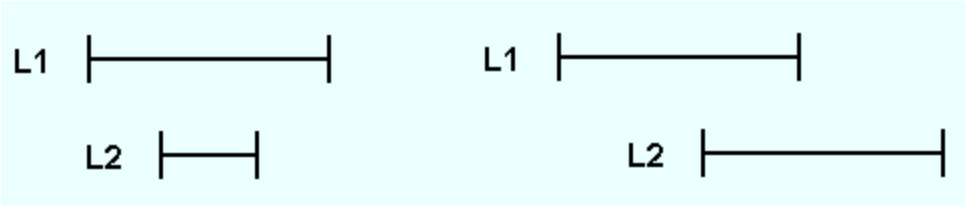

FIG. 9.4 – Configurations des recouvrements de parole : L1> (à gauche) ou L2> (à droite)

Dans l'épisode *Travel* Les chiffres sont sans doute plus parlants à l'échelle du seul extrait de la conversation entre Celia et James. Comme le montre le Tableau 9.1, le nombre de segments produits simultanément se monte à 16 par locuteur. La fréquence d'occurrence est très proche de celle calculée pour l'ensemble de la conversation. Une partie non négligeable du temps de parole total des locuteurs dans l'épisode est occupée par les deux locuteurs en même temps. On note cependant une grande disparité entre James et Celia : celle-ci partage près de la moitié de son temps de parole avec son interlocuteur, contre près de 10% pour James. Cette tendance se retrouve lorsque l'on s'intéresse à la proportion de mots prononcés en recouvrement dans le discours des locuteurs (8% pour James, contre 38% pour Celia, et 12% en moyenne).

Ces quelques calculs rapides tendent à montrer que les chevauchements de parole sont loin d'être un phénomène marginal dans la conversation spontanée telle qu'elle s'illustre dans notre corpus.

9.3.1.3 Deux configurations

Nous relevons deux configurations possibles pour décrire les prises de parole simultanées, celles-ci étant très rarement simultanées du début à la fin. La première configuration, illustrée par le schéma de gauche de la Figure 9.4, voit les propos du locuteur en place (L1) recouverts partiellement par ceux de L2. C'est L1 qui conserve cependant la parole (d'où notre notation : L1>). Dans le second cas, la prise de parole de L2 a pour effet de faire taire L1 (L2>).

Les deux configurations sont également représentées dans notre extrait de corpus. Il apparaît que le locuteur en place avant le recouvrement a 50% de chances de conserver la parole à l'issue d'un recouvrement de parole.

Le déséquilibre de temps de parole entre les deux locuteurs est reflété par le nombre de fois où ils gardent ou prennent la parole dans ces conditions. La configuration L1> voit très largement James dans le rôle du locuteur en place avant et après le chevauchement. Dans la seconde configuration, les rôles sont mieux répartis puisque Celia parvient plusieurs fois à conserver la parole à l'issue du chevauchement.

9.3.2 Un phénomène perturbateur ?

Les chevauchements de parole sont potentiellement perturbateurs à plusieurs titres. En effet, lorsque deux locuteurs parlent simultanément, l'écoute de l'autre mais également l'écoute de soi-même est rendue difficile. Considérons tout d'abord dans quelle mesure les plans syntaxique, sémantique et informationnel peuvent s'en trouver affectés.

9.3.2.1 Durée des segments recouverts

Les segments concernés sont de courte durée. Dans notre extrait, la durée moyenne des segments chevauchés se monte à près de 0,7 s et correspond à environ 2,5 syllabes. Le segment le plus court dure 0,3 s (C9.3 : *bear*/J9.3 : *er*), contre plus de 1,3 s pour le plus long (C9.6 : *-sier wouldn't it*/J9.6 : *yeah*). Le nombre de mots contenus dans les segments prononcés en situation de recouvrement dans notre extrait varie de 1 à 5, avec une moyenne autour de 2,2.

L'exemple suivant montre que le corpus Celia&James voit certains chevauchements de parole durer bien au-delà de ce qui s'illustre dans notre extrait. La durée du chevauchement se monte ici à plus d'une seconde et demie :

J *they theirs is an* undergraduates' subject so they have

C yes ***so they:: they do have quite a lot***

J ***well I've (?) met about fifteen a yea::r*** over four years or something <h> {0,50}
I guess *in total* number of people perhaps we're even

C *mm*

(Celia&James : China)

D'une manière générale, cependant, les segments concernés se situent en dessous de dix syllabes et en moyenne autour de 2 à 3 syllabes. La courte durée des chevauchements de parole tend à relativiser leur caractère perturbateur.

9.3.2.2 Caractéristiques syntaxiques

Si, comme nous venons de le voir, un énoncé entier peut être prononcé en situation de chevauchement de parole, les segments recouverts correspondent rarement à des groupes syntaxiques complets. Citons simplement les deux exemples suivants (J9.2 et C9.12) :

J [ex.J9.2] *and **a lot of the*** gappers' parents have <h> {0,70} gone out to visit them {0,95}

C *but **who's not*** studying [ex.C9.12] *Chinese*

(Celia&James : Travel)

Ces exemples montrent que la structure syntaxique n'est pas forcément perturbée par le chevauchement. En dehors de toute situation de ce type, l'intégrité des syntagmes n'est pas systématiquement conservée en anglais, contrairement au français. Le chevauchement de parole n'est donc pas particulièrement perturbateur à cet égard.

Cependant, le recouvrement des propos de L1 par L2 provoque parfois une difficulté de formulation chez celui qui conserve ou obtient la parole à l'issue du chevauchement. C'est le cas dans l'exemple J9.4 où le début de l'énoncé de James est recouvert par le rire de Celia. James abandonne en effet la construction du groupe nominal amorcée avec l'adjectif possessif. Une fois que les rires de Celia ont cessé, il reprend entièrement la construction du groupe syntaxique. Notons également que ce faux départ permettant la réinitialisation de la structure syntaxique est précédée par une pause pleine,

marque du travail de formulation :

J [ex.J9.4] *and then **my er* my sister** got a travel grant from the university to go to China

(Celia&James : Travel)

Sur le plan syntaxique, les perturbations engendrées par les chevauchements de parole restent donc limitées.

9.3.2.3 Caractéristiques sémantiques

Difficultés de compréhension limitées En situation de recouvrement de parole, la conversation est susceptible de devenir quelque peu cacophonique. La compréhension du message peut être rendue plus difficile. D'une part, L1, qui voit ses paroles recouvertes, tend à exercer un contrôle moindre sur l'émission de son discours. Ce phénomène nous semble dû au fait qu'il soit potentiellement gêné par l'intervention de L2 mais également par le fait qu'il doive exercer la double activité de celui qui parle et celui qui écoute de façon simultanée. Cette remarque est valable également pour L2.[20]

Faible sémantisme des segments recouverts Sur le plan sémantique, nous remarquons cependant que les éléments recouverts ont la plupart du temps un apport informationnel faible. C'est le cas des segments prononcés par L1 autant que ceux de L2.

[20]Dans notre transcription, ce phénomène se lit dans nos difficultés passagères, matérialisées par des points d'interrogation, à identifier le sens de certains segments chevauchés. Cette difficulté est sans doute due au fait que deux signaux de parole s'entremêlent. Plus essentiellement nous semble-t-il, elle est une conséquence de ce phénomène de double activité cognitive qui occasionne des prononciations moins soignées.

J yeah Tsing Tao well it's one of the *best known*
C *mm*
J {0,20} Chinese beers <h> {0,70} er *y- you get in (?)numerous(?) country*
C *m- so what's the food like*

(Celia&James : Beer)

Dans cet exemple, *numerous* nous semble douteux. D'un point de vue grammatical, l'adjectif entre en outre en contradiction avec le nombre du nom qu'il qualifie (*country* et non *countries*). Notons néanmoins que ces passages douteux en situation de recouvrement de parole sont rares dans notre corpus, ce qui tend encore à relativiser l'impact perturbateur des chevauchements.

Nombre de mots grammaticaux, et notamment de pronoms, figurent en effet dans les segments produits en situation de chevauchement de parole. Mentionnons les exemples J9.1 (*which is*) et C9.7 (*but she hasn't*), qui constituent des reprises anaphoriques d'éléments au contenu sémantique déjà posé. Dans l'exemple J9.1, James amorce une relative dont l'antécédent est l'énoncé précédent, tandis que le pronom sujet utilisé par Celia a pour référent la sœur de James dont il est question depuis près d'une minute (depuis la fin de l'épisode *China*).

De même, la référence de certains mots lexicaux recouverts se retrouve dans l'ensemble des connaissances partagées établi par les deux locuteurs. C'est le cas de *Chinese* dans le passage suivant :

J *yeah* she's the young one who's in China on a gap year but not studying [ex.J9.11] ***Chinese***

C *but who's not* studying [ex.C9.12] ***Chinese***

J *but* she's <rires>

C but she's learnt these phrases anyway

(Celia&James : Travel)

La langue chinoise parcourt toute cette partie de la conversation. James dit que sa sœur maîtrise le vocabulaire de base qui lui sert notamment à marchander dans les supermarchés. Le terme *Chinese* apparaît également plusieurs fois dans le proche contexte avant :

J s- since she actually speaks **Chinese** that's really quite helpful <h>

et, plus loin,

C *but she hasn't* learnt any **Chinese**

(Celia&James : Travel)

Ainsi, bien que recouvert, le terme n'a pas besoin d'être redit pour être compris parfaitement par l'interlocuteur.

Nous remarquons également en deux points dans notre extrait de corpus que le cœur du mot tend à être épargné par le chevauchement. Dans la remarque de Celia en C9.6 (*easi*er wouldn't it**), de même qu'en C9.10 (*young*er one**), James ne

recommence à parler qu'après l'énonciation de la racine sur laquelle le suffixe du comparatif se greffe.[21] Dans les deux cas, la quantité d'information des éléments recouverts est très faible : un *tag question* occupant le postrhème et un pronom anaphorique. L'identification du sens des propos de Celia n'est donc pas rendue problématique par le chevauchement.

Dans certains cas sur lesquels nous allons revenir, les locuteurs jugent nécessaire de répéter certains termes recouverts. Il apparaît cependant généralement que le chevauchement de parole ne brouille pas de façon rédhibitoire la compréhension du message.

9.3.3 Un ciment dans l'édifice conversationnel

Les conséquences perturbatrices des chevauchements de parole sont à nos yeux négligeables par rapport au rôle que ceux-ci jouent dans la construction de l'édifice conversationnel, sur les plans discursif, intersubjectif et interactionnel. Les situations de parole simultanée illustrent également certains principes fondamentaux de la gestion de l'interaction.

9.3.3.1 Sur le plan de l'organisation discursive

Concurrence des projets discursifs Les chevauchements de parole peuvent témoigner de projets discursifs concurrents. Dans le passage suivant, Celia et James comparent les pays d'Asie du point de vue de leurs attraits touristiques. Le propos roule autour de la Chine, que Celia trouve plus effrayante que les autres pays. James commence un énoncé abondant dans le sens de son interlocutrice, mais l'interrompt en cours de route. Après une longue pause, les deux locuteurs prennent la parole simultanément et proposent chacun d'orienter la conversation en suggérant un thème différent. Si les deux locuteurs proposent de caractériser la Chine par comparaison avec

[21] Si notre unité de référence est celle du mot, il est nécessaire dans ces deux cas de considérer la pertinence d'une unité de taille inférieure.

d'autres pays, James adopte une perspective politique abstraite (libéralisme contre communisme) tandis que Celia adopte une perspective linguistique et préfère citer un exemple concret illustrant en creux le cas de la Chine :

C but China I've never I m- China's a bit scarier
J well that's what I was about to say it's such a {0,70}
and in a liberal country I mean
C *and of course India {0,25} I mean* India English is there's no language problem at all

(Celia&James : Goa 2, Languages)

Au terme du chevauchement de parole, James a abandonné son projet au profit de celui de Celia. Celui-ci s'en trouve nous semble-t-il consolidé par la bataille par laquelle il s'est imposé. James le prend rapidement à son compte et complète le raisonnement de Celia en appliquant lui-même l'argument linguistique à la Chine :

C *though* China is is a different *kettle of fish altogether*
J *yes because you* {0,50} most people don't speak English

(Celia&James : Languages)

Remarquons que le thème linguistique permet ensuite à James de réorienter la conversation sur l'apprentissage de l'anglais en Chine tel que sa sœur a pu le lui décrire. Il semble donc que, loin d'être mise en péril par des projets concurrents, l'organisation discursive se trouve resserrée par le fait qu'un nouveau thème s'impose au terme d'une lutte pour la parole. Le chevauchement de parole semble constituer un test pour la viabilité d'un nouveau thème.

Emergence d'un nouveau thème Nous avons vu plus haut que l'impact des recouvrements de parole sur la compréhension du contenu sémantique du message était relativement limité, du fait de la faible quantité d'information véhiculée par les termes concernés. Dans certains cas, cependant, les mots recouverts sont porteurs d'un sémantisme d'autant plus fort qu'ils constituent un nouveau thème vers lequel s'oriente la conversation. C'est le cas, dans l'exemple précédent, du mot *India*. Dans un nouvel exemple, James répond à la question de Celia concernant le lien existant

entre la Chine et la famille de James (dans l'épisode *China*) en évoquant d'abord le cas de ses deux sœurs. A la charnière entre cet épisode et celui qui nous intéresse, il enchaîne sur le cas de ses parents. Le segment *and my parents* est cependant recouvert par les paroles de Celia. James répond alors à la réflexion de Celia sur les propos tenus précédemment et reprend le fil de son discours en posant à nouveau le thème qu'il a tenté d'introduire plus haut :

J *and **my parents***
C *well she probably* heard a lot about it
J yeah I ff- I {0,30} possibly I don't know if that had much to do with it
and **my parents** <h> {1,00} had the idea of going on holiday there partly to visit Elspeth while she was out there

(Celia&James : China, Travel)

La deuxième mention de *my parents* s'enchaîne immédiatement avec ce qui précède. Notons ici le fonctionnement de l'intensité qui souligne le groupe nominal et de la longue pause qui le suit. Les deux paramètres nous semblent opérer à deux niveaux. D'une part, un surcroît d'intensité tend à mettre en relief un segment sur le plan thématique, tandis que la pause contribue à unifier ce segment et créer une attente pour ce qui suit. D'autre part, sur le plan interactionnel, le surcroît d'intensité tend à prévenir toute tentative de prise de parole par Celia, en dépit de la longue pause qui suit et qui pourrait suggérer qu'une place transitionnelle a été atteinte.

En ce sens, les chevauchements de parole, loin d'être nuisibles, contribuent au dynamisme récursif de la conversation dont nous avons pu voir qu'il constituait un principe essentiel de la construction du discours à l'oral spontané.[22]

9.3.3.2 Sur le plan intersubjectif

Sur le plan intersubjectif, les chevauchements de parole sont susceptibles de créer une cacophonie nuisible à la construction du consensus. Dans quelle mesure ce phénomène se vérifie-t-il dans nos données ?

[22] Voir Section 6.3, p.300.

Construction du consensus fragilisée Dans le passage suivant, James peine à démontrer à Celia que le cursus de chinois à Oxford est particulièrement exigeant. Le groupe nominal *hard work* revient par trois fois dans la bouche de James, dont deux en situation de recouvrement de parole :

J so **it's**
C *mm*
J ***sort of hard work*** <**h**> {**0,45**} **all the time** so it's really {0,55} **it was really *quite hard work***
C *it's a different* scene for them *altogether*
J *so there's three* years two terms of Oxon plus <h> {0,60} another
C *mm*
J *five months* of intensive instruction in Taiwan actually it's not in China
C mm
J hm {1,40} **it's er quite hard work**

(Celia&James : Travel)

Hard work fait l'objet d'un investissement intersubjectif de la part de James, lisible dans la modalisation du syntagme par *sort of*, appuyé de *all the time*, puis *really* et *quite*, et enfin simplement *quite*. Cependant, James ne semble pas réussir à faire admettre ce fait dans le paysage qu'il partage avec Celia. Après la marque d'écoute de Celia, relative, à nos yeux, à l'énoncé précédent, James reformule son énoncé. Celui-ci n'est pas ponctué à son tour par une marque consensuelle mais par une exclamation s'émerveillant du dépaysement procuré par un semestre de cours en Chine. James reprend alors sa recherche de consensus, initiée une troisième fois par le ligateur *so*. Il développe ce qu'il entend par *hard work* et obtient par deux fois une marque d'écoute signalant un consensus autour des représentations mentales engagées avant de conclure en réitérant son propos initial.

Les chevauchements de parole semblent ici constituer un contexte défavorable à l'élargissement du paysage partagé par les interlocuteurs. Cependant, les données de notre corpus indiquent plus généralement que le caractère consensuel de la relation intersubjective se trouve renforcé par les situations de parole recouverte.

Construction du consensus renforcée Un certain nombre de formules à interprétation co-énonciative font partie des segments que l'on retrouve fréquemment en situation de recouvrement de parole. De sémantisme faible, elles jouent un rôle essentiel sur le plan intersubjectif.

Dans notre extrait de corpus, plus du tiers des segments chevauchés impliquent la marque d'écoute *mm*. Dans l'épisode *Beer* analysé dans un chapitre précédent, tous les *mm* sont prononcés en situation de recouvrement de parole. Comme nous avons pu le voir à l'occasion de cette analyse,[23] ces segments témoignent de la position consensuelle adoptée par le locuteur par rapport au contenu de pensée qu'il attribue au locuteur en place.

Nous pensons ici aux exemples suivants, qui constituent autant de marques de consensus ou de demande de consensus de la part des deux locuteurs :

C *oh ri-* [ex.9.8]

J *you know like* no that's ridiculous <rires> *you know* [ex.9.9]

C *mm*

(Celia&James : Travel)

La marque d'acquiescement de Celia vient ponctuer la clarification précédemment sollicitée par l'énoncé **but she hasn't* learnt any Chinese*. Elle témoigne donc d'une concordance des représentations mentales mises en jeu et de l'agrandissement du paysage partagé à partir duquel l'échange peut continuer à se construire. Cette marque est recouverte par un ponctuant faisant appel à des représentations anticipées comme consensuelles, au sein de ce paysage partagé. Le ponctuant réapparaît plus loin, à son tour recouvert par une marque d'écoute. Il est remarquable que l'intensité sur ces segments soit particulièrement faible : les chevauchements de parole n'ont pas ici pour objectif de permettre à l'un des locuteurs de conserver la parole. Ils n'entrent pas en jeu dans une stratégie de lutte pour la parole, mais de construction du consensus. Le consensus semble en effet se consolider dans la texture dense créée par les

[23] Voir Sections 7.1.3.1, p.322, et Section 7.2.2, p.332.

recouvrements de parole.

Co-construction du discours Citons un dernier exemple appuyant notre hypothèse selon laquelle les chevauchements de parole sont en fait à bien des égards un ciment de l'interaction verbale. Ils ont la propriété essentielle à nos yeux de permettre une co-construction du discours. Sur le point d'arrêter l'enregistrement, Celia et James dressent un bilan qualitatif de leur conversation. Les chevauchements de parole sont particulièrement nombreux dans ce passage :

J *I was going to sort* of be careful and not mumble like I usually do but {0,45} they probably want {0,30} th- *genuine m-*
C *this is* normal *RP speech*
J *the genuine* mm genuine *me yes th-*
C *this is genuine* RP speech
guaranteed genuine R- P speech
J *genuine RP speech except that* {1,00} I should open *my mouth more*
C *free conversation*
J <rires>
C <rires>

(Celia&James : Genuine RP speech)

Deux motifs parallèles se développent ici : celui de James qui dit avoir été véritablement lui-même et celui de Celia qui estime que leur conversation est un échantillon de véritable anglais standard.

D'un point de vue discursif, le thème du passage tarde à se stabiliser, tant les interventions sont enchevêtrées. Deux projets concurrents semblent être à la lutte. Cependant, contrairement à nos exemples précédents où l'un des projets est finalement abandonné, les deux finissent ici par n'en faire plus qu'un. Tout d'abord, on assiste à un rapprochement des deux thèmes : lancé par James pour décrire ses propres productions verbales, le qualificatif *genuine* est repris à son compte par Celia qui l'inclut dans son projet initial au détriment de *normal*. Ensuite, James opère la synthèse entre les deux projets : **genuine RP speech except that* {1,00} I should open *my mouth more** :

il s'agit de véritable anglais standard avec le véritable James.

Nous retrouvons ici la dynamique singulière de la conversation spontanée. Les locuteurs font l'expérience étonnante (ponctuée par des rires) d'une construction du discours à deux, où les termes de l'un deviennent ceux de l'autre, et d'une correspondance parfaite des représentations mentales engagées à ce moment-là.

Loin de mettre en péril la compréhension du message, l'organisation discursive ou la relation intersubjective, les chevauchements de parole contribuent au dynamisme récursif propre à la conversation spontanée. Ils ne sont donc pas nécessairement le lieu d'une lutte pour la parole appuyée par une intensité forte et une mélodie montante, témoin, au niveau interactionnel d'un accordage difficile des représentations mentales. Ils favorisent au contraire l'émergence de thèmes productifs et consolident le paysage partagé au sein duquel le consensus se bâtit.

Conclusion

Tout au long de notre travail s'est illustré le fait qu'il est méthodologiquement utile mais véritablement illusoire de vouloir séparer le plan segmental et le plan suprasegmental. Les deux domaines ne sont en effet pas séparés dans l'expérience perceptive du locuteur engagé dans une conversation spontanée. Ils forment un tout complexe dont les composantes entrent dans un rapport de complémentarité ou parfois de redondance qui leur permet de contribuer tout deux à la structuration du discours. Pour conclure cette thèse, nous souhaitons revenir sur les hypothèses sur lesquelles nous nous sommes fondée. Il s'agit de rappeler les éléments que notre travail a apportés afin de valider celles-ci.

Le discours conversationnel comme objet fractal

L'hypothèse fractale, issue de la théorie mathématique, n'a pas pour vocation ici de s'appliquer de façon stricte. Il s'agit davantage de l'utiliser comme un modèle permettant de concevoir l'unité d'un ensemble à travers sa complexité et d'y chercher des régularités expliquant son organisation. Le discours conversationnel est riche de multiples facettes. Notre compréhension de son fonctionnement ne peut donc que s'enrichir d'une analyse qui cherche à rendre compte de sa complexité en décrivant comment ses composantes s'articulent à différents niveaux.

C'est pourquoi nous avons développé une réflexion en terme de niveau de perti-

nence des phénomènes rencontrés, pour refléter une division de notre objet d'étude en différents couches où la question de la structuration se pose. Ces niveaux sont relatifs à la taille des unités examinées ainsi qu'à leur portée, c'est-à-dire à la couche à laquelle ils se rapportent dans l'édifice conversationnel. Nous avons ainsi relevé différents types de phénomènes aux niveaux syntaxique (par exemple, la subordination), sémantique (notamment, le sens des marqueurs de subordination causale) et prosodique (impliquant les différents paramètres intonatifs). A ces phénomènes de divers ordres nous avons attribué une valeur spécifique sur les plans thématique ou plus largement discursif, mais également pragmatique, intersubjectif, et enfin interactionnel.

Ainsi, à travers l'exemple des marqueurs de subordination, notre étude de la subordination circonstancielle à l'oral et en particulier des circonstancielles de cause a révélé dans notre deuxième partie que les mêmes phénomènes se retrouvent à différentes échelles. En outre, nous avons montré dans la troisième partie de ce travail que la pertinence des mêmes paramètres suprasegmentaux varie en fonction du niveau d'analyse où on les considère. Ajuster constamment la précision de notre analyse à la taille de l'unité qui nous intéresse permet selon nous de contribuer à mettre au jour la texture particulière de notre objet d'étude et facilite la description de ses principes de structuration.

Le caractère fractal du discours conversationnel est apparu dans notre analyse des quelques points sur lesquels s'est portée notre attention. Cette hypothèse s'est trouvée en divers endroits étroitement liée à nos deux autres hypothèses, relatives à la syntaxe de l'oral ainsi qu'à la contribution de la prosodie à la hiérarchisation des constituants discursifs.

La syntaxe spécifique de l'oral

Notre travail a apporté un certain nombre d'éléments permettant de valider l'hypothèse d'une syntaxe de l'oral significativement différente de celle de l'écrit. Ces éléments concernent la subordination circonstancielle, et en particulier la subordination causale.

Réduction du champ des possibles

L'une des caractéristiques de la syntaxe de l'oral telle qu'elle s'illustre dans notre corpus est la faible variété des structures et des marqueurs utilisés, en regard de la grande diversité offerte par la grammaire.

Un phénomène récurrent

Nous avons montré au Chapitre 4 que ce phénomène s'illustre à plusieurs échelles, conformément à l'hypothèse fractale. Il est apparu tout d'abord que les catégories sémantiques des propositions subordonnées circonstancielles les plus représentées à l'oral sont les causales et les conditionnelles, au détriment de toutes les autres. La fréquence de ces deux types de circonstancielles est une caractéristique de la conversation spontanée, qui la distingue de l'écrit.

Nous nous sommes ensuite intéressée à l'éventail des conjonctions de subordination que l'on retrouve le plus fréquemment à l'oral. A cet effet, nous avons comparé des données issues de notre corpus, du corpus de la *Longman Grammar of Spoken and Written English* (LGSWE) de D. Biber *et al.*[24] et le *British National Corpus* (BNC). En nous limitant aux seules conjonctions les plus simples (*if, because/cos, when, before, while, although, until, since* et *after*), nous avons constaté que trois d'entre elles,

[24]BIBER, Douglas *et al.* (1998). *Longman Grammar of Spoken and Written English.* Londres : Longman.

if, when et *because/cos*, représentaient à elles seules 90% des conjonctions simples utilisées.

Ensuite, comme nous avons pu le voir au chapitre suivant, le même phénomène se vérifie lorsque l'on considère les marqueurs susceptibles d'être utilisés pour exprimer un rapport causal entre deux propositions. Le relevé des formes de notre corpus a permis de comparer l'utilisation des marqueurs *since, as, for, the reason why* et *because/cos*. Ici encore, le champ des possibles se réduit de façon extrêmement sensible aux seuls *because/cos* (plus de 90%).

Enfin, notre analyse de la subordination en *because/cos* révèle également une grande uniformité dans l'ordre dans lequel se présentent les propositions. Nous avons ici comparé le corpus London-Lund utilisé par R. Quirk *et al.*[25] et notre propre corpus. Dans plus de 90% des cas, la proposition subordonnée apparaît en position finale.

Un phénomène structurel

Après avoir constaté la récurrence de ce phénomène de réduction du champ des possibles à différentes échelles, nous avons formulé deux hypothèses explicatives concurrentes. La première, structurelle, est celle que nous retenons et dont nous nous faisons l'écho dans le présent développement.

Elle s'oppose à une hypothèse conjoncturelle qui explique le faible emploi de certaines formes par rapport à d'autres par un effet de corpus. Le contenu thématique des conversations constituant notre source de données, ou la relation intersubjective qui s'y construit, justifierait l'utilisation privilégiée de *if, when* et *because/cos*. De même, au sein des marqueurs de causalité, la préférence des locuteurs pour *because/cos* proviendrait du contenu des propos échangés ou du type de relation intersubjective engagée. Cette hypothèse, sans doute au moins partiellement recevable, est remise en

[25] QUIRK, Randolph *et al.* (1985). *A Comprehensive Grammar of the English Language*. Londres : Longman.

cause par un certain nombre de faits. Les différences relevées parmi les textes de notre corpus quant à la proportion des conjonctions simples examinées ne reflètent pas de différence tangible sur le plan thématique ou intersubjectif. En outre, le sémantisme véhiculé par les marqueurs favoris ne leur est pas spécifique : la causalité peut s'exprimer aussi bien par *because/cos* que par les autres marqueurs considérés, à quelques nuances près.

Ces nuances constituent précisément un point important sur lequel s'appuie l'hypothèse selon laquelle ce phénomène est bien une caractéristique structurelle de la conversation spontanée. *If, when* et *because/cos* constituent en effet les marqueurs archétypiques permettant d'exprimer la condition, la concession, la temporalité ou la causalité. Ils se situent, du point de vue des marqueurs de subordination, au centre organisateur de la notion en question.

Nous avons rappelé quelles nuances distinguent par exemple *since, as* et *because/cos* quant au repérage des propositions. Elles nous semblent également en mesure d'expliquer la très forte représentation de *because/cos* par opposition aux autres marqueurs de causalité. Il semble que, de façon structurelle, les locuteurs tendent à se diriger vers un marqueur générique, polyvalent, au détriment d'un marqueur plus spécialisé, et, pour certains, polysémique. Ainsi, *because* est préféré à *since* ou *as*, de la même façon que *when* est préféré à *before* ou *after*.

Un marqueur supplémentaire : *cos*

L'oral est riche d'un marqueur supplémentaire par rapport à l'écrit : *cos*. Notre travail montre en effet que, en dépit d'un certain nombre de faits le rapprochant de *because*, il s'agit d'un marqueur à part entière dont certains contextes, liés à des critères distributionnels spécifiques, favorisent l'apparition.

Critères distributionnels spécifiques

Un faisceau de critères distributionnels permet de distinguer les deux marqueurs.

Nous pensons tout d'abord aux sujets syntaxiques utilisés dans la proposition principale des énoncés en *because* et *cos*. Notre étude se limite aux pronoms personnels dont le référent est un animé humain. Nous montrons que les énoncés en *cos* contiennent de façon privilégiée un sujet à la première ou la deuxième personne du singulier ou du pluriel, tandis que *because* est lié à la troisième personne.

Ensuite, la place des pauses silencieuses constitue un critère distributionnel permettant de distinguer les deux marqueurs. On les trouve de façon caractéristique entre la principale et la subordonnée lorsque celle-ci est introduite par *cos*, alors qu'elles se situent de préférence dans la subordonnée avec *because*.

Typiquement, les énoncés en *because* et *cos* se présentent sous la forme suivante, conformément à nos observations concernant la place des pauses silencieuses et le type de sujet syntaxique de la principale :

He/she/they + V ... + **because** + **{pause}** + S + V ...

I/you/we + V ... + **{pause}** + **cos** + S + V ...

Enfin, la présence de marques du travail de formulation (parmi lesquelles les pauses silencieuses d'hésitation, pauses pleines, faux départs) et de prise en charge (dont *I mean, you know*) caractérise les énoncés en *because*, qui sont en cela significativement différents des énoncés en *cos*. Non seulement les énoncés en *cos* sont moins nombreux à contenir de telles marques, mais celles-ci y sont moins denses.

Contextes privilégiés

Ces critères distributionnels différents correspondent à des contextes d'apparition spécifiques à l'un et l'autre des marqueurs.

La place typique de la pause dans un énoncé en *because* renseigne sur la fonction

de celle-ci : il s'agit quasiment invariablement d'une pause d'hésitation, marque du travail de formulation. L'élaboration du propos dans les énoncés en *because* semble plus difficile que dans les énoncés en *cos* où nous analysons la relation entre p et q comme davantage congruente. Cette différence est également reflétée selon nous par la proximité des référents des sujets syntaxiques typiquement associés avec l'un et l'autre des marqueurs. La sphère du « eux », caractéristique de *because*, témoigne d'un plus grand éloignement entre les représentations mentales engagées par les parties prenantes du discours et les actants du procès dont il est question. Avec la sphère du « nous », davantage associée à *cos*, la proximité des énonciateurs et des actants rend la validation moins problématique.

La prise en charge des validations se fait en T_0 avec *because*, alors qu'elle a tendance à être déjà acquise avec *cos*. Couramment associé à *so*, *because* apparaît dans des contextes où il est appelé à faire montre d'une force argumentative supérieure à *cos*. Ce dernier est quant à lui plus souvent lié à *but*. Sur le plan de la relation inter-sujets, *because* apparaît dans un contexte où l'énonciateur anticipe une possible divergence de représentations, tandis que *cos* est davantage lié à un contexte où le contenu de pensée de l'autre est anticipé comme consensuel ou n'est pas pris en compte.

Réinvestissement des structures existantes

La syntaxe de l'oral se caractérise également par une utilisation différente des formes habituelles, notamment en ce qui concerne la subordination causale.

Assouplissement des principes de la subordination

Notre étude de la subordination en *because/cos* montre que le lien entre subordonnée et principale tend à être relativement lâche. Une forme originale apparaît de façon récurrente, que nous désignons sous le terme de « collaboration », dans laquelle

subordonnée et principale ne sont pas prononcées par le même locuteur. Proposition repère et proposition repérée proviennent de sources différentes.

D'autre part, nous relevons peu de formes dont la syntaxe indique que la subordonnée apporte une restriction à la relation prédicative exprimée dans la principale. Cependant, les énoncés dont la subordonnée constitue un commentaire métalinguistique sont également relativement peu nombreux. Dans une large partie de nos exemples, la relation entre p et q peut être interprétée comme appositive. La subordonnée énonce simplement une circonstance non restrictive incidente à la validation de la relation prédicative dans p, sans pour autant relever véritablement du métalinguistique.

Gestion du statut discursif des éléments reliés

La créativité de la langue s'illustre enfin dans son emploi des marqueurs de la subordination afin de signaler le statut discursif respectif des éléments liés. Ce n'est donc pas de dépendance syntaxique qu'il s'agit mais d'une dépendance thématique. Il est significatif que l'emploi de *because/cos* ne soit que très marginalement lié à l'interrogatif *why* (moins de 1%).

Le sémantisme causal induit par des marqueurs comme *because/cos* ou *so* se trouve singulièrement affaibli dans la conversation spontanée. Les marqueurs tissent un réseau relatif à l'organisation discursive, qui oppose des éléments de premier plan et des éléments d'arrière plan. Ainsi, *because/cos* introduit un élément secondaire, tandis que *so* signale typiquement un segment relatif à la ligne principale du discours. La hiérarchisation des constituants tend donc à glisser du plan syntaxique, où s'illustrent les conjonctions, vers celui de l'organisation discursive, domaine des ligateurs.

Notre travail a également montré que l'un des principes de cette organisation se donne dans le linéaire sous la forme d'une spirale. Dans un contexte argumentatif comme dans un contexte descriptif, la progression discursive admet un certain nombre

de retours en arrière. Elle s'appuie notamment sur les marqueurs *because/cos* et *so* pour élargir le domaine des connaissances partagées et construire le discours sur un consensus. Les structures concernées par ce dynamisme récursif sont du type *A because B so A'*. Ce phénomène illustre encore le caractère fractal du discours conversationnel, dans la mesure où il s'observe à l'échelle de l'énoncé ou du paragraphe oral, aussi bien qu'à celle de l'épisode thématique, ou qu'à celle de la conversation dans son ensemble.

Différente et créative, la syntaxe de l'oral s'appuie sur une dimension supplémentaire essentielle : celle des paramètres suprasegmentaux.

Signification et rôle des paramètres prosodiques

La dernière hypothèse sur laquelle notre étude s'est fondée concerne la dimension suprasegmentale de la conversation spontanée en anglais. Ici encore, notre travail a apporté des éléments permettant d'évaluer la signification conventionnelle des paramètres prosodiques telle qu'elle s'illustre dans notre corpus, ainsi que leur rôle dans la construction de l'édifice conversationnel.

Pour une méthodologie de l'analyse prosodique

Notre travail plaide pour l'établissement d'une méthodologie de l'analyse des phénomènes prosodiques dans la conversation spontanée. Le linguiste de l'écrit est en effet relativement démuni devant la complexité de la représentation de la réalité physique à laquelle les outils logiciels lui donnent accès.

L'enregistrement et la transcription d'un corpus original adéquat (compte tenu de nos objectifs de recherche) constitue un préliminaire essentiel à notre étude. Dans un deuxième temps, il s'est agi d'établir les données expérimentales sur lesquelles fonder nos analyses. De même qu'enregistrement et transcription suivent un protocole relativement strict, il nous a paru important d'étendre cette réflexion méthodologique

au choix des paramètres suprasegmentaux de la voix parlée à prendre en compte, ainsi qu'à l'interprétation des valeurs numériques et à la représentation graphique fournies par les logiciels. Une définition préalable permettant de cerner la nature des phénomènes acoustiques nous a permis de retenir les paramètres suivants : durée, intensité et fréquence fondamentale. Obtenue par calcul, cette dernière met devant des choix supplémentaires à faire par rapport à la durée et à l'intensité, simplement mesurées.

Afin de pouvoir donner une lecture linguistique des variations prosodiques, il nous a fallu déterminer lesquelles d'entre elles sont véritablement pertinentes du point de vue de la perception. La psychoacoustique nous a renseigné sur quelques principes fondamentaux de la perception humaine, notamment que celle-ci n'est pas linéaire, et qu'une valeur prend essentiellement sens par rapport aux autres valeurs de son contexte immédiat. En tenant compte des seuils de perception différentielle et de l'échelle logarithmique sur laquelle ils s'inscrivent, nous avons rappelé ou calculé quels ordres d'écart entre une valeur et la valeur suivante (dans les limites correspondant à la voix parlée) sont réellement perçues par l'oreille humaine. Il est apparu que, devant la multitude d'informations présentées sur une courbe intonative, les phénomènes de faible amplitude et de petite durée sont à négliger au profit d'autres plus clairement discernables. Nous avons également montré que la déformation provoquée par notre utilisation d'une échelle linéaire pour représenter des faits perçus selon une échelle logarithmique est négligeable. Cependant, nos expériences du Chapitre 7 sur la conduite de la ligne mélodique tiennent compte de cette différence d'échelle.

Enfin, nous avons cherché à évaluer dans quelle mesure les différents paramètres évoluent indépendamment les uns des autres sur le plan physiologique, et comment la perception de l'un influence celle d'un autre. Il s'est agi de déterminer si les variations des paramètres sont susceptibles d'être analysés séparément. Du point de vue de la production, si la pression sous-glottique contrôle à la fois la fréquence fondamentale

et l'intensité, la gestion des deux paramètres est suffisamment décorrélée pour que leurs variations puissent faire sens en elles-mêmes. En outre, il apparaît que fréquence fondamentale et intensité ont tendance à baisser à mesure que l'on s'éloigne de la dernière pause respiratoire. Cependant, cette tendance n'exclut pas l'apparition de modulations mélodiques ou de pics d'intensité à n'importe quel moment. Sur le plan de la perception, certaines interférences existent également, qui autorisent néanmoins une prise en compte séparée des phénomènes, mais engage à également les considérer en combinaison.

Niveaux de pertinence et délinéarisation de la chaîne

La contribution des paramètres intonatifs à la hiérarchisation des constituants discursifs est multiple. Notre travail met en évidence l'existence de différents plans sur lesquels les variations des paramètres intonatifs sont pertinents. Ces variations contribuent à délimiter des unités dans la chaîne parlée et à les organiser en réseaux à plusieurs niveaux de lecture. Ceux-ci opèrent indépendamment de l'ordre linéaire dans lequel les unités sont agencées. Les variations mélodiques et intensives étant mobiles, ce que nous avons appelé une « délinéarisation » de la chaîne syntaxique est à l'œuvre. Celle-ci consiste en fait à signaler le statut discursif, intersubjectif et interactionnel d'éléments de tailles diverses indépendamment de leur statut syntaxique. Nous avons distingué plus haut trois niveaux de pertinence pour les phénomènes intonatifs. Seuls deux relèvent véritablement de l'échelle d'analyse que nous avons choisie : le niveau discursif et intersubjectif, et le niveau interactionnel.

Statut discursif et intersubjectif

Ce niveau de pertinence des paramètres intonatifs concerne des segments relativement courts, la plupart du temps des mots isolés ou des syntagmes. Les différents

paramètres que nous avons envisagés y ont leur rôle à jouer.

Hauteur mélodique Notre Chapitre 7, s'appuyant sur une double hypothèse statistique et linguistique, a montré dans quelle mesure les variations de la ligne mélodique renseignent sur l'état de la relation intersubjective. Selon notre hypothèse, que nos résultats tendent à valider, la plage mélodique moyenne correspond à une posture neutre, non conflictuelle, entre les parties prenantes du discours. Elle témoigne du fait que l'énonciateur anticipe un consensus entre le contenu de pensée qu'il attribue à l'autre et le sien. La plage haute est liée au contraire à l'anticipation d'un écart de représentations désignée par les postures de discordance et de forçage de consensus. La plage mélodique basse correspond enfin à un état de la co-énonciation où la pensée de l'autre n'est pas prise en compte. Elle témoigne d'une posture égocentrée.

Un approfondissement de notre hypothèse doit permettre de distinguer les deux postures correspondant à la bande supérieure, en prenant en compte un faisceau de paramètres suprasegmentaux et/ou segmentaux. Nous avons également montré que, si la plage mélodique basse correspond à la posture égocentrée, la réciproque n'est pas toujours vraie. La prise en compte conjointe des paramètres prosodiques (et notamment de l'intensité en plus de la mélodie) permet d'obtenir de meilleurs résultats.

Intensité L'intensité contribue à la gestion de l'espace intersubjectif dans la mesure où le faible volume sonore d'un segment situé dans la plage mélodique moyenne permet d'interpréter celui-ci non pas comme consensuel mais comme égocentré. Le rôle de l'intensité est néanmoins essentiellement discursif. Des travaux précédents[26] ont montré que les éléments faisant l'objet d'un cadrage thématique tendent à être soulignés par un pic d'intensité. Ce paramètre contribue ainsi à la construction des objets

[26]PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (2000). Analyse d'une composante articulatoire : l'intensité en anglais oral spontané. *Journée d'étude, Laboratoire d'informatique et de mécanique pour les sciences de l'ingénieur (LIMSI-CNRS), séminaire de recherche de J.-B. Berthelin.* 14 mars 2000. Communication affichée.

de discours. Les pics peuvent affecter des mots apparaissant pour la première fois, des mots déjà mentionnés ou appartenant à un domaine notionnel déjà abordé. La mise en relief intensive permet dans les deux derniers cas de procéder à un recadrage :

> Le locuteur s'appuie sur les éléments mis en avant par l'intensité pour orienter le discours dont ils sont l'objet vers une nouvelle direction.[27]

Ce phénomène se vérifie dans un certain nombre d'extraits de corpus (notamment dans les épisodes *Thesis* et *Teaching* de Nick&Matthew ou encore *Pat* de Mike&Hannah) et mériterait une analyse de plus grande envergure dans le cadre de recherches à venir. Il correspond en outre en partie aux observations de M. Candea & F. Lefeuvre sur les récits en français, dans lesquels les actants du discours et les circonstants sont fréquemment soulignés par un pic d'intensité.[28]

Pause silencieuse Enfin, la pause silencieuse a également un rôle important au niveau discursif et intersubjectif. Pause d'hésitation, elle a pour fonction de signaler l'état de formulation du contenu de pensée que l'énonciateur cherche à exprimer. Pause de segmentation, elle unifie le segment précédent et délimite des unités plus ou moins larges qui peuvent ensuite être mises en relation les unes avec les autres. Enfin, pause de focalisation, elle isole le segment suivant et le met en relief. Elle le rend ainsi propice à une interprétation marquée sur le plan intersubjectif. C'est sur ce dernier type de pause, absent de l'extrait de corpus analysé au Chapitre 8, que nous souhaiterions nous attarder dans des travaux à venir.

Statut interactionnel

Les paramètres intonatifs, isolés et en combinaison, sont pertinents au niveau de l'organisation matérielle du dialogue. Ils renseignent ainsi sur le statut interactionnel

[27] *ibid.*

[28] CANDEA, Maria & LEFEUVRE, Florence (à paraître). Les pics d'intensité dans le récit oral spontané. *Travaux linguistiques du CERLICO* 17.

des unités dégagées.

Pause silencieuse La pause silencieuse constitue une charnière potentielle permettant l'alternance des tours de parole. Notre Chapitre 8 montre que la perception des pauses, partiellement corrélée avec leur durée et leur fonction, est facilité par leur inscription dans un contexte prosodique propice à l'alternance des tours de parole. Ainsi, les pauses situées avant une réinitialisation de la ligne de déclinaison et un rehaussement du niveau d'intensité générale bénéficient d'un meilleur taux de perceptibilité que les autres et soulignent des charnières importantes au niveau interactionnel.

Hauteur mélodique La ligne mélodique, conformément au code de la production de C. Gussenhoven,[29] tend à décliner en fin d'énoncé. Une baisse de F_0 tend ainsi à s'interpréter au niveau interactionnel comme un signal de fin de tour. Au contraire, une montée mélodique sur les dernières syllabes d'un énoncé a une valeur continuative. La conduite de la ligne mélodique indique donc dans quelle mesure la prise de parole de l'autre est la bienvenue à la place transitionnelle suivante. Cependant, comme nous avons pu le montrer dans des travaux antérieurs, une montée mélodique peut également être liée à une interrogative dont le propos est précisément d'inviter l'autre à prendre la parole.[30] L'interprétation continuative doit donc tenir compte du statut intersubjectif de l'énoncé et peut également s'appuyer sur le niveau de l'intensité sur les segments concernés par la montée mélodique

Intensité L'intensité générale d'un segment renseigne également sur son statut interactionnel. En cas de chevauchement de parole, nous avons montré dans notre dernier chapitre que les segments à l'intensité faible ne constituaient pas une menace

[29]CHEN, Aoju, GUSSENHOVEN, Carlos & RIETVELD, Toni (2002). Language-specific Uses of the Effort Code. in Bel, Bernard & Marlien, Isabelle (éds.), *Proceedings of the Speech Prosody 2002 conference* : 215-218.

[30]PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (2001). Les interrogatives en anglais oral spontané : intonation et énonciation. *Anglophonia-Sigma* 8 : 125-152.

pour le tour de parole du locuteur en place. C'est ainsi que les segments prononcés en situation de recouvrement de parole tels que les marques d'écoute ou les ponctuants phatiques, dont l'objet est de renforcer le locuteur en place dans sa position d'énonciateur principal ; sont pour la plupart prononcés avec une intensité relativement faible. Les segments marqués par une intensité forte témoignent au contraire de la part du locuteur d'une revendication de son droit à la parole.

Ces remarques formulées à partir des segments chevauchés rencontrés dans un extrait de notre corpus mériteraient d'être approfondies par l'étude d'un plus grand nombre d'exemples. En particulier, l'analyse de davantage d'exemples dans lesquelles une relation conflictuelle entre les interlocuteurs occasionne des chevauchements de parole dont l'issue est décidée par le niveau d'intensité déployé par chaque locuteur.

Au terme de ce travail, nous espérons avoir apporté suffisamment d'éléments permettant de valider nos hypothèses de départ, ou, tout du moins, de confirmer leur intérêt dans le cadre d'une étude conjointe de faits syntaxiques et de phénomènes intonatifs de l'anglais oral spontané. Il s'est agi d'éclairer un certain nombre de principes organisationnels s'illustrant dans un corpus conversationnel original. Notre analyse en plusieurs niveaux a notamment permis de rendre compte de la texture du système de l'oral à laquelle segmental et suprasegmental contribuent à part égale. Si l'analyse séparée des deux dimensions a fait apparaître des phénomènes intéressants, la prise en compte conjointe de données contextuelles, sémantiques, syntaxiques et intonatives s'est montrée nécessaire et productive aux niveaux d'analyse qui nous intéressent.

Notre travail contribue ainsi à montrer dans quelle mesure la syntaxe de l'oral, étudiée à travers le prisme de la subordination circonstancielle de cause, réinvestit les formes de l'écrit pour créer un système autre. Caractérisée par un dynamisme récursif, l'organisation du discours s'appuie sur les divers paramètres intonatifs dont la portée dépasse le niveau thématique pour s'étendre à la gestion de l'espace intersubjectif

ainsi qu'à celle de l'interaction.

En convoquant des outils et méthodes aussi divers que ceux du traitement du signal, de la phonétique et de la statistique, notre objectif a été de servir la description et l'analyse linguistique d'un domaine particulièrement riche et exigeant. Nous espérons avoir ainsi pu éclairer certains phénomènes relatifs à la hiérarchisation des constituants discursifs, propres, sinon à la langue orale, du moins à un corpus d'anglais oral spontané.

Annexe A

Annexe au chapitre 3

Nous rappelons ici les caractéristiques techniques du matériel utilisé pour l'établissement de notre corpus. L'enregistrement, la numérisation et l'édition des fichiers sonores, et enfin l'analyse prosodique ont nécessité l'utilisation d'outils spécialisés.

A.1 Enregistrement du corpus

Le matériel, de qualité professionnelle, qui a servi à enregistrer notre corpus nous a été gentiment prêté par le centre de recherches de M.-A. Morel. Pour chaque enregistrement, nous avons utilisé :

- deux magnétophones Marantz à réglage manuel PMD 201 professional
- deux micros-cravate omnidirectionnels Vivanco EM116
- deux cassettes audio normales (durée 60 minutes)

A.2 Numérisation du corpus et édition des fichiers sonores

Les caractéristiques des outils de numérisation que nous avons utilisés sont les suivantes :

- logiciel d'édition de fichiers au format WAV : Cool Edit version 96 en shareware. Syntrillium Software Corporation. Concepteur : David Johnson.
- résolution optimale : 44 kHz, 16 bit

Les fichiers au format MP3 présentés sur le CD-ROM ont subi une compression au taux de 192 kbps.

A.3 Analyse prosodique

Après avoir envisagé quatre logiciels différents pour l'analyse prosodique, notre choix s'est fixé sur WinPitch Classic, avec la possibilité de recourir occasionnellement à Praat :

- logiciel WinPitch Classic : Real Time Analysis for Speech : version : v.1.92, 1996-1997 pour Windows 95/98. Pitch Instruments Inc., Toronto, Ontario, Canada. Concepteur : Philippe Martin.
- résolution optimale : 44kHz, 16bit
- algorithme pour l'analyse de Fo : fast comb.
- occasionnellement : logiciel Praat : doing phonetics by computer : version : v.4.0.49, 1992-2003. Concepteurs : Paul Boersma et David Weenink.

Les avantages et inconvénients comparés des logiciels WinPitch Classic, Praat, Anaproz et WinSnoori sont présentés au Tableau A.1.

AVANTAGES	INCONVÉNIENTS
Anaproz	
* affichage simultané de l'intensité et de F_0 * nombreux réglages de l'affichage * possibilité de saisir la transcription d'une séquence sur l'écran du tracé correspondant * possibilité de conversion des tracés au format image	* courbes d'intensité peu lisibles * impossibilité d'afficher le spectrogramme * impossibilité d'afficher la valeur précise de F_0 en un point * nécessité pour l'ordinateur de disposer d'une mémoire très importante pour un fonctionnement rapide
Praat v.4.0.49	
* compatible avec différentes plate-formes (Windows, Unix, Linux, MacOS) * plusieurs algorithmes de calcul de F_0 disponibles * possibilité de lier un fichier texte d'étiquetage aux fichiers son et courbes pour l'alignement automatique * liste de diffusion permettant d'entrer en contact avec d'autres utilisateurs * fonction zoom pour l'analyse fine d'un segment * logiciel ouvert, possibilité d'ajouter des fonctionnalités à l'aide de scripts	* utilisation complexe et peu intuitive * impressions à grande échelle non prévues
WinSnoori (démo)	
* grande finesse dans les échelles d'analyse	* peu intuitif * version testée instable * nécessité pour l'ordinateur de disposer d'une mémoire très importante pour un fonctionnement rapide
WinPitch Classic v.1.92	
* affichage simultané de I et de F_0 en temps réel * possibilité de choisir l'algorithme de calcul de F_0 * possibilité de choisir le type et l'étendue de l'échelle de représentation de F_0 et de I * possibilité de régler le nombre de secondes de parole analysées représenté par écran * possibilité de saisir la transcription d'une séquence sur l'écran du tracé correspondant * réécoute de segments par sélection directement sur l'écran des tracés, en regard de la transcription * possibilité de modifier les tracés par la synthèse vocale * possibilité d'afficher les valeurs précises de F_0 et de I en un point * module statistique et possibilité d'extraire des données chiffrées * supporte des fichiers volumineux (max. 2 minutes) * possibilité d'imprimer les tracés dans un fichier * possibilité de conversion des tracés au format image	* impression rendue difficile par une gestion des tracés écran par écran * gestion de la transcription d'autant plus difficile que le fichier son est volumineux (longueur optimale : 20 s) * impossibilité de manier un fichier courbes sans le fichier son attaché * impossibilité de lier un fichier texte à un fichier son * nécessité de saisir la transcription dans le fichier courbes * fonctionne exclusivement sous Windows, version WinPitch Classic 1.92 incompatible avec Windows 2000

TAB. A.1 – Comparaison de quatre logiciels d'analyse prosodique

Annexe B

Annexe au chapitre 4

Cette section a pour objectif de clarifier la façon dont nous avons établi les données utilisées au Chapitre 4 dans le cadre de notre étude distributionnelle des marqueurs de subordination. Notre analyse est fondée sur le dépouillement de trois corpus différents : la grammaire de D. Biber *et al.* (LGSWE),[1] le *British National Corpus* (BNC) et une grande partie de notre corpus (corpus FP).

Outre la différence de taille, plusieurs facteurs rendent d'importantes précautions nécessaires pour analyser et comparer les trois corpus avec rigueur. Tout d'abord, l'accès aux données est différent dans chacun d'entre eux. Ensuite, et par voie de conséquence, le type de données accessibles est différent puisqu'il relève d'un choix parfois déjà opéré par les auteurs du corpus. Comment avons-nous procédé pour sélectionner des données chiffrées aptes à servir de point de départ à une étude distributionnelle comparative ?

[1] BIBER, Douglas *et al.* (1998). *Longman Grammar of Spoken and Written English*. Londres : Longman.

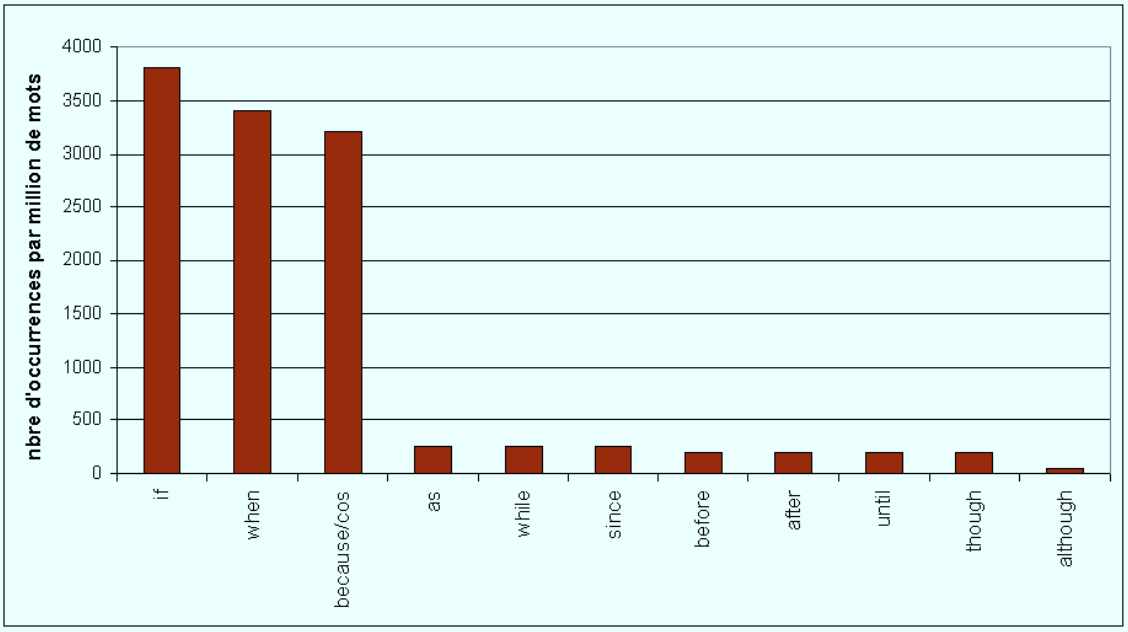

FIG. B.1 – Dénombrement des conjonctions simples dans le corpus de la LGSWE

B.1 Données issues de la LGSWE

La *Longman Grammar of Spoken and Written English* (LGSWE) propose une étude quantitative concernant les occurrences des conjonctions à partir d'une classification sémantique retenant les catégories suivantes : temps, manière, raison, concession et condition. Un tableau en deux parties a pour but de différencier en fonction de la fréquence des différentes conjonctions, d'une part, les quatre types de corpus utilisés par les auteurs (bulletin d'information, fiction, conversation et communication scientifique), et, d'autre part, la conversation en anglais britannique et en anglais américain. Nous retiendrons ici essentiellement la partie concernant la conversation en anglais britannique, afin de pouvoir comparer les tendances relevées avec celles de notre corpus.

Les conjonctions répertoriées par la LGSWE sont les suivantes, par ordre décroissant de fréquence : *if, when, because/cos* représentent à elles seules près de 90%

des conjonctions relevées, tandis que *as, while, since, before, after, until, though*, et *although* sont très peu représentées.

Notons que les auteurs dénombrent séparément les occurrences de la conjonction *as* selon qu'elle introduit une subordonnée temporelle ou de manière, et de *since* en fonction de son emploi temporel ou causatif. Dans l'histogramme que nous présentons ici à la Figure B.1, le nombre d'occurrences de chacune des conjonctions est indépendant du sémantisme de la subordonnée qu'elle introduit (les chiffres des deux catégories ont donc été additionnés dans chaque cas). Ainsi, une erreur est introduite pour les deux conjonctions qui ont fait l'objet de deux opérations de dénombrement.

La précision de l'étude quantitative menée ici par la LGSWE et retranscrite par nos soins sous la forme d'un tableau de valeurs est la suivante : les données relatives aux conjonctions dans ce corpus ont une granularité de 200 occurrences. Le dénombrement effectué par les auteurs ne donne en outre pas une image très fidèle de la quantité absolue des différentes conjonctions dans la totalité de leur corpus. Le nombre observé étant ramené à un million de mots, il ne s'agit donc que d'une proportion. Au total, il apparaît que le corpus de la LGSWE contient environ 12 000 conjonctions simples par million de mots (soit 1,2%).

B.2 Données issues du BNC

Les conjonctions dont nous avons effectué le relevé dans le BNC oral sont délibérément les mêmes que celles que nous avons choisi d'étudier dans notre corpus. En plus de celles répertoriées par la LGSWE, nous comptons *once* et *till*. En outre, *cos* fait l'objet d'un dénombrement séparé par rapport à *because*. Seul *unless*, très peu fréquent dans la LGSWE, a été écarté ici.

Quelques précautions ont été nécessaires afin d'obtenir les données présentées ici. Tout d'abord, pour rendre la comparaison possible avec notre corpus et avec la partie

conversationnelle du corpus de la LGSWE, nous avons restreint notre champ d'investigation à la seule portion du BNC qui a été produite à l'oral. Ensuite, l'étiquetage syntaxique des mots du BNC permet de dénombrer précisément les occurrences de chacun des mots qui nous intéressent employés en tant que conjonctions de subordination. Cependant, certaines ambiguïtés subsistent, signalées comme telles par les auteurs du corpus. Ainsi, une requête concernant *when* révèle cinq formes différentes. Les différents *when* sont décrits comme adverbe interrogatif (AVQ), conjonction de subordination (CJS), inclassable (UNC), ou relevant de l'une des deux premières classes (combinaison entre AVQ et CJS) sans pouvoir véritablement trancher entre les deux. Seules les deux premières étiquettes témoignent donc d'une appartenance non ambiguë à une classe. Nous avons décidé de prendre ici en considération uniquement les occurrences pour lesquelles l'identification en tant que conjonction de subordination n'est pas sujette à caution, c'est-à-dire exclusivement celles portant l'étiquette CJS. En outre, nous avons pu nous assurer que le mot *cause* n'a pas été utilisé dans le BNC pour transcrire ce que nous transcrivons par *cos*. Aucune occurrence de ce mot en tant que conjonction de subordination n'est présente dans le corpus oral du BNC (*cuz* en est également absent).[2]

Le dépouillement des 10 365 464 mots du BNC oral y révèle la présence de 133 239 conjonctions simples (soit 1,28%). Comme le montre la Figure B.2, les conjonctions simples relevées dans le BNC sont inégalement représentées. *If, when*, puis *because, cos* et *as* sont les plus fréquentes. Comme dans le cas du corpus de la LGSWE, ces quatre conjonctions rassemblent plus de 90% des occurrences.

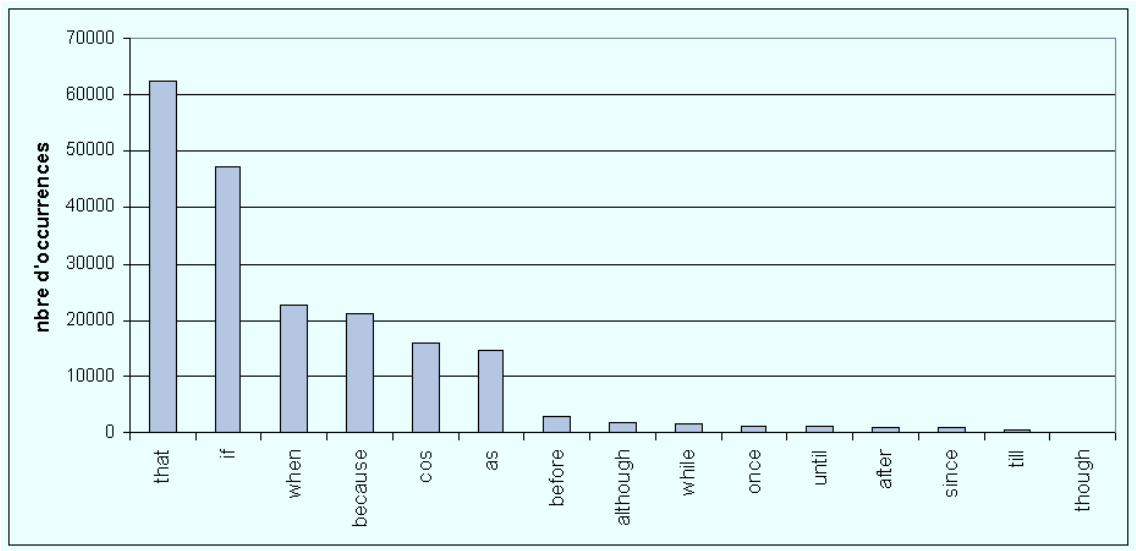

FIG. B.2 – Dénombrement des conjonctions simples dans le BNC

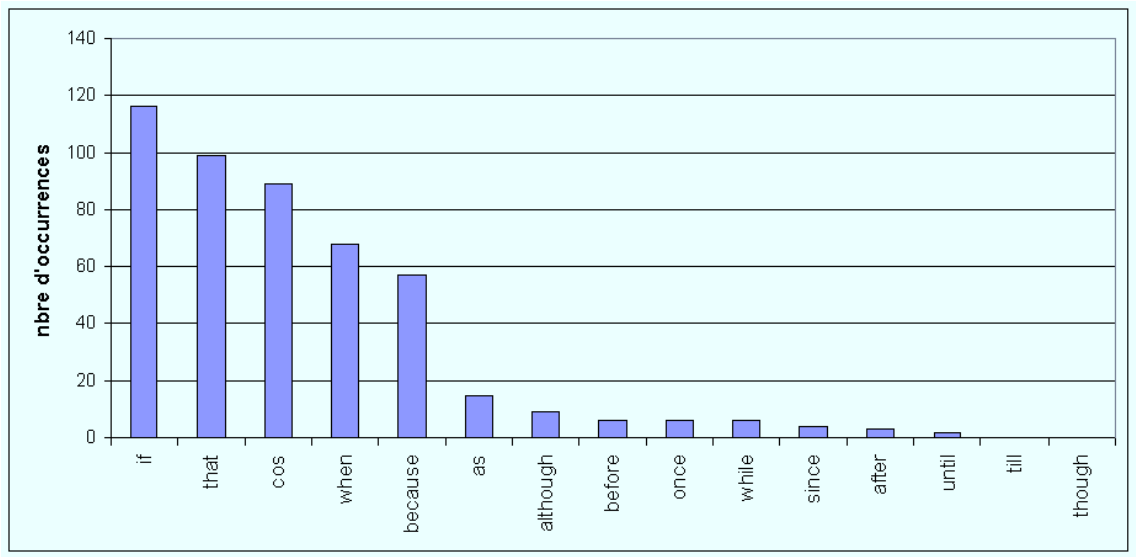

FIG. B.3 – Dénombrement des conjonctions simples dans notre corpus

B.3 Données issues du corpus FP

Le relevé systématique des conjonctions simples dans le corpus FP, conformément à la liste proposée plus haut,[3] révèle les tendances suivantes. Des 482 occurrences rencontrées se détachent d'abord *if*, puis *cos, when* et *because*, qui totalisent près de 90% des occurrences. Les autres marqueurs, *although, before, once, while, since, after, until*, et dans une moindre mesure *as*, apparaissent exceptionnellement. Comme l'illustre la Figure B.3, notre corpus ne contient aucune occurrence des conjonctions simples *till* et *though*.

B.4 Bilan : données comparables

Comme nous avons pu le constater plus haut, le nombre et la nature des conjonctions relevées ne sont pas les mêmes dans les trois corpus.

Par exemple, *because* et *cos* ne sont pas différenciés dans la LGWSE. *Till* ne figure pas non plus dans la liste de cette grammaire. En outre, les requêtes par parties du discours permettent le dénombrement de formes précises de marqueurs dans le BNC. Dans la LGSWE, en revanche, il n'est pas possible d'isoler le nombre d'occurrences d'une forme spécifique d'un marqueur lorsque le calcul a été fait pour plusieurs formes du même marqueur. C'est le cas de la conjonction *as*, qui dans notre corpus et dans notre dépouillement du BNC a été dénombré seul, alors que la LGSWE présente le nombre d'occurrences de ce marqueur y compris dans les locutions *as if* et *as though*.

En conséquence, le Tableau 4.1 (p.207) et la Figure 4.2 (p.208) ne présentent pas les chiffres comparés de l'occurrence de *as* dans les trois corpus. De la même façon, les données de la LGSWE concernant *though* comprennent les occurrences de

[2]Ironiquement, les 18 occurrences de *cause* en tant que conjonction se trouvent dans des dialogues de deux textes écrits.

[3]Nous ajoutons à la liste de la Section 296 du Chapitre 4(p.206) une distinction entre *because* et *cos*.

la locution *even though*, ce qui n'est pas le cas dans les deux autres corpus étudiés. Faute de pouvoir comparer des phénomènes de même ordre, notre analyse ne prend donc pas en compte ces deux conjonctions. Est également écartée de la comparaison la conjonction *till*.

Restent donc neuf marqueurs pour lesquels, le dénombrement ayant été fait sur une même base, une étude comparative peut être engagée : *if, because/cos, when, before, while, although, until, since* et *after*.

Annexe C

Annexe au chapitre 5

Nous présentons ici les tableaux d'effectifs ayant servi aux tests statistiques évoqués au Chapitre 5. Le test de Pearson, ou χ^2 (khi2)

> [...] sert à apprécier en probabilité l'écart constaté entre une observation et un modèle théorique, quel que soit le nombre des variables.[1]

Accompagné d'un degré de confiance, il donne un indice statistique de la différence entre deux échantillons en fonction de variables à déterminer. Dans le cadre de notre recherche, il nous a permis d'établir des différences significatives quant à la distribution de *because* et *cos* dans un certain nombre de configurations.

Ces tests ont été effectués sur les 123 énoncés de notre corpus que nous considérons suffisamment aboutis pour faire l'objet d'une analyse.

C.1 Référence du sujet de la principale

Les effectifs présentés au Tableau C.1 témoignent, d'après le test du χ^2, d'une distribution significativement différente des marqueurs *because* et *cos* quant à la référence du sujet grammatical de la proposition principale : $\chi^2 = 7{,}691$ pour 2 ddl, $p \leqslant 0{,}025$.

[1] MULLER, Charles (1992). *Initiation aux méthodes de la statistique linguistique.* Paris : Honoré

	BECAUSE	COS
sphère du « nous »	19	38
sphère du « eux »	8	2
autres	22	32
TOTAL	49	72

TAB. C.1 – Référence du sujet de la principale

La catégorie « nous » couvre les sujets première et deuxième personne du singulier et du pluriel à référence humaine, « eux » correspond aux sujets troisième personne à référence humaine, et « autre » rassemble tous les autres sujets grammaticaux.[2]

C.2 Position des pauses silencieuses

Le Tableau C.2 présente les données permettant de déterminer une différence significative entre les deux marqueurs quant à la distribution des pauses silencieuses. Notre calcul tient compte des places suivantes : entre la principale et la subordonnée (site 1), dans la subordonnée (site 2) et ailleurs (c'est-à-dire, pour les pauses uniques, dans la principale, ou pour les pauses multiples, à deux sites ou plus).

	BECAUSE	COS
Pas de pause :	19	40
Pause unique :		
Entre principale et subordonnée (site 1)	2	10
Dans subordonnée (site 2)	18	13
Autres sites :	10	11
TOTAL :	49	74

TAB. C.2 – Position des pauses silencieuses

Champion. (p.116).

[2]L'effectif total des énoncés en *cos* tient compte du fait que deux des propositions en *cos* de notre corpus ne sont pas rattachées à une proposition principale.

La distribution des pauses silencieuses à la charnière syntaxique, à l'intérieur de la subordonnée et l'absence de pause dans les énoncés en *because* et *cos* varient de façon significative en fonction du marqueur : $\chi^2 = 8{,}969$ pour 3 ddl, $p \leqslant 0{,}05$.

C.3 Présence et densité des marques du travail de formulation

C.3.1 Présence de marques du travail de formulation

La présence des marques du travail de formulation dans les énoncés en *because* les distingue de façon significative des énoncés en *cos*. Fondé sur les effectifs du Tableau C.3, le test indique en effet les résultats suivants : $\chi^2 = 15{,}220$ pour 1 ddl, $p \leqslant 0{,}001$.

	BECAUSE	COS
absence de marque	11	43
présence de marque	38	31
TOTAL :	49	74

TAB. C.3 – Présence de marques du travail de formulation

C.3.2 Densité des marques du travail de formulation

Plus présentes dans les énoncés en *because*, les marques du travail de formulation y sont également plus denses. Complément du précédent, le Tableau C.4 présente les effectifs réels utilisés pour le test. Ceux-ci correspondent au nombre d'énoncés en *because* et *cos* comportant 0, 1, 2, 3, et 4 marques ou plus. Ici encore, la différence de distribution est significative : $\chi^2 = 19{,}352$ pour 4 ddl, $p \leqslant 0{,}001$.

	BECAUSE	COS
Enoncés sans marque	11	43
Enoncés avec 1 marque	7	12
Enoncés avec 2 marques	10	8
Enoncés avec 3 marques	8	4
Enoncés avec 4+ marques	13	7
TOTAL	49	74

TAB. C.4 – Densité des marques du travail de formulation

C.4 Force argumentative : *so* et *but*

Dans le cadre de notre réflexion sur la force argumentative comparée des deux marqueurs, nous avons évoqué la plus grande probabilité pour voir *so* apparaître à droite d'une proposition en *because* et *but* à droite d'une proposition en *cos*. Cette remarque est fondée sur les chiffres du Tableau C.5, témoins d'une distribution significativement différente ($\chi^2 = 6{,}073$ pour 2 ddl, $p \leqslant 0{,}05$).

	BECAUSE	COS
Enoncés suivis de *so*	13	8
Enoncés suivis de *but*	3	10
Autres énoncés	33	56
TOTAL	49	74

TAB. C.5 – Densité des marques du travail de formulation

Références bibliographiques

J. of Pragmatics : *Journal of Pragmatics*

J. of Phonetics : *Journal of Phonetics*

P. U. : Presses Universitaires

CUP : Cambridge University Press

ADAMCZEWSKI, Henri & DELMAS, Claude (1982, 1990). *Grammaire linguistique de l'anglais.* Paris : Armand Colin.

ARMSTRONG, Lilias Eveline & WARD, Ida Caroline (1926). *A Handbook of English Intonation.* Londres : Heffer.

AUROUX, Sylvain *et al.* (1996). *La philosophie du langage.* Paris : PUF.

BADDELEY, Alan (1990). *Human Memory. Theory and Practice.* Boston : Allyn and Bacon.

BADER, Françoise (1986). Structure de l'énoncé en Indo-Européen. *Bulletin de la Société de Linguistique de Paris* Tome LXXXI : 71-120.

BARNSLEY, Michael F. (1988). *Fractals everywhere.* San Diego, CA : Academic Press.

BATLINER, Anton & OPPENRIEDER, Wilhelm (1988). Rising intonation: not passed away but still alive. *J. of Pragmatics* 12 : 227-233.

BEAUGENDRE, Frédéric (1995). Modèles de l'intonation pour la synthèse. in *Ecole thématique « Fondements et perspectives en Traitement Automatique de la Parole ».* CNRS, Marseille-Luminy.

BECKMAN, Mary & AYERS, Gayle (1994). *Guidelines for ToBI Labelling, version 2.0.* Linguistics Department. Ohio State University.

BELL, Alexander Melville (1859). *The elocutionary manual.* Londres : Hamilton, Adams and Co.

BELL, Alexander Melville (1867). *Visible Speech, the science of universal alphabetics, or self-interpreting physiological letters.* Londres : Simpkin, Marshall and Co.

BENVENISTE, Emile (1966, 1974). *Problèmes de linguistique générale, I & II.* Paris : Gallimard.

BIBER, Douglas (éd.) (1999). *Longman Grammar of Spoken and Written English.* Harlow, Essex : Longman.

BILGER, Mireille *et al.* (1997). Transcription de l'oral et interprétation. Illustration de quelques difficultés. *Recherches sur le français parlé* 14 : 57-86.

BILGER, Mireille (1997). Coordination : analyses syntaxiques et annotations. *Recherches sur le français parlé* 14 : 256-272.

BLANCHE-BENVENISTE, Claire & JEANJEAN, Colette (1987). *Le français parlé : transcription et édition.* Paris : Didier.

BLANCHE-BENVENISTE, Claire (1997). Transcriptions et technologies. *Recherches sur le français parlé* 14 : 87-99.

BOLINGER, Dwight (éd.) (1972). *Intonation, Selected Readings.* Harmondsworth : Penguin Books.

BOLINGER, Dwight (1949). Intonation and analysis. *Word* 5 (3) : 248-254.

BOLINGER, Dwight (1951). Intonation: Levels versus configurations. *Word* 7 (3) : 199-210.

BOLINGER, Dwight (1955). Intersections of stress and intonation. *Word* 11 (2) : 195-203.

BOLINGER, Dwight (1986). *Intonation and its Parts, Melody in Spoken English.* Londres : Edward Arnold.

BOLINGER, Dwight (1989). *Intonation and its Uses, Melody in Grammar and Discourse.* Londres : Edward Arnold.

BOLINGER, Dwight (1998). Intonation in American English. in Hirst, Daniel & Di Cristo, Albert (éds.), *Intonation Systems, A Survey of Twenty Languages* : 45-55. Cambridge : CUP.

BOOMER, Donald (1965). Hesitation and grammatical encoding. *Language and Speech* 8 : 148-158.

BOOMER, Donald & DITTMAN, Allen (1962). Hesitation pauses and juncture pauses in speech. *Language and Speech* 5 : 215-220.

BOUQUIAUX, Luc (1986). Syntaxématique. Définition et classement des unités syntaxiques ou catégories grammaticales ou parties du discours. *Modèles Linguistiques* 8 (1) : 63-68.

BOURCET, Patrice & LIENARD, Pierre (1987, 2e éd. 1994). Acoustique fondamentale. in Mercier, Denis (éd.), *Le Livre des techniques du son* : 1 : 13-43. Paris : Eyrolles.

BRENET, Michel (1926). *Dictionnaire pratique et historique de la musique.* Paris : Armand Colin.

BUTCHER, Andrew (1980). Pause and Syntactic Structure. in Dechert, Hans W. & Raupach, Manfred (éds.), *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler* : 85-90. La Haye : Mouton.

BUTLER, Charles (1634, 1910). *English Grammar.* Niemeyer : Halle A. S.

CANDEA, Maria (1997). Peut-on définir la pause dans le discours comme un lieu d'absence de toute marque ? *Travaux linguistiques du CERLICO* 10 : 231-244.

CANDEA, Maria (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits « d'hésitation » en français oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle.

CANDEA, Maria & LEFEUVRE, Florence (à paraître). Les pics d'intensité dans le récit oral spontané. *Travaux linguistiques du CERLICO* 17.

CAPPEAU, Paul (1997). Données erronées : quelles erreurs commettent les transcripteurs ? *Recherches sur le français parlé* 14 : 117-126.

CHAFE, Wallace (1980). Some Reasons for Hesitating. in Dechert, Hans W. & Raupach, Manfred (éds.), *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler* : 169-180. La Haye : Mouton.

CHOMSKY, Noam (1965). *Aspects of the Theory of Syntax.* Cambridge, MA : MIT Press.

COLLINGE, N.E. (éd.) (1990). *An Encyclopædia of Language.* Londres : Routledge.

COOK, Mark (1971). The incidence of filled pauses in relation to part of speech. *Language and Speech* 14 : 135-139.

COPPIETERS, René (1997). Quelques réflexions sur la question des données : corpus et intuitions. *Recherches sur le français parlé* 14 : 21-41.

COTTE, Pierre (éd.) (1999). *Langage et linéarité.* Villeneuve-d'Ascq : P. U. du Septentrion.

COUPER-KUHLEN, Elizabeth & SELTING Margret (éds.) (1996). *Prosody in Conversation. Interactional Studies.* Cambridge : CUP.

COUPER-KUHLEN, Elizabeth (1996). Intonation and Clause Combining in Discourse: the Case of Because. *Pragmatics* 6 (3) : 389-426.

CROMPTON, Andrew S. (1980). Intonation: Tonetic stress versus levels versus configurations. *Word* 31 : 151-198.

CRUTTENDEN, Alan (1986, 1997). *Intonation.* Cambridge : CUP.

CRYSTAL, David (1969). *Prododic Systems and Intonation in English.* Cambridge : CUP.

CRYSTAL, David (1972). The Intonation System of English. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 110-136. Harmondsworth : Penguin Books.

CULIOLI, Antoine (1973). Sur quelques contradictions en linguistique. *Communications* 20 : 83-91.

CULIOLI, Antoine (1978). Valeurs modales et opérations énonciatives. *Le Français Moderne* 46 (4).

CULIOLI, Antoine (1985). *Notes du séminaire de DEA 1983-1984.* Poitiers : Atelier de Reprographie de l'Université de Poitiers.

CULIOLI, Antoine (1990, 1999). *Pour une linguistique de l'énonciation.* Gap : Ophrys.

CUTLER, Anne & LADD, D. Robert (éds.) (1983). *Prosody: Models and Measurements.* Berlin : Springer.

DANEŠ, František (1960). Sentence intonation from a functional point of view. *Word* 16 : 34-54.

DANEŠ, František (1972). Order of Elements and Sentence Intonation. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 216-232. Harmondsworth : Penguin Books.

DANON-BOILEAU, Laurent & MOREL, Mary-Annick (éds.) (1999). *Oral-Ecrit, Formes et théories. Faits de Langue* 13. Gap : Ophrys.

DANON-BOILEAU, Laurent & MOREL, Mary-Annick (1996). Intonation et intention. Du suprasegmental au verbal. « Le malheur de la question, c'est la réponse ». in *Actes du colloque international sur « Le questionnement social »*. Rouen.

DANON-BOILEAU, Laurent, MOREL, Mary-Annick & RIALLAND, Annie (1992). Intonation et structure de l'énoncé oral. in *Actes du Séminaire Dialogue, GDR-PRC Communication Homme-Machine*. Dourdan.

DANON-BOILEAU, Laurent (1987). *Enonciation et référence*. Paris : Ophrys.

DANON-BOILEAU, Laurent (1992). Ce que « ça » veut dire : les enseignements de l'observation clinique. in Morel, Mary-Annick & Danon-Boileau, Laurent (éds.), *La deixis* : 415-427. Paris : PUF.

DE COLA-SEKALI, Martine (1991a). *Connexion inter-énoncés et structuration des relations temporelles et argumentatives en anglais contemporain. Une étude énonciative des connecteurs polyvalents SINCE et FOR*. Thèse de Doctorat de 3e cycle. Université Paris III – Sorbonne Nouvelle.

DE COLA-SEKALI, Martine (1991b). Connexions inter-énoncés et relations intersubjectives : l'exemple de *because, since*, et *for* en anglais. *Langages* 104 : 62-78.

DECHERT, Hans W. & RAUPACH, Manfred (éds.) (1980). *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler*. La Haye : Mouton.

DEESE, James (1980). Pauses, Prosody, and the Demands of Production in Language. in Dechert, Hans W. & Raupach, Manfred (éds.), *Temporal Variables in Speech. Studies in Honour of Frieda Goldman-Eisler* : 69-84. La Haye : Mouton.

DELATTRE, Pierre (1972). The Distinctive Function of Intonation. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 159-174. Harmondsworth : Penguin Books.

DELECHELLE, Gérard (1980). Because. *Etudes anglaises* 76 : 19-35.

DELECHELLE, Gérard (1984). La valeur causale de *as* et *with*. *Etudes anglaises* 90 : 249-258.

DELECHELLE, Gérard (1989). *L'expression de la cause en anglais contemporain. Etude de quelques connecteurs et opérations.* Thèse de doctorat d'Etat. Université Paris III.

DELMAS, Claude, ADAMS, Philip, DELECHELLE, Gérard, *et al.* (1993). *Faits de langue en anglais. Méthode et pratique de l'explication grammaticale.* Paris : Dunod.

DELMAS, Claude (1987). *Structuration abstraite et chaîne linéaire en anglais contemporain.* Paris : CEDEL.

DELOMIER, Dominique & MOREL, Mary-Annick (1986). Réflexions sur les systèmes de notation de l'oral. *DRLAV* 34-35 : 158-160.

DI CRISTO, Albert (1985). *De la microprosodie à l'intonosyntaxe.* Aix-en-Provence : Publications de l'Université de Provence.

DUBOIS, Jean *et al.* (1973). *Dictionnaire de linguistique.* Paris : Larousse.

DUCHET, Jean-Louis (1991). *Code de l'anglais oral.* Gap : Ophrys.

DUCROT, Oswald & SCHAEFFER, Jean-Marie (1995). *Nouveau dictionnaire encyclopédique des sciences du langage.* Paris : Le Seuil.

DUEZ, Danielle (1982). Silent and non-silent pauses in three speech styles. *Language and Speech* 25 : 11-28.

DUEZ, Danielle (1985). Perception of Silent Pauses in Continuous Speech. *Language and Speech* 28 (4) : 377-389.

DUEZ, Danielle (1991). *La pause dans la parole de l'homme politique.* Paris : Editions du CNRS.

DUEZ, Danielle (1995). Silence et pouvoir : une comparaison de trois discours de François Mitterrand. *Travaux du Cercle Linguistique d'Aix-en-Provence* 13 : 153-164.

DUEZ, Danielle (1999). La fonction symbolique des pauses dans la parole de l'homme politique. *Faits de Langue* 13 : 91-97.

DUEZ, Danielle (2001). Caractéristiques acoustiques et phonétiques des pauses remplies dans la conversation en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage* 20 : 31-48.

DUNCAN, Starkey D., Jr. (1974). Some Signals and Rules for Taking Speaking Turns in Conversations. in Weitz, Shirley (éd.), *Non-verbal Communication* : 298-311. New York : Oxford University Press.

EAGLES (1996). *Preliminary recommendations on corpus typology.* EAG-TCWG-CTYP/P. Pise : Consiglio Nazionale delle Ricerche. Istituto di Linguistica Computazionale.

ERPICUM, D. & PAGE, M. (éds.) (1986). Dimensions de l'interaction dans une conversation. *Cahiers du Département de Psychologie de l'Université de Montréal* 5.

FAURE, Georges (1962). *Recherches sur les caractères et le rôle des éléments musicaux dans la prononciation anglaise.* Paris : Didier.

FERRE, Gaëlle (2003). Discursive, Prosodic and Gestural Marking of Focalisation Pauses in British English. in Mettouchi, Amina & Ferré Gaëlle (éds.), *Actes de IP2003, Interfaces Prosodiques* : 265-270. Nantes : Université de Nantes.

FEUILLET, Jack (1992). Typologie de la subordination. *Travaux linguistiques du CERLICO* 5 : 7-28.

FÓNAGY, Ivan (1983). *La vive voix : Essais de psycho-phonétique.* Paris : Bibliothèque Scientifique Payot.

FRAISSE, Paul (1974). *Psychologie du rythme.* Paris : PUF.

FREI, Henri (1954). Critères de délimitation. *Word* 10 (2) : 136-145.

FRY, Dennis Butler (1979). *The Physics of Speech.* Cambridge : CUP.

FUCHS, Catherine (1993). Position, portée et interprétation des circonstants. in Guimier, Claude (éd.), *1001 circonstants* : 253-283. Caen : P. U. de Caen.

GAMMA, Erich *et al.* (1995). *Design Patterns: Elements of Reusable Oject-Oriented Software.* Boston : Addison-Wesley.

GARDNER, Rod (2002). *When Listeners Talk: Response Tokens and Listener Stance.* Amsterdam/Philadelphia : John Benjamins.

GARVIN, Paul L. & MATHIOT, Madeleine (1958). Fused units in prosodic analysis. *Word* 14 (2-3) : 178-186.

GELUYKENS, Ronald (1987). Intonation and speech act type: An experimental approach to rising intonation in queclaratives. *J. of Pragmatics* 11 : 483-494.

GELUYKENS, Ronald (1989). R(a)ising Questions: Question Intonation Revisited. *J. of Pragmatics* 13 : 567-575.

GOLDMAN-EISLER, Frieda (1968). *Psycholinguistics: Experiments in Spontaneous Speech.* New York : Academic Press.

GOLDMAN-EISLER, Frieda (1972). Pauses, Clauses, Sentences. *Language and Speech* 15 : 103-113.

GOODWIN, Charles (1981). *Conversational Organization: Interaction between Speakers and Hearers.* New York : Academic Press.

GRABE, Esther (1998). *Comparative Intonational Phonology: English and German.* Wageningen : Ponsen and Looijen bv.

GRIBENSKI, André (1957). La hauteur des sons : sa mesure, sa perception. *La Nature* 3269 : 359-363.

GRICE, Herbert Paul (1957). Meaning . *Philosophical Review* 66 : 377-388.

GROSJEAN, François & DESCHAMPS, Alain (1972). Analyse des variables temporelles du français spontané. *Phonetica* 26 : 129-156.

GROSJEAN, François & DESCHAMPS, Alain (1973). Analyse des variables temporelles du français spontané II : Comparaison du français oral dans la description avec l'anglais (description) et avec le français (interview radiophonique). *Phonetica* 28 : 191-226.

GROSJEAN, François & DESCHAMPS, Alain (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica* 31 : 144-184.

GUAÏTELLA, Isabelle (1999). Réitérations et énumérations : approche des fonctions discursives et interactives de l'intonation. *Faits de Langue* 13 : 150-156.

GUERON, Jacqueline (1993). La grammaire générative. in Cotte, Pierre *et al.* (éd.), *Les théories de la grammaire anglaise en France.* Paris : Hachette.

GUILLAUME, Gustave (1939). Comment se fait un système grammatical. in *Conférences de l'Institut de Linguistique de l'Université de Paris* 7 : 43-57. Paris : Boivin.

GUILLAUME, Gustave (1952). La langue est-elle ou n'est-elle pas un système ? *Cahiers de Linguistique Structurale* 1 : 3-30.

GUIMIER, Claude (1993). L'établissement d'un corpus de circonstants. in Guimier, Claude (éd.), *1001 circonstants* : 11-45. Caen : P. U. de Caen.

GUSSENHOVEN, Carlos (1984). *On the grammar and semantics of sentence accents.* Dordrecht : Foris.

GUSTAFSON-CAPKOVA, Sofia & MEGYESI, Beata (2002). Silence and Discourse Context in Read Speech and Dialogues in Swedish. in Bel, Bernard & Merlien, Isabelle (éds.), *Proc. Speech Prosody 2002 conference* : 363-366. Aix-en-Provence : Publications de l'Université de Provence.

HADDING-KOCH, Kerstin (1961). *Acoustico-Phonetic Studies in the Intonation of Southern Swedish.* Lund : C.W.K. Gleerup.

HAEGEMAN, Liliane & GUERON, Jacqueline (1999). *English Grammar: a Generative Perspective.* Oxford : Blackwell.

HALLIDAY, Michael A. K. (1967). *Intonation and Grammar in British English.* La Haye : Mouton.

HALLIDAY, Michael A. K. & HASAN, Ruqaiya (1976). *Cohesion in English.* Londres : Longman.

HARRIS, Zellig Sabbettai (1951). *Methods in Structural Linguistics.* Chicago : University of Chicago Press.

HARRIS, Zellig Sabbettai (1954). Distributional structure. *Word* 10 (2) : 146-162.

HART, John (1569). *An orthographie: Conteyning the Due Order and Reason Howe to Write and Paint Thimages of Manes Voices, most like to Life or Nature.* Londres : William Seres.

HAWKINS, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech* 14 : 277-288.

HAZAEL-MASSIEUX, Marie-Christine (1983). Partie de la phrase ou parties du discours ? Contribution à une recherche sur les unités du discours. *Travaux du Cercle Linguistique d'Aix-en-Provence* 1 : 87-103.

HIRST, Daniel & DI CRISTO, Albert (éds.) (1998). *Intonation Systems, A Survey of Twenty Languages.* Cambridge : CUP.

HIRST, Daniel & DI CRISTO, Albert (1998). A survey of intonation systems. in Hirst, Daniel & Di Cristo, Albert (éds.), *Intonation Systems, A Survey of Twenty Languages* : 1-44. Cambridge : CUP.

HIRST, Daniel (1974). Intonation and Context. *Linguistics* 141 : 5-16.

HIRST, Daniel (1998). Intonation in British English. in Hirst, Daniel & Di Cristo, Albert (éds.), *Intonation Systems, A Survey of Twenty Languages* : 56-77. Cambridge : CUP.

HOCKETT, Charles F. (1942). A System of Descriptive Phonology. *Language* 18 : 31-21.

HOGG, Richard, BLAKE, Norman, LASS, Roger & ROMAINE, Suzanne (éds.) (1992-1999). *The Cambridge History of the English Language.* Cambridge : CUP.

HOUSE, David (2003). Hesitation and Interrogative Swedish Intonation. *PHONUM* 9 : 185-188.

HUGGINS, A.W.F (1964). Distortions of the temporal pattern of speech. *Journal of the Acoustical Society of America* 36 : 1055-1064.

JAKOBSON, Roman & WAUGH, Linda (1979, trad. fr. 1980). *La charpente phonique du langage.* Paris : Editions de Minuit.

JOHANNSON, Stig & HOFLAND, Knut (1989). *Frequency Analysis of English Vocabulary and Grammar, vol.1 & vol.2.* Oxford : Clarendon Press.

JONES, Daniel (1909, 1956). *The Pronunciation of English.* Cambridge : CUP.

JONES, Daniel (1950). *The Phoneme, its Nature and Use.* Cambridge : Heffer.

KEENAN, Edward & COMRIE, Bernard (1977). Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry* 8 : 62-100.

KERBRAT-ORECCHIONI, Catherine (1990, 1998). *Les interactions verbales, Tome 1 : Approche interactionnelle et structure des conversations.* Paris : Armand Colin.

KINGDON, Roger (1958). *The groundwork of English intonation.* Londres : Longman.

KITANTOU, Mpaya (1987, 2e éd. 1994). La perception auditive. in Mercier, Denis (éd.), *Le livre des techniques du son* : 1 : 155-181. Paris : Eyrolles.

KRECKEL, Marga (1981). Tone units as message blocks in natural discourse: segmentation of face-to-face interaction by naive, native speakers. *J. of Pragmatics* 5 : 459-476.

LADD, D. Robert (1996). *Intonational Phonology.* Cambridge : CUP.

LAPAIRE, Jean-Rémi & ROTGÉ, Wilfrid (1991). *Linguistique et grammaire de l'anglais.* Toulouse : P. U. du Mirail.

LAPAIRE, Jean-Rémi & ROTGÉ, Wilfrid (1992). *Réussir le commentaire grammatical de textes : CAPES, Agrégation, Anglais.* Paris : Ellipses.

LAZARD, Gilbert (1986). Deux échelles de transitivité. *Actances* 2 : 59-68.

LAZARD, Gilbert (1995). La notion de distance actancielle. in Bouscaren, J., Franckel, J.-J. & Robert S. (éds.), *Langues et langage. Problèmes et raisonnement en linguistique – Mélanges offerts à Antoine Culioli* : 135-146. Paris : PUF.

LE QUERLER, Nicole (1993). Subordination, thématisation, rhématisation. *Travaux linguistiques du CERLICO* 6 : 97-121.

LEHISTE, Ilse (1970). *Suprasegmentals.* Cambridge, MA : MIT Press.

LEHISTE, Ilse (1976). Influence of Fundamental Frequency Pattern on the Perception of Duration. *J. of Phonetics* 4 : 113-117.

LEKOMCEV, Ju. K. (1974). On certain formal characteristics of linguistic hierarchies. *Linguistics* 131 : 39-48.

LEON, Pierre R. & MARTIN, Philippe (1972). Machines and Measurements. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 30-47. Harmondsworth : Penguin Books.

LEROY, Christine (1985). La notation de l'oral. *Langue Française* 65 : 6-16.

LIEBERMAN, Philip (1965). On the acoustic basis of the perception of intonation by linguists. *Word* 21 : 40-54.

LOCAL, John, WELLS, W.H.G, & SEBBA, M. (1985). Phonology for Conversation: Phonetic Aspects of Turn Delimitation in London Jamaican. *J. of Pragmatics* 9 : 309-330.

MACLAY, Howard & OSGOOD, Charles E. (1959). Hesitation Phenomena in Spontaneous English Speech. *Word* 15 : 19-44.

MANDELBROT, Benoît (1975). *Les objets fractals : forme, hasard et dimension.* Paris : Flammarion.

McENERY, Tony & WILSON, Andrew (1996). *Corpus Linguistics.* Edimbourg : Edinburgh University Press.

MERCIER, Denis (éd.) (1987, 1994). *Le livre des techniques du son (3 tomes).* Paris : Eyrolles.

MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation, L'exemple du français.* Gap : Ophrys.

MOREL, Mary-Annick & RIALLAND, Annie (1993). L'énoncé oral complexe : les relatives en QUI. in Muller, Claude & Roulland, Daniel (éds.), *Travaux de linguistique du CERLICO* 6 : 221-243. Rennes : P. U. de Rennes.

MOREL, Mary-Annick (1992). L'opposition thème/rhème dans la structuration des dialogues oraux. *Journal of French Language Studies* 2 : 85-98.

MOREL, Mary-Annick (1992). Intonation et thématisation. *L'information grammaticale* 54.

MOREL, Mary-Annick (1995). Valeur énonciative des variations de hauteur mélodique en français. *Journal of French Language Studies* 5 : 189-202.

MOREL, Mary-Annick (1999). Complémentarité des indices du plan segmental et du plan suprasegmental dans l'oral spontané en français. in *Actes des journées d'Etudes sur les Corpus.* Copenhague.

MOREL, Mary-Annick (1999). Morphosyntaxe et intonation : Complémentarité des indices dans l'oral spontané en français. in *Actes du Congrès des Romanistes de Septembre 1998.* Mayence.

MOREL, Mary-Annick (2003). Parleur et écouteur. Formes intonatives couplées dans l'échange oral en français. in Mettouchi, Amina & Ferré, Gaëlle (éds.), *Actes de IP2003, Interfaces Prosodiques* : 293-300. Nantes : Université de Nantes.

MOREL, May-Annick & DANON-BOILEAU, Laurent (2000). Les productions sonores de l'écouteur du récit : coopération ou subversion ? *Revue Québécoise de Linguistique* 29 (1).

MORGENSTERN, Aliyah & SEKALI, Martine (à paraître). De la parataxe aux premiers connecteurs : les prémices de l'argumentation chez le jeune enfant. *Cahiers de l'ENS-LSH.*

MULLER, Charles (1992). *Initiation aux méthodes de la statistique linguistique.* Paris : Honoré Champion.

NICAISE, Alain & GRAY, Martin (1998). *L'intonation de l'anglais.* Paris : Nathan Université.

O'CONNOR, Joseph Desmond & ARNOLD, Gordon Frederick (1961, 1973). *Intonation of Colloquial English.* Londres : Longman.

OHALA, John J. (1994). The frequency code underlies the sound symbolic use of voice pitch. in Hinton, L., Nichols, J., & Ohala, J. J. (éds.), *Sound symbolism* : 325-347. Cambridge : CUP.

PADUČEVA, E. V. (1974). On the Structure of the Paragraph. *Linguistics* 131 : 49-58.

PALMER, Harold E. (1924). *A Grammar of Spoken English.* Cambridge : Heffer.

PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (2000). Analyse d'une composante articulatoire : l'intensité en anglais oral spontané. in *Journée d'étude, Laboratoire d'informatique et de mécanique pour les sciences de l'ingénieur (LIMSI-CNRS), séminaire de recherche de J.-B. Berthelin.* Communication affichée. Orsay.

PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (à paraître). The Intonation, Context and Syntax of Questions in Spoken French and English: an utterer-centred approach. in Hoey, Michael & Scott, Nelia (éds.), *Questions : Multiple Perspectives on a Common Phenomenon.* Liverpool : University of Liverpool Press.

PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (à paraître). Y a-t-il une intonation propre aux questions en anglais ? in Richard-Zapella, Jeanine (éd.), *Aspects pragmatiques du questionnement.* Paris : L'Harmattan.

PASSOT, Frédérique, SZLAMOWICZ, Jean & VIALLETON, Elodie (2001). Les interrogatives en anglais oral spontané : intonation et énonciation. *Anglophonia-Sigma* 8 : 125-152.

PASSOT, Frédérique (1999). *Accent de mot, accent de phrase : Aspects de l'accentuation et de l'intonation de l'anglais oral spontané.* Mémoire de DEA. Université de Paris III – Sorbonne Nouvelle.

PASSOT, Frédérique (2003). Statistiques suprasegmentales et hypothèses linguistiques : problèmes et perspectives. in *INTO-01, Intonation, Notation et Transcription de l'Oral.* Communication orale. Université de Rouen.

PIERREHUMBERT, Janet & HIRSCHBERG, Julia (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. in Cohen, Philip R., Morgna, Jerry et al. (éds.), *Intentions in Communication* : 271-311. Cambridge, MA : MIT Press.

PIERREHUMBERT, Janet (1980). *The Phonology and Phonetics of English Intonation.* Thèse de doctorat. MIT.

PIKE, Kenneth L. (1972). General Characteristics of Intonation. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 53-82. Harmondsworth : Penguin Books.

PIKE, Kenneth, L. (1945). *The Intonation of American English.* Ann Arbor, MI : University of Michigan Press.

PIOT, Mireille (1988). Coordination-Subordination : une définition générale. *Langue Française* 77 : 5-18.

QUIRK, Randolph *et al.* (1985). *A Comprehensive grammar of the English language.* Londres : Longman.

REBOUL, Anne & MOESCHLER, Jacques (1998a). *La pragmatique aujourd'hui : une nouvelle science de la communication.* Paris : Le Seuil.

REBOUL, Anne & MOESCHLER, Jacques (1998b). *Pragmatique du discours : de l'interprétation de l'énoncé à l'interprétation du discours.* Paris : Armand Colin.

REMI-GIRAUD, Sylvianne (1988). Le complément circonstanciel : problèmes de définition. in Rémi-Giraud, S. & Roman, A. (éds.), *Autour du circonstant* : 65-113. Lyon : P. U. de Lyon.

REY, Alain & REY-DEBOVE, Josette (dir.) (1994). *Le Petit Robert 1 : dictionnaire alphabétique et analogique de la langue française.* Paris : Dictionnaires Le Robert.

ROSSARI, Corinne (2000). *Connecteurs et relations de discours : des liens entre cognition et signification.* Nancy : P. U. de Nancy.

ROSSI, Mario (1971). Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 24 : 129-161.

ROSSI, Mario (1972). Le seuil différentiel de durée. in Valdman, Albert (éd.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre* : 435-450. La Haye : Mouton.

ROSSI, Mario (1977). L'intonation et la troisième articulation. *Bulletin de la Société de Linguistique de Paris* Tome LXII : 55-68.

ROSSI, Mario (1978). The perception of non repetitive intensity glides on vowels. *J. of Phonetics* 6 : 9-18.

ROSSI, Mario *et al.* (1981). *L'intonation, de l'acoustique à la sémantique.* Paris : Klincksieck.

ROSSI, Mario & CHAFCOULOFF, M. (1972a). Les niveaux intonatifs. *Travaux de l'Institut de Phonétique d'Aix en Provence* 1 : 167-176.

ROSSI, Mario & CHAFCOULOFF, M. (1972b). Recherche sur le seuil différentiel de fréquence fondamentale de la parole. *Travaux de l'Institut de Phonétique d'Aix en Provence* 1 : 179-185.

RUSH, James (2e éd. 1833). *The Philosophy of the Human voice, Embracing its Psychological history, together with a system of principles by which criticism in the Art of Elocution may be Rendered intelligible.* Philadelphie : Grigg and Elliott.

RUTHERFORD, William E. (1970). Some observations concerning subordinate clauses in English. *Language* 46 (1) : 97-115.

SACKS, Harvey, SCHEGLOFF, Emanuel A., JEFFERSON, Gail (1974). A simplest systematics for the organization of turn-taking for conversation. *Language* 50 (4) : 696-735.

SAUSSURE, Ferdinand de (1972). *Cours de linguistique générale.* Paris : Payot.

SAUVAGEOT, Aurélien (1992). *La structure du langage.* Aix-en-Provence : Publications de l'Université de Provence.

SAVILLE-TROIKE, Muriel (1985). The Place of Silence in an Integrated Theory of Communcation. in Tannen, Deborah & Saville-Troike, Muriel (éds.), *Perspectives on Silence* : 3-18. Norwood, NJ : Ablex.

SCHIFFRIN, Deborah (1987). *Discourse Markers.* Cambridge : CUP.

SCHLEPPEGRELL, Mary (1991). Paratactic *because. J. of Pragmatics* 16 : 323-337.

SCHUETZE-COBURN, Stephan, SHAPLEY, Marian & WEBER, Elizabeth (1991). Units of Intonation in Discourse: A Comparison of Acoustic and Auditory Analyses. *Language and Speech* 34 (3) : 207-234.

SELTING, Margret (1996). On the Interplay of Syntax and Prosody in the Constitution of Turn-Constructional Units and Turns in Conversation. *Pragmatics* 6 (3) : 357-388.

SHANNON, Claude E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal* 27 : 379-423, 623-655.

SHANNON, Claude E. (1950). Communication Theory — Exposition of Fundamentals. *IRE Transactions Information Theory* 1 : 44-47.

SHANNON, Claude E. & WEAVER, Warren (1949, trad. 1975). *Théorie mathématique de la communication.* Paris : Retz.

SHRIBERG, Elizabeth, BATES, Rebecca, STOLCKE, Andreas, *et al.* (1998). Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech? *Language and Speech* 41 (3-4) : 443-492.

SILVERSTEIN, Michael (1976). Hierarchy of Features and Ergativity. in Dixon, R.M.W. (éd.), *Grammatical Categories in Australian Languages* : 112-171. Canberra : Australian Institute of Aboriginal Studies.

SIMON, Anne Catherine (2002). *Segmentation et segmentation prosodiques du discours. Une approche multidimensionnelle et expérientielle.* Thèse de doctorat. Louvain-la-Neuve.

SIMPSON, John & WEINER, Edmund (1989, 2e éd.). *Oxford English Dictionary.* Oxford : Clarendon Press.

SPERBER, Dan & WILSON, Deirdre (1986, 1995). *Relevance – Communication and Cognition.* Oxford, Cambrige MA : Blackwell.

STEELE, Joshua (1775, 2e éd. 1971). *Prosodia Rationalis, or An Essay towards Establishing the Melody and Measure of Speech to be Expressed and Perpetuated by Peculiar Symbols.* Olms : Georg Publishers.

SVARTVIK, Jan & QUIRK, Randolph (1980). *A Corpus of English Conversation.* Lund : C.W.K. Gleerup.

SWERTS, Mark, STRANGERT, Eva & HELDNER, Mattias (1996). F_0 Declination in read-aloud and spontaneous speech. in *Proceedings of the International Conference on Spoken Language Processing* 3. Philadelphia.

SZLAMOWICZ, Jean (1997). *Etude des indices segmentaux et suprasegmentaux de structuration de l'anglais oral spontané.* Mémoire de DEA. Université Paris III – Sorbonne Nouvelle.

SZLAMOWICZ, Jean (2001). *Contribution à une approche intonative et énonciative du rôle des ligateurs dans la construction du discours en anglais oral spontané.* Thèse de doctorat. Université Paris III – Sorbonne Nouvelle.

SZLAMOWICZ, Jean (2002). Les pauses en anglais : de la faillite du silence à la structuration linguistique, ou de l'iconique au conventionnel. in Delmas, Claude & Roux, Louis (éds.), *Correct, incorrect en linguistique anglaise. CIEREC Travaux 113* : 157-174. Saint-Etienne : Publications de l'Université de Saint-Etienne.

TAGLICHT, Josef (1982). Intonation and the Assessment of Information. *Journal of Linguistics* 18 : 213-230.

TANNEN, Deborah & SAVILLE-TROIKE, Muriel (1985). *Perspectives on Silence.* Norwood, NJ : Ablex.

TANNEN, Deborah (1985). Silence: Anything But. in Tannen, Deborah & Saville-Troike, Muriel (éds.), *Perspectives on Silence* : 93-111. Norwood, NJ : Ablex.

TESNIERE, Lucien (1959, 1988). *Eléments de syntaxe structurale.* Paris : Klincksieck.

TOURATIER, Christian (1995). Oral et écrit, deux utilisations d'une même langue. *Travaux du Cercle Linguistique d'Aix-en-Provence* 13 : 55-61.

TRAGER, George L. (1972). The Intonation System of American English. in Bolinger, Dwight (éd.), *Intonation, Selected readings* : 83-86. Harmondsworth : Penguin Books.

TRAGER, George Leonard & SMITH, Henry Lee (1951). *An outline of English Structure.* Norman, OK : Battenberg Press.

VAISSIERE, Jacqueline (1983). Language-Independent Prosodic Features. in Cutler, Anne & Ladd, Robert D. (éds.), *Prosody : Models and Measurements.* Berlin : Springer.

VALIN, Roch (1964). La méthode comparative en linguistique historique et en psychomécanique du langage. *Cahiers de psychomécanique du langage* 6 : 1-57.

VALLEE, Michaël (2003). *Because* et *for* partagent-ils les mêmes propriétés énonciatives ? *CORELA — Cognition, Représentation, Langages* 1 (2).

VION, Robert (1992). Enonciation et interaction. *Sigma* 15 : 79-104.

VION, Robert (1995). Les unités du dialogue. *Travaux du Cercle Linguistique d'Aix-en-Provence* 13 : 165-180.

WALKER, John (1781, 1974). *Elements of elocution.* Londres : Scolar Press.

WALKER, John (1787, 1974). *The Melody of Speaking, Delineated; or Elocution taught like music, by visible signs.* Londres : Scolar Press.

WELLS, John Christopher (1990). *Longman Pronunciation Dictionary.* Londres : Longman.

WIENER, Norbert (1948). *Cybernetics.* New York : J. Wiley.

WITTING, C. (1962). A method of evaluating listeners' transcriptions of intonation on the basis of instrumental data. *Language and Speech* 15 (3) : 138-150.

WYLD, Henry (2001). *Subordination et énonciation.* Gap : Ophrys.

YULE, George (1980). Speakers' Topics and Major Paratones. *Lingua* 52 : 33-47.

ZELLNER, Brigitte (1994). Pauses and the temporal structure of speech. in Keller, E. (éd.), *Fundamentals of speech synthesis and speech recognition* : 41-62. Chichester : John Wiley.

250 références

Index

Armstrong, L.E., 57

Arnold, G.F., 57

Artstein, R., 97, 98

Auroux, S., 33, 34

Ayers, G., 68

Baddeley, A., 356, 373

Barnsley, M., 140

Beaugendre, F., 74

Beckman, M., 68

Bell, A.M., 56, 57

Benveniste, E., 27, 42, 106

Biber, D., 148, 153, 203, 205, 234, 237, 422

Bilger, M., 178

Blake, N., 283

Blanche-Benveniste, C., 172, 173, 175, 178

Boersma, P., 180

Bolinger, D., 62, 63, 65

Boomer, D., 366, 367

Bouquiaux, L., 36

Bourcet, P., 75

Brenet, M., 52

Butler, C., 53

Campione, E., 345

Candea, M., 180, 344–346, 367, 376, 377, 383, 432

Cappeau, P., 177

Chafcouloff, M., 77, 78

Chafe, W., 372, 373

Chaucer, G., 229

Chen, A., 137, 325, 338, 341, 379, 400, 433

Chomsky, N., 93, 95, 372

Collinge, N.E., 96

Colombo, F., 179

Comrie, B., 95

Conrad, S., 148

Coppieters, R., 153, 178

Coulthard, M., 103

Couper-Kuhlen, E., 249, 266, 270, 275, 276, 280, 285

Crompton, A., 60–64

Cruttenden, A., 53, 55–58

Crystal, D., 57, 61, 189
Culioli, A., 27, 28, 106, 200, 318

Daneš, F., 188
Danon-Boileau, L., 41, 44, 68–70, 108–118, 173, 318–320, 323, 356
Delattre, P., 39, 68
Deléchelle, G., 259, 261, 262
Delmas, C., 33
Deschamps, A., 242, 261, 345, 348, 349, 352, 378
De Cola-Sekali, M., 201, 202, 218, 222–224, 298, 299, 301, 302, 318–320
Dittman, A., 366, 367
Di Cristo, A., 78, 80–82, 85, 87, 88
Dubois, J., 27
Ducrot, O., 19, 90, 92, 94, 144
Duez, D., 345, 346, 350–352, 361, 364
Duncan, S., 394, 395, 405

Einstein, A., 143
Erpicum, D., 36

Faure, G., 52, 54, 55, 76, 79
Ferré, G., 369, 378
Feuillet, J., 197, 198
Fraisse, P., 79, 82, 345
Frei, H., 189

Gamma, E., 21
Gardner, R., 322
Garvin, P., 188
Goldman-Eisler, F., 81, 348, 350, 361
Goodwin, C., 51
Gougenheim, G., 144
Grabe, E., 151
Gribenski, A., 77
Grice, H.P., 98, 100
Grosjean, F., 242, 261, 345, 348, 349, 352, 378
Gsell, R., 119
Guéron, J., 140, 141, 197
Guillaume, G., 20, 42, 141
Gussenhoven, C., 65, 136, 137, 325, 338, 341, 379, 400, 433

Hadding-Koch, K., 51
Haegeman, L., 140, 141, 197
Halliday, M.A.K., 284, 310
Harris, Z., 90–92, 388, 389
Hart, J., 53
Hasan, R., 284, 310
Hirschberg, J., 66
Hockett, C., 31
Hockett, C.F., 90
Hofland, K., 149, 152, 203
Hogg, R., 283

House, D., 137

Huggins, A.W.F., 81

Jakobson, R., 39

Jeanjean, C., 178

Jefferson, G., 393, 397, 399, 401, 403, 405, 406

Johannson, S., 149, 152, 203

Jones, D., 57, 62, 76, 155

Jouët-Pastré, G., 115

Keenan, E., 95

Kerbrat-Orecchioni, C., 29, 36, 103, 104, 154, 176, 395, 403

Kingdon, R., 57, 61

Kitantou, M., 77, 79, 80, 83, 85, 86

Ladd, D.R., 65, 66

Lapaire, J.-R., 206, 297

Lass, R., 283

Lazard, G., 96

Leech, G., 149

Lefeuvre, F., 432

Lehiste, I., 80, 88

Lekomcev, J., 32

Leroy, C., 147, 171, 173, 174

Lieberman, P., 78

Liénard, P., 75

Local, J., 389, 390

Mandelbrot, B., 139, 142, 143

Marlowe, C., 229

Martin, P., 180, 181

Mathiot, M., 188

McEnery, T., 146

Milgram, M., 73

Moeschler, J., 20, 23, 24, 26, 28

Morel, M.-A., 41, 44, 68–70, 108–118, 173, 356, 436

Morgenstern, A., 301, 302

Muller, C., 213, 327, 328, 362

Nolan, F., 151

O'Connor, J.D., 57

Ohala, J., 136

Page, M., 36

Palmer, H.E., 57

Passot, F., 115, 119, 121, 126, 337, 431, 433

Perrin, J., 139

Pierrehumbert, J., 65, 66

Pike, K., 52, 58, 63

Platon, 33

Quirk, R., 150, 195, 198–200, 203, 205, 233

Reboul, A., 20, 23, 24, 26, 28

Rémi-Giraud, S., 197
Rey, A., 33
Rey-Debove, J., 33
Rietveld, T., 137, 325, 338, 341, 379, 400, 433
Romaine, S., 283
Rossi, M., 31, 77, 80–82, 87, 88
Rotgé, W., 206, 297
Roulet, E., 103
Rush, J., 56
Rutherford, W., 262

Sacks, H., 393, 397, 399, 401, 403, 405, 406
Sampson, G., 150
Saussure, F. de, 19, 20
Schaeffer, J.-M., 19, 90, 92, 94, 144
Schegloff, E., 393, 397, 399, 401, 403, 405, 406
Schiffrin, D., 221, 285–287, 289–292, 294, 295, 298
Schleppegrell, M., 268, 269, 281, 282
Shannon, C., 98, 99, 101, 102
Silverstein, M., 96
Simon, A.C., 69, 70, 105
Simpson, J., 175, 228
Sinclair, J., 103, 173
Smith, H.L., 58, 62, 394
Sperber, D., 23, 100, 101, 103
Steele, J., 54
Svartvik, J., 150
Szlamowicz, J., 115, 118, 121, 129, 180, 431, 433

Tannen, D., 375
Tesnière, L., 43
Tinothai, K., 119
Trager, G.L., 58, 62, 394

Valin, R., 20, 23, 26
Véronis, J., 330, 345
Vialleton, E., 121, 154, 162, 180, 431, 433

Wagner, P., 52
Walker, J., 54, 57
Ward, I.C., 57
Waugh, L., 39
Weaver, E., 98
Weiner, E., 175, 228
Wiener, N., 99
Wilson, A., 146
Wilson, D., 23, 100, 101, 103
Witting, C., 80
Wyld, H., 196, 200, 201

Zarader, J.L., 73