

THESE DE DOCTORAT DE

L'UNIVERSITE DE NANTES
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 602
Education, Langages, Interaction, Cognition, Clinique
Spécialité : *Sciences du Langage*

Par

Manon LELANDAIS

Expression multimodale de la subordination en anglais

Thèse présentée et soutenue à Nantes, le 2 juillet 2019
Unité de recherche : UMR 6310 LLING
Thèse N° :

Rapporteuses avant soutenance :

Sophie HERMENT Professeur des Universités, Université Aix-Marseille
Laure LANSARI Maître de Conférences, Université Paris Diderot

Composition du Jury :

Présidente du jury : Sabrina BENDJABALLAH
Rapporteuse : Sophie HERMENT
Rapporteuse : Laure LANSARI
Examineur : Maarten LEMMENS
Directrice de thèse : Catherine COLLIN
Co-directrice de thèse : Gaëlle FERRE

Directrice de Recherche, Université de Nantes
Professeur des Universités, Université Aix-Marseille
Maître de Conférences, Université Paris Diderot
Professeur des Universités, Université de Lille
Professeur des Universités, Université de Nantes
Maître de Conférences, Université de Nantes

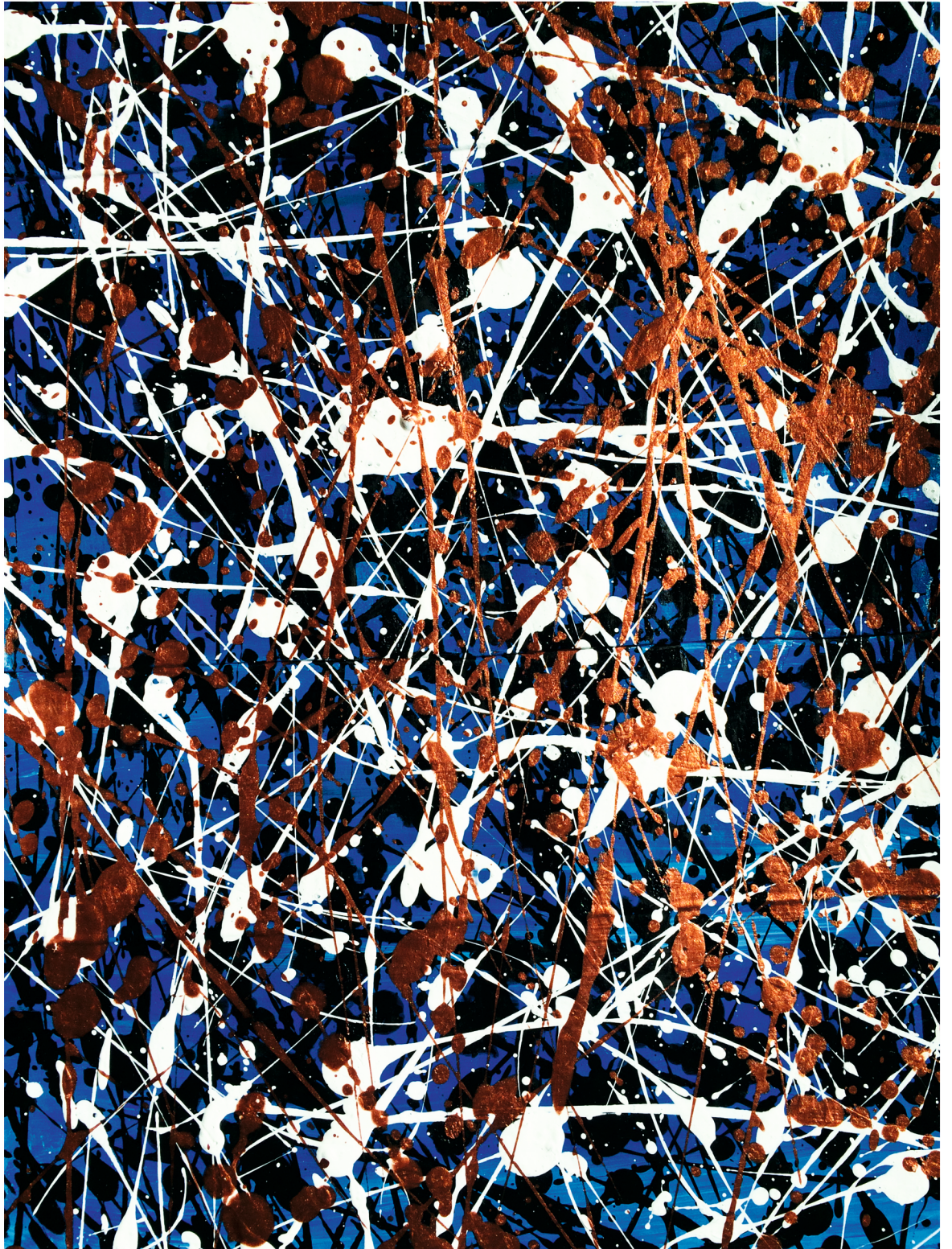


Illustration : Amélie Richard

N° attribué par la bibliothèque

--	--	--	--	--	--	--	--	--	--

THÈSE

pour obtenir le grade de

Docteur de l'UNIVERSITÉ DE NANTES

Spécialité : **Anglais**

préparée au laboratoire **Laboratoire de Linguistique de Nantes (LLING)**

dans le cadre de l'École Doctorale **Éducation, Langages, Interactions,
Cognition, Clinique (ELICC)**

présentée et soutenue publiquement

par

Manon LELANDAIS

le 2 juillet 2019

Titre:

Expression multimodale de la subordination en anglais

Directrice de thèse: **Catherine Collin**

Co-directrice de thèse: **Gaëlle Ferré**

Jury

SABRINA BENDJABALLAH	Université de Nantes	Présidente du jury
SOPHIE HERMENT	Université Aix-Marseille	Rapporteure
LAURE LANSARI	Université Paris Diderot	Rapporteure
MAARTEN LEMMENS	Université de Lille	Examineur
CATHERINE COLLIN	Université de Nantes	Directrice de thèse
GAËLLE FERRÉ	Université de Nantes	Co-directrice de thèse

Résumé

À partir d'un corpus vidéo de conversation spontanée en anglais, notre travail de thèse s'attache à déterminer si plusieurs types syntaxiques de constructions subordonnées opérant au niveau de la modification (i.e. appositives, circonstancielles, relatives déterminatives) expriment le même degré d'intégration à leur environnement co-textuel. Dans la littérature syntaxique, les constructions subordonnées sont décrites comme des structures dépendantes, qui spécifient ou élaborent le contenu d'une autre proposition. Si de nombreux travaux se consacrent à leurs caractéristiques verbales ou prosodiques, peu d'études choisissent d'examiner l'articulation entre les différentes modalités communicatives lors de leur production, et de fournir une vision plus nuancée de leur intégration. Or, le développement d'outils et de procédés analytiques pouvant être mis en relation facilite la description des subordonnées en tant que phénomènes multimodaux. En montrant que les constructions sous étude n'expriment pas une dépendance uniforme à leur environnement selon la façon dont les locuteurs utilisent les modalités prosodique et gestuelle pour exprimer plus ou moins de démarcation, les résultats suggèrent d'une part que les appositives sont produites avec davantage de rupture que les autres types syntaxiques, et d'autre part que la création d'une rupture s'appuie majoritairement sur des moyens davantage prosodiques que gestuels. Deux tests de perception auprès de participants naïfs anglophones étoffent ces résultats en montrant d'une part qu'une plus grande rupture est perçue dans les appositives que dans les autres types, et d'autre part que la prosodie joue un plus grand rôle dans la perception de la rupture que la gestualité. Une représentation schématique de l'impact des constructions subordonnées sur la construction de l'interaction est ensuite introduite. Le modèle multiparamétrique de la subordination que nous proposons est basé sur le fait que ces constructions sont des pratiques interactionnelles qui offrent une reconstruction interprétative du discours.

Abstract

Based on a video recording of conversational British English and within the framework of Multimodal Discourse Analysis, this thesis tests whether three different syntactic types of subordinate structures operating at the level of modification (i.e. appositive clauses, adverbial clauses, restrictive relative clauses) are evenly integrated to their environment. Subordinate constructions have been described in linguistics as dependent forms elaborating on primary elements of discourse. Although their verbal and vocal characteristics have been deeply analysed, few studies have focused on the articulation of the different communicative modalities in their production or provided a qualified picture of their integration. The main hypothesis is based on the capacity of subordinate constructions to show distinct forms of autonomy depending on their syntactic type, thus expressing different degrees of break. Beyond showing that subordinate constructions are not evenly dependent on their environment depending on how speakers use the prosodic and kinetic modalities to express greater (in)dependency, the results suggest on the one hand that appositive clauses show a greater break than the other syntactic types, and on the other hand that the creation of a break mainly relies on prosodic cues. As far as the perception of subordinate constructions is concerned, two experiments with naïve speakers of British English show that a greater demarcation is perceived in appositive clauses than in the two other syntactic types of subordinate constructions, and that prosody plays a bigger role than that of hand gestures in the perception of demarcation in discourse. A schematic representation of the impact of subordinate clauses on interaction is eventually proposed. The multiparameter framework for modelling subordination in spontaneous conversation is based on the fact that subordinate constructions are practices in interaction that offer an interpretative reconstruction of discourse.

Remerciements

Je suis très reconnaissante à Catherine Collin pour avoir accepté de diriger ma thèse, et pour sa confiance depuis mes années de Master. Je remercie également ma co-directrice de thèse, Gaëlle Ferré, pour ses conseils, ses encouragements, et la finesse de ses enseignements tout au long de mon parcours universitaire. Je tiens à ce qu'elle m'a transmis et à nos échanges.

Merci à Sabrina Bendjaballah, Laure Lansari, Maarten Lemmens, et à Sophie Herment d'avoir accepté de participer à mon jury.

Je suis très fière d'avoir travaillé sous les conseils éclairés et la bienveillance de Camille Debras et de Roxane Bertrand, membres de mon comité de suivi de thèse. Ce sont deux personnes que j'admire au sein de ma communauté scientifique.

Mes remerciements vont également aux locutrices et locuteurs du corpus ENVID, pour leur entrain et leur ressource. Ce corpus collaboratif m'a par ailleurs permis de bénéficier d'une plus grande collection de données, et je suis reconnaissante envers toutes les personnes ayant contribué à sa construction. L'une des interactions a notamment été enregistrée grâce à Stéphane Dupont et Jean-François Surget, ingénieurs au PAM de l'Université de Nantes. Merci également à David Imbert pour ses conseils sur les statistiques, et à Romain Legé pour sa participation aux stimuli de test.

Cette thèse a été réalisée grâce à un contrat doctoral, obtenu par une allocation ministérielle. Je remercie les membres du laboratoire du LLING pour m'avoir encouragée lors de ma préparation aux auditions des bourses de thèse. Je remercie également Anne-Laure Besnard et Jiyoung Choi pour leur aide et leur générosité.

Je salue les formidables personnes que j'ai rencontrées au fil des conférences, ainsi que nos discussions, que j'espère reprendre bientôt : Alice Cravotta, Aurélie Goujon, Claudia Ribeiro, Corine Astésano, David House, Katherina Walper-Gormaz, Laury

Garnier, Leonardo Contreras, Mike Stevens, Rafèu Sichel-Bazin, Rémi Lamarque, Richard Bonnie. Merci également à Sotaro Kita de m'avoir fait découvrir le cinéma de Kore-Eda.

Merci à mes parents, Maryse Morisot et Loïc Lelandais, et à ma famille pour leur soutien infaillible. Merci à Adeline R, Aline V, Alison R, Axelle D, Benoit D, Flavie E, Florence L, Julie B, Lucie H, Maëlle J, Marion T, Silvan G, et Virginie G. Merci à la famille d'Amélie. La bise aux Usées Coutumes, à Marine, Laura, Chloé, Joy, Paula, Justine, Aurélie, Paillette, et Jessy. Bises aussi aux Culbuteuses.

Enfin, merci à Amélie, pour tout le reste.

Table des matières

Résumé	iii
Abstract	iv
Remerciements	v
Table des matières	vii
Table des figures	xi
Liste des tableaux	xv
Conventions de transcription	xvii
Introduction	1
I Première partie : cadre théorique et méthodologie	21
1 Cadre théorique	23
1.1 La linguistique interactionnelle	23
1.2 L'Analyse de Discours Multimodale	24
1.3 La grammaire des constructions	25
1.4 La grammaire cognitive	29
1.5 Unités théoriques d'analyse	34
1.5.1 La proposition	35
1.5.2 Les <i>tone-units</i>	37
1.5.3 Les unités gestuelles	40
1.6 Travaux antérieurs sur la subordination	42
1.6.1 La subordination en syntaxe	43
1.6.2 La subordination dans le discours	49
1.6.3 La subordination prosodique	54
1.6.4 La subordination gestuelle	55
1.7 Hypothèses	58
2 Corpus et méthodologie	61
2.1 Enregistrement du corpus	61
2.2 Transcription et annotation du corpus	63

2.2.1	Le groupe intonatif en conversation spontanée	63
2.2.2	Segmentation et étiquetage intonatif du signal	64
2.2.3	Classification des gestes	66
2.3	Traitement et annotation des constructions subordonnées	76
2.3.1	Localisation, nombre, et distribution des constructions subordonnées dans le corpus	76
2.3.2	Annotation des constructions subordonnées dans le corpus	79
2.4	Mise en relation des différentes annotations	81
2.5	Nature des différentes analyses	82
II	Deuxième partie : résultats et analyse	85
3	Frontières verbales	87
3.1	Mesure des frontières verbales	87
3.2	Résultats	89
3.2.1	Relatives déterminatives	89
3.2.2	Circonstanciennes	103
3.2.3	Appositives	117
3.3	Discussion et conclusion	133
4	Frontières prosodiques	139
4.1	Mesure des frontières prosodiques	139
4.2	Résultats	140
4.2.1	Relatives déterminatives	140
4.2.2	Circonstanciennes	148
4.2.3	Appositives	155
4.3	Discussion et conclusion	169
5	Perception des frontières prosodiques	173
5.1	Introduction	173
5.2	Etudes antérieures sur les frontières en prosodie perceptive	174
5.2.1	La notion de frontière : théorique ou perceptuelle?	174
5.2.2	Les indices prosodiques liés à la perception d'une frontière	180
5.2.3	La force des frontières	182
5.3	Corpus et stimuli	183
5.3.1	Préparation des stimuli	183
5.3.2	Méthodes de manipulation des stimuli	186
5.4	Participants et procédure	188

5.4.1	Participants	188
5.4.2	Le support de test	189
5.4.3	Procédure	190
5.4.4	Analyse des données	191
5.5	Résultats	192
5.6	Discussion	197
5.7	Conclusion	204
6	Frontières gestuelles	207
6.1	Mesure des frontières gestuelles	207
6.2	Résultats	208
6.2.1	Relatives déterminatives	208
6.2.2	Circonstanciennes	222
6.2.3	Appositives relatives	239
6.3	Discussion et conclusion	260
7	Perception des frontières gestuelles	267
7.1	Introduction	267
7.2	Etudes antérieures sur les frontières perçues en gestualité	268
7.2.1	La notion de frontière en gestualité	268
7.2.2	Les tests de perception en gestualité	274
7.2.3	Les indices gestuels liés à la perception d'une frontière	276
7.2.4	La force des frontières	279
7.3	Corpus et stimuli	280
7.3.1	Préparation des stimuli	280
7.4	Participants et procédure	284
7.4.1	Participants	284
7.4.2	Le support de test	285
7.4.3	Procédure	285
7.4.4	Analyse des données	286
7.5	Résultats	287
7.6	Discussion	289
7.7	Conclusion	293
	Conclusion	297
	Bibliographie	311
	Index	343

A	Formulaire de consentement éclairé	347
B	Indications des tests de perception	351
C	Liste des subordonnées sélectionnées	355

Table des figures

1.1	Mouvements de tête et haussement de sourcils de Zoe pendant la réalisation de l'exemple (4).	34
1.2	Mouvement intonatif emphatique (montant-descendant) sur la syllabe accentuée de l'item lexical "hyper".	34
2.1	Schémas représentant l'emplacement des locuteurs et de la caméra.	62
2.2	Elena réalisant un geste de négation (shake) en corrélation avec le groupe intonatif "no it's been fine".	67
2.3	Sourcils haussés de Zoe pendant la production d'un <i>backchannel</i> (réaction) au discours de Michelle.	67
2.4	Geste emblématique de NÉGATION : index levé, le poignet oscillant de gauche à droite.	70
2.5	Geste iconique affilié à un élément lexical spécifique.	70
2.6	Deux gestes iconiques mimétiques.	71
2.7	Pointage référentiel abstrait.	71
2.8	Pointage concret de Beth.	72
2.9	Deux gestes métaphoriques apportant respectivement de l'information modale et interactionnelle.	73
2.10	Battement de la main associé à une récupération lexicale dans le discours verbal de Rhianna.	74
2.11	Moment sélectionné lors d'un adaptateur d'auto-contact produit par Tom.	74
2.12	Capture d'écran du logiciel ELAN.	75
3.1	Continuum de dépendance syntactico-sémantique des constructions subordonnées.	89
3.2	Allongement syllabique final sur SC.	96
3.3	Mouvement prosodique emphatique réalisé par Michelle.	120
3.4	Deux réalisations prosodiques différentes du marqueur de discours "and".	125

3.5	Contour final descendant sur l'appositive d'Alex.	128
4.1	Contour intonatif de l'exemple (3).	143
4.2	Unité prosodique commune pour L et SC.	144
4.3	Allongement syllabique final sur SC.	145
4.4	Deux contours de continuation dans l'exemple (6).	146
4.5	Regroupement de L et de SC sous un même groupe intonatif, et contour continuatif en fin de SC.	148
4.6	Contour intonatif de l'exemple (9).	149
4.7	Mouvement mélodique sous forme d'abaissement sur SC.	150
4.8	Technique de <i>rush-through</i> entre L et SC.	152
4.9	Large rupture prosodique entre SC et R.	153
4.10	Circonstancielle temporelle présentant plus de modulation que son co-texte.	154
4.11	Deux contours similaires (L et SC) où SC montre un registre intonatif plus bas.	159
4.12	Ton descendant-montant sur SC.	161
4.13	<i>Backchannel</i> de Tim produit entre SC et R.	161
4.14	Abaissement de hauteur F0 entre L et SC.	162
4.15	Intervalles non-neutres entre les tons frontière.	165
4.16	Co-occurrence d'une pause silencieuse inter-consituant et d'un raccourcissement de la durée sur SC.	166
4.17	Allongement syllabique final très significatif de L.	168
5.1	Interface de la troisième page du test de perception réalisé sous eSurv.	191
5.2	Interface de la troisième page du test de perception, montrant des stimuli délexicalisés.	192
6.1	Deux battements successifs réalisés en co-occurrence avec la relative déterminative.	208
6.2	Battement manuel sur la relative déterminative et inclinaisons de tête.	210
6.3	Série de gestes manuels répétés faisant office de <i>catchment</i>	214
6.4	Deux pointages référentiels montrant des trajectoires distinctes.	217
6.5	Différents gestes manuels au cours de la séquence (5), où deux configurations sont récurrentes.	221
6.6	Geste métaphorique produit en co-occurrence avec une circonstancielle conditionnelle.	223
6.7	Pointage réalisé en co-occurrence avec la circonstancielle temporelle.	225

6.8	Geste métaphorique réalisé en corrélation avec la circonstancielle, présentant une trajectoire opposée à celle des autres gestes manuels.	228
6.9	Rehaussement initial.	229
6.10	Pointage en corrélation avec la circonstancielle, présentant une trajectoire opposée à celle du pointage précédent.	230
6.11	Série de pointages dans la séquence (10).	232
6.12	Deux gestes manuels aux configurations distinctes, malgré la reprise anaphorique d'un même référent.	234
6.13	Battement de tête produit en co-occurrence avec une circonstancielle.	236
6.14	Geste métaphorique produit en chevauchement sur les groupes intonatifs de L, SC, et R.	238
6.15	Iconique produit sur L suivi d'un métaphorique sur SC.	240
6.16	Geste déictique en chevauchement sur les groupes intonatifs de L et de SC, suivi d'un second déictique.	242
6.17	Métaphorique contrastant avec une série de gestes manuels récurrents.	244
6.18	Changements de direction du regard de Joey.	246
6.19	Mouvements de tête et haussement de sourcils dans la séquence (18).	250
6.20	Haussement de sourcils et différents pointages manuels.	251
6.21	Changement de main se produisant entre SC et R.	253
6.22	Deux gestes métaphoriques localisés sur SC.	256
6.23	Changements de direction du regard au long de la séquence (22).	258
6.24	Changement de main entre le geste localisé sur SC et le geste suivant.	260
7.1	Croix de fixation apparaissant à l'écran.	281
7.2	Position de repos des personnes gestualisantes.	282
7.3	Préparation, réalisation, et rétraction du geste iconique.	283
7.4	Interface de la troisième page du test de perception réalisé sous eSurv.	286
7.5	Placement des trois types de subordinées sur un continuum d'autonomie.	299
7.6	Identité des cadres interprétatifs entre la principale et la subordonnée.	303
7.7	la subordonnée élargit le cadre interprétatif par rapport à celui de la proposition principale.	304
7.8	La subordonnée ouvre un nouveau cadre interprétatif par rapport à celui de la proposition principale.	305
7.9	Graphe représentant les unités prosodiques, gestuelles et discursives pour les propositions relatives.	307
7.10	Graphe représentant les unités prosodiques, gestuelles et discursives pour les circonstancielles.	308

7.11 Graphe représentant les unités prosodiques, gestuelles et discursives pour les appositives.	308
---	-----

Liste des tableaux

2.1	Récapitulatif des types et valeurs/fonctions principales des gestes annotés.	69
2.2	Distribution (%) des constructions subordonnées tous types confondus par locuteur dans le corpus.	76
2.3	Distribution (%) des constructions subordonnées tous types confondus par interaction dans le corpus.	77
2.4	Distribution (%) des circonstanciées par locuteur, interaction, et dans le corpus.	77
2.5	Distribution (%) des relatives déterminatives par locuteur et interaction.	78
2.6	Distribution (%) des appositives par locuteur et interaction.	78
2.7	Récapitulatif du nombre de constructions subordonnées par locuteur.	79
2.8	Récapitulatif des entités syntaxiques couvertes par les étiquettes.	81
4.1	Distribution des pauses silencieuses intra-constituant dans les relatives déterminatives.	141
4.2	Hauteur moyenne des tons de frontière pour les appositives.	164
5.1	Tableau récapitulatif des différents stimuli composant le test de perception des frontières.	185
5.2	Interactions significatives entre les types syntaxiques et les scores de frontière dans la parole filtrée et la parole non-filtrée	193
5.3	Interactions significatives entre les indices prosodiques en isolation et les scores de frontière.	194
5.4	Corrélations significatives entre les indices prosodiques en isolation dans la perception des frontières.	196
6.1	Distribution des changements de regard dans les séquences comprenant une appositive.	248
7.1	Interactions significatives entre les types syntaxiques et les scores de frontière.	287

7.2 Interactions significatives entre les conditions gestuelles et les scores
de frontière. 288

Conventions de transcription

Une ligne de transcription correspond à un groupe intonatif.

(h)	reprise de souffle audible
#	pause silencieuse
(...)	activité vocale (rires, déglutition, soupirs)
[...]	activité gestuelle illustrée
-	construction interrompue
L	co-texte gauche
SC	construction subordonnée
R	co-texte droit

Introduction

S'appuyant sur un corpus vidéo de conversation spontanée en anglais britannique, cette thèse propose une étude de la subordination dans un cadre d'Analyse de Discours Multimodale. Notre travail s'attache à décrire l'intégration de plusieurs types syntaxiques de constructions subordonnées à leur environnement discursif, ainsi que celle des différents signaux verbaux, vocaux, et gestuels déployés lors des processus de production et de perception de ces constructions.

Notre analyse porte sur le regroupement de constructions verbales, et spécifiquement sur la façon dont des constructions verbales peuvent s'organiser pour constituer des entités plus grandes, articulant une séquence "principale" à une séquence décrite par la littérature comme "dépendante". Ce travail s'intéresse aux séquences discursives comprenant des constructions subordonnées opérant au niveau de la modification syntaxique (e.g. Huddleston et Pullum, 2002).

Tels qu'ils sont traditionnellement définis en syntaxe et en Analyse du Discours, les modifieurs font référence à des éléments dont le contenu plus ou moins référentiel est décrit comme modifiant ou spécifiant un élément d'une construction primaire adjacente (Matthews, 1981; Halliday, 1985; Delaveau, 2001). Les modifieurs sont souvent décrits comme des additions associées à un autre contenu propositionnel par un lien plus ou moins lâche dans la structure hôte ou enchâssante (Quirk et al., 1985, p. 1058; Huddleston et Pullum, 2002, p. 1048). Deux types sémantiques sont distingués dans la description des relations de dépendance (van Rijn, 2017). Alors que certaines têtes syntaxiques requièrent de façon inhérente une référence à un élément dépendant, qui peut être considéré comme l'argument de la tête syntaxique, d'autres dépendants ne sont pas présupposés de façon inhérente par leur tête et sont considérés comme des modifieurs. Ils fournissent une caractérisation sémantique plus poussée du référent ou de l'état de faits exprimée par la tête, ou complètent la tête avec d'autres informations. Les propositions relatives et les circonstancielles font par exemple partie de la classe des modifieurs (van Rijn, 2017).

Au lieu de comparer des propositions subordonnées à des propositions non-subordonnées dans le discours, cette étude a pour but d'identifier des différences entre trois types syntaxiques de constructions subordonnées en termes de démarcation, à travers leur expression multimodale. Nous proposons que les constructions subordonnées, particulièrement la catégorie des modifieurs, ne sont pas aussi unifiées et dépendantes que telles qu'elles sont décrites, et que différents degrés de rupture verbale, prosodique, et gestuelle peuvent être distingués selon leur type syntaxique. Les constructions subordonnées sous étude comprennent les trois types syntaxiques de propositions fonctionnant comme modifieurs les plus répandues dans notre corpus oral de conversation spontanée : les circonstancielles, les relatives déterminatives, et les appositives. Ces trois types syntaxiques sont respectivement illustrés dans les extraits (1), (5), et (6).

Les propositions circonstancielles nuancent le procès ou l'état de faits décrit dans la proposition principale par rapport à des agences telles que le temps, les moyens, la condition, la cause, ou le but, souvent avec un élément spécifiant la nature de leur relation (Langacker, 2008, p. 419-420). Ce travail se focalise sur les circonstancielles introduites par “when” et “if”. Ces deux structures ont des fonctions se rejoignant et se chevauchant dans le discours (Dancygier et Sweetser, 2000)¹, et sont d'ailleurs les plus répandues dans notre corpus. Les résultats et les conclusions données dans cette étude pour les circonstancielles ne concernent que ces propositions. Dans l'exemple (1) ci-dessous, la circonstancielle spécifie les circonstances pour lesquelles la relation prédicative “I tried driving once in her car” est réalisée, en localisant dans le temps la situation exprimée par le verbe et ses compléments. Ses éléments référentiels sont stabilisés dans le sens où leur portée est définie. L correspond au groupe intonatif (ou partie de groupe intonatif) précédant immédiatement la relative déterminative, SC désigne la construction subordonnée (nous utilisons la notation de l'anglais pour

1. Les circonstancielles en “when” et en “if” ont en commun leur usage “prédicatif” (Dancygier et Sweetser, 2000) : elles permettent toutes deux d'élaborer des prédictions d'événements ou d'états de faits, mais seulement par rapport à un événement ou un état de faits non-réalisé. De même, les deux circonstancielles établissent ou modifient des espaces mentaux (Fauconnier, 1984), i.e. elles mettent en place un type de raisonnement qui s'appuie sur la création d'un cadre/espace mental hypothétique, et la prédiction des conséquences de cet espace, basée sur les connaissances du monde, les informations introduites précédemment dans le discours, et les relations typiques de cause à effet. La différence entre les deux circonstancielles se joue sur le plan des prises de position épistémiques qu'elles permettent. Alors que “when” engage le locuteur à la réalité de l'espace mental construit dans la principale (même si cette réalité a déjà eu lieu ou aura lieu dans le futur), “if” ne l'engage pas. Alors que le locuteur identifie des croyances avec la conclusion de la proposition principale avec “when” (exprimant un pôle épistémique positif), “if” exprime le désengagement du locuteur sur les événements/états de faits qu'il décrit. Le locuteur n'exprime pas une position vers le pôle négatif de la réalisation de la relation prédicative, mais maintient une position plutôt “neutre”. Le locuteur peut émettre des hypothèses avec un énoncé en “if”, mais ne présuppose pas la validation ou la non-validation de la relation prédicative exprimée.

subordinate construction), et R caractérise le groupe intonatif lui faisant suite. Les conventions de transcription sont disponibles p. xvii.

- (1) Rhianna L i tried driving once in her car
 SC **when we were on a # little road in the countryside #**
 R and hem (swallows) she said well turn left #

Dans l'exemple (2), Rhianna décrit les conditions de sécurité stipulées par la loi européenne à l'égard des provisions de carburant pour les avions. Elle explique que tous les avions doivent décoller avec suffisamment de carburant afin d'être en mesure de changer de destination en cours de vol, et donne un exemple. La proposition de R n'est valable que pour la situation décrite dans la circonstancielle.

- (2) Rhianna L so that for example
 SC **if there's a terrorist attack**
 R (h) hem # there are loads of solutions

Les circonstanciels sont considérées comme extérieures au cadre construit par les propositions principales. La relation sémantique entre les circonstanciels et la prédication qu'elles modifient est sujette à débat (Muller, 2008), comme les deux autres types syntaxiques de notre étude. Selon Gosselin (1990), les circonstanciels sont liées à la proposition qu'elles modifient par un connecteur qui indique leur statut adverbial. Les connecteurs adverbiaux ne lient pas les portions de discours par la rection ou l'enchâssement, ni par séquentialité linéaire pour Blüdhorn (2008). Les connecteurs lient plutôt les portions de discours par la référence. Des connexions adverbiales anaphoriques (orientées vers le discours précédent) et cataphoriques (orientées vers le discours à suivre) peuvent être distinguées, selon l'endroit où l'information nécessaire est placée. En anglais conversationnel, les circonstanciels temporels ont tendance à suivre la proposition qu'elles modifient (Miller et Weiner, 1998), comme il est visible dans l'exemple (1), caractérisé par une connexion adverbiale anaphorique, alors que les circonstanciels conditionnelles ont tendance à précéder les propositions auxquelles elles sont liées (Dancygier et Sweetser, 2000), comme dans l'exemple (2).

La fonction d'une circonstancielle est de "signaler que plusieurs propositions apparaissant au fil d'un texte² entretiennent la même relation à un certain critère, et peuvent donc être groupées dans des unités" appelées des cadres (Péry-Woodley, 2000, p. 62; Charolles, 2003). Les circonstanciels ont un effet sur les processus

2. Un texte est ici considéré comme une entité sémiotique structurée de façon cohérente.

cognitifs du co-participant. Dans l'exemple (3) ci-dessous, la circonstancielle est en position initiale. "I passed" en tant qu'événement fonctionne comme repère cognitif, en cadrant l'état cognitif temporaire exprimé dans "I didn't think I was a very good driver".

- | | | | |
|-----|-----|----|------------------------------------|
| (3) | Tim | L | although |
| | | SC | when i passed |
| | | R | i didn't think i was a good driver |

Les circonstancielles ont été décrites comme ayant un fonctionnement macro-syntaxique soit de préfixe à la proposition qu'elles modifient, soit de suffixe. Notre travail comporte un nombre égal de circonstancielles initiales et finales. La portée interprétationnelle des circonstancielles initiales et finales a été étudiée par de nombreuses études dans la littérature d'Analyse du Discours³. Les circonstancielles initiales sont utilisées par les locuteurs afin d'éviter d'asserter des informations considérées comme déjà connues (i.e. comme faisant partie des connaissances communes; Muller, 2008). Dans l'exemple (3), Tim ne déclare pas directement qu'il a obtenu son permis de conduire, mais utilise cette information en tant que partie des connaissances communes entre les participants pour établir un contraste avec la proposition suivante. Les circonstancielles finales suggèrent l'exhaustivité devant une éventuelle question du co-locuteur à propos de la proposition précédente (Muller, 2008). Il arrive que certains espaces mentaux soient construits d'emblée à l'aide du contexte, et qu'une proposition structurant l'espace mental explicitement serve, du point de vue du locuteur, à clarifier ou confirmer un espace mental déjà construit (Dancygier et Sweetser, 2000). Dans l'exemple (1) plus haut, la circonstancielle répond à une éventuelle question de la co-locutrice à propos des détails dans lesquels "I tried driving once in her car" s'est déroulé.

La grammaire traditionnelle ne précise pas en grands détails le lien syntaxique entre les circonstancielles et la proposition qu'elles modifient (Auer, 2005), et se trouve bien souvent à définir les circonstancielles uniquement par leur connecteur dans une démarche analytique circulaire, ou par justifier des théories à l'aide

3. Chafe (1976) déclare d'abord qu'elles établissent un cadre spatial, temporel, ou individuel dans lequel la prédication principale est valable. Thompson et Longacre (1985) argumentent en faveur de leur contribution pertinente à la structuration des paragraphes discursifs. Les circonstanciennes initiales sont des "signaux grammaticaux" qui indiquent l'ouverture d'une nouvelle unité de discours, pour laquelle elles définissent le cadre (Bestgen, 2009). Thompson (2002) ajoute que les circonstanciennes initiales soulèvent un "problème" à propos des attentes satisfaites par les segments de discours précédents, et que les prochains segments discursifs vont y apporter des solutions. Les circonstanciennes finales jouent un rôle plus restreint et local, en explicitant la scène spatio-temporelle dans laquelle l'action/l'état de faits décrit dans la proposition précédente est menée à bien (Muller, 2006).

d'exemples inventés. Comme les deux autres types syntaxiques de constructions sous étude, les circonstancielle se voient attribuer un statut différent des propositions complétives, qui fonctionnent soit comme sujet, soit comme complément d'un verbe, i.e. un argument. Les circonstancielle sont décrites comme des "satellites", en tant que dépendances "plus externes" aux prédicats et leurs arguments (Dik, 1989). Blanche-Benveniste (1990, p. 19) explique que certains éléments peuvent s'agglutiner autour du noyau verbal d'une proposition, mais sont seulement "associées" à la construction verbale. Harris (1982) considère les circonstancielle comme des prédicats dans lesquels l'un des agents est le procès/l'état de faits verbal en lui-même. Comme les autres prédicats, ils ne sont pas inclus dans les choix se déroulant à un niveau inférieur, tels que le choix de leurs arguments. Lazard (1994, p. 81) estime que les circonstancielle sont "nécessaires en tant que partie de la cible communicative", mais pas du point de vue de l'agentivité grammaticale. Pour Muller (2006), elles lient des segments de parole sans nécessairement impliquer leur prédicat dans une relation syntaxique de subordination. Du point de vue de la grammaire de l'énonciation, elles opèrent une spécification secondaire sur les prédications. Alors que les notions essentielles (i.e. sujet, verbe, complément) sont liées à la prédication par des propositions "principales", les circonstancielle stabilisent les éléments référentiels, en cadrant leur portée interprétative (Wyld, 2003). Les circonstancielle sont également dites à l'arrière-plan du discours, liées aux prédictions de la proposition qu'elles modifient (Dancygier et Sweetser, 2000).

Alors qu'une proposition circonstancielle modifie une autre proposition, une proposition relative modifie soit une expression nominale, soit une autre proposition. Une relation de co-référence est créée entre le référent nominal et un participant dans le procès décrit dans la proposition relative. Ce participant, ou "pivot", a un rôle sémantique à la fois dans la relative et dans la proposition enchâssante qui contient le référent nominal faisant l'objet d'une modification (Langacker, 2008, p. 424). Selon Muller (2006) les relatives sont à peu près équivalentes aux adjectifs étant donnée leur fonction de compléments adnominaux. Les relatives présentent soit des fonctions de caractérisation exclusives (relatives déterminatives), soit des fonctions de caractérisation non-exclusives (appositives). Ces fonctions de caractérisation sont toutefois opérées par un syntagme verbal, à la différence des adjectifs. Emonds (1979) et Cotte (2008) décrivent les relatives déterminatives comme des structures relatives "attachées" par comparaison aux appositives, qui sont "détachées", marquant iconiquement la distance de l'objet. D'autres études (e.g. Muller, 2006; Krifka, 2007) signalent les limites de cette distinction et déclarent que la différence tient en l'utilisation de la co-référence. Les structures relatives peuvent être utilisées comme des

déterminants complexes comme les adjectifs, si le verbe est réduit à sa fonction spécifiante ou qualifiante (voir les exemples (4) et (5)). Cependant, dans le discours, les priorités communicatives peuvent être inversées en utilisant seulement la fonction co-référentielle d'un antécédent. Le dernier segment de discours est prolongé, mais une telle extension peut posséder sa propre valeur pragmatique (voir les exemples (6), (7), et (9)). Quelques études n'adhèrent pas à l'opposition traditionnelle entre les relatives déterminatives et les appositives. Cependant, elles soutiennent toutes le fait que plusieurs grands types de propositions relatives peuvent être distingués (Kleiber, 1980)⁴.

Dans une construction relative déterminative, une expression nominale spécifie un type basique, et une proposition aide à l'identification d'une instance particulière de ce type. L'antécédent est le membre d'une classe qui peut seulement être identifié par l'information donnée par la modification. Dans l'exemple (4), la relative déterminative augmente la pertinence de "the reasons", en créant une sous-catégorie pour ce référent.

- | | | | |
|-----|-----|----|-------------------------|
| (4) | Tim | L | the reasons |
| | | SC | they gave # |
| | | R | i mean Blair was also # |

Cette étude se concentre sur les relatives déterminatives introduites par "that" et par "Ø" en tant que pronoms relatifs. "That" et "Ø" sont analysés comme fonctionnant principalement à la définition de l'antécédent (Cotte, 2008). À l'inverse de "which", qui lie le référent qu'il modifie à la situation contextuelle (Melis, 2008), "that" et "Ø" ont une orientation objectale (i.e. définition d'une propriété de l'objet) caractérisée par une neutralité énonciative. Le contenu d'une relative déterminative est estimé plus "objectif" que subjectif, preuve de l'effort du locuteur d'éviter toute fonction commentative (ibid.). Cette construction permet aux locuteurs de fournir au co-locuteur plus d'informations complexes à propos de l'antécédent que ces derniers ne pourraient traiter dans des structures non-relatives. L'antécédent ouvre un fichier informationnel à propos du référent qu'il décrit, et "that" ou "Ø" indique

4. Kuroda (1968) analyse la différence en termes de structure profonde, par un contraste entre les déterminants. Le Goffic (1979) distingue cinq types de relatives selon les opérations d'identification qu'elles marquent. Larreya (1979) identifie également cinq types de structures relatives en s'inspirant du schéma d'identification de Culioli (1999). Van den Broeck (1973) remplace la distinction binaire par une échelle de huit degrés, allant d'une "restriction totale" à une "non-restriction totale". Le point du milieu correspond à des occurrences dans lesquelles le co-texte et le contexte n'éclaircissent pas l'ambiguïté. Thompson (2002) casse le cadre grammatical traditionnel, en ne considérant plus les relatives comme des structures subordonnées, mais comme des phénomènes de conjonction.

que le fichier informationnel à propos de ce référent va être complété (Muller, 2006). Dans l'exemple (5), “the little bits” ouvrent un fichier informationnel, complété par “stick out” en tant que trait définitoire introduit par “that”.

- | | | | |
|-----|-------|----|---|
| (5) | Joey | L | you know and the little bits |
| | | SC | that stick out # |
| | Elena | | yeah |
| | Joey | R | you have those
like with the pollen on # |

Bien que certaines soient également introduites par un pronom relatif, les appositives ne sont pas invoquées afin d'isoler un référent nominal, mais pour y apporter un commentaire additionnel (Langacker, 2008, p. 429). Leur portée de modification varie d'un seul référent nominal à un syntagme verbal ou une proposition entière (Longacre, 1996). Dans l'exemple (6), l'appositive évalue qualitativement “black pudding”, qui peut toutefois être identifié indépendamment.

- | | | | |
|-----|-----|----|-------------------------------------|
| (6) | Tom | | but then again |
| | | L | i'll happily eat black pudding |
| | | SC | which i know is disgusting # |
| | | R | blood and guts and # |

Dans l'exemple (6), l'appositive est introduite par le pronom relatif “which”. À la différence de “that”, “which” introduit un commentaire “subjectif” qui porte sur la relation entre deux propositions. Dans l'exemple (6), Tom lie “black pudding” à l'adjectif non-neutre “disgusting”. Ce commentaire subjectif peut ajouter un événement qui participe à la succession des événements dans le discours. Dans l'exemple (7), le procès < do on a computer > est consécutif à < get an assessed piece of work >.

- | | | | |
|-----|-----|----|--------------------------------------|
| (7) | Tim | L | you get an assessed piece of work |
| | | SC | which you do on a computer |
| | | R | using a program called author catway |

En termes de dynamisme communicatif, la construction relative en “which” est une assertion qui continue le récit créé par la première proposition, en décrivant son résultat sur le référent que la relative commente (Muller 2006, p. 331). En établissant une distinction fonctionnelle entre plusieurs types d'appositives, Melis (2008) évoque les appositives commentatives (comme dans l'exemple (6)) et les appositives continuatives (comme dans l'exemple (7)). Dans ce dernier type, le pronom relatif intervient seulement en tant que relateur inter-propositionnel.

Toutefois, toutes les appositives ne sont pas introduites par “which”, ni même par un pronom relatif. Dans la conversation spontanée, les appositives sont aussi des syntagmes juxtaposés à leur antécédent nominal⁵, comme dans l'exemple (8).

- (8) Joey but he said #
 L Keith #
 SC **the guy that lives there** #
 R said he'd help us
 doing the stuff #

Joey insère cette appositive au milieu d'une description de son jardin. Elle précise l'identité de “Keith” par “the guy that lives here” dans un souci d'exactitude référentielle. Les deux groupes nominaux renvoient au même référent, sans que la construction appositive soit indispensable au bon déroulement de la séquence.

Neveu (2000) et Loock et O'Connor (2013) expliquent que l'apposition est contrainte à évoluer dans la zone de localité de son antécédent par des lois d'interprétabilité. Neveu (2000, p. 112) parle du “système appositif” comme l'appariement d'un apport et d'un support, réunis dans un même cadre syntactico-sémantique que définit un mécanisme d'incidence.

Selon Loock et O'Connor (2013), au niveau discursif, la construction appositive implique la juxtaposition de deux unités discursives, U1 et U2, où U2 fournit de l'information supplémentaire à propos d'U1. À travers leur référence et leur liaison à un antécédent particulier, le type d'information ajoutée par U2 indexe une relation prédicative implicite entre U1 et U2, principalement de type U1 < BE > U2. Loock et O'Connor (2013) parlent de U1 en termes d'ancre ou d'antécédent et de U2 en termes de supplément d'après Huddleston et Pullum (2002), ou de “noyau”

5. Cette étude se concentre sur un nombre égal d'appositives relatives (introduites par une conjonction) et d'appositives canoniques (syntagme apposé), toutes deux considérées comme des appositives en raison de leur caractéristiques communes. Les points communs tiennent à la possibilité d'étendre une appositive canonique en une appositive relative, au lien supposément lâche entre l'appositive et la proposition qu'elle modifie (Quirk et al., 1985), et à la nécessité de la coréférence entre le pronom en “wh-” ou l'appositive canonique et le syntagme antécédent (Loock, 2013). Pour que deux unités soient appositives, Biber et al. (1999) remarquent que les références qu'elles véhiculent doivent être identiques, ou l'une des références doit être incluse dans l'autre, ce qui est le cas pour les appositives canoniques et les appositives relatives. Certains chercheurs (Burton Roberts, 1975; McCawley, 1996) proposent que l'appositive canonique est une appositive relative réduite (Halitsky, 1974; Doron, 1994). D'autres chercheurs proposent l'inverse : l'appositive relative pourrait être une version réduite d'un syntagme nominal placé en apposition. Loock (2013) souligne que cette hypothèse se retrouve dans des cadres très éloignés, comme la théorie des opérations énonciatives (Cotte, 1998) et des théories plus formelles (Canac Marquis et Tremblay, 1996; De Vries, 2002, 2006). Radford (1981) les traite toutes les deux comme des compléments de nom.

et de “modifieur” (Lago, 1994). Heringa (2007) décrit également l’appositive comme le postmodifieur non-restrictif de l’ancre. Bien que la combinaison des deux unités crée une relation de prédication, les deux propositions ont chacune une valeur de vérité indépendante.

La littérature s’accorde sur le fait que les appositives montrent plusieurs caractéristiques typiquement associées aux propositions non-subordonnées (Krifka, 2007). Leur capacité à former des actes illocutoires distincts (e.g. Peterson, 2004; Holler, 2005) est l’une de ces propriétés. Quirk et al. (1985) considèrent les appositives comme sémantiquement équivalentes aux coordonnées. Si aucune clarification n’est fournie sur ce que les auteurs entendent par “sémantiquement équivalentes”, cette classification est surtout valable pour les appositives “sententielles”, où l’appositive modifie la proposition entière en tant qu’antécédent. Dans l’exemple (9) ci-dessous, la séquence pourrait être glosée avec “it’s only one shop for the whole thing now, and that’s quite bad”.

- | | | | |
|-----|-----|----|--|
| (9) | Tim | L | it’s only one shop for the whole thing now |
| | | SC | which is quite bad # |
| | Tom | | one shop |
| | Tim | R | yeah # |

Alors que la distinction fonctionnelle entre les appositives continuatives (qui font progresser le discours et qui contribuent à son premier plan) et les appositives commentatives (qui apportent de l’information secondaire au discours) est largement acceptée (Lambrecht, 1996; Holler, 2005; Loock, 2007), une récente étude (Lytvynova et Dao, 2014) la remet en cause, en soutenant qu’aucune de ces deux catégories ne peut être apparentée à des unités de discours autonomes, d’un point de vue grammatical comme pragmatique. Potts (2005) décrit également le contenu des appositives comme non-asserté, qui ne peut pas être directement questionné.

En résumé, notre étude porte sur les circonstancielles, les relatives déterminatives, et les appositives. Ces trois types syntaxiques de constructions subordonnées sont choisis pour leurs similarités formelles et fonctionnelles, ainsi que pour leur distribution dans notre corpus. Du côté formel, les trois types syntaxiques possèdent un syntagme nominal ou verbal pour antécédent, et sont introduites par les pronoms “wh-”, “th-”, “if”, et “Ø”. Du côté fonctionnel, les trois types fonctionnent syntaxiquement en tant que modifieurs, dans le sens où ils provoquent une (ré)orientation de la lecture du discours dans lequel ils sont insérés. Les trois types de subordonnées établissent une (ré)analyse d’un cadre interprétatif, que ce cadre soit référentiel

et/ou spatio-temporel. De plus, elles sont toutes considérées comme syntaxiquement “optionnelles” par la littérature.

Nous avons abordé le fait que les subordonnées sous étude sont différentes des subordonnées complétives, qui sont utilisées en tant que sujet, complément du sujet, ou complément d’objet direct dans la proposition principale. Les subordonnées complétives sont considérées comme essentielles à la structure des propositions principales, à la différence de nos subordonnées. En revanche, nos subordonnées sont également différentes des propositions comparatives et de degré, bien que ces dernières appartiennent également à la catégorie des modifieurs (Biber et al., 1999). Les propositions comparatives et celles de degré sont généralement introduites par les morphèmes “as + as” ou “more/less/-er + than”. Elles peuvent également être contraintes au schéma structurel “so/such” + “that”. La contrainte sémantique est donc observable, pour ces structures, sur des morphèmes produits dans la proposition principale, qui ne sont pas que des conjonctions. Nous pouvons également différencier nos subordonnées des propositions commentatives (i.e. *comment clauses*, e.g. “most of London is accessible with the tube **I think**”). Les subordonnées sous étude sont différentes de ces dernières car elles montrent des fonctions pragmatiques beaucoup plus diverses que de marquer une position modale (épistémique, qualitative, déontique, radicale). Les propositions commentatives sont de plus issues d’un processus de grammaticalisation (Brinton, 2008), soit d’un adverbe à un marqueur pragmatique, soit d’une proposition principale à une commentative. Nous distinguons également les parenthétiques (e.g. “i’d kind of be more interested in uh **well there’s no English equivalent** but uh commissaire de police”; Lelandais et Ferré, 2014) des constructions sous étude, puisque les parenthétiques ne sont pas nécessairement des modifieurs syntaxiques, et peuvent être des propositions principales simplement juxtaposées à d’autres propositions dans le discours. De même, nos subordonnées sont assez proches de ce que Biber et al. (1999) appellent les “propositions sans verbe” (e.g. “**although not a classic** this film is worth watching”) ⁶, à la différence que les “propositions sans verbe” se trouvent très majoritairement produites dans des corpus écrits. Les “propositions sans verbe” sont de plus limitées à des expressions formulaires selon Biber et al. (1999).

Les constructions subordonnées ⁷ sont généralement définies comme dépendantes

6. Les subordonnées sous étude et les “propositions sans verbe” partagent des similarités, puisque ces dernières contiennent une relation prédicative implicite qui les lie avec un élément de la proposition principale sous la forme < élément de la proposition principale BE proposition sans verbe >.

7. Nous ne proposons de définition générativiste de la subordination, car nous ne souscri-

d'une autre prédication, d'un point de vue micro- ou macro-syntaxique (e.g. Tomlin, 1985; Lehmann, 1988). Lehmann (1988, p. 181) propose une définition de la notion de dépendance, liée à celle de l'intégration : "une relation grammaticale R qui connecte les syntagmes X et Y est une relation de dépendance si X occupe une position grammaticale de Y ou vice-versa. Dans une relation de dépendance, Y dépend de X si X détermine la catégorie grammaticale du complexe et par là-même ses relations externes". En résumé, une proposition B est dépendante d'une proposition A si B manque d'autonomie fonctionnelle, et est lié à l'un des composants de A : B est hiérarchiquement déclassé dans son rapport à A.

Dans cette étude portant sur le discours spontané, nous considérons qu'une structure B est dépendante d'une structure A si B est intégrée à A, dans le sens où il n'existe pas de marque de frontière entre A et B. Au contraire, une structure A est considérée autonome si A possède des marques de frontière. Les différentes marques de frontière prises en compte pour déterminer la dépendance/l'autonomie dans notre corpus sont décrites entièrement dans les hypothèses du Chapitre 1 p. 58.

Alors que les subordinées sont vaguement définies comme dépendantes, la littérature montre très peu de consensus pour définir des portées et des frontières claires à ces structures (voir Chapitre 1 p. 47). Haiman et Thompson (1984) et Fox et Thompson (1990) ont figuré parmi les premières études à appeler à l'abandon du label de "subordination", jugé trop équivoque, et à l'identification d'autres paramètres plus précis pour l'analyse de la combinaison de propositions. Le label de subordination nous paraît surtout équivoque en ce qu'il mêle deux modes d'organisation discursive : d'une part le mode linéaire, où se mettent en place des continuités et des ruptures, et d'autre part le mode thématique, où se construisent le premier plan et l'arrière-plan, et se signalent le focus ou la compression.

Dans cette thèse, nous étudions les différents types d'autonomie des constructions sous étude uniquement sur le mode linéaire, et nous distinguons les cas dans lesquels elles sont dépendantes du co-texte, comme le décrit la littérature traditionnelle, et ceux dans lesquels elles se démarquent du co-texte, grâce à une approche interactionnelle et multimodale du discours.

vons pas à cette approche innéiste. Notons simplement que la subordination a été très largement abordée dans cette sphère théorique, depuis l'approche transformationnelle de Chomsky et les règles de récursivité (e.g. 1965). Everett (2005) et Nevins et al. (2009) offrent une discussion des problématiques liées à la subordination et à la récursion. Un travail à partir des verbes enchâssants a été initié par Kratzer (2006). Dans ce champ théorique également, le débat à propos des différences syntaxiques, sémantiques, et pragmatiques entre les propositions "principales" et les subordinées est florissant (e.g. Corum et al., 1973; Aelbrecht et al., 2012).

Nous considérons le langage comme un système organisé essentiellement établi par la pratique en contexte, et décrivant la contribution des éléments langagiers construits de manière collaborative à des significations sociales plus larges. De ce point de vue théorique découle notre vision du discours en tant que production essentiellement co-construite de manière incrémentale dans un espace interactionnel dynamique. Cette perspective correspond à la description que fait la linguistique interactionnelle des multiples stratégies de construction du sens en termes de ressources modulaires et émergentes, élaborées dans la temporalité des séquences interactionnelles et des ajustements entre les participants (Schegloff, 1996; Ford et al., 1996; Thompson et Couper-Kuhlen, 2005; Mondada, 2008).

Les constructions subordonnées témoignent de stratégies interactionnelles particulières; leur analyse s'est pourtant souvent concentrée sur la parole, traitée en priorité sur d'autres modalités communicatives. Si un nombre important de travaux sur ce type de construction s'est attaché à examiner leurs relations co-textuelles de la perspective de la syntaxe (Haiman et Thompson, 1984; De Vries, 2006; Haegeman, 2010) ou de la prosodie (Bolinger, 1984; Couper-Kuhlen, 1986), le développement d'outils et de procédés analytiques (e.g. Sloetjes et Wittenburg, 2008; Benzitoun et al., 2009; Bigi, 2012; Boersma et Weenink, 2013) pouvant être mis en relation facilite la description des constructions subordonnées en tant que phénomènes multimodaux.

Dans la conversation face-à-face, les participants négocient le sens à travers des contributions multimodales, dans lesquelles les ressources linguistiques de la parole communiquent avec des modalités⁸ kinétiques, comme la direction du regard, les mouvements de sourcils et de tête, ainsi que les gestes manuels (bien que la posture et les expressions faciales soient utilisées par les locuteurs d'une façon communicative, elles ne sont pas prises en compte dans ce travail, de façon à ne pas multiplier les paramètres). Ces différentes modalités ne fonctionnent pas indépendamment les unes des autres, bien que l'une d'entre elles puisse être plus proéminente que les autres à un moment donné (Norris, 2004). Dans la parole en elle-même, la modalité vocale ou verbale peut être plus proéminente à un moment donné. Cette analyse vise à établir des liens entre la prosodie et la gestualité dans l'organisation du discours, permettant de voir le processus de production des constructions subordonnées à travers le prisme du système organisé et composite que représente le langage.

8. *Modes* en anglais, i.e. des ensembles de ressources sémiotiques organisées socialement et culturellement.

L'Analyse de Discours Multimodale représente un champ de recherche diversifié et grandissant qui inclut des travaux dans nombre de disciplines comme la sémiotique sociale ou la (psycho ou socio-) linguistique. En linguistique, la plupart des études "multimodales" sont en fait bimodales, puisqu'elles se focalisent sur la relation de phénomènes verbaux à la gestualité. Dans cette discipline, très peu d'études adoptent une perspective réellement multimodale : le travail de Loehr (2004) peut être mentionné, puisqu'il porte sur les relations temporelles entre les unités gestuelles et intonatives dans le discours. Les études de Swerts et Kraemer (2005; 2008) sont également multimodales, en s'intéressant à la prosodie audiovisuelle. Dohen et Lœvenbruck (2009) analysent les indices audiovisuels dans les études de perception à propos de phénomènes discursifs. Ferré (2014) propose une approche multimodale du marquage dans le discours. Cependant, aucune étude n'a été réalisée sur la subordination en Analyse de Discours Multimodale à notre connaissance.

Considérer que le langage intègre la parole et des gestes dans un système organisé permet d'étudier la production des subordonnées en se focalisant sur les frontières (i.e. des marqueurs de démarcation en termes d'indépendance) à plusieurs niveaux. L'hypothèse principale de travail est basée sur la capacité de ces constructions à montrer des formes distinctes d'autonomie en fonction de leur type syntaxique. Différents degrés de frontière sont alors identifiés de la nouvelle perspective apportée par la multimodalité, en fournissant une image nuancée de l'insertion multimodale des constructions subordonnées dans le discours.

L'étude de la subordination en conversation spontanée d'un point de vue multimodal donne de nouvelles perspectives sur la production du discours, particulièrement sur la flexibilité de la planification et de la modélisation du discours. Davantage d'information sur la production discursive en temps réel avec un focus particulier sur les frontières contribue également aux champs d'études du traitement automatique du langage naturel et de l'acquisition du langage (Gilbert et al., 2016).

Le corpus que nous utilisons pour cette étude est une collection vidéo de dialogues semi-spontanés⁹ en anglais britannique, ENVID, comptant 5 paires de locuteurs. Leurs échanges sont libres mais restent tout de même provoqués; ils se déroulent dans des studios d'enregistrement. Les données audio et vidéo recueillies, représentant deux heures et dix minutes d'interaction, ont été rendues exploitables

9. Nos dialogues sont semi-spontanés dans le sens où les participants se trouvent dans des studios d'enregistrement. Le contexte physique des interactions ne correspond donc pas à un milieu écologique (i.e. un milieu où les participants décident naturellement d'interagir, comme un domicile, un café, ou un lieu de travail).

après la segmentation, la transcription, et l'annotation du signal audiovisuel à plusieurs égards (syntaxique, prosodique, et gestuel grâce à différentes pistes d'annotation). La définition d'un ensemble d'unités aussi bien verbales (e.g. les propositions), vocales (les groupes intonatifs), et gestuelles (les unités gestuelles) nous permet notamment d'élargir le faisceau de traits propre à chaque type syntaxique de subordonnée. Nous avons sélectionné **55 occurrences** de chaque type (circonstancielle, relative déterminatives, appositives), considérées comme des échantillons prototypes. Notre étude porte donc sur un total de **165 constructions** sélectionnées.

Le travail que nous présentons est scindé en deux parties. La première partie est consacrée à l'ancrage théorique et à la méthodologie de notre étude. Le Chapitre 1 définit le cadre théorique de ce travail. Il décrit les approches théoriques ayant influencé notre étude, en détaillant notre relation à la linguistique interactionnelle, à l'Analyse de Discours Multimodale, à la grammaire des constructions, et à la grammaire cognitive. Les unités théoriques servant à la segmentation du discours que nous utilisons sont également définies. Notre travail s'appuie sur trois grandes unités théoriques, pouvant se chevaucher lors de la production langagière. La proposition représente le format de segmentation et de traitement du flux verbal, dans une tradition plutôt fonctionnelle. Les groupes intonatifs (i.e. *tone-units*) sont utiles à la représentation et l'analyse des données vocales, et proviennent de la tradition britannique. Nous ajoutons quelques paramètres à cette vision, ce qui permet de nous apporter un point de vue à la fois phonétique, prosodique, et phonologique pour certains phénomènes. Enfin, dans un cadre davantage fonctionnel que formel, nous segmentons les mouvements manuels et faciaux à l'aide de la notion d'unité gestuelle. Nous abordons également dans ce chapitre les études antérieures sur la subordination. Nous expliquons que la littérature syntaxique et discursive portant sur les subordonnées considère ces structures comme dépendantes de la structure hôte, bien que peu de consensus existe entre les chercheurs sur le label même de subordination, et sur le niveau d'intégration de ces constructions. Puisqu'un parallèle est établi entre structure syntaxique et structure discursive, les constructions subordonnées sous étude sont décrites comme intégrées à la structure discursive du co-texte, mais subsidiaires du point de vue de leur fonction. Bien qu'aucun travail dans le cadre de la multimodalité n'ait à notre connaissance été réalisé sur les subordonnées, les chercheurs en prosodie parlent d'emboîtement intonatif de certaines unités par d'autres. En ce qui concerne la gestualité, certaines unités discursives peuvent également être liées par une seule et même unité gestuelle. Certaines configurations gestuelles servent également à signaler la dépendance, par opposition au signal de la présence d'une frontière dans le discours. Le chapitre se termine par la

formulation de nos hypothèses, qui représentent une solide liste d’indices de frontière verbaux, prosodiques, et gestuels pour différencier les types syntaxiques des subordinées.

La description du corpus et l’exposé de notre méthodologie se trouvent dans le Chapitre 2, où nous décrivons les étapes de transcription et d’annotation des données. Nous décrivons le corpus ENVID (Lelandais et Ferré, 2017), une collection de dialogues en anglais britannique. Nous détaillons ensuite la transcription orthographique et la segmentation du signal audio en groupes intonatifs à l’aide du logiciel Praat (Boersma et Weenink, 2013) et la segmentation du signal image en unités gestuelles. Enfin, nous justifions les aspects quantitatifs et qualitatifs de notre approche analytique quant aux données récoltées, en soulignant que ces deux types d’analyse sont tous deux nécessaires et complémentaires dans le traitement de données multimodales venant de la conversation spontanée.

La deuxième partie est consacrée aux résultats et à leur discussion. Le Chapitre 3 décrit la mesure des frontières verbales produites dans nos trois types de constructions subordinées, et propose une analyse des résultats obtenus. Nous nous intéressons dans ce chapitre à la relation des constructions subordinées à leur entourage immédiat d’un point de vue verbal; à savoir si la majorité des structures est concernée par la présence d’une rupture syntaxique et/ou discursive, ou si les subordinées sont préférentiellement intégrées au co-texte gauche et/ou droit. Si les relatives déterminatives sont clairement intégrées à leur co-texte (gauche), les circonstancielles présentent des caractéristiques plus mixtes. Alors que leur macro-syntaxe est flexible, elles combinent certains traits discursifs démarcatifs à d’autres indices plutôt cohésifs. La disjonction ne consiste pas à isoler la circonstancielle en tant qu’unité discursive, mais plutôt à signaler un changement d’ordre interprétatif sur le discours précédent et/ou subséquent. Les appositives, quant à elles, sont clairement autonomisées des points de vue macro-syntaxique et discursif. L’encodage d’un procès/état de fait distinct de ceux du co-texte (i.e. qui n’exprime pas une action/état de fait simultané à ceux du co-texte) dans la subordinée, ainsi que la dotation d’une force illocutoire propre à la subordinée constituent les moyens macro-syntaxiques les plus communs d’autonomisation des constructions subordinées dans notre corpus. En ce qui concerne les indices discursifs, les ressources les plus utilisées par les locuteurs sont l’introduction d’une rupture de thème/participant au sein de la macro-séquence, ainsi que la production de marqueurs de discours de structuration. Du point de vue de la réception, les *backchannels* minimaux sont également des indicateurs très répandus. Pour autant, ces indices se combinent rarement.

Le Chapitre 4 se consacre à la production des frontières vocales. Les subordonnées les plus intégrées à leur entourage sont les relatives déterminatives, ne présentant que deux indices de rupture réguliers, tous deux de nature rythmique. Viennent ensuite les circonstancielles, qui comptent trois marques de frontière. À la différence des relatives déterminatives, la rupture s'établit par le biais de ressources intonatives. Enfin, les appositives sont clairement autonomisées en rassemblant huit indices disruptifs, mêlant paramètres rythmiques et intonatifs. Si les traits rythmiques et les rehaussements intonatifs représentent les indices les plus mobilisés dans le marquage d'une frontière prosodique, les marques les plus combinées à d'autres à l'intérieur des séquences sont les pauses, aussi bien remplies (dans le cas des circonstancielles) que silencieuses (intra-constituant dans le cas des relatives déterminatives, inter-constituant dans le cas des appositives). Ces pauses sont le plus souvent articulées à des marques suprasegmentales de disjonction telles qu'un raccourcissement de la durée des segments ou bien un contour intonatif descendant précédant la subordonnée.

Le Chapitre 4 est mis en regard avec le Chapitre 5, qui décrit les résultats quant à la perception de ces frontières prosodiques. Le test de perception que nous mettons en place a pour but de confirmer (1) que des frontières de nature vocale sont perceptibles par des auditeurs naïfs dans les séquences contenant des subordonnées, (2) que différents degrés de frontière peuvent être perçus, et (3) que la rupture est plus perceptible dans les séquences contenant des appositives relatives. Si les pauses silencieuses sont les plus utilisées en tant que marqueurs de rupture, le test doit également déterminer si elles représentent l'indice le plus fiable dans la perception des frontières prosodiques. Nous faisons d'abord un tour d'horizon des travaux autour de la notion de frontière prosodique, sous son acception théorique et son acception perceptuelle. Nous montrons que les indices acoustiques corrélés à la perception d'une frontière prosodique dans la littérature sont liés à la présence d'une pause silencieuse, à l'allongement syllabique final, à la présence de tons, et à un rehaussement initial de la ligne de déclinaison intonative. Nous expliquons ensuite la construction de notre test, à partir d'un ensemble d'extraits audio contenant une subordonnée de notre corpus ENVID, et nous décrivons la procédure selon laquelle nos 52 stimuli ont été présentés à 30 participants dans un ordre aléatoire via une interface web, ainsi que le choix des participants. Les résultats obtenus permettent de confirmer que (1) la perception des frontières prosodiques varie selon les différents types syntaxiques de subordination ; (2) les différents types syntaxiques ne montrent pas les mêmes ordres et niveaux d'interaction avec les mêmes indices prosodiques

de frontière. Notre test confirme également (3) l'importance de la pause silencieuse en tant qu'indice de frontière, à la fois en isolation et dans les combinaisons.

Le Chapitre 6 est consacré à la production des frontières gestuelles dans les séquences discursives comportant une construction subordonnée. Les relatives déterminatives sont très nettement les plus intégrées à leur entourage. Elles ne présentent qu'une seule marque significative de rupture. Cette marque est de nature rythmique et concerne les gestes manuels. Les circonstancielles comptent quant à elles deux indices de rupture significatifs, accompagnés de trois tendances n'atteignant pas la significativité. Comme les relatives déterminatives, la rupture s'établit par le biais des gestes manuels. Enfin, les appositives sont autonomisées par quatre indices, mêlant des paramètres venant de plusieurs articulateurs (mains, regard, sourcils). La démarcation gestuelle s'exprime majoritairement par des ressources correspondant aux gestes manuels. Cependant, les différences entre les trois types syntaxiques quant à la nature et à la distribution des indices suggèrent qu'aucun indice commun n'est systématiquement utilisé pour la démarcation des subordonnées. La présence des battements (manuels dans les relatives déterminatives, de tête dans les circonstancielles) et des haussements de sourcils souligne en revanche l'importante contribution de l'aspect prosodique de la gestualité. Le rôle du rythme dans la démarcation gestuelle des subordonnées montre d'une part que la gestualité fait écho à la modalité vocale dans la production des subordonnées et d'autre part que la prosodie n'est pas uniquement exprimée par les phénomènes vocaux. Si nous constatons que l'aspect prosodique des indices de démarcation se reflète à la fois dans les phénomènes gestuels et dans les phénomènes vocaux, et que l'identification d'indices gestuels de segmentation du discours sur le plan linéaire est à la fois faisable et nécessaire, nous soulignons que la quantité et la distribution de ces indices gestuels sont moindres par rapport à celles des indices vocaux. En relevant des indices de rupture, nous constatons que les subordonnées sont surtout marquées par une gestualité d'ajustement (métadiscursif, discursif, représentationnel, interactionnel) et de modélisation pragmatique du discours. Nous affirmons donc que les subordonnées introduisent une rupture lorsqu'elles établissent une position assertive différente de celle du groupe intonatif précédent. Pour éviter un décalage entre l'apport du locuteur et les représentations du co-locuteur, la gestualité marque cette rupture, mais est également en mesure d'indiquer la valeur informationnelle de cette rupture dans la séquence discursive en cours.

De la même manière que pour les résultats prosodiques, le Chapitre 7 confronte les résultats du Chapitre 6 aux données perceptives, recueillies grâce à la mise en

place d'un test autour de la rupture dans le discours exprimée par des moyens gestuels. En comprenant des constructions totalement intégrées à leur co-texte du point de vue vocal, ce test déterminera si les indices gestuels sont aussi fiables dans la perception des frontières que les indices vocaux. Notre test de perception porte uniquement sur les gestes manuels, en raison de la nature des stimuli utilisés. Le test de perception que nous mettons en place a pour but de confirmer que (1) des frontières exprimées par des moyens gestuels sont perceptibles par des participants naïfs dans les séquences contenant des subordinées; (2) différents degrés de frontière peuvent être perçus; (3) la rupture est plus perceptible dans les séquences contenant des appositives relatives. Nous faisons d'abord un tour d'horizon des travaux autour de la notion de frontière gestuelle, qu'elle soit liée à la prosodie ou à la syntaxe. Nous montrons que peu d'études perceptives ciblent directement le rôle démarcatif des gestes manuels, mais qu'il a été prouvé en production que deux gestes manuels consécutifs montrant des configurations et des coordonnées totalement différentes peuvent signaler une frontière dans le discours. Nous expliquons ensuite la construction de notre test, réalisé à partir d'extraits audio du corpus ENVID. Nous décrivons ensuite la procédure selon laquelle nos 24 stimuli ont été présentés à 30 participants dans un ordre aléatoire via une interface web, ainsi que le choix des participants. Les résultats obtenus permettent de confirmer que (1) la perception des frontières gestuelles varie selon les différents types syntaxiques de subordination; (2) les différents types syntaxiques ne montrent pas les mêmes niveaux d'interaction avec les indices gestuels contenus dans les différentes conditions expérimentales; (3) la rupture est davantage perceptible dans les séquences contenant des appositives relatives. Toutefois, nous soulignons de manière cruciale, à partir de nos résultats, que si la gestualité est bel et bien impliquée dans la perception des frontières, les frontières exprimées de manière gestuelle sont perçues à un moindre degré par rapport aux frontières exprimées de manière vocale, de la même façon qu'en production, les frontières sont exprimées avec davantage de moyens prosodiques que gestuels.

Une conclusion synthèse p. 297 propose la représentation des trois types syntaxiques de subordinées sur un continuum, allant de la structure la moins autonome (les relatives déterminatives) à la structure la plus autonome (les appositives). Nous revenons également sur le fait que les frontières s'expriment majoritairement avec des moyens davantage prosodiques que gestuels. Une représentation schématique de l'impact des constructions subordinées sur la construction de l'interaction est introduite, en lien avec notre ancrage théorique. Le modèle multiparamétrique de la subordination que nous proposons est basé sur le fait que ces constructions sont des pratiques interactionnelles qui offrent une reconstruction interprétative du discours.

Nous nous attachons ainsi à montrer que les constructions subordonnées témoignent de modes d'organisation distincts en fonction de leur type syntaxique. Nous proposons qu'elles constituent des moyens d'apporter des contributions modales composites, à la fois propositionnelles, textuelles, et interactionnelles.

Première partie

Chapitre 1

Cadre théorique

Notre travail témoigne de plusieurs approches multimodales, complémentaires à l'étude de la conversation spontanée. Nous commençons par identifier nos différents ancrages théoriques pour ensuite définir nos unités de segmentation et d'analyse. Nous terminons ce chapitre par une revue des travaux sur la subordination, qu'elle soit syntaxique, discursive, prosodique, ou gestuelle.

1.1 La linguistique interactionnelle

Notre approche se situe à la fois dans l'Analyse du Discours et dans l'Analyse Conversationnelle en considérant le langage comme un système organisé essentiellement établi par la pratique, et en étudiant comment des éléments langagiers contribuent à des significations sociales plus larges.

Nous savons par Halliday et Hasan (1976) que le sens se construit grâce à l'enchâssement ou l'imbrication des structures selon différentes échelles, et qu'il se construit également grâce aux chaînes de cohésion qui garantissent la liaison et la solidarité entre différentes unités d'un message.

Ce message, dans une perspective qui conçoit le discours comme une forme coopérative d'organisation sociale (Goodwin, 1981; Goodwin et Goodwin, 1992), fait partie d'une activité communicative. Nous considérons d'après Lemke (1991) que les structures d'activités communicatives sont caractéristiques d'une communauté. Ce sont des séquences d'options régulières, répétables et répétées, qui sont grandement déterminées par le contexte de la structure d'activité dans laquelle elles sont réalisées.

Toute communication linguistique s’inscrit dans un contexte au sens large (contexte discursif et situationnel) qu’elle-même aide à créer (Gumperz, 1982; 1992). Cette approche implique un regard particulier sur les ressources langagières, mettant en avant les procédés des locuteurs mis en œuvre pour construire le sens, assurer l’intercompréhension, et ainsi mener à bien l’activité conversationnelle plutôt que la description des formes pour en établir une grammaire.

Les contributions langagières, issues d’un système linguistique intégrant parole et gestualité de manière organisée, sont donc vues comme dynamiques et émergentes, élaborées dans la temporalité des séquences interactionnelles et dans leur adéquation au contexte. Le sens est le produit des contributions langagières des participants dans un contexte particulier, et celui de l’anticipation et de la réflexion des participants sur les “objets de discours” (Mondada, 2011a) qu’ils construisent incrémentalement (Mondada, 2011b). Si un contexte d’entrée est induit par le texte¹ qui précède, les énoncés fournissent également un contexte de sortie, qui peut être exploité par le texte qui suit. Nous envisagerons donc le sens d’un énoncé avec son potentiel de changement du contexte. L’un des objectifs de notre analyse sera de se demander comment la langue se manifeste et se constitue en même temps dans les pratiques. C’est-à-dire comment les pratiques langagières, et notamment les activités interactionnelles des participants, identifient, exploitent, et par là configurent les ressources de ce qui sera désigné comme étant la langue (Mondada, 2008).

1.2 L’Analyse de Discours Multimodale

L’analyse multimodale consiste à mettre en relation des informations linguistiques produites dans différentes modalités (*modes* en anglais, i.e. des ensembles de ressources sémiotiques organisées socialement et culturellement), chacune d’elles contribuant à l’élaboration et à la perception du message communiqué. Ces différentes modalités ne fonctionnent pas indépendamment les unes des autres, bien que l’une d’entre elles puisse être plus préminente que les autres à un moment donné (Norris, 2004).

Les études multimodales s’intéressent à la communication linguistique comme relevant à la fois des modalités verbales, vocales, et gestuelles, tant sur le plan segmental (gestes articulatoires qui produisent des phonèmes) que sur le plan supra-segmental (intonation et gestes co-verbaux), qui agissent au niveau de l’énoncé ou

1. “texte” désigne ici une entité sémiotique cohérente structurée. Un texte peut donc être oral.

du discours, même si les gestes co-verbaux ne sont pas obligatoires dans la communication linguistique (Ferré, 2014).

De nombreux chercheurs postulent un lien étroit entre la parole et certains gestes, dits co-verbaux, qui participent à la communication linguistique. Pour McNeill (1992; 2005) par exemple, la gestualité co-verbale n’apporte pas un complément d’information à la parole, support du message, mais le message linguistique est le résultat d’une imagerie cérébrale complexe qui intègre des informations gestuelles et langagières, ainsi que des informations environnementales. C’est aussi la direction indiquée par une certaine branche des sciences cognitives, travaillant sur la cognition “incarnée” (Iverson et Thelen, 1999; Bernardis et Gentilucci, 2006; Barsalou, 2008; Streeck, 2009) : les activités cognitives (dont la communication) se nourrissent du contexte (l’environnement physique, culturel, l’histoire conversationnelle, les connaissances partagées) et sont étroitement liées à la perception, à l’action, et à l’introspection. Dans cette perspective, les mains ne sont par exemple pas seulement des instruments d’action et d’expression, mais aussi de cognition et d’acquisition de connaissances. La parole et la gestualité ont co-évolué à la fois naturellement et culturellement pour se compléter et se suppléer, se coordonner et s’intégrer mutuellement. De récents travaux confirment cette position (Roustand, 2012; Ferré, 2014; Ruth-Hirrel et Wilcox, 2018), et suggèrent que la communication linguistique s’opère donc par un couplage des modalités vocale et gestuelle.

En ce sens, nous considérons que le visuel, le vocal et le verbal créent des unités de sens composites qui s’organisent mutuellement (Lemke, 2002; 2005). Par une constellation d’éléments de sens reliés entre eux, les locuteurs dépassent les contraintes propres à chaque mode et peuvent combiner ou séquencer les éléments selon des logiques (Martinec et Salway, 2005), des espaces et des temporalités différentes (Debras, 2017). L’interaction de plusieurs ressources sémiotiques à des échelles différentes crée donc des informations de forme complexe, qui peuvent à leur tour être réutilisées (i.e. servir de base) à la création de nouvelles significations (Goodwin, 2012).

1.3 La grammaire des constructions

Dans la grammaire des constructions (e.g. Fillmore, 1985; Kay et Fillmore, 1999) et la grammaire des propriétés (e.g. Blache, 2016), l’unité fondamentale du langage prend la forme de construction. Les constructions sont des associations entre forme et propriétés fonctionnelles sémantiques dont le degré de complexité est variable. Ce

sont des entités théoriques servant à construire les éléments de la langue (e.g. les mots, les syntagmes, ou les propositions). Les morphèmes individuels, les mots de plusieurs morphèmes, les syntagmes, les propositions, et les complexes propositionnels sont tous considérés comme des constructions, d’une complexité variable. Les constructions complexes (i.e. toutes les constructions de plus d’un seul morphème) sont également des associations entre forme et fonction. Toutes les constructions possèdent un ensemble de propriétés et un ensemble de contraintes. L’exemple (1) représente plusieurs constructions² formant une construction complexe (i.e. une expression nominale), pour l’occurrence “the Spanish girls that were there”.

- (1) the Spanish girls that were there
 [CON the] [CON Spanish] [CON girls] [CON that] [CON were] [CON there]
 [CON the Spanish girls] [CON that were there]
 [CON the Spanish girls that were there]

Cette représentation ne comporte pas d’information hiérarchique, mais est obtenue à partir des propriétés syntactico-sémantiques de chaque construction. Les propriétés peuvent être décrites selon plusieurs grandes lignes (Blache, 2016) : l’ordre linéaire entre les catégories formelles de constructions, la co-occurrence obligatoire entre deux catégories, l’exclusion de co-occurrence entre deux catégories, l’impossibilité pour une catégorie de se répéter, et la dépendance syntactico-sémantique entre deux catégories.

Dans la grammaire des constructions, le lexique et la grammaire sont considérés comme faisant partie d’un continuum. Les constructions (qui correspondent à des modules indépendants dans les théories génératives du langage, comme le lexique ou la syntaxe) sont différenciées par le degré de complexité et de schématicité. Cette vue est intéressante en ce qu’elle exprime la nécessité de considérer les modules linguistiques en tant que continua ou en tant que composants hybrides et perméables. Dans ce travail, nous parlons en conséquence d’aspects syntaxiques, discursifs, et pragmatiques dans les constructions, mais pas de catégories exclusives.

La grammaire des constructions vise à une représentation de la langue décentralisée de la grammaire, grâce à un formalisme prenant en compte syntaxe, phonologie, morphologie, sémantique et pragmatique. Avec cette approche, les expressions idiomatiques et les expressions figées ne sont pas à la périphérie de la langue, mais en font partie au même titre que les autres constructions. La gram-

2. Chaque construction est représentée entre crochets. Les constructions sont réparties sur plusieurs lignes dans un souci de lisibilité.

maire des constructions permet de prendre en compte plusieurs niveaux de structuration linguistique à la fois dans la description des unités et de rendre compte de leur fonctionnement contextuel. C'est également une grammaire non transformationnelle : une forme passive a par exemple le même statut dans la grammaire que la forme active, et n'en est pas dérivée. La grammaire n'est plus vue comme le moyen de générer le langage au moyen d'un ensemble d'opérations, mais plutôt comme le moyen de construire toutes sortes de déclarations au moyen d'un ensemble de descriptions (Blache, 2016).

Les constructions peuvent être motivées par d'autres constructions, partiellement dérivées d'autres constructions, ou incorporer d'autres constructions en tant que composants (Lakoff, 1987; Goldberg, 1995). En conséquence, les constructions forment des réseaux structurés. Cet aspect permet de ne pas nécessairement placer la parole au centre du modèle de communication. La phase gestuelle possède également son ensemble de propriétés et de contraintes, et peut être groupée à d'autres phases pour acquérir des caractéristiques sémantiques et constituer un geste³. La grammaire des constructions est alors extrêmement utile à la multimodalité, où la question de la représentation du langage est essentielle. Les modalités ne peuvent en effet être articulées ou conçues comme des réseaux unidirectionnels de dépendance comme des arbres syntaxiques. La représentation en arbre n'est possible que dans chaque modalité, et paraît peu représentative. La grammaire des propriétés permet la modélisation de données complexes de corpus différents, avec par exemple leur mise en regard sur un axe temporel commun (Bird et Liberman, 2001).

Par ailleurs, toutes les constructions à partir du niveau des mots jusqu'au niveau du discours témoignent d'un certain degré de conventionnalité, car les participants sont capables de se comprendre. La créativité est expliquée par une diminution de l'emploi d'arrangements d'unités linguistiques conventionnelles à des fins innovantes. Cette conception de la créativité dans les constructions de la parole a des implications pouvant améliorer la compréhension de la créativité dans la gestualité. Hirrel (2018) souligne que les gestes co-verbaux avec des propriétés témoignant de créativité ou imprévisibles ont souvent été traités comme pleinement idiosyncratiques et *ad hoc* (McNeill, 1992). Ces descriptions ont récemment été remises en cause et les études reconnaissent que même les gestes créatifs possèdent des propriétés conventionnelles (Cienki et Mittelberg, 2013; Kok et Cienki, 2016). Si l'on

3. La réalisation d'un geste manuel implique différentes phases : la phase de réalisation du geste est parfois précédée d'une phase de préparation, et peut être suivie par la tenue du geste (où l'articulateur reste statique dans la configuration obtenue après la réalisation du geste), et par une phase de rétraction.

envisage que les gestes peuvent constituer des expressions complexes, une extension des théories de la grammaire des constructions peut améliorer la description des composants conventionnels et créatifs des gestes.

Hirrel (2018) souligne que la multimodalité en grammaire des constructions commence seulement à être explorée. Elle cite la ressource du *Distributed Little Red Hen Lab* (DLRHL), qui inclut une collection de plus de 250 000 heures de vidéos recherchables et étiquetées. Le DLRHL est l'une des premières ressources permettant aux chercheurs d'appliquer des méthodologies de linguistique de corpus à l'étude des expressions multimodales (Turner et Steen, 2013). La gestualité a fait l'objet de plusieurs travaux théoriques en grammaire des constructions (e.g. Zima, 2014; Brône et Sambre, 2015). Turner et Steen (2013) appréhendent par exemple les significations linguistiques comme des mélanges intermodaux qui synthétisent rapidement certains traits choisis de l'information en de nouvelles entités.

Clark (1996) et Enfield (2009; 2017) ont également étudié la nature des expressions composites mêlant gestualité et parole. Clark (1996) développe la notion de signaux composites, qui sont des actions significatives combinant plusieurs signaux, différents en nature. Enfield (2009) suit une approche similaire avec le concept d'énoncés composites. Les énoncés composites comprennent des signes linguistiques conventionnels et des unités gestuelles, ainsi que des moyens conventionnels offerts par l'environnement physique dans lequel se situe l'interaction. Clark et Enfield conçoivent tous les deux les expressions composites en tant que points de départ à l'analyse plutôt que d'appréhender la signification d'un point de vue des composants en blocs de constructions. Cette vue est analogue aux conceptions de la grammaire des constructions. Les constructions sont les unités basiques du langage à partir desquelles les catégories de signification sont dérivées.

Néanmoins, Hirrel (2018) explique également que les efforts pour intégrer la gestualité aux théories de la grammaire des constructions ont été ralentis par quelques problématiques. Schoonjans (2017) aborde quelques-uns de ces problèmes, liés au fait qu'il est difficile d'intégrer la gestualité dans les théories linguistiques en raison des présupposés à propos de la signification des gestes. Schoonjans (2017) note qu'un haut degré de variation existe dans l'expression formelle des gestes récurrents. Les gestes ne sont jamais produits à l'identique à travers plusieurs occurrences, même ceux qui partagent un cœur formel (e.g. les gestes cycliques). En revanche, la variation dans la production d'une expression n'est pas un phénomène spécifique à la gestualité. Les propriétés phonétiques et phonologiques de la parole sont également

variables, même à travers les occurrences d'un seul locuteur. Un autre problème identifié par Schoonjans (2017) concerne la difficulté d'établir si un schéma multimodal est assez fréquent pour être considéré comme une construction. Cette problématique est liée à la nécessité, dans les grammaires des constructions, qu'un schéma soit conventionnalisé afin de pouvoir être considéré comme une construction (Croft, 2001; Goldberg, 2006). La conventionnalisation est une question de degré, et la fréquence n'est pas une condition suffisante pour établir la conventionnalité (Bybee, 2008). Schoonjans (2017) note que le problème de la fréquence n'est pas spécifique à la gestualité mais est un problème général dans l'approche de la grammaire des constructions.

Si la grammaire des constructions est extrêmement utile à la représentation de données multimodales, elle présente néanmoins également des limites pour l'instant sur les données de la conversation spontanée. Si cette grammaire permet la prise en compte d'un contexte, ce contexte est encore trop peu élargi. La définition des constructions, ainsi que le fait de représenter et de tracer leurs mouvements, ne permettent pas la description en détail de certains phénomènes langagiers du spontané tels que les troncations, les ellipses, les répétitions, et les abandons.

1.4 La grammaire cognitive

Notre approche s'inspire également de la grammaire cognitive en envisageant la grammaire en tant que partie intégrante de la cognition. Plus largement, les structures linguistiques mobilisent et révèlent une multitude de phénomènes et processus cognitifs (comme l'attention, la perception, et la mémoire) dont elles ne peuvent être dissociées (Langacker, 2001). La grammaire cognitive invoque fréquemment des principes de la psychologie de la forme (Gestalt) et établit des analogies entre la structure linguistique et certains aspects de la perception visuelle.

La grammaire cognitive présuppose également que le langage s'appuie sur des capacités humaines plus spécifiques pour l'abstraction et la catégorisation. Grâce à une approche fonctionnelle, elle a démontré que le rôle de capacités conceptuelles telles que la métaphore, l'intégration, ou la construction d'espaces mentaux (e.g. Lakoff, 1980; Fauconnier, 1984; Fauconnier et Turner, 1996) est au centre du langage. L'exemple (2)⁴ présentant deux appositives (la première sur L, l'autre sur SC) montre que les formes grammaticales sont liées à des enjeux conceptuels pour les

4. Les crochets représentent les chevauchements de parole.

locuteurs.

- | | | | |
|-----|-----|----|--------------------------------------|
| (2) | Tom | | who do- who played stuff |
| | Tim | | hem well just the regulars |
| | | L | the locals # |
| | | SC | the band called the misfits # |
| | | | [and |
| | Tom | | what kind] of stuff # |
| | | | what kind of stuff are they playing |
| | Tim | | mostly old stuff but # |

Tim répond à une question de Tom à propos du bar dans lequel il est allé la veille, et des musiciens jouant ce soir-là. SC constitue ici une réponse à la question de Tom au même titre que L et le segment précédant L. L apporte une équivalence sémantique au segment le précédant : “who played stuff? The locals”. L est plus précis que le segment précédent en permettant de répondre en même temps à une potentielle sous-question à partir de ce segment : “who are the regulars?”. Dans un souci d’exhaustivité, Tim fournit ici à Tom un élément de réponse supplémentaire sur SC : “the band called the misfits”. Ces éléments de réponse sont sémantiquement équivalents, et tout à fait interchangeables : “hem well just the band called the misfits, the locals, the regulars”. L’élément de L n’a en revanche pas la même fonction pragmatique que SC. Alors que l’élément de L effectue un renvoi sur “the regulars” en explicitant ses propriétés, le contenu de SC crée un emboîtement de point de vue à la manière d’un *zoom* (Langacker, 2001). “The band called the misfits” décrit une entité sémantique plus petite que celle de L : ce groupe est inscrit à l’intérieur de celui dénoté par L. Son inversion avec L n’a alors pas d’incidence sur les instructions d’interprétation données au co-participant sur l’ensemble de l’extrait, qui prend des apparences de repère constitutif en créant un schéma d’accès cognitif à l’information (i.e. regulars > locals > band called the misfits).

La grammaire cognitive est essentielle à la description des unités de la parole spontanée, et propose une vision permettant à la fois l’analyse de constructions typiques du spontané (truncations, répétitions, mouvements de constituants inter- et intra-propositionnels) et l’analyse de phénomènes à travers plusieurs prismes simultanément (e.g. structure attentionnelle, gestion de la structure informationnelle, charge cognitive).

Si les théories génératives du langage postulent des frontières rigoureuses entre le lexique, la grammaire et la morphologie, la grammaire cognitive envisage plutôt une gradation, avec des niveaux intermédiaires. De la même manière, la grammaire cognitive utilise préférentiellement des conceptions continues ou prototypiques plutôt

que la décomposition en unités discrètes ou la hiérarchisation uniforme. Elle vise à unifier la prise en compte d’aspects du langage habituellement traités séparément par d’autres théories, notamment en recourant à un niveau de schématicité ou d’abstraction élevé. L’exemple (3) montre que l’emploi d’un format grammatical a des répercussions lexicales et pragmatiques.

- | | |
|----------|--|
| (3) Joey | but the thing is |
| | <u>the guy</u> packed the bag |
| L | and hem # he |
| SC | like the Sainsbury’s man (laughs) |
| R | and hem # the pizza bases have got a crack in them |

SC modifie “the guy”, codé en surface par un fléchage puis par un pronom personnel, correspondant traditionnellement aux entités identifiables parmi d’autres du même type, alors qu’il est totalement nouveau dans le discours et dans la mémoire à court terme de la co-locutrice. L’appositive modifie donc cet élément thématique par un autre fléchage précisant le type parmi lequel il est singularisé (i.e. les employés de la grande surface Sainsbury’s). SC réalise alors un processus d’ancrage informationnel, de “stabilisation référentielle” pour l’élément thématique de la séquence nouvellement introduit sur L. Au long de cette séquence, les objets sont repérés par des entités humaines (représentant un degré plus élevé sur une échelle de reconnaissance grâce au phénomène d’empathie; Fox et Thompson, 1990; Langacker, 2008), et les entités humaines sont amenées dans la sphère des participantes, avec l’allusion à une chaîne de supermarchés faisant partie des connaissances partagées. Cette appositive est déclenchée par l’incapacité du locuteur à définir le statut informationnel de l’antécédent par rapport au co-locuteur. On observe d’ailleurs ici un décalage entre le statut informationnel et cognitif de l’item servant d’antécédent à cette apposition et sa forme lexicale de surface. Il arrive en effet fréquemment que les locuteurs encodent de l’information totalement nouvelle qui mobilise des concepts non-activés et non-accessibles pour le co-participant sous une forme lexicale réduite suggérant au contraire une familiarité du co-locuteur à l’information, ainsi qu’un concept déjà activé dans la mémoire à court terme du co-participant ou facilement accessible. L’appositive mobilise alors une forme lexicale de surface dont le rôle est de faciliter l’accès au référent présenté sur L.

Par ailleurs, la grammaire est considérée comme porteuse de sens en elle-même, et non comme un système formel autonome. Pour Langacker (2008), les représentations du langage des locuteurs incluent trois types de structures se manifestant dans les expressions linguistiques : les structures sémantiques, phonologiques, et symboliques. Les structures sémantiques sont des conceptualisations mobilisées pour la significa-

tion dans le langage. Les structures phonologiques sont les éléments perceptibles du langage. Elles incluent les sons, les gestes, et les représentations orthographiques. Les structures symboliques résultent de l'association entre les structures sémantiques et phonologiques. Les structures symboliques associent donc forme et signification. Les structures sémantiques appartiennent au pôle sémantique et les structures phonologiques sont associées au pôle phonologique. Le pôle sémantique inclut non seulement le contenu explicite d'une expression, mais aussi l'intégralité de la signification contextuelle, y compris les inférences. Les structures symboliques sont quant à elles décrites comme bipolaires.

Selon Langacker (2008, p. 61), la plupart des expressions linguistiques employées sont symboliquement complexes, dans le sens où ce sont des assemblages d'autres éléments symboliques plus petits. Puisque les expressions complexes sont des combinaisons de structures symboliques, elles sont équivalentes aux constructions de la grammaire des constructions. Les assemblages symboliques bénéficient d'un degré de conventionnalité varié. Alors que certains composants individuels de la structure peuvent être très conventionnels, l'intégration de ces composants en tant qu'assemblage symbolique peut être plus créatif.

Dans le cadre de la grammaire cognitive, la gestualité est associée avec le canal (articulatoire) de la vocalisation. En d'autres termes, les gestes sont des éléments perceptibles qui font partie du pôle phonologique. La grammaire cognitive présente toutefois quelques limites relatives à la multimodalité. Si la théorie intègre des données à la fois verbales, vocales, et gestuelles, elle fournit des catégorisations extrêmement générales pour le vocal et le gestuel. Hirrel (2018) propose que la gestualité n'est pas associée à un, mais à au moins trois canaux de vocalisation. Cette proposition inclut des canaux séparés pour les gestes manuels, les gestes faciaux, et le positionnement du corps. Hirrel (2018) propose également que les composants des structures symboliques vocalisés par les gestes forment des constructions gestuelles plus complexes (i.e. des assemblages symboliques gestuels).

Les constructions gestuelles sont supposées schématiques sur le pôle sémantique car elles nécessitent typiquement des structures symboliques de la parole pour leur élaboration. Les composants symboliques et les assemblages symboliques de la gestualité s'intègrent aux assemblages symboliques de la parole pour former des constructions multimodales. Les principes de la grammaire cognitive sont donc utiles pour rendre compte de la façon dont les composants symboliques de la gestualité sont intégrés aux constructions gestuelles et de la façon dont ces constructions sont

intégrées aux constructions de la parole.

L'exemple (4) associé à la figure 1.1⁵, où les images (a) et (b) correspondent à différents moments de sa production, et à la figure 1.2 permet d'illustrer le schéma proposé par Langacker (2008).

- | | | |
|-----|----------|--|
| (4) | Michelle | i just hope we don't get stuck this time |
| | Zoe | oh god that would be awful |
| | L | i'd [(a) panic |
| | SC | if] we got stuck under the tunnel |
| | R | [(b) i'm hyper claustro][(c) phobic #] |

Michelle et Zoe vont prochainement prendre l'Eurostar ensemble. Michelle, qui a déjà été bloquée dans le tunnel à cause d'une panne, vient de mentionner cet événement et espère qu'elle ne sera pas de nouveau bloquée. Zoe se projette quant à elle dans cette situation. Zoe exprime sa surprise, et prend position de façon négative sur cette éventualité, reléguée au domaine de l'irrealis avec la circonstancielle conditionnelle. Deux occurrences de "would" expriment le rejet de la réalisation de < get stuck > dans le domaine de la realis, et l'adjectif "awful" exprime la prise de position qualitative de Zoe. En plus des éléments lexicaux, les structures grammaticales employées par Zoe participent à la construction du sens. Ils appartiennent en ce sens au pôle sémantique. En revanche, ces éléments lexicaux et ces structures grammaticales sont verbalisées et gestualisées. Le pôle phonologique est également mobilisé. Toutefois, l'emploi d'un pic intonatif emphatique sur l'adverbe "hyper" (visible dans la figure 1.2 ci-dessous) ne mobilise pas que le pôle phonologique : le pic intonatif et l'adverbe participent à la structure de l'information, et font appel à une gradation. À travers cette séquence, Zoe n'utilise pas de gestes manuels participant à la construction du sens. En revanche, elle produit deux mouvements de tête et un haussement de sourcils. Zoe secoue la tête de gauche à droite alors qu'elle produit "panic" (a). Ce geste de négation intensifie sa prise de position modale (McClave, 2000) : Zoe rejete toutes les autres possibilités de réaction face à l'éventualité de l'arrêt du train sous le tunnel. Si la forme du geste en elle-même fait partie de la structure phonologique, le geste de négation représente également une structure symbolique en participant à la signification de façon très conventionnelle. Il en va de même pour l'inclinaison de tête de Zoe (b), puisqu'elle incarne la consternation que lui provoque ces pensées. Les deux haussements de sourcils (l'un co-occurent à l'inclinaison de tête de Zoe, l'autre visible sur (c)) font un travail semblable à celui du pic intonatif sur "hyper", en mettant en valeur une partie informative de l'énoncé.

5. L'unité gestuelle est représentée par les crochets.

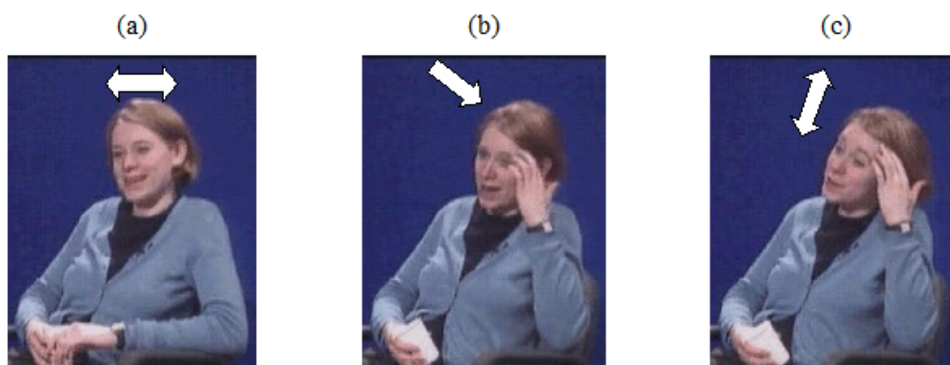


FIGURE 1.1 – Mouvements de tête et haussement de sourcils de Zoe pendant la réalisation de l'exemple (4).

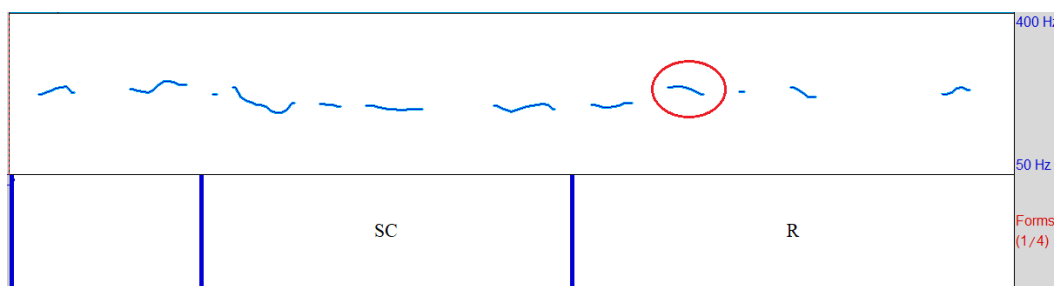


FIGURE 1.2 – Mouvement intonatif emphatique (montant-descendant) sur la syllabe accentuée de l’item lexical “hyper” dans l’exemple (4), illustré par la courbe prosodique dans Praat (la piste de transcription montre les segments - L, SC - R dans la séquence).

En résumé, notre étude est ancrée dans plusieurs approches théoriques, toutes utiles à la description et l’analyse des données multimodales de la conversation spontanée. Nous nous appuyons sur des approches multiples car les richesses qu’offrent ces différentes théories sont orientées vers des données de nature différente. La linguistique interactionnelle et l’Analyse de Discours Multimodale représentent les piliers de notre analyse. La grammaire des constructions présente des avantages et de grandes avancées quant à la représentation et l’appréhension des données multimodales, tandis que la grammaire cognitive permet une description très flexible des composants spécifiques à la conversation spontanée.

1.5 Unités théoriques d’analyse

Nous présentons ici le contexte théorique des unités principales de segmentation du discours utilisées dans ce travail. Bien que la segmentation de nos données relève

davantage de la méthodologie, chacun de nos choix en termes d'unités de langage possède des implications théoriques qu'il convient de mentionner dans ce chapitre.

1.5.1 La proposition

Une unité syntaxique et discursive ?

Les chercheurs dans le domaine de la conversation spontanée ont été nombreux à proposer des hypothèses en concurrence à propos de ce qui constitue une unité élémentaire de la parole (Carlson et al., 2001). Si certains choisissent des unités prosodiques (Pierrehumbert et Hirschberg, 1990; Chafe, 1994; Hirschberg et al., 1995), des tours de parole (Sacks, 1995), des phrases (Polanyi, 1978), ou des segments de discours définis par une intention (Grosz et Sidner, 1986), un grand nombre de linguistes considèrent l'unité discursive élémentaire sous forme de proposition (Grimes, 1975; Givón, 1984; Longacre, 1996).

Halliday (1985) et Chafe (1988) considèrent la proposition comme l'unité syntaxique la plus pertinente à l'analyse des situations et de leurs participants, l'analyse de la structure informationnelle (sujet-prédicat, information ancienne-nouvelle, focus), et à l'analyse de l'interaction linguistique, appelée "échange" par Halliday (1985), i.e. l'analyse des constructions syntaxiques par lesquelles les locuteurs déclarent, posent des questions, délivrent des ordres. Dans cette conception traditionnelle, la structure syntaxique véhicule une prédication à propos d'une entité et constitue le locus du réseau le plus dense de relations distributionnelles et de dépendance.

Thompson et Couper-Kuhlen (2005, p. 484) avancent également que les segments de parole les plus projetés (i.e. prévus par les locuteurs pour être véhiculés dans le discours) et traités comme complets sont les formats grammaticaux. Elles observent que le type de format le plus fréquent est composé d'un prédicat et des "expressions qui l'accompagnent". Ce format est celui de la proposition (*clause* en anglais) : "la proposition est le locus de l'interaction, dans le sens où elle représente le format grammatical le plus fréquent pris comme repère par les locuteurs en projetant les actions des énoncés des autres et en agissant à partir de ces projections." (p. 481)

Les propositions en conversation spontanée

Les chercheurs travaillant sur l'oral (e.g. Sacks et al., 1974; Blanche-Benveniste, 1997; Morel et Danon-Boileau, 1998) sont unanimes pour constater qu'il est difficile de parler de phrases à l'oral car le travail de formulation est apparent (reprises, faux-départs, marques du travail de formulation), mais aussi dans une conversation, qui comporte des interruptions ou encore la collaboration au travail d'énonciation (Bertrand et al., 2013).

Bien qu'elle soit l'unité syntaxique la plus opérationnelle à l'analyse de l'oral, la proposition n'est pas une unité discursive. Il faut donc souvent ajuster la définition de la proposition pour une description adaptée des segments discursifs autour desquels s'organise la parole. Certains segments de discours ne rentrant pas dans la définition traditionnelle de la proposition constituent des unités discursives qu'il convient d'individualiser. C'est le cas des amorces, de certaines répétitions ou encore des segments abandonnés ou repris qui ne comportent pas de verbe. Les segments interrompus en cours d'élaboration sont donc considérés comme des propositions dans cette étude s'ils sont séparés par des pauses ou possèdent leur propre contour prosodique.

L'hypothèse selon laquelle le discours oral spontané comporterait moins de hiérarchisation que l'écrit, bien qu'émise par de nombreux chercheurs, n'est pas confirmée par de nombreuses études. Certaines études (e.g. Poole et Field, 1976) ont même démontré le contraire. Sans proposer de synthèse de ces travaux, il convient ici de signaler que le discours oral est d'une complexité différente de celle de l'écrit (Halliday et Hasan, 1976). Alors que l'écrit est décrit comme un mode d'organisation compact avec une syntaxe relativement simple mais saturé lexicalement, l'oral possède une structure syntaxique compliquée avec une grande proportion de "subordonnées" mais une densité lexicale réduite. Blanche-Benveniste (1990) montre que certaines structures syntaxiques propres à l'oral comme les effets de regroupement s'expliquent par les stratégies du locuteur à l'oral : le locuteur peut introduire plusieurs éléments de lexique organisés en paradigme, dont le dernier semble entraîner une précision lexicale, ou introduire des effets d'hyponymie, où une liste comporte un terme générique qui sert à classer les autres termes énumérés.

1.5.2 Les *tone-units*

En ce qui concerne la prosodie et plus particulièrement l’intonation⁶, notre approche s’inscrit dans la tradition britannique (Halliday, 1967; Crystal, 1969; Wells, 2006), avec l’addition de certaines variables. La tradition britannique se base sur des domaines et considère l’intonation en termes de contours intonatifs dynamiques. L’accent est donc mis sur le ton, particulièrement sur le mouvement intonatif le plus conséquent dans chaque unité (i.e. les *tone-units*⁷), ayant lieu sur une syllabe accentuée et souvent équivalent au mouvement nucléaire. La syllabe nucléaire ne porte toutefois pas nécessairement le ton. Nous avons donc choisi d’aligner la parole avec le signal sonore au niveau du *tone-unit* (l’*Intonation Phrase* définie par Selkirk (1978)). Cette unité prosodique correspond à ce que Pierrehumbert et Hirschberg (1990) appellent l’*Intermediate Phrase* (syntagme intermédiaire).

Selon Halliday (1985), les unités informationnelles sont réalisées sous forme de groupes intonatifs. L’unité intonative prototypique est une séquence de mots combinés sous un seul contour intonatif cohérent, appelée le *tone-unit*. Ces unités sont généralement séparées par de brèves pauses, souvent celles causées par une légère rupture de tempo (Chafe, 1988). Les *tone-units* comportent au moins une syllabe nucléaire, i.e. une syllabe sur laquelle se réalise le mouvement intonatif maximal du groupe. La syllabe nucléaire marque le focus de l’information nouvelle et correspond généralement à la syllabe accentuée du dernier élément lexical majeur du groupe tonique. On parle pour cette syllabe de véritable “accent” et non plus de *stress*, utilisé pour les autres syllabes pleines des mots accentués du groupe. Cela correspond au focus non-marqué de l’information nouvelle. La tonique peut être placée ailleurs, signalant alors un focus d’information nouvelle marqué. Les groupes intonatifs coïncident habituellement avec les propositions, mais certaines raisons forcent parfois les locuteurs à étendre la proposition à plusieurs unités intonatives (Chafe, 1994).

Les *tone-units* peuvent être groupés dans des unités plus larges : les paragraphes (appelés également paratons par Wennerstrom (2001)). Les paragraphes sont des groupes de *tone-units* qui forment un mouvement intonatif global : la hauteur intonative décline progressivement d’un *tone-unit* à l’autre, jusqu’à ce qu’elle soit réinitialisée à la formation d’un nouveau paragraphe (Cruttenden, 1986). Les para-

6. L’intonation est le contour mélodique de la parole due à la hauteur et à l’intensité des voyelles. Elle structure la parole en organisant temporellement la hauteur intonative perçue dans les énoncés. L’étude de l’intonation fait partie de la prosodie, avec celle du volume, de l’intensité, du tempo, et du rythme (accentuation, débit de parole, hésitations, pauses).

7. Nous utilisons le terme “groupe intonatif” en tant que synonyme de *tone-unit* dans ce travail.

graphes font partie des éléments permettant de prévoir/prédire un tour de parole conversationnel (Sacks et al., 1974), et peuvent être délimités par des pauses, une rupture mélodique et des allongement syllabiques.

Selon la tradition britannique, la partie la plus importante et le constituant majeur du contour est le noyau (i.e. nucleus) (N), la syllable la plus fortement marquée du *tone-unit*. Le nucleus est la seule partie obligatoire du *tone-unit*. Ce sont ses caractéristiques toniques qui définissent le *tone-unit*. Un *tone-unit* doit avoir au moins une syllabe accentuée.

Les autres parties du *tone-unit* sont l'avant-tête (*pre-head*, P), formé par les syllabes non-accentuées avant le premier accent, la tête (*head*, H), qui va de la première syllabe accentuée jusqu'au nucleus, et la queue (*tail*, T), qui suit le nucleus. L'exemple (5) tiré de notre corpus montre la déconstruction d'un *tone-unit* (TU) en ses sous-parties :

- (5) TU[PRE-HEAD[and so she dis] HEAD[owned everything that she could associate my] NUCLEUS[na] TAIL[na]]

La tradition britannique prête également des fonctions pragmatiques aux contours intonatifs, sans pour autant établir de relation figée entre forme et fonction. Les tons descendants sont généralement terminatifs (i.e. ils indiquent que le *tone-unit* en cours ne présuppose pas l'existence d'un *tone-unit* à venir complétant l'information qu'il apporte; Bolinger, 1984). Au contraire, les tons montants sont plutôt indicateurs d'une dépendance informationnelle au *tone-unit* qui le suit, ou d'un appel au co-locuteur pour les tons montant très haut par rapport au registre du locuteur (ibid.). Les tons montants-descendants sont des tons emphatiques, alors que les tons descendants-montants impliquent ou présupposent une information non explicitée (pragmatique dans la plupart des cas, telle qu'une hésitation ou un sous-entendu; Wells, 2006).

Bien que certains chercheurs aient noté quelques problèmes d'identification floue des éléments constituant la pré-tête et la queue des *tone-units*, nous estimons que cette approche est adéquate à la description de l'anglais conversationnel, à condition d'un certain nombre d'ajustements de la théorie à nos données (cf. Chapitre 2). Nous complétons également l'apport théorique de la tradition britannique par une approche paramétrique avec les outils Momel-Intsint (Hirst, 2007; Bigi, 2012; Bigi

et Hirst, 2012), qui fournissent plusieurs niveaux de représentation utiles à l’analyse de nos données.

L’algorithme Momel fournit une représentation phonétique de la parole. D’un point de vue formel, cet algorithme permet une représentation automatique discrète de la courbe brute de la fréquence fondamentale (i.e. F0, qui participe à la perception de la hauteur intonative) sous forme d’une séquence de points cibles. Ces points cibles peuvent également être considérés comme des cibles phonologiques. Les points cibles sont convertis automatiquement en symboles tonals selon l’alphabet INTSINT. Ces symboles renvoient à des tons relatifs – *upstep* (réhaussement), *downstep* (abaissement) – utiles à la classification de niveaux de frontière. D’autres paramètres servent également à classifier des degrés de frontière et des degrés de prééminence (i.e. repérage des frontières intonatives grâce à la mesure de la durée, du débit de parole, de l’allongement syllabique).

Le choix de la théorie britannique plutôt que celui du modèle autosegmental de phonologie intonative (i.e. AM, Pierrehumbert, 1980; Beckman et Pierrehumbert, 1986; Ladd, 2008)⁸ est motivé par plusieurs paramètres. Bien que le modèle autosegmental de phonologie intonative soit extrêmement complet (notamment par ses correspondances entre tons et effets sémantico-pragmatiques), la théorie britannique offre une représentation assez iconique de l’intonation, en ce qu’elle se base sur les contours. De plus, l’AM est extrêmement utile pour les objets d’étude ayant des enjeux phonologiques, alors que la théorie britannique possède des implications d’ordre pragmatique et informationnel, puisque ses unités de segmentation de la parole (i.e. les *tone-units*) sont également des unités informationnelles (Halliday, 1985; Chafe, 1994). La prise en compte de ces aspects sont notamment essentiels en Analyse du Discours. Le *tone-unit* représente en outre une unité de segmentation plus petite que le syntagme intonatif, pertinente à l’observation des frontières et des dépendances. Un dernier critère méthodologique est également entré en compte. Nous avons choisi les *tone-units* et la tradition britannique pour leur compatibilité avec les outils d’an-

8. Dans le modèle autosegmental de phonologie intonative (i.e. AM; Pierrehumbert, 1980; Beckman et Pierrehumbert, 1986; Ladd, 2008), la parole est considérée comme une structure hiérarchisée de domaines prosodiques. Les énoncés sont découpés en constituants qui sont hiérarchiquement dominés par des unités prosodiques de niveau supérieur dans la hiérarchie prosodique. Si nous nous intéressons à l’intonation en particulier, la plus grande unité prosodique marquée intonativement est l’IP (*Intonational Phrase* ou syntagme intonatif). À travers les langues étudiées jusqu’ici, une IP est définie par un ton de frontière à son extrémité droite (parfois à celle de gauche également), un allongement syllabique final, et une pause optionnelle après l’unité (Jun et Fletcher, 2014; Katsika et al., 2014). L’unité inférieure dans cette représentation phonologique est l’ip (*intermediate phrase* ou syntagme intermédiaire en français). L’ip possède au moins un accent nucléaire, et est bornée à sa droite par un accent de syntagme (*phrase accent*; Beckman et al., 2005). L’IP est donc constituée d’une ou plusieurs ips, et se termine par un ton de frontière.

notation Momel-Intsint. Le modèle de l'AM nous a néanmoins influencée par son attention particulière portée aux tons de frontière. Dans une volonté de déployer l'éventail de critères prosodiques pertinents à l'analyse du discours, nous avons mesuré manuellement la hauteur intonative des tons frontière (situés sur les syllabes accentuées initiales et finales de chaque *tone-unit*), ainsi que la hauteur intonative moyenne sur l'ensemble du *tone-unit* pour chaque segment étudié. Les travaux de l'AM soulignent en effet que la hauteur mélodique joue un rôle important au niveau de la structure du discours et de l'information.

De même, le choix des *tone-units* plutôt que des unités interpausales (segments de parole entre des pauses de plus de 200 millisecondes), couramment utilisées en Analyse Conversationnelle, est également motivé par la plus petite taille des *tone-units*, davantage adéquate à l'observation des frontières. Les *tone-units* permettent également la prise en compte de contraintes plus variées que la gestion des tours de parole, comme la structure informationnelle et les contours mélodiques.

1.5.3 Les unités gestuelles

Dans le cadre de l'Analyse de Discours Multimodale, nous estimons que la communication s'établit à travers des gestes oraux, manuels, et faciaux. Les structures syntaxiques, sémantiques, prosodiques, pragmatiques, et interactionnelles employées par les locuteurs ne constituent pas des entités séparées aux frontières bornées, et sont employées par les différents articulateurs. Il a été montré que les gestes manuels produits en co-occurrence avec la parole sont synchronisés avec cette dernière sémantiquement (e.g. McNeill et Duncan, 2000; McNeill, 2005), temporellement (e.g. Chui, 2005), et structurellement (e.g. Kita et Özyürek, 2003; Lewandowski et Öskaliskan, 2018). En raison de cette coordination, les processus de production de la parole et de la gestualité sont liés à tous les niveaux (e.g. Iverson et Thelen, 1999; Bernardis et Gentilucci, 2006; Roustand, 2012; Ferré, 2014).

Le rôle essentiel des gestes co-verbaux a été démontré à travers le prisme de nombreux champs théoriques, notamment par la pragmatique (e.g. Lascarides et Stone, 2009), la linguistique cognitive (e.g. Sweetser, 2006), et la psycholinguistique (e.g. McNeill, 2005). Les gestes référentiels (i.e. ceux qui participent à la description ou à la représentation d'objets, d'actions, d'idées) comme non-référentiels (i.e. ceux qui véhiculent de l'information abstraite à propos de la grammaire, du discours et/ou de l'interaction) sont bénéfiques à la fois au locuteur avec la facilitation de la mise

en mots et l'amélioration de la fluidité de la production vocale du locuteur (Krauss et al., 2000) et au co-locuteur, avec l'apport d'informations supplémentaires à la parole (e.g. McNeill, 1992 ; 2005 ; Goldin-Meadow, 2015), l'amélioration de la perception de descriptions (e.g. Driskell et Radtke, 2003), la désambiguïsation lexicale de certains mots (e.g. Holle et Gunter, 2007), l'amélioration de la compréhension dans des conditions de communication difficiles (e.g. Sueyoshi et Hardison, 2005), ou encore l'intérêt de la perception de gestes coverbaux sur la mémoire à long terme (e.g. Church et al., 2004).

Plusieurs théories psycholinguistiques ont proposé des motivations cognitives à la co-occurrence des gestes à la parole (e.g. McNeill, 1992; Krauss et al., 2000; Kita et Özyürek, 2003; Goldin-Meadow, 2005; de Ruiter, 2007; Hostetter et Alibali, 2010). Les chercheurs s'accordent sur le fait que les gestes semblent provenir de représentations cognitives d'informations spatiales.

En ce qui concerne les gestes manuels et faciaux, notre travail s'appuie sur la définition du geste proposée par Kendon (2004) : est considérée comme geste toute action visible d'une partie du corps utilisée comme énoncé ou partie d'un énoncé. Cette définition présente plusieurs implications. Elle ne s'appuie pas sur la présomption d'une intention de communication et évite donc la question de la gestualité "délibérée" (i.e. consciente ou inconsciente) des locuteurs, qui produisent nombre de gestes sans s'en rendre compte sans que cela empêche les co-locuteurs de les interpréter (Debras, 2018). Avec cette définition un geste n'a pas non plus besoin d'être produit en co-occurrence avec la parole pour contribuer à l'interaction. De plus, cette définition souligne que le critère premier pour considérer un mouvement kinésique comme un geste correspond à la façon dont il s'insère dans une succession d'autres actions ou dimensions, constitutives de l'interaction en cours (Debras, 2018). Un geste est donc un mouvement du corps qui devient signifiant en s'inscrivant dans le contexte simultané et séquentiel des actions des participants (cf. Chapitre 2).

Nous limitons cette étude aux gestes manuels, aux mouvements de tête et de sourcils, et à la direction du regard. Le cadre formel de Kendon révèle toutefois quelques limites quant à l'observation et à l'analyse des subordonnées dans la conversation spontanée, qui nécessitent une démarche davantage fonctionnelle. Nous ancrons notre travail dans l'approche de McNeill (1992; 2005), qui s'intéresse à la fonction de chaque geste par rapport à la parole (plus ou moins) co-occurrence. À chaque unité gestuelle (commençant à l'ébauche du geste et finissant au retour à la position de repos s'il y en a une ou à un changement de coordonnées spatiales/trajectoire)

correspond une dimension ou fonction principale vis-à-vis de la parole co-produite (cf. Chapitre 2).

L'approche de McNeill est basée sur la synchronie entre geste et parole, l'ensemble formant un système selon lui complètement intégré. La théorie du *Growth Point* (McNeill, 1992; McNeill et Duncan, 2010) se base par exemple sur l'observation selon laquelle parole et gestualité co-expriment des unités d'idée, utilisant chacun une forme différente. Alors que la gestualité emploie une forme de représentation globale et synthétique, la parole utilise une forme de représentation analytique et combinatoire (McNeill, 1992). Nous nuancions cette pensée en considérant les gestes et la parole comme un système effectivement intégré, mais chacun possédant des spécificités propres. De même, le geste n'est pas nécessairement produit en relation à la parole co-occurrence : il arrive que le geste soit en relation avec la parole immédiatement produite après, ou que l'apogée du geste soit alignée temporellement avec la syllabe proéminente, mais que geste et item lexical ne correspondent pas sémantiquement.

En résumé, notre travail s'appuie sur trois grandes unités théoriques, pouvant se chevaucher lors de la production langagière. La proposition représente le format de segmentation et de traitement du flux verbal, dans une tradition plutôt fonctionnelle (e.g. Halliday, 1985; Chafe, 1988; Thompson et Couper-Kuhlen, 2005). Les *tone-units* sont utiles à la représentation et l'analyse des données vocales, et proviennent de la tradition britannique (Halliday, 1967; Crystal, 1969; Wells, 2006). Nous ajoutons quelques paramètres à cette vision, nous permettant un point de vue à la fois phonétique, prosodique, et phonologique pour certains phénomènes. Enfin, dans un cadre davantage fonctionnel que formel (McNeill, 1992; McNeill et Duncan, 2010), nous segmentons les mouvements manuels et faciaux à l'aide de la notion d'unité gestuelle.

1.6 Travaux antérieurs sur la subordination

La littérature syntaxique et discursive portant sur les subordonnées établit clairement le postulat selon lequel ces constructions syntaxiques sont dépendantes de la structure hôte, et constituent un support informatif pour les éléments primaires (qui, eux, représentent un apport informatif), en les localisant, en les définissant, et/ou en les cadrant. Une correspondance est également généralement établie entre la construction syntaxique hiérarchique comme elle est décrite classiquement (prin-

cipale + subordonnée) et les valeurs énonciatives des relations (dans le même ordre, prédication d’une relation + reprise d’une relation stabilisée). Ces constructions sont donc couvertes d’un voile uniforme de dépendance, de supplémentarité, ou d’adjonction. Aucun travail dans le cadre de la multimodalité n’a à notre connaissance été réalisé sur les subordonnées. Nous nous limitons à passer en revue la littérature à propos de l’organisation linéaire des subordonnées (dépendance vs. autonomie), bien que les subordonnées sous étude soient également vues comme de l’arrière-plan informationnel du point de vue de leur organisation thématique (premier plan informationnel vs. arrière-plan; Lelandais et Ferré, 2017).

1.6.1 La subordination en syntaxe

D’emblée, le concept de subordination implique à la fois l’intégration et la supplémentarité par rapport au co-texte, mêlant deux modes/niveaux d’organisation des constituants linguistiques. Les premières études partent de la supplémentarité pour tirer des inférences à propos de l’intégration. Ainsi, la description grammaticale des subordonnées influence grandement leur description syntaxique (et vice-versa), si bien que l’un des deux modes/niveaux d’organisation (linéaire et thématique) trouve toujours une correspondance dans l’autre.

Le “nécessaire” et le “secondaire” grammatical

Sur le plan grammatical, certains éléments du message peuvent apparaître nécessaires au sens sans pour autant s’imposer comme parties constitutives de l’énoncé. Cette problématique correspond à ce que Chervel (1979) appelle l’énoncé minimum lorsqu’il retrace l’historique des circonstanciels dans la grammaire :

“Les mots qui marquent les circonstances ne sont que des adjoints, que les mots précédents n’exigent pas nécessairement. Il faut donc bien distinguer les déterminations nécessaires d’avec celles qui n’influent en rien à l’essence de la proposition grammaticale, en sorte que sans ces adjoints on perde à la vérité quelques circonstances de sens, mais la proposition n’en soit pas moins telle proposition” (Dumarsais cité d’après Chevalier (2006, p. 699)).

Si certains compléments sont classifiés comme aussi importants que le sujet et le verbe car ils sont “nécessaires grammaticalement” à la correction de l’énoncé (Chervel, op. cit., p. 10), ceux qui pourraient disparaître sans affecter la gramma-

ticalité de la phrase sont “nécessaires logiquement”, mais non grammaticalement. Les premiers sont appelés compléments prochains, les seconds compléments éloignés. Chervel (ibid.) observe que cette analyse est l’extension de celle opérée pour classer les subordinées : si certaines sont indispensables à la complétude de l’énoncé (e.g. les complétives), d’autres ne sont requises que par le “sens” (e.g. les appositives). L’énoncé resterait “correct” même si on les supprimait. Mais comment définir la nécessité “logique” et le “sens” mentionnés par ces grammairiens? Dans la grammaire catégorielle, le “circonstanciel” va s’opposer à “l’essentiel” :

“Une proposition subordonnée ou ajoutée est circonstancielle quand elle n’exprime qu’un fait ou une réflexion secondaire, que l’on pourrait retrancher sans nuire beaucoup au sens” (Chemin-Dupontès, 1811, p. 39).

Gross (2005) explique que l’analyse classique des complexes de propositions consiste à les décomposer en deux sous-ensembles inégaux et complémentaires : une principale et une subordonnée. La principale est considérée comme fondamentale, la subordonnée n’apportant que des informations de nature accessoire ou périphérique. Cette subordonnée est rattachée à la principale par une conjonction ou une locution conjonctive (ou prépositive).

La “subordonnée” est donc analysée comme un constituant de la proposition. Elle peut apparaître sous chacun des constituants de la proposition (groupe verbal, adjectival, propositionnel, nominal), mais nous nous intéresserons ici principalement à celles dominées par un groupe nominal (relatives) et propositionnel (circonstancielles). Reste à savoir si cette domination est celle d’un complément sous-catégorisé (c’est le cas des complétives, auxquelles nous ne nous intéresserons pas ici), ou celle d’un modifieur, i.e. un constituant dit “facultatif”. C’est cette dernière catégorie qui fait l’objet de notre attention. La modification par enchâssement d’un constituant dans un autre est le propre du phénomène de la récursivité, décrit par Halliday (1985).

Pour Halliday (ibid.), les propositions sont des unités qui possèdent un noyau (*nucleus* en anglais) correspondant à un événement, un procès ou un état, et une périphérie correspondant à la localisation, l’environnement, et/ou les événements précédents. Alors que le noyau et les éléments du cœur (*core*) sont des composants obligatoires sémantiquement et que leurs composants syntaxiques correspondants sont également indispensables (au moins une partie du cœur en fonction du prédicat), la périphérie est “optionnelle”. Les locutions adverbiales comme “yesterday mor-

ning” ou “à l’heure de pointe” peuvent être ajoutées à la proposition mais appartiennent à sa périphérie. Halliday (1985) parle d’éléments “primaires” et d’éléments “secondaires” (p.195) dans une paire de propositions liées par l’élaboration, où une proposition élabore le sens de l’autre par la spécification ou la description. La proposition secondaire n’introduit pas un nouvel élément mais fournit une caractérisation plus poussée de ceux déjà présents.

En décrivant les complexes propositionnels, Halliday (1985) fait donc l’exposition d’une structure relationnelle montrant comment chaque unité qui la compose est liée au reste du texte par un réseau de ces relations. Il distingue notamment deux types majeurs de relations syntaxiques. Alors que la parataxe relie des éléments de statut égal (où l’élément initiateur et l’élément continuatif sont libres et peuvent constituer chacun un ensemble fonctionnel), l’hypotaxe lie des éléments de statut inégal (où l’élément dominant est libre, mais celui dépendant ne l’est pas). Ce sont des relations générales à tout complexe, qu’il soit formé par des mots, des groupes ou des propositions.

Un fonctionnement paratactique n’aurait pas les mêmes implications sémantico-pragmatiques qu’un fonctionnement hypotactique : en considérant ces deux fonctionnements comme des façons différentes d’encoder le sens, Lakoff (1984) propose que l’hypotaxe est un mode de fonctionnement centré sur le locuteur alors que la parataxe témoigne plutôt d’une orientation vers le co-locuteur, qui se voit confié un plus grand travail d’interprétation pragmatique.

Les modifieurs et la dépendance

Les éléments facultatifs dans le complexe propositionnel sont communément appelés des modifieurs (Delaveau, 2001), qu’ils fonctionnent au niveau du syntagme ou de la proposition⁹. Dans la division catégorielle des complexes propositionnels en une classification qui comprend deux sous-groupes inégaux et complémentaires, i.e. une proposition principale et une subordonnée, les modifieurs sont vus comme des constituants optionnels qui fonctionnent au niveau du syntagme ou de la proposition. Cette classification découle du concept des énoncés minimaux : certains éléments du message sont estimés sémantiquement nécessaires sans pour autant représenter des

9. Nous utiliserons “proposition” en tant que synonyme à *clause* dans ce travail, bien que les deux termes soient d’orientations différentes : tradition française et/ou connotation sémantique pour le premier, tradition anglo-saxonne et/ou orientation syntaxique pour le second.

éléments constitutifs.

En syntaxe, les modifieurs sont souvent décrits comme des ajouts au contenu propositionnel de la structure hôte ou enchâssante (Quirk et al., 1985; Huddleston et Pullum, 2002). Deux types sémantiques sont distingués dans la description des relations de dépendance (van Rijn, 2017). Alors que certaines têtes de constituants nécessitent par nature de faire référence à un élément dépendant, que l'on considère comme un argument de ces têtes, d'autres éléments dépendants ne sont pas présumés de manière inhérente par une tête et sont considérés comme des modifieurs. Ils fournissent simplement une caractérisation sémantique plus poussée d'un référent (ou état de faits) exprimé par la tête, ou complètent la tête avec des informations supplémentaires. Un exemple de modifieur attesté dans la littérature est la proposition relative (van Rijn, 2017).

Les constructions subordonnées sont généralement définies comme dépendantes d'une autre prédication d'un point de vue micro- ou macro-syntaxique (Tomlin, 1985; Lehmann, 1988). Lehmann (1988, p. 181) propose une définition de la notion de dépendance, liée à celle de l'intégration. "Une relation grammaticale R connectant les syntagmes X et Y est une relation de dépendance si X occupe une place grammaticale de Y ou vice-versa. Dans une relation de dépendance, Y dépend de X si X détermine la catégorie grammaticale du complexe and ses relations externes." En somme, une proposition B est dépendante d'une proposition A si B manque d'autonomie fonctionnelle, et est liée à l'un des composants de A : B est hiérarchiquement déclassée dans son lien avec A.

Alors que les subordonnées sont largement définies comme dépendantes, la littérature montre peu de consensus sur la délimitation de portées et de frontières claires pour ces structures. Alors que certains chercheurs appréhendent les constructions relatives comme des éléments enchâssés (Jackendoff, 1977), d'autres (Jespersen, 1927; Fabb, 1990; Peterson, 1999) considèrent que certaines constructions relatives sont extérieures à la structure syntaxique de la proposition principale. Plus spécifiquement, la différenciation entre les relatives déterminatives et les appositives relatives sur un plan syntaxique est problématique, et sujette à débat (Borsley, 1992; Arnold et Borsley, 2008). Alors que les appositives relatives dérivent de la coordination pour Burton-Roberts (1999) et De Vries (2006), Quirk et al. (1985, p. 1257) et Biber et al. (1999, p. 135) recommandent des niveaux de représentation spécifiques dans la subordination, avec respectivement des notions telles que les "relatives télescopées" et les "éléments périphériques". Les appositives relatives sont

aussi classifiées en tant qu’adverbiaux par Biber et al. (1999, p. 853). Enfin, dans le riche système des relations syntaxiques proposé par Matthiessen et Thompson (1988, p. 238), les circonstancielles et les relatives non-restrictives sont “moins subordonnées” que d’autres structures, appartenant à des relations de combinaison hypotactiques, où propositions dominantes et dépendantes sont des sœurs syntaxiques, plutôt que de l’enchâssement où la subordonnée est un constituant de la proposition principale.

Cristofaro (2003) et Langacker (2008) signalent que la subordination sémantique et illocutoire n’est pas nécessairement alignée sur la subordination syntaxique, et que la notion de subordination se comprend beaucoup mieux en termes de conceptualisation dynamique (i.e. les points de départ et repères cognitifs pris par les locuteurs pour délivrer leur message, qui changent de construction en construction). Il a également été proposé qu’un continuum de subordination existe au sein d’un seul et même type syntaxique de proposition (Tao et McCarthy, 2001), et que certains types de propositions subordonnées ne soient pas décrits de la meilleure manière en tant que tels, surtout les circonstancielles et les appositives relatives (Depraetere, 1996; Thompson, 2002).

Le fait que certains éléments du message soient estimés sémantiquement nécessaires sans pour autant représenter des éléments constitutifs a été remis en question par un grand nombre de linguistes (e. g. Haiman et Thompson, 1984; Chafe, 1988; Smessaert et al., 2005). L’acception syntaxique du terme “subordination”, i.e. élément en relation de dépendance grammaticale morphologiquement marquée par une conjonction de subordination, est elle-même remise en cause :

“Un récent article de Haiman et Thompson (1984) a attiré l’attention sur la nature incohérente des divers phénomènes traditionnellement associés à la subordination, et a suggéré que le concept soit abandonné puisqu’il échoue à adopter tout ensemble unifié d’observations¹⁰” (Chafe 1988, p. 18).

La subordination, en tant qu’outil de description de la langue conversationnelle contemporaine, est décrite comme inadéquate pour une finesse d’analyse, remise en question par des distinctions morphosyntaxiques imprécises, particulièrement quant à la nature des éléments introducteurs (Aarts et McMahan, 2008). Quelques études ont montré qu’elle n’est pas un type de connexion phrastique homogène, car elle inclut plusieurs niveaux, entre les arguments d’une proposition et ses adjonctions,

10. Notre traduction.

qu'elle considère comme liés plus ou moins de la même façon à une structure principale.

Haiman et Thompson (1984) argumentent donc en faveur de l'abandon de ce label trop équivoque pour se concentrer sur d'autres paramètres évoqués pour une meilleure analyse de la combinaison de propositions : l'identité de sujet, de temps, ou d'aspect entre les deux propositions, la réduction de l'une des deux, la portée ou le cadre de l'une agissant sur l'autre, ou la non-chronologie temporelle entre les deux propositions.

D'autres critères sont également suggérés pour l'évaluation des combinaisons de propositions, dans une hiérarchie de relations syntactico-sémantiques : une relation sémantique étroite entre deux propositions corréle un lien syntaxique étroit (Van Valin, 1984; Van Valin et La Polla, 1997)¹¹. Les propositions sont des unités comprenant un noyau essentiel (qui contient le prédicat, correspondant à un événement, procès, ou état, et ses compléments essentiels), et une périphérie optionnelle (qui correspond au cadre spatio-temporel comme la localisation ou l'environnement; Van Valin, 1984; Halliday, 1985). Une proposition attachée aux composants nucléaires d'une autre proposition montre un lien plus fort qu'une proposition attachée aux éléments périphériques d'une autre proposition. Cette typologie des liens syntaxiques donne un continuum, allant de la cosubordination nucléaire (marquant le lien syntaxique le plus serré) à la coordination périphérique (encodant le lien le plus faible). Tous les procédés de hiérarchisation offrent l'image d'un dégradé. Le pôle supérieur est la parataxe (Lehmann, 1988) avec la conservation de toutes les catégories verbales d'aspect, de temps, et de mode; une haute transitivity, et une grande dynamisme (Longacre, 1996). Le pôle inférieur est la perte de tous les renseignements caractéristiques du verbe avec réduction et/ou nominalité (Lehmann, 1988; Ramm et Fabricius-Hansen, 2005).

Dans la grammaire cognitive, la dépendance est expliquée en termes de conceptualisation. Selon Langacker (2008, p. 199), certaines structures "requièrent le support d'autres structures [...] pour leur propre manifestation. En conséquence, une

11. Van Valin (1984, p. 555) et Van Valin et LaPolla (1997) postulent une hiérarchie des relations entre propositions dans l'interface syntaxe/sémantique. Selon eux, plus la relation sémantique entre deux propositions est proche, plus elles sont liées syntaxiquement. Ces chercheurs s'appuient sur le principe d'iconicité, qui établit un parallélisme entre l'ordre du monde et l'ordre de la langue. Trois domaines distincts sont établis dans la proposition : le noyau (prédicat), le cœur (arguments fondamentaux du noyau –modalités sujet et adjectivales) et la périphérie (éléments obliques et éléments définissant le cadre spatio-temporel). Une unité reliée dans la périphérie est moins liée à la principale qu'une unité reliée dans le cœur ou le noyau.

structure dépendante fait schématiquement référence à une structure autonome support en tant qu’aspect intrinsèque de sa propre caractérisation.” Un modifieur, par exemple, est dépendant d’une tête autonome (p. 203). La dépendance est une relation asymétrique entre deux structures connectées, avec une structure primant sur l’autre. Par exemple, dans “the flowers I bought today are for my sister”, le champ référentiel de la proposition principale prime sur celle de la proposition relative “ \emptyset I bought today”. Le centre structurel de la proposition principale (i.e. le prédicat “are for my sister”) domine également le prédicat dans la proposition subordonnée, puisque le prédicat dans la proposition subordonnée élabore à partir d’un composant de la proposition principale. La structure complexe est une organisation à plusieurs niveaux dans laquelle une relation propositionnelle incorpore l’autre en tant que participant.

Dans la grammaire des propriétés (e.g. Blache, 2016), la dépendance syntactico-sémantique est représentée par différents sous-types de relations de dépendances, dont la modification. Un objet modifie l’organisation syntaxique d’un autre et modifie également sa structure sémantique.

Ce sont autant de critères qui poussent à évaluer le type de connexion entre les propositions selon un éventail plus large et plus détaillé de paramètres syntactico-sémantiques (Van Valin, 1984), ou à sortir du cadre de la micro-syntaxe tout en établissant des distinctions plus fines, entre fonctionnement régi ou non-régi, autonomie énonciative, autonomie modale (Benzitoun, 2013; Debaisieux, 2006; 2016), autonomie pragmatique (Thompson, 2002; Heringa, 2007).

1.6.2 La subordination dans le discours

Alors que les constructions syntaxiques sous étude sont analysées comme dépendantes d’un point de vue syntaxique dans la littérature, elles sont décrites comme réalisant des relations d’élaboration au niveau de la structure discursive. Ces relations d’élaboration contrastent avec les relations de continuation du discours.

Les travaux fondateurs de Matthiessen et Thompson (1988) émettent l’hypothèse que la systématisme de la structure grammaticale découle de la systématisme des demandes fonctionnelles placées sur le langage. Il y aurait donc un lien entre syntaxe et organisation textuelle à travers des notions telles que les noyaux

et les satellites, une similarité entre la combinaison des propositions et les textes. Ils soutiennent que ce lien est une analogie fondamentale, malgré la distinction entre l'organisation rhétorique du discours et l'organisation syntaxique des propositions en combinaisons. Deux grands points sont abordés : les mêmes relations générales gouvernent les relations entre les propositions dans les combinaisons de propositions et celles entre les unités textuelles de plus haut rang, et les propositions ont la même structure combinatoire que les unités textuelles de plus haut rang, comme membres de liste ou comme noyau et satellite.

Selon la théorie développée par Asher et Lascarides (2003), les relations de discours sont créées par les différentes relations sémantiques entre les propositions. La subordination discursive a pour effet de permettre à la proposition suivant immédiatement la subordonnée de s'attacher directement à la proposition dominante dans le discours, créant ainsi une relation de cohérence. À l'intérieur d'une séquence narrative par exemple, gouvernée par une séquence temporelle d'événements, les propositions sont coordonnées discursivement si elles dénotent des événements indépendants, ordonnés temporellement. Des descriptions statiques, quant à elles, peuvent interférer avec la séquence d'événements, créant une relation d'arrière-plan avec les événements temporels qui l'entourent. Elles sont donc subordonnées.

Dans la *SDRT* (*Segmented Discourse Representation Theory*), l'arrière-plan désigne, au sens large, toutes les relations de discours pouvant être subordonnées à d'autres, comme l'élaboration (e.g. Lascarides et Asher, 2007). Cette même notion désigne toutes les relations entre noyaux et satellites en *RST* (*Rhetorical Structure Theory*, Mann et Thompson, 1988). Il peut être exemplifié par une séquence de propositions comme dans l'extrait (6) de notre corpus où la deuxième proposition décrit un état qui pourrait empiéter temporellement sur l'événement introduit dans la première ; c'est-à-dire que SC véhicule une description d'arrière-plan relative à L.

- (6) Alex L so i did the second year of hem # like English # here
 Rhianna mhmh
 Alex SC **(h) (laughs) which was interesting**
 R (h) hem yeah and then i just graduated now #

La fonction du satellite est d'augmenter la capacité du co-locuteur à comprendre ce qui est présenté dans le noyau (Taboada et Mann, 2006). La définition de la *RST* ne restreint pas l'ordre d'occurrence du noyau et du satellite, i. e. le satellite peut être placé avant ou après le noyau (lien cataphorique vs. lien anaphorique), et la relation de discours peut joindre les deux segments dans les deux sens. Ces fonctions

sont présentes dans tout type de séquence (narration, argumentation, description) et ont des corrélations dans la structure intentionnelle du discours (Grosz et Sidner, 1986) : si la coordination veut que l'intention communicative du premier segment (le *DSP-Discourse Segment Purpose*) soit satisfaite avant celle du second, la subordination implique quant à elle que la satisfaction de l'intention communicative du segment subordonné contribue à satisfaire celui du segment dominant. Selon Carston (1993), les constituants subordonnés et conjoints ne forment un ensemble de pertinence cohérent qu'avec les éléments superordonnés, comme une seule unité.

De nombreuses études discursives se sont intéressées aux morphèmes introducteurs des subordonnées. Dès lors que le discours est pris en compte, la conjonction est décrite comme multifonctionnelle (e.g. Debaisieux, 2006; Berrendonner, 2008). Les dispositifs tels que “because” peuvent être utilisés pour spécifier des relations grammaticales dans les constructions dominées par une catégorie (i.e. ils peuvent être fonctionnels micro-syntaxiquement, spécifiant une relation verbe-circonstancielle comme causale). Ce même “because” peut tout aussi bien spécifier la relation de cette même signification conventionnelle entre deux segments de discours. Dans un type de construction plus large, la même signification conventionnelle (i.e. la cause) peut être spécifiée entre deux actions communicatives dans une interaction, reliant non seulement les actions réalisées par la parole mais aussi des actions co-verbales impliquées dans le même schéma interactionnel. En somme, “because” peut tout aussi bien spécifier des relations syntaxiques que discursives et/ou pragmatiques. La présence de marqueurs (à l'origine) grammaticaux tels que “parce que” ou “donc” restreint la portée des fonctions sémantiques et pragmatiques que le segment peut remplir. Par exemple, un segment introduit par “because” ne peut pas prendre de valeur concessive sur le discours. L'autonomie pragmatique d'une unité est vérifiée s'il est possible de reconstruire autour de cette unité une structure prédicative.

Les subordonnées font partie des constituants “enrichis pragmatiquement” (Blakemore, 2006) : elles agissent sur les contraintes de traitement et d'interprétation du discours, notamment par le biais des relations de cohérence en déclenchant certaines opérations d'un système inférentiel. Si l'on considère comme Mann et Thompson (1988) ou Grosz et al. (1995) que les relations entre les portions successives du discours sont caractérisées en termes d'inférences que l'on peut dégager de chaque portion, les constructions subordonnées sous étude facilitent la charge de traitement des co-locuteurs en structurant le message d'une façon à ce qu'ils en déduisent les bonnes inférences rapidement. Ces segments de discours permettent un contrôle à la fois sur

le dire et sur le dit, sur la signification que le co-locuteur attribue aux énoncés. Si la supposition commune selon laquelle les sujets ont tendance à généraliser à partir de ce qu'ils apprennent est valide, l'un des rôles de ces relations de cohérence est de permettre à la locutrice de promouvoir ou d'inhiber ces généralisations. Le contexte discursif filtre l'ensemble des relations rhétoriques et temporelles qui peuvent être utilisées pour attacher l'énoncé à la représentation du texte. Ces segments ont donc un rôle procédural et cognitif dans le discours : ils règlent les opérations de mobilisation de connaissances et répartissent les contenus propositionnels dans des blocs homogènes.

À ce niveau pragmatique, les structures faisant l'objet de notre étude entrent dans la définition des signaux de contextualisation (*contextualisation cues*, Gumperz, 1992). Ce sont des segments de parole relativement courts qui forment l'arrière-plan de l'interprétation d'un énoncé/d'une séquence. Ces séquences invoquent le contexte, dans le sens où ce sont des portions qui fournissent des données additionnelles permettant au co-locuteur d'évaluer ou de mesurer les événements principaux d'un texte (par exemple en précisant la sélection de la référence, en clarifiant la structure rhétorique, ou en indiquant la force illocutoire). Ce sont donc des données dites de "second rôle", qui soutiennent l'architecture du message mis en avant en s'appuyant sur la reconstruction inférentielle du co-locuteur pour en évoquer l'arrière-plan interprétatif essentiel. Si ces ressources sont donc des procédés discursifs exploités par le locuteur pour construire, modifier et accéder au contenu des modèles mentaux d'un discours en train de s'énoncer représentés chez les locuteurs et co-locuteurs, ce sont également des outils pour coordonner l'attention des participants tout au long du flux discursif, produit dans un certain contexte dont ils font partie (Clark et Krych, 2004; Loock et O'Connor, 2013).

Gumperz (op. cit.) avance que les signaux de contextualisation affectent la façon dont les participants produisent des inférences à trois niveaux différents : la gestion de l'interaction, comme la régulation de la prise de parole et le signal du degré de pertinence informationnelle, le séquençage, par exemple en indiquant des implications et en désambiguïsant l'intention du locuteur dans les énoncés, et enfin le cadrage, en générant des attentes quant à la nature de l'interaction et en accordant son ton ou son ambiance. Levinson (2003, p. 29) précise en conséquence que ces signaux véhiculent souvent du contenu non-propositionnel, comme des informations métalinguistiques, rhétoriques, modales, ou encore interactionnelles.

Toutefois, Thompson (2002) revient sur le consensus selon lequel l'information

apportée par les subordinées sont stockées et utilisées en tant que fragments schématiques épistémiques, évidentiels ou évaluatifs (appelés e/e/e), en donnant une perspective sur la proposition finie avec laquelle ils sont associés. Elle ajoute que la partie de l'énoncé construite comme servant de cadre e/e/e est une question d'inférence et de négociation. Ce n'est donc en aucun cas nécessairement la subordinée qui construit le cadre. C'est aussi le point de vue de Langacker (2008), dans le cadre de la grammaire cognitive, qui figure parmi l'un des rares chercheurs à explicitement affirmer que bien des propositions subordinées ou secondaires, malgré leur fonction les marquant comme des éléments modificateurs ou spécifiques d'une autre entité, ont un contenu aussi important et focalisé que les propositions coordonnées ou indépendantes. C'est selon lui une conception trop hiérarchique du discours qui fait de la subordinée un complément alors qu'elle est souvent la proposition stable indispensable à la construction de l'énoncé et non un simple ajout plus ou moins superflu. Il prend l'exemple des énoncés tels que L et SC dans la séquence (7), tirée de notre corpus.

- | | | | |
|-----|---------|----|--|
| (7) | Alex | L | you know there's like announces |
| | | SC | that come on like # every five minutes |
| | | | [...] |
| | | L | you know they have the like cabin baggage restriction |
| | Rhianna | | mh |
| | Alex | SC | that says that all the stuff fits in the overhead # lockers |

Dans ces énoncés, les propositions “principales” dans L servent uniquement à cadrer et introduire l'information vraiment nouvelle. Un résultat cohérent et un discours équivalent est obtenu lorsqu'on omet les principales :

- (8) Announces come on like # every five minutes. All the stuff must fit in the overhead # lockers

Il apparaît que ces propositions ne sont que subordinées dans l'exemple (7) dans le sens où elles sont encodées en tant que participants dans les procès/états véhiculés dans les propositions enchâssantes.

Les constructions subordinées sont utiles pour examiner la production et la compréhension langagière en temps réel, et présentent des implications pour la modélisation et la représentation du discours. Pourtant, leur analyse se focalise souvent sur la parole uniquement. Alors que nombre de travaux ont étudié leur apport syntaxique (e.g. Lyttle, 1974) ou leur apport prosodique (e.g. Couper-Kuhlen,

1986), le développement d’outils et de procédés analytiques (e.g. Sloetjes and Wittenburg, 2008; Benzitoun et al., 2009; Bigi, 2012; Boersma and Weenink, 2013) facilite désormais une description des subordonnées en tant que phénomènes multimodaux.

Si aucune étude à notre connaissance ne s’est attachée à analyser les constructions subordonnées dans une perspective multimodale, la subordination prosodique a été abordée par certains travaux (e.g. Bolinger, 1984; Chafe, 1984; Local, 2007), avant que d’autres études sur la gestualité co-verbale n’exposent des moyens gestuels de subordination (Enfield, 2009; Streeck, 2009; Calbris, 2011), plus ou moins corrélés à la subordination verbale et/ou vocale.

1.6.3 La subordination prosodique

La subordination prosodique est essentiellement réalisée par l’intonation (Bolinger, 1984). Si Bolinger (1984) observe une certaine corrélation entre le marquage intonatif et la structuration syntaxique, il est aussi l’un des premiers à nuancer cette vision en rappelant que le système intonatif est indépendant du système syntaxique. La subordination prosodique peut être réalisée sans parler de subordination syntaxique. L’approche de Mertens (2008) montre également que “les unités prosodiques ne sont pas nécessairement coextensives aux unités syntaxiques traditionnelles” (p. 124) : les subordonnées syntaxiques sont autonomes prosodiquement lorsque les stratégies du locuteur décident d’en faire des unités sémantiques, des cadres ou des repères. L’intonation peut également lier deux entités indépendantes syntaxiquement pour les faire appartenir à un ensemble plus grand, et détermine si “l’unité informationnelle est achevée” (p. 125).

Au fil d’un paragraphe vocal, la hauteur intonative (i.e. F0, la fréquence fondamentale) décline progressivement. Les locuteurs peuvent jouer dans un discours sur la juxtaposition d’unités intonatives indépendantes et les faire contraster avec des unités emboîtées. Les unités intonatives simplement descendantes, qui suivent la déclinaison naturelle, apparaissent comme non-marquées intonativement (Morel et Rialland, 1992). Ce phénomène s’observe également entre les hauteurs des syllabes initiales et finales des paragraphes intonatifs (Wennerstrom, 2001)¹². Un ton abaissé

12. Le travail de Wennerstrom (2001) met en valeur la notion de “paraton”, i.e. le paragraphe intonatif décrit par Yule (1980) et repris dans les analyses de Couper-Kuhlen (1986a; 1986b). Des unités phonologiques plus grandes que les *tone-units* sont identifiables dans la parole uniquement sur des critères intonatifs, correspondant à des paragraphes conceptuels. Le paragraphe intonatif suit la ligne de déclinaison naturelle du discours. Alors qu’un paragraphe dont l’attaque se situe

par rapport à un ton haut précédent correspond par là à la relation neutre entre deux groupes prosodiques, i.e. l'abaissement (*downstep*; Wichmann, 2000; Wells, 2006). De même, un segment de discours montrant une syllabe finale basse (contour terminatif) n'emboîte pas le segment suivant, et est autonome par rapport à la suite. À l'inverse, un segment de discours caractérisé par une syllabe finale haute (contour continuatif) emboîte le constituant qui le suit et constitue un repère pour la suite (Morel, 1992; Martin, 2013). Wells (2006) mentionne également une montée intonative sur la syllabe finale en tant que marqueur de subordination, mais sur la syllabe finale du segment subordonné lui-même. Cette montée indique que ce segment préface d'autres segments à venir, qui continuent le paragraphe et le sujet en cours de traitement (Bolinger, 1984; Wells, 2006).

Par contraste, une variation de hauteur sur la syllabe initiale, autre que le *downstep*, signale une frontière (Bolinger, 1984). Les frontières dans la parole peuvent aussi être créées par des variations de tempo (Crystal, 1969; Priva, 2017). Les pauses silencieuses participent également à la segmentation du discours (Local, 1992), tout comme l'allongement syllabique final (Oller, 1973; Wightman et al., 1992; Cho, 2006; Auran et Loock, 2011).

1.6.4 La subordination gestuelle

En suivant la perspective des recherches récentes en gestualité, les gestes et la parole sont étroitement liés pendant la production à tous les niveaux (Iverson et Thelen, 1999; Bernardis et Gentilucci, 2006; Roustand, 2012; Ferré, 2014). Une hiérarchisation ou une interruption dans le processus de production de parole devraient donc avoir un effet sur les gestes, et vice-versa (Kita et al., 2007).

De nombreuses études (Cassel et McNeill, 1990; Gullberg, 2006; Eisenstein et al., 2008; Calbris, 2011; Perniss et Özyürek, 2015; Azar et al., 2016; Fredericksen, 2016; Sekine et Kita, 2017) ont montré que certains traits gestuels participent à la création et/ou au maintien de la cohérence et de la cohésion (Halliday et Hasan, 1976) dans la parole. De même, la représentation de référents par les gestes manuels est un processus cumulatif, souvent mené à bien à travers une série de plusieurs unités gestuelles (Streeck, 2009). La référence est maintenue par la cohésion dans l'espace, le choix de

dans le registre bas d'un locuteur suggère une incise, celui présentant une hauteur d'attaque dans le registre haut est associé à une frontière et, possiblement à une nouvelle orientation de topique discursif. Cette analyse confirme que la hauteur mélodique, à prendre en compte sur plusieurs éléments lexicaux, joue, comme tous les éléments du rythme, un rôle important au niveau de la structure du discours et de l'information.

la main utilisée pour réaliser les gestes, et/ou la forme du mouvement, dont le style (McNeill et Levy, 1993; Streeck, 2009). Deux segments de parole peuvent également être liés par leur production en co-occurrence avec une seule unité gestuelle (Enfield, 2009). De même, la tenue d'un geste (i.e. le maintien statique d'une ou des deux mains dans la configuration obtenue juste après la réalisation d'un geste, avant un changement de position pour la réalisation d'un autre geste ou un retour à la position de repos) permet entre autres d'ajuster le discours en temps réel en gardant la synchronie entre parole et geste (McNeill, 1992).

En montrant qu'un geste manuel peut être structurellement lié par sa forme et sa fonction à un geste du discours précédent ou au contraire que l'information peut être segmentée linéairement par les gestes, Enfield (2009, p. 210) remarque que trois techniques gestuelles sont particulièrement favorables à l'ajustement du discours, et particulièrement à la subordination. Le *hold-and-drag* consiste à tenir la configuration d'un geste manuel récemment réalisé (ancré à un référent particulier du discours) pour le déplacer à un autre endroit de l'espace gestuel. Le *hold-and-work-with-free-hand* décrit la tenue d'un geste d'une main alors que l'autre main travaille à l'extension ou à la refonte d'une partie du schéma (gestuel et référentiel). Enfin, le *move-body-into-new-space* représente un processus où le corps est orienté dans une nouvelle direction pendant que le schéma obtenu par la configuration manuelle, lui, ne bouge pas et reste des coordonnées spatiales absolues, recontextualisées par le nouvel espace.

Il en va de même pour la “construction symétrie-dominance” (p. 114), dans laquelle les locuteurs réalisent un geste symétrique à deux mains dans une première phase. Ils gardent ensuite une main tenant cette position, représentant l'information de la première phase mise en arrière-plan, alors que l'autre main exécute un nouveau geste, représentant l'information mise au premier plan (Fredericksen, 2016). Alors que la première phase construit un ancrage ou un cadre pour un apport d'information imminent à la fois dans le discours et dans l'espace, la dépendance syntaxique entre les deux phases de cette construction fonctionne simultanément dans ces deux domaines, garantissant la cohérence à la fois de l'organisation du discours et de la représentation de la structure spatiale. La main tenant le premier geste fonctionne comme un point de référence stable comme support sémiotique, en signalant que certaines informations continuent d'être pertinentes à ce qui est en train d'être énoncé.

Calbris (2011, p. 47) étudie comment les gestes liés à des unités verbales rendent une organisation visuelle hiérarchique à travers des schémas. Si le maintien d'une

forme gestuelle correspond au “maintien de l’idée qu’elle représente”, un changement concomitant dans la gestualité segmente le discours. “Le changement crée une rupture dans le continuum et opère une fonction démarcative, alors que le type de changement fournit des informations et opère une fonction référentielle” (p. 46). Les frontières dans le discours sont donc communément marquées par un changement de la main utilisée pour réaliser un geste, ou bien par un changement des coordonnées spatiales où sont réalisés les gestes (Streeck, 2009). De même, Enfield (2009) observe que la posture de repos pour les deux mains peut revêtir un aspect prosodique, en signifiant une frontière dans le discours, une “mise au repos” du topique en cours (p. 123) simplement par contraste après une séquence gestuelle.

Une frontière peut également être marquée par un changement de point de vue lors de la production de gestes iconiques mimétiques comme lors d’une narration ou d’une description (McNeill, 2005). Les gestes mimétiques peuvent être réalisés selon deux points de vue : celui de l’observateur (une partie du corps réalise le mouvement alors que le locuteur observe le geste) ou celui du personnage (l’intégralité du corps produit le geste). Le locuteur peut interrompre une séquence réalisée selon la perspective de l’observateur par un bref passage réalisé sous le point de vue du personnage.

En ce qui concerne les autres articulateurs, les locuteurs quittent souvent du regard les co-locuteurs pour l’élaboration du discours dès lors que le tour de parole est pris et garanti (Beattie, 1978; De Kok et Heylen, 2009). Un changement de direction du regard vers le co-participant annonce une frontière dans le discours (De Kok et Heylen, 2009; Streeck, 2014) ou un appel vers le co-locuteur (Goodwin, 1979; Knoeferle et Kreysa, 2012; Holler et al., 2014). Lorsqu’une séquence discursive est clôturée avec succès, ses participants détournent systématiquement le regard les uns des autres. En revanche, l’un des partis ou tous maintiennent le regard vers les autres lorsque la séquence nécessite une expansion (Rossano, 2013).

De même, les gestes de battement¹³ peuvent individualiser certaines entités (Cavé et al., 1996). Ces battements sont connectés à la structure discursive par leur fonction (Kendon, 1972; De Kok et Heylen, 2009). Les mouvements de sourcils, particulièrement les hausses, peuvent également segmenter le discours (Ekman et Friesen, 1969; Granström et House, 2005).

13. Les battements, manuels ou de tête, sont de rapides mouvements vers le bas, marquant la structuration rythmique du discours.

En résumé, la littérature syntaxique et discursive portant sur les subordonnées considère ces structures comme dépendantes de la structure hôte, bien que peu de consensus existe entre les chercheurs sur le label même de subordination, et sur le niveau d'intégration de ces constructions. Puisqu'un parallèle est établi entre structure syntaxique et structure discursive, les constructions subordonnées sous étude sont décrites comme intégrées à la structure discursive du co-texte, mais subsidiaires du point de vue de leur fonction. Aucun travail dans le cadre de la multimodalité n'a à notre connaissance été réalisé sur les subordonnées. Toutefois, les chercheurs en prosodie parlent d'emboîtement intonatif de certaines unités par d'autres. En ce qui concerne la gestualité, certaines unités discursives peuvent également être liées par une seule et même unité gestuelle. Certaines configurations gestuelles servent également à signaler la dépendance, par opposition au signal de la présence d'une frontière dans le discours.

1.7 Hypothèses

L'enjeu consiste à montrer que l'on ne peut pas rendre compte des procédés de regroupement des constructions subordonnées sous étude par une grammaire unique en postulant une opposition simple entre deux pôles (principale et subordonnée) et qu'il est nécessaire de poser des "grammaires multiples" pour rendre compte de la diversité des phénomènes. L'analyse des différents types de regroupements fait apparaître une multitude d'interactions entre syntaxe, morphologie, sémantique, et pragmatique, s'opposant à la vision réductrice d'intégration vs. non intégration qui présuppose une régularité parfaite entre les différents modules de la communication linguistique.

Notre étude ne s'attache pas à comparer les constructions subordonnées à des constructions non-subordonnées, mais s'attache plutôt à différencier chaque type syntaxique de subordonnée sous étude du point de vue de son intégration au co-texte. Il s'agira d'étudier les différents types d'autonomie des constructions sous étude, et de distinguer dans quels cas elles sont intégrées aux unités adjacentes, comme le décrit la littérature traditionnelle, et dans quels cas elles sont démarquées de leur co-texte.

Notre hypothèse de travail principale, motivée par les travaux antérieurs, est donc fondée sur la capacité des subordonnées à recouvrir une diversité de fonctionnements en s'appuyant sur plusieurs types d'autonomies. Cette hypothèse conductrice

structure notre travail d'analyse, et s'appuie sur une liste d'indices syntaxiques, discursifs, prosodiques, et gestuels pour rendre compte de différents types de frontières.

Si les subordonnées ne sont pas autonomes, elles devraient être régies syntaxiquement (i.e. elles devraient être sous la portée du prédicat principal, ce qui détermine leur position syntaxique), ou montrer une autonomie modale et/ou pragmatique limitée, puisque ces traits sont associés aux subordonnées (Van Valin, 1984). De même, la production des constructions subordonnées ne devrait pas être accompagnée de marqueurs de discours jouant un rôle de transition (Redeker, 2006). Ces subordonnées devraient également susciter peu de réactions du co-locuteur, que celles-ci soient verbales ou non (Bertrand et al., 2007; Bertrand et al., 2009).

Au niveau prosodique, on peut s'attendre à ce que les subordonnées enchâssées soient intégrées dans le même groupe intonatif que celui de la proposition principale (Bolinger, 1984), ou qu'elles présentent des contours continuatifs (i.e. montée intonative finale impliquant que le locuteur va garder son tour; Hirschberg et Grosz, 1992). Les subordonnées devraient être produites dans un registre bas ou moyen, la ligne de déclinaison habituelle du paragraphe intonatif ne devant pas être interrompue (Wennerstrom, 2001). Les constructions subordonnées ne devraient pas non plus causer de changement rythmique important, et comporter peu de pauses (Local, 1992).

En ce qui concerne la gestualité, les constructions dépendantes devraient être produites avec des moyens cohésifs comme l'usage de la même unité gestuelle que celle de l'environnement séquentiel (Enfield, 2009). Aucun battement ne devrait a priori être produit en co-occurrence avec une structure subordonnée. Si les subordonnées sont intégrées au co-texte, elles devraient être réalisées sans changement de direction du regard vers le co-locuteur (Beattie, 1979; De Kok et Heylen, 2009), et sans haussement de sourcils (Cavé et al., 1996).

Chapitre 2

Corpus et méthodologie

2.1 Enregistrement du corpus

Le corpus que nous utilisons pour cette étude est celui du corpus ENVID (Lelandais et Ferré, 2017), une collection de dialogues en anglais britannique. Ce corpus collaboratif inclut des enregistrements vidéo collectés par des chercheurs et des étudiants de Master de l'Université de Nantes entre 2000 et 2012. Nous avons personnellement collaboré à ce corpus en enregistrant le dialogue entre les locutrices Alex et Rhianna en 2012, d'une durée de 30 minutes.

Cinq dialogues ont été sélectionnés pour cette étude, totalisant deux heures et dix minutes d'interaction entre dix locuteurs. Ils sont composés des dyades suivantes : Alexandra et Rhianna (30 minutes), Michelle et Zoe (30 min), Tim et Tom (30 min), Elena et Joey (30 min), ainsi que Kate et Beth (20 minutes de tâche interactionnelle non utilisées + 10 min de conversation spontanée).

Chaque dialogue a été enregistré dans les studios d'enregistrement de différentes universités, afin d'en garantir le traitement prosodique. Les participants sont tous britanniques et sont âgés de 20 à 23 ans. Dans chaque paire, les participants sont amis ou se sont déjà rencontrés, et ont reçu pour seule consigne de discuter comme ils le font d'habitude. Ces choix visent à obtenir des interactions aussi naturelles que possible, notre étude portant sur les stratégies des locuteurs lors d'une conversation spontanée. Les participants ont été informés qu'ils participaient à une étude de linguistique¹, mais aucune mention n'a été faite de la prosodie ou des gestes. Ils ont été laissés seuls pendant les enregistrements afin de ne pas inhiber la conversation, déjà considérée influencée, du moins pendant les premières minutes, par le simple

1. Le formulaire de consentement utilisé pour notre contribution au corpus est disponible en annexe p. 347.

fait de savoir qu'ils étaient filmés (Labov, 1972).

Afin de rendre l'atmosphère plus conviviale et familière, certains studios sont aménagés en salons : Alexandra et Rhianna ainsi que Zoe et Michelle sont installées dans des fauteuils. Les autres participants sont assis autour d'une table. Certains locuteurs ont été autorisés à garder les gobelets de boisson avec lesquels ils sont arrivés : loin d'être gênante, la manipulation d'objets fait partie intégrante de l'étude de la gestualité.

Tous les participants sont filmés par un plan fixe avec un angle de vue assez large, face à leur interlocuteur ou tournés de trois-quarts vers ce dernier. Ils sont visibles au moins de la tête au buste, les caméras fixes placées assez près pour capturer les mouvements faciaux mais permettant une vue de l'ensemble du haut de leur corps. Pour deux dialogues (Alexandra et Rhianna, Kate et Beth), une caméra fixe est placée face aux deux participantes, tournées de trois-quarts l'une vers l'autre. Pour les trois autres dialogues, une caméra fait face à chaque participant.



FIGURE 2.1 – Schémas représentant (a) l'emplacement des locutrices et de la caméra pendant l'enregistrement (dialogues Alexandra/Rhianna et Kate/Beth) et (b) l'emplacement des locuteurs et des caméras pendant l'enregistrement (dialogues Michelle/Zoe, Tim/Tom, Joey/Elena).

La figure 2.1² montre ces deux types de configuration avec les schémas (a) et (b).

Chaque participant était équipé d'un micro-cravate, permettant deux pistes audio séparées pour chaque dialogue. Un fichier audio a été créé pour chaque locuteur au format WAV, afin de faciliter l'analyse des chevauchements de parole. Les enregistrements vidéo natifs ont été convertis en MP4 à une fréquence de 25 images par

2. L1 et L2 désignent les locuteurs (locuteur 1, locuteur 2), tandis que les figures grises représentent la ou les caméras.

seconde. Un fichier vidéo MP4 a donc été créé pour chaque dialogue, en juxtaposant les images des deux caméras pour ceux filmés avec une caméra face à chaque participant. Ce choix de matériel et de formats est considéré comme le meilleur compromis entre la nécessité d'un enregistrement de qualité, permettant des analyses prosodique et gestuelle détaillées, et le souhait de données aussi naturelles et authentiques que possible.

2.2 Transcription et annotation du corpus

Toute analyse repose sur une étape de segmentation ayant pour objectif d'identifier des unités de traitement (qui ne seront pas nécessairement conservées pour l'analyse). On segmente ainsi généralement le signal audio en unités inter-pausales (segments de parole entre des pauses de plus de 200 millisecondes) ou la transcription en tokens (unités lexicales pouvant être complexes). La segmentation peut concerner des domaines et des niveaux différents : segmentation en propositions, en tours de parole, etc. (Adamou et al., IRCOM 2014). La ponctuation peut ici jouer un rôle important à l'écrit. Cette notion est en quelque sorte transposée à l'oral, avec un balisage de la transcription (qu'il soit basé sur des critères prosodiques, syntaxiques ou pragmatiques). Nous nous intéressons dans cette étude à la segmentation du signal audio en groupes intonatifs et à la segmentation du signal image en unités gestuelles.

2.2.1 Le groupe intonatif en conversation spontanée

Les dialogues ont tous été transcrits orthographiquement manuellement sous Praat (Boersma et Weenink, 2013), la parole alignée avec le signal sonore au niveau du groupe intonatif³ (l'*Intonation Phrase* définie par Selkirk (1978)) selon la tradition britannique (Halliday, 1967; Crystal, 1969; Wells, 2006). Nous avons vu dans le Chapitre 1 que la tradition britannique se base sur des domaines et considère l'intonation en termes de contours intonatifs dynamiques. Cette unité prosodique correspond à ce que Pierrehumbert et Hirschberg (1990) appellent syntagme intermédiaire (*Intermediate Phrase*).

Les frontières des groupes intonatifs correspondent la plupart du temps aux frontières syntaxiques et prosodiques, la fin d'un groupe intonatif étant parfois marquée par une pause. En revanche, la segmentation de l'anglais spontané en

3. En anglais, *tone-unit*.

groupes intonatifs remet quelquefois en cause la définition traditionnelle des groupes intonatifs évoquée dans le Chapitre 1, qui doit être adaptée. Une conversation implique un certain nombre de répétitions, d’hésitations, et de reprises dues au travail de formulation ou aux bruits divers. Ces phénomènes linguistiques montrent une syntaxe différente de l’écrit, qui ne coïncide pas avec les pauses ou les silences. La prosodie doit alors servir de critère décisif pour segmenter le flux de la parole, avec des indices définis par Selkirk (1978) tels que les changements de rythme, ceux de registre intonatif, et les pauses. En conséquence, certains groupes intonatifs ne correspondent pas à la définition traditionnelle, comme dans l’exemple (1), adapté de notre corpus.

(1)	TU1	Alex	oh if you go to P aris
	TU2	Rhianna	for my
	TU3	Alex	where would that be
	TU4	Rhianna	well for my ma sters cos #
	TU5		hem
	TU6	Rhianna	i don’t really know what i want to do yet

Le groupe intonatif de Rhianna “For my” est considéré comme une unité même si le groupe ne contient pas de noyau (nucleus), puisqu’elle est en fait interrompue par Alex. Les interruptions créent des groupes intonatifs différents lorsqu’un locuteur commence un tour et que le co-locuteur commence un autre tour en l’interrompant. Ce genre de groupe intonatif est très fréquent dans le corpus.

Dans cet extrait (1), les groupes intonatifs d’Alex sont plus proches de la définition classique : ils comprennent tous les deux un noyau, et un changement intonatif a lieu entre le TU1 et le TU3 (“where” est plus bas que “Paris” et le ton continue de baisser, comme c’est le cas pendant les questions en “wh-”). Le “hem” de Rhianna est isolé du TU4 à cause de la pause (#) suivant “cos”. Il y a aussi un changement de hauteur intonative : “hem” est énoncé plus haut que “cos”. Les TU5 et TU6 auraient pu former un seul et même contour, mais “hem” forme un mouvement en lui-même. Le début du TU6 est plus haut, en corrélation avec un changement de rythme important : le TU6 montre un débit plus rapide.

2.2.2 Segmentation et étiquetage intonatif du signal

Une partie du corpus est également alignée au niveau des mots et des phonèmes. Cet alignement a été obtenu automatiquement avec le logiciel SPPAS (Bigi, 2012) et corrigé manuellement. La transcription phonémique utilise l’alphabet *SAMPA*

(Wells, 1997) au lieu de l'alphabet phonétique international (i.e. *IPA*) afin d'éviter l'incompatibilité entre divers outils d'annotation.

L'algorithme Momel-Intsint (Hirst, 2007; Bigi, 2012) a été utilisé pour l'annotation automatique des points cibles de F0 (fréquence fondamentale) dans le signal. Les annotations sont faites à deux égards : l'algorithme note la hauteur intonative (en Hz) sur des syllabes cibles, ce qui permet le calcul de valeurs moyennes de F0 pour des segments spécifiques. L'algorithme code aussi des valeurs (relatives) mélodiques symboliques, dans lesquelles chaque hauteur de F0 mesurée comparée aux précédentes, i.e. des changements significatifs dans la courbe de F0, soit par rapport à la plage intonative du locuteur (*Top*, *Bottom*), soit par rapport aux tons ou séquences de tons voisins (*Upstep*, *Downstep*, *Same*, *Low*, *High*). Dans cette étude, nous nous intéressons particulièrement aux valeurs qui indiquent une réinitialisation significative de la hauteur intonative (*Top*, *Bottom*) ou un changement significatif de registre intonatif (*Upstep*, changement vers un registre plus haut; *Downstep*, vers un registre plus bas). Nous nous intéressons également à la valeur *Same*, qui indiquerait, si trouvée en plus grand nombre dans nos séquences, un contour plat et une absence de rupture entre les différents constituants de la séquence.

Des annotations manuelles complémentaires ont été effectuées au niveau intonatif autour de nos échantillons, afin de transformer en données observables les données brutes fournies par les tracés mélodiques.

Le registre intonatif de chaque locuteur a été identifié sur une longue séquence de parole prélevée en cours d'échange. Les hauteurs de F0 minimale et maximale ont été repérées afin de déterminer, par une gradation des valeurs propres à chaque locuteur et par le calcul d'une moyenne, trois plages (haute, moyenne, et basse) pour l'interprétation globale des valeurs énonciatives des constituants. Dans chaque séquence contenant une construction subordonnée de notre échantillon, nous avons étiqueté ces différentes hauteurs de registre intonatif au segment entier et aux syllabes initiales et finales. La hauteur intonative des constituants initiaux des constructions subordonnées est alors mesurée en fonction du rehaussement ou de l'abaissement qu'elle provoque par rapport au ton final de l'unité précédente. Si notre annotation ne prend pas directement en compte la déclinaison des paragraphes, elle permet de cette manière l'interprétation des exemples à cet égard : les constructions sous étude sont contextualisées dans le registre du locuteur.

De la même façon, la nature du mouvement intonatif est étiquetée à chaque

syllabe nucléaire de la séquence sous étude, différenciant entre les mouvements descendant, descendant-montant, montant, montant-descendant, et plat.

Afin d'établir une fiabilité dans la classification des contours intonatifs nucléaires, un second codeur expert a jugé 20% des données annotées par nos soins. Le taux d'accord entre les annotateurs est de 81.9%.

2.2.3 Classification des gestes

Les gestes ont été transcrits à l'aide du logiciel ELAN (Sloetjes et Wittenburg, 2008) après que les pistes de transcription réalisées sous Praat y aient été importées, de façon à optimiser la visualisation de l'alignement parole-gestes. Cette transcription comprend la direction du regard, les mouvements de tête et de sourcils, ainsi que les gestes manuels, et suit les paramètres proposés par Bressem et Ladewig (2011).

Annotation descriptive

L'annotation descriptive est basée sur les unités gestuelles, comme décrites dans le Chapitre 1, plus particulièrement sur les phrases gestuelles (Kendon, 2004). Chaque unité gestuelle considérée ici, à la manière des phrases gestuelles, commence à l'ébauche du geste, c'est-à-dire avant l'image de la première excursion par rapport à la position de repos (onset) et finit juste après le retour à cette position de repos s'il y en a un. Dans le cas très fréquent de deux gestes réalisés à la suite, la première unité gestuelle finit à un changement significatif de forme et/ou de trajectoire. Une dimension ou fonction principale quant à la parole coproduite correspond à chaque unité gestuelle.

Sur une série de pistes différentes en fonction des articulateurs, les mouvements de tête sont étiquetés en tant que *nods* (acquiescements), *shakes* (gestes de négation), *tilts* (inclinaisons de la tête sur le côté), *beats* (mouvements du menton vers le bas sans valeur d'acquiescement) et *jerks* (rejets de la tête vers l'arrière). La figure 2.2 ci-dessous représente l'un de ces gestes de tête. Elena commence par secouer la tête sur un axe horizontal vers la droite (image 2.2a), puis vers la gauche (image 2.2b), et revient à droite, en corrélation avec un haussement de sourcils (image 2.2c).

La direction du regard est annotée soit vers le co-participant soit ailleurs, auquel cas figurent les coordonnées directionnelles (*higher* (plus haut)/*lower* (plus bas)).



FIGURE 2.2 – Elena réalisant un geste de négation (shake) en corrélation avec le groupe intonatif “no it’s been fine”.

Les mouvements de sourcils distinguent entre les deux positions *rising* (sourcils haussés) et *frowning* (sourcils froncés). La figure 2.3 montre un exemple de sourcils haussés, fonctionnant comme partie d’un *backchannel* (réaction de Zoe en tant que co-locutrice au discours de Michelle, l’encourageant à poursuivre son tour). Zoe a également le regard dirigé sur Michelle.

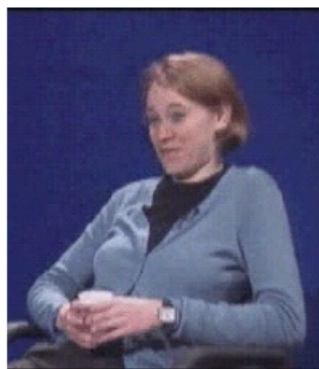


FIGURE 2.3 – Sourcils haussés de Zoe pendant la production d’un *backchannel* (réaction) au discours de Michelle.

Les gestes manuels, plus complexes, sont transcrits sur deux pistes différentes pour chaque participant : si une première piste est consacrée à la description formelle du geste (forme et position de la main, geste réalisé à une ou deux mains, type de mouvement et direction), la seconde piste indique la relation fonctionnelle du geste à la parole, en suivant la typologie établie par McNeill (1992; 2005) et en considérant les liens entre gestes et affiliés lexicaux (Kipp et al., 2007). Elle relève les fonctions suivantes :

- emblèmes : gestes conventionnels
- iconiques : “images d’entités ou d’actions concrètes”
- métaphoriques : “images de l’abstrait”, impliquant un usage métaphorique des formes ou de l’espace (McNeill, 2005, p. 39)
- battements : gestes liés au rythme de la parole (McNeill, 1992, p. 80)
- pointages : gestes déictiques⁴
- butterworths : gestes désorganisés réalisés lors de recherches lexicales
- adaptateurs : gestes d’auto-contact réalisés généralement pour des raisons liées au confort des participants, mais inclus à l’étude en raison de leur nombre important et du fait qu’ils fournissent des informations aux annotateurs sur l’organisation des tours, car ces gestes sont plus fréquents quand les participants écoutent.

Lorsqu’un geste possède deux dimensions, ce qui est fréquent, la dimension principale ou dominante est notée dans cette piste et la dimension secondaire est inscrite sur une piste séparée.

Afin d’établir une fiabilité dans la classification des types gestuels, un second co-deur expert a jugé 20% des données annotées par nos soins. Le taux d’accord entre les annotateurs est de 100% pour la direction du regard, 96.4% pour les mouvements de sourcils, 81.3% pour les mouvements de tête, et 72.1% pour les gestes manuels.

Catégorisation des fonctions gestuelles

Nous nous sommes ensuite intéressée aux différentes fonctions que permet chaque type sémiotique de geste manuel décrit dans la littérature, et les avons intégrés dans une liste de sélection afin de déterminer si les constructions sous étude s’appuient sur une fonction préférentielle parmi les autres. Les gestes manuels sont ici classés selon deux grandes fonctions : un rôle référentiel (décrire ou représenter des objets, des actions, des idées) et un rôle non-référentiel⁵ (véhiculer de l’information abstraite à

4. Si le terme “pointage” correspond au type sémiotique du geste, la fonction linguistique de celui-ci est désignée par le terme “déictique”.

5. La classe non-référentielle distingue des informations à propos de différents composants : grammaire, modalité, discours/séquence, et interaction. Elle correspond à ce que Kendon (1995) appelle les gestes pragmatiques (*pragmatic gestures*, qu’il définit comme “toutes les façons dont lesquelles les gestes peuvent être associés à des traits significatifs d’un énoncé qui ne font pas partie de sa signification référentielle ou de son contenu propositionnel” (notre traduction, Kendon, 2004, p. 158)). Nous n’appelons pas cette classe “gestes pragmatiques” car à notre sens, tout ce qui n’est pas référentiel n’est pas pragmatique. Les gestes ayant un aspect grammatical et/ou syntaxique

Type/Valeur	référentiel	non-référentiel
iconique	X	X
métaphorique	X	X
pointage	X	X
battement		X
emblème	X	
adaptateur		X
butterworth		X

TABLE 2.1 – Récapitulatif des types et valeurs/fonctions principales des gestes annotés.

propos de l’organisation du discours et/ou de l’interaction).

Comme les gestes manuels peuvent avoir plusieurs dimensions (McNeill, 2005), deux valeurs sont parfois notées dans la même étiquette et ont été comptées toutes les deux le cas échéant; l’une dominante, l’autre secondaire.

Le tableau 2.1 ci-dessous présente une récapitulation des gestes manuels annotés, montrant que les différents types gestuels peuvent appartenir à la même valeur, et qu’une même valeur peut à l’inverse être obtenue par différents types de gestes.

La série de figures suivante présente des exemples de différents types de gestes, associés à leur fonction en contexte. Elles comportent un geste emblématique (figure 2.4), trois gestes iconiques (figure 2.5), figure 2.6), deux pointages (figure 2.7, figure 2.8), deux métaphoriques (figure 2.9), un battement (figure 2.10), et un adaptateur (figure 2.11).

La signification de cet emblème est indépendante de la parole dans la mesure où il peut être produit et compris sans être complété par aucune parole. Il complète d’ailleurs le discours verbal d’Alex au moment de la séquence montrée dans la figure 2.4, énumérant les raisons pour lesquelles elle ne veut pas conduire en France. En expliquant que les ronds-points sont inversés par rapport à l’Angleterre (“yeah but it’s on the other side and #”), elle interrompt son discours verbal et produit cet emblème, qu’elle accompagne d’un éclat de rire. Il remplit une fonction référentielle dans le discours, dans la mesure où sa forme et sa signification sont conventionnelles, fixées culturellement pour symboliser le concept de NÉGATION. Il permet également de renforcer l’assertion d’Alex, qui vient apporter une contradiction à une intervention précédente de Rhianna avec le “but” adversatif.

sont donc inclus de la même façon dans notre appellation.



FIGURE 2.4 – Geste emblématique de NÉGATION : index levé, le poignet oscillant de gauche à droite.

Le geste iconique sur la figure 2.5 est utilisé pour décrire la taille d'un référent verbal : il est descriptif. Alex prend un point de vue d'observatrice pour réaliser ces gestes : une partie de son corps est mobilisée pour la construction des gestes alors que le reste est observateur.



luggage restriction

FIGURE 2.5 – Geste iconique affilié à un élément lexical spécifique.

Le geste sur la figure 2.5 est purement référentiel, donnant des informations plus précises sur le référent.

Les iconiques peuvent également être mimétiques, comme dans 2.6a, b, et c ou Michelle raconte à Zoe la fois parmi ses allers-retours entre la France et l'Angleterre où son train est resté coincé dans le tunnel sous la Manche. Elle produit un premier

geste iconique représenté dans (a) et (b) où elle tire plusieurs fois sur son écharpe en mimant les voyageurs autour d'elle qui commencent à devenir anxieux, idée verbalisée par l'interprétation de ce geste par Zoe, le groupe adjectival “a little bit stressed”, beaucoup moins imagé. C'est cette fois-ci le point de vue du personnage qui est pris : Michelle utilise tout son corps pour réaliser le geste et se mettre ainsi à la place des passagers du train, agrippant ensuite les accoudoirs de son siège pour accompagner la description verbale “people getting sort of white knuckles, gripping the sides of their seats” sur (c).



FIGURE 2.6 – Deux gestes iconiques mimétiques.



FIGURE 2.7 – Pointage référentiel abstrait.

Dans la figure 2.7, Alex pointe du doigt vers un référent correspondant à un objet concret; cependant cet objet n'est pas présent physiquement à l'endroit et au moment où la conversation se déroule : ce geste désigne les coffres à bagage en cabine d'un avion. Ce geste fournit des coordonnées spatiales à un référent situé dans l'avion, établi comme un environnement métaphorique de la conversation. C'est donc

un geste abstrait dans le sens où, bien qu’Alex pointe spécifiquement (i.e. une entité spécifique peut être assignée au geste et les “cabin overhead lockers” décrivent un objet pouvant être trouvé dans le monde extérieur), cet objet n’est pas présent physiquement au moment où le geste est produit. Bien que les coffres à bagages soient réels, ceux indiqués ici sont une construction du discours : le monde réel est filtré par une situation discursive créée à partir de représentations mentales. Un pointage référentiel est donc vu comme abstrait si l’entité montrée n’est pas présente dans l’environnement physique contextuel. Les pointages en relation à l’environnement mis en place par la parole (mots, segments discursifs, interaction) sont en ce sens encore plus abstraits, comme les quelques pointages sans référent identifiable. Ces derniers sont en revanche non-référentiels (i.e. ils organisent plutôt un composant du discours ou de l’interaction). À l’inverse, un geste de pointage concret montre une entité directement sélectionnée du “monde extérieur”, comme une personne ou un objet physiquement présent (figure 2.8). Le contexte dans lequel le pointage (qu’il soit abstrait ou concret) est produit permet de définir celui-ci comme un pointage référentiel ou non-référentiel.



FIGURE 2.8 – Pointage concret de Beth (à droite) montrant à Kate (à gauche) où elle doit placer l’une des formes du jeu.

Le geste manuel métaphorique dans la figure 2.9a exprime un désengagement de Rhianna sur son assertion, comme si elle mettait physiquement une barrière entre elle-même et son discours. Au cours d’une argumentation dans laquelle l’avion est décrit comme un moyen de transport angoissant, elle avance que Ryanair ne respecte pas la législation européenne à propos des normes de réserve de carburant; elle ajoute néanmoins “they got caught so maybe they do now”. Le geste métaphorique renforce à la fois le manque d’engagement épistémique exprimé verbalement par “maybe” et indexe ce segment discursif additionnel comme une concession nuançant, sinon stop-

pant son argumentaire.

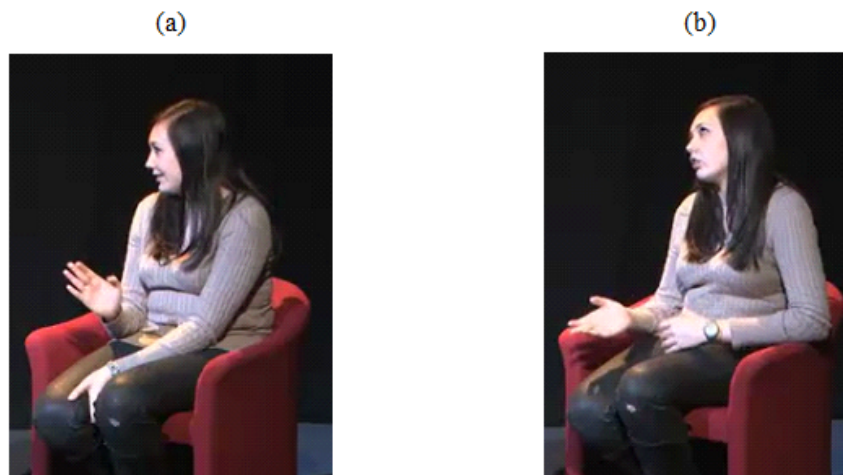


FIGURE 2.9 – Deux gestes métaphoriques apportant respectivement de l’information modale et interactionnelle.

Dans un autre contexte, le métaphorique produit sur (b) dans la figure 2.9 inscrit un corrélat verbal (i.e. “London”) dans les connaissances partagées entre Alex et Rhianna. La paume ouverte de Rhianna est dirigée vers l’espace conversationnel. C’est un appel vers la co-locutrice fonctionnant comme un rappel puisque Londres a déjà été évoqué plus tôt dans la conversation. Par ailleurs, ce geste marque cette partie du discours réintroduisant Londres comme concessive, faisant partie d’une modification libre de sa propre ligne d’argumentation (i.e. les parisiens pensent que toutes les autres villes sont petites) : “well actually I didn’t have that thing with London”. Ces deux gestes remplissent donc une fonction non-référentielle (i.e. ici, organisationnelle) dans le discours de Rhianna.

La figure 2.10 montre un battement de la main produit par Rhianna, correspondant à une recherche lexicale lors d’une séquence explicative à propos de la formation suivie par les hôtesses de l’air. Rhianna lève d’abord la main droite sur (a), sa main et ses doigts en position verticale, avant de descendre rapidement sur (b) lorsque Rhianna retrouve l’objet de sa recherche (i.e. récupération des informations/des termes recherchés), la main droite s’arrêtant avec un léger rebond près de ses genoux (c). L’activité gestuelle illustrée par les images ci-dessus est représentée entre crochets dans l’exemple (2).

(2) Alex wha- don’t you have to do like a special training for that

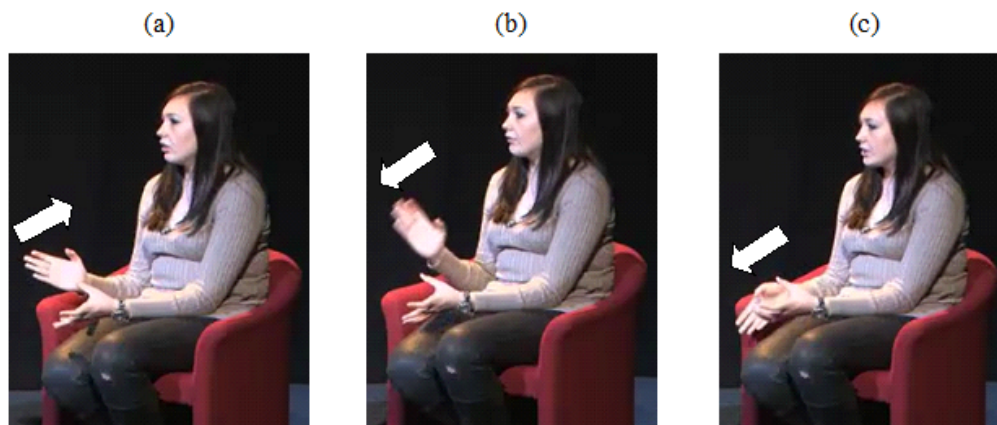


FIGURE 2.10 – Battement de la main associé à une récupération lexicale dans le discours verbal de Rhianna.

Rhianna (h) # yeah but it's only three days and
 Alex really
 Rhianna for students
 (h) they will do like student jobs for the summer #
 in this case
 (h) i think it's
 [(h) **now you have a medical visit**]
 in # between three and six days of training #
 and then you can fly for the summer #

Marquant la structuration rythmique de l'énoncé, ce geste permet à Rhianna d'organiser son discours en indiquant (et peut-être même provoquant) la fin du processus de recherche, initié vocalement par des pauses d'hésitation placées après "case" et "it's", ainsi que verbalement avec "I think".



FIGURE 2.11 – Moment sélectionné lors d'un adaptateur d'auto-contact produit par Tom.

Dans la figure 2.11 Tim se frotte l'espace au-dessus de l'œil en position de locuteur, continuant une séquence débutée par Tom (son co-locuteur) à propos de la politique de George W. Bush. Cet adaptateur fonctionne en tant que marqueur d'abandon de tour de parole (jouant un rôle dans l'organisation de l'interaction) alors que les pauses silencieuses et les intonations de *trail-off* (Local, 1992) des deux participants suggèrent que le topique conversationnel a été assez exploité.

Visibilité des annotations

Toutes les annotations ont été rédigées en anglais pour une plus large diffusion des résultats et, à plus long terme, une plus grande diffusion des données. La figure 2.12 donne un aperçu des annotations gestuelles dans le logiciel ELAN (Sloetjes et Wittenburg, 2008), alignées temporellement sur la vidéo.

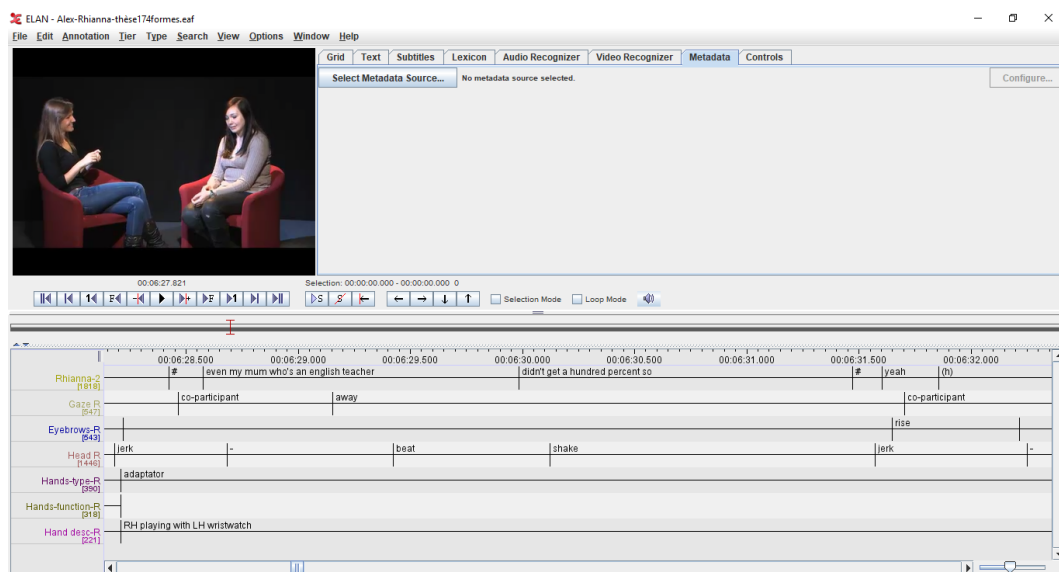


FIGURE 2.12 – Capture d'écran du logiciel ELAN, montrant la segmentation du discours de Rhianna en groupes intonatifs sur la première piste d'annotation, la direction du regard de Rhianna sur la seconde, ses mouvements de sourcils sur la troisième, ainsi que ses mouvements de tête sur la quatrième. Les cinquième, sixième, et dernière pistes sont consacrées respectivement au type de geste manuel produit par Rhianna, à la fonction de ce geste manuel, et à sa description.

2.3 Traitement et annotation des constructions subordonnées

Les constructions subordonnées, identifiées comme des phénomènes langagiers, ont d’abord été identifiées et notées sur la transcription orthographique de chaque interaction, avant d’être annotées sous Praat.

2.3.1 Localisation, nombre, et distribution des constructions subordonnées dans le corpus

Un total de 303 constructions a d’abord été localisé dans la totalité du corpus : 59 appositives (37 appositives relatives et 22 appositives canoniques), 83 relatives déterminatives, et 161 circonstancielles (89 circonstancielles de temps et 72 circonstancielles conditionnelles).

Nous avons calculé la distribution des subordonnées en fonction du temps d’élocution⁶ du locuteur (tableau 2.2), de celui de l’interaction, et celui du corpus entier (tableau 2.3). Les subordonnées, tous types syntaxiques confondus, représentent 10.09% du temps d’élocution du corpus (correspondant à un débit de 2.68 constructions/min) comme le montre le tableau 2.3.

Alex	6.61	Rhianna	7.36
Joey	5.78	Elena	2.68
Tom	9.29	Tim	4.92
Michelle	7.26	Zoe	4.47
Kate	4.33	Beth	3.39

TABLE 2.2 – Distribution (%) des constructions subordonnées tous types confondus par locuteur dans le corpus.

C’est le locuteur Tom qui montre la plus grande distribution de constructions subordonnées (9.29% du temps d’élocution de ce locuteur), avec 64 occurrences. À l’inverse, Elena montre la plus petite distribution (2.68% du temps d’élocution de cette locutrice). Elena ne produit que 12 constructions subordonnées.

Au niveau des interactions, le dialogue entre Tim et Tom comprend la part la plus importante de constructions subordonnées (7.72% du temps d’élocution dans

6. Le temps d’élocution d’un locuteur correspond au temps de la totalité du discours de celui-ci, dont sont exclues les pauses silencieuses.

Corpus total	10.09
Tim & Tom	7.72
Alex & Rhianna	7
Michelle & Zoe	5.71
Joey & Elena	4.56
Kate & Beth	4.03

TABLE 2.3 – Distribution (%) des constructions subordonnées tous types confondus par interaction dans le corpus.

l'interaction).

La série de tableaux suivants (tableau 2.4, tableau 2.5, et tableau 2.6) présente respectivement la distribution détaillée des circonstancielles, des relatives déterminatives, et des appositives.

temps circonstancielles/temps d'élocution du locuteur et des interactions				
		Locuteurs		Interaction
Alex	6.1	Rhianna	4.53	5.25
Tom	6.15	Tim	1.41	4.44
Kate	3.3	Beth	1.4	2.69
Joey	5.3	Elena	0.83	2.5
Michelle	1.55	Zoe	1.93	1.76

TABLE 2.4 – Distribution (%) des circonstancielles par locuteur, interaction, et dans le corpus.

Les circonstancielles correspondent au type syntaxique de subordonnée le plus répandu dans le corpus (3.46% du temps d'élocution total). Le tableau 2.4 montre que Tom, en tant que locuteur, consacre le plus de temps d'élocution à des circonstancielles, avec 6.15% de son temps d'élocution. A l'inverse, Elena leur consacre le moins de temps d'élocution, puisqu'elles représentent 0.83% de son temps d'élocution. En fonction des interactions, c'est en revanche dans le dialogue entre Alex et Rhianna que les circonstancielles sont les plus présentes avec 5.25% du temps d'élocution dans l'interaction.

Les relatives déterminatives représentent 1.88% du temps d'élocution total dans le corpus et forment le second type de construction subordonnée le plus répandu. C'est Michelle qui montre la plus grande distribution de relatives déterminatives, avec 4.65% de son temps d'élocution. À l'inverse, Kate consacre le moins de temps aux relatives déterminatives, avec 0.37% de son temps d'élocution. Au sein des interactions, c'est la dyade de Michelle et Zoe qui consacre le plus de temps aux relatives

temps relatives déterminatives/temps d'élocution du locuteur et des interactions				
		Locuteurs		Interaction
Michelle	4.65	Zoe	2.08	3.22
Tom	3.29	Tim	1.52	2.66
Joey	1.41	Elena	0.79	1.25
Alex	0.62	Rhianna	1.33	1
Kate	0.37	Beth	1.3	0.84

TABLE 2.5 – Distribution (%) des relatives déterminatives par locuteur et interaction.

déterminatives (3.22% du temps d'élocution).

temps appositives/temps d'élocution du locuteur et des interactions				
		Locuteurs		Interaction
Tom	1.84	Tim	1.52	1.81
Alex	0.78	Rhianna	1.99	1.43
Joey	1.28	Elena	0.99	1.17
Kate	0.62	Beth	0.78	0.67
Michelle	0.68	Zoe	0.54	0.56

TABLE 2.6 – Distribution (%) des appositives par locuteur et interaction.

Les appositives représentent 1.23% du temps d'élocution du corpus entier. Ce sont les appositives de Rhianna qui montrent la distribution la plus élevée (1.99% de son temps d'élocution total). Zoe, à l'inverse, en produit le moins (0.54% de son temps d'élocution). Les appositives sont les plus présentes dans l'interaction entre Tim et Tom, constituant 1.81% du temps d'élocution total du dialogue.

On observe à travers cette série de tableaux que le plus grand écart de distribution des constructions subordonnées est celui entre les circonstancielles et les appositives.

Nous avons décidé de restreindre notre étude à un nombre limité de constructions subordonnées, et avons sélectionné 55 occurrences de chaque type (circonstancielles, relative déterminatives, appositives), considérées comme des échantillons prototypes. Notre étude porte donc sur un total de 165 constructions sélectionnées. La liste de ces constructions subordonnées est disponible en annexe p. 355. Notre choix s'est porté sur les constructions encadrées par un co-texte immédiat (le groupe intonatif intégrant la construction subordonnée est précédé et suivi par un autre groupe intonatif du locuteur au lieu d'un silence donnant lieu à un changement de tour de parole). Nous nous sommes également assurée que la sélection de constructions syntaxiques soit équilibrée entre les locuteurs. Cet aspect est notamment important pour

Locuteurs	AP	C	RD	TOTAL
Tom	11	11	12	34
Rhianna	8	9	6	23
Joey	9	5	9	23
Alex	4	14	2	20
Zoe	4	6	7	17
Michelle	3	1	11	15
Tim	9	3	3	15
Elena	4	2	2	8
Kate	2	3	2	7
Beth	1	1	1	3
TOTAL	55	55	55	165

TABLE 2.7 – Récapitulatif du nombre de constructions subordonnées par locuteur.

l’analyse de la gestualité, afin d’éviter tout biais lié à la variabilité inter-locuteur. De même, notre sélection porte autant que possible sur les constructions appartenant à des séquences discursives impliquant une production de gestes. Le tableau 2.7 offre un récapitulatif du nombre de constructions subordonnées par locuteur (le terme AP désigne les appositives, C désigne les circonstanciées, et RD désigne les relatives déterminatives).

2.3.2 Annotation des constructions subordonnées dans le corpus

Une première piste d’annotation dans Praat (permettant de coder des informations définies par l’annotateur, alignées temporellement sur le signal audio; Boersma et Weenink, 2013) précise le type syntaxique de construction subordonnée et distingue donc entre appositive (3), circonstancielle (4), et relative déterminative (5) dans le fichier d’annotation (i.e. textgrid) de chaque locuteur. Si cette première piste d’annotation dans Praat renseigne le type de construction subordonnée repérée, une seconde piste prend en compte le contexte immédiat de cette construction : le groupe intonatif du co-texte précédant immédiatement la construction repérée est annoté L. SC correspond à la construction subordonnée, et R marque le groupe intonatif suivant immédiatement la subordonnée sous étude. Chaque type syntaxique de construction est illustré dans la série d’exemples ci-dessous. Dans chaque exemple, L correspond au groupe intonatif (ou partie de groupe intonatif) précédant immédiatement la constructions subordonnée, SC désigne la construction subordonnée (nous utilisons la notation de l’anglais pour *subordinate construction*, et R caractérise le groupe intonatif lui faisant suite.

- (3) Michelle L i never see any of my French people
 SC **my French hem flatmates**
 R they never talk to me you know

Dans l'appositive représentée par SC dans l'exemple (3), Zoe précise le référent de L “my French people” avec un autre groupe nominal.

- (4) Tim L although
 SC **when i passed**
 R i didn't think i was a very good driver

Dans la circonstancielle représentée dans l'exemple (4), Tim précise les circonstances pour lesquelles le prédicat de R est valide.

- (5) Tom L he ended up having lessons from everyone
 SC **he knew**
 R you know you can # give lessons
 if you're over twenty-one

La relative déterminative représentée par SC dans l'exemple (5) est nécessaire à l'identification du référent, renvoyant non pas à “everyone” mais à “everyone \emptyset he knew”.

Les séquences contenant une construction subordonnée sont donc annotées “L-SC-R”, et comprennent le plus souvent trois groupes intonatifs. Ces étiquettes désignent des groupes intonatifs comprenant des entités syntaxiques diverses selon la position des constructions subordonnées dans la macro-séquence. Les exemples qui suivent permettent de les présenter.

- | | | | | |
|-----|-----|----|--------------------------------|---|
| (6) | Tom | L | but some of them # | proposition noyau amorcée |
| | | SC | a couple of my friends | insertion de groupe nominal en préfixe |
| | | R | just won't drive on a motorway | proposition noyau de L reprise et continuée |

Dans la séquence (6) contenant une appositive positionnée en préfixe à la structure macro-syntaxique, L n'est qu'un constituant de proposition, SC est un simple groupe nominal apposé, et R désigne la fin de la proposition commencée sur L, contenant le verbe principal (“drive”).

- | | | | | |
|-----|------|----|--------------------------------|--|
| (7) | Joey | L | you know Jamie's brother | proposition noyau |
| | | SC | Tom # | insertion de groupe nominal en suffixe |
| | | R | he's had like a thing with her | autre proposition |

L	SC	R
proposition 1	proposition 2	proposition 3
constituant de proposition 1	groupe nominal apposé	fin de proposition 1
proposition 1	groupe nominal apposé	proposition 2
constituant de proposition 1	proposition 2	fin de proposition 1

TABLE 2.8 – Récapitulatif des entités syntaxiques couvertes par les étiquettes selon la position des constructions subordonnées dans la macro-séquence.

Néanmoins, dans la séquence (7), contenant une appositive positionnée en suffixe, L est cette fois-ci une proposition entière contenant le verbe principal (“know”), SC est un simple groupe nominal apposé, et R est une autre proposition.

(8)	Alex	L	but everyone else	proposition amorcée
		SC	i’ve spoken to	proposition insérée en préfixe
		R	is having trouble finding something #	proposition noyau de L reprise et continuée

A l’intérieur de la séquence (8), contenant une relative déterminative positionnée en préfixe, L désigne un constituant de la proposition qui sera reprise et continuée en R, qui contient le prédicat principal (“is having trouble”).

(9)	Zoe	L	do you remember that first night	proposition noyau
		SC	when everyone turned up	proposition insérée en suffixe
		R	and i # insulted what’s his name Freddy #	autre proposition

A l’inverse, dans la séquence (9), contenant une relative déterminative positionnée en suffixe, placée après le verbe principal (“remember”), L et R sont des propositions distinctes.

Le tableau 2.8 offre un récapitulatif des entités syntaxiques couvertes par les étiquettes selon la position des constructions subordonnées dans la macro-séquence.

Afin d’établir une fiabilité dans la classification des types syntaxiques de subordonnées (appositive, circonstancielle, relative déterminative), un second codeur expert a jugé 20% des données annotées par nos soins. Le taux d’accord entre les annotateurs est de 100%.

2.4 Mise en relation des différentes annotations

Bien que toutes ces annotations soient réalisées à partir de plusieurs niveaux (syntaxique, discursif, prosodique, gestuel), elles peuvent être mises en relation grâce

à leur alignement temporel (i.e. la correspondance temporelle entre les ressources vidéo, audio, la transcription et les annotations) dans le logiciel ELAN. Cet alignement des données permet de spécifier une relation entre chaque type de phénomène, chaque piste étant superposée aux autres dans le logiciel.

2.5 Nature des différentes analyses

Notre approche analytique envers les données récoltées témoigne à la fois de démarches quantitatives et qualitatives, toutes deux nécessaires et complémentaires dans le traitement de données multimodales venant de la conversation spontanée.

En ce qui concerne les aspects quantitatifs de notre méthodologie, les données sont exportées directement dans des fichiers Excel et triées. Nous évaluons l'autonomie verbale, vocale, et gestuelle de trois types de constructions subordonnées. Nous testons si ces constructions diffèrent dans leur nombre de facteurs verbaux (Chapitre 3), prosodiques (Chapitre 4), et gestuels (Chapitre 6) d'indépendance. De plus, nous comparons les résultats de production ainsi obtenus à des résultats de perception : nous testons si les constructions diffèrent dans leur nombre de facteurs d'indépendance perçus (Chapitre 5 et Chapitre 7). Pour répondre à nos questions de recherche, nous utilisons une série de modèles linéaires mixtes généralisés (*GLMMs* estimés par la méthode du maximum de vraisemblance) avec le programme R 3.4.0 (R Core Team, 2017) et le *package* *lme4* (Bates et al., 2017), afin d'expliquer le degré d'influence de chaque facteur d'indépendance sur les constructions subordonnées (cf. Chapitres 5, 7). Du fait de la variation entre les locuteurs et les dialogues dans la production des constructions subordonnées, nous incluons systématiquement le locuteur (Speaker) et le dialogue (Dialogue) en tant que facteurs aléatoires dans les modèles. Nous utilisons également le F-test de Fisher dans les Chapitres 3, 4, et 6, puisque le nombre et la nature des données requièrent l'utilisation d'un test non-paramétrique plus robuste.

L'analyse quantitative tend à s'assurer du fait que les tendances observées dans les résultats soient généralisables ou comparables à d'autres données. À l'appui de modélisations statistiques, l'analyse quantitative permet de vérifier si les résultats obtenus reflètent d'authentiques caractéristiques linguistiques, et que les tendances observées ne sont pas dues à d'autres facteurs ou au hasard. Les résultats des analyses quantitatives sont souvent dits plus objectifs que ceux des analyses qualitatives, car ils sont (en théorie) répliquables, à partir d'hypothèses réfutables (Debras, 2018).

En revanche, pour des raisons statistiques, les analyses quantitatives s'appuient sur des classifications souvent rigides ou très étroites. Or, nombre de phénomènes linguistiques montrent des aspects analysables au prisme de différentes catégorisations. La richesse des données peut alors être occultée par la nécessité de séparer différents jeux de données en de grands ensembles d'occurrences.

L'analyse qualitative vise au contraire à identifier des usages en contexte à travers la description fine et complète d'extraits du corpus. Notre étude comporte en ce sens un nombre important d'extraits authentiques, dont certains phénomènes syntaxiques, discursifs, prosodiques et/ou gestuels font l'objet d'une description détaillée à la lumière des différents co-textes et contextes. Ces descriptions permettent des distinctions fines et des nuances dans la catégorisation de nos données, indépendamment du fait que ces descriptions soient généralisables à un grand nombre d'occurrences, ou du fait qu'elles représentent une tendance significative dans les données. Les Chapitres 3, 4, et 6 comportent par exemple nombre d'analyses qualitatives d'occurrences. Ces occurrences sont présentées par des statistiques descriptives (i.e. fréquence des phénomènes dans le corpus). Puisque les extraits analysés sont choisis par l'analyste, cette approche fait face à un reproche majeur, notamment par les tenants des méthodes quantitatives : les cas les plus problématiques pouvant remettre en question la ligne d'analyse choisie peuvent aisément être mis de côté (Debras, 2018). Les études qualitatives permettent toutefois indiscutablement de caractériser la variété de fonctions ou de traits multimodaux d'une construction à travers divers contextes d'usage. De même, les méthodes qualitatives font preuve d'une plus grande validité écologique (i.e. elles tentent de s'approcher au plus près de la complexité des pratiques humaines en contexte) que les méthodes quantitatives, qui visent avant tout la reproductibilité des résultats et s'appuient sur des modèles simplifiés des phénomènes observés.

Deuxième partie

Chapitre 3

Frontières verbales

3.1 Mesure des frontières verbales

Nous nous intéressons dans ce chapitre à la relation des constructions subordonnées à leur entourage immédiat d'un point de vue verbal; à savoir si la majorité des constructions est concernée par la présence d'une rupture syntaxique et/ou discursive, ou si les subordonnées sont préférentiellement intégrées au co-texte gauche et/ou droit.

Nous évaluons l'autonomie verbale des constructions subordonnées, et nous testons si nos trois types syntaxiques de constructions sont différents dans leur nombre d'indices verbaux d'indépendance. Notre relevé d'indices a pour but de fournir une vision nuancée du degré d'insertion des constructions subordonnées dans la chaîne verbale, tant au niveau syntaxique qu'au niveau discursif. Or, certains travaux d'analyse sur l'oral remarquent quelques ambiguïtés lors de la différenciation de ces deux niveaux d'organisation (Corminbœuf, 2014). Une courte vue d'ensemble de nos choix d'analyse syntaxique est par conséquent proposée avant la description des résultats.

Si la littérature établit la rection grammaticale¹ comme le moyen de mise en dépendance syntaxique le plus puissant (Van Valin et La Polla, 1997), nous avons choisi de porter également attention au statut sémantique des syntagmes verbaux (e.g. verbe de procès, verbe d'état, verbe cognitif) et à la présence ou non de marques de liaison telles que les pronoms relatifs ou les conjonctions. Nous expliquons brièvement les principaux ajustements dérivant de l'analyse des données afin de contextualiser les résultats.

1. La rection désigne la configuration grammaticale dans laquelle le syntagme verbal principal détermine la position syntaxique d'un autre syntagme et lui impose des contraintes fonctionnelles, comme celles de temps, de mode, de négation et/ou d'accord (Debaisieux, 2006).

Certaines études analysent comme faits de rection toutes les relations d'implication non-symétrique existant entre deux éléments au sein d'un énoncé, en utilisant les effets d'agrammaticalité provoqués par la suppression de l'un d'eux (i.e. le segment A régit le segment B si on ne peut pas avoir B sans A; Groupe de Fribourg 2012). Or, il est souvent difficile de distinguer les relations d'implication commandées par la syntaxe des catégories de celles relevant d'un autre niveau d'organisation (Sabio, 2013) : si l'impossibilité de supprimer un élément d'un énoncé prouve son caractère "essentiel", il ne permet pas de distinguer les nécessités grammaticales des nécessités discursives, ni de déterminer si ces éléments constituent malgré tout une unité distincte. Nous intégrons par conséquent dans ce travail des choix de description tenant compte du domaine de la macro-syntaxe (Benzitoun et al., 2009). Selon Blanche Benveniste (1990), les constructions régies sont équivalentes à une proforme² (i.e. remplaçable par un pronom; Debaisieux 2016), et reçoivent les modalités assertive, négative ou interrogatives du syntagme verbal principal. Les constructions régies sont également clivables/topicalisables (e.g. c'est A que/qui B; Benzitoun et al. 2009).

De même, les éléments régis grammaticalement sont normativement définis comme ceux se trouvant sous la portée du syntagme verbal principal (Gaatone, 2005). L'application de ce critère aux subordonnées nécessite toutefois la décomposition de ce phénomène en plusieurs degrés, placés sur un continuum allant de l'intégration à l'autonomie. Ce continuum est illustré par la figure 3.1 plus bas. Nous distinguons dans ce travail quatre cas de figure :

1. l'élément est pleinement régi grammaticalement et est réalisé sous la même unité pragmatique : il est directement intégré au noyau et ne constitue pas de préfixe ou de suffixe; il ne constitue pas non plus un acte de langage à lui seul. Cette configuration représente 37 % des subordonnées du corpus.
2. l'élément est régi grammaticalement mais il n'est pas réalisé sous la même unité pragmatique; la construction subordonnée est encodée sous la portée du verbe principal à l'aide d'un pronom relatif ou d'une conjonction pour les cas les plus explicites, mais constitue un acte de langage à elle-seule. 23 % des subordonnées sont dans ce cas de figure.
3. l'élément est non-régi grammaticalement, mais est réalisé sous la même unité

2. Nous appliquons les tests de l'Approche Pronominale avec une seule réserve : un test fonctionnel sur le français impliquant des déplacements dans l'unité syntaxique et des manipulations modales et/ou illocutoires peut être indicateur d'un phénomène différent sur l'anglais.

pragmatique; il ne constitue pas d'acte de langage à lui seul. Cette configuration concerne 12.1 % des subordonnées.

4. l'élément est non-régi et constitue un acte de langage. 26.7 % des subordonnées sont réalisées de cette manière et sont donc pleinement autonomes.



FIGURE 3.1 – Continuum de dépendance syntactico-sémantique des constructions subordonnées.

Après avoir identifié des indices de rupture dans les différents composants tels que nous les avons formulés dans nos hypothèses du Chapitre 1 p. 58, nous sommes en mesure de placer les trois types de subordonnées sur un continuum entre intégration et autonomie, ainsi que de déterminer les indices les plus mobilisés dans le marquage d'une rupture. Les relatives déterminatives sont les plus intégrées à leur entourage, ne présentant qu'un seul indice de rupture régulier, de nature discursive. Les circonstancielles comptent trois marques de disjonction, réparties entre les composants syntaxique et discursif. Les appositives relatives sont finalement très nettement autonomisées, totalisant sept paramètres de nature discursive en grande partie. Le détail des caractéristiques de chaque type de subordonnée est donné de la construction la moins autonome à celle la plus en rupture.

3.2 Résultats

3.2.1 Relatives déterminatives

Les relatives déterminatives représentent le type syntaxique le plus intégré au contexte. Si cette liaison est totale du point de vue syntaxique, les relatives déterminatives sont distinguées de leur entourage par un indice discursif.

Macro-syntaxe

En ce qui concerne l'intégration syntaxique des relatives déterminatives, les cas de rection totale concernent 60 % d'entre elles. Les relatives déterminatives sont

Cette relative déterminative est insérée dans une séquence où Zoe exprime son stress à l'idée de devoir enseigner une heure entière à l'aide de matériel informatique, matériel avec lequel elle ne se sent pas très familière. Les constituants de SC font bien partie de la relation prédicative de L, en tant qu'extension de l'item "people" sur L. En revanche, "people" n'est pas tout à fait en relation immédiate avec le syntagme verbal "explain", dans le sens où ce constituant joue le rôle de complément d'objet indirect. Si SC montre bien une dépendance syntaxique, cette dépendance n'est pas exactement la même que celle entre un syntagme verbal et l'un de ses arguments directs, comme dans l'exemple (1) vu précédemment. D'ailleurs, l'ensemble formé par le dernier constituant de L et l'intégralité de SC n'est pas clivable : *It's to people that obviously use computers a lot more than me that how do we explain. La *Noun Phrase Accessibility Hierarchy*, proposée par Keenan et Comrie (1977), mesure le coût cognitif demandé par la perception d'une structure relative. Selon cette échelle, l'intégration de deux propositions par le biais d'un item qui a un rôle syntaxique d'objet rendrait la relative plus complexe et moins accessible qu'un pronom relatif sujet. La structure de l'exemple (2) serait donc d'une complexité supérieure par rapport à une relative introduite par un pronom relatif sujet. Par ailleurs, SC est caractérisé par deux adverbes intensifieurs "obviously" et "a lot more", qui renforcent et modalisent l'assertion principale du passage. La localisation de ces deux éléments sur SC est une trace de la forte valeur pragmatique du segment, comparé à L et à R. La co-locutrice pourrait d'ailleurs réagir au contenu de L, qui est une question, mais aussi au contenu de SC, qui transforme cette question en un énoncé très orienté. Elle pourrait produire un énoncé du type "maybe they don't use a computer more than you do", ou bien répondre directement à l'implicature que Zoe attache à SC "oh come on, you know how to use a computer, you'll be fine". Cette orientation énonciative très marquée se voit également avec un changement dans l'utilisation des pronoms (Zoe passe du "we" et du général avec "people" à la première personne avec "me" en fin de SC).

L'autonomie syntactico-sémantique totale (décrite par la configuration (4) sur le schéma de dépendance syntaxique) ne concerne que 12.7 % des relatives déterminatives, ce qui représente la proportion la plus faible des types syntaxiques de subordinées. Ce résultat est significativement différent des appositives ($F(54,54) = 2.3, p < 0.001$), mais pas des circonstancielles. Les occurrences autonomes d'un point de vue syntactico-sémantique sont par conséquent particulièrement remarquables. Deux catégories de relatives déterminatives non-régies se distinguent, entre les occurrences en réelle juxtaposition (comme l'exemple (3) ci-dessous) et les constructions marquées par un mouvement syntaxique (à l'image de l'exemple (4) plus bas).

- (3) Zoe L i'm coming back for someone
 SC **that's not even there (laughs)**
 R (h) it was awful #

Les relations macro-syntaxiques et sémantiques de cette relative déterminative ne correspondent pas à la relation traditionnellement exprimée par le marqueur “that”. Si “that” marque traditionnellement une relation de subordination entre deux propositions, Zoe présente ici un élément adversatif relevant davantage de la coordination par le biais d'un morphème subordonnant. La réalisation du prédicat < not be there > n'est en fait pas antérieure ou concomitante à < come back > sur L ni à < be awful > sur R, et SC pourrait être glosé par “I'm coming back for someone, but that someone is not even there”. Le participant sujet de SC est distinct du “I” de L. Le test du contraste sous la portée d'un opérateur verbal modal est difficilement réalisable du fait de la charge modale de l'adverbe “even” : *I'm not coming back for someone that's not even there, but for someone that's even there”. Les tests de l'Approche Pronominale (e.g. Debaisieux, 2016) signalent néanmoins que l'indépendance du segment est davantage pragmatique/modale que syntaxique à proprement parler. On peut en effet tout à fait substituer l'ensemble formé par SC et le dernier item de L à un pronom (“I'm coming back for him”), mais c'est au niveau du discours que le segment perd sa pertinence (*“I'm coming back for him (h) it was awful #”).

Les cas de non-rection entre deux segments successifs peuvent également être impliqués dans un mouvement syntaxique servant au marquage d'une proposition. Ce mouvement peut consister en une structure présentative ou existentielle (i.e. du type “it's X that Y”/“s(he)'s X who Y”/there's X who Y”, où Y comporte le syntagme verbal principal de la séquence; Fox et Thompson, 1990), ou bien en une structure relative expérientielle/de perception (Kuno, 1974; Lambrecht, 2000), comme dans l'exemple (4).

- (4) Tom did you hear the news
 L recently
 SC **that they published this # er surv-**
 oh not survey
 R like a # a piece of work

Dans cette séquence, la relative déterminative est attachée au groupe nominal complément d'objet direct “the news”, dont elle déploie le sémantisme. SC est néanmoins séparé de son antécédent par l'adverbe temporel “recently”, qui porte

sur le prédicat “hear the news” et est réalisé sous un groupe intonatif distinct. Ce dernier prédicat est composé d’un verbe de perception, “hear”. Un tel appariement au verbe d’expérience “hear” permet de placer le procès décrit par SC (< publish survey >) en élément rhématique, et de lui donner ainsi davantage de dynamicité. SC a donc une fonction pleinement prédicative (Lambrecht, 2000), par rapport à la fonction cadrative du syntagme verbal avant L, dont la voix est davantage passive (contrainte par le sémantisme du verbe). Bien que SC implique un sujet autre que son antécédent, cette structure présente de nombreuses similarités avec ce que Grosu et Landman (1998) appellent les *internally headed relatives*, dans le sens où “hear the news” est difficilement interprétable comme syntagme recteur de SC, et que la licéité de SC dans la macro-séquence semble intrinsèque. Dans cet extrait, la relation prédicative avant L sert de repère, et fait de l’association de SC à son antécédent une entité à la fois objet de perception par rapport au centre perceuteur (“you”) et matière à prédication (SC). “Hear” ne contient pas “publish”, mais aide à augmenter sa dynamicité. La séquence peut d’ailleurs être glosée sans le morphème subordonnant, comme l’exemple (5) ci-dessous :

- (5) Tom did you hear the news
 L recently
 SC **they published this # er surv-**
 oh not survey
 R like a # a piece of work

Ces constructions spécifiques ont l’apparence d’unités dépendantes du verbe mais n’en ont pas les propriétés : elles ne sont pas équivalentes à une proforme, pas cliables, et peuvent posséder une modalité propre.

Si les relatives déterminatives témoignent en majorité d’une forte intégration syntactico-sémantique à leur entourage, les indices morphosyntaxiques spécifient également un lien étroit avec le co-texte gauche. 69.1 % des occurrences sont introduites par les pronoms relatifs “that” (47.3 % des relatives déterminatives) et \emptyset (21.8 % des relatives déterminatives), qui marquent une relation syntactico-sémantique très serrée entre deux propositions, et dénotent une orientation objectale (i.e. neutre; Melis, 2008; Khalifa, 2004). L’exemple (6) illustre cette caractéristique.

- (6) Joey L you know and the little bits
 SC **that stick out #**
 Elena yeah
 Joey R you have those
 like with the pollen on #

Dans cette séquence descriptive, le pronom relatif “that” définit une propriété du référent “the little bits”. Cette propriété objectale est déployée analytiquement à partir du contenu sémantique de l’antécédent, et ne dépend pas de ce que peut en dire Joey en tant qu’énonciatrice. Le pronom relatif “that” témoigne ici d’une neutralité énonciative.

Un autre facteur lié aux morphèmes introducteurs participe à l’intégration co-textuelle des relatives déterminatives. Il s’agit de leur grande proximité temporelle avec le constituant qu’elles modifient lors de leur production. Ces segments sont contraints, par leurs morphèmes introducteurs, à exprimer des relations avec l’un des derniers constituants du co-texte gauche au sein des séquences macro-syntaxiques dans lesquelles elles s’inscrivent, comme dans l’exemple (7).

- (7) Tom so burglary was quite common in #
 L the street
 SC **i lived in**
 R and well they’ve- #
 you could tell
 which ones would get burgled

Une pause silencieuse démarcative avant L permet de lier SC au constituant de L qu’il modifie (i.e. “the street”). Le contenu propositionnel des deux segments L et SC est alors réparti en un bloc cognitif homogène et fluide. La tentative de lier R à cet ensemble est à l’inverse un échec, se soldant en un abandon de structure.

En résumé, les relatives déterminatives sont majoritairement caractérisées par une dépendance macro-syntaxique totale. La plupart d’entre elles sont introduites par des morphèmes introducteurs spécifiant un lien serré avec le co-texte gauche.

Discours

Au niveau discursif, les relatives déterminatives ne comptent qu’un seul indice de rupture. Ces constructions sont les plus fréquemment suivies par un *backchannel*³ minimal⁴ du co-locuteur, avec 50 % des relatives déterminatives suivies par

3. Le terme de *backchannel* correspond à l’ensemble des signaux verbaux, vocaux, et/ou gestuels produits par le co-locuteur pour montrer sa réception (son écoute, sa compréhension, son accord, son désaccord) au discours produit (Yngve, 1970).

4. Un *backchannel* minimal est un court signal non-intrusif (Tolins et Fox Tree, 2014), qu’il soit verbal (e.g. item du type “yeah”), vocal (e.g. “mhmh”), et/ou gestuel (e.g. un hochement

un *backchannel* de ce type, comme l'illustre l'exemple (8)⁵. Cette proportion n'est toutefois pas significativement différente de celle des appositives et de celle des circonstancielles.

(8)	Rhianna		no it's a # special program
		L	uh # for (h) people
		SC	who speak good English
	Alex		[yeah NOD]
	Rhianna	R	so hopefully#

À la suite de cette relative déterminative réalisée dans une séquence de questions-réponses, Alex produit un *backchannel* verbal minimal avec l'item "yeah", suivi d'un mouvement d'acquiescement de la tête (nod). En ne produisant aucun autre signal plus intrusif (sûrement en relation avec le contour intonatif final montant de Rhianna sur "English", indiquant clairement une suite), Alex montre son alignement⁶ à la séquence. Un consensus structurel s'établit entre les deux participantes, par rapport au rôle actuel de locutrice pris par Rhianna : Alex indique que les unités de discours précédentes ont été admises et comprises, à la manière d'un accusé-réception (Bertrand et al., 2009), et encourage Rhianna à continuer son discours (Schegloff, 1982; Goodwin, 1986; Bangerter et Clark, 2003). Bien que peu intrusifs, le fait que ces *backchannels* soient produits entre SC et R est une autre preuve de l'ensemble syntactico-sémantique et informationnel formé par L et SC. Benus et al. (2007) suggèrent que les *backchannels* mineurs sont élicités par la production de contours intonatifs montants du locuteur. Ces contours indiquent la continuation du discours et/ou représentent un appel au co-locuteur. Contrairement à ce résultat, les *backchannels* mineurs produits après les relatives déterminatives de notre corpus ne sont pas élicités par la production d'un contour intonatif montant sur SC, puisque seulement 24 % de ces *backchannels* minimaux sont immédiatement précédés d'un tel mouvement intonatif. En revanche, notons que 84 % des *backchannels* sont précédés d'un allongement significatif sur la syllabe finale de SC. La frontière entre SC et R, bien marquée par un allongement syllabique final, est l'endroit préférentiel des co-locuteurs pour la réalisation d'un signal de réponse. Ce signal de réponse peut être orienté vers le contenu de L, celui de SC, ou bien à celui de l'ensemble L-SC, puisque 80 % des *backchannels* minimaux sont produits après une relative déterminative di-

de tête, un sourire). Une distinction s'opère traditionnellement entre *backchannels* génériques et *backchannels* spécifiques (Goodwin, 1986; Bavelas et al., 2000). Alors que les génériques expriment simplement l'attention et la compréhension du co-locuteur, les spécifiques répondent au contenu spécifique du tour précédent, e.g. par une évaluation du type "oh wow" (Stivers, 2008).

5. L'unité gestuelle est représentée par les crochets. La nature de l'unité gestuelle figure en majuscules.

6. Le co-participant montre son alignement à la séquence en cours lorsqu'il produit un signal encourageant sa progression (Stivers, 2008).

rectement liée à L sous un même contour intonatif. Les *backchannels* minimaux sont donc plutôt réalisés à la suite de groupes intonatifs relativement longs, comme le montre l'exemple (9).

(9)	Michelle		i dunno must have been about #	
		L	ten # girls	1.5 sec
		SC	who were # in their final year #	3.2 sec
	Zoe		[NOD]	
	Michelle	R	so they must have been like eighteen year olds	1.5 sec

Dans l'exemple (9), Michelle décrit les élèves sous sa charge lors de son année en tant qu'assistante dans un internat de Rochefort. Une pause démarcatrice isole le référent "ten girls" du discours précédent. Ce référent est lié à SC sans aucune pause, L et SC ne formant qu'un seul groupe intonatif. Ce groupe intonatif est relativement long (4.7 secondes). La syllabe finale de SC est significativement allongée, comme le montre la figure 3.2 ci-dessous.

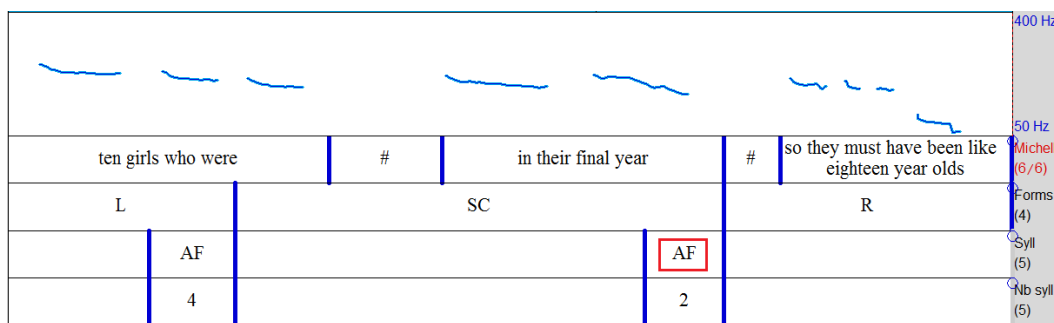


FIGURE 3.2 – Allongement syllabique final sur SC dans l'exemple (9), illustré dans la courbe intonative de Praat et les repérages syllabiques (AF = accentuée finale) et phonémiques (4 = syllabe composée de quatre phonèmes, 2 = syllabe composée de deux phonèmes) en comparaison avec la durée syllabique finale de L.

Le contour intonatif du dernier élément lexical de "girls" est relativement plat et similaire aux autres mouvements de son co-texte. En revanche, le contour du dernier item lexical du groupe prosodique de SC, "year", est descendant. Au premier abord, cette syllabe finale accentuée présente une durée comparable à la dernière syllabe accentuée de L. Pourtant, "year" ne comporte que deux phonèmes alors que la dernière syllabe accentuée de L en comporte quatre : elle est très significativement allongée (durée phonémique de 0.139 sec sur la finale de L vs. 0.257 sec sur la finale de SC). Zoe répond à cet ensemble par un hochement de tête. 68 % des *backchannels* minimaux réalisés à la suite d'une relative déterminative comportent un tel mouvement de tête, qu'il soit combiné ou non à d'autres traits gestuels ou à des traits vocaux/verbaux. Cette proportion en fait le marqueur le plus utilisé parmi les *backchannels* minimaux. De même, les *backchannels* gestuels non combinés à des

traits vocaux ou verbaux figurent en légère majorité (52 %) dans ces contextes. Cosnier (1988) remarque qu’une réponse minimale tel qu’un *backchannel* gestuel permet à la co-locutrice de marquer qu’elle a bien pris en compte le désir de la locutrice de poursuivre son tour, en utilisant un *backchannel* interférant le moins possible avec le discours produit (Bertrand et al., 2009). Michelle continue d’ailleurs son tour dans l’exemple (9), en proposant une élaboration de SC à l’aide du marqueur de discours inférentiel “so” (Fraser, 1999). Ce marqueur est également utilisé à titre résomptif puisque Michelle développe le même topique⁷ discursif (i.e. les élèves sous sa charge à l’internat), en explicitant l’âge des élèves.

Contrairement à ce qui était prévisible compte tenu de leur faible nombre d’indices de rupture, les relatives déterminatives et leurs macro-séquences hôtes comptent une proportion médiane de disfluences⁸ (e.g. répétitions d’items, abandon de structures) entre les trois types syntaxiques sous étude. Notons toutefois que ce taux de disfluences n’est pas significativement important (20 % des séquences contenant une relative déterminative présentent une ou plusieurs disfluences). Ces disfluences se produisent majoritairement dans le co-texte droit (i.e. R) et comptent légèrement plus d’abandons de structures (54.5 %) que de répétitions (45.5 %), comme dans le récit (10).

- (10) Tom so burglary was quite common in #
 L the street
 SC **i lived in**
 R and well they’ve- #
 you could tell
 which ones would get burgled

Alors que L et SC sont énoncés d’une manière homogène et fluide, la tentative de lier R à cet ensemble donne lieu à un abandon de structure. Tom éprouve des difficultés à intégrer le segment de discours R. Il combine deux marqueurs de discours résomptifs⁹ (“and” et “well”; Redeker, 2008) et change d’élément thématique¹⁰ en

7. Le topique d’une séquence discursive correspond à un référent établi, en tant qu’objet de discours, comme ce dont traite la séquence conversationnelle actuelle par le locuteur et le co-locuteur (Grosz et Sidner, 1986; Mondada, 2011).

8. Bien que nous n’ayons pas pris en compte les constructions subordonnées comportant des interruptions majeures (du locuteur ou du co-locuteur), il arrive que le co-texte présente des répétitions lexicales ou des structures grammaticales inachevées, auxquelles sont tout de suite greffées de nouvelles structures. Nous considérons que ces disfluences “modérées” sont représentatives d’un corpus oral spontané et qu’elles sont tout à fait analysables.

9. Les marqueurs de discours, spécifiant des relations procédurales (i.e. d’interprétation) entre plusieurs segments de discours, permettent de déterminer les endroits où des mouvements particuliers sont susceptibles d’être mis en place.

10. “Thème” est dans ce chapitre employé au sens de Halliday (1985), i.e. l’élément à partir duquel des informations sont prédiquées.

faveur d'un "they" indéfini. Tom retrouve la fluidité dans le groupe intonatif à la suite de R, à l'aide d'un autre élément thématique (i.e. "you"). Le fait que les disfluences se trouvent sur R indique encore une fois que les unités de L et de SC ont tendance à être planifiées ensemble. Il indique aussi que l'éloignement entre antécédent et modifieur, ainsi que l'insertion de *backchannels*, ne peuvent être les seuls facteurs des difficultés de calcul et/ou d'élocution : l'éloignement entre antécédent et subordonnée est quasiment nul dans cet exemple, et cette occurrence est tout à fait intégrée d'un point de vue macro-syntaxique. Le co-locuteur ne produit pas non plus de *backchannel*.

Hormis la présence de *backchannels* minimaux et des disfluences assez courantes dans les séquences, la forte intégration discursive des relatives déterminatives est visible à travers plusieurs paramètres. Elles expriment d'abord en intégralité une relation anaphorique, en étant liées à L. Cette reprise sémantique d'un élément référentiel du segment de discours précédent participe à la construction de la cohérence et évoque une continuité de l'expérience cognitive (De Beaugrande et Dressler, 1981), comme le montre l'exemple (11)¹¹.

- | | | | | |
|------|-------|----|--|--|
| (11) | Joey | L | you know these pots | |
| | | SC | that have like # | |
| | | | holes in them # | |
| | Elena | | [yeah we've got one at home | |
| | Joey | | like funnely bits] yeah | |
| | | R | they're good for strawberry growing apparently # | |

Cette relative déterminative permet de définir l'élément référentiel "these pots". SC lui propose une caractéristique définitoire et utilise un pronom ("them") désignant le même référent. L'intervention d'Elena continue la chaîne référentielle avec "one", qui finit avec l'utilisation de "they" par Joey en tant qu'élément thématique sur R. 45 % des relatives déterminatives comportent un pronom continuant la chaîne référentielle de la séquence. Si cette proportion n'est pas la plus élevée, elle est cependant notable.

La continuité séquentielle assurée par les relatives déterminatives est également garantie par leur position au sein du paragraphe discursif. À l'image de l'exemple (12), 76.4 % d'entre elles sont insérées en position de suffixe, position la moins intrusive.

- | | | | | |
|------|----------|---|-------------------------------------|-------|
| (12) | Michelle | L | it was better than that other place | [NOY] |
|------|----------|---|-------------------------------------|-------|

11. Les crochets représentent les chevauchements de parole.

SC	i was at #	[SUF1]
R	Rochefort # (laughs)	[SUF2]
	sur Mer	[SUF3]

Michelle s'apprête à décrire le collège de Rochefort sur Mer. SC précise la référence de "that other place", introduite dans le rhème de L avec un démonstratif distal. Hormis le démonstratif "that", plusieurs paramètres participent à la mise en valeur de cet élément référentiel. Le syntagme verbal principal de la séquence (i.e. "was"), en tant que copule, lui sert de faire-valoir en donnant aux constituants rhématiques le rôle de complément du sujet, sujet dont le sémantisme est vide. Michelle appose en suffixe un syntagme nominal (i.e. "Rochefort sur Mer") à la suite de SC, désignant le même référent. Elle décompose même ce syntagme en deux groupes intonatifs par un rire. Le résultat est une construction de la référence par une succession de groupes intonatifs, similaire à la construction d'un repère constitutif. Alors qu'un repère constitutif concerne les constituants thématiques, l'arrangement des constituants de cette séquence place l'élément en cours d'élaboration en position rhématique.

Par ailleurs, nous avons identifié trois types de rupture communément marquées par les subordonnées au niveau discursif : les constructions sortent des attentes séquentielles (i.e. la subordonnée présente un ou plusieurs éléments ne répondant pas aux attentes que le co-locuteur peut projeter de la séquence par rapport à son type spécifique, e. g. récit, argumentation, questions-réponses, description); les constructions présentent une rupture chronologique; les constructions présentent une rupture de thème/participant. Ces critères établis, 43.6 % des relatives déterminatives ne présentent aucune rupture parmi ces trois catégories. Cette proportion est plus importante que celle des circonstancielles et des appositives ne présentant pas de rupture, bien qu'elle ne soit pas significative. Les cas d'homogénéité sont illustrés dans l'exemple (13).

- (13) Kate but i like sort of #
 i like cakes
 L and everything else
 SC **that's sweet**
 R like chocolate and anything #

Dans cette relative déterminative "maximaliste" (Grosu et Landman, 1998) donnant une propriété contrastive à "everything else", SC contribue à la séquence argumentative dans laquelle elle est insérée sans présenter d'élément inattendu. SC constitue en effet une partie de l'un des arguments consécutifs justifiant l'opinion

argumentative énoncée antérieurement par Kate (“I’m a sweet tooth”). Le sujet principal reste “I” au long de la séquence discursive et SC est énoncé au présent simple comme le reste.

45.4 % des relatives déterminatives changent néanmoins de thème ou de participant au procès décrit dans l’énoncé. Nous ne considérons pas que ce trait soit spécifiquement représentatif de ce format syntaxique, puisqu’un autre type de construction présente exactement la même proportion de changement thématique. La relative déterminative dans la séquence (14) montre que ce genre de rupture est réalisée à des fins discursives spécifiques.

- (14) Alex L but everyone else
 SC **i’ve spoken to**
 R is having trouble finding something #

Dans un contexte d’argumentation co-construite avec Rhianna, la co-locutrice, Alex montre combien il est difficile de trouver un job d’été sans l’aide d’une connaissance. Elle vient de finir un récit personnel se concluant sur la façon dont elle-même a trouvé un travail, et revient à la ligne principale de l’argumentation en réassertant l’opinion à partir de laquelle Alex et Rhianna élaborent des mouvements discursifs. L’utilisation du “I” sert à la fois à marquer la transition entre récit personnel et ligne argumentative directrice (impliquant un changement de coordonnées énonciatives) et à restreindre la référence de L. Pour des raisons de politesse et de face conversationnelle¹² (Goffman, 1967), Alex ne peut pas affirmer que tout le monde a des difficultés à trouver un travail saisonnier. SC sert alors à augmenter la précision du référent “everyone else” en veillant à ne pas diminuer la cohérence argumentative, tout en permettant à Alex de protéger sa face conversationnelle. Sans SC, l’assertion pourrait impliquer une généralisation hâtive de la part d’Alex.

Les marqueurs de discours, dont nous avons déjà mentionné la pertinence dans le marquage des mouvements discursifs, figurent assez peu dans les séquences contenant une relative déterminative. De manière générale, ils se trouvent en plus grand nombre dans R (48.9 % des marqueurs de discours apparaissant dans une séquence contenant une relative déterminative). Selon le modèle de cohérence de Redeker (2006) à la suite du travail de Schiffrin (1987), les marqueurs de discours peuvent être classés en deux grands ensembles : les “opérateurs” qui structurent le discours à un niveau séquentiel, et les “particules”, qui modulent l’interprétation des énoncés.

12. La face conversationnelle est un concept de sociolinguistique. Selon Brown et Levinson (1987), la face correspond à l’image publique que chaque individu construit, maintient, et tente de protéger.

Seulement 25 % des séquences contenant les relatives déterminatives possèdent un marqueur de structuration. La séquence (15) ne comporte aucun marqueur; les deux participantes discutent de Brecon, une ville galloise qu'elles connaissent toutes les deux. Joey, la co-participante, demande à Elena si elle s'y est déjà rendue.

- | | | | |
|------|-------|----|------------------------------------|
| (15) | Elena | | i went # |
| | | | i think # |
| | | L | the first year |
| | | SC | my brother was in uni # |
| | Joey | | ooh |
| | Elena | R | think we went there just for lunch |

Cette relative déterminative s'insère dans le discours d'Elena sans que celle-ci ait recours à un marqueur de discours structural. Au lieu de présenter l'information demandée par Joey à l'aide d'une référence temporelle factuelle et absolue, Elena utilise un point de référence personnel pour décrire le moment où elle a visité Brecon. Elle lie un évènement décrit à la première personne à une temporalité exprimée à la troisième personne en faisant référence à "my brother". Cette relative déterminative présente donc en revanche une rupture d'élément thématique. Joey marque d'ailleurs l'information d'un *backchannel* vocal minimal, indiquant une mise à jour de ses connaissances (Heritage, 1984). Malgré cette intervention, Elena reprend le fil de sa séquence sans disflue, toujours sans marqueur.

En résumé, les relatives déterminatives sont caractérisées par leur grande intégration verbale, particulièrement liées à leur co-texte gauche. Elles sont majoritairement régies par le syntagme verbal de gauche, et sont produites dans la localité de l'expression nominale qu'elles modifient. Leur morphème introducteur leur dénote une orientation objectale, à l'image de l'ouverture d'un fichier informationnel précisant l'identité du référent antécédent. Les relatives déterminatives ne créent guère davantage de rupture discursive. La seule interruption qu'elles occasionnent prend la forme de *backchannels* minimaux du co-locuteur; ceux-ci figurent en remarquable proportion dans leur séquence hôte et sont positionnés à leur suite, entre SC et R. La plupart de ces *backchannels* expriment néanmoins un accord structural et/ou interactionnel entre locuteur et co-locuteur, et encouragent le locuteur à continuer sa séquence.

La grande intégration des relatives déterminatives à leur co-texte est également visible dans l'absence de corrélation significative entre les différents indices de rupture évoqués. La corrélation la plus élevée entre les indices concerne l'interaction entre les ruptures thématiques et les *backchannels* minimaux : 52 % des séquences

dans lesquelles le co-locuteur produit un *backchannel* comportent une relative déterminative changeant de thème, comme dans la séquence (15) précédente. Cette interaction suggère que le co-locuteur exprime légèrement davantage sa réception du discours lorsque le locuteur prédique des informations à partir d'un objet différent du discours précédent. Cette corrélation ne représente toutefois que 23.6 % des relatives déterminatives du corpus, et reste donc une tendance peu significative.

Compte-tenu des caractéristiques des relatives déterminatives, les interactions entre les différentes marques d'intégration sont davantage représentatives des séquences contenant ce type syntaxique. Il n'est par exemple pas rare qu'un cas de rection macro-syntaxique totale ne présente également aucune rupture discursive parmi les types que nous avons identifiés (51.5 % des relatives déterminatives totalement régies d'un point de vue syntaxique ne montrent également pas de rupture discursive; cette configuration représente 30.9 % des relatives déterminatives). La séquence (16) suivante illustre cette intégration totale sur les deux niveaux d'organisation verbale :

- (16) Zoe L i hate speaking to people
 SC **that i don't know that well #**
 Michelle yeah it's a real strain after a while
 isn't it
 Zoe R yeah but every si-
 it's nice
 like # from time to time

Zoe décrit ici les avantages et les inconvénients de son année passée en France, et exprime la lassitude qu'elle éprouvait à force de loger chez des familles différentes tous les weekends. D'un point de vue macro-syntaxique, SC est remplaçable par une proforme ("I hate speaking to them"/"I hate speaking to those people") et est également clivable avec son antécédent ("It's to people that I don't know that well that I hate speaking"). De même, les deux locutrices viennent de plaisanter sur le fait que Zoe changeait de famille tous les weekends, "being passed around all the families". Elles se sont mis d'accord par le rire sur le fait que Zoe allait de famille en famille, par incidence toutes totalement étrangères à Zoe. SC n'est donc pas un acte de langage autonome, et est encodé comme de l'information présupposée : la co-locutrice ne peut pas objecter à SC isolément de L (*"oh but you knew them well"). D'un point de vue discursif, SC comporte un verbe cognitif au présent, comme le reste de la séquence. L'élément thématique reste "I" comme sur L, et SC contribue comme L à l'évaluation du récit de Zoe.

3.2.2 Circonstanciellles

Présentant trois indices de rupture, les circonstanciellles sont caractérisées par une macro-syntaxe relativement flexible et dynamique. Les deux indices discursifs sont de telle nature qu'ils renforcent l'idée de flexibilité des circonstanciellles, aussi bien dans leurs positions que dans leurs fonctions. Si elles montrent une capacité à ouvrir ou clôturer des cadres de discours valables pour plusieurs groupes intonatifs à la suite, la grande proportion de marqueurs de modulation du discours à leur suite confirme que les circonstanciellles marquent un changement d'ordre interprétatif plutôt qu'une rupture linéaire dans les séquences et dynamiques discursives.

Macro-syntaxe

D'un point de vue macro-syntaxique, seulement 24 % des circonstanciellles sont pleinement régies au sein de leur séquence hôte. Cette proportion relativement basse est partagée avec les appositives. La différence avec les relatives déterminatives n'est pas significative. Les constructions pleinement contrôlées peuvent être antéposées et postposées, comme le montre l'exemple (17) :

- | | | | |
|------|-------|----|---------------------------------------|
| (17) | Joey | L | i'd- i used to get honey out of these |
| | | SC | when i was little # |
| | Elena | | did you |
| | | | [can you do that |
| | Joey | R | yeah if you] take hem # |

Dans cette séquence, “when I was little I'd- I used to get honey out of these” est parfaitement envisageable. Cette flexibilité est liée au fait que SC est encodée comme une présupposition et ne constitue pas un acte de langage indépendamment de L : la co-locutrice ne peut pas contredire SC indépendamment du co-texte (*“but you were never little”). De même, le prédicat < be little > est complémentaire à celui de L (< get honey >), sans construire de chronologie. SC peut être clivée (“it's when I was little that I used to get honey out of these”) et le test de la négation (Deulofeu, 2013) est également réalisable (“I didn't use to get honey out of these when I was little, but when I was on holiday”).

Si peu de circonstanciellles témoignent d'une si grande rection, la majorité des constructions ne sont pas totalement autonomes pour autant. 54.5 % d'entre elles sont en fait partiellement autonomes, montrant un mode d'insertion intermédiaire. À l'image de l'exemple (18), 31 % des circonstanciellles ne sont pas contrôlées par les syntagmes verbaux de leur séquence hôte, mais font partie de la même unité

énonciative que leur co-texte (i.e. elles ne constituent pas des actes de langage distincts).

- (18) Rhianna L and then
 SC **when i arrived #**
 R i heard of that # French and British programme

Dans cette séquence comprenant une circonstancielle servant de cadre au discours de Rhianna, l'action dénotée par SC n'est pas contrôlée par le syntagme verbal de R. SC participe même à la chronologie de la séquence, en exprimant un procès antérieur à l'événement de R. Le procès de SC est inscrit dans une succession. La circonstancielle pourrait d'ailleurs être glosée sans grande différence sémantique par "I arrived and I heard of that # French and British programme". Cette circonstancielle temporelle respecte l'ordre des événements, conformément à la tendance de ce type à contribuer à un ordre expérientiel iconique, identifiée par Crompton (2009). SC impose toutefois des contraintes d'ordre temporel sur le syntagme verbal de R : le procès de R ne peut être acté qu'au prétérit (*"when I arrived I hear of that French and British Programme"). Par ailleurs, SC est intégrée au niveau de son contenu illocutoire : il est difficile d'imaginer la co-locutrice réagir au contenu de SC sans le lier à R (*"but you didn't arrive"). SC est encodée comme de l'information présupposée, et ne représente pas un acte de langage autonome.

À l'inverse, 23.6 % des circonstancielles constituent des actes de langage, mais font partie de la même unité prédicative que celle de leur macro-séquence. C'est le cas de l'exemple (19) ci-dessous.

- (19) Michelle L because sometimes #
 SC **(h) when everybody went home**
 R i'd stay for a weekend # and #

Dans cette circonstancielle servant de cadre au discours de Michelle, le prédicat "everybody went home" est mis en parallèle dans une relation d'opposition au prédicat de R, "stay for a weekend". Il reste d'ailleurs possible d'inverser l'articulation des deux propositions sans grand impact sémantique : "I'd stay for a weekend when everybody went home". La co-locutrice peut toutefois remettre en question SC, avec un énoncé du type "it can't be everybody, someone had to stay".

L'extrait (20) témoigne de la capacité de certaines circonstancielles à porter une valeur pragmatique très forte, telle qu'une demande :

- (20) Zoe what time do you get back home today
 Michelle so we can get some # synchronization going on
 Zoe L but
 SC **if you # ring me #**
 R you can buy it today

Zoe tente de persuader Michelle d’acheter au plus vite son billet de train pour l’Angleterre, pour qu’elles puissent rentrer ensemble pour les vacances. On peut tout à fait interpréter SC “traditionnellement”, comme projetant une condition sur la réalisation de R (Hakulinen et Selting, 2005) : avec l’adverbe modal “can” sur R, la conditionnelle joue un rôle sur la relation de possibilité entre “you” et la réalisation du prédicat “buy it”. SC détermine la déonticité de R. De plus, l’adverbe “today” ancre la réalisation de ce prédicat dans l’immédiat et le *realis*, plus que dans le possible. En revanche, le sens d’optionnalité de la condition apportée par SC est en fait neutralisé. L’action < you buy it today >, que Zoe s’applique à inscrire dans le champ du possible et même de l’imminent, ne sera pas réalisée si < you ring me > s’avère du domaine de l’*irrealis*. SC est alors une directive : Zoe voudrait que Michelle lui téléphone. La séquence pourrait être glosée sans morphème subordonnant : “but ring me and you can buy it today”. De même, la séquence est proche des tournures en “if A, then B”, exprimant d’abord une condition, puis une conséquence (Linström et al., 2016) : “if you ring me, then you can buy it today”. La force pragmatique de cette circonstancielle qui vise à produire une action de la part de la co-locutrice fait écho à celles exposées par les travaux de Ford (e.g. 1997), de Evans (2007), et de Ehmer (2016). Si ces structures ont également été montrées dans des environnements interactionnels spécifiques où la conditionnelle n’est pas précédée ou suivie d’un autre syntagme verbal (Lindström et al., 2016), les circonstanciels précédés et/ou suivies d’autres syntagmes verbaux comme dans la séquence (20) possèdent une force pragmatique qui nous paraît remarquable, émergeant des contingences de la conversation spontanée.

Seulement 20 % des circonstanciels sont autonomes, ce qui représente une faible proportion par rapport à la distribution de constructions jouissant d’une autonomie partielle (i.e. 55 %). Cette différence n’est toutefois pas significative. Si nous nous intéressons aux deux sous-classes sémantiques que nous avons prises en compte (i.e. circonstanciels temporelles et circonstanciels conditionnelles), nous observons une légère tendance vers l’autonomie totale chez les circonstanciels de condition par rapport aux temporelles (12.7 % des circonstanciels conditionnelles en autonomie vs. 9.1 % des circonstanciels temporelles). Cette tendance est toutefois loin d’être significative, puisque la proportion totale de constructions en autonomie n’est de toute manière pas élevée. Dans la séquence (21), Joey vient d’expliquer à Elena

comment obtenir du miel de fuchsia. Elena remarque que le procédé est le même que pour le chèvrefeuille, à l'aide d'une circonstancielle conditionnelle totalement autonome du point de vue macro-syntaxique.

- (21) Elena that's like in the
 L that with the honeysuckle #
 Joey yeah
 Elena SC **if you take it off**
 R and just squeeze down
 [you #
 Joey yeah]

SC exprime une action antérieure à R : il faut d'abord enlever les panicules de la fleur ("take it off") avant de presser cette dernière ("squeeze down") pour obtenir le miel. L'inversion des procès de SC et R donnerait une séquence d'action illogique : *"and just squeeze down if you take it off"). La "règle" de la précédence impose ici de tenir compte dans la suite de ce qui est mémorisé antérieurement (Muller, 2008). La position de SC n'est donc pas modifiable. S'il est vrai que SC précise les conditions pour lesquelles la relation prédicative de R peut être réalisée, R est coordonné à SC par "and". Réelle conjonction de coordination ou marqueur de discours, on peut dans tous les cas lire R comme la continuation de l'enchaînement prétendument hypothétique, sans que ces deux propositions soient suivies par une proposition principale. Le fait que deux lectures soient possibles témoignent de l'important degré de liberté de SC. Cela n'empêche absolument pas la co-locutrice de réagir à la séquence : elle a compris les tours d'Elena et exprime même son accord avec le contenu propositionnel. De même, la co-locutrice pourrait objecter à la marche à suivre et plus particulièrement à SC, avec une intervention comme "yeah but you can't really take them off, can you?". SC est donc un acte de langage à lui-seul. Ce cas de conditionnelle "insubordonnée" (Lindström et al., 2016) est très proche d'une directive ou d'un conseil, moins directe (Dancygier, 1998) et par conséquent plus polie qu'une injonction ou qu'une instruction réalisée à l'aide d'une proposition principale (Ford, 1993; Couper-Kuhlen et Thompson, 2009; Laury et Suzuki, 2011). Le résultat de toutes ces actions aurait pu être énoncé après R ("get honey" suivant possiblement le sujet "you"), mais la co-locutrice n'y est pas sensible : elle produit le *backchannel* "yeah" en chevauchement total avec la fin du discours d'Elena. Ce *backchannel* est d'une telle intensité qu'il met fin à la fois au discours d'Elena et à la séquence autour du miel de fleurs.

Au niveau des marqueurs morphologiques, les circonstanciels sont traditionnellement contraintes à exprimer la simultanéité ou la correspondance pour les temporelles, l'*irrealis* ou l'hypothétique pour les conditionnelles. Cependant, certaines

occurrences (18.2 % des circonstancielle), temporelles comme conditionnelles, expriment une opposition, comme la circonstancielle suivante insérée dans la séquence argumentative (22).

- (22) Tim L although
 SC **when i passed**
 R i didn't think i was a good driver

Dans cet extrait, seul SC comporte un verbe de procès (i.e. “I passed”); ces verbes de procès caractérisent 58 % des circonstancielle. Bien que SC soit introduit par une conjonction et que “I passed” serve de repère à l'état cognitif temporaire exprimé par R, la relation sémantique entre les deux propositions est celle d'une opposition : le segment pourrait être glosé par “I didn't think I was a very good driver but I passed”, ou bien “although I passed, I didn't think I was a very good driver”. L'absence de parallélisme entre marques de surface syntaxiques et relations logiques et discursives est mise en avant par Blühdorn (2008), qui démontre que l'éventail de fonctions des connecteurs morphosyntaxiques ne soutient pas la vision d'un parallélisme entre structure syntaxique et structure discursive. D'autres chercheurs, comme Shyldkrot Hava (1995, p. 146), parlent de “polysémie” et de grammaticalisation des connecteurs qui introduisent ces subordonnées. Nous considérons pour notre part ce phénomène en termes de multiplicité de fonctionnements du même morphème introducteur (Debaisieux, 2006; Berrendonner, 2008) : “when” et “if” ne sont effectivement pas très spécifiants du type de relation qu'ils mettent en place entre les deux segments de discours qu'ils lient. Dans le sens où la diversité des usages des marqueurs se joue sur plusieurs niveaux (syntaxique, discursif), nous n'irons pas jusqu'à parler de pragmatization (e.g. Traugott, 1999), puisque les différents niveaux d'interprétation (i.e. syntaxique et discursif) sont simultanément accessibles (Debaisieux, 2016). En revanche, comme certains marqueurs de discours, la possibilité de lecture grammaticale des conjonctions restreint la portée des fonctions sémantiques et pragmatiques que le segment peut remplir. Par ailleurs, si les conjonctions sont polyfonctionnelles, elles sont contraintes à apparaître à gauche de la proposition qu'elles introduisent, à la différence d'une majorité de marqueurs de discours. La polyfonctionnalité des morphèmes introducteurs est alors intéressante du point de vue des mécanismes cognitifs par lesquels les locuteurs connectent certaines structures à d'autres, que les unités qu'elles lient soient syntaxiques ou discursives.

En somme, les circonstancielle présentent davantage d'occurrences témoignant d'une macro-syntaxe flexible que d'occurrences pleinement régies. Les syntagmes ver-

baux qu’elles encodent se caractérisent par un dynamisme relativement important par rapport aux autres formats syntaxiques. Bien que leurs marqueurs introducteurs les contraignent en majorité à exprimer un faisceau fini de relations sémantiques par rapport au co-texte, certaines circonstancielles expriment des relations plus complexes, dépassant la simultanéité pour les temporelles, ou l’*irrealis* pour les conditionnelles.

Discours

Les circonstancielles comptent deux indices de rupture discursive. Le premier indice donne des informations sur les potentiels emplacements de rupture lors de la production de ce type de construction. À l’image de l’exemple (23) plus bas, les séquences contenant des circonstancielles contiennent la proportion la plus élevée de marqueurs de discours “particule” (Redeker, 2006), indicateurs d’une relation procédurale attitudinale. En revanche, cette proportion n’est pas significativement différente de celle des relatives déterminatives ou de celle des appositives. Ces marqueurs se situent préférentiellement à la fin du co-texte gauche (40 % des marqueurs de modulation produits dans une séquence contenant une circonstancielle). Les circonstancielles sont alors introduites par des marqueurs spécialisés non pas dans l’indication d’un changement particulier dans l’organisation discursive (i.e. l’indication d’une relation procédurale textuelle), mais dans la modulation de l’interprétation du discours (i.e. l’indication d’une relation procédurale attitudinale – épistémique ou pragmatique).

Deux marqueurs spécifiques sont récurrents, et n’apparaissent pas aux mêmes endroits. Le marqueur de modulation le plus commun est “like” (34.3 % de tous les marqueurs de modulation, tous emplacements dans les séquences contenant une circonstancielle), préférentiellement sur R, suivi par “you know” (17.1 % de tous les marqueurs, tous emplacements), préférentiellement sur L. Le fait que le marqueur le plus répandu (i.e. “like”) n’apparaisse pas sur L, qui comporte pourtant davantage de marqueurs de discours de modulation que R, indique que L comporte une plus grande diversité de marqueurs que R. Il serait par contre très hâtif de conclure que les relations procédurales indiquées par les marqueurs sont plus diverses lorsque les circonstancielles sont introduites que lorsqu’elles sont clôturées : un seul et même marqueur peut recouvrir une multiplicité de fonctions en fonction du contexte. La séquence (23) illustre l’utilisation des deux marqueurs “you know” et “like”.

(23) Alex L you know

SC **when there's turbulence #**
 R (h) i always look at them i'm like
 if they're standing there smiling
 then it's fine (laughs)
 whereas if they ever look worried
 i'm always a bit like
 why- why me
 why now (laughs)

Cette circonstancielle servant de cadre au discours d'Alex se trouve dans une description de l'angoisse qu'éprouve cette dernière à bord d'un avion. En se positionnant parallèlement au contenu propositionnel de l'énoncé (Clark, 1996), "you know" sur L opère un pointage sur le discours à suivre plutôt que de constituer un véritable appel à l'autre (Fox Tree et Schrock, 2002); cette forme n'est d'ailleurs pas suivie de *backchannel*. Il invite la co-locutrice à tirer des inférences de ce qui va être dit, en identifiant la pertinence et les implications de l'énoncé suivant, ici à un niveau représentationnel : les deux locutrices ont précédemment abondamment discuté des vols en avion, et classé l'habitude des vols à turbulences dans leurs connaissances partagées. Alex invite donc ici Rhianna à interpréter les segments de discours qui vont suivre avec l'histoire conversationnelle de ce sujet en tête. "You know" agit donc davantage sur le plan métacommunicatif que sur celui du contenu propositionnel, en améliorant la distribution du message plutôt que de porter un message en lui-même. En fin de R, "like" débute une reconstitution du comportement d'Alex dans la situation de discours qu'elle a créée (Clark et Gerrig, 1990), à bord d'un avion pendant que des turbulences secouent l'appareil. Ce "like" sert à introduire le glissement du discours d'Alex-énonciatrice vers celui d'Alex-personnage de la situation décrite, et sert à augmenter la valeur dramatique de la description (Romaine et Lange, 1991; Buchstaller, 2001).

À l'image de ce dernier exemple (23), 33.3 % des "like" insérés à la suite des circonstanciels introduisent la reconstruction par représentation d'une situation. Ces situations sont dans tous les cas du corpus centrées sur le locuteur et mettent en scène des expériences personnelles. Il serait trop rapide de conclure que ces séquences sont centrées sur l'interaction et élicitent un certain alignement, voire une certaine affiliation du co-locuteur (une reconstruction représentationnelle d'une situation passée ou fictive est vivante et s'accompagne souvent de rires, de montée intonative, et de gestualisation), car le co-locuteur produit très peu de *backchannels* pendant ces séquences. Schiffrin (1987), puis Fox Tree et Schrock (2002) remarquent bien que certains marqueurs ne sont que superficiellement tournés vers le co-locuteur, et servent davantage le discours du locuteur. C'est le cas de cet usage de "like". Dans ces séquences, il est très intéressant d'observer que l'ouverture ou la délimitation d'un

premier espace mental avec une circonstancielle s’accompagne de l’ouverture d’un second, dans lequel les locuteurs se mettent en scène à l’intérieur de la situation de discours qu’ils décrivent. Dans les deux cas, le co-locuteur reçoit des instructions quant à l’interprétation d’un ou plusieurs segments de discours. Si l’espace mental ouvert par les subordinées porte sur les conditions dans lesquelles les actions se réalisent et/ou dans lesquelles les états de faits sont vrais, l’espace mental ouvert par “like” se situe davantage aux niveaux des plans énonciatifs (on change de locuteur, de coordonnées spatio-temporelles) et discursifs (on passe du discours à une représentation de discours).

33.3 % des “like” insérés à la suite d’une circonstancielle remplissent une fonction de *filler* pendant des disfluences, à l’image de l’exemple (24) plus bas. Les circonstanciels semblent occasionner de légères disfluences, que le locuteur n’a visiblement pas de mal à traiter, puisque les séquences comprenant des circonstanciels ne sont pas caractérisées par des marques de disfluences plus majeures, tels que des abandons ou des répétitions de structures. Les disfluences sont légèrement plus marquées dans les circonstanciels en “if” (54.5 % des disfluences), et correspondent à des hésitations ou pseudo-hésitations, comme dans l’exemple (24).

- (24) Alex especially like
 L in London
 SC **if you’re paying London rent**
 R it’s gonna be like what #
 (h) six hundred pounds a month

Dans cette séquence argumentative, Alex explique qu’à Londres, les prêts étudiants ne suffisent pas pour vivre, et qu’il faut généralement articuler un travail aux études. Elle utilise une circonstancielle de type “prédictive” (Dancygier et Sweetser, 2000), en mettant en place un type de raisonnement qui s’appuie sur la création d’un cadre/espace mental hypothétique, et la prédiction des conséquences de cet espace, basée sur les connaissances du monde, les informations introduites précédemment dans le discours, et les relations typiques de cause à effet. SC montre l’emploi d’une forme verbale au présent, et R comporte un semi-modal (i.e. “gonna”). Cette configuration aspectuelle et temporelle est prototypique de cet usage, selon Fillmore (1990). Le “like” énoncé sur R a un double rôle. Ce marqueur sert d’abord de remplisseur et donne à Alex le temps d’estimer le prix d’un loyer à Londres, combiné à une pause silencieuse, puis à une reprise de souffle audible. Alex rompt le rythme de son discours, mais marque sa volonté de garder le tour de parole. Si ce segment de discours est montré comme mobilisant des efforts de réflexion particuliers de la part de la locutrice, il permet surtout à Alex de mieux projeter le discours

qui va suivre (Buchstaller, 2001). D’ailleurs, la position rhématique de cette occurrence de “like”, la forme interrogative métadiscursive “what”, et les pauses servent également à la création d’un focus sur la seconde partie du rhème de R. La réponse à la question métadiscursive d’Alex est “poussée” sous un groupe intonatif distinct par la combinaison de ces trois éléments, renforçant l’effet de dialogisme créé par Alex. “What” indique que la suggestion du référent qui suit R correspond à une approximation, mais permet aussi d’intensifier l’argument d’Alex : “six hundred pounds a month” pour le loyer rend clairement les prêts étudiants insignifiants. Le “like” de R n’est alors pas si éloigné du “like” réalisé avant le groupe intonatif de L, exemple prototypique du “like” de focus (Underhill, 1988) qui met en valeur “in London” en tant que dernier constituant du groupe adverbial.

Nous avons vu que les circonstancielles sont annoncées par une variété de marqueurs de discours procéduraux qui servent majoritairement à centrer l’attention sur le discours à venir, mais sont clôturées par un marqueur préférentiel, qui délivre des instructions d’interprétation très différentes en fonction de la séquence discursive. Du point de vue de la position des circonstancielles à l’intérieur des macro-séquences, 60 % des occurrences sont préfixées à la structure principale (84.6 % des conditionnelles, 54 % des temporelles), comme dans la séquence (25). Cette proportion est significativement différente de celle des appositives préfixées ($F(54,54) = 13.4$, $p < 0.0001$), et des relatives déterminatives préfixées ($F(54,54) = 13.4$, $p < 0.0001$). Si la littérature est divisée sur la tendance générale pour les circonstancielles (plutôt postposées pour Miller et al., 1998; plutôt préfixées pour Dancygier et Sweetser, 2000), le fait que la majorité de notre corpus se place en position initiale souligne la pertinence des circonstancielles à l’ouverture de paragraphes discursifs. Le locuteur choisit très fréquemment ce format syntaxique pour orienter son discours de manière prospective.

(25)	Alex	it’s really funny cos hem	
	L	(h) you know	
	SC	when you’re not allowed to laugh #	[PREF]
	R	(h) and then there’s like a massive silence # especially in class when someone says something funny and i’m always # dying for ages it makes it so much worse # when you know you’re not allowed to laugh #	[NOY] [SUF1] [SUF2]

Située en début d’interaction entre Alex et Rhianna, cette circonstancielle préfixée sert de cadre à l’interprétation de presque toute la séquence. Alex démontre que la

nervosité se termine toujours en rire nerveux pour elle, en décrivant une situation à titre d'exemple qu'elle-même a vécu de nombreuses fois. L'opérationnalité du *nextness principle* (Ochs, 1979, p. 62), principe de proximité, est limitée ici, puisque l'interaction démarre : peu de contributions informationnelles servent de guide à la co-locutrice pour interpréter le discours d'Alex. Par sa position initiale, SC indique que la co-locutrice doit intégrer un nouveau segment dans sa représentation du discours, ce qui lui permet d'éviter la recherche de continuité (par exemple par la recherche dans les informations situationnelles, puisque les informations discursives sont ici limitées) et diminue le coût de traitement de l'information. En délimitant une portion de texte, cette circonstancielle signale à la fois l'ouverture d'un nouvel espace et la fermeture (ou du moins la suspension de validité) d'un espace précédent ou concurrent (Combettes, 2005), comme celui formé par les informations contextuelles et situationnelles précédentes et contemporaines au tour d'Alex. À l'inverse, la circonstancielle est répétée à la fin de la séquence discursive, en position de suffixe avec une intonation de *trailoff* (Local, 1992; très basse et très plate). Cette occurrence n'ouvre aucun nouvel espace cognitif, et clôture seulement la séquence en abandonnant le tour de parole.

Avec cette majorité de circonstanciels préfixés à la macro-séquence, ce format syntaxique n'est pas le premier choisi par les locuteurs du corpus pour une action de *self-monitoring* (Clark et Krych, 2004), i.e. une action de modification en temps réel de leur propre discours, par un segment aussitôt apposé à la suite de celui qu'ils souhaitent modifier. La position médiane, considérée comme la plus intrusive puisqu'elle interrompt le segment noyau, caractérise seulement 4 % des circonstanciels. Si les locuteurs estiment que l'interprétation de leur segment de discours n'est pas optimale pour le co-locuteur, la modification des espaces mentaux se réalise donc en fin de macro-séquence. Face à cette particularité, nous pensons à l'hypothèse de Dancygier et Sweetser (2000), selon laquelle les constructions insérées en milieu de séquence violent certains principes de la structuration des espaces mentaux/cognitifs, qui est traditionnellement prospective. Les auteurs considèrent les constructions postposées comme de simples élaborations sur des espaces mentaux déjà existants. Il arrive que certains espaces mentaux soient construits d'emblée à l'aide du contexte (situationnel, représentationnel, discursif), et qu'une proposition structurant explicitement l'espace mental serve, du point de vue du locuteur, à clarifier, confirmer, ou modifier un espace mental déjà construit. En raison de leur nombre relativement important dans notre corpus, nous préférons parler dans cette première partie de la préférence des locuteurs pour la "progressivité" dans le discours et l'interaction (Stivers et Robinson, 2006), par rapport au rythme discontinu

que créent les insertions en position médiane, et par rapport au retour cognitif que créent les postpositions.

Nous avons vu que les circonstancielles ouvrent facilement de nouveaux espaces cognitifs et modulent l'interprétation du discours. 45.4 % d'entre elles changent de participant ou d'élément thématique. Nous ne considérons pas ce trait comme spécifiquement propre aux circonstancielles, puisque les relatives déterminatives présentent la même proportion de segments changeant d'élément de départ. Ce phénomène s'observe toutefois aussi bien sur les temporelles (46.4 % des circonstancielles temporelles présentent ce type de rupture) que les conditionnelles (44.4 % des circonstancielles conditionnelles). L'exemple (26) survient juste après une séquence argumentative, dans laquelle Joey a avancé que quatre de leurs amis se disputent particulièrement souvent.

- (26) Joey but it's all okay now
 cos it's all forgotten but #
 there's always like a bit of a #
 Elena oh yeah
 Joey L and that's happened before as well
 SC **when # Holly and Beth were really drunk #**
 R and Gemini drove them home
 in Holly's car

Joey initie un récit à partir de L, qui permet d'élaborer sa thèse argumentative (leurs amis se disputent beaucoup). L et les segments de discours précédents utilisent un élément thématique général ("it's", "there's", "that's") à la manière de plusieurs constats successifs, où seuls les constituants rhématiques des segments de discours importent. En revanche, SC entre dans le récit avec la description d'une scène : "Holly and Beth" sont clairement identifiées, et sont liées au complément du sujet "really drunk". Si un changement d'élément thématique s'opère, Joey fait en sorte de garder son tour, en réalisant une pause silencieuse démarcative seulement après avoir initié la circonstancielle, à la suite de la conjonction "when". Cette pause permet de retarder le commencement du récit, créant un horizon d'attente pour la co-locutrice. Il s'agit de la phase d'orientation du récit de Joey (Labov et Waletzky, 1967), donnant les informations nécessaires à la co-locutrice (personnes, temps, et situation comportementale; *ibid.*) avant de décrire les événements (e.g. "Gemini drove them home"). Après les constats successifs de Joey, SC initie un compte-rendu de manière beaucoup plus frappante en représentant directement les référents en tant que protagonistes du récit, alors qu'ils étaient jusqu'ici objets de constats. L'adverbe intensifieur "really" est une autre preuve du rôle de SC par rapport au but discursif de Joey. Cet adverbe donne une valeur modale à la phase d'orientation,

exposant habituellement des détails objectifs de manière factuelle. Le fait que cette valeur modale porte sur le degré de validation de la relation prédicative désamorce une plus grande prise de position (e.g. qualitative), mais renforce la pertinence de l'argumentation de Joey.

En revanche, les circonstancielles sont également caractérisées par une grande propension à continuer les chaînes référentielles déjà existantes : à l'image de l'exemple (27), 69.1 % d'entre elles présentent au moins un pronom entretenant une relation anaphorique avec le co-texte. Cette proportion est remarquable dans le sens où les circonstancielles montrent plus de continuité pronominale que les relatives déterminatives, type syntaxique exprimant le moins de rupture verbale. Cette proportion est également significativement différente des appositives ($F(54,54) = 2.2$ $p < 0.005$).

- (27) Elena L oh it's really shiny
 SC **when it's all straight**
 Joey mh
 Elena R and like it looks really nice

Elena livre un récit à propos du lissage de ses cheveux chez le coiffeur, et relate le point de vue de la personne l'ayant coiffée. SC, comme L et R, présente le pronom "it" en élément thématique, recouvrant le topique de la séquence discursive, mentionné au début de celle-ci par Elena : "my hair". Le pouvoir cohésif des circonstancielles par la référence à travers des formes lexicales réduites comme les pronoms va dans le sens des observations de Blüdhorn (2008) à l'égard de ce type syntaxique. Selon lui, les connecteurs adverbiaux n'expriment majoritairement pas de liens envers leur co-texte par la rection grammaticale ou la séquence linéaire, mais connectent surtout deux unités sémantiquement, par le biais de la référence. Ce procédé permet de lier le contenu de la circonstancielle à un référent précédent (que ce référent soit celui qui précède immédiatement la circonstancielle ou un référent mentionné plusieurs tours auparavant; Mazeland et Huiskes, 2001) ou à un référent ultérieur. En plus d'impliquer ce référent sous la même forme réduite dans un rôle identique, cette circonstancielle s'intègre remarquablement dans la séquence, ne présentant aucune disjonction temporelle ou de rupture par rapport au type de la séquence. SC apporte un élément descriptif comme L et R, et l'intégralité de la séquence est réalisée au présent simple. L'espace de projection cognitive est le même pour toute la séquence, l'ensemble ne comportant aucune pause ou disfluente intra-constituant.

En somme, les circonstancielles sont caractérisées par un statut

syntactico-sémantique intermédiaire. Le fait que la dépendance pragmatique soit dissociée de la dépendance syntaxique dans une majorité de cas est intéressant, car caractéristiques pragmatiques et caractéristiques syntaxiques ont longtemps été spécifiquement associées (Lehmann, 1988; Matthiessen et Thompson, 1988) pour décrire les circonstancielles. Leurs marqueurs morphologiques peu variés ne correspondent pas forcément aux liens sémantiques qu'ils établissent traditionnellement entre deux unités. Au niveau discursif, les circonstancielles se distinguent clairement par leur rôle sur le plan interprétatif, ouvrant et fermant des espaces cognitifs valides pour un ou plusieurs segments de discours. Cette capacité à moduler l'interprétation des co-locuteurs est visible à travers leur position et les marqueurs de discours "particule" qui caractérisent la séquence. Les circonstancielles sont en revanche très intégrées du point de vue de la référence (montrant un grand pouvoir cohésif), et de leur contribution aux séquences hôtes. Les co-locuteurs ne sont pas extrêmement réceptifs à ce type de construction, du point de vue de la faible part de *backchannels* qu'elles suscitent. Les indices verbaux ne sont pas assez nombreux pour parler d'autonomie, mais sont de telle nature que l'on ne peut parler d'intégration. Les circonstancielles recouvrent une remarquable flexibilité de fonctionnement. Ni récursives, ni autonomes, nous concluons que les circonstancielles présentent une stratégie d'insertion intermédiaire, en présentant un changement dont la valeur est plus cognitive que syntaxique. En guidant l'interprétation et en projetant des unités discursives, elles œuvrent à la construction de zones d'alignement en optimisant le traitement du discours.

En ce qui concerne l'articulation des différents indices au sein d'une même séquence, une seule corrélation est observable. Les verbes de procès ont tendance à apparaître dans les circonstancielles ouvrant une séquence cadrative dans le discours, plutôt que dans celles clôturant un espace mental : 71.8 % des circonstancielles comportant un verbe de procès marquent l'ouverture d'un nouveau contexte interprétatif. Cette configuration concerne 41.8 % des circonstancielles de notre corpus, et concerne autant les conditionnelles (20.7 % des circonstancielles) que les temporelles (21.1 % des circonstancielles). Elle est illustrée par la séquence (28) :

- (28) Joey it's always Gemini and Charlotte
 against Holly and Beth #
 Elena mh #
 Joey L cos someone #
 SC **when they had that party #**
 R do you remember that party on #
 couple of weeks ago

Bien que SC comporte un élément thématique anaphorique (“they” renvoyant à “Gemini and Charlotte” et à “Holly and Beth”), le complément d’objet direct “that party” ne fait pas référence à un objet de discours dont les participantes ont déjà discuté. Le pronom démonstratif distal “that” participe à la mise en focus de ce nouvel élément. Joey s’assure d’ailleurs par une question directe sur R que ce nouvel espace mental est activé chez Elena. Elle fait de nouveau référence à “that party”, en l’ancrant dans le discours grâce à la locution adverbiale (“on # couple of weeks ago”) qui lui est attachée. SC comporte un syntagme verbal plutôt dynamique : le procès < have a party > peut être borné dans le temps et le syntagme verbal est réalisé au prétérit. Les syntagmes verbaux de L et R contrastent avec cette configuration, puisque L comporte un verbe copule dénotant un état et R montre un verbe cognitif. Le fait que les verbes de procès apparaissent en co-occurrence avec l’ouverture d’espaces cadratifs dans nos circonstancielle suggère que les locuteurs privilégient la dynamique des syntagmes verbaux pour introduire des repères discursifs qui vont jouer un large rôle textuel, en restant valides pour plusieurs segments de discours. Ici encore, la valeur du changement se joue sur le plan cognitif. Selon Fauconnier (1991), l’utilisation du prétérit dans la délimitation d’espaces cognitifs permet non seulement de spécifier un lien temporel entre les relations prédicatives du segment cadré et du segment cadratif, mais aussi d’établir une perspective. La fonction de ces marques serait d’établir une “configuration complexe de domaines dans le discours et de structurer cette configuration à plusieurs niveaux simultanément” (p. 245). Une autre observation se trouve en lien avec nos résultats. Pour Dancygier et Sweetser (2000), les circonstancielle en “when” témoignent d’un fort engagement de l’énonciateur quant à la réalité de l’espace décrit par cette circonstancielle. La dynamique des syntagmes verbaux et leur encodage par du prétérit ne font alors que renforcer cet engagement. Nous pourrions alors imaginer une relation entre l’engagement du locuteur quant à la réalité décrite dans la circonstancielle et la facilitation d’ouverture d’un espace mental dans le discours. En revanche, selon ces mêmes auteurs, les conditionnelles en “if” expriment le manque d’engagement épistémique positif du locuteur quant à la réalité décrite par la circonstancielle. Or, nous observons autant de dynamique dans les prédicats des conditionnelles. Si la tendance de notre corpus va dans le sens de l’hypothèse de Fauconnier (1991) mais reste trop faible pour justifier toute assertion, l’ensemble de nos observations sur les circonstancielle de notre corpus suggèrent que certains aspects grammaticaux et morphosyntaxiques servent de véritables balises discursives, plus particulièrement dans le déploiement et la gestion des espaces mentaux (Col, 2010). Les circonstancielle de notre corpus témoignent à la fois de la pré-construction du discours et de l’émergence de certaines structures : les locuteurs conçoivent leur séquence prospectivement en fonction du

but discursif qu'elle vise à réaliser. Ils choisissent à cet effet la position des informations présentées pour orienter la compréhension du co-locuteur. Ces choix sont sensibles aux contingences de leur production, et peuvent être modifiés à tout moment.

3.2.3 Appositives

Totalisant sept indices de rupture, les appositives se démarquent clairement et témoignent d'une autonomie syntaxique comme discursive. Les indices montrent toutefois plus de variation au niveau du discours.

Macro-syntaxe

Si la majorité des subordonnées était jusque-là concernée par une intégration ou un statut syntactico-sémantique mixte, les appositives marquent le passage vers l'autonomie, souvent en contradiction avec la lecture descriptive de leur marqueur introducteur.

Bien qu'une majorité des appositives soient encodées à l'aide d'un pronom relatif, 49.1 % d'entre elles montrent une organisation macro-syntaxique autonome, comme dans l'exemple (29). Cette proportion est significativement différente de celle des relatives déterminatives ($F(54,54) = 2.3$, $p < 0.002$), mais pas de celle des circonstancielles.

- (29) Tim L you get an assessed piece of work
 SC **which you do on a computer**
 R using a program called author catway

Tim explique dans cette séquence descriptive la façon dont les étudiants urbanistes sont évalués. Le procès décrit par SC (< do on a computer >) n'est pas inclus la relation prédicative de L (< get an assessed piece of work >), mais est consécutif. Cet extrait pourrait être glosé en supprimant la conjonction : "You get an assessed piece of work. You do it on a computer". De même, le co-locuteur peut réagir au contenu de SC sans pour autant remettre en cause L, par une question comme "do you draw everything on a computer, then?". D'ailleurs, l'ensemble formé par le dernier item de L et par SC n'est pas remplaçable par un pronom, cette substitution vidant la macro-séquence de son sens (*"you get that. Using a program called author

catway”).

De même, les appositives comptent légèrement plus de constructions montrant un statut macro-syntaxique intermédiaire (25.4 % des constructions; 86.7 % de ces dernières sont intégrées syntaxiquement, mais forment un contenu illocutoire distinct) que d’occurrences montrant une réaction totale (23.6 % des appositives). Le statut intermédiaire est plus marqué chez les appositives relatives que chez les appositives canoniques. Alors que 45.5 % des appositives canoniques sont pleinement régies, seulement 9.1 % des appositives relatives montrent cette configuration syntaxique. L’exemple (30) montre une appositive relative dont le statut macro-syntaxique est intermédiaire.

(30)	Zoe	L	well for the train	[NOY1]
		SC	that i hope # you’re gonna get and i’ve got	[MED]
		R	is the nine o’ clock #	[NOY2]

Michelle, la co-locutrice, vient de demander à Zoe à quelle heure partait son train. Dans cette séquence “agrammaticale” où Zoe éprouve des difficultés de calcul des unités, Zoe encode SC en tant qu’élément thématique de l’état de fait principal, situé sur R. SC est introduit par une conjonction (“that”), qui ne correspond pas à l’orientation objectale (i. e. neutre) qui lui a été prêtée par la littérature (e.g. Melis 2008). SC est une unité énonciative autonome : bien que ce segment ne soit pas directement contestable par la co-locutrice (*“no, you don’t hope I’m gonna get it”; *“no, you haven’t got that train”), il vise à produire un fort effet sur cette dernière en l’incitant à prendre un billet pour le même train au plus vite. SC est alors très orienté. Cette intentionnalité diffère de celle de L et de R. Au lieu de répondre immédiatement à la demande d’information de Michelle concernant l’horaire de son train pour Londres, Zoe scinde sa contribution en deux parties (NOY1 et NOY2 – les deux parties du noyau) et rappelle le thème de son argumentation précédente (Michelle doit prendre un billet pour le même train) en position médiane (MED). Bien que l’élément référentiel “the train” occupe une fonction grammaticale de sujet sur L et R, il se trouve complément d’objet direct sur SC si l’on suit la trace du mouvement syntaxique (“I hope you’re going to get the train” devenu “the train that I hope you’re going to get \emptyset ”). Contrairement au reste de la séquence, l’énonciatrice prend position sur le contenu propositionnel en le liant fortement à “you”. SC permet alors de renvoyer à la situation interactionnelle, puisque Zoe a déjà encouragé Michelle à prendre un billet.

La majorité des appositives relatives est donc caractérisée par un phénomène de

“greffe”¹³ d’une structure à une autre (Deulofeu, 2010), plutôt interprété comme une rupture de construction, dans le sens où elle constitue une violation des contraintes de sous-catégorisation d’ordinaire imposées par une catégorie sur les éléments qu’elle régit : les appositives relatives, bien que mises en relation avec une autre prédication, “débordent” de l’organisation micro-syntaxique de cette première prédication.

Si les appositives canoniques comptent davantage de constructions pleinement régies lorsqu’elles sont comparées aux appositives relatives, 50 % d’entre elles représentent des actes illocutoires distincts, comme dans la séquence (31).

- | | | | |
|------|-------|----|---|
| (31) | Joey | | and yeah Crickhowell is really nice
it’s got hem
i was going for charity shop there |
| | | L | they’ve got some really nice crafty shops |
| | | SC | gifted shops # |
| | Elena | | ah yeah |
| | Joey | R | yeah # |

Joey fait allusion à quelques boutiques de Crickhowell, hameau du pays de Galles connu des deux participantes. SC peut faire l’objet d’une remise en question métadiscursive de la locutrice (“are they really gifted shops though?”), ou de la co-locutrice (“yes but they’re not really gifted shops, are they?”). Il semble donc que l’association de modalité qualitative à une prédication implicite (ici, < Crickhowell’s “nice crafty shops” BE gifted shops >; Neveu, 2000; Heringa, 2007) aide la construction appositive à acquérir une force illocutoire propre.

La rareté des appositives canoniques totalement autonomes (3.6 % des canoniques) s’explique par leur agencement syntaxique. Pour gagner en autonomie, elles doivent en effet à la fois modifier un groupe nominal périphérique de la proposition les précédant, tel qu’un complément d’objet indirect ou un complément circonstanciel, et témoigner d’une forte prise de position sur ce constituant, comme dans l’exemple (32).

- | | | | |
|------|----------|----|---|
| (32) | Michelle | L | i was living with this Scottish girl called Nicola # |
| | | SC | Nicola (laughs) |
| | | R | (h) and just upstairs
we had the two hem surveillantes and # |

Cette construction appositive est remarquable pour plusieurs raisons. Elle forme un constituant périphérique à la structure syntaxique de L, puisqu’elle constitue une

13. La greffe est le procédé qui consiste à remplir des positions syntaxiques à l’aide d’une autre catégorie que celle attendue (Deulofeu, 1999).

répétition apportant une équivalence sémantique à “this Scottish girl” qui fait partie d’un complément circonstanciel. Elle n’est pas non plus contrôlée par le syntagme verbal de R. La seconde raison est encore plus saisissante : l’acte de langage implicite qu’elle véhicule n’est pas encodé verbalement, mais bien vocalement, Michelle imitant d’une façon hyperbolique l’accent écossais. Ce mouvement provoque le rire des deux participantes, et pourrait tout à fait faire l’objet d’une remise en question humoristique de la part de la co-locutrice, du type “oh come on, it could’t be that strong” ou bien “it’s not even a Scottish accent”. Le mouvement prosodique emphatique réalisé par Michelle, placé sur la seconde syllabe de “Nicola”, est visible dans la figure 3.3.

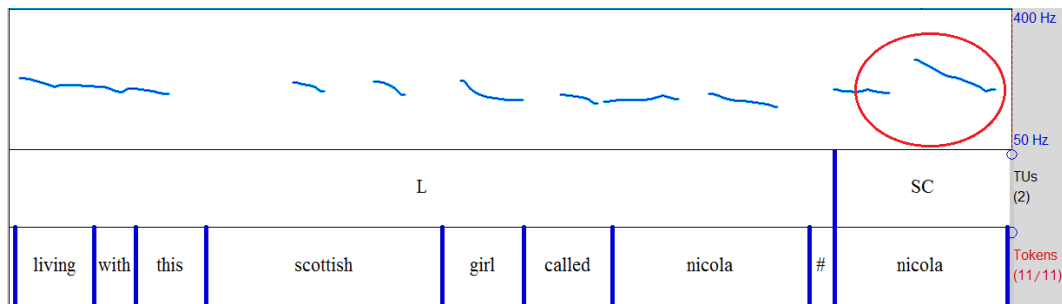


FIGURE 3.3 – Mouvement prosodique emphatique réalisé par Michelle sur la seconde occurrence de “Nicola” dans l’exemple (32), illustré par la courbe prosodique dans Praat (la première piste de transcription montre les segments - L, SC - dans la séquence, la seconde piste montre le découpage de la séquence en tokens lexicaux).

En ce qui concerne les morphèmes introducteurs des appositives, 90.9 % des appositives relatives sont introduites par les pronoms relatifs “which” (81.8 % des appositives relatives) et “who” (9.1 %), dont le statut est intermédiaire. Contrairement à “that” (introduisant 6.1 % des appositives relatives) et “∅” (3 %), ces éléments ont la capacité d’établir un lien non plus objectal mais situationnel : ils permettent un pivot dans la construction syntaxique et discursive, souvent caractérisé par un changement des rôles grammaticaux ainsi que des paramètres énonciatifs. Les mêmes référents sont impliqués dans d’autres relations par le biais du pronom relatif, permettant un “suivi référentiel” (i.e. le *reference tracking* de Fox et Thompson, 2009), comme dans l’exemple (33) :

- (33) Beth L and then we went into # a place called Tropicana Comp. circonstanciel
 SC **which was horrible (laughs) #** Sujet
 Kate where is that
 R it’s on Saint Mary street Sujet
 near the castle

Par le biais de “which”, l’élément référentiel de L passe de complément circonstanciel de lieu à sujet sur SC. SC construit alors un autre type de relation à SC que celui que nous avons vu jusqu’ici : bien que l’on observe une reprise des référents, le focus est mis sur l’introduction d’un point de vue que l’on lie à ces éléments référentiels par le changement de rôle grammatical. L’élément référentiel modifié garde ensuite la fonction grammaticale prise sur SC, mais ne suscite plus de position énonciative.

Les appositives canoniques constituent le seul groupe à ne pas être introduit par un marqueur verbal : elles ne sont établies par aucune conjonction (coordonnante ou subordonnante), pronom relatif, ou marqueur de discours. Nous pouvons alors parler de juxtaposition des syntagmes nominaux en apposition dans la chaîne parlée. Si 63.6 % des appositions canoniques sont placées immédiatement à la suite du constituant qu’elles modifient, certaines constructions sont un peu plus éloignées de leur antécédent. Un éloignement important entre modifieur et modifié cause d’ailleurs des difficultés de traitement et/ou d’élocution sur R, comme dans la séquence (34).

- (34) Joey but the thing is
 the guy packed the bag
 L and hem # he
 SC **like the Sainsbury’s man (laughs)**
 R and hem # the pizza bases have got a crack in them

Deux syntagmes sujets (i.e. “he” sur L, puis “the Sainsbury’s man” sur SC) sont ajoutés à “the guy” utilisé lors de la construction initiale. Joey anticipe une question potentielle de son interlocutrice par un phénomène de *self-monitoring* (Clark et Krych, 2004, p. 64) : Elena pourrait trouver que la pertinence de “the guy” n’est pas optimale. Ces ajustements portant sur le participant d’un procès comporte des points communs avec une opération de thématisation, en présentant une caractéristique du sujet grammatical qui a pourtant déjà été inscrit à l’origine d’un processus. Avec les disfluences marquées par “hem” et une pause silencieuse, Joey reprend le fil de son récit sur R en changeant de sujet grammatical. Par son détachement, l’apposition a du mal à se constituer en caractérisation actancielle : Joey indique avec difficulté que “the Sainsbury’s man” étend le poste actanciel défini par “the guy”, en utilisant “like” comme marqueur de réparation.

Si les appositives canoniques modifient strictement un groupe nominal, la portée des appositives relatives est plus variable. 33.3 % des appositives relatives modifient l’intégralité de la proposition précédente, au lieu d’un seul syntagme. L’exemple (35) illustre ces appositives relatives dont l’antécédent est moins restreint.

Dans une discussion autour du permis de conduire, Rhianna explique que sa grand-mère paternelle a décidé de payer les leçons de conduite de ses petits-enfants. L'insertion de l'appositive dans la séquence provoque un dédoublement du syntagme sujet apparaissant dans L (Kim, 2012). En complétant la partie gauche (i.e. pré-verbale) de l'énoncé de Rhianna, ce groupe nominal en apposition retarde l'apparition du syntagme verbal, qui est réalisé sur R. Le dédoublement du sujet semble avoir engendré un coût cognitif important, puisque Rhianna abandonne l'élaboration de la structure propositionnelle pour en débiter une nouvelle, à l'aide du marqueur de discours résomptif “basically”. Le sujet désigne le même item référentiel, mais est cette fois réalisé à l'aide d'une forme lexicale réduite (“she”).

Les appositives présentent également la plus grande proportion d'occurrences réalisées en position macro-syntaxique médiane, avec 24 % de ces constructions. Ce phénomène s'observe davantage sur les appositives canoniques (avec 61.5 % des occurrences en position médiane) que sur les appositives relatives (38.5 %). L'apposition dans la séquence (37) suivante est insérée dans la description que fait Joey de son jardin :

(37) Joey		but he said #	[NOY1]
	L	Keith #	[NOY1]
	SC	the guy that lives there #	[MED]
	R	said he'd help us	[NOY2]
		doing the stuff #	[NOY2]

La caractéristique du sujet “the guy” encodée par le syntagme verbal de SC est insérée entre deux arguments (séparés chacun en deux groupes intonatifs) de la prédication matrice du mouvement discursif principal. Deux syntagmes sujets sont ajoutés au pronom “he” utilisé lors de la construction initiale. Ces occurrences successives témoignent de la mise en place relativement laborieuse d'un élément thématique, faisant l'objet de plusieurs ajustements rapprochés où les coordonnées énonciatives sont recalculées à chaque assertion. Cette manipulation portant sur le participant d'un procès se rapproche d'une opération de thématisation, en présentant une caractéristique du sujet grammatical mais ne l'inscrivant pas à l'origine d'un processus.

Par ailleurs, 76.4 % des appositives créent au moins l'un des trois types de rupture que nous avons identifiés. Ce résultat n'est pas significativement plus important que les autres formats syntaxiques, mais il n'en reste pas moins très élevé. 44 % des appositives créent un décalage par rapport aux attentes liées au type de séquence dans lesquelles elles sont inscrites. Cette proportion représente significati-

vement plus que les relatives déterminatives ($F(54,54) = 3.6, p < 0.0001$). Certaines de ces constructions créent même une nouvelle séquence insérée dans la première, comme dans l'exemple (38).

- (38) Tom i really enjoy Alpha
 L i look forward to it every week
 SC **which is how we were feeling last year**
 when we were on it
 R we'd get to like # a Saturday
 go play football together

Cette séquence débute avec le récit de Tom à propos d'un groupe de discussion. Il cite les propos tenus par l'une des participantes à l'organisateur, un ami de Tom qui cherchait à savoir si le groupe de parole était apprécié. Au lieu de continuer le récit de la perspective de son ami et de la participante, il change de point de vue et de temporalité pour expliquer que c'était l'état d'esprit dans lequel lui-même était lorsqu'il participait à ce groupe. R n'est donc pas la continuation du premier récit, mais le commencement d'un second, centré sur Tom et ses amis et enchâssé dans le premier.

En ce qui concerne les emplacements de rupture pour les appositives, ces constructions présentent la proportion la plus importante de marqueurs de discours de structuration annonçant une transition (31 % des séquences contenant une appositive comportent au moins un marqueur de ce type). Cette proportion n'atteint toutefois pas le seuil de significativité. Comme pour les relatives déterminatives dans une moindre mesure, ces marqueurs de discours suivent la construction subordonnée, et ouvrent R (71.4 % des marqueurs de discours de structuration dans les séquences hôtes des appositives se trouvent sur R). Deux marqueurs sont récurrents : "and" représente 51.4 % des occurrences à l'image de l'exemple (39), et "but" correspond à 34.3 % des marqueurs, comme le montre l'exemple (40).

- (39) Michelle L i was living with this Scottish girl called Nicola #
 SC **Nicola (laughs)**
 Zoe (laughs)
 Michelle R (h) and just upstairs
 we had the two hem surveillantes and #
 Zoe well was she an assistant as well

Michelle décrit son année en tant qu'assistante en France. Alors que L et R représentent des contributions factuelles portant sur l'agencement du bâtiment, SC juxtapose un trait humoristique vocal au syntagme nominal "this Scottish girl called Nicola". Michelle mime un accent écossais, caractérisant le référent qu'elle vient

d'introduire dans le rhème de L à l'aide du démonstratif proximal "this". Cette imitation apporte de la modalité à la séquence et introduit une variation dans cette description plutôt homogène. Cette contribution est marquée du rire des deux participantes. Afin de conserver son tour de parole et de reprendre le fil de sa description pour avancer son agenda conversationnel, Michelle combine une reprise de souffle audible au marqueur continuatif "and" en début de R. "And" agit donc à la suite de l'appositive en tant que *pop marker* (Redeker, 2006, p. 346), en fermant l'espace ouvert par SC et en indiquant une transition : le début d'un autre segment de discours (Grosz et Sidner, 1985)¹⁴. R retrouve une neutralité énonciative et décrit un autre endroit, avec l'introduction de nouveaux référents dans le segment suivant R. Ce segment comporte également un "and", qui joue un rôle très différent dans le discours de Michelle : celui-ci clotûre son discours et abandonne le tour de parole. Cette occurrence correspond au *trailoff* décrit par Local (2007), dont la position au sein du segment de discours, le manque de modulation, et le manque de hauteur intonative indiquent la fin d'une unité conversationnelle. Ces deux fonctions très différentes du même marqueur sont visibles sur la courbe intonative de R et du segment qui suit dans la figure 3.4; alors que la première occurrence de "and", le *pop-marker*, présente un contour très distinct du reste et modulé, le *trailoff-and* est plat et allongé, indiquant une frontière prosodique imminente.

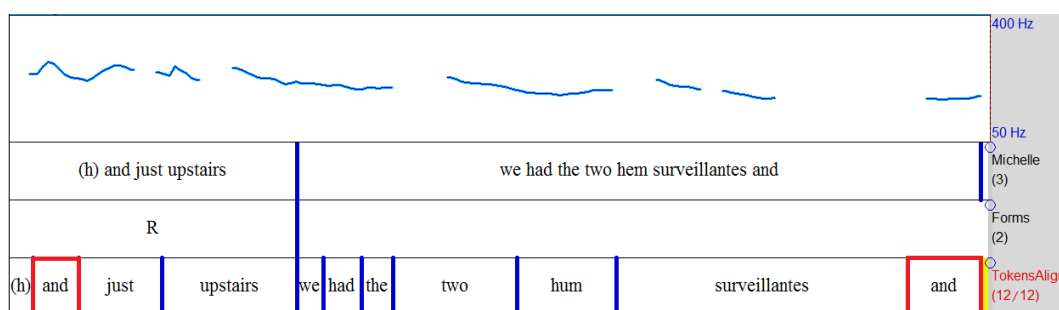


FIGURE 3.4 – Deux occurrences dans l'exemple (39) du même marqueur de discours "and" illustrées par la courbe Praat, montrant deux réalisations prosodiques très différentes.

L'exemple (40) illustre l'utilisation du marqueur contrastif "but" (e.g. Fraser 1999), également très présent autour des appositives.

(40) Tom L it's got blue carpet down now segment de discours 1

14. Selon le schéma de structure discursive de Grosz et Sidner (1985), le discours est constitué d'une série de séquences, chacune d'elle accompagnée d'un plan communicatif avec des objectifs et intentions précises. Ces intentions sont divisibles en de plus petites unités hiérarchisées, accompagnant chaque sous-topique : les *Discourse Segment Purposes*.

SC	which is supposed to be all s-all-purpose	segment de discours 2
R	<u>but</u> # you can't skate on it obviously #	<u>segment de discours 3</u>

Cette appositive relative est insérée au début d'un récit, où se trouve un court résumé de la situation que Tom s'apprête à raconter en détails : la salle de sport municipale a un jour été recouverte d'un revêtement rendant impossible toute partie de hockey; ce changement a provoqué une grande réaction chez le locuteur et ses amis. Ce court extrait montre une structure argumentative, où un fait est exposé (la salle est désormais recouverte de "blue carpet"), sa cause est évoquée (les salles doivent toutes progressivement devenir "all-purpose"), et les conséquences pour Tom et ses amis sont développées : on ne peut plus patiner dans cette salle. Plus que la mise en opposition entre deux idées (la salle est désormais multisports, mais exclut le patin; "but" exprime un contraste entre SC et l'une de ses interprétations, qui est éliminée), "but" indique un changement de repère énonciatif pour la prédication : alors que SC est factuel, R correspond à une assertion propre à l'énonciateur (Tom et ses amis), montrant alors un conflit d'intérêt/d'interprétation de la situation entre les deux groupes. R constitue bien un mouvement discursif autre que L, en avançant un point très distinct de la structure argumentative. Tom élabore donc trois segments discursifs différents dans cet extrait, correspondant chacun à un but discursif. "But" marque également un retour au topique conversationnel (i.e. le patin), après une explication inhérente à l'item référentiel récemment introduit (i.e. le "blue carpet"). Si cet objet est l'élément perturbateur du récit, "but" initie le retour au cœur du sujet de la narration, en liant une caractéristique du nouvel item référentiel au topique de la séquence. Ce marqueur est donc utilisé ici pour sa fonction cohésive, en liant un objet récemment inséré au fil principal du discours.

Du point de vue du co-locuteur et de ses signaux de réponse, les appositives comportent la plus haute proportion de *backchannels* substantiels¹⁵ (i.e. les *backchannels* dont la fonction n'est pas seulement d'encourager le locuteur à continuer son tour en montrant sa compréhension, mais aussi de montrer une affiliation¹⁶ au point de vue modal du locuteur sur son discours, de produire une évaluation du discours ou un jugement, ou de réaliser une demande de précision; Bertrand et al., 2007; Bertrand et al., 2009). 22 % des appositives sont suivies d'un *backchannel* substantiel

15. Contrairement aux *backchannels* minimaux, les substantiels ne consistent pas en un simple item verbo-vocal du type "yeah", "mhmh", "ok" ou "alright" (que cet item soit couplé ou non avec de la gestualité, comme un hochement de tête ou un haussement de sourcils).

16. Le co-participant montre son affiliation interactionnelle en adoptant la position modale attendue par le locuteur et/ou en reconnaissant l'importance de la nouvelle information présentée (Stivers, 2008).

du locuteur, ce qui représente un taux relativement important. Cette proportion est significativement plus importante que celle des relatives déterminatives ($F(54,54) = 2.5$, $p < 0.0005$), et que celle des circonstancielles ($F(54,54) = 2.1$, $p < 0.005$). Les relatives déterminatives, apparaissant également systématiquement à la suite de leur antécédent, se voient ponctuées de *backchannels* minimaux. Les appositives relatives, contenant beaucoup de prise de position modale, élicitent des *backchannels* contenant également une prise de position, répondant (positivement ou négativement) à celle du locuteur. L'un des facteurs expliquant cette tendance pourrait être lié au phénomène de convergence interactionnelle¹⁷ (e.g. Bavelas et al., 2000; Bertrand et al., 2009; Holler et Wilkin, 2011).

Cependant, une nette disparité s'observe : 91.7 % de ces *backchannels* substantiels sont produits à la suite des appositives relatives. Cette disproportion s'explique facilement, puisque les appositives relatives sont beaucoup plus à même de représenter des actes illocutoires et d'encoder un contenu modal. Il est alors plus facile et naturel pour le co-locuteur de prendre position par rapport à un contenu orienté, qui de plus est encodé par une proposition et non pas par un seul groupe nominal. De même, les appositives relatives contiennent un verbe fini, indicateur de la complétion d'une unité syntaxique (Koiso et al., 1998). Les canoniques, elles, sont clairement des syntagmes projetant d'autres unités syntaxiques à leur suite. Nous avons vu que les contours intonatifs jouent également un rôle dans la production de *backchannels* (e.g. Benus et al., 2007). Canoniques et relatives présentent le même type de contour majoritaire (i.e. descendant-montant), mais contrairement à ce que nous attendions, ce contour ne participe pas à cette grande production de *backchannels* substantiels. Le contour descendant-montant possède une charge pragmatique complexe, confiant des implications au référent associé au contour. Le co-locuteur doit alors recontextualiser ce référent, et aurait pu indiquer au locuteur que le référent a bien changé de statut dans sa représentation mentale, ou bien prendre position par rapport à la charge pragmatique qu'il implique par un *backchannel* substantiel. Or, seulement 25 % des *backchannels* substantiels en question sont produits à la suite d'une appositive présentant un contour descendant-montant. Au contraire, 75 % de ces *backchannels* suivent un contour soit terminatif (i.e. descendant; 50 %), soit emphatique (i.e. montant-descendant; 25 %). Sur notre corpus, les appositives au contour terminatif sont donc plus suivies de *back-*

17. La convergence interactionnelle correspond au phénomène de répétition (*mimicry*, *alignment* ou *accomodation*) dans l'interaction, par lequel les participants indiquent la prise en compte des contributions du co-participant en produisant des indices ou schémas de contributions similaires, et favorisent ainsi la co-construction de l'interaction par l'alignement des représentations à de multiples niveaux (Bertrand et al., 2009).

channels substantiels; le co-locuteur perçoit SC comme une unité potentiellement achevée. Si aucun *backchannel* substantiel n'est produit après un contour montant, ce phénomène est lié aux caractéristiques prosodiques générales des appositives, dont seulement 9.1 % sont caractérisées par un contour final montant.

Alex produit l'exemple (41) dans un récit à propos des trains Eurostar, dont Rhianna, la co-locutrice, vient de déplorer les tarifs.

- (41) Alex i got my # hem return journey
 L for fifty pounds
 SC **which was really amazing #**
 Rhianna [NOD yeah it's not bad]
 Alex R but then that was #
 it was really fun-
 uh i- ironically
 it was the weekend
 when they had all the snow on the lines

L'appositive relative propose une évaluation du prédicat formé par les deux groupes intonatifs "I got my # hem return journey for fifty pounds". L'item final de L "fifty pounds" représente clairement un appel au co-locuteur, présentant un contour final montant, visible dans la figure 3.5. L'appositive est au contraire réalisée avec un contour descendant, et est suivie d'un *backchannel* de Rhianna.

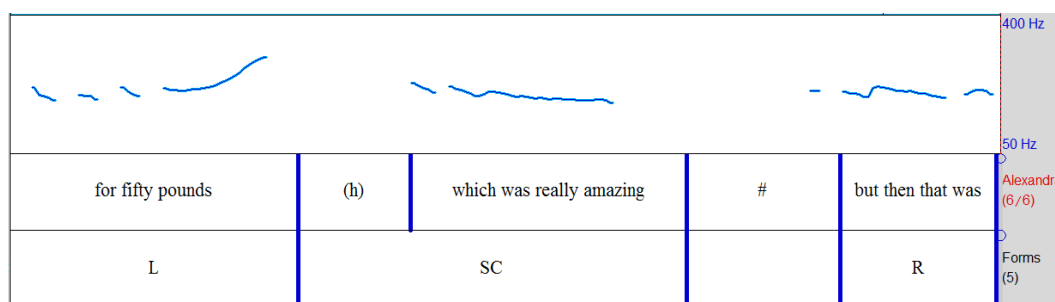


FIGURE 3.5 – Contour final descendant sur l'appositive relative d'Alex dans l'exemple (41), illustré par la courbe intonative Praat.

Rhianna produit un *backchannel* complexe¹⁸, comportant des éléments verbaux et gestuels, et témoigne d'une affiliation au point de vue que propose Alex sur son propre discours. Ce *backchannel* correspond à la configuration décrite par Goodwin (1996, p. 391) pour les évaluations (*assessments*), composée d'un pronom et d'un

18. Le terme "complexe", utilisé pour qualifier un *backchannel*, ne fait pas allusion à la substance plus ou moins sophistiquée du message ou à son effort de production, mais désigne simplement un *backchannel* produit à l'aide de plus d'une modalité à la fois (e.g. vocal et gestuel, verbal et gestuel) dans ce travail.

verbe copule, suivi d'un intensifieur optionnel et d'un adjectif évaluatif. Si cette occurrence ne comporte pas d'intensifieur, ces derniers sont également plutôt rares dans le corpus de conversation spontanée de Jurafsky et al. (1998). Cependant, la modalité qualitative de l'énoncé de Rhianna ("it's not bad") ne fait pas tout à fait écho à celle d'Alex ("really amazing"). La prise de position est très atténuée : au lieu d'un adjectif dont le degré très élevé de scalarité par rapport à l'adjectif "good" est renforcé par l'adverbe "really", Rhianna utilise une tournure négative, qui renvoie à l'adjectif "bad". Ce détachement par rapport au segment très assertif et orienté (Ducrot, 1980) d'Alex lui donne une position énonciative plus neutre et mesurée, à la manière du discours d'expert. Par ailleurs, Rhianna quitte Alex du regard pour énoncer son *backchannel*, les yeux dans le coin inférieur gauche, puis revient immédiatement vers Alex. Elle acquiesce également en continu pendant son intervention. Le changement de direction du regard donne à Rhianna une certaine neutralité qui renforce son détachement de la situation d'énonciation, à laquelle Rhianna consacre un moment de réflexion. Le hochement de tête, lui, valide à la fois la contribution d'Alex par son couplage avec l'élément vocal "yeah" et opère une sorte de retour cognitif sur l'ensemble de la séquence : Rhianna est pleinement dans le mimétisme d'un discours d'expert, prenant le temps d'évaluer ce qu'on lui présente avec mesure. Alex récupère le tour de parole à l'aide de la combinaison de deux marqueurs de discours, "but then". "Then" fonctionne comme un déictique pour établir un point de référence temporel, alors que "but" lie les états de faits précédents à un segment imminent de manière adversative. Ces deux marqueurs indiquent, à un niveau séquentiel, la mise en parallèle d'un état de fait perturbateur qui bouleverse le cours du récit et les faits établis. Alex a des difficultés à introduire ce nouvel état de fait, interrompant sa première construction présentative pour en commencer une seconde.

58.3 % de ces *backchannels* substantiels sont, à l'image du dernier exemple (41), des évaluations. Si Tolins et Fox Tree (2014) notent que les *backchannels* substantiels n'entraînent pas le commencement d'un nouveau segment de discours par le locuteur, mais l'élaboration du segment qui vient d'être présenté, notre corpus est davantage en demi-teinte et ne permet pas de le confirmer. La production d'un *backchannel* substantiel par le co-locuteur suscite le commencement d'un segment présentant de nouvelles informations ou servant un nouveau but discursif dans 50 % des cas. L'exemple (41) précédent illustre ce premier cas de figure : R initie la description de l'élément perturbateur majeur de son récit. Alex ne commente plus le prix du billet pour le voyage retour, mais raconte le voyage retour en lui-même. L'autre moitié des séquences comportant un *backchannel* substantiel est caractérisée

par la confirmation des informations explicitées dans le dernier segment par une élaboration ou une explication, à l'image de l'exemple (42).

- (42) Rhianna L i mean my mum's pushing me to get my license
 SC (h) uh which i guess i [NOD should] #
 Alex it's a good thing [to have
 Rhianna R (h) but] #
 Alex if you ever need it i mean
 Rhianna well first of all
 for the moment
 i'm just not interested enough

Produite dans une séquence argumentative autour du permis de conduire, cette appositive relative est un commentaire de la locutrice sur le segment de discours précédent (la mère de Rhianna aimerait qu'elle passe son permis; Rhianna reconnaît la validité "objective" de ce point de vue par "I guess" et le modal "should"). La contribution d'Alex se scinde en deux tours : dans le premier groupe intonatif, elle participe à la concession apportée à la thèse de la séquence (même si Rhianna ne souhaite pas passer son permis pour le moment, il est pratique d'avoir son permis en cas de besoin). Avant que Rhianna ne continue le fil de sa séquence, Alex modifie son *backchannel* par le biais d'une circonstancielle de condition, qui permet d'adoucir la force assertive de sa contribution. Par ce second groupe intonatif, Alex fait preuve de diplomatie et donne implicitement une certaine valeur argumentative à la thèse défendue par Rhianna (si elle n'en a pas besoin, elle a raison de ne pas vouloir passer son permis). Rhianna a par ailleurs indiqué vouloir continuer son tour, avec une reprise de souffle audible et le marqueur de discours adversatif "but" : Alex s'applique à éviter l'infélicité conversationnelle.

En résumé, les appositives expriment une rupture assez nette d'un point de vue discursif. L'ensemble des appositives introduisent de la variation dans une séquence homogène. Leurs séquences hôtes sont riches en disfluences et en marqueurs de discours structurants. Nous avons mis en lumière la capacité des appositives relatives à véhiculer du contenu modal, où le locuteur prend position sur les relations exprimées par le co-texte. Les appositives permettent alors des glissements de points de vue. Les co-locuteurs y sont sensibles, répondant par des *backchannels* denses. Ces *backchannels* n'expriment pas simplement leur alignement à la séquence en cours, mais davantage leur affiliation (ou leur non-affiliation) au point de vue exprimé par le locuteur (Stivers, 2008). Les appositives canoniques, quant à elles, sont fréquemment juxtaposées dans la macro-séquence en position médiane, et reviennent sur les référents établis dans le co-texte gauche pour les modifier. Ce sont les subordonnées les plus représentatives du phénomène de *self-monitoring* de la part du locuteur, évaluant

et modifiant une partie très restreinte de son discours en temps réel. Les appositives canoniques illustrent le compromis réalisé par le locuteur entre progressivité et compréhension : le coût représenté par l'arrêt momentané du discours est jugé moins important par le locuteur que celui estimé de la reconstruction interactionnelle qu'entraînerait le non-alignement du co-locuteur à une séquence discursive entière. Tout doute quant à la compréhension du co-locuteur est traité localement par l'insertion d'une forme minimale qui n'appelle pas un développement. Cette logique de "contingence locale" voit l'absence de conjonction ou d'autre marque cohésive verbale, et les locuteurs ne sont pas particulièrement sensibles aux ajustements qu'elles réalisent : les *backchannels* sont clairement moins denses que ceux suscités par les appositives relatives.

En ce qui concerne les corrélations entre les différents indices de rupture, il est étonnant d'observer que si des interactions existent au titre de tendances, elles ne sont pas très représentatives de l'ensemble des appositives. La corrélation la plus grande s'observe entre l'autonomie macro-syntaxique totale et les marqueurs de discours de structuration : 50 % des séquences contenant une appositive autonome d'un point de vue macro-syntaxique comportent au moins un marqueur de discours jouant sur la structuration de la séquence. Cette configuration, qui représente 23.6 % des séquences contenant une appositive dans le corpus, est illustrée par la séquence (43) :

- (43) Alex L and (h) it took us to Limoges #
 SC **which was like # what like # four hours away from here #**
 R (h) so yeah that was pretty grim

SC représente un acte de langage à part entière, et véhicule un contenu propositionnel bien distinct de son entourage. La co-locutrice pourrait très bien réagir au contenu de SC indépendamment de celui de L et de R ("no it's actually three hours away from Nantes"). La valeur informative de SC n'est pas moindre que celle de R, bien que le prédicat de SC porte sur un item informationnel moins large et plus délimité (i.e. l'antécédent "Limoges") que celui sur lequel porte R (i.e. l'ensemble de la séquence formé par plusieurs propositions précédentes). Cet extrait se situe à la fin d'un récit à propos d'un vol particulièrement éprouvant vécu par Alex. SC est montré comme mobilisant des efforts de réflexion particuliers de la part de la locutrice. Il comporte deux occurrences du marqueur de discours modulateur "like" et présente également des pauses silencieuses intra-constituant. Si la première occurrence de "like" s'apparente à un *filler* pendant les difficultés d'Alex, il est suivi d'un second "like", qui permet de mettre en focus (Underhill, 1988) le rhème de l'énoncé ("four hours away from here"), déjà annoncé par une interrogation métadiscursive

(“what”). Cet élément permet surtout d’intensifier le récit : l’avion a modifié son plan de vol en raison des conditions météorologiques au moment de l’arrivée, et décide au-dessus de l’aéroport de Tours de partir atterrir à Limoges. Limoges est loin de Tours et de Nantes, et prolonge considérablement le voyage. Après une pause silencieuse séparant SC de R et une pause remplie indiquant qu’Alex continue son tour, le marqueur de discours “yeah” en début de R est couplé au marqueur de discours structurant “so”. Ensemble, ils témoignent d’une reconsidération par la locutrice de son discours : en entérinant le commentaire de SC avec “yeah”, Alex marque en même temps un retour cognitif sur l’ensemble de la séquence par “so”, en signalant que des inférences vont être tirées. Alex ouvre l’évaluation finale de son récit.

Les marqueurs de discours structurants ont également tendance à apparaître dans les séquences contenant des appositives en rupture avec les attentes séquentielles : 50 % des appositives sortant des attentes liées au type de séquence comportent ou sont suivies par un tel marqueur. Notons toutefois que cette tendance n’est pas forte, représentant seulement 21.8 % des appositives du corpus. Cette caractéristique concerne d’ailleurs surtout les appositives relatives, puisque 62.5 % d’entre elles présentent les deux indices. Cette corrélation est illustrée dans l’exemple (44) :

- (44) Joey my mum’s school
 it’s the school i used to go to
 have this place
 L in # Gwyn Fechan
 SC **which is # just a place**
 it’s not even a village #
 R hem # which # and the nearest town is Crickhowell

Joey décrit un hameau situé dans la campagne du pays de Galles, non-loin d’un village que les locutrices connaissent bien toutes les deux, Crickhowell. Elle interrompt sa séquence descriptive pour produire un commentaire évaluatif à propos du dernier référent introduit, “this place in # Gwyn Fechan”. La tentative d’enclencher une nouvelle prédication sur R en gardant cet élément référentiel par le biais du pivot “which” est soldée d’un abandon et d’un recours au marqueur discursif “and”. Il marque la reprise d’un élément référentiel qui n’est pas activé sur L, mais dans les groupes intonatifs précédents : Crickhowell est le topique de la séquence. Dans ce passage, la succession d’une présentative avec un mouvement de négation dans le second groupe intonatif de SC à une assertion positive sur R indique la coexistence de deux informations, différentes par la nature de leur pertinence au topique de la séquence : SC ne donne pas d’informations avançant la description de Joey. La structure présentative de SC fait écho à celle produite plus haut (“it’s the

school i used to go to”), ne contribuant pas non plus à la réalisation du but descriptif de la séquence. La co-locutrice n’est d’ailleurs pas réceptive à ces contributions en décalage par rapport aux attentes séquentielles : elle ne produit aucun *backchannel* d’alignement ou d’affiliation interactionnelle.

3.3 Discussion et conclusion

Les caractéristiques verbales des subordonnées permettent de différencier les trois types syntaxiques quant à leur mode d’insertion linéaire. Les relatives déterminatives sont clairement intégrées à leur co-texte (gauche). Les circonstancielles présentent des caractéristiques plus mixtes. Alors que leur macro-syntaxe est flexible, elles combinent certains traits discursifs démarcatifs à d’autres indices plutôt cohésifs. La disjonction ne consiste pas à isoler la circonstancielle en tant qu’unité discursive, mais plutôt à signaler un changement d’ordre interprétatif sur le discours précédent et/ou subséquent. Les appositives, quant à elles, sont clairement autonomisées des points de vue macro-syntaxique et discursif.

Les relatives déterminatives sont les constructions les plus intégrées à leur co-texte (gauche) verbalement. Cette intégration se fait au moyen de ressources macro-syntaxiques liées à la rection grammaticale, et par des ressources discursives surtout associées à l’anaphoricité qu’elles développent et à leur position au sein du paragraphe discursif. Les relatives déterminatives se démarquent toutefois par la proportion de *backchannels* minimaux réalisés à leur suite. Elles présentent également un taux médian de disfluences. Ces observations indiquent que les disjonctions d’ordre verbal, si présentes, sont majoritairement réalisées entre SC et R dans les séquences comprenant une relative déterminative. Par ailleurs, rares sont les relatives déterminatives insérées dans une séquence comptant deux indices de rupture : les marques disjonctives apparaissent très rarement en corrélation, même en mêlant indices macro-syntaxiques et indices discursifs.

Les circonstancielles présentent une macro-syntaxe partiellement autonome la plupart du temps, ainsi qu’une certaine dynamité en ce qui concerne le statut sémantique de leurs prédicats. La diversité de fonctionnement de leurs morphèmes introducteurs (i.e. “if” et “when”) est également remarquable (Debaisieux, 2006; Berrendonner, 2008). Hormis leur capacité à projeter et/ou clôturer un cadre d’interprétation pour plusieurs groupes intonatifs consécutifs et leur mobilité dans la macrostructure, les circonstancielles se démarquent finalement moins que ce qui était

prévisible au niveau du discours. Elles sont d'ailleurs introduites par des marqueurs spécialisés non pas dans l'indication d'un changement particulier dans l'organisation discursive (i.e. l'indication d'une relation procédurale textuelle), mais dans la modulation de l'interprétation du discours. En termes de corrélations entre les indices de rupture, une seule combinaison est assez récurrente : les verbes de procès ont tendance à apparaître dans les circonstanciels ouvrant une séquence cadrative dans le discours. Ici encore, l'enjeu se trouve sur le plan cognitif, cette corrélation suggérant que les syntagmes verbaux dynamiques seraient plus aisément choisis par les locuteurs en tant que repères textuels.

Enfin, les appositives marquent le passage à une autonomie majoritaire. Alors que les relatives se trouvent souvent en contradiction avec la lecture descriptive de leur marqueur introducteur, les appositives canoniques témoignent d'une disjonction par leur position au sein de la macro-séquence. Par ailleurs, certaines appositives canoniques se trouvent très éloignées de leur antécédent. Cette tendance n'est pas définitoire car elle concerne une minorité d'occurrences, mais est assez rare dans la littérature (e.g. Forsgren, 1993; Neveu, 2000; Furuwaka, 2005) pour constituer un trait non-négligeable du corpus. Les appositives sont les seules constructions à présenter un aussi large faisceau d'indices au niveau discursif. Elles se démarquent notamment par leur décalage par rapport aux attentes liées à la nature de leur séquence hôte, par les marqueurs de discours de structuration intervenant à leur suite, et par la proportion de disfluences que ces constructions occasionnent. Les co-locuteurs sont sensibles à la rupture créée par les appositives relatives, puisque la frontière entre celles-ci et leur co-texte droit représente un locus préférentiel pour la production de *backchannels* substantiels. Au lieu d'un décalage au niveau des représentations comme celui créé par les circonstanciels, les appositives relatives introduisent majoritairement une forte prise de position énonciative du locuteur sur son discours. Les co-locuteurs, possiblement par un effet miroir ou tout autre phénomène d'imitation, réagissent à cette insertion de modalité en prenant également position sur les événements et états de faits du discours. Étonnamment, les corrélations entre indices ne sont pas très marquées : elles sont légèrement plus variées que celles caractérisant les circonstanciels, mais représentent plus faiblement notre corpus lorsqu'elles sont considérées individuellement.

Le mode d'insertion linéaire des appositives comporte des similarités avec celui des segments parenthétiques (Mondada et Zay, 1999; Duvallon et Routarinne, 2005; Blakemore, 2009). Si les parenthétiques constituent des cas de parataxe (i.e. la juxtaposition d'une structure prédicative sans marqueur de coordination ou su-

bordination; Lelandais et Ferré, 2014) et présentent donc une rupture encore plus manifeste dans la chaîne verbale, elles sont clôturées par des marqueurs de discours comme les appositives et occasionnent la même proportion de disfluences. Comme les appositives canoniques, elles sont souvent insérées en position médiane, et découlent du phénomène de *self-monitoring* du locuteur en manipulant des objets de discours/référents en majorité présents dans le co-texte gauche. Appositives canoniques comme parenthétiques ouvrent un focus attentionnel sur un objet, qui doit être ré-analysé par le co-locuteur en raison d'une contradiction anticipée entre le discours du locuteur et les attentes du co-locuteur (Lelandais et Ferré, 2016). En revanche, les appositives canoniques sont contraintes à porter sur un objet beaucoup plus restreint que celui sur lequel agissent les parenthétiques, de taille et de nature variables. Comme les appositives relatives, les parenthétiques sont majoritairement utilisés pour véhiculer une prise de position du locuteur sur le contenu de son discours. Toutefois, les parenthétiques se différencient des appositives relatives par leur grande capacité à établir des liens cataphoriques avec le co-texte. Les formes prospectives prédisposent la compréhension du co-participant à un certain type d'action : elles peuvent suggérer la sorte de réception de la part du co-locuteur qui est attendue par le locuteur pour sa séquence, afin de parer à d'éventuelles marques de désalignement ou de désaffiliation. Par ces traits cadratifs, elles sont un peu plus proches des circonstancielles à ce niveau.

L'encodage d'un procès/état de fait distinct (i.e. qui n'exprime pas une action/état de fait simultané) de ceux du co-texte dans la subordonnée, ainsi que la dotation d'une force illocutoire propre à la subordonnée constituent les moyens macro-syntaxiques les plus communs d'autonomisation des constructions subordonnées dans notre corpus. En ce qui concerne les indices discursifs, les ressources les plus utilisées par les locuteurs sont l'introduction d'une rupture de thème/participant au sein de la macro-séquence, ainsi que la production de marqueurs de discours de structuration. Du point de vue de la réception, les *backchannels* minimaux sont également des indicateurs très répandus. Pour autant, ces indices se combinent rarement. La proportion d'indices d'intégration dans les séquences sous étude telles que le travail cohésif des références lexicales réduites (i.e. les pronoms) sont également très répandus. Cette dernière observation peut servir à classifier nos trois types de subordonnées : les relatives déterminatives montrent plus de signes d'intégration que de démarcation, tant sur le plan macro-syntaxique que sur le plan discursif. Les circonstancielles sont des constructions mixtes sur le plan macro-syntaxique; elles œuvrent davantage à la modulation du discours qu'à sa démarcation d'une manière linéaire. Les appositives comptent finalement beaucoup plus d'indices démarcatifs

que de marques cohésives. Les corrélations entre indices de rupture, absentes sur les relatives déterminatives, sont de nature différente pour les circonstancielles et les appositives.

Dans ce chapitre, nous avons proposé que la subordination peut être représentée par un plus grand nombre d'indices macro-syntaxiques et discursifs que par le simple repérage de la construction subordonnée par rapport à un verbe régissant situé dans le co-texte. À travers les récurrences que nous avons pu dégager, l'insertion des subordonnées dans la chaîne verbale témoigne de la plasticité des unités syntaxiques et discursives à l'intérieur des séquences sous étude, que cette insertion soit une intégration ou une démarcation. Dans leur ensemble, les subordonnées témoignent de la temporalité complexe du discours, faite à la fois de projections (circonstancielles), d'émergence (appositives canoniques), et d'orientations rétrospectives (relatives déterminatives et appositives relatives). Les co-locuteurs n'ont pas moins le choix entre l'acceptation et le rejet de ces démarcations. S'ils manifestent davantage leur alignement à la séquence/l'activité en cours par la production de *backchannels* minimaux, les constructions les plus en rupture et montrant le plus de prise de position modale reçoivent, elles, une grande proportion d'affiliation avec la perspective prise par le locuteur, avec des rires ou des contributions personnelles très courtes.

Les caractéristiques des subordonnées permettent d'observer que les locuteurs se montrent davantage préoccupés par l'économie et le bon déroulement du discours que par la présentation chronologique des événements : si la pertinence (plus spécifiquement l'idée du niveau de pertinence construite par le locuteur de son propre discours), pour mener à bien les divers objectifs discursifs, détermine le choix de l'ordre des énoncés (les subordonnées évoluent globalement dans la localité de l'élément qu'elles modifient), ces structures tendent également à l'établissement d'un consensus entre locuteur et co-locuteur (que ce consensus porte sur un référent, le discours, ou l'interaction), par lequel progresse le discours. La forme courte prototypique des subordonnées, plus ou moins séparées de leur entourage selon les types syntaxiques, est en adéquation avec les besoins et contraintes discursives dans l'espace et le contexte interactionnels.

À partir de cette analyse verbale, des tendances sont observables pour la production des subordonnées. Dans le Chapitre 4 suivant, ces résultats seront mis en perspective par rapport aux analyses des moyens de démarcation prosodique, puis aux moyens de démarcation gestuelle dans le Chapitre 6. Plus précisément, nous serons en mesure d'étudier l'articulation temporelle des différentes frontières

dans les séquences sous étude (e.g. complémentarité ou redondance). Nous pourrions également juger du marquage entre les différents types syntaxiques, à savoir si les relatives déterminatives sont également, dans la modalité vocale puis dans la modalité gestuelle, moins en rupture que les circonstancielles, elles-mêmes moins autonomes que les appositives relatives.

Chapitre 4

Frontières prosodiques

4.1 Mesure des frontières prosodiques

Nous évaluons dans ce chapitre l'autonomie prosodique des subordonnées sur un mode d'organisation linéaire (intégration vs. démarcation). Nous déterminons si ces constructions créent globalement une rupture ou bien si elles sont préférentiellement intégrées au co-texte gauche et/ou droit.

Après avoir identifié et mesuré les indices de rupture prosodique les plus pertinents tels que nous les avons formulés dans les hypothèses du Chapitre 1 p. 58, nous sommes en mesure de placer les trois types syntaxiques sur un continuum, de l'intégration à l'autonomie prosodique, ainsi que d'identifier les indices les plus mobilisés dans le marquage d'une rupture. Les subordonnées les plus intégrées à leur entourage sont les relatives déterminatives, ne présentant que deux indices de rupture réguliers, tous deux de nature rythmique. Viennent ensuite les circonstancielles, qui comptent trois marques de frontière. À la différence des relatives déterminatives, la rupture s'établit par le biais de ressources intonatives. Enfin, les appositives relatives sont clairement autonomisées en rassemblant huit indices de rupture, mêlant paramètres rythmiques et intonatifs. Le détail des caractéristiques de chaque type syntaxique de subordonnée est donné de la catégorie la moins autonome à celle la plus en rupture.

4.2 Résultats

4.2.1 Relatives déterminatives

Les relatives déterminatives ne comportent que deux indices de rupture prosodique. Ce type syntaxique exprime donc relativement peu de démarcation du point de vue vocal. Ces marques sont de nature rythmique. Les RD présentent d’abord une durée distinctive, SC étant significativement le segment le plus long de la séquence dans laquelle il s’inscrit (L : $F(54, 54) = 3.01$, $p < 0.0001$; R : $F(54, 54) = 4.48$, $p < 0.0001$). L’exemple (1) illustre cette tendance, qui représente 43.6 % des relatives déterminatives du corpus :

(1)	Tim	L	had to rely on false # information	1.3 sec
		SC	that they thought was correct #	<u>1.5 sec</u>
		R	i mean #	0.5 sec

Cette relative déterminative se situe au milieu d’une argumentation autour de dissimulations politiques, dans laquelle Tim exprime sa difficulté à prendre parti. Il oppose la notion de mensonge (“intentional lie” énoncé auparavant) à celle de méprise (“false information that they thought was correct”). Il consacre deux groupes intonatifs à cette dernière notion, introduite dans le rhème de L avec une pause silencieuse pouvant s’apparenter à de la démarcation (elle permet de lier SC à L dans un même souffle, le tout formant une unité de conscience; Chafe, 1994), mais aussi à de la focalisation (elle apparaît en fin de groupe intonatif pour introduire un seul élément lexical bien que le terme soit très générique, et un ralentissement du débit s’observe après la pause; Candea, 2000; Ferré, 2004). Le contenu sémantique de SC est crucial pour permettre l’opposition à “intonational lie”, créant une sous-classe contrastive pour “false information” (les fausses informations que le gouvernement savait pertinemment fausses vs. les fausses informations que le gouvernement croyait vraies). Bien que ce segment ne comporte pas de pause intra-constituant, SC est plus long que L et que R.

Les séquences contenant des RD présentent par ailleurs le plus de pauses silencieuses intra-constituant (les pauses silencieuses de cette catégorie représentent 42.2 % du temps de pause silencieuse intra-constituant total, totalisant 12.7 secondes) parmi celles contenant les trois types de constructions subordonnées. Ces pauses sont plus présentes au sein des relatives déterminatives en elles-mêmes (représentant 68.5 % du temps de pause dans les séquences contenant une relative déterminative, pour un total de 8.7 secondes) que dans le co-texte gauche (7.1 % du temps de pause, totalisant 0.9 secondes) et droit (24.4 % du temps de pause,

totalisant 3.1 secondes). Cette différence n'est toutefois pas significative statistiquement. La présence de pauses silencieuses à l'intérieur des relatives déterminatives est illustrée par l'exemple (2) :

- (2) Tom L check the fact
 SC **that it says # eighty percent # good meat**
 R kind of i don't know what it would say

Alors que L et R ne sont pas caractérisés par des pauses silencieuses (intra-constituant comme inter-constituant), Tom réalise deux pauses silencieuses à l'intérieur de SC. Ces deux pauses ne correspondent pas à des frontières syntaxiques à proprement parler. La première pause sépare le prédicat “says” de son complément d'objet direct. N'intervenant pas en début de groupe en isolant une conjonction ou un sujet grammatical, cette pause prend surtout son sens au niveau du discours, puisqu'elle marque le passage du discours de Tom à la lecture de l'étiquette de la viande en question. La pause signale ce changement d'ordre énonciatif et prend alors une fonction démarcative. De plus, elle permet de marquer un lien serré entre la fin de L et le début de SC. L'unité interactionnelle en cours est alors signalée comme non-finie et Tom s'assure par-là de conserver son tour de parole. La seconde pause de SC est différente; apparaissant en fin de groupe, elle sépare le groupe adjectival de son extension adverbiale. Le débit est considérablement ralenti après cette dernière. C'est une pause de focalisation, qui met une partie du rhème en valeur et marque l'information la plus importante de la séquence à un niveau pragmatique.

Nous avons vu que les pauses silencieuses intra-constituant jouent plusieurs rôles dans les séquences contenant les relatives déterminatives. Nous sommes également en mesure d'observer qu'elles n'occupent pas la même place au sein des séquences en fonction de ces différents rôles. Le tableau 4.1 détaille la distribution des pauses silencieuses intra-constituant par nombre d'occurrences dans les séquences des relatives déterminatives, et montre que ces pauses ne sont pas mobilisées de la même façon à travers le temps.

	démarcation	focalisation	hésitation
L	6	2	2
SC	8	7	7
R	4	5	9
total	18	14	18

TABLE 4.1 – Distribution par nombre d'occurrences des différents types de pauses silencieuses intra-constituant dans les relatives déterminatives.

Si les pauses silencieuses situées sur L et sur R ont des rôles clairement différents (pauses de démarcation sur L et pauses d’hésitation sur R), les pauses situées sur les relatives déterminatives elles-mêmes démontrent une grande variabilité de fonctionnement, ainsi qu’un usage très équilibré de ces trois différentes fonctions. L’exemple (3) illustre ces emplois :

- (3) Michelle L (h) and so she # disowned everything
 SC **that # she could associate my nana #**
 R for example her accent # and #

Dans cet extrait, les deux premières pauses intra-constituant (l’une sur L, l’autre sur SC) sont utilisées à une même fin interactionnelle : une première pause sépare le sujet grammatical du prédicat de L alors qu’une seconde pause sépare le pronom relatif de SC de la proposition qu’il introduit. Elles ne coïncident donc pas avec les frontières syntaxiques, et ne sont pas accompagnées d’accélération ou de ralentissement significatifs du débit de parole : ce sont des pauses démarcatives déplacées (Ferré, 2004), i.e. des pauses déplacées permettant au locuteur de garder le tour de parole tout en effectuant une recherche lexicale ou un calcul des unités). La frontière entre L et SC aurait effectivement pu être perçue comme une *TRP*¹ par Zoe, la colocutrice, si Michelle ne l’avait pas désamorçée par un contour descendant unique et une technique de *rush-through* (i.e. Michelle n’allonge aucune syllabe de “everything” avant d’enchaîner rapidement par “that”), visible dans la figure 4.1². En revanche, R est caractérisé par une pause assimilable à de l’hésitation, bien que celle-ci ne soit pas accompagnée de marque lexicale du travail de formulation : le débit est ralenti avant cette pause et la syllabe finale de l’item “accent” est allongée. Malgré la réalisation de cet item avec un contour de liste (légèrement montant-descendant) et cette pause facilitant la réflexion et/ou le travail d’encodage, Michelle produit finalement la conjonction “and” de façon post-rhématique, en plaçant cet élément lexical en queue de groupe intonatif, avec un fondamental très bas et sans mouvement mélodique propre. Cette configuration correspond au *trailoff-and* de Local (2007) et abandonne le tour.

La conjonction introduisant SC est apposée à L, en suivant la déclinaison F0 de la finale de L : le fondamental continue de chuter, de 188 Hz en fin de L à 176 Hz en début de SC (valeurs *Low* à la suite l’une de l’autre). Si Michelle ne permet pas

1. *Transition Relevance Places* (i.e. *TRPs*), points de complétude sémantique, syntaxique, et intonative du discours, signalant la fin potentielle d’un tour de parole (Sacks et al., 1974).

2. Nous n’avons pas inclus le texte de transcription dans la figure, car la séquence discursive est trop longue pour faire apparaître la transcription intégrale de l’exemple; c’est également le cas dans la majorité des figures de ce chapitre.

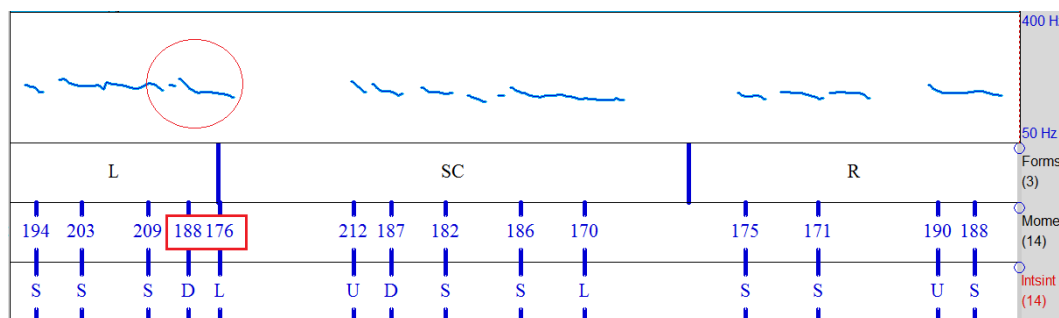


FIGURE 4.1 – Contour intonatif de l'exemple (3) dans Praat (la première piste de transcription montre les segments – L, SC, R – dans la séquence, la deuxième piste présente les valeurs de hauteur F0 MOMEL, et la troisième piste montre les valeurs symboliques de INTSINT), où une technique de *rush-through* joint les items finaux de L aux items initiaux de SC.

l'insertion d'un *backchannel* entre L et SC en liant les deux segments, elle ne rend pas impossible la perception d'une *TRP* en fin de SC, en marquant même une pause silencieuse. Sans réaction de Zoe, Michelle apporte alors une élaboration au segment de discours qu'elle a mené à bien en L/SC, en l'exemplifiant de façon à ce qu'un lien soit explicitement marqué entre ce segment et le topique en cours : les accents. Elle augmente de cette façon la pertinence du segment de discours face à l'absence de réaction de Zoe, qui suggère que l'information n'a pas été développée en totalité.

Une corrélation en lien avec les techniques de *rush-through* décrites ci-dessus s'observe dans les séquences comprenant une relative déterminative : 62.5 % des relatives déterminatives possédant au moins une pause silencieuse intra-constituant sont directement intégrées au groupe intonatif de gauche. D'une manière générale, les relatives déterminatives sont d'ailleurs le seul type syntaxique dont les occurrences sont en majorité (55.3 %) directement liées au co-texte gauche sous une même unité intonative, par des contours plats sur les items lexicaux du co-texte gauche. Dans l'exemple (4) associé à la figure 4.2, L et SC partagent une même unité prosodique et prédicative, dont le noyau prosodique (souligné dans le texte) est sur SC. Une rupture importante est en revanche marquée par rapport à R, à l'aide d'une pause silencieuse inter-constituant :

- (4) Zoe L and the daughter
 SC **that she's had all her life #**
 R the # cos there's the black one

Produite dans une argumentation à propos de la qualité dramatique d'un film, cette relative déterminative présente un contour commun avec L. Le dernier item lexical de L (“daughter”) est caractérisé par un contour relativement plat, à l'image

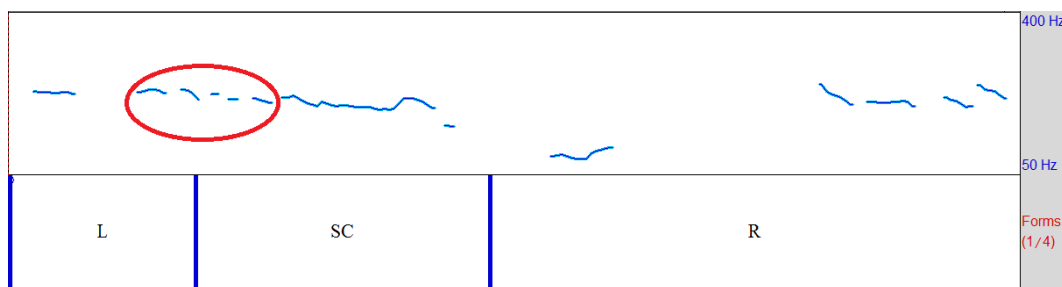


FIGURE 4.2 – Courbe intonative de l'exemple (4) dans Praat, montrant une unité prosodique commune pour L et SC; la piste de transcription indique les segments - L, SC, R - dans la séquence.

des constituants qui le précèdent. En revanche, une rupture importante est marquée par rapport à R : Zoe tente d'abord une reprise anaphorique nominale de l'élément référentiel affiné en SC ("the"), suivie d'un abandon au profit de la sélection d'un autre élément référentiel placé en opposition à "the daughter that she's had all her life".

La rupture importante marquée entre SC et R est confirmée par l'allongement syllabique sur les relatives déterminatives : de manière générale, SC présente un allongement final significatif par rapport à la durée phonémique moyenne de la syllabe finale de L ($F(80, 55) = 3.91, p < 0.001$; durée phonémique moyenne de 0.095 sec sur la syllabe finale de SC vs. 0.069 sur la finale de L). Cet allongement de la syllabe finale concerne 52.7 % des relatives déterminatives sous étude, et est visible dans la séquence suivante, associée à la figure 4.3 :

- | | |
|--------------|--|
| (5) Michelle | i dunno must have been about # |
| L | ten # girls |
| SC | who were # in their final year # |
| R | so they must have been like eighteen year olds |

Dans l'exemple (5), Michelle décrit les élèves sous sa charge lors de son année en tant qu'assistante dans un internat de Rochefort. Elle produit une pause démarcative avant le dernier item lexical de L, lui permettant de lier L à SC. Le contour intonatif du dernier élément lexical de "girls" est relativement plat et similaire aux autres mouvements de son co-texte. En revanche, le contour du dernier item lexical du groupe prosodique de SC, "year", est descendant. Au premier abord, cette syllabe finale accentuée présente une durée comparable à la dernière syllabe accentuée de L. Pourtant, "year" ne comporte que deux phonèmes alors que la dernière syllabe accentuée de L en comporte quatre. En prenant en compte la durée phonémique, la différence devient alors très nette : la durée phonémique de la dernière syllabe

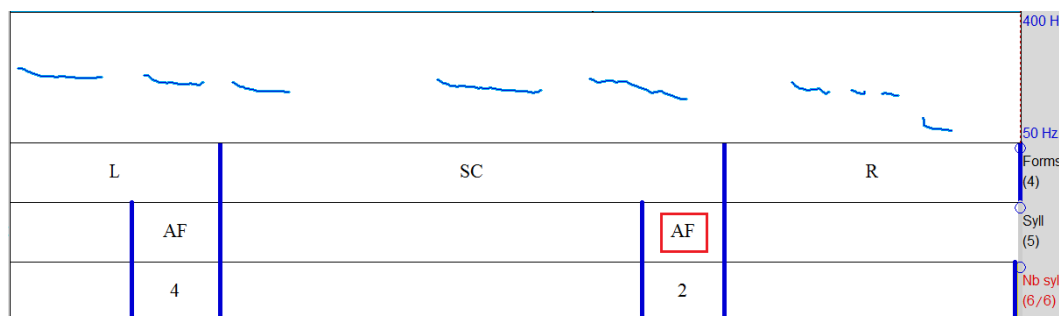


FIGURE 4.3 – Allongement syllabique final sur SC dans l’exemple (5), illustré dans la courbe intonative de Praat et les repérages syllabiques (AF = accentuée finale) et phonémiques (4 = syllabe composée de quatre phonèmes, 2 = syllabe composée de deux phonèmes) en comparaison avec la durée syllabique finale de L.

accentuée de L est de 0.139 sec, celle de la dernière syllabe accentuée de SC est de 0.257 sec. Cette syllabe atteint le second palier d’allongement calculé à partir de nos données, représentant un allongement très significatif. Une pause silencieuse inter-constituant corrobore la frontière entre SC et R, également marquée lexicalement par le marqueur de discours résomptif “so” (Schiffrin, 1987; Redeker, 2006). Alors qu’aucun intervalle de hauteur ne caractérise la syllabe finale de L et l’initiale de SC, un léger rehaussement intonatif s’observe entre la finale de SC et l’initiale de R.

Par ailleurs, en ce qui concerne les relatives déterminatives produites sous un groupe intonatif distinct du co-texte, leurs séquences hôtes montrent la plus haute distribution de contours continuatifs³ (58 % sur L seulement; 16 % à la fois sur L et sur SC comme l’illustre l’exemple (6); 9 % sur SC seulement).

- (6) Rhianna it’s basically #
 L uh the school #
 SC **hem which is right # below**
 R that uh rents # the flat out

Dans cette séquence de questions-réponses où Rhianna décrit son nouvel appartement, une relative déterminative est produite au milieu d’une structure présentative.

Si la séquence présente des marques du travail de formulation (“uh” sur L et sur R, “hem” sur SC) et des pauses qui impactent la fluidité de la séquence, L et SC sont caractérisés par des contours montants finaux, et les syllabes initiales de SC et de R sont abaissées (valeur “D” de Intsint). Bien qu’ils soient réalisés sous des groupes intonatifs distincts, ces trois segments sont liés par des procédés suprasegmentaux.

3. Un contour continuatif est un contour final montant ouvrant un lien avec le segment qui suit — R dans le cas des relatives déterminatives.

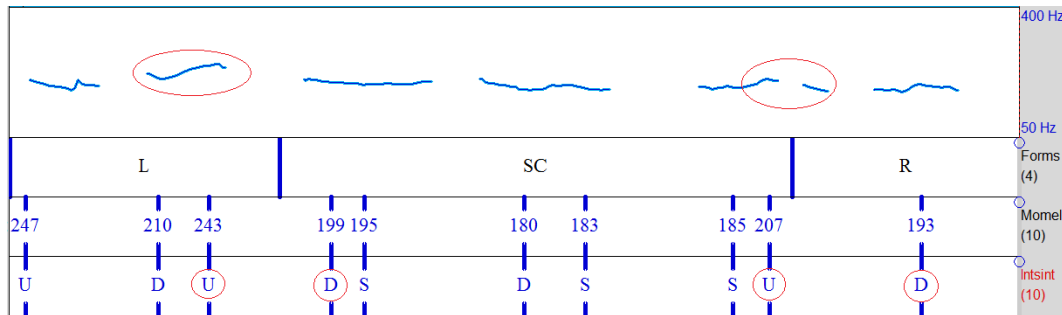


FIGURE 4.4 – Deux contours de continuation dans l'exemple (6) illustrés par la courbe intonative de Praat et les valeurs codées de Intsint (U = Upstep, D = Downstep).

En résumé, bien que les relatives déterminatives se démarquent par des caractéristiques rythmiques (durée, pauses silencieuses intra-constituant), leur intonation signale bien la dépendance. Elles ne sont majoritairement pas réalisées sous des groupes intonatifs distincts, témoignant d'une forte intégration directe au co-texte gauche. Cette intégration intonative peut également être signalée par des contours continuatifs.

Si nous venons de déterminer les indices les plus récurrents lors de la production des relatives déterminatives, rien ne permet d'affirmer que ces indices apparaissent en co-occurrence lors des séquences comprenant ce type de subordonnée. En prenant en compte les corrélations entre les différentes marques, la corrélation la plus forte s'observe entre un allongement de la durée sur le segment SC et la présence de pauses silencieuses intra-constituant dans la séquence (93.8 % des relatives déterminatives comprenant une pause silencieuse intra-constituant sont également plus longues que leur co-texte). Dans la mesure où ces pauses silencieuses sont produites à l'intérieur du groupe intonatif et en allongent la durée, cette corrélation est naturelle. En revanche, cette configuration ne représente que 27.2 % des relatives déterminatives : une telle démarcation rythmique ne reste observable que sur moins d'un tiers de notre échantillon. Aucun effet n'est visible sur le débit : si un ralentissement du débit plus important entre L et SC s'observe sur les moyennes (le débit passant de 6.88 syll/sec sur L à 6.01 syll/sec sur SC, puis à 5.54 syll/sec sur R), la différence n'est pas significative et SC ne possède pas le débit le plus lent de la séquence. Ces tendances sont illustrées dans l'exemple (7) :

(7) Joey	L	live in a place	0.5 sec; 8 syll/sec
	SC	where you can be # like self-sufficient #	2.3 sec; 3.91 syll/sec
	R	and have your own # garden #	2.2 sec; 2.73 syll/sec

Si la durée de SC s'allonge par rapport au co-texte gauche et droit, le débit est plus lent sur R que sur la relative déterminative. Ce contraste s'explique par la différence de durée entre les deux pauses silencieuses intra-constituant de la séquence, qui sont pourtant de même nature. La première occurrence est présente sur SC, la seconde sur R. La pause silencieuse à l'intérieur de SC est insérée en fin de groupe, entre le verbe copule et l'adjectif servant de complément du sujet. C'est une pause de focalisation (d'une longueur de 0.29 sec), mettant en valeur le constituant rhématique du groupe, également marqué par un ralentissement du débit et un marqueur de discours à valeur de focus, "like" (Underhill, 1988). La pause silencieuse insérée en fin de R sert également la focalisation : elle met en valeur le groupe nominal rhématique, bien que celui-ci soit un terme plutôt générique, et le débit est ici aussi ralenti après la pause. En revanche, cette pause est un peu plus longue que celle de SC, d'une durée de 0.31 sec. La différence de durée entre les pauses silencieuses est légèrement plus marquée que celle entre les segments. La combinaison des deux contrastes de durée (SC plus long avec une pause silencieuse plus courte, R moins long avec une pause silencieuse plus longue) suffit pour donner à R le débit le plus lent de la séquence.

La corrélation entre la présence de pauses silencieuses intra-constituant et un allongement syllabique final très significatif sur SC n'est pas extrêmement marquée (66 % des relatives déterminatives énoncées avec une pause silencieuse intra-constituant présentent un allongement de la syllabe finale très significatif), et ne représente que 18 % des relatives déterminatives.

En somme, bien que des types de combinaisons soient déductibles à partir des tendances liées à la nature et à l'emplacement des indices de rupture, les séquences contenant une relative déterminative présentent rarement plus d'un indice de rupture à la fois. Ces combinaisons d'indices sont de nature strictement rythmique. Si l'on s'intéresse aux caractéristiques prosodiques générales des relatives déterminatives et non plus aux marques de rupture, la très forte intégration intonative des relatives déterminatives à leur environnement s'observe très nettement : 73.7 % des relatives déterminatives produites avec un contour continuatif sont des occurrences directement intégrées au groupe intonatif de gauche. Cette configuration concerne 23.6 % des relatives déterminatives, ce qui n'est pas une proportion beaucoup plus faible que celle décrivant la corrélation la plus fréquente entre indices de rupture dans ce type de subordonnée. En plus d'être intégrées au groupe de gauche, un nombre non-négligeable de ces constructions créent donc un lien vers le groupe intonatif de

droite, visible dans l'exemple (8) ci-dessous, associé à la figure 4.5.

- (8) Joey L the Spanish girls
 SC **that were there #**
 R on our second one

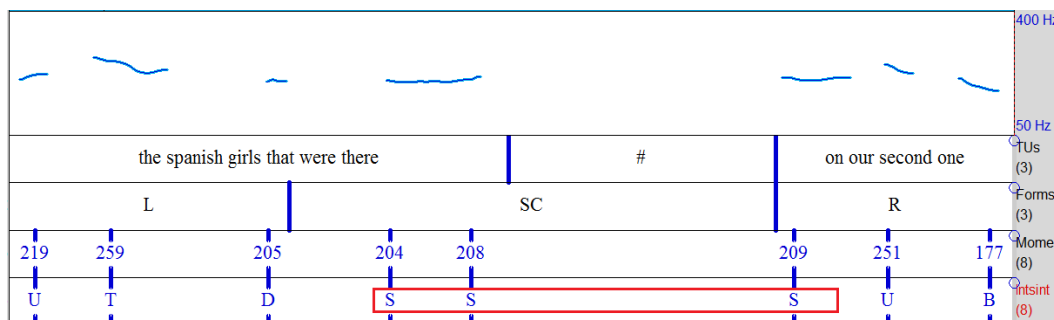


FIGURE 4.5 – Regroupement de L et de SC sous un même groupe intonatif, et contour continuatif en fin de SC dans l'exemple (8), illustrés dans la courbe intonative de Praat, la transcription des groupes intonatifs et les valeurs codées de Intsint (S = Same).

Malgré la pause silencieuse inter-constituant qui marque une frontière entre SC et R, SC, partageant un groupe intonatif commun avec L, est réalisé à l'aide d'un contour plat dont le final est légèrement montant. R débute à la même hauteur intonative que la fin de SC (à une hauteur de 209 Hz contre 208 Hz en fin de SC), comme l'indiquent les trois valeurs Intsint "S" successives. Ici, le suprasegmental prend toute sa valeur en indiquant que le tour de Joey n'est pas fini malgré une rupture de rythme.

4.2.2 Circonstanciellles

Du fait de leur mobilité dans la macrostructure (i.e. elles peuvent être préposées ou postposées à la configuration syntaxique nucléaire principale), les propositions circonstanciellles montrent une configuration prosodique diversifiée en ce qui concerne les frontières et la variation intonative interne. Certains marqueurs de disjonction sont néanmoins communs à toutes ces subordinées.

Présentant trois marqueurs de rupture prosodique, les circonstanciellles se démarquent par des ressources intonatives intra-constituant. Contrairement à ce qui était prévisible, ces constructions sont produites avec une plus grande variation de hauteur F0 parmi les tons d'une même unité que les autres types syntaxiques (avec

SC présentant un maximum de 116 valeurs Intsint démarcatives dans la séquence, contre 107 pour les appositives et 99 pour les relatives déterminatives). SC est donc réalisé avec plus de mouvement mélodique que les autres segments, comme le montre la figure 4.6, associée à l'exemple (9) :

- (9) Tom L no one #
 SC **when people go into your house**
 R normally you can

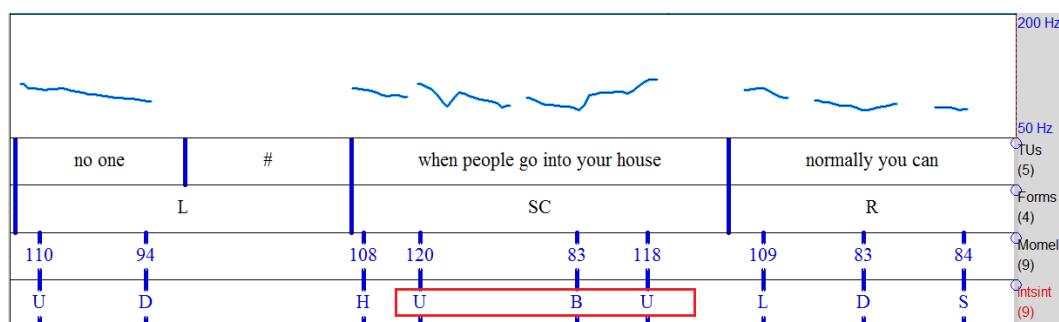


FIGURE 4.6 – Contour intonatif de l'exemple (9) illustré par la courbe intonative de Praat, la transcription des groupes intonatifs, et les valeurs Intsint (U = Upstep; D = Downstep), où SC comporte plus de mouvement illustré par les valeurs démarcatives que L et R (L = Low; S = Same).

SC possède son propre contour intonatif, et sa syllabe finale est plus élevée que l'initiale (108 Hz vs. 118 Hz). Ce segment comporte plus de variation intra-constituant (valeurs Intsint "U", "B") que son co-texte. Cette SC préposée constitue un repère pour ce qui suit et enchâsse R, qui présente un peu moins de variation interne. En revanche, il convient de noter que cette variation n'est pas observable sur les tons frontière de cette construction par rapport au co-texte : la syllabe initiale de SC n'est que légèrement plus haute que la syllabe finale de L (valeur Intsint "H"), SC commençant plus bas que l'initiale de L. De même, le début de R n'est que légèrement plus bas (valeur Intsint "L") que la finale de SC et commence à la même hauteur que l'initiale de SC. L'intervalle entre SC et R peut être qualifié de neutre. La variation est donc bien marquée à l'intérieur de la circonstancielle et non à l'interface avec le co-texte.

Plus spécifiquement, les variations de hauteur intonative intra-constituant des circonstanciels consistent majoritairement en des abaissements de registre sur les syllabes (*Downsteps*; 64 valeurs détectées), l'emportant sur les rehaussements (*Upsteps*; 48 valeurs détectées). À l'image des autres indices décrits dans cette partie, ce phénomène s'observe sur les deux sous-types sémantiques pris en compte (cir-

constancielle temporelle, circonstancielle conditionnelle), mais cette spécificité est légèrement plus marquée sur les conditionnelles, bien que la différence avec les temporelles ne soit pas significative. Parallèlement à ces tons de rupture, les circonstanciels sont plus modulés que leur co-texte droit : elles comportent moins de valeur Intsint *Same* que R. Cette différence n'est toutefois pas significative (47 valeurs détectées sur SC vs. 52 valeurs sur R). L'exemple (10) associé à la figure 4.7 montre les abaissements successifs de F0 sur les circonstanciels :

- (10) Joey L i hope so #
 SC **if i go to dance tonight**
 R i'll ask her

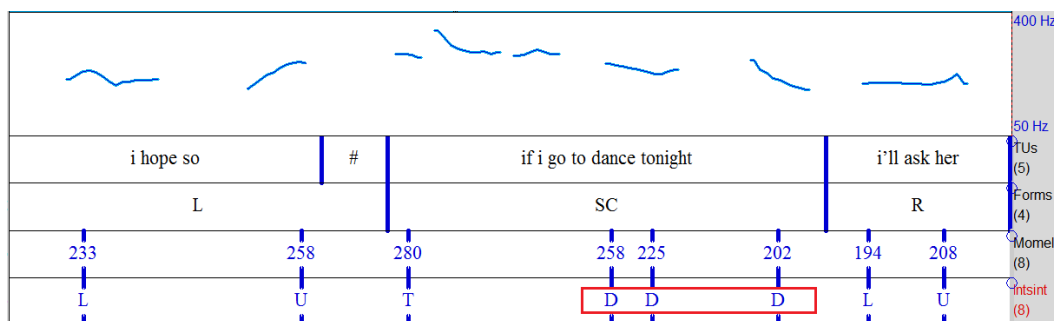


FIGURE 4.7 – Contour intonatif de l'exemple (10), où l'on observe davantage de mouvement mélodique sous forme d'abaissement ($D = Downstep$) sur SC, illustrés dans la courbe Praat, la transcription des groupes intonatifs, et les valeurs codées Intsint ($U = Upstep$, $T = Top$).

Dans cette circonstancielle de condition, une grande montée intonative s'observe sur la syllabe finale de L. Joey commence à répondre à une question de la co-locutrice à propos de la soirée qu'elle organise ("is Suzie going"); le ton montant véhicule la préférence de Joey pour la réalisation d'une réponse positive, en même temps que son désengagement de la prise en charge épistémique de la réponse : Joey aimerait que Suzie vienne, mais rien ne lui permet de l'affirmer pour le moment. La hauteur de réalisation de ce ton, pour sa part, représente un appel à la co-locutrice; Suzie est une connaissance commune des deux participantes, et Joey manifeste son intérêt pour l'apport d'Elena de ce référent dans le discours : la plage haute est utilisée pour reconnaître la pertinence de la question d'Elena. Après une pause silencieuse inter-constituant, SC reprend plus haut que L : la co-locutrice n'a pas manifesté l'envie de prendre la parole malgré la pause silencieuse entre L et SC, laissant Joey élaborer davantage sa réponse. Peut-être juge t-elle la contribution de L insuffisante du point de vue de la quantité et de sa substance. Avec le maintien de la plage haute, Joey indique à la co-locutrice que ce nouveau segment répond également à sa

question, et va représenter une contribution importante. En effet, alors que L exprimait un désengagement épistémique auquel Elena n'a pas souhaité réagir, SC, bien que conditionnelle, initie un engagement déontique de la part de Joey, et décrit une action dont la réalisation se situe sur le plan de l'habitude. Par une série d'abaissements intonatifs, Joey termine SC en revenant à un registre moyen, qui constitue la bande neutre du discours. Le ton descendant final donne une inflexion déterminée à son segment, préparant la valeur d'engagement de R.

Les circonstancielles se distinguent également de leur entourage par des stratégies rythmiques. Si elles ne présentent pas de durée ou de débit de parole significativement différent, les séquences dans lesquelles elles sont incluses montrent la plus haute distribution de pauses remplies (39.6 % de la totalité des pauses remplies produites dans les séquences sous étude). Cette distribution doit toutefois être nuancée, représentant 20 % des circonstancielles de notre corpus : les pauses remplies à la suite des circonstancielles correspondent à une tendance plutôt qu'à un phénomène clairement marqué. Ces pauses remplies se situent en majorité immédiatement après leur production (47.4 % des pauses remplies produites dans les séquences contenant une circonstancielle), i.e. dans le co-texte droit. Les reprises de souffle audibles (h) sont les plus fréquentes (68.4 % des pauses remplies), illustrées dans l'exemple (11).

- (11) Rhianna L so that for example
 SC **if there's a terrorist attack**
 R (h) hem # there are loads of solutions

Alors qu'une réalisation fluide et une technique de *rush-through* visibles dans la figure 4.8 marquent L et SC comme faisant partie d'une même unité cognitive, R représente une plus grande charge de calcul. Rhianna combine une reprise de souffle audible à la marque de travail de formulation "hem" suivie par une pause silencieuse d'hésitation. Bien que physiologique, la reprise de souffle audible est employée à des fins interactionnelles : Rhianna indique sa volonté de garder le tour de parole malgré un travail d'encodage imminent.

Cette reprise de souffle complète l'action du contour descendant-montant final de SC, qui est plus complexe : ce contour suggère non seulement l'incomplétude, mais aussi que certaines implications sont confiées au dernier item lexical du groupe. La reprise de souffle vient alors confirmer que si des inférences sont à tirer de ces derniers énoncés, Rhianna n'a pas encore mené à bien son but discursif. À cet égard, les intervalles de hauteur intonative entre les différents segments sont intéressants dans cet exemple, en ce qu'ils sont sensiblement inexistantes. Ces intervalles non-marqués

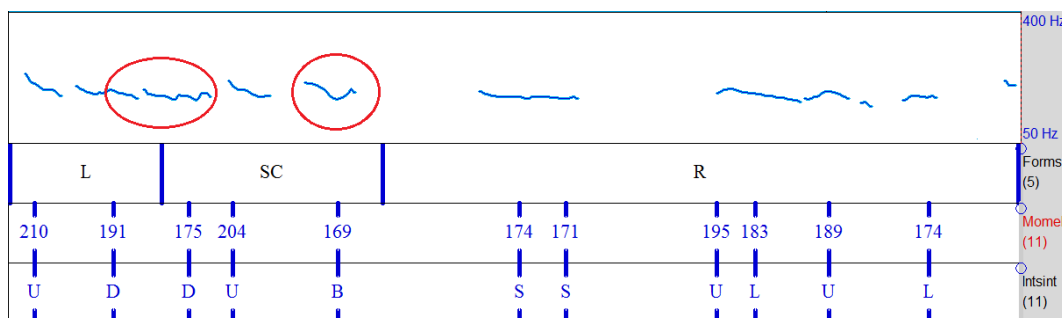


FIGURE 4.8 – Contour prosodique de l'exemple (11) présentant un *rush-through* entre L et SC ainsi qu'un contour nucléaire descendant-montant sur SC, suivi par moins de mouvement prosodique en R, illustrés dans la courbe Praat et les valeurs codées Intsint (U = Upstep, D = Downstep, B = Bottom, S = Same).

indiquent que cette suite de groupes intonatifs fait partie du même paragraphe oral, malgré les disjonctions créées par les pauses marquant le travail de formulation.

À l'image de ce dernier exemple (11), “hem” apparaît également préférentiellement dans le co-texte droit, et est plus fréquente que la marque “uh” (avec 5 occurrences contre 2). “Hem” indique une frontière moins importante que “uh” (Swerts, 1998; Benus et al., 2006), visible dans l'exemple (12) à titre comparatif, associé à la figure 4.9 :

- | | | | |
|------|---------|----|---|
| (12) | Rhianna | L | she was an airforce pilot |
| | | SC | when she was young |
| | Alex | | really |
| | Rhianna | R | (h) she's # <u>uh</u> # |
| | | | she was the first woman to g- get into the # army |

Inserée dans une description à propos de la mère de la locutrice, cette SC post-posée est réalisée de manière fluide jusqu'au *backchannel* minimal d'Alex, qui exprime verbalement et gestuellement (par un haussement de sourcils) sa surprise face à l'information qui vient d'être énoncée. Le fait qu'Alex montre son affiliation⁴ entre SC et R est un indicateur en lui-même : l'unité L/SC est perçue comme un ensemble informationnel par Alex, qui n'intervient pas avant sa complétude. Rhianna indique par une reprise de souffle audible (réalisée en chevauchement avec le *backchannel* d'Alex) que d'autres segments de discours sont à venir, s'assurant de garder le tour conversationnel. Le *backchannel* minimal d'Alex, bien que court et peu intrusif, a perturbé sa description : Rhianna combine une pause silencieuse au marqueur

4. Le co-participant montre son affiliation interactionnelle en adoptant la position modale attendue par le locuteur et/ou en reconnaissant l'importance de la nouvelle information présentée (Stivers, 2008).

d'hésitation “uh”, avant d'abandonner la structure syntaxique en cours pour reprendre son sujet grammatical dans une nouvelle unité intonative. Dans ce cas, un abaissement de la hauteur intonative est perceptible sur R (valeur Intsint “D”), suivi par un rehaussement dans le groupe intonatif suivant (valeur Intsint “U”).

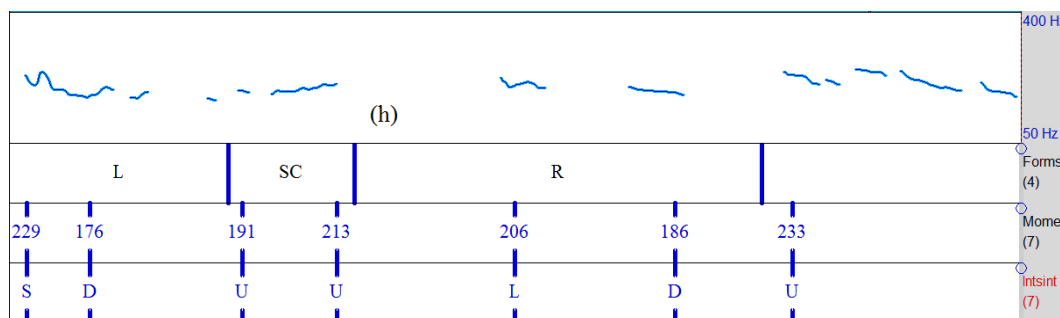


FIGURE 4.9 – Contour prosodique correspondant à l'exemple (12) où une large rupture se crée entre SC et R, illustré dans la courbe Praat et les valeurs codées Intsint.

Pour récapituler, un changement est signalé par des variations de hauteur F0 intra-constituant dans les circonstanciellles. Ces subordonnées sont d'ailleurs plus modulées que leur co-texte droit. Des pauses remplies sont également produites à leur suite, marquant selon leur nature la volonté du locuteur de garder son tour, et/ou des difficultés de calcul ou d'encodage des unités. Dans ce type syntaxique, les indices vocaux signalent l'autonomie.

Cependant, rares sont les circonstanciellles présentant plus d'un indice de rupture dans la même séquence : seulement 14.5 % d'entre elles sont à la fois plus modulées que le co-texte et suivies d'une pause remplie. Un lien entre les deux indices s'établit toutefois, puisque 72.7 % des circonstanciellles suivies par des pauses remplies sont également plus modulées que leur entourage. La production de certaines circonstanciellles nécessiterait alors une fenêtre d'attention et un effort communicatif particuliers, à l'image de l'ouverture d'un cadre à partir duquel les locuteurs mettent en place des stratégies de conservation de leur tour de parole. En décomposant les circonstanciellles selon leur sémantisme, on observe que les circonstanciellles de temps sont plus fréquemment suivies de pauses remplies (29.6 % des temporelles sont suivies de pauses, alors qu'elles sont produites à la suite de 11.1 % des conditionnelles). La combinaison de la variation de hauteur intonative à la présence de pauses remplies à la suite des circonstanciellles temporelles est visible dans l'exemple (13) ci-dessous, associé à la figure 4.10.

- (13) Alex L you know
 SC **when there's turbulence #**
 Rhianna mhmh
 R (h) i always look at them i'm like #

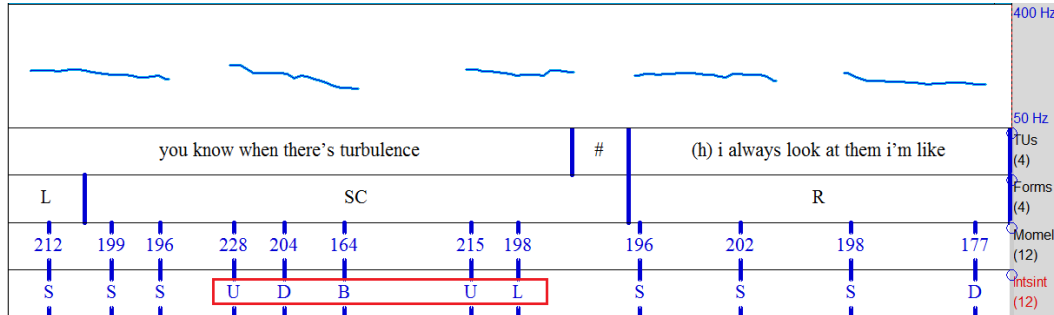


FIGURE 4.10 – Circonstancielle temporelle présentant plus de modulation que son contexte, suivie par une reprise de souffle audible, représentée par la courbe intonative de Praat et les valeurs codées de Intsint (B = Bottom, U = Upstep, D = Downstep, S = Same, H = High, L = Low).

Dans la séquence (13), Alex décrit son comportement typique lors d'un vol, correspondant à l'angoisse que lui procure l'avion en général. Cette circonstancielle sert de cadre au discours d'Alex, placée avant le noyau de la structure macrosyntaxique (i.e. le prédicat "i always look at them"). Cette position lui permet de conditionner non seulement l'interprétation de R, mais aussi celle des segments suivants, tous repérés par rapport aux coordonnées de la situation énoncée avec la circonstancielle. Il invite la co-locutrice à interpréter les segments de discours qui vont suivre avec les coordonnées énonciatives imposées par la situation énoncée de SC. L et SC sont liés sous un même contour intonatif : cela est en partie dû à la courte durée de L, fonctionnant comme un marqueur de discours afin de centrer l'attention de la co-locutrice. En se positionnant parallèlement au contenu propositionnel de l'énoncé (Clark, 1996), "you know" opère un pointage sur le discours à suivre plutôt que de constituer un véritable appel à l'autre (Fox Tree et Schrock, 2002); cette construction n'est d'ailleurs pas suivie d'un *backchannel* (dont une occurrence sera insérée plus tard dans la séquence), et est énoncée très rapidement (20 syll/sec), à l'aide d'un contour plat sans grande montée intonative suggérant un appel. Il invite la co-locutrice à tirer des inférences de ce qui va être dit, en identifiant la pertinence et les implications de l'énoncé suivant, ici à un niveau représentationnel : les deux locutrices ont précédemment abondamment discuté des vols en avion, et classé l'habitude des vols à turbulences dans leurs connaissances partagées. Alex invite donc ici Rhianna à interpréter les segments de discours qui vont suivre avec l'histoire conversationnelle de ce sujet en tête.

En revanche, le débit ralentit considérablement sur SC, passant à 4.54 syll/sec. Le contour intonatif de ce segment est plutôt modulé, passant de 228 Hz à la valeur minimale de 164 Hz (valeur Intsint “B” pour *Bottom*), pour ensuite remonter aux alentours de 200 Hz (valeur Intsint “U” suivie de “L” pour *Low*). Rhianna produit un *backchannel* minimal vocal et gestuel entre SC et R (composé de l’élément vocal “mhmh” accompagné d’un mouvement d’acquiescement de la tête), indiquant que les connaissances partagées sont bien alignées entre les deux participantes (Stivers, 2008) et encourageant Alex à continuer. Une reprise de souffle audible (h) entre la circonstancielle et la suite permet à Alex de mettre en valeur ce qu’elle vient d’énoncer grâce à une rupture de rythme et de prendre en compte le *backchannel* de Rhianna, tout en indiquant que son tour de parole n’est pas fini et que des segments vont être liés à cette ouverture de cadre : une situation va être explicitée à l’intérieur des coordonnées venant d’être posées. R reprend à la même hauteur que la fin de SC (valeur Intsint “S” pour *Same*), mais son contour est moins modulé que celui de la circonstancielle.

4.2.3 Appositives

Les indices prosodiques permettant d’autonomiser les appositives sont plus nombreux et variés que ceux des autres constructions traitées jusqu’ici. Elles comportent le plus grand nombre d’indices de frontière prosodique avec huit marques de rupture, segmentales comme suprasegmentales.

Les appositives sont d’abord les constructions les plus courtes et les plus rapides. Elles sont significativement les plus courtes au sein de leurs séquences hôtes (L : $F(54, 54) = 2$ $p < 0.02$; R : $F(54, 54) = 1.9$, $p < 0.02$). Elles sont également plus courtes que les relatives déterminatives ($F(54,54) = 3.5$, $p < 0.0001$) et que les circonstancielle (cependant $p > 0.05$). Bien qu’elles soient énoncées plus rapidement que leur co-texte, les différences ne sont pas significatives en ce qui concerne l’accélération du débit. Ce dernier résultat est donc à prendre avec précaution, et rejoint ceux de Auran et Loock (2006), qui ne trouvent pas non plus de différence de débit significative entre les appositives relatives et leur co-texte. Ces caractéristiques sont illustrées dans l’exemple (14) :

(14)	Tim	L	and they played this Irish # tune	1.7 sec; 4.12 syll/sec
		SC	which was awesome #	0.8 sec; 5 syll/sec
		R	this guy with <u>hem</u> # a whistle	1.5 sec; 4.6 syll/sec

Dans cet extrait, Tim répond à une question du co-locuteur portant sur la soirée de la veille dans un pub, et sur les musiciens jouant ce soir-là. Il interrompt ici le fil de sa réponse pour qualifier “this Irish tune” avec un terme quasi-hyperbolique, “awesome”, dans un commentaire de type appréciatif. En revanche, cette appositive ne se voit pas accorder le même temps de production : elle est nettement plus courte que son co-texte et le débit y est accéléré.

Les séquences hôtes des appositives montrent également la plus haute distribution de pauses silencieuses inter-constituant (58.2 % du temps de pause silencieuse total dans les trois types syntaxiques, totalisant 39.3 sec). Ce résultat est significativement supérieur à celui des relatives déterminatives ($F(63,26) = 3.22, p < 0.001$) et à celui des circonstancielles ($F(63,18) = 2.70, p < 0.02$). 68.7 % de ces pauses silencieuses inter-constituant suivent immédiatement le segment subordonné (i.e. elles se situent entre SC et R), comme dans l'exemple (15) ci-dessous :

- (15) Zoe L and i'm from Durham
 SC **which is in the north-east of England #**
 R (h) have you ever been there

Cette appositive est énoncée en tout début d'interaction. La séquence de présentation des deux locutrices ayant été élicitée, Zoe, encore très consciente du fait qu'elle est enregistrée, utilise cette construction subordonnée pour anticiper les attentes de récepteurs externes, potentiellement non-familiers avec la ville de Durham et sa localisation. Elle décrit alors factuellement où se situe sa ville d'origine, prenant en compte le potentiel décalage entre les représentations des deux partis (i.e. les locutrices amies *online* vs. les potentiels récepteurs ultérieurs *offline*) sans pour autant les mettre en avant dans une forme autonome. Zoe articule habilement cette contribution à une demande en dehors de l'interaction à l'avancement de son dialogue direct avec Michelle, la co-locutrice, en utilisant le pronom “you” dans le co-texte droit. Néanmoins, la transition n'est pas si lisse : elle est bien marquée à l'aide d'une pause silencieuse se situant entre SC et R. Zoe mobilise également une reprise de souffle audible au début de R en tant que marqueur de continuation de son tour de parole.

Si la majorité des pauses silencieuses de ce type se situent, comme pour les autres types syntaxiques, entre la construction subordonnée et le co-texte droit, le fait que le reste de ces pauses (31.3 %) soient localisées entre L et SC donne à ce type syntaxique une bien meilleure répartition des marques segmentales dans la séquence hôte que celle des autres types; contrairement aux relatives déterminatives et aux

circonstancielle, il n'est pas rare qu'une discontinuité sous forme de pause s'observe entre le co-texte gauche et la construction subordonnée. C'est le cas dans l'exemple (16), où l'appositive relative ajoute une qualification permettant de distinguer "Jenna and Cath" des autres éléments référentiels établis dans cette séquence :

- | | | |
|-----------|----|--|
| (16) Joey | L | she's good friends with Jenna and Cath # |
| | SC | who she does live with # |
| | R | but Cath always goes out with the other optometrists # |

Cette qualification permet d'établir un élément concessif au thème argumentatif co-construit par Joey et Elena : une de leurs amies ne construit pas beaucoup de liens avec ses colocataires. Elle est marquée à chaque frontière par des pauses silencieuses démarcatives, correspondant ici aux frontières syntaxiques des énoncés.

L'allongement syllabique se fait une autre preuve de rupture entre L et SC : si l'allongement du co-texte gauche n'est pas significatif par rapport aux items finaux de SC (durée phonémique de 0.095 sec sur L vs. 0.091 sec sur SC) car l'allongement syllabique est conséquent pour les deux emplacements, la durée sur L est significativement plus importante que celle du co-texte gauche des relatives déterminatives ($F(69, 55) = 3.92$ $p < 0.0001$, avec une durée phonémique moyenne de 0.095 sec pour le co-texte gauche des appositives relatives vs. 0.069 sec pour celui des relatives déterminatives), et que celle du co-texte gauche des circonstancielles ($F(69, 58) = 3.9$ $p < 0.05$, avec une durée phonémique de 0.095 sec pour le co-texte gauche des appositives relatives vs. 0.077 sec pour celui des circonstancielles). 52.7 % des séquences contenant les appositives relatives de notre corpus possèdent un allongement syllabique final très significatif (second palier d'allongement de notre classification) sur L. Cet allongement est moins marqué sur SC, avec 32.7 % des appositives relatives présentant un allongement syllabique final très significatif. Le placement de l'allongement syllabique est alors à l'inverse de celui des pauses silencieuses inter-constituant, dont la durée la plus importante se trouve entre SC et R.

Sur ce dernier exemple (16), la syllabe finale de L (i.e. "Cath"), qui est accentuée, présente une durée phonémique de 0.207 sec. L dépasse de très loin le second palier d'allongement de notre classification, dénotant un allongement très significatif. À titre comparatif, l'avant-dernière syllabe de SC, qui est elle aussi une syllabe accentuée, comporte une durée phonémique de 0.083 sec. C'est également la durée phonémique de la dernière syllabe de SC (qui en revanche n'a pas le même poids, puisque non-accentuée), qui n'atteint pas le premier palier d'allongement défini pour ce type de syllabe finale non-accentuée. L'allongement syllabique à la fin du groupe

prosodique de L est donc considérable, même en prenant en compte le fait que “Cath” se termine par une consonne allongeante. Pourtant, L et SC sont morphologiquement liés à l’aide du pronom relatif “who”, descriptivement plus liant et objectal que le relateur “which” : ce dernier marque un lien situationnel, exposant la construction complexe à des pivots syntaxiques et discursifs (Melis, 2008). Les informations segmentales dans cette séquence sont alors ici en contradiction avec les marqueurs morphologiques des constructions L et SC. Grâce à ces informations, SC revêt un caractère additif et paratactique, à mi-chemin entre le commentaire et la concession dans la séquence discursive de Joey.

De plus, les appositives comportent un taux relativement élevé de pauses remplies (33.3 % des pauses remplies produites dans les séquences contenant des subordinées; 25.5 % de ces séquences comportent au moins une pause remplie) majoritairement situées à la suite de leur production. Elles montrent donc un rythme très irrégulier. Ces tendances sont illustrées dans l’exemple (17) :

(17)	Alex	L	so i did the second year of # <u>hem</u> # like English # here (h) (laughs)	2.8 syll/sec
		SC	which was interesting	4.5 syll/sec
		R	(h) <u>hem</u> # yeah and then i just graduated now	3.4 syll/sec

Cette appositive apparaît dans une séquence de questions-réponses où Alex décrit son parcours universitaire à Rhianna. Caractérisée par de nombreux changements de rythme, cette séquence n’est pas fluide. L comporte deux pauses silencieuses intra-constituant, la première apparentée à de l’hésitation étant donné son couplage à la marque de travail de formulation “hem”. La seconde peut être interprétée dans la continuité de cette hésitation : “here” est un terme commun et l’hésitation porterait alors non pas sur le terme à utiliser (hésitation au niveau de la verbalisation) mais plutôt sur l’idée que souhaite exprimer Alex (hésitation au niveau de la conceptualisation; Ferré, 2004). Cette seconde pause peut également être rapprochée d’une pause de focalisation, qui met l’item “English” en relief a posteriori en le couplant à “here” : les deux items comportent des syllabes toniques. Le rire d’Alex sépare L de SC, servant à mettre en valeur le potentiel décalage entre la situation interactionnelle et son discours (alors qu’Alex et Rhianna conversent en anglais, leur langue natale, Alex annonce qu’elle a étudié en faculté d’anglais pendant une année en France). Le rire d’Alex sert également à atténuer (*hedging*) le caractère fragmenté de son discours. Elle énonce le commentaire appréciatif SC avec un débit accéléré, passant de 2.8 syll/sec sur L à 4.5 syll/sec sur SC. Une pause inter-constituant sert à démarquer SC du début de R. Elle laisse à Rhianna le temps de tirer les implications comiques

de “interesting”, qui qualifie “the second year of # hem # like English here”, et de bien les assimiler : Alex fait encore une fois appel à leur histoire conversationnelle partagée, où leur identité commune de “native English speakers living in France” est devenue un leitmotiv. Alex a des difficultés à reprendre le fil de sa réponse : elle combine le marqueur de continuation de tour (h) à une marque d’hésitation suivie d’une pause silencieuse, puis s’appuie sur “yeah” et “and then”, deux marqueurs de discours résomptifs (Schiffrin, 1987; Redeker, 2006), pour construire un nouveau prédicat et avancer sa description.

Du point de vue de l’intonation, 98.3 % des occurrences sont réalisées sous des groupes intonatifs séparés. Ces subordonnées sont donc très largement autonomisées. Ces résultats sont en accord avec ceux de Nespor et Vogel (1986), trouvant également une grande majorité d’appositives relatives énoncées sous des contours intonatifs indépendants. Les appositives relatives montrent par ailleurs moins de contours montants intégratifs que les autres types syntaxiques : 82 % des séquences dans lesquelles elles se trouvent ne comportent aucun contour montant, comme le montrent l’exemple (18) et la figure 4.11 :

- (18) Rhianna L even compared to Easyjet
 SC **which is another low cost company #**
 R yeah i hate Ryanair #

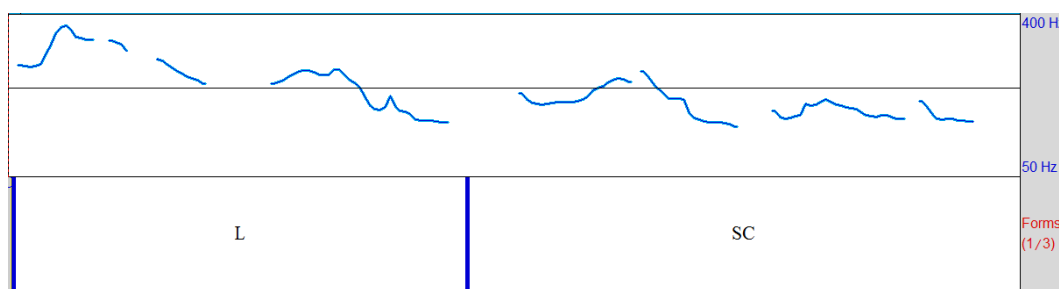


FIGURE 4.11 – Deux contours similaires (L et SC) dans un extrait de l’exemple (18) où SC montre un registre intonatif plus bas, illustré par la courbe Praat.

Cette SC qualifie le référent “Easyjet” au cours d’une argumentation, en permettant de distinguer cet élément référentiel parmi d’autres items de même classe sémantique. Si l’absence de pause entre L et SC suggère la proximité de ces deux énoncés au niveau du contenu propositionnel/référentiel, le fait que les deux segments soient réalisés à l’aide de contours définitifs indiquent qu’ils constituent deux mouvements distincts au niveau discursif : alors que L corrobore le thème principal (le personnel de Ryanair est très désagréable), SC fait pivoter ce nouvel argument

en élément concessif en explicitant le point commun entre les deux companies. SC introduit un changement de point de vue. Un léger rehaussement entre la syllabe finale de L et l'initiale de SC indique également le changement au niveau discursif, mais signale aussi que les deux segments font bien partie du même paragraphe oral. Alors que le registre intonatif de SC est plus bas, nous observons ici la construction d'une similarité par le contour mélodique.

Ce dernier exemple est également prototypique du corpus en ce qui concerne une autre ressource intonative : la nature du mouvement nucléaire, en particulier celui de L, caractérisé par un ton descendant (38.3 % des tons nucléaires sur L); celui de SC est à l'inverse plus souvent descendant-montant (48.3 % des tons sur SC). Ce résultat diffère des conclusions de Astruc et Nolan (2007), qui décrivent une duplication majoritaire du ton de L sur l'appositive relative (i.e. SC est réalisé avec un ton de même nature que celui de L). La duplication du ton de L sur les appositives relatives reste toutefois la seconde combinaison intonative la plus répandue (avec un nombre égal de descendant sur L/descendant sur SC, et de descendant-montant sur L/descendant-montant sur SC). Le contour descendant-montant crée un lien discursif et/ou implicationnel (i.e. pragmatique) entre SC et le co-texte (e.g. signifiant un contraste, une réserve, un avertissement, une hésitation, la politesse; Ward et Hirschberg, 1985; Hirschberg et Pierrehumbert, 1986). Les contours descendants-montants sur SC sont illustrés par l'exemple (19) associé à la figure 4.12 ci-dessous. Notons que la distribution des contours montants à la fois sur les segments précédant et correspondant aux appositives reste la moins importante de toutes les constructions subordonnées (12.1 % des tons nucléaires de L et SC) : si un lien avec le co-texte est bien exprimé avec le contour descendant-montant, ce lien possède une charge pragmatique plus complexe qu'une continuation ou qu'un fort appel au co-locuteur.

- (19) Tom L there's North Pentwyn
 SC **which is not North of Pentwyn (giggles) #**
 Tim okay
 Tom R so why d'they call it North Pentwyn

Tom explique dans cette séquence la topographie de son quartier. Après avoir énoncé "Pentwyn" à l'aide d'une première structure présentative réalisée avant L, il y oppose "North Pentwyn", où "North" est réalisé avec un fort accent contrastif montant-descendant, la différence alors bien mise en avant. SC revêt une double fonction discursive : cette subordonnée continue pleinement la description géographique de Tom, mais permet d'émettre un commentaire correctif humoristique. "Pentwyn", item lexical devenu une référence récurrente et connue dans le discours, est cette fois

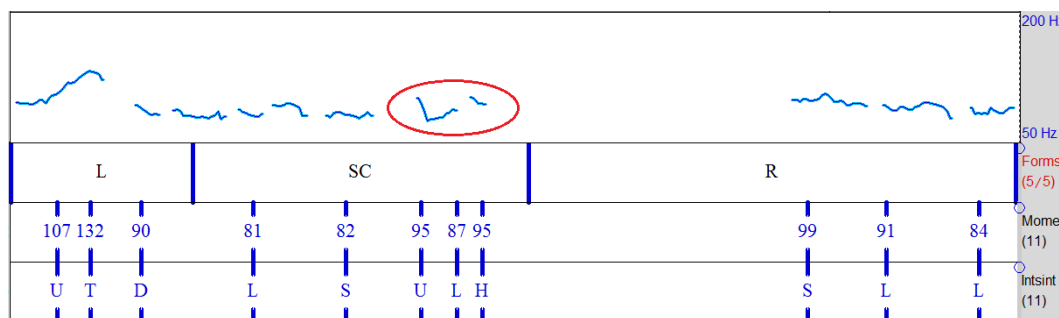


FIGURE 4.12 – Ton descendant-montant sur SC, marquant une différence claire avec le ton contrastif montant-descendant de L dans la séquence (19), illustré par la courbe Praat

réalisé avec un contour descendant-montant. Ce contour est à l'image du changement de rôle grammatical du référent, désormais sujet d'un commentaire. Tom signale que des implications sont à tirer sur le nouveau statut du référent. Ce référent doit être recontextualisé par le co-participant. Tim réagit à ce changement de statut par un *backchannel* lexical minimal : l'information à propos de ce référent a bien changé de statut dans la représentation mentale de Tim (Heritage, 1984), avec l'addition d'une nouvelle perspective (Fox Tree et Schrock, 1999). Ce *backchannel* se fait d'ailleurs le reflet intonatif du mouvement nucléaire de la subordonnée, réalisé avec un contour également descendant-montant, visible dans la figure 4.13.

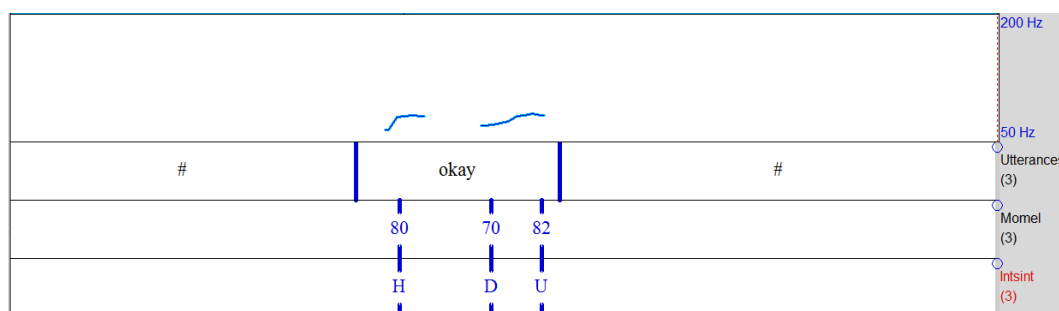


FIGURE 4.13 – *Backchannel* de Tim produit entre SC et R lors de la séquence (19), réalisé à l'aide d'un contour montant-descendant illustré par la courbe Praat et les valeurs Intsint (H= High, D= Downstep, U=Upstep).

Notons que ce *backchannel* comporte un mouvement très distinct, bien qu'il soit réalisé à une hauteur inférieure à celle des groupes intonatifs de Tom. Benus et al. (2007) proposent que ces *backchannels*, plus articulés qu'un simple terme affirmatif ou marqueur de discours, tendent à être élicités par un contour montant du locuteur principal sur l'élément précédent. Cet effet miroir sur le mouvement intonatif du *backchannel* nous évoque davantage l'établissement vocal d'un consensus dans la séquence : les représentations faisant partie des connaissances communes ont été

mises à jour. Par le biais de ce contour intonatif, Tim ne marque pas seulement son alignement⁵ à la séquence en cours, mais marque également de manière locale et économique son affiliation au point de vue de Tom sur les dernier référents énoncés. Sa réserve/surprise quant à l'information fait écho à celle de Tom, tout en encourageant une élaboration de la part de ce dernier.

Malgré une longue pause silencieuse de Tom entre SC et R, R reprend à la même hauteur que la finale de SC (99 Hz vs. 95 Hz, valeur Intsint "S" pour *Same*). En dépit du changement de mode énonciatif (R est une question), R est bien la continuation du même topique, également lié aux groupes intonatifs précédents par le marqueur résomptif "so". Ce marqueur de discours indique que les items lexicaux précédents sont sujets à d'autres inférences à propos du même topique, inférences qui avaient été mises en attente dans l'agenda du locuteur (Fraser, 1999) après la disjonction de la subordonnée, du rire, et de la réaction de Tim.

L'abaissement de registre représente l'une des autres stratégies disjonctives prototypiques sur les appositives relatives. Alors qu'une discontinuité de hauteur F0 marque le segment à la suite de la construction subordonnée dans les autres types, créant une disjonction de manière rétrospective, 60 % des locuteurs réalisent les appositives avec un changement simultané de hauteur du fondamental (> 20 Hz; $p < 0.05$ pour ces locuteurs), comme dans l'exemple (20) associé à la figure 4.14 :

- (20) Joey L in # Gwyn Fechan
 SC **which is # just a place**
 it's not even a village #
 R hem # which # and the nearest town is Crickhowell

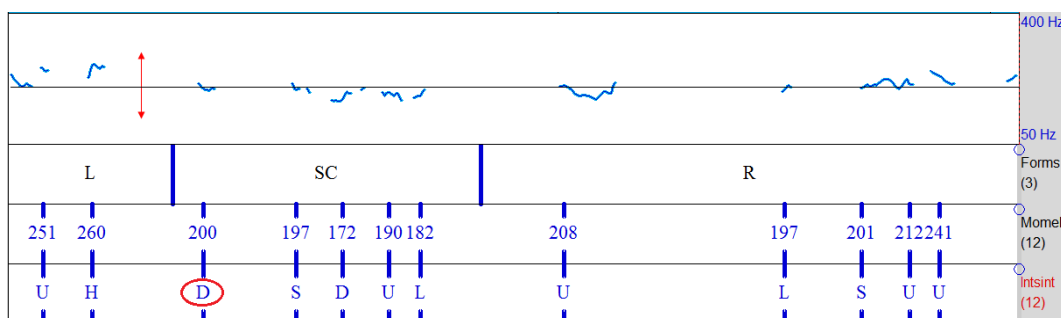


FIGURE 4.14 – Abaissement de hauteur F0 entre L et SC dans la séquence (20), illustré par la courbe Praat et les valeurs Intsint (D = Downstep).

Joey décrit à Elena un hameau situé non-loin de Crickhowell, endroit qui s'avère

5. Le co-locuteur produit un signal encourageant la progression du locuteur (Stivers, 2008).

connu des deux locutrices. SC est également un commentaire évaluatif. Alors que le registre haut de L indique que “Gwyn Fechan” est une information majeure d’un point de vue pragmatique (mise en avant par une pause silencieuse de focalisation), SC est énoncé à un registre inférieur, prenant une valeur concessive dans le discours, auquel Joey n’accorde pas la même place et le même poids. Le décalage de hauteur est visible sur toute la construction, agissant au niveau du registre intonatif. L et R montrent des hauteurs moyennes de 208.3 Hz et 202.8 Hz respectivement, alors que SC est énoncé à une hauteur moyenne de 185.4 Hz. Ce décrochage de F0 en plage basse suggère un retrait de la locutrice de l’espace co-énonciatif pour un commentaire métadiscursif, centré sur son propre discours : il s’agit d’un retour sur un élément de l’énoncé précédent. Un changement s’opère parallèlement dans la construction verbale, par une situation de coréférence. L’expression “just a place” sur SC obscurcit la référence “Gwyn Fechan”, devenue moins précise par l’indétermination avec l’emploi du pronom indéfini et du terme général “place”. Ce segment commente la pertinence du référent lexical de L dont la désignation s’est brouillée. En revanche, SC n’est pas caractérisé par un contour plat; le contour nucléaire descendant-montant donne une valeur corrective à ses éléments lexicaux. Le rehaussement de F0 sur l’ensemble du groupe intonatif suivant, R, indique bien la coexistence de deux informations, différentes par la nature de leur pertinence au topique de la séquence (i.e. “Crickhowell”).

Dans ce même exemple (20), les marques de discontinuité sont davantage de nature segmentale sur R, avec un grand nombre de pauses silencieuses, une marque de travail de formulation et un abandon de structure grammaticale en début de segment. La reprise difficile est également visible sur la hauteur intonative de la marque de travail de formulation “hem”, qui reste énoncée dans la plage basse. Bien qu’elle indique la volonté de ne pas quitter son tour de parole, Joey exprime une hésitation liée à la conception/projection d’unités plutôt qu’une hésitation portant sur la verbalisation : elle n’est pas sûre de la façon dont elle va structurer la suite de son discours. Alors que l’abandon de la structure en “which” montre que Joey avait opté pour la continuation de son commentaire, elle revient finalement à l’élaboration de la structure principale de sa séquence discursive, qu’elle souhaite mener à bien. La hauteur intonative regagne progressivement la plage moyenne, marquant son retour dans la sphère coénonciative, où les informations s’échangent et se négocient entre locuteur et co-locuteur (Morel et Danon-Boileau, 1998). Ce retour au consensus co-énonciatif est également marqué par des procédés lexicaux, tel que l’utilisation du nom “Crickhowell”, faisant partie des connaissances partagées des deux participantes.

Ce dernier exemple permet de montrer que le manque de modulation reste extrêmement sporadique parmi les appositives relatives, SC ne comptant pas plus de valeurs Intsint *Same* que son co-texte ou les autres types de constructions subordonnées. Si certaines de ces constructions sont réalisées à l’aide d’une véritable “intonation parenthétique” (Hirschberg et Grosz, 1992; Wichmann, 2001; Local, 2007) en présentant à la fois un abaissement de hauteur F0 et peu de modulation, elles sont extrêmement peu nombreuses (représentant 0.07 % des appositives relatives du corpus), et le contour plat n’est pas un trait prototypique des appositives relatives. Il constitue davantage, à notre sens, une ressource pragmatique et/ou informationnelle ponctuelle sur laquelle nous reviendrons dans le second volet d’analyse de ce travail.

Comme il est également visible dans l’exemple (20), les segments suivant les relatives appositives s’accompagnent majoritairement d’un registre rehaussé (cependant, $p > 0.05$). Pour le calcul de la hauteur du registre, nous avons identifié trois paliers intonatifs selon le registre de chaque locuteur; un ton haut se produit dans la partie supérieure du registre, un ton moyen se produit dans la partie neutre, et un ton bas se produit en plage basse. Ces valeurs lexicales ont ensuite été codées en données numériques en remplaçant les tons réalisés en plage haute par la valeur 3, 2 pour ceux produits en plage moyenne, 1 pour ceux en plage basse. Des moyennes ont ensuite été calculées, facilitant la comparaison d’une unité à l’autre. Pour les appositives relatives, nous obtenons une hauteur de registre de 1.48 sur L, 1.31 sur Sc, et 1.40 sur R. Un rehaussement initial sur la syllabe d’attaque de R corrobore par ailleurs un intervalle non-neutre entre SC et R. Le tableau 4.2 répertorie les hauteurs (normalisées) des tons de frontière⁶ sur l’ensemble de la séquence L - SC - R pour les appositives.

initiaux L	finaux L	initiaux SC	finaux SC	initiaux R	finaux R
1.57	1.33	1.36	1.33	1.45	1.35

TABLE 4.2 – Hauteur moyenne (normalisée) des tons de frontière pour les appositives en fonction du segment de la séquence.

Ce tableau présente des données normalisées selon la méthode utilisée pour la hauteur du registre. Ces données permettent d’observer que l’intervalle entre la finale de SC et l’initiale de R exprime une plus grande différence de hauteur que celle entre L et SC (+ 0.12 de hauteur vs. 0.03 respectivement). Un rehaussement initial

6. Les tons de frontière se trouvent en début (syllabe initiale) et à la fin (syllabe finale) des groupes prosodiques.

s'opère bien à la suite de ces subordonnées, beaucoup plus net et significatif que celui caractérisant la syllabe initiale de la subordonnée elle-même. Cette ressource intonative liée aux tons de frontière est alors complémentaire et non-redondante par rapport aux informations données par la hauteur du fondamental à l'intérieur des constituants. Les rehaussements initiaux sont observables dans l'exemple (21) associé à la figure 4.15.

- (21) Joey L but someone #
 SC **which is a bit silly**
 R had broken into Beth's room #

SC est inséré au milieu d'une narration relativement longue de Joey, au sujet d'une fête particulièrement mouvementée ayant récemment eu lieu chez un groupe d'amis installés en colocation. Cette construction subordonnée est remarquable dans le sens où son antécédent est verbalisé seulement après son énonciation, bien qu'elle présente toutes les caractéristiques morphologiques de l'appositive relative. Avant d'introduire un nouvel événement décisif quant à la tournure qu'a pris la soirée, Joey prend position qualitativement en tant qu'énonciatrice sur l'événement qu'elle s'apprête à énoncer. Cet ajout de modalité appréciative permet de retarder la narration.

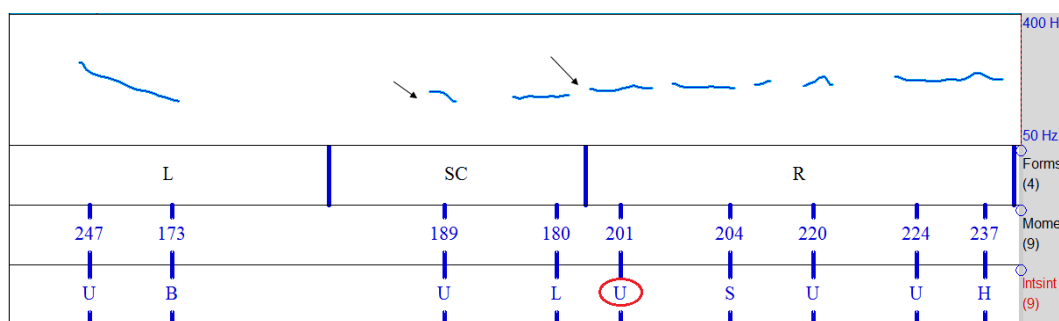


FIGURE 4.15 – Intervalles non-neutres entre les tons frontière de la séquence de l'exemple (21), illustrés par la courbe Praat et les valeurs Intsint (U = Upstep).

Bien que la proximité temporelle de SC et R indique que leur contenu propositionnel est directement lié par une relation référentielle cataphorique (i.e. les items “which” et “silly” renvoient au prédicat de R), on observe dans la figure 4.15 un rehaussement à la fois de la syllabe d'attaque de SC (valeur “U”), et de la syllabe d'attaque de R. Le rehaussement au début de R (indiquant une différence de hauteur de 21 Hz) est un peu plus conséquent que celui du début de SC (impliquant une différence de hauteur de 16 Hz). L'information apportée par ces rehaussements n'est pas incompatible avec le fait qu'un décrochement intonatif s'opère dès le début

de SC par sa hauteur F0 moyenne. Alors que L est énoncé à une hauteur moyenne de 211.9 Hz, SC plonge à 183.6 Hz. La séquence retrouve sur R une hauteur similaire aux segments qui précèdent l'appositive relative avec une moyenne de 215.9 Hz.

En résumé, les appositives relatives montrent une autonomie prosodique totale, avec des indices de disjonction segmentaux (pauses silencieuses inter-constituant, groupes intonatifs distincts) et suprasegmentaux (durée et débit distinctifs, allongement syllabique très significatif et contour descendant sur L, abaissement de registre). La rupture est marquée à la fois entre L et SC et entre SC et R.

En ce qui concerne l'articulation des différents indices au sein d'une même séquence, la plus grande corrélation s'observe entre la présence de pauses silencieuses inter-constituant et un raccourcissement de la durée de SC (85.2 % des occurrences où SC est plus court que le co-texte sont également marquées par des pauses silencieuses inter-constituant). Cette caractéristique représente 41.8 % des appositives relatives. L'exemple (22) en est à l'image, où SC présente une durée de 0.9 sec (1.5 sec avec le rire de Beth), en contraste net avec la durée de L (i.e. 3.7 sec) et celle de R (1.9 sec).

(22)	Beth	L	and then we went into # a place called Tropicana	3.7 sec
		SC	which was horrible (laughs) #	0.9 sec
		R	it's on Saint Mary Street near the castle	1.9 sec

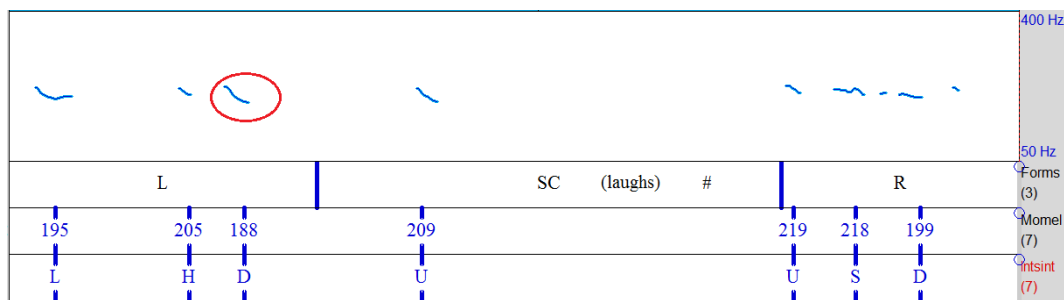


FIGURE 4.16 – Co-occurrence d'une pause silencieuse inter-constituant et d'un raccourcissement de la durée sur SC dans la séquence (22), également caractérisée par des intervalles non-neutres entre les tons de frontière des groupes intonatifs illustrés par la courbe Praat et les valeurs Intsint (U = Upstep).

Dans cette séquence, Beth fait le récit de son weekend, en réponse à une question de la co-participante (“what did you do this weekend”). L’item “a place called Tropicana” est énoncé à l’aide d’un contour final descendant sur L, malgré le fait que celui-ci soit l’objet d’une prise de position qualitative imminente de l’énonciatrice,

sous forme d'appositive relative. Ce contour est lié à la situation discursive et interactionnelle : étant donné la question de Kate, Beth indique par le contour terminatif de L que ce segment répond aux attentes de cette dernière (L est composé d'un verbe d'action, décrivant l'évènement majeur du samedi soir de Beth). Si ce contour aurait pu être interprété par Kate comme une *TRP*, le fait que SC suive L sans aucune pause désamorce une potentielle prise de tour de la part de la co-locutrice, et le tour de Beth est sauvegardé. Beth ponctue SC par un rire, atténuant sa forte prise de position : un contraste s'établit immédiatement avec la juxtaposition du terme lexical quasi hyperbolique "horrible" à son rire. Ce décalage propre à l'humour guide le travail interprétatif de la co-locutrice. La frontière entre SC et R est d'autant plus grande que Beth fait suivre une pause silencieuse à son rire, avant d'élaborer sur le même référent. La rupture créée par le rire et la pause silencieuse de Beth s'observe sur la hauteur intonative de R : la syllabe initiale de R est rehaussée par rapport à la finale de SC (valeur Intsint "U" pour *Upstep*), et la moyenne de hauteur intonative de R (i.e. 211 Hz) est légèrement plus haute que celle de L (i.e. 203 Hz). Une réinitialisation du fondamental a donc eu lieu à la suite de l'appositive relative.

Cet extrait illustre une seconde corrélation. La présence de pauses silencieuses inter-constituant (majoritairement entre SC et R) est également liée à la réalisation de L avec un contour intonatif descendant/montant-descendant (88 % des séquences réalisées avec un contour descendant ou montant-descendant sur L présentent des pauses silencieuses inter-constituant). Cette configuration concerne 40 % de nos appositives relatives. Nous observons donc une collaboration entre certains indices segmentaux et suprasegmentaux. Cette correspondance entre indices de différente nature suggère que les frontières sont marquées de manière véritablement prosodique et variée dans les séquences comprenant des appositives relatives, l'organisation de ces indices montrant une temporalité bien spécifique à l'intérieur de ces séquences (contour descendant final sur L, pause silencieuse entre SC et R). Ces résultats sont assez proches de ceux de Paya (2002), qui présente quelques caractéristiques prosodiques des "incidentes" en catalan, dans le sens où ses résultats indiquent également la présence de pauses inter-constituant et l'intonation descendante sur L. En revanche, aucune mention n'est faite d'un raccourcissement de la durée dans ses analyses sur les incidentes.

Les pauses silencieuses inter-constituant sont également fréquemment combinées à un allongement syllabique final très significatif sur L : 34.5 % des appositives relatives présentent ces deux caractéristiques (79 % des appositives introduites après un allongement syllabique très significatif sur L présentent une pause silencieuse

inter-constituant. L'une de ces pauses est à la suite de l'allongement, i.e. entre L et SC, dans 57.9 % des cas de co-occurrence entre les deux indices). La présence de la pause induit sans doute l'allongement syllabique. Sans tenir compte de n'importe quelle corrélation, l'appositive relative est séparée de son co-texte par deux pauses silencieuses dans 32.7 % de nos séquences. C'est le cas dans la séquence (23) ci-dessous, associée à la figure 4.17.

- (23) Beth L Holly went out #
 SC **which is a bit unsocial really but #**
 R hem # and # Beth had locked her room

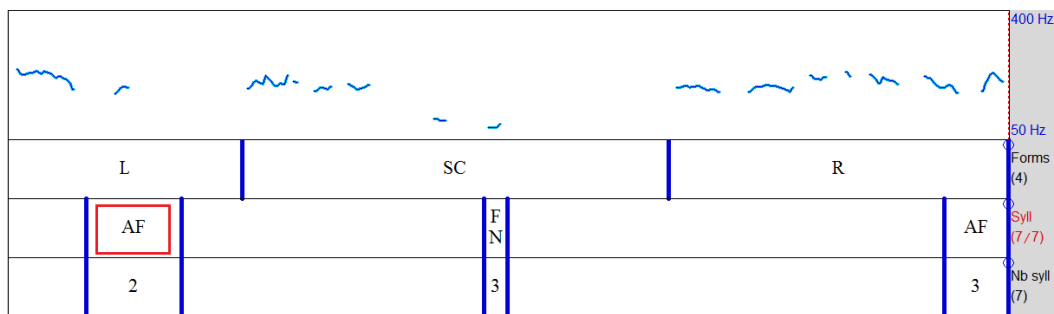


FIGURE 4.17 – Réalisation prosodique de la séquence (23) illustrée par la courbe Praat, où le repérage des syllabes finales avec leur poids phonémique montre l'allongement syllabique final très significatif de L (AF : syllabe accentuée finale; FNA : syllabe finale non-accentuée).

Avant d'introduire un nouvel événement décisif dans le récit d'une soirée chez des amis, Joey prend position qualitativement en tant qu'énonciatrice sur l'événement qu'elle vient d'énoncer. Une frontière entre la narration et son commentaire sous forme d'appositive relative est d'abord marquée par l'allongement de la dernière syllabe de L. "Out" est une syllabe accentuée composée de deux phonèmes, dont la durée phonémique est de 0.299 sec; cette durée atteint le second palier d'allongement pour les syllabes finales accentuées selon notre classification, traduisant un allongement très significatif. Cette syllabe présente un contour intonatif montant après une descente sur l'ensemble de L. L'ajout de cette légère montée permet au contour intonatif de jouer un rôle complémentaire aux marques rythmiques : bien qu'un changement soit annoncé, ce changement ne concerne pas les tours de parole. Après une pause silencieuse inter-constituant, Joey énonce SC beaucoup plus rapidement que le co-texte (à une vitesse de 5.29 syll/sec alors que L et R sont respectivement énoncés aux vitesses de 3.64 syll/sec et de 3.81 syll/sec), et sépare SC de R par une autre pause silencieuse inter-constituant. L'effet disruptif est visible sur la reprise de R, qui est difficile : le début de R est marqué par des pauses d'hésitation (remplies

et silencieuses).

Les corrélations évoquées jusqu'ici mettent en avant l'importance du rôle des pauses silencieuses inter-constituant en tant que frontières des appositives relatives, et permettent d'affirmer que les indices rythmiques sont clairement les plus impliqués dans les co-occurrences d'indices. La corrélation entre les trois indices rythmiques précédemment décrits (i.e. allongement syllabique significatif sur la syllabe finale de L, raccourcissement de la durée sur SC, et présence de pause silencieuse à la suite de SC) en est d'ailleurs une autre preuve : ces trois indices jouent un rôle conjoint dans 32.7 % des appositives relatives (66 % des appositives avec une durée plus courte possèdent les deux autres indices).

(24)	Rhianna	L	i mean my mum's pushing me to get my license	2.4 sec
		SC	(h) uh which i guess i should #	2.1 sec
		R	(h) but # well first of all #	2.6 sec

Alors que Rhianna sépare SC de L par une reprise de souffle audible (h), elle combine une pause silencieuse inter-constituant à une seconde reprise de souffle audible pour séparer l'appositive relative de R. SC est également plus courte que le reste de la séquence : le commentaire concessif de Rhianna portant sur le prédicat de L est bien réalisé avec un changement de rythme. Comme dans l'exemple (23), l'effet de cette rupture est visible sur R, où Rhianna insère une pause silencieuse intra-constituant assimilable à une pause d'hésitation entre deux marqueurs de discours résomptifs ("but" et "well"), après la reprise de souffle audible.

4.3 Discussion et conclusion

Notre analyse confirme que les différents types syntaxiques des subordonnées sous étude peuvent être distingués en fonction de leur degré d'autonomie prosodique. Ils peuvent être positionnés sur un continuum, de l'intégration (relatives déterminatives) à l'autonomie (appositives), en passant par des stratégies combinatoires intermédiaires (circonstancielle).

Les relatives déterminatives sont les constructions les plus intégrées à leur contexte (gauche) vocalement. Cette intégration se fait au moyen de ressources intonatives, soit par un phénomène direct de regroupement des unités, soit par des contours montants finaux à valeur continuative. Les relatives déterminatives se démarquent néanmoins par des moyens rythmiques, qui ne se combinent majoritairement pas à

l'intérieur d'une même séquence (une relative déterminative prototypique de notre corpus ne présente pas à la fois des pauses silencieuses intra-constituant et un allongement de la durée).

Les circonstancielles sont plus libres, à l'image de leur capacité pragmatique à projeter un cadre interprétatif sur plusieurs segments consécutifs ou bien à clore une unité. Leur intonation véhicule un changement; elle indique une altération des schémas interprétatifs et/ou informationnels plus qu'elle ne segmente les circonstancielles en unités isolées. Les formes postposées présentent des variantes non négligeables à partir de la réalisation prototypique en *afterthought* (Chafe, 1984), i.e. en post-rhème (Morel et Danon-Boileau, 1988). Les locuteurs montrent également plus de difficultés de calcul après la réalisation de ce type de subordonnée, surtout après les circonstancielles temporelles. Les segments à leur suite sont néanmoins davantage caractérisés par "hem", signalant une interruption locale, que par "uh", indexant une discontinuité plus forte (Shriberg et Lickley, 1993; Swerts, 1998; Clark et Fox Tree, 2002). À l'image des relatives déterminatives, nous n'observons pas de grande combinaison d'indices de rupture dans les occurrences.

Enfin, les appositives sont les constructions les plus indépendantes, la plupart d'entre elles montrant une autonomie prosodique totale. Ce sont les seules subordonnées à présenter un éventail aussi vaste d'indices de disjonction, tant au niveau segmental qu'au niveau suprasegmental. La rupture est donc plus perceptible, avec la plus grande différence significative de hauteur F0, et la plus faible distribution de contours continuatifs. Il n'est pas rare qu'une appositive relative présente à la fois un contour descendant sur L, une durée plus courte, et une pause silencieuse inter-constituant entre SC et R. Ce type de subordonnée est également fréquemment réalisé avec un décrochage de F0. Si la différence de hauteur intonative et la présence de pauses inter-constituant précédant ou à la suite de ces subordonnées est une caractéristique également prototypique des segments parenthétiques (Hirschberg et Grosz, 1992; Lelandais et Ferré, 2014), les appositives relatives se dissocient des parenthétiques en présentant rarement de compression, qu'elle soit rythmique (une durée réduite en même temps qu'un débit accéléré) ou intonative (une réduction de la gamme intonative).

La démarcation est majoritairement exprimée par des ressources rythmiques, dans tous les types de constructions subordonnées. Plus spécifiquement, les locuteurs s'appuient préférentiellement sur la durée et sur les pauses silencieuses inter-constituant pour marquer une rupture. En ce qui concerne les procédés in-

tonatifs, les rehaussements initiaux sont largement utilisés. La présence de ces indices suggère qu'une prosodie délimitative plus qu'intégrative (Mertens, 2008) est à l'œuvre dans les séquences comprenant des constructions subordonnées; leur placement indique quant à lui que les frontières prosodiques sont préférentiellement marquées rétrospectivement, immédiatement après le segment subordonné. La disjonction prosodique invite le co-locuteur à recontextualiser les éléments précédents. Seules les appositives relatives, constructions syntaxiques les plus disruptives, marquent directement la démarcation.

Si les traits rythmiques et les rehaussements intonatifs représentent les indices les plus mobilisés dans le marquage d'une frontière prosodique, les marques les plus combinées à d'autres à l'intérieur des séquences sont les pauses, aussi bien remplies (dans le cas des circonstancielles) que silencieuses (intra-constituant dans le cas des relatives déterminatives, inter-constituant dans le cas des appositives relatives). Ces pauses sont le plus souvent articulées à des marques suprasegmentales de disjonction telle qu'un raccourcissement de la durée des segments ou bien un contour intonatif descendant précédant la subordonnée. La répartition du travail de marquage de cette rupture est organisée temporellement entre les indices : dans les appositives relatives, type syntaxique où la rupture est la plus marquée, le contour conclusif se trouve majoritairement sur L lorsque les pauses silencieuses inter-constituant se trouvent davantage entre SC et R.

La combinaison des indices vocaux à l'intérieur de ces séquences, sensible à la temporalité et aux ajustements en temps réel du discours spontané, témoigne d'une flexibilité de planification et de production discursive, présentant des logiques prosodiques variées quant aux modes de présentation des subordonnées et des segments composant leur co-texte. Il a été vu que le co-texte des subordonnées est modulable en fonction des occurrences : les subordonnées peuvent être intégrées directement au groupe prosodique gauche, précédées d'un contour continuatif sur le co-texte, ou bien précédées de marques de rupture comme un contour intonatif terminatif ou une pause silencieuse. À l'instar de Dehé (2009) et de son étude des parenthétiques, nous soulignons que l'affirmation selon laquelle les subordonnées, qu'elles soient intégrées ou en rupture, n'affectent pas l'intonation des segments hôtes (De Vries, 2007) n'est pas envisageable : le simple fait que le co-texte gauche présente un ton continuatif ou un ton descendant terminal est l'une des manières par lesquelles le co-texte est impliqué dans l'insertion de la subordonnée, qu'elle consiste en une intégration ou en une juxtaposition.

Cette analyse s'est attachée à montrer que la prise en compte d'un vaste faisceau d'indices prosodiques démarcatifs permet une nouvelle appréhension de la subordination, dérivant de nombreuses interactions entre les indices segmentaux et suprasegmentaux. Le relevé et l'analyse des ressources prosodiques utilisées pour l'intégration (dans le cas des relatives déterminatives) et pour la démarcation (dans le cas des appositives relatives) des différents types de subordonnées mettent en valeur les groupes intonatifs en tant qu'unités plastiques et dynamiques. À l'intérieur des séquences sous étude, nous avons voulu mettre en valeur la façon dont les groupes intonatifs se forment dans la temporalité des séquences discursives et interactionnelles, ainsi que dans les ajustements entre les locuteurs. Notre relevé d'indices suggère que le marquage d'une rupture prosodique est un processus élaboré localement, sensible à la dynamique discursive, en adéquation avec le contexte référentiel, discursif et/ou interactionnel.

À partir de cette analyse prosodique, des tendances sont observables pour la production des subordonnées. Ces résultats seront confrontés aux données perceptives recueillies grâce à la mise en place d'un test autour de la rupture prosodique dans le Chapitre 5. Ce test de perception devra plus précisément confirmer que des frontières de nature vocale sont perceptibles dans les séquences contenant des subordonnées, que différents degrés de frontière peuvent être perçus, et que la rupture est plus perceptible dans les séquences contenant des appositives relatives. Si les pauses silencieuses sont les plus utilisées en tant que marqueurs de rupture, le test déterminera si elles représentent également l'indice le plus fiable dans la perception des frontières prosodiques.

Chapitre 5

Perception des frontières prosodiques

5.1 Introduction

Ce chapitre confronte les résultats de notre analyse prosodique sur la production des subordonnées aux données perceptives, recueillies grâce à la mise en place d'un test autour de la rupture prosodique. Après avoir introduit les enjeux majeurs de notre test, nous proposons un bilan de l'état actuel des études perceptives sur les frontières prosodiques. Nous exposons ensuite notre méthodologie, avant de donner nos résultats et leur interprétation.

À travers l'analyse de nos données sur une étude de production, nous avons vu que les locuteurs sont capables d'exprimer une disjonction entre deux constituants de manière audible. Les indices prosodiques les plus récurrents sont les pauses silencieuses inter-constituants, la variation de durée, ainsi que des rehaussements intonatifs initiaux. Ces indices sont le plus souvent employés seuls, bien que les combinaisons s'observent sur les constructions syntaxiques présentant un plus haut degré de rupture. Les co-locuteurs, quant à eux, utilisent supposément ces indices pour segmenter le flux de parole en unités et ainsi traiter l'information (Ostendorf et al., 1990; Collier et al., 1993). En revanche, les données recueillies sur la production des subordonnées ne permettent pas de déterminer l'influence des indices vocaux par rapport aux indices verbaux (i.e. syntaxiques, lexicaux) sur le traitement et l'interprétation des subordonnées, ni de savoir si les indices les plus utilisés par les locuteurs sont également les plus utilisés par les co-locuteurs lors de la segmentation du discours.

Les études récentes sur la perception prosodique des subordonnées sont peu nom-

breuses, mais Fromont et al. (2017) montre que les indices prosodiques seuls ne suffisent pas à modifier l'interprétation préférentielle d'une proposition. Cependant, cet article cible seulement l'optionalité des frontières prosodiques dans certains cas, puisqu'une structure syntaxique donnée peut avoir plusieurs phrasés intonatifs tout aussi acceptables (Watson et Gibson, 2005).

Le test de perception a donc pour but de répondre à deux questions principales : les auditeurs peuvent-ils assigner des degrés de force de frontière perçue entre deux propositions (l'une étant syntaxiquement subordonnée), et si oui, quelle est la relation de ces différents degrés d'une part aux indices prosodiques segmentaux et suprasegmentaux, et d'autre part aux différents types de constructions syntaxiques ?

Ce test de perception a pour but de confirmer (1) que des frontières de nature vocale sont perceptibles par des auditeurs naïfs dans les séquences contenant des subordonnées, (2) que différents degrés de frontière peuvent être perçus, et (3) que la rupture est plus perceptible dans les séquences contenant des appositives relatives. Si les pauses silencieuses sont les plus utilisées en tant que marqueurs de rupture, le test doit également déterminer si elles représentent l'indice le plus fiable dans la perception des frontières prosodiques.

5.2 Etudes antérieures sur les frontières en prosodie perceptive

5.2.1 La notion de frontière : théorique ou perceptuelle?

De nombreux corpus oraux comprennent une segmentation des données en unités prosodiques diverses, qui dépend en partie d'une annotation de "frontières" prosodiques (e.g. Svartvik et Quirk, 1980; Auran et al., 2004). Ces unités sont soit annotées manuellement par des experts, soit détectées automatiquement à partir de traits acoustiques. Bien qu'elles soient de taille variable en fonction des diverses théories phonéto-prosodiques, ces unités sont décrites comme essentielles à l'étude des relations entre les phénomènes prosodiques, syntaxiques, et discursifs.

L'utilisation des frontières dans la tradition britannique

La tradition britannique (e.g. Halliday, 1967; Crystal, 1969; Wells, 2006) considère que les unités informationnelles de la parole sont réalisées sous forme de groupes intonatifs (Halliday, 1985). L'unité intonative prototypique est une séquence d'items lexicaux et grammaticaux combinés sous un seul contour intonatif cohérent, le *tone-unit*. Les frontières ne sont abordées que pour définir et délimiter ces unités. Selon la tradition britannique, les *tone-units* sont délimités par de brèves pauses, souvent causées par une légère rupture de tempo (Chafe, 1988). Il est également observé que les *tone-units* coïncident avec les propositions syntaxiques, mais que certaines raisons forcent parfois les locuteurs à étendre une proposition à plusieurs unités intonatives (Chafe, 1994). Les *tone-units* peuvent être groupés dans des unités plus larges, les paragraphes (également paratons chez Wichmann, 2000). Les paragraphes sont des groupes de *tone-units* qui forment un mouvement intonatif global : la hauteur intonative décline progressivement d'un *tone-unit* à l'autre, jusqu'à ce qu'elle soit réinitialisée à la formation d'un nouveau paragraphe (Cruttenden, 1986). Les paragraphes font partie des éléments permettant de prévoir un tour de parole, et peuvent être délimités par des pauses, une rupture mélodique, et des allongements syllabiques (ibid.). Yule (1980) et Wichmann (2000) approfondissent cette notion et distinguent les paragraphes majeurs, délimités par une longue pause et caractérisés par une attaque haute, des paragraphes mineurs, qui montrent des pauses plus courtes et des attaques moins hautes. Yule (1980) trouve également un ton descendant très bas à la fin d'un paragraphe majeur, contrastant avec un ton à peine descendant à la fin d'un paragraphe mineur (Couper-Kuhlen, 1986). Bolinger (1972) remarque de son côté que l'ampleur des mouvements intonatifs peuvent établir une hiérarchie d'importance des unités de la parole. En revanche, Bolinger (1989, p. 81) réduit également la fonction de la démarcation prosodique à "établir une hiérarchie informelle de commencements et de fins, où les constituants majeurs peuvent être distingués des constituants mineurs". En revanche, la plupart de ces travaux portent sur une annotation de parole lue, et la confrontation de ces unités à la segmentation de la conversation spontanée remettent en cause les définitions traditionnelles (répétitions, hésitations, reprises), y compris celle des frontières décrites.

L'utilisation des frontières dans le modèle autosegmental de phonologie intonative

La notion de frontière possède également un sens particulier pour certaines théories phonologiques, abondamment utilisée en production comme en perception.

Dans la théorie de la phonologie prosodique (Selkirk, 1984; Nespor et Vogel, 1986) et ses différentes acceptions (Ladd et Campbell, 1991; Gussenhoven et Rietveld, 1992) ainsi que dans dans le modèle autosegmental de phonologie intonative (i.e. AM; Pierrehumbert, 1980; Beckman et Pierrehumbert, 1986; Ladd, 2008), la parole est considérée comme une structure hiérarchisée de domaines prosodiques. Les énoncés sont découpés en constituants qui sont hiérarchiquement dominés par des unités prosodiques de niveau supérieur dans la hiérarchie prosodique. Si nous nous intéressons à l'intonation en particulier, la plus grande unité prosodique marquée intonativement est l'IP (Intonational Phrase ou syntagme intonatif). À travers les langues étudiées jusqu'ici, une IP est définie par un ton de frontière à son extrémité droite (parfois à celle de gauche également), un allongement syllabique final, et une pause optionnelle après l'unité (Jun et Fletcher, 2014; Katsika et al., 2014). L'unité inférieure dans cette représentation phonologique est l'ip (intermediate phrase ou syntagme intermédiaire en français)¹. L'ip possède au moins un accent nucléaire, et est bornée à sa droite par un accent de syntagme (phrase accent; Beckman et al., 2005). L'IP est donc constituée d'une ou plusieurs ips, et se termine par un ton de frontière. De nombreuses études montrent que des indices corrèlent avec la force des frontières, et observent également que la force phonologique des indices de frontière est plus importante que la présence de ces indices elle-même (Frazier, et al., 2006). Pour Pierrehumbert et Hirschberg (1990), l'abaissement final reflète par exemple le degré de "finalité" d'un énoncé. De même, il a été montré que le degré de glottalisation à la fin des unités correspond à la hiérarchie prosodique (Fougeron et Keating, 1997), et que les pauses jouent un plus grand rôle à des frontières de haut niveau (Clifton et al., 2002). Cependant, l'accent est une nouvelle fois mis sur les unités phonologiques et leur hiérarchie. En ce sens, les annotations ne reflètent pas toujours les productions physiques et physiologiques sur lesquelles elles s'appuient.

À l'interface entre syntaxe et prosodie, les travaux de Delais-Roussarie (e.g. 2000; 2004; 2005; 2008; 2011) portent également sur les frontières prosodiques en français, avec un intérêt particulier pour l'organisation métrique des énoncés. En partant du constat que les structures syntaxiques et prosodiques ne sont pas isomorphes sans pour autant être complètement indépendantes l'une de l'autre, Delais-Roussarie propose un modèle de structuration prosodique impliquant un plan formel et un plan phonologico-phonétique, tous deux mis en relation. Le plan formel est en lien avec la structure syntaxique et repose sur l'assignation d'indices de force aux frontières syntaxiques (ou syntactico-prosodiques). Le plan phonologico-phonétique rend compte

1. L'existence du syntagme intermédiaire en anglais a fait l'objet de débats. cf. Ladd (1996) pour une discussion.

des réalisations prosodiques effectives possibles, et des découpages observés et/ou observables en surface. En remettant en cause l'existence de catégories prosodiques figées (e.g. Intonational Phrase, Accentual phrase) l'approche de Delais-Roussarie permet de dépasser des débats relatifs à la force et à la réalisation des frontières mais aussi d'avoir un regard nouveau sur la façon de gérer les interfaces (en particulier phonologie-syntaxe, et phonologie-phonétique). Cette approche est le fruit d'études montrant que l'interface syntaxe/prosodie est influencée par des contraintes de nature prosodique telle que la taille des constituants (Delais-Roussarie, 1995), ainsi que des études portant sur des structures syntaxiques spécifiques montrant une très forte corrélation avec une frontière prosodique majeure, telles que des incidentes (Delais-Roussarie, 2006) ou des disloquées (Delais-Roussarie et Post, 2008).

L'utilisation des frontières dans les approches acoustiques

L'approche IPO (*Dutch Institute for Perception Research*² a développé une théorie générale de la structure intonative, davantage basée sur les mouvements intonatifs que sur les groupes intonatifs (Cohen et 't Hart, 1967; 't Hart et Collier, 1975). Les indices considérés comme pertinents à l'identification des unités sont les pauses, le rehaussement de la ligne de déclinaison intonative, et l'allongement final (Hermes, 2006). Les travaux plus tardifs combinent l'analyse intonative des IPO et les approches autosegmentales de la structure prosodique (e.g. De Pijper et Sanderman, 1994), et observent grâce à des tests de perception que les frontières les plus fortes sont réalisées avec plus d'indices phonétiques de frontière (pauses et discontinuité mélodique). En revanche, ces tests portent également sur de la parole lue par des professionnels. Blaauw (1994) a également étudié les marqueurs de frontières prosodiques en comparant la parole lue à la conversation spontanée, particulièrement en ce qui concerne la localisation et la longueur des pauses. Elle trouve une corrélation positive entre le nombre de marqueurs prosodiques de frontière et la taille du constituant syntactico-prosodique marqué.

Barth-Weingarten (2016) propose une analyse et une description paramétrique des "césures" dans la parole, plutôt que des "frontières", terme présupposant selon elle l'existence d'unités à bornes fixes. Les césures sont créées par des discontinuités dans les paramètres prosodiques et phonétiques de la parole, qui se groupent plus ou moins à certains moments et sont de force variable. Elles incluent des discontinuités

2. L'institut néerlandais de recherche en perception se trouve à Eindhoven. Cet institut a développé un système de codage de l'intonation du néerlandais, adapté et appliqué ensuite à d'autres systèmes intonatifs comme celui de l'anglais (Willems et al., 1988), du français (Beaugendre, 1994) ou de l'italien (Quazza 1991).

de fréquence fondamentale, d'intensité, et de tempo. En se focalisant uniquement sur les discontinuités de la parole, cette approche a pour but d'éviter le problème des "frontières floues" (i.e. lorsqu'une frontière entre deux segments de parole est difficile à localiser, ou lorsqu'il est difficile de décider combien d'unités intonatives sont comprises dans un seul segment de parole) et celui de la circularité dans la description des unités intonatives et de leur interface avec d'autres dimensions organisationnelles langagières (e.g. unités syntaxiques, discursives, informationnelles). En utilisant comme point d'entrée les emplacements précédant les interventions des co-participants dans la parole, Barth-Weingarten (2016) observe que les césures incluent de manière minimale une syllabe suffisamment proéminente et un mouvement intonatif notable en fin d'énoncé, avec des tons descendants accompagnés de glottalisation. D'autres paramètres récurrents incluent un allongement syllabique notable en fin d'énoncé et une baisse d'intensité (souvent couplés à un relâchement audible des articulateurs lors de la production d'une plosive en fin d'énoncé), une configuration ouverte du conduit vocal, une expiration ou une respiration audible, ou un silence. L'étude ajoute donc davantage de traits phonético-prosodiques et évoque la variation dans le type et l'amplitude des changements de paramètres. Le continuum entre présence claire, présence réduite, et absence d'un indice permet de comprendre chaque changement de paramètre pour chaque indice isolé dans le flux de parole. Cependant, en prenant comme repère l'unité de tour de parole, la granularité de cette étude repose elle aussi sur une position théorique et méthodologique. Une hiérarchie entre les césures est également créée : celles réalisées à l'intérieur des tours de parole sont par définition plus faibles si leur domaine d'action local est inférieur (i.e. si la césure ne donne pas lieu à un changement d'ordre interactionnel). La mesure de ce dernier paramètre dépend également des positions théoriques et méthodologiques d'un annotateur expert.

Si ces approches ont grandement contribué à l'étude des unités prosodiques et ont permis de mettre en valeur le rôle de certains indices de frontières, toutes ces approches nécessitent et découlent d'une expertise en prosodie, qui va influencer l'annotation selon les théories adoptées. Le caractère méta-linguistique, dans un test, ne permet pas un jugement divorcé de toute théorie linguistique. Ce ne sont pas non plus des prises de décisions spontanées reflétant les mécanismes de réception en temps réel des co-locuteurs lors d'une conversation.

L'utilisation des frontières dans les études de perception

Dans cette recherche, nous appréhendons la force des frontières en tant que notion perceptuelle en elle-même, comme un phénomène mesuré directement dans un test sur des auditeurs naïfs, qui ne peuvent explicitement faire référence à des structures et phénomènes syntaxiques, phonologiques, ou prosodiques.

Notre test a été inspiré par divers travaux portant principalement sur la perception de frontières prosodiques par des auditeurs naïfs (De Pijper et Sanderman, 1994; Cole et al., 2008; Mo et al., 2008), visant à évaluer la corrélation entre un nombre d'indices prosodiques et les jugements quant à la présence de frontières prosodiques. Dans cette littérature, le terme de frontière prosodique est le plus souvent employé pour désigner une rupture perceptuelle (Cho et Hirst, 2006) produite par des moyens vocaux entre deux unités.

La perception des frontières en conversation spontanée a été étudiée en néerlandais (Strefkeerk et al., 1997; Buhmann et al., 2002), en suédois (Swerts, 1997), en anglais américain (Yoon et al., 2004; Mo et al., 2008; Cole et al., 2010; Cole et al., 2014), en français (Smith, 2009; Astésano et al., 2012; Roux et al., 2016; Simon et al., 2016), en kabyle et en hébreu (Mettouchi et al., 2007), ainsi qu'en coréen (Cho et Hirst, 2006) et en mandarin (Yang et Wang, 2002; Li et Yang, 2009). Bien que certains chercheurs comme Ladd (2008, p. 288) considèrent que les frontières prosodiques restent difficiles à décrire et à identifier avec constance, l'ensemble de ces travaux montrent que les auditeurs perçoivent les frontières de manière efficace et cohérente en parole spontanée.

Les tests de perception réalisés par des auditeurs naïfs sont extrêmement informatifs sur la segmentation du discours et son interprétation. Les auditeurs naïfs permettent d'éviter tout biais théorique, puisque les sujets ne sont pas formés à la tâche. Ils révèlent également des prises de décision spontanées dans un contexte relativement contrôlé. En revanche, toute demande explicite aux participants d'indiquer des frontières prosodiques élicite des jugements méta-linguistiques. Ces jugements dépendent de la conception de la tâche que se fait le participant. Certains travaux ont comparé les annotations des auditeurs naïfs à celles d'experts (Buhmann et al., 2002; Amir et al., 2004), travaillant généralement avec le système d'annotation ToBI (Silverman et al., 1992; cf. *infra* p. 182). Ces études permettent de tester la robustesse ainsi que la cohérence des annotations, et montrent généralement un fort taux d'accord entre naïfs et experts (e.g. Pagel et al., 1995; Auran et al., 2005; Simon et Christodoulides, 2016). L'étude de Roy et al. (2017) montre de plus que les

facteurs prosodiques influençant la perception de frontières pour des auditeurs naïfs sont les mêmes facteurs repérés en production dans les études réalisées en conditions expérimentales.

5.2.2 Les indices prosodiques liés à la perception d'une frontière

Les travaux dont nous nous inspirons montrent que les indices acoustiques corrélés à la perception d'une frontière prosodique sont liés à la présence d'une pause silencieuse (e.g. Yoon et al., 2007), à l'allongement syllabique final (e.g. Mo, 2008), à la présence de tons (Portes, 2002), et à un rehaussement initial de la ligne de déclinaison intonative (Wagner et Watson, 2010, p. 907-910). Les indices proviennent donc de phénomènes rythmiques et de phénomènes mélodiques (Astésano et al., 2012; Barth-Weingarten, 2016). Bien que l'intensité, la glottalisation, et la configuration articulaire du conduit vocal jouent également un rôle dans la perception des frontières (Mo, 2008; Barth-Weingarten, 2016), nous n'évaluons pas ces indices dans ce travail.

L'un des traits les plus utilisés à travers les langues pour la présence d'une frontière prosodique est la présence d'une pause silencieuse (Wightman et al., 1992; Carlson et Swerts, 2003; Yoon et al., 2007; Wagner et Watson, 2010; Roy et al., 2017). En revanche, beaucoup d'études antérieures se basent sur de la parole lue (e.g. de Pijper et Sanderman, 1994). La corrélation entre la présence d'une pause silencieuse et la perception d'une frontière est moins directe en parole spontanée car la pause silencieuse peut aussi être une marque d'hésitation (Mertens et Simon, 2013). Dans des extraits de map-task et des extraits de débats télévisés en français, Smith (2009) montre que les pauses (silencieuses et remplies de plus de 150 ms) favorisent la perception d'une frontière, mais ne constituent pas des indicateurs fiables. Dans des travaux sur le français (Duez, 1985; 1991) et sur le coréen (Cho et Hirst, 2006), la pause silencieuse devient un indice décisif lorsqu'elle est d'une durée supérieure à 200 ms. Ce seuil de 200 ms³ est fréquemment repris (e.g. Candea, 2000) depuis Fraisse (1974). Plus la pause s'allonge, plus la frontière perçue est

3. Grosjean et Deschamps (1973; 1975) proposent toutefois pour leur étude comparative de l'anglais et du français la durée fixe de 250 ms comme seuil minimum (corpus radiophonique); Lacheret-Dujour et Victorri (2002) utilisent le seuil minimum de 300 ms pour noter les pauses silencieuses, toujours sur un corpus radiophonique. Campione et Véronis (2004) classifient quant à eux les pauses silencieuses en trois catégories. Les pauses brèves de leur corpus en français spontanée sont les pauses égales ou plus courtes que 200 ms, les moyennes sont égales à ou entre 200 ms et 1000 ms, et les longues sont égales ou plus longues que 1000 ms. Les mêmes auteurs (2002) notent que les pauses silencieuses longues n'apparaissent qu'en conversation spontanée.

importante : Swerts (1997) observe sur des monologues spontanés suédois que les pauses longues (plus de 250 ms) ont tendance à être associées à des frontières plus fortement perçues. Pour Duez (1985), la durée de la pause silencieuse est d'ailleurs l'indicateur le plus important, et l'influence d'autres indices augmente dès lors que la durée de la pause diminue. Lehiste (1979, p. 200) montre ainsi que les indices suprasegmentaux interagissent dans la perception d'une rupture prosodique, et que ces indices se contrebalancent. La difficulté générale d'identifier le degré d'influence des autres indices prosodiques dans la perception des frontières est confirmée par Mo et Cole (2010).

Un second indice rythmique joue un rôle important dans la segmentation du discours. En anglais conversationnel (Kreiman, 1982; Mo, 2008), les voyelles accentuées positionnées avant les frontières sont significativement plus longues que celles positionnées ailleurs. Cette durée vocalique est fortement corrélée à la perception d'une frontière (ibid.). Sur une étude en hébreu conversationnel (Amir et al., 2004), l'allongement syllabique final se trouve l'indice le plus prévalent. En revanche, il est encore ici difficile de déterminer le poids de cet indice par rapport aux autres : les auteurs signalent que l'allongement syllabique apparaît en co-occurrence avec d'autres indices (i.e. pause silencieuse, rehaussement initial de la F0) dans une grande majorité de cas.

En ce qui concerne le poids des indices mélodiques, les études examinant le rôle des contours intonatifs indiquent que la corrélation entre contour descendant vs. contour montant et la perception de frontière est instable (Simon et Christodoulides, 2016). De Pijper et Sanderman (1994) observent que la discontinuité mélodique est le seul indice vocal à se produire en isolation pour marquer des frontières. En français, Portes (2002) montre que les frontières prosodiques perçues comme faibles sont associées à des contours montants alors que les frontières perçues comme fortes sont liées aux contours descendants. Smith (2009) constate quant à elle que l'amplitude du mouvement intonatif se trouve meilleur corrélat de la force de la frontière perçue que la direction du ton.

Au long d'un paragraphe vocal (i.e. un groupe de *tone-units* formant un mouvement intonatif global; Cruttenden, 1986), la hauteur de F0 décline naturellement de manière progressive. Le rehaussement de cette ligne de déclinaison est également un indice de frontière ('t Hart et al., 1990).

L'intensité, la glottalisation, et la configuration articulaire du conduit vocal

ne sont pas mesurés dans ce travail bien qu'ils jouent également un rôle dans la perception des frontières. Les disfluences sont également exclues à la fois des stimuli et de l'analyse, étant donné la courte durée de nos stimuli et leur grand nombre de variables. Barth-Weingarten (2016) montre également que la présence ou l'absence de certains de ces paramètres prosodiques combinés les uns aux autres participe à la perception de frontières plus ou moins fortes dans la parole conversationnelle.

5.2.3 La force des frontières

L'une des questions motivant notre test de perception concerne le nombre de niveaux de frontières pouvant être distingués. Certaines études demandent aux auditeurs de choisir sur une échelle à dix degrés, mais le nombre de catégories est généralement bas (de Pijper et Sanderman, 1994). Grover et al. (1998) et Auran et al. (2005) montrent qu'une échelle de quatre valeurs permet de transcrire la force des frontières de manière plus cohérente. Le nombre de valeurs dépend généralement des principes théoriques à propos des frontières prosodiques. L'influence grandissante exercée par le modèle métrique autosegmental de phonologie (AM; e.g. Pierrehumbert, 1980; Beckman et Pierrehumbert, 1986) a conduit à l'élaboration d'un système de notation de l'intonation : ToBI (Tone and Break Indices) qui intègre à la fois la notation tonale proposée par Pierrehumbert (voir Pierrehumbert, 1980) et celle des degrés de frontière, selon une échelle à cinq niveaux (Silverman et al., 1992). Cet outil est d'abord développé pour l'annotation prosodique de l'anglais américain pour une utilisation en synthèse vocale (Silverman et al., 1992; Beckman et al., 2005; Bruggos et al., 2006). La force de la frontière prosodique perçue entre les constituants, exprimée par les tons, est notée sur une échelle de 0 à 4 (i.e. Break Index). En anglais américain, [0] représente la cohésion entre les items (comme le phénomène de flapping⁴); [1] représente le petit degré de jointure que l'on attend entre les mots; [2] est réservé à l'incertitude; [3] et [4] représentent respectivement les frontières de l'ip (syntagme intermédiaire; cf. supra p. 176) et de l'IP (syntagme intonatif; cf. supra p. 176).

D'autres études s'intéressant au degré de consensus dans la perception des frontières parmi des unités variées choisissent d'identifier des "frontières de consensus" aux emplacements où une certaine proportion de participants ont identifié une

4. Le flapping est un phénomène consonantique observé dans de nombreuses variétés de l'anglais, comme en anglais américain, en anglais australien, ou en anglais néo-zélandais. Le flapping neutralise par réduction le contraste entre /t/ et /d/ à une battue voisée, le bout de la langue frappant la crête alvéolaire. Ce phénomène est particulièrement observable lorsque ces occlusives sont produites entre deux voyelles.

frontière (Auran et al., 2005; Smith, 2009). La force des frontières peut également être calculée comme la proportion de sujets indiquant avoir perçu une frontière à un emplacement donné, et exprimée comme une valeur entre 0 et 1 (Cole et al., 2010; Simon et Christodoulides, 2016).

Dans ce test, nos hypothèses sont les suivantes : (1) la force des frontières perçue varie en fonction du type syntaxique de subordination; (2) la perception des frontières ne s’appuie pas sur les mêmes indices prosodiques en fonction du type syntaxique de subordination.

5.3 Corpus et stimuli

Un ensemble d’extraits audio de notre corpus de conversation semi-spontanée en anglais britannique, ENVID, a été choisi. Les extraits sont d’une durée variable de 0.7 à 6.5 secondes (durée moyenne de 4 secondes), et proviennent de sept locuteurs différents de notre corpus (cinq femmes, deux hommes). La préparation des stimuli est détaillée ci-dessous.

5.3.1 Préparation des stimuli

Deux sets de stimuli distincts ont été sélectionnés :

- 3 courts extraits sans aucun trait prosodique assimilé à la création d’une rupture (2 extraits de locutrices, 1 extrait d’un locuteur; les extraits comportant tout type de marqueur d’hésitation vocal ou verbal ont été exclus)
- 3 courts extraits présentant un phénomène de rupture claire (1 extrait d’une locutrice, 2 extraits de locuteurs) combinant chacun trois indices de rupture (i.e. pause silencieuse, allongement syllabique final, et rehaussement intonatif initial) au même endroit.

Les critères de sélection de ces extraits authentiques étaient les mêmes pour les deux sets : la parole provient d’un seul locuteur, le signal est clair et ne présente aucun bruit. La séquence comprend trois unités : une proposition syntaxique subordonnée (un extrait avec une appositive, un extrait avec une circonstancielle, et un extrait avec une relative déterminative), ainsi qu’une proposition précédant cette subordonnée (i. e. L, le co-texte gauche), et une proposition suivant cette subor-

donnée (i.e. R, le co-texte droit). Chaque extrait est donc analysable en tant que séquence L - SC - R. Pour le set de stimuli présentant un phénomène de rupture claire, la rupture se situe entre la proposition précédant la subordonnée (L) et la subordonnée (SC). La rupture est donc réalisée par une pause silencieuse entre L et SC, un allongement de la syllabe accentuée finale de L, et un rehaussement de F0 initial sur la première syllabe de SC.

Afin de s'assurer que ces deux sets de stimuli soient clairement perçus comme sans rupture pour le premier, et clairement avec rupture pour le second, un pré-test auprès de six auditeurs naïfs anglophones natifs a été mis en place. Ce pré-test avait pour but de valider les stimuli avant toute manipulation. Les deux sets de stimuli ont été assemblés et randomisés en un seul bloc. Au signe des auditeurs, chaque extrait son a été lu deux fois sur ordinateur par l'expérimentatrice. Les auditeurs disposaient d'une transcription orthographique sur papier pour chaque extrait. Cette transcription ne comportait aucune ponctuation ni aucun saut de ligne (sauf nécessaire à la largeur de la page) afin de ne donner aucun indice structurel. La transcription du groupe intonatif de L (avant la subordonnée) figurait sur le côté gauche de la feuille, la transcription des groupes intonatifs de SC (la construction syntaxique subordonnée) et de R (après la subordonnée) sur le côté droit. Un seul espace blanc figurait donc sur la ligne, entre L et SC. Les auditeurs ont reçu pour consigne de noter une barre verticale entre les groupes intonatifs L et SC s'ils percevaient une rupture ou frontière dans la parole. Pour ce test, le taux d'accord entre participants est de 100 % : aucune rupture n'est perçue pour le premier set d'extraits, et tous les extraits présentant une rupture claire sont perçus comme tels.

Le set de stimuli sans rupture a donné lieu à cinq manipulations différentes à l'aide du logiciel Praat (Boersma et Weeninck, 2013), détaillées dans la section 5.3.2. Pour l'ensemble du set (1 appositive relative, 1 circonstancielle, 1 relative déterminative), les modifications suivantes ont été réalisées par synthèse pour créer de nouveaux énoncés :

- ajout d'une pause silencieuse de 500 ms entre L et SC
- ajout d'un ton descendant sur la syllabe nucléaire du groupe intonatif de L
- ajout d'un ton montant sur la syllabe nucléaire du groupe intonatif de L
- allongement significatif de la syllabe accentuée finale de L
- réinitialisation de la hauteur intonative entre la syllabe finale de L et la syllabe initiale de SC

Les différents sets donnent 21 extraits et peuvent être listés comme tels :

- 3 extraits authentiques sans rupture (séquence L - SC - R)
- 3 extraits authentiques avec rupture
- 3 extraits manipulés avec une pause silencieuse de 500 ms
- 3 extraits manipulés avec un ton descendant
- 3 extraits manipulés avec un ton montant
- 3 extraits manipulés avec un allongement syllabique final significatif
- 3 extraits manipulés avec une réinitialisation de la hauteur intonative

Une version délexicalisée de ces différents sets (21 extraits au total) a également été créée à l'aide des filtres de Praat (pass Hann band, from 0 to 500 Hz pour les extraits dont les locutrices sont des femmes, 0 to 450 Hz pour les extraits dont les locuteurs sont des hommes). Le filtrage passe-bas permet d'éliminer la majeure partie des indices segmentaux : le contenu sémantique de la production n'est pas accessible. Seul le contour mélodique de la production est perceptible. Les versions délexicalisées doublent le nombre de sets, totalisant 42 extraits.

10 distracteurs viennent compléter les stimuli. Ces distracteurs ont été choisis selon les mêmes critères que les constructions authentiques (i.e. parole provenant d'un seul locuteur, signal clair et sans bruit), à la différence que les traits prosodiques des distracteurs ne sont pas contrôlés. Ils se décomposent également en séquences L - SC - R. Le tableau 5.1 récapitule les différents types de stimuli utilisés dans le test, ainsi que leur nombre.

	lexicalisés	délexicalisés	TOTAL
authentiques sans rupture	3	3	6
authentiques avec rupture	3	3	6
pause silencieuse de 500 ms	3	3	6
ton descendant	3	3	6
ton montant	3	3	6
allongement syllabique final	3	3	6
réinitialisation de la F0	3	3	6
distracteurs	5	5	10
TOTAL	26	26	52

TABLE 5.1 – Tableau récapitulatif des différents stimuli composant le test de perception des frontières.

Les 52 fichiers sons (12 stimuli authentiques et 30 stimuli manipulés + 10 distracteurs) ont été randomisés dans un tableau Excel. Un numéro a été assigné à

chaque son (e.g. sound 001, sound 002). Les autres colonnes du tableau détaillent le type syntaxique auquel appartient la construction SC dans chaque extrait, son locuteur, ainsi que son nombre d'indices de rupture et leur détail.

5.3.2 Méthodes de manipulation des stimuli

Les différentes manipulations réalisées à l'aide du logiciel Praat (Boersma et Weenink, 2013) sont centrées autour de deux techniques principales : la synthèse vocale et le filtrage. Nous justifions ici l'emploi de ces méthodes ainsi que les différents choix de calibrage des indices de rupture sur les stimuli.

La synthèse

La synthèse permet de combiner la prosodie d'origine des locuteurs à des variations mélodiques et temporelles artificielles et localisées. Le contenu sémantique de la parole reste accessible. En revanche, l'un des désavantages majeurs de cette technique porte sur le rendu perceptif, qui manque quelquefois de naturel. Deux manipulations principales ont été réalisées grâce à la synthèse (sélection d'un son dans Praat, puis Analyse Periodicity, puis To Manipulation). Nous avons pu d'une part modifier la nature des tons nucléaires (descendant, montant), et créer une réinitialisation de la courbe mélodique entre L et SC. D'autre part, nous avons manipulé la durée de certains constituants. Une pause silencieuse de 500 ms entre L et SC a été créée grâce à la fonction édition du logiciel Praat, et la durée syllabique de la syllabe finale accentuée de L a été allongée grâce à l'ajout de points de durée dans la fonction manipulation de Praat.

En ce qui concerne les manipulations des tons nucléaires, la trajectoire de la F0 est modifiée sur une syllabe cible. Nous avons stylisé la courbe sous Praat, de façon à ce que les contours soient les plus clairs possibles. L'excursion réalisée commence à l'onset de la syllabe en question, et termine après la coda. L'ampleur de l'excursion est définie en demi-tons d'après les travaux de t'Hart (1981), Rietveld et Gussenhoven (1985), Traunmüller et Eriksson (1995), et ceux de Kakouros et Rasanen (2016). Rietveld et Gussenhoven (1985) montrent qu'une variation de 1.5 tons est perceptible. Nous avons choisi une ampleur d'excursion variable entre 1.5 et 4 demi-tons selon les extraits, de façon à ce qu'une différence soit perceptuellement remarquable mais aussi à ce que le mouvement reste naturel au sein de l'énoncé. Des ajustements de hauteur ont été réalisés sur l'ensemble de l'extrait lorsque le change-

ment de ton influait sur d'autres paramètres mélodiques (e.g. réinitialisation de la courbe mélodique, saut d'octave). La hauteur F0 moyenne de chaque extrait a été mesurée avant et après manipulation, de façon à limiter tout changement perceptif global sur la courbe de l'extrait.

La rehaussement initial est créé sur la syllabe initiale de SC. La différence de hauteur entre la syllabe finale de L et l'initiale de SC est également définie en demi-tons, mais varie entre 4 et 6 demi-tons selon les extraits. Après le test de plusieurs seuils, le minimum de 4 demi-tons nous a paru créer la différence la plus claire et la plus naturelle (Kakouros et Rasanen, 2016). À l'image des modifications relatives aux tons nucléaires, des ajustements de hauteur ont été réalisés sur l'ensemble de l'extrait dès lors que le rehaussement de F0 influait sur d'autres paramètres mélodiques (e.g. saut d'octave).

En ce qui concerne l'allongement de la durée syllabique, nous nous appuyons sur les travaux de Duez (1985; 1987, p. 222), montrant qu'un allongement de 50 % est perceptible en tant que frontière par l'auditeur. La durée syllabique de chaque syllabe accentuée finale de nos stimuli originaux a été extraite et allongée de 50 %, grâce à l'ajout de points de durée dans l'outil manipulation de Praat. Dans notre analyse prosodique de la production des subordonnées, nous avons défini deux paliers d'allongement à partir de nos données, en fonction du poids phonémique de chaque syllabe. Nous nous sommes assurée que chaque syllabe allongée par la synthèse atteigne le second palier d'allongement (très significatif). Les quelques syllabes allongées de 50 % n'atteignant pas ce second palier ont été allongées de 75 %.

Les pauses silencieuses ont été ajoutées entre L et SC grâce à l'éditeur de son Praat. Puisque chaque stimulus est extrait d'un dialogue de 30 min, une pause silencieuse de 500 ms provenant du dialogue original correspondant (sans prise de parole ni bruit) a été insérée entre L et SC dans chaque extrait. Ce choix résulte de l'observation des pauses silencieuses extra-constituants dans notre propre corpus (d'une durée moyenne de 0.56 s, et d'une médiane de 0.55 s), ainsi que de notre volonté d'un indice de frontière à la fois représentatif du corpus et sans équivoque. Nous suivons donc les résultats antérieurs selon lesquels la pause silencieuse devient un indice décisif lorsqu'elle est d'une durée supérieure à 200 ms. Étant donnée la courte durée de nos stimuli (1 seconde pour les plus brefs), nous avons choisi une pause relativement longue sans pour autant que cette dernière soit d'une durée plus importante que celle de la parole dans l'extrait.

Le filtrage

L'écoute d'énoncés dont la transcription est disponible donne des indices sur la façon dont les auditeurs naïfs perçoivent différentes constructions syntaxiques réalisées avec différents traits intonatifs. En revanche, les auditeurs ont accès aux informations verbales et vocales : il est impossible de distinguer l'influence des indices verbaux de celle des indices vocaux sur le choix de segmentation des auditeurs. Le moyen de jauger l'importance des deux modalités est de rendre les informations verbales inaccessibles. Le filtrage nous a paru la méthode la plus appropriée à la délexicalisation de la parole. Le test porte sur la contribution des indices prosodiques dans leur ensemble, et cette méthode ne dissocie pas rythme et intonation : à l'écoute d'énoncés filtrés, les auditeurs ont accès aux indices mélodiques (e.g. tons nucléaires) et temporels (e.g. pauses, débit, durée). En revanche, l'influence de certains indices segmentaux est neutralisée et l'identification des segments n'est donc pas possible. Puisque nous ne manipulons ni ne contrôlons l'intensité dans cette étude, certains indices phonotactiques peuvent toutefois transparaître avec les intensités relatives des segments. Nous avons donc varié la fréquence filtrée en fonction de chaque énoncé. L'un des inconvénients du filtrage est la difficulté de contrôler en détail les informations filtrées par rapport à celles qui parviennent aux auditeurs (Crystal, 1969, p. 74; Horgues, 2010).

Si le signal sonore obtenu par le filtrage nécessite quelques instants d'adaptation aux auditeurs naïfs, il est semblable au bruit étouffé d'une conversation entendue d'une autre pièce. Nous avons utilisé le filtrage Pass Hann band dans Praat, avec une fréquence de 0 à 450/500 Hz. Le signal sonore obtenu étant très bas en amplitude, chaque extrait a ensuite été amplifié (5 dB) dans Audacity.

5.4 Participants et procédure

5.4.1 Participants

Nos 52 stimuli ont été présentés à 30 participants dans un ordre aléatoire via une interface web (cf. infra p. 189). Une moitié des participants a été recrutée en ligne via des listes de diffusion mail ainsi que des réseaux sociaux. La seconde moitié des participants a été recrutée en personne à Bristol (Royaume-Uni). Pour cette seconde moitié, le test a été réalisé en présence de l'expérimentatrice. Des questions préliminaires sur l'interface web nous ont permis de nous assurer que tous les participants étaient anglophones natifs, qu'ils n'avaient aucun trouble auditif à

leur connaissance, qu'ils n'avaient aucune expérience d'annotation prosodique, et qu'ils avaient des écouteurs à disposition. Si au total 30 participants britanniques ont réalisé la tâche (âgés de 19 à 45 ans, moyenne : 24 ans), les réponses de 8 participants (tous recrutés en ligne) ont été exclues de l'analyse. Ces 8 participants ne remplissaient pas l'une ou plusieurs des conditions précédemment évoquées.

5.4.2 Le support de test

Le test a été réalisé sur eSurv (2017), une plateforme open-source de réalisation de questionnaires. Cette plateforme a pour avantage de présenter un éditeur de questions simple et maléable, ainsi que de supporter du contenu multimédia intégré dans les questions. Elle propose également l'hébergement gratuit du questionnaire réalisé, accessible et diffusable aux participants par un simple lien. Les participants sont identifiés par leur adresse IP. Une fonctionnalité permet de bloquer l'accès aux adresses IP une fois le test réalisé, ce qui empêche les participants de répondre au test deux fois. Les réponses des participants sont exportables au format CSV.

La diffusion d'un test en ligne permet d'obtenir un nombre de participants supérieur à ce que permettraient les moyens physiques d'une réalisation en laboratoire. Le recrutement de participants anglophones est facilité par internet, comparé au nombre d'anglophones natifs disponibles le même jour pour se rendre à l'Université de Nantes. Une réalisation en ligne peut également augmenter la diversité de l'échantillon des auditeurs en termes d'âge et de catégorie socio-professionnelle, si le test est diffusé par le biais de réseaux suffisamment variés. En revanche, l'échelle de temps de complétion de l'étude est différente d'une réalisation en laboratoire : la diffusion par réseaux (e.g. mailing lists et réseaux sociaux) prend davantage de temps et nécessite plusieurs relances. Les autres désavantages de la diffusion en ligne par rapport à un test réalisé en laboratoire sont généralement liés au manque de contrôle sur les paramètres précis de présentation des stimuli (e.g. volume, port d'écouteurs), ainsi qu'au manque de supervision du test par l'expérimentateur. Les auditeurs disposent par exemple de matériel informatique dont la qualité peut varier. Si notre test nécessite un casque pour des conditions de réalisation optimales, nous ne disposons d'aucun moyen de nous assurer que les auditeurs suivent à la lettre les instructions d'écoute. De même, le test de perception est relativement long à compléter : si le texte explicatif demande aux participants de s'assurer qu'ils ont au moins 40 minutes à consacrer au test, nous n'avons aucun moyen de tester la fatigue auditive et attentionnelle des participants sur les dernières questions. La

réalisation d'un test en ligne facilite aussi les abandons : la capacité de concentration des participants est moins importante lors de la complétude d'un test sur internet, en isolation dans un environnement familial. Il est également possible que le test de perception n'occupe qu'une seule des nombreuses fenêtres ouvertes dans leur navigateur internet. Un certain nombre d'études décrites dans Sprouse (2011) et dans Woods et al. (2005) ont toutefois comparé les réponses de participants d'études en ligne à celles de participants d'études réalisées en laboratoire. La grande majorité des résultats indique une différence non-significative entre les deux groupes de réponses (la seule différence concerne les expériences où le temps de présentation des stimuli est très court et contrôlé), ce qui confirme la fiabilité des études en ligne. L'un des arguments en faveur de la réalisation de tests en ligne consiste à augmenter le nombre de participants au test de façon à réduire le taux d'erreur et à pallier le manque de contrôle sur la présentation des stimuli (Woods et al., 2015). L'étude d'Ito et al. (2017) montre également qu'une population hétérogène de participants naïfs recrutés en ligne et présentant de grandes variations individuelles peut donner lieu à des résultats similaires à ceux d'une expérience réalisée en laboratoire lors d'une tâche.

Pour une plus grande commodité sur l'interface de test, les 52 fichiers sons composant nos stimuli ont été transformés en vidéos MP4 avec Movie Maker, et déposées sur une playlist non-répertoriée sur Youtube. Chaque vidéo reprend le nom des fichiers sons randomisés (e.g. sound 001, sound 002).

5.4.3 Procédure

Les participants cliquant sur le lien du questionnaire sont dirigés sur la première page de test, où un texte explicatif expose brièvement les enjeux de l'étude. Ce test explicatif est disponible en annexe p. 351.

Bien que le questionnaire soit anonyme, les participants répondent sur la seconde page à six questions d'informations personnelles (pays de résidence, langue maternelle, âge, sexe, port d'écouteurs, présence/absence de trouble auditif). Ces questions ont pour but de filtrer les participations d'un public non-cible (e.g. un participant anglophone non-natif).

Les questions débutent sur la troisième page. La structure est identique pour chacune d'entre elles et les questions sont au nombre de dix par page. La figure 5.1

plus bas illustre l’interface de test. Chaque question présente la vidéo de l’extrait son numéroté (e.g. sound 001), ainsi qu’une transcription orthographique des extraits lexicalisés face à la vidéo. La transcription ne présente aucune ponctuation ni saut de ligne. L’endroit de réalisation potentielle d’une frontière prosodique (i.e. entre L et SC) est marqué d’un astérisque rouge dans la transcription (e.g. “with this guy * who does it he does like a top gear programme”). Sous la vidéo (à gauche de la fenêtre de test) et la transcription (à droite de la fenêtre), une échelle de Likert à cinq niveaux présente les différents boutons : pas de frontière (no boundary); frontière faible (weak boundary); incertain (uncertain); frontière (boundary); frontière forte (strong boundary). Cette échelle s’inspire des niveaux de perception définis par l’outil ToBI (Silverman et al., 1992) et ceux de Carlson et Swerts (2003) en proposant une notation relativement précise. Les degrés de force sont néanmoins adaptés aux auditeurs naïfs, en ne faisant référence à aucune unité ou structure prosodique. La consigne figure en chapeau à chaque page du test : “veuillez noter la présence ou l’absence de frontière à l’endroit *” (*Please rate the presence or absence of a boundary at **).

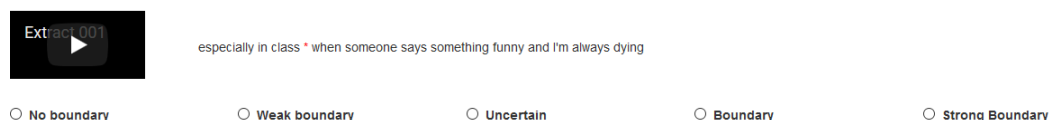


FIGURE 5.1 – Interface de la troisième page du test de perception réalisé sous eSurv, où figurent les premières questions. Chaque question présente la vidéo d’un stimulus ainsi que sa transcription, suivi de l’échelle graduelle de force de frontière perçue à cinq niveaux.

En ce qui concerne les stimuli délexicalisés, l’extrait son n’est composé que de L et de SC, et la transcription à droite de la vidéo est la suivante : “X * X”, où chaque “X” représente un groupe intonatif. La figure 5.2 présente l’affichage d’un stimulus délexicalisé.

La page présentant les dix dernières questions se termine par un espace ouvert aux commentaires des participants, sous la forme d’une question (i.e. “Voudriez-vous laisser un commentaire?”; *Is there any comment you would like to make?*).

5.4.4 Analyse des données

Des ratios ont été calculés successivement pour chaque variable en respectant l’échelle à cinq niveaux. Une série de modèles linéaires mixtes généralisés (*GLMMs*

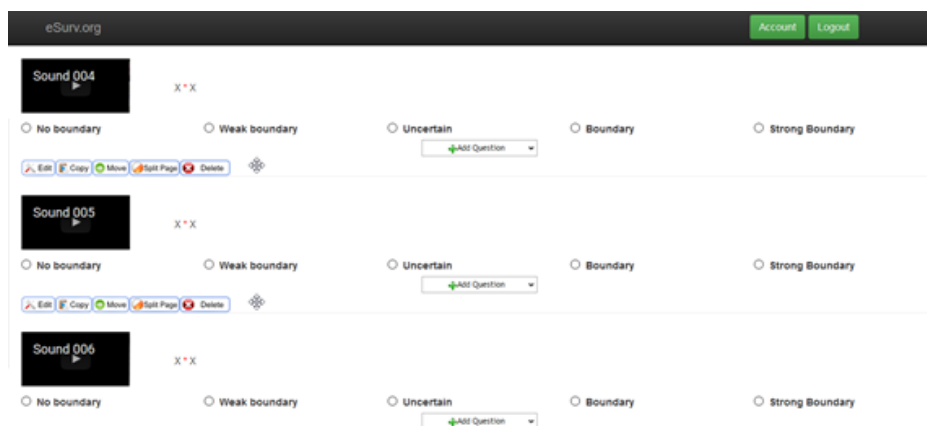


FIGURE 5.2 – Interface de la troisième page du test de perception réalisé sous eSurv, où figurent des questions portant sur des stimuli délexicalisés. Chaque question présente la vidéo d’un stimulus, suivi de l’échelle graduelle de force de frontière perçue à cinq niveaux. La transcription contient seulement deux “X” (un pour chaque groupe intonatif) séparés d’un astérisque marquant l’endroit de potentielle réalisation d’une rupture prosodique.

estimés par la méthode du maximum de vraisemblance) a également été réalisée avec le programme R 3.4.0 (R *Core Team*, 2017) et le *package* lme4 (Bates et al., 2017), afin d’expliquer le degré d’influence de chaque variable sur la perception des frontières. Les corrélations entre les indices ont également été calculées avec le programme R 3.4.0. Nous ne mesurons pas l’accord inter-annotateur à l’aide du Kappa de Fleiss : étant donné le nombre de catégories à choisir pour les participants (cinq) et le nombre relativement réduit de stimuli, la pertinence de cet outil pour décrire les tendances d’accord est mise en cause (Gross, 1986; Powers, 2012; Noël, 2013; Ito et al., 2017; Roy et al., 2017).

5.5 Résultats

Le premier résultat à noter est que les différences de scores entre les stimuli authentiques et les stimuli manipulés ne sont significatives pour aucune des variables. Cela signifie que nos participants ont jugé du poids des frontières de manière égale dans les deux types de stimuli.

Nous avons d’abord exploré les interactions potentielles entre les trois types de constructions syntaxiques (facteur fixe = Type; valeurs = appositive, circonstancielle, relative) et le jugement du poids des frontières (facteur fixe = Scores; valeurs = frontière forte (strong boundary), frontière (boundary), frontière faible (weak boundary), incertain (uncertain), pas de frontière (no boundary)). Des interactions

ont été trouvées pour les scores “frontière forte” (*strong*) et “faible” (*weak*). L’effet principal des jugements “frontière forte” (facteur fixe = Strong; valeurs = oui, non) est significatif pour les appositives ($\beta = 1.72$, $SE = .27$, $p < 0.0001$) seulement. En revanche, les différences avec les circonstancielles et les relatives ne sont pas significatives. Toutefois, comme le montre le tableau 5.2, ce même effet principal sur les appositives n’est pas significatif pour les mêmes stimuli où la parole est filtrée. Ce résultat est donc propre aux stimuli comportant à la fois les informations prosodiques et les informations lexico-syntaxiques.

De même, l’effet principal des jugements “frontière faible” (facteur fixe = Weak; valeurs = oui, non) est significatif pour les circonstancielles ($\beta = 0.36$, $SE = .17$, $p < 0.05$) seulement. Bien que la différence ne soit significative qu’avec les appositives ($\beta = -1.46$, $SE = .27$, $p < 0.0001$), ce même effet principal sur les circonstancielles n’est pas significatif pour les mêmes stimuli filtrés. Aucune interaction significative n’a été trouvée pour les relatives déterminatives. Cependant, dans les mêmes stimuli contenant de la parole filtrée, l’effet principal des jugements “frontière forte” est significatif pour les relatives déterminatives ($\beta = 0.65$, $SE = .31$, $p < 0.05$).

	filtré	non filtré
appositives	frontière forte ***	n.s
circonstancielles	frontière faible *	n.s
relatives déterminatives	n.s	frontière forte *

TABLE 5.2 – Interactions significatives entre les types syntaxiques et les scores de frontière dans la parole filtrée et la parole non-filtrée (***) = $p < 0.0001$; (*) = $p < 0.05$; n.s = non significatif.

En résumé, alors que les appositives sont significativement associées à la perception de frontières fortes et que les circonstancielles sont significativement associées à la perception de frontières faibles, les relatives déterminatives ne sont pas associées à la perception de frontières. Les frontières prosodiques ne sont donc pas perçues avec la même intensité dans les trois constructions subordonnées. Une gradation s’observe dans la perception de la force des frontières entre les appositives relatives, les circonstancielles, et les relatives déterminatives. Cette gradation disparaît dans la parole filtrée.

Nous avons ensuite exploré les interactions entre les indices prosodiques en isolation et le jugement du poids des frontières, dans nos trois types syntaxiques de constructions subordonnées. Comme le montre le tableau 5.3, les appositives et les circonstancielles sont assez similaires en termes d’interactions, mais les rela-

tives déterminatives présentent des différences dans l'ordre et le niveau des interactions. Les pauses silencieuses représentent l'indice le plus fort dans la perception des frontières pour les trois types syntaxiques. Cependant, l'effet principal des pauses silencieuses est significatif pour les frontières "fortes" dans le cas des appositives ($\beta = 2$, $SE = .40$, $p < 0.0001$) et des circonstancielles ($\beta = 1.67$, $SE = .39$, $p < 0.0001$), alors que cet effet est seulement significatif pour le degré "frontière" dans le cas des relatives déterminatives ($\beta = 1.04$, $SE = .32$, $p < 0.005$). La perception de la force des frontières prosodiques est donc impactée par le jugement syntaxique en ce qui concerne les pauses silencieuses. Dans les stimuli filtrés, l'effet principal des pauses silencieuses est diminué dans le cas des appositives et des circonstancielles (seulement significatif pour le degré "pas de frontière"; appositives : $\beta = -1.38$, $SE = .48$, $p < 0.005$; circonstancielles : $\beta = -2.53$, $SE = .67$, $p < 0.0001$), alors que cet effet reste similaire dans le cas des relatives déterminatives (significatif pour le degré "frontière"; $\beta = 0.79$, $SE = .40$, $p < 0.05$).

	appositives	circonstancielles	relatives déterminatives
(1)	pause *** "frontière forte"	pause *** "frontière forte"	pause ** "frontière"
(2)	allongement final * "frontière forte"	allongement final * "frontière forte"	rehaussement initial * "frontière"
(3)	n/a	ton descendant * "frontière faible"	allongement final * "pas de frontière"

TABLE 5.3 – Interactions significatives entre les indices prosodiques en isolation et les scores de frontière dans chaque type syntaxique (***) = $p < 0.0001$; ** = $p < 0.005$; * = $p < 0.05$).

Le second effet principal est également différent pour les relatives déterminatives. Dans le cas des appositives et des circonstancielles, l'effet principal de l'allongement syllabique final est significatif pour le jugement "frontière forte" (appositives : $\beta = 0.78$, $SE = .38$, $p < 0.05$; circonstancielles : $\beta = 1.13$, $SE = .42$, $p < 0.05$). En revanche, dans le cas des relatives déterminatives, l'effet principal de l'allongement syllabique final est seulement significatif pour le jugement "pas de frontière" ($\beta = -1.09$, $SE = .39$, $p < 0.01$). Cet effet est donc très inférieur dans le cas des relatives déterminatives. Pour cet indice prosodique également, la perception de la force des frontières prosodiques est impactée par le jugement syntaxique.

Le second effet principal le plus important pour les relatives déterminatives concerne le rehaussement initial. Cet effet s'observe dans une moindre mesure que les résultats sur les autres constructions syntaxiques, puisqu'il est seulement signifi-

catif pour le jugement “frontière” ($\beta = 0.8$, $SE = .37$, $p < 0.05$).

Aucun autre effet principal des indices de frontière pour les jugements n’a été trouvé dans le cas des appositives. Les circonstancielles et les relatives déterminatives montrent une troisième interaction. Dans le cas des circonstancielles, l’effet principal des tons descendants est significatif pour le jugement “frontière faible” ($\beta = 1.5$, $SE = .48$, $p < 0.05$). Un ton descendant est donc moins fiable qu’une pause silencieuse ou qu’un allongement syllabique final dans la perception des frontières, mais est tout de même identifié en tant qu’indice de frontière dans ce type syntaxique.

Dans le cas des relatives déterminatives, l’effet principal de l’allongement syllabique final est significatif pour le jugement “pas de frontière” ($\beta = -1.09$, $SE = .39$, $p < 0.01$) : il y a moins de jugements “pas de frontière” dans les stimuli présentant un allongement syllabique final que dans ceux sans allongement. L’allongement syllabique final ne joue donc pas de rôle dans le jugement du poids des frontières, mais est toutefois utile dans l’identification des frontières.

En résumé, l’exploration des interactions entre les indices prosodiques en isolation et le jugement du poids des frontières montre que certains indices sont utilisés par les auditeurs naïfs pour percevoir des frontières prosodiques dans les trois constructions syntaxiques, mais à une force différente. Ces indices sont les pauses silencieuses et l’allongement syllabique final. D’autres indices ne sont utilisés que pour une construction syntaxique particulière parmi les trois. Les tons descendants se trouvent associés à la perception d’une frontière faible uniquement dans le cas des circonstancielles, et le rehaussement initial se trouve associé à la perception d’une frontière uniquement dans le cas des relatives déterminatives.

Nous avons enfin étudié les corrélations entre les indices prosodiques en isolation dans chaque construction syntaxique. Comme le montre le tableau 5.4, les trois types syntaxiques de subordonnées montrent différentes corrélations significatives entre les indices.

Nous observons au premier abord que dans chaque type syntaxique, le premier et/ou le second indice le plus fortement associé à la perception des frontières en isolation (décrit pour chaque type syntaxique dans le tableau 5.3) est impliqué dans une corrélation. L’intensité de la liaison des indices dans la perception d’une frontière prosodique diffère donc en fonction du type syntaxique.

	appositives	circonstancielle	relatives
(1)	allongement + rehaussement r = 0.68	pause + allongement r = 0.71	pause + rehaussement r = 0.61
(2)	pause + allongement r = 0.6	pause + rehaussement r = 0.61	pause + allongement r = 0.6

TABLE 5.4 – Corrélations significatives entre les indices prosodiques en isolation dans la perception des frontières pour chaque type syntaxique.

Cependant, la plus forte corrélation d'indices dans le cas des appositives n'implique pas les pauses silencieuses. Pour les appositives, la plus forte interdépendance entre les indices prosodiques ne correspond donc pas au poids des indices en isolation dans la perception d'une frontière prosodique. La combinaison de l'allongement syllabique final et du rehaussement intonatif représente la corrélation la plus élevée. De plus, les appositives représentent le seul cas de figure où ces deux indices sont corrélés dans la perception d'une frontière prosodique. La seconde corrélation la plus importante correspond en revanche aux deux indices les plus forts isolément dans la perception d'une frontière prosodique.

En ce qui concerne les circonstancielle au contraire, la corrélation la plus forte (impliquant les pauses silencieuses et l'allongement) correspond bien à la hiérarchie de poids des indices en isolation dans la perception d'une frontière prosodique vue dans le tableau 5.3. En revanche, la seconde corrélation implique une nouvelle fois la pause, combinée non pas au troisième indice le plus important isolément dans la perception d'une frontière (i.e. les tons descendants), mais au rehaussement de la courbe intonative. Si le rehaussement de la courbe intonative n'est pas un indice utilisé dans la perception des frontières dans le cas des circonstancielle lorsqu'il est en isolation, cet indice devient impliqué dans la perception des frontières lorsqu'il est combiné à une pause. Les tons descendants, en revanche, ne montrent pas de liaison particulière avec un autre indice. Dans les circonstancielle, le degré de corrélation entre les indices ne correspond que partiellement à la hiérarchie d'importance des indices en isolation dans la perception d'une frontière prosodique.

Dans le cas des relatives déterminatives, la corrélation d'indices la plus forte (impliquant les pauses silencieuses et le rehaussement) correspond également à la hiérarchie de poids des indices en isolation dans la perception d'une frontière prosodique vu dans le tableau 5.3. La seconde corrélation implique une nouvelle fois la pause, combinée au troisième indice le plus important isolément dans la percep-

tion d'une frontière prosodique, i.e. l'allongement syllabique final. Pour les relatives déterminatives, la force des corrélations entre les différents indices prosodiques correspond donc strictement à la hiérarchie de poids des indices en isolation dans la perception d'une frontière.

En résumé, les corrélations entre les indices de rupture prosodique dans les appositives ne correspondent que partiellement au poids des indices prosodiques en isolation dans le test de perception : la première corrélation d'indices mesurée dans ce test ne correspond pas. La correspondance entre perception et poids des indices en isolation est également partielle pour les circonstancielles, cette fois en raison de la seconde corrélation d'indices la plus forte mesurée dans notre test. En ce qui concerne les relatives déterminatives, la correspondance entre perception et poids des indices en isolation est stricte.

5.6 Discussion

En ce qui concerne les interactions potentielles entre les trois types de constructions syntaxiques (facteur fixe = Type; valeurs = appositive, circonstancielle, relative) et le jugement du poids des frontières (facteur fixe = Scores; valeurs = frontière forte (strong boundary), frontière (boundary), frontière faible (weak boundary), incertain (uncertain), pas de frontière (no boundary)), nous avons vu avec le tableau 5.2 qu'alors que les appositives sont significativement associées à la perception de frontières fortes et que les circonstancielles sont significativement associées à la perception de frontières faibles, les relatives déterminatives ne sont pas associées à la perception de frontières. Ces résultats ne peuvent être expliqués en termes de longueur, de position séquentielle, de statut discursif, ou d'indices prosodiques des constituants, puisque ces paramètres ont été contrôlés. Cependant, la complexité et le poids des constituants restent des facteurs potentiels. Bien que la prosodie et la syntaxe ne soient pas nécessairement en correspondance, nos résultats sont en accord avec les résultats macro-syntaxiques, puisque les appositives expriment une rupture macro-syntaxique par rapport à la séquence hôte (cf. Chapitre 3 p. 117).

Le fait que les circonstancielles sont associées à la perception de frontières faibles trouve aussi un écho sur le plan macro-syntaxique, puisque nous avons vu dans le Chapitre 3 p. 104 que les circonstancielles montrent une grande flexibilité macro-syntaxique. Elles montrent en majorité une autonomie macro-syntaxique partielle, où l'autonomie syntaxique ne correspond pas nécessairement à l'autonomie pragma-

tique.

Les relatives déterminatives ne sont pas associées à la perception de frontières, alors que leur équivalent filtré est associé à la perception de frontières fortes. Etant donné le fait que les sujets qui annotent des frontières prosodiques dans leur langue maternelle perçoivent généralement plus de frontières que les sujets qui annotent de la parole délexicalisée (Mettouchi et al., 2007), ce résultat suggère que les relatives déterminatives sont négativement corrélées à la perception des frontières. En termes de macro-syntaxe, nous avons vu dans le Chapitre 3 p. 89 que les relatives déterminatives sont majoritairement pleinement intégrées à leur co-texte gauche, régies syntaxiquement par la proposition du co-texte gauche.

En résumé, les frontières prosodiques ne sont pas perçues avec la même intensité dans les trois constructions subordonnées. Une gradation s’observe dans la perception de la force des frontières entre les appositives relatives, les circonstancielles, et les relatives déterminatives. Cette gradation disparaît dans la parole filtrée, et suit le même schéma que la gradation de l’autonomie macro-syntaxique observée lors de la production des trois constructions subordonnées. Pour ce résultat, une correspondance s’établit entre les résultats en perception et ceux en production.

Nous avons ensuite exploré les interactions entre les indices prosodiques en isolation et le jugement du poids des frontières, dans nos trois types syntaxiques de constructions subordonnées. Comme l’a montré le tableau 5.3, les pauses silencieuses représentent l’indice le plus fort dans la perception des frontières pour les trois types syntaxiques. Cependant, l’effet principal des pauses silencieuses est significatif à un moindre degré dans le cas des relatives déterminatives. Ce résultat est en accord avec le précédent : les frontières sont moins perçues dans les relatives déterminatives, même avec l’indice prosodique le plus fort. La perception de la force des frontières prosodiques est donc impactée par le jugement syntaxique en ce qui concerne les pauses silencieuses.

Le second effet principal est également différent pour les relatives déterminatives. Dans le cas des appositives et des circonstancielles, l’effet principal de l’allongement syllabique final est significatif pour le jugement “frontière forte”. En revanche, dans le cas des relatives déterminatives, l’effet principal de l’allongement syllabique final est seulement significatif pour le jugement “pas de frontière”. Cet effet est donc très inférieur dans le cas des relatives déterminatives. Pour cet indice prosodique également, la perception de la force des frontières prosodiques est impactée par le

jugement syntaxique.

Nous avons également vu que le second effet principal le plus important pour les relatives déterminatives concerne le rehaussement initial. Cet effet s’observe dans une moindre mesure que les résultats sur les autres constructions syntaxiques, puisqu’il est seulement significatif pour le jugement “frontière”. En conversation spontanée, nous avons vu dans le Chapitre 4 p. 145 que le rehaussement initial est peu commun entre L et SC dans les séquences contenant une relative déterminative. Alors qu’elles sont majoritairement produites sous le même groupe intonatif que L, les autres sont typiquement précédées d’un contour de continuation. La fin de L et le début de SC sont donc habituellement produits à la même hauteur intonative. L’interaction concernant le rehaussement intonatif n’est donc pas basée sur un effet d’apprentissage en production qui impliquerait une correspondance stricte entre un type syntaxique et un trait prosodique.

Aucun autre effet principal des indices de frontière pour les jugements n’a été trouvé dans le cas des appositives. Cela signifie que les indices de frontière utilisés dans les appositives jouent uniquement un rôle dans la perception d’une frontière “forte”. L’étude de production ainsi que le travail de Auran et Loock (2011) ont montré que les pauses silencieuses et l’allongement syllabique final sont des traits typiques des séquences contenant une appositive, contrairement à celles contenant une circonstancielle et celles contenant une relative déterminative. Cependant, bien que les appositives soient également très couramment précédées d’un ton descendant sur la syllabe nucléaire de L et qu’elles soient souvent produites avec un changement de hauteur intonative ou avec un rehaussement initial (ibid.), les tons descendants ou montants ne jouent aucun rôle significatif dans la perception des frontières dans la présente étude. Le rehaussement intonatif en tant qu’indice en isolation n’est pas utilisé non plus.

Les circonstancielles et les relatives déterminatives montrent une troisième interaction. Le fait que ces deux types syntaxiques présentent un nombre important d’interactions malgré leur plus faible association aux frontières prosodiques en général prouve la sensibilité des auditeurs naïfs aux indices prosodiques (Cole et al., 2014). Dans le cas des circonstancielles, l’effet principal des tons descendants est significatif pour le jugement “frontière faible”. Un ton descendant est donc moins fiable qu’une pause silencieuse ou qu’un allongement syllabique final dans la perception des frontières, mais est tout de même identifié en tant qu’indice de frontière dans ce type syntaxique. Dans la parole conversationnelle, les circonstancielles sont com-

munément précédées d'un ton descendant sur L, souvent lié à la gestion des cadres interprétatifs et aux fonctions de discours (Chafe, 1984). Nous avons vu dans le Chapitre 3 que les circonstancielles peuvent être analysées comme des moyens de grouper plusieurs propositions, liées par le fait qu'elles doivent être interprétées par le prisme d'un critère véhiculé dans une circonstancielle (Brown et Yule, 1983). Cela implique que les co-locuteurs gardent à l'esprit l'introducteur de cadre pour le traitement du segment hôte et au-delà, jusqu'à ce qu'un indicateur vienne annoncer la fin de sa portée. Un ton descendant sur le segment précédent peut servir d'indicateur.

Dans le cas des relatives déterminatives, l'effet principal de l'allongement syllabique final est significatif pour le jugement "pas de frontière" : il y a moins de jugements "pas de frontière" dans les stimuli présentant un allongement syllabique final que dans ceux sans allongement. L'allongement syllabique final ne joue donc pas de rôle dans le jugement du poids des frontières, mais est toutefois utile dans l'identification des frontières. Cependant, nous avons vu dans le Chapitre 4 p. 143 qu'en production, l'allongement syllabique final est rarement produit sur L dans les séquences contenant une relative déterminative. L'interaction concernant l'allongement syllabique n'est donc pas non plus basée sur un effet d'apprentissage en production.

En résumé, l'exploration des interactions entre les indices prosodiques en isolation et le jugement du poids des frontières montre que certains indices sont utilisés par les auditeurs naïfs pour percevoir des frontières prosodiques dans les trois constructions syntaxiques, mais à une force différente. Ces indices sont les pauses silencieuses et l'allongement syllabique final. En montrant que la pause silencieuse en tant qu'indice influençant la perception de frontières prosodiques correspond aux indices de frontière mis à jour sur des données de production, nos résultats sont en accord avec la récente étude de Roy et al. (2017). Notre étude établit également l'allongement syllabique final comme un tel indice. D'autres indices ne sont utilisés que pour une construction syntaxique particulière parmi les trois. Les tons descendants se trouvent associés à la perception d'une frontière faible uniquement dans le cas des circonstancielles, et le rehaussement initial se trouve associé à la perception d'une frontière uniquement dans le cas des relatives déterminatives. Si les tons descendants trouvent une correspondance avec les circonstancielles en production, le rehaussement initial n'est pas utilisé pour la production des relatives déterminatives.

Nous avons enfin étudié les corrélations entre les indices prosodiques en isolation dans chaque construction syntaxique. Comme l'a montré le tableau 5.4, les trois

types syntaxiques de subordonnées montrent différentes corrélations significatives entre les indices.

Nous avons observé au premier abord que dans chaque type syntaxique, le premier et/ou le second indice le plus fortement associé à la perception des frontières en isolation (décrit pour chaque type syntaxique dans le tableau 5.3) est impliqué dans une corrélation. L'intensité de la liaison des indices dans la perception d'une frontière prosodique diffère donc en fonction du type syntaxique.

Cependant, la plus forte corrélation d'indices dans le cas des appositives n'implique pas les pauses silencieuses. Ce résultat est surprenant étant donné le grand poids des pauses silencieuses dans la perception des frontières prosodiques, à la fois dans notre test en tant qu'indice en isolation et tel qu'il est décrit dans les études antérieures (Wightman et al., 1992; Carlson et Swerts, 2003; Yoon et al., 2007; Wagner et Watson, 2010). Pour les appositives, la plus forte interdépendance entre les indices prosodiques ne correspond donc pas au poids des indices en isolation dans la perception d'une frontière prosodique. La combinaison de l'allongement syllabique final et du rehaussement intonatif représente la corrélation la plus élevée. De plus, les appositives représentent le seul cas de figure où ces deux indices sont corrélés dans la perception d'une frontière prosodique. La seconde corrélation la plus importante correspond en revanche aux deux indices les plus forts isolément dans la perception d'une frontière prosodique. Par ailleurs, en ce qui concerne les données de production, les deux corrélations observées en perception représentent des configurations assez fréquentes en production. Nous avons non seulement vu dans le Chapitre 4 que les appositives sont très fréquemment produites avec plusieurs indices de frontière en co-occurrence, mais que ces indices impliquent souvent la combinaison d'une disjonction rythmique à une disjonction d'ordre mélodique. Nous en concluons que pour les appositives, les corrélations d'indices dans la perception d'une frontière prosodique sont assez proches des corrélations d'indices de frontière en production, hormis l'absence de la pause silencieuse dans la plus forte corrélation. Le fait que la pause silencieuse apparaisse presque systématiquement dans les combinaisons d'indices prosodiques en production prohibe toute correspondance stricte.

En ce qui concerne les circonstancielles au contraire, il a été vu que la corrélation la plus forte (impliquant les pauses silencieuses et l'allongement) correspond bien à la hiérarchie de poids des indices en isolation dans la perception d'une frontière prosodique (visible dans le tableau 5.3). En revanche, la seconde corrélation n'implique pas le troisième indice le plus important isolément dans la perception d'une

frontière (i.e. les tons descendants), mais le rehaussement de la courbe intonative. Le rehaussement de la courbe intonative n'est pas un indice utilisé dans la perception des frontières dans le cas des circonstanciellelles lorsqu'il est en isolation, mais cet indice devient impliqué dans la perception des frontières lorsqu'il est combiné à une pause. Les tons descendants, en revanche, ne montrent pas de liaison particulière avec un autre indice. Nous en avons conclu que dans les circonstanciellelles, le degré de corrélation entre les indices ne correspond que partiellement à la hiérarchie d'importance des indices en isolation dans la perception d'une frontière prosodique. En ce qui concerne les données de production, nous avons vu dans le Chapitre 4 que les circonstanciellelles sont rarement produites avec plus d'un indice de rupture en combinaison. La seule combinaison d'indices visible au sein d'une seule séquence implique une plus grande modulation mélodique sur la circonstanciellelle que sur le co-texte, et une pause remplie sur R. Or, ces deux indices de rupture ne sont pas mesurables dans un test impliquant le type de stimuli court choisi. S'il est difficile de parler de correspondance entre les corrélations d'indices en production et en perception, nous remarquons toutefois que les indices de rupture et leurs (rares) combinaisons incluent une discontinuité d'ordre mélodique aussi bien en production qu'en perception.

Dans le cas des relatives déterminatives, la corrélation d'indices la plus forte (impliquant les pauses silencieuses et le rehaussement) correspond à la hiérarchie de poids des indices en isolation dans la perception d'une frontière prosodique vue dans le tableau 5.3. La seconde corrélation implique une nouvelle fois la pause, combinée au troisième indice le plus important isolément dans la perception d'une frontière prosodique, i.e. l'allongement syllabique final. Pour les relatives déterminatives, la force des corrélations entre les différents indices prosodiques correspond donc strictement à la hiérarchie de poids des indices en isolation dans la perception d'une frontière. De la même manière que pour les circonstanciellelles, nous avons vu dans le Chapitre 4 qu'en production, les relatives déterminatives sont très rarement produites avec plus d'un indice de rupture à la fois. Nous ne parlerons donc pas de correspondance entre production et perception. De plus, les rares corrélations d'indices prosodiques observées en production concernent uniquement des indices de nature rythmique. Les corrélations mesurées en perception dans ce chapitre nous donnent néanmoins des informations supplémentaires sur des relatives déterminatives modélisées (i.e. artificiellement produites avec une rupture importante).

Si nous nous intéressons à la nature des indices impliqués dans des corrélations,

deux indices de nature rythmique sont corrélés (pause + allongement syllabique final) dans 50% des cas de figure. Dans l'autre moitié des cas de figure, un indice de nature rythmique (pause ou allongement) est lié à une discontinuité mélodique (rehaussement de la ligne intonative). La présence de la pause dans 83.3 % des corrélations confirme à la fois la forte liaison des deux autres indices à la pause et le poids de la pause silencieuse en tant qu'indice (isolé et combiné) dans la perception d'une frontière prosodique.

Les corrélations mettent également en lumière le rôle du rehaussement de la courbe intonative. Si le rehaussement de la courbe intonative n'est pas un indice utilisé dans la perception des frontières dans le cas des appositives et des circonstancielles lorsqu'il est en isolation, cet indice montre une forte liaison à l'allongement syllabique final dans le cas des appositives, et à la pause silencieuse dans le reste des cas. En définissant ces trois traits prosodiques comme les plus pertinents dans la perception d'une frontière prosodique, ce résultat est proche de ceux des travaux de Cohen et 't Hart (1967), de 't Hart et Collier (1975) et de Hermes (2006), bien que l'approche IPO auxquels ils appartiennent ne porte pas sur des données issues de la conversation spontanée. Nos résultats confirment également les observations de Collier et al. (1993) et de De Pijper et Sanderman (1994) sur la forte interdépendance entre le rehaussement de la ligne de déclinaison, les pauses silencieuses, et l'allongement syllabique final. Nos résultats précisent en revanche que ces trois indices ne sont pas liés de la même manière en fonction de la syntaxe.

En résumé, les corrélations entre les indices de rupture prosodique dans les appositives ne correspondent que partiellement à la fois au poids des indices prosodiques en isolation dans le test de perception, et aux corrélations d'indices réalisées par les locuteurs en production : la première corrélation d'indices mesurée dans ce test ne correspond ni à l'un, ni à l'autre. La correspondance entre perception, poids des indices en isolation, et production est également partielle pour les circonstancielles, cette fois en raison de la seconde corrélation d'indices la plus forte mesurée dans notre test. En ce qui concerne les relatives déterminatives, la correspondance entre perception et poids des indices en isolation est stricte. En revanche, la correspondance entre production et perception ne peut être établie si nous gardons à l'esprit que les relatives déterminatives sont rarement produites avec plus d'un indice. Indépendamment de la syntaxe, la pause est l'indice prosodique le plus présent dans toutes les corrélations observées dans la perception des frontières prosodiques. Peu repéré par les participants en tant qu'indice de frontière en isolation, le rehaussement intonatif joue également un rôle non négligeable dans la perception des

frontières lorsqu'il est combiné à autre indice. L'étude de Roy et al. (2017) souligne que les corrélations entre les indices de frontière prosodique dans les tests de perception en général font l'objet d'une grande variabilité interpersonnelle. En dépit de cette variabilité, nous montrons que les schémas de corrélation sont sensibles au format syntaxique du stimulus.

5.7 Conclusion

Nous avons présenté dans ce chapitre une expérience visant à étudier la perception des frontières prosodiques par des auditeurs naïfs, dans des extraits de parole spontanée contenant une construction subordonnée syntaxiquement. Nous avons confirmé les hypothèses suivantes : (1) la perception des frontières prosodiques varie selon les différents types syntaxiques de subordination; (2) les différents types syntaxiques ne montrent pas les mêmes ordres et niveaux d'interaction avec les mêmes indices prosodiques de frontière. Notre test confirme également l'importance de la pause silencieuse en tant qu'indice de frontière, à la fois en isolation et dans les combinaisons.

Nous avons vu que les frontières prosodiques ne sont pas perçues avec la même intensité en fonction des types syntaxiques : les appositives sont significativement associées à la perception de frontières fortes, les circonstancielles à la perception de frontières faibles, et les relatives déterminatives ne sont pas associées à la perception de frontières. Cette gradation dans la perception des frontières correspond à la gradation de l'autonomie macro-syntaxique des différents types syntaxiques de constructions subordonnées traitée dans le Chapitre 3, et à la gradation de l'autonomie prosodique des différents types syntaxiques vue dans le Chapitre 4. En ce qui concerne les interactions des types syntaxiques avec les indices prosodiques de frontière en isolation, nous avons observé que les pauses silencieuses et l'allongement syllabique final sont utilisés par les auditeurs naïfs pour percevoir des frontières dans les trois constructions prosodiques, mais à un degré de force différent. En revanche, les circonstancielles et les relatives déterminatives montrent des associations avec des indices spécifiques : les tons descendants sont associés à la perception de frontières faibles uniquement dans le cas des circonstancielles, alors que le rehaussement de la courbe intonative est associé à l'identification d'une frontière uniquement dans le cas des relatives déterminatives. Les correspondances entre la pertinence de ces indices en perception et leur fréquence d'occurrence en production sont assez irrégulières. La pause silencieuse et l'allongement syllabique final ne correspondent qu'à la descrip-

tion de la production des appositives. Si une correspondance s'établit en production entre les circonstancielles et la présence de tons descendants sur L, les relatives déterminatives sont très rarement précédées d'un rehaussement intonatif.

En ce qui concerne les corrélations entre les indices, l'intensité de la liaison entre les indices dans la perception d'une frontière prosodique diffère en fonction du type syntaxique. De plus, les appositives et les circonstancielles montrent des degrés de corrélations différents de la hiérarchie de poids des différents indices prosodiques en isolation dans la perception d'une frontière. Seules les relatives déterminatives montrent une correspondance stricte entre le poids des indices en isolation et leurs degrés de corrélation. En ce qui concerne les correspondances entre les corrélations en production et perception, les appositives montrent des similitudes non négligeables entre production et en perception, bien que l'absence de la pause représente une différence fondamentale. Les circonstancielles et les relatives déterminatives ne sont habituellement pas produites avec plus d'un indice de frontière à la fois. La pertinence des résultats concernant les corrélations d'indices de frontière en perception se trouve donc davantage sur le plan de la modélisation. Néanmoins, le fait que ces corrélations soient différentes entre les circonstancielles et les relatives déterminatives étoffe la confirmation de notre seconde hypothèse (les différents types syntaxiques ne montrent pas les mêmes ordres et niveaux d'interaction avec les mêmes indices prosodiques de frontière) d'une autre manière.

En mettant à jour des différences dans la perception des frontières prosodiques en fonction du type syntaxique de subordination, ce test contribue à l'exploration de l'interface entre syntaxe et prosodie. Si notre test ne reflète pas les prises de décision spontanées et en temps réel des co-locuteurs en situation de conversation en raison de l'isolation des indices prosodiques, il permet néanmoins d'exposer des auditeurs naïfs à un stimulus ressemblant à des bribes de conversation spontanée.

L'utilisation de modèles linéaires mixtes généralisés permet d'étudier l'influence de chaque variable sur la perception des frontières avec un traitement statistique robuste malgré le petit nombre de participants et de stimuli (accentuant notamment la différence entre les stimuli produits par des femmes et ceux produits par des hommes), ainsi que la potentielle hétérogénéité du groupe d'auditeurs. En revanche, si les modèles linéaires mixtes généralisés amènent dans la modélisation des contraintes subtiles et purement distributionnelles, certains paramètres situationnels relatifs aux conditions expérimentales ne sont pas pris en compte. La fatigue attentionnelle représente l'un de ces paramètres, étant donnée la relative longueur

de notre test (40 minutes) pour des participants naïfs. De même, la variation individuelle joue un grand rôle dans la perception des frontières prosodiques, surtout au niveau de l'assignation du poids des différents indices dans la perception de ces frontières (Roy et al., 2017). Si ces paramètres n'invalident ni nos observations ni nos conclusions, ils sont à prendre en considération dès lors que nous évoquons la pertinence de nos résultats.

Chapitre 6

Frontières gestuelles

6.1 Mesure des frontières gestuelles

Nous évaluons dans ce chapitre l'autonomie gestuelle des subordonnées sur un mode d'organisation linéaire (intégration vs. démarcation). Nous déterminons si ces constructions créent globalement une rupture ou bien si elles sont préférentiellement intégrées au co-texte gauche et/ou droit.

Après avoir identifié et mesuré les indices de démarcation gestuelle les plus pertinents tels que nous les avons formulés dans les hypothèses du Chapitre 1 p. 58, nous sommes en mesure de placer les trois types syntaxiques sur un continuum, de l'intégration à l'autonomie gestuelle, ainsi que d'identifier les indices les plus mobilisés dans le marquage d'une rupture. Les relatives déterminatives sont très nettement les plus intégrées à leur entourage. Elles ne présentent qu'une seule marque significative de rupture. Cette marque est de nature rythmique et concerne les gestes manuels. Une tendance non-significative est également relevée pour les relatives déterminatives, mais concerne trop peu d'occurrences pour constituer un réel indice. Les circonstancielles comptent quant à elles deux indices de rupture significatifs, accompagnés de trois tendances n'atteignant pas la significativité. Comme les relatives déterminatives, la rupture s'établit par le biais des gestes manuels. Enfin, les appositives relatives sont autonomisées par quatre indices, mêlant des paramètres venant de plusieurs articulateurs (mains, regard, sourcils). Le détail des caractéristiques de chaque type syntaxique de subordonnée est donné de la catégorie la moins autonome à celle la plus en rupture.

6.2 Résultats

6.2.1 Relatives déterminatives

Les relatives déterminatives ne comportent qu'un seul indice de rupture gestuelle. Ce type syntaxique exprime donc relativement peu de démarcation du point de vue visuel. De nature rythmique, cette marque concerne les gestes manuels. Les relatives déterminatives comptent le plus grand nombre de battements isolés (i.e. les battements manuels ne faisant pas partie d'une série de gestes répétés dans une séquence discursive), par rapport aux autres types syntaxiques (circonstancielles : $F(89, 96) = 2.15, p < 0.0002$; appositives relatives : $F(89, 75) = 2.33, p < 0.0002$), mais aussi au sein des séquences dans lesquelles elles sont produites (L : $F(89, 69) = 1.74, p < 0.005$; R : $p > 0.5$). 37.7 % des relatives déterminatives de notre corpus sont produites en co-occurrence avec au moins un battement manuel. La séquence (1), associée à la figure 6.1¹ où les images (a), (b), (c), et (d) correspondent à différents moments de sa production, illustre cette tendance :

- | | | | |
|-----|----------|----|---|
| (1) | Michelle | L | [(a) but i put it on the bit |
| | | SC | where hem (cough) (h) they were uh#] |
| | | | [(b) in the] [(c) garden] |
| | | R | [(d) and they were talking] |

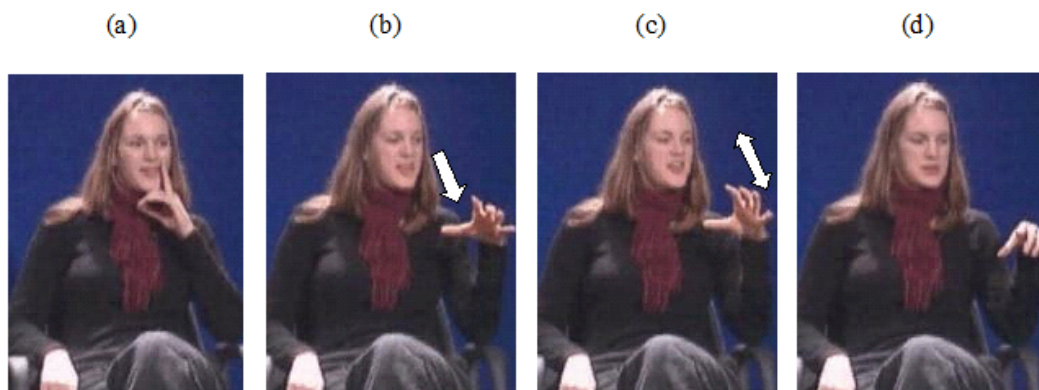


FIGURE 6.1 – Deux battements successifs dans l'exemple (1), réalisés en co-occurrence avec la relative déterminative.

Michelle vient de mentionner un film que Zoe a déjà vu, récemment diffusé à la télévision. Ayant pris le film en cours de route, Michelle décrit plus précisément le moment où elle a commencé son visionnage. Sur le plan prosodique, le début de SC

1. L'unité gestuelle est représentée par les crochets.

est directement intégré au groupe intonatif de L par une technique de *rush-through* (cf. Chapitre 4 p. 142 pour une définition du *rush-through*). SC est en revanche immédiatement caractérisé par des disfluences vocales liées à une recherche, scindé en deux unités par une pause silencieuse relativement importante. Sur le plan gestuel, le second groupe intonatif de SC se démarque du reste de la séquence par ses deux battements successifs (b) et (c). En produisant ces deux battements en co-occurrence avec “in the garden” Michelle indique pragmatiquement le contenu informationnel le plus pertinent de la séquence (Cassell et McNeill, 1990), encodé comme le résultat d’une importante recherche liée à la planification des unités et/ou des concepts, et/ou liée au lexique. Les battements ont pour effet de marquer qu’il existe une relation entre les termes récupérés et l’hésitation qui les précède (McNeill, 1992), même si la recherche porte sur un autre plan que le plan lexical. La configuration main ouverte paume vers le bas des battements prend une valeur déictique abstraite alors que Michelle s’est efforcée de localiser temporellement une scène exacte, et choisit d’isoler un lieu. L’information locative “in the garden” permet en effet de mieux répondre à la question “which moment?” que le contenu de L, dont l’incomplétude ne permet pas une sélection précise, et que le contenu de R, décrivant une action non-spécifique et récurrente, qui caractérise l’intégralité du film. Le second groupe intonatif de SC est également marqué par une F0 très haute sur le plan intonatif, créant un grand rehaussement entre les deux groupes intonatifs. Michelle rétracte partiellement sa main ouverte paume vers le bas sur R, en baissant le poignet pour trouver une nouvelle position de repos (d) qu’elle gardera pendant les prochains tours de parole. Aucun autre battement n’est produit dans la séquence. Si la F0 moyenne de R est considérablement abaissée par rapport au second groupe intonatif de SC, elle est également légèrement plus basse que sur L. Cette baisse d’effort articulatoire général creuse le contraste avec les moyens de marquage de SC. La densité modale (Norris, 2004) est nettement plus importante sur le second groupe intonatif de SC que sur R.

Notons que les battements manuels ne sont pas systématiquement liés à la récupération d’une recherche de la part du locuteur : seulement 10.9 % des relatives déterminatives produites avec un battement sont caractérisées par une recherche, quelle qu’elle soit. Nous avons également vu dans le Chapitre 4 que les relatives déterminatives ne présentent pas la part la plus importante de marques vocales d’hésitation. Très fluide d’un point de vue vocal, la description (2) suivante est caractérisée par un battement visible dans la figure 6.2.

- (2) Tom have you seen the #
 there’s some programmes on that
 [(a) on BBC #]

L with [(b)this guy]
 SC **who does [(c) it#]**
 R he does like a [(d) Top] Gear programme
 but he also does this kind of #

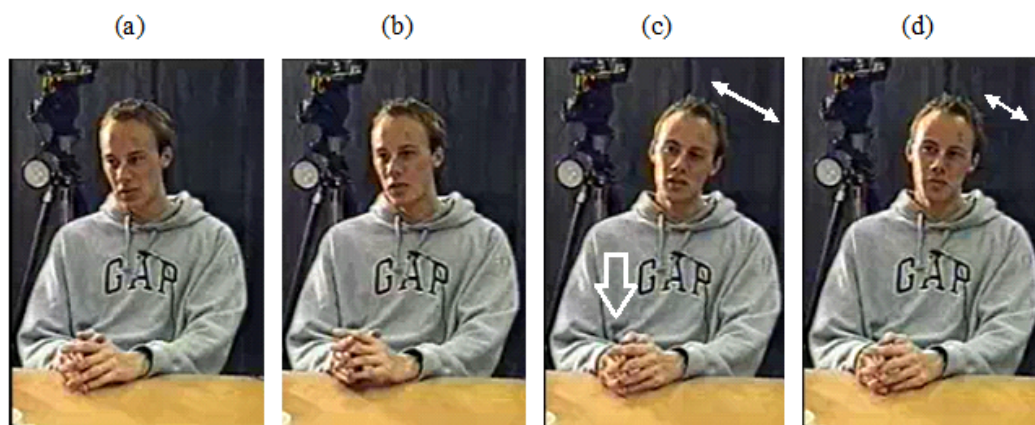


FIGURE 6.2 – Battement manuel sur la relative déterminative et inclinaisons de tête dans l'exemple (2).

Tim et Tom viennent de comparer un documentaire de Michael Moore au style des émissions télévisées de Channel Five. Tom commence la description d'une autre émission d'investigation (a). D'un point de vue informationnel, la série de groupes intonatifs jusqu'à la fin de SC repousse la frontière droite de "there's some programmes on that", en y accrochant un nouvel élément rhématique de groupe intonatif en groupe intonatif. Tom élabore de manière incrémentale sur l'émission qu'il essaie de faire identifier par Tim. Ne connaissant pas le nom du présentateur, il lève les yeux vers Tim sur L (b) en introduisant une nouvelle entité référentielle (i.e. "this guy") à l'aide d'un syntagme prépositionnel et d'un présentatif ("this"). Cette mise en focus constitue un appel au co-locuteur. Face à l'absence de réaction de celui-ci, Tom prédique à propos du même référent sur SC à l'aide de la relative déterminative. Cette séquence ne montre pas de disflue d'un point de vue vocal. Tom ne montre pas de difficulté à encoder les unités de discours, et ne marque pas la séquence comme une recherche : le fait qu'il ne connaisse pas le nom du présentateur est clairement marqué comme tel. Il fait cependant appel aux connaissances partagées entre les participants afin d'évaluer la pertinence de sa contribution à la séquence, portant sur le journalisme d'investigation. Sur le plan prosodique, L et SC forment deux groupes intonatifs distincts. Cette séparation est remarquable dans le sens où L et SC sont très courts, et dans le sens où nous avons vu dans le Chapitre 4 que les relatives déterminatives sont le plus souvent intégrées au groupe intonatif de L. Les

items “guy” et “does” sont également marqués d’un accent emphatique montant-descendant. Si une réplique du même mouvement intonatif s’observe, celui de SC est légèrement moins haut que celui de L. D’un point de vue verbal, le focus informationnel reste sur l’élément thématique “this guy” repris elliptiquement avec le pronom relatif “who”, puisque l’élément rhématique “it” est déjà connu et que l’item verbal “does” reste très générique en n’apportant pas de précision. D’un point de vue gestuel en revanche, seulement le dernier élément de SC, “it”, est accompagné d’un battement (c). Tom penche également la tête vers la gauche, en maintenant le regard vers Tim. Toujours sans réaction du co-locuteur, Tom prédique à propos du même élément thématique sur R, encodé cette fois-ci par le pronom “he”. Le décalage entre les entités lexicales de surface faisant référence au présentateur et les connaissances partagées entre les deux participants est encore d’actualité : “he” réfère à une entité qui reste à identifier par Tim. Tom fait une nouvelle fois appel aux connaissances partagées entre les deux participants en liant “he” à “a Top Gear programme”, qu’il marque d’une autre inclinaison de tête (d), et d’un léger rehaussement de F0. Tom encode le même référent thématique (i.e. le présentateur de l’émission) sous plusieurs formes lexicales de surface différentes, mais reste sans réponse du co-locuteur. SC constitue l’appel le plus marqué avec une inclinaison de la tête et un battement manuel, que Tom fait suivre d’une pause silencieuse. Cette pause constitue une *TRP*², le regard de Tom maintenu sur le co-locuteur. Cette *TRP* restera non prise en compte par ce dernier.

Le fait que les battements manuels en grand nombre soient une spécificité des relatives déterminatives est compatible avec les résultats de la littérature, qui aborde surtout les battements manuels pour leur fonction de désambiguïsation des structures syntaxiques complexes (Holle et al., 2012), et pour leur rôle emphatique dans le discours (Ferré, 2014; Dimitrova et al., 2016). Puisque ces structures syntaxiques sont significativement plus longues que leur co-texte (cf. Chapitre 4 p. 140) et relativement complexes, il est tout à fait vraisemblable que les locuteurs mobilisent les battements pour guider l’interprétation des co-locuteurs, en démarquant des items lexicaux/verbaux spécifiques (Fromont et al., 2010; Biau et Soto-Faraco, 2013; Wang et Chu, 2013; Dimitrova et al., 2016). La sélection gestuelle de certains items de la modalité verbale permet en effet d’activer des chemins d’accès cognitifs chez les co-locuteurs (Hubbard et al., 2009; Dimitrova et al., 2016), et facilite la recherche d’informations pertinentes de ce dernier dans le discours du locuteur (Krahmer et Swerts, 2007; Kelly et al., 2008).

2. *Transition Relevance Place*, point de complétude sémantique, syntaxique, et intonative du discours, cf. Chapitre 4 p. 142 pour une définition complète des *TRPs*.

En anglais, les propriétés temporelles jouent un rôle important en plus des propriétés tonales dans le groupement des unités prosodiques en phrases, ainsi que dans leur proéminence (Breen et al., 2010). La recherche sur le lien entre les gestes (mains, tête, regard) et les frontières prosodiques soutient l'idée selon laquelle la parole et les gestes d'autres articulateurs participent ensemble à la formation d'une frontière prosodique (Barkhuysen et al., 2008, Ishi et al., 2008, Krivokapic, 2017). L'une des manifestations centrales de la structure prosodique dans la parole est l'allongement temporel des gestes phonatoires aux frontières des unités. Si les effets de la structure prosodique s'étendent au-delà du conduit vocal, les gestes produits avec d'autres articulateurs devraient montrer les mêmes propriétés temporelles que ceux du conduit vocal. Or, seulement 50 % de ces battements manuels sont produits en co-occurrence avec une frontière prosodique (i.e. la syllabe finale du groupe intonatif de L et/ou la syllabe finale du groupe intonatif de SC) dans les séquences contenant une relative déterminative. La durée de ces battements aux frontières des groupes intonatifs n'est pas significativement différente de celle des battements réalisés ailleurs à l'intérieur des groupes intonatifs. La durée des battements en co-occurrence avec la syllabe finale de L n'est pas non plus significativement différente de celle des battements en co-occurrence avec la syllabe finale de SC. Nous concluons que les battements qui caractérisent les relatives déterminatives n'ont pas spécifiquement pour fonction de marquer une frontière prosodique, que cette frontière soit avec le co-texte gauche ou le co-texte droit.

Un autre phénomène mérite d'être mentionné ici, bien qu'il ne constitue pas un indice de rupture en lui-même, concernant les propriétés temporelles d'un type de geste dont l'emploi n'est pas caractéristique des relatives déterminatives. Nous venons de constater qu'aucune différence significative ne permet d'aller dans le sens de la construction d'une frontière prosodique pour les battements manuels en co-occurrence avec les relatives déterminatives, mais nous avons vu que ces battements ne sont qu'en partie produits aux frontières des groupes intonatifs. En revanche, 56.6 % des battements de tête réalisés dans les séquences contenant une relative déterminative sont produits en co-occurrence avec une frontière prosodique. Si les relatives déterminatives ne présentent pas la plus grande proportion de battements de tête et que leur emploi n'est pas une spécificité de ce type syntaxique, nous observons tout de même un allongement significatif de la durée des battements de tête produits en co-occurrence avec la syllabe finale de SC, par rapport à la durée de ceux produits sur la syllabe finale de L ($F(6, 5) = 6.13, p < 0.05$). Ce résultat va dans le sens des résultats prosodiques décrits dans le Chapitre 4 à propos des

relatives déterminatives, où l’allongement syllabique caractérise la syllabe finale de SC, mais pas la syllabe finale de L. Bien que les battements de tête ne soient pas une spécificité des relatives déterminatives, nous observons tout de même leur participation au marquage d’une frontière prosodique.

Hormis la présence des battements manuels dans les séquences et l’allongement significatif des battements de tête en co-occurrence avec la syllabe finale de SC, la forte intégration gestuelle des relatives déterminatives est visible à travers plusieurs paramètres. Les relatives déterminatives sont caractérisées par davantage de marqueurs signalant leur intégration que de marqueurs signalant leur démarcation.

Les relatives déterminatives sont d’abord accompagnées de gestes manuels faisant le plus souvent partie d’un *catchment*³. 25 % des relatives déterminatives de notre corpus sont produites en co-occurrence avec un ou plusieurs gestes manuels inscrits dans un *catchment*. Cela représente significativement plus que les circonstancielles ($F(54, 54) = 1.71, p < 0.05$) et que les appositives relatives ($F(54, 54) = 3.68, p < 0.0001$). Ce résultat est extrêmement intéressant dans la mesure où les *catchments* ont longtemps été étudiés dans des séquences où la charge sémantique propositionnelle du discours est telle que le nombre de gestes référentiels est très élevé (e.g. McNeill et al., 2001). Or, les séquences contenant le plus de relatives déterminatives prises en compte dans cette étude sont les séquences argumentatives (41 % des séquences contenant une relative déterminative sont des argumentations). Dans l’extrait (3) suivant, Rhianna explique que les Parisiens jugent leur ville bien meilleure que les autres. Alex, la co-locutrice, produit un *backchannel* en comparant l’engouement des Parisiens à celle des Londoniens pour leur ville. Rhianna rejette cette proposition et intensifie ses arguments avec l’adverbe “ever”. Elle produit également une série de gestes bimanuels au long de la séquence.

3. Un *catchment* (McNeill, 2000; McNeill et Duncan, 2000) désigne la récurrence formelle de certains traits gestuels à travers plusieurs groupes intonatifs dans le discours, formant une unité discursive thématique. Cette notion a d’abord été définie avec des corpus de récits élicités, facilitant la production de gestes référentiels tels que les iconiques. Dans notre corpus semi-spontané, les iconiques figurent de manière moins systématique et les séquences discursives sont plus hétéroclites. Nous repérons et annotons les *catchments* uniquement dans les cas de répétition formelle d’un ou plusieurs traits gestuels lors d’une série de plusieurs gestes manuels, que cette répétition se trouve dans des gestes consécutifs ou non. Pour que des récurrences forment un *catchment*, la similarité doit s’observer sur l’inclinaison, l’orientation, et la trajectoire de la main réalisant le geste, ainsi que sur les coordonnées de ce même articulateur dans l’espace de gesticulation du locuteur. Ces critères plus formels sont adoptés d’après Hogrefe et al. (2011), et d’après leur notion de *hammering distance*, qui fournit une grille de traits formels servant à quantifier la différence de configuration entre deux gestes manuels. Les gestes de très petite ampleur ou faiblement articulés comme une ouverture de la paume vers le haut sont exclus, puisque ces gestes possèdent une très faible variabilité et nécessitent peu d’effort, articuloire comme communicatif.

- (3) Rhianna L it's like
 (h) [(a) Paris is the] [(b) only place]
 SC **you** [(c) **could ever live**] **in #**
 Alex (laughs)
 Rhianna R [(d) any other town]
 e[(e)ven Lyon]
 (h) or Marseille
 [(f) is a small town]
 [(g) even Lon- well] actually
 i didn't have that thing with London

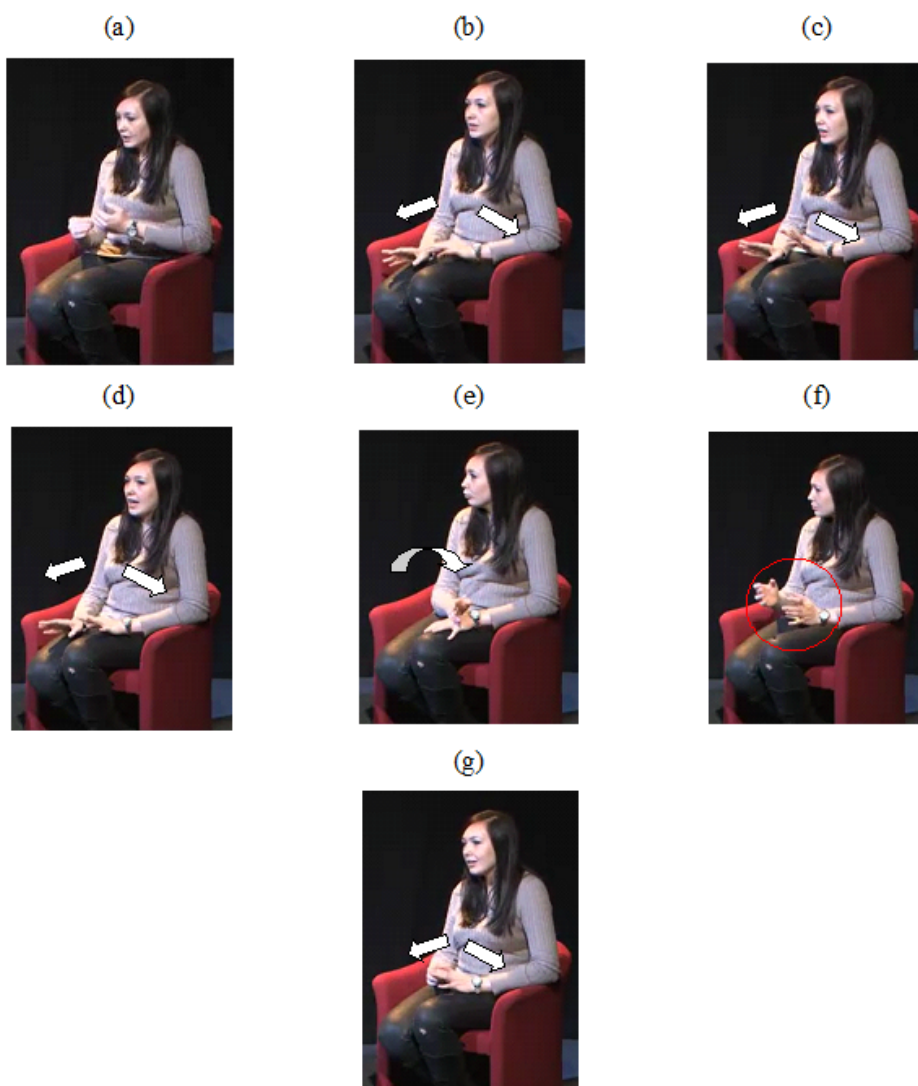


FIGURE 6.3 – Série de gestes manuels au cours de la séquence (3), parmi lesquels la configuration bimanuelle paumes vers le bas, trajectoire oblique des deux côtés de l'espace de gestualisation est répétée, faisant office de *catchment* dans la séquence. Ce *catchment* est visible dans les vignettes (b), (c), (d), et (g).

La figure 6.3 montre le fil conducteur visuo-spatial tissé avec l'argumentation

de Rhianna. En commençant la séquence par “it’s like”, Rhianna annonce un glissement de rôles énonciatifs : elle se met à place des Parisiens dans le reste de la séquence et les cite directement en les imitant de manière hyperbolique. Après avoir produit un adaptateur (a), cette dernière réalise un premier geste métaphorique (b) sur L en co-occurrence avec “the only place”, à l’aide de ses deux mains, placées de manière symétrique au-dessus de ses cuisses. Les paumes sont orientées vers le bas, légèrement obliques. Les deux mains effectuent une coupe latérale de chaque côté de l’espace de gestualisation de Rhianna. Le sémantisme de ce métaphorique autour de la négation fait écho à l’adjectif “only”, indiquant à la fois l’exclusivité de Paris par rapport aux autres villes (i.e. la négation d’autres équivalences; Calbris, 1985) et le rejet de la proposition d’Alex comparant Paris à Londres. Un second métaphorique (c) est produit sur SC, faisant cette fois-ci écho à l’adverbe intensifieur “ever”. En revanche, la configuration, l’orientation, et la trajectoire de ce second métaphorique sont similaires au premier. Après une pause silencieuse, Rhianna reprend cette posture hyperbolique en proposant un exemple à l’argument de l’exclusivité Parisienne, et compare Paris à d’autres grandes villes de France. Le déterminant indéfini “any” est cette fois-ci utilisé, tandis qu’un troisième métaphorique (d) présentant des traits semblables aux deux autres est produit sur la totalité du groupe intonatif. “Any” et “ever” montrent un parcours mental des différentes possibilités, parcours représenté par le mouvement latéral des mains. Ces différents groupes intonatifs sont accompagnés de battements de têtes, ainsi que de gestes de négation : la densité modale est très forte en ce qui concerne la gestualité. La répétition de ces unités gestuelles (des mains et de la tête) est ici emphatique (Streeck, 2009) et participe à l’imitation hyperbolique de Rhianna. Alex est sensible à cette exagération : en riant ostensiblement, elle reconnaît le tour de parole de Rhianna comme humoristique. Les trois groupes intonatifs suivant R, commençant avec “even Lyon” et finissant avec “is a small town”, présentent des gestes manuels différents. L’introduction du référent “Lyon” est précédée d’un pointage abstrait de la main droite (e), plaçant ce référent tout à gauche de l’espace de gestualisation de Rhianna, devenu une carte de France. Le référent suivant, “Marseille”, ne fait en revanche pas l’objet d’un placement sur cette même carte. Un iconique (f) et un battement de tête sont produits en co-occurrence avec “small town”, Rhianna plaçant ses deux mains de manière symétriques, paumes orientées l’une vers l’autre, semblant tenir un petit objet. Alors qu’elle introduit une nouvelle ville en continuant la liste d’exemples, Rhianna reprend la configuration des métaphoriques précédents (paumes vers le bas, coupe latérale). Ce geste (g) est néanmoins de plus petite ampleur, et sa production se trouve stoppée lorsque Rhianna s’interrompt pour modifier sa ligne d’argumentation : “London” devient le support d’un élément concessif, nuancant les

propos de Rhianna. Le fait que Rhianna ait recours à la même configuration gestuelle (métaphorique au sémantisme négatif, paumes vers le bas, coupe latérale) à différents moments de ce passage suggère que Rhianna prédique toujours à propos de la ville de Paris, et tient toujours le rôle énonciatif d'une parisienne de manière hyperbolique. La cohésion thématique apportée par les gestes structure l'argumentation de Rhianna. Produit en co-occurrence avec l'un de ces gestes manuels, SC est bien marqué comme partie intégrante de la séquence discursive en cours. Le tableau est différent sur le plan vocal : si les différents groupes intonatifs de la séquence sont systématiquement marqués d'un contour emphatique montant-descendant, celui du groupe intonatif correspondant à SC est moins important. Alors que la prosodie ne participe ni à l'intégration, ni à l'isolation de SC, la gestualité l'inscrit pleinement dans la séquence en cours en lui confiant un mouvement propre, mais à l'exacte image des autres.

Par ailleurs, si la proportion des gestes manuels réalisés en chevauchement entre L, SC, et R n'est pas significativement différente entre les séquences contenant des relatives déterminatives et celles contenant des circonstancielles, le chevauchement est tout de même important (40 % des séquences contenant des relatives déterminatives sont produites avec des gestes manuels en chevauchement sur deux groupes intonatifs ou plus). Contrairement aux autres types syntaxiques, ce chevauchement est très localisé : 72.7 % des chevauchements observés sont des gestes manuels produits à cheval sur L et SC. Les locuteurs choisissent alors fréquemment d'inclure les déterminatives dans la même unité articulatoire que leur co-texte gauche. Ce choix est intéressant dans la mesure où nous avons vu dans le Chapitre 4 d'une part que les relatives déterminatives sont très fréquemment intégrées au groupe intonatif précédent, et d'autre part que les relatives déterminatives sont significativement plus longues que leur co-texte (gauche comme droit). Ce partage de certaines structures articulatoires est visible dans la séquence (4)⁴ suivante, représentée par la figure 6.4.

(4)	Joey		[(a) the you know that small table #
		L	that's an Ikea] [(b) one
		SC	they've bought them][(c)selves#
	Elena		NODS
	Joey	R	like like little bits and pieces of it]

À l'aide de questions-réponses, Joey et Elena passent en revue les éléments de décoration qu'elles ont remarqués la veille lors d'une soirée chez des amis en collocation. Joey centre l'attention d'Elena sur un nouveau référent dans le discours à

4. Dans cet exemple, l'activité gestuelle de la co-locutrice représente un *backchannel* et est décrite en majuscules.

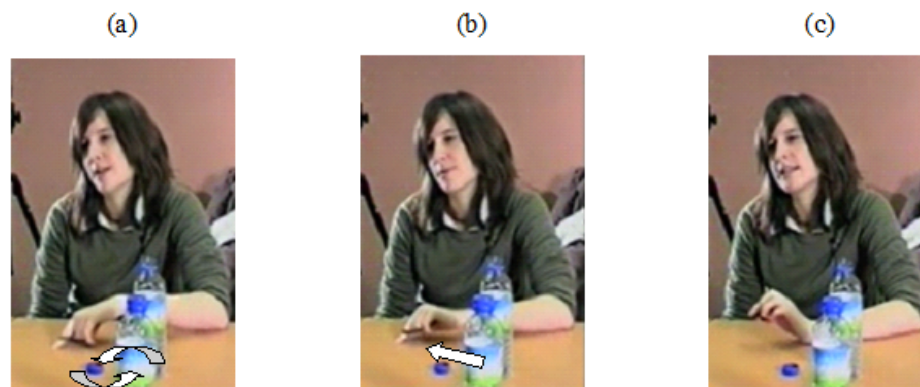


FIGURE 6.4 – Deux pointages référentiels montrant des trajectoires très distinctes au cours de la séquence (4). Le second pointage est réalisé en chevauchement sur la fin du groupe intonatif de L et les trois-quarts du groupe intonatif de SC.

l'aide du groupe intonatif précédant L. Elle active le référent “small table” par un démonstratif distal (“that”), suggérant que ce dernier est accessible : l'information est à récupérer parmi les éléments visuels faisant partie de leurs connaissances partagées (les objets étaient physiquement présents dans l'espace interactionnel de la veille). L, SC, et R maintiennent l'activation de ce même référent : une variation de formes lexicales réduites font allusion à “that small table”, comme “one” et “it”. Une chaîne référentielle est construite, participant à la construction de la cohérence et évoquant une continuité de l'expérience cognitive (De Beaugrande et Dressler, 1981, p. 36). SC propose à “that small table” une caractéristique contrastive. D'un point de vue gestuel, L et le groupe intonatif précédent sont groupés dans une seule unité gestuelle (a) : Joey réalise un pointage abstrait de sa main gauche, l'index tendu. Ce pointage n'est pas fixe : Joey produit en même temps un mouvement circulaire du poignet en continu. Plutôt que d'ancrer le référent “table” dans l'espace en un point fixe, Joey en délimite les contours en fonction de l'espace pris dans la pièce. L'adjectif “small” fait alors davantage l'objet de sa description que la forme de la table en elle-même. Elle stoppe le mouvement circulaire avant l'énonciation de “one”, le dernier item lexical de L, et déplace son pointage en traçant une ligne oblique sur la table (b), partant du centre vers le bord, et sur le côté droit : elle amène le pointage vers elle. Cette unité gestuelle est en chevauchement sur la fin de L et les trois-quarts du groupe intonatif de SC. La configuration et la direction de ce mouvement donnent des informations distinctes et complémentaires. La configuration, par le maintien de l'index tendu, participe à la cohérence du passage : on prédique sur le même référent. En revanche, la direction du mouvement indique que le focus s'établit sur un autre référent en train d'être lié à “table” : “themselves”. Cet item lexical est

réalisé avec un ton montant-descendant soulignant le contraste. La direction de ce geste manuel place donc le focus sur les participants de l'action. En dirigeant son pointage vers elle, la locutrice prend le rôle de ces participants. Cette déterminative voit le focus étroit opéré sur “themselves” représenté gestuellement, par la direction du mouvement. Bien avant la verbalisation de “bought”, ce geste initie les traits sémantiques d'une acquisition, où un objet est amené vers un espace personnel. Le mouvement bref suggère une action ponctuelle. Cette action est toutefois répartie sur les groupes intonatifs de L et de SC. Après une pause silencieuse entre SC et R, Joey reprend la manipulation de l'élastique autour de son poignet (c), qu'elle avait stoppée au profit du premier geste manuel de l'extrait. La construction de sens manuelle jugée nécessaire à cette unité communicationnelle est achevée.

Le fait que les relatives déterminatives soient régulièrement intégrées dans la même unité articulatoire que leur antécédent permet une discussion du ratio théorique entre gestes manuels et propositions syntaxiques, fixé par McNeill (1992) à la moyenne d'un geste manuel par proposition. Si les relatives déterminatives respectent ce ratio en majorité (i.e. 60 %), cette majorité est extrêmement fragile. La contrainte temporelle ne peut entrer en ligne de compte pour expliquer ce phénomène, puisque nous avons vu que les relatives déterminatives sont les constructions syntaxiques les plus longues en termes de production. En termes de contraintes articulatoires, la durée de ces constructions n'est donc aucunement trop courte pour la production d'un geste, puisque nous constatons que le débit ralentit significativement sur les relatives déterminatives. Cependant, plusieurs facteurs jouent en faveur d'une révision de ce ratio. McNeill (1992) ne fournit d'abord pas de détails sur sa définition de la proposition. Cette définition ne permet par exemple pas de déterminer si les propositions subordonnées sont appréhendées séparément des propositions non-subordonnées (i.e. “principales”) adjacentes, ou si l'ensemble “principale” - subordonnée compte comme une proposition complexe. Les résultats de McNeill (1992) sont par ailleurs basés sur des productions en contexte exclusivement narratif. Or, une myriade d'autres facteurs influencent la production gestuelle en production spontanée, tels que la situation interactionnelle, les connaissances partagées entre les participants, ou la diversité des buts discursifs des locuteurs. De plus, l'étude de McNeill (1992) ne prend en compte que les gestes référentiels, laissant les battements ou certains métaphoriques de côté. Dans notre corpus, les chevauchements se situent en majeure partie dans les séquences descriptives (45 % des chevauchements) et dans les séquences argumentatives (35 %). Nous concluons que le ratio établi par McNeill (1992) dans les séquences narratives n'est pas valable pour ces deux autres types de séquence en ce qui concerne les relatives déterminatives. L'un des premiers facteurs

explicatifs de cette différence est à notre sens lié aux buts communicatifs des locuteurs dans ces deux types de séquence, et à l'apport sémantique des items lexicaux et verbaux, de nature différente de celui des récits. Les buts communicationnels des séquences descriptives et argumentatives sont en effet moins susceptibles d'être en relation avec la description d'une succession d'actions ponctuelles.

Ce dernier extrait (4) permet également de montrer que les relatives déterminatives font partie de séquences présentant peu de changements de direction du regard. 27 % des séquences contenant des déterminatives ne présentent même aucun changement, à l'image de l'exemple (4) où le regard de Joey reste fixé sur Elena, la co-locutrice. L'une des potentielles explications demeure dans l'appel continu de Joey aux connaissances partagées entre les deux participantes dans cet extrait. Pour autant, les déterminatives et leurs séquences ne possèdent pas la plus haute proportion de regards vers le co-participant (ni sur l'ensemble de la séquence, ni sur la subordonnée). La tendance à utiliser les connaissances partagées dans les séquences contenant des relatives déterminatives reste néanmoins à tester sur l'ensemble des séquences. Si cette tendance s'avère une régularité spécifique aux séquences des déterminatives, le lien entre le maintien de la direction du regard et les connaissances partagées pourra être plus spécifiquement interrogé.

En résumé, les caractéristiques gestuelles des relatives déterminatives les intègrent beaucoup plus à leurs séquences hôtes qu'elles ne les démarquent. Leur seul indice de rupture tient dans leur grande proportion de battements manuels isolés (i.e. ne faisant pas partie d'un *catchment*), ceux-ci significativement plus nombreux que dans leur co-texte (gauche) et que dans les autres types syntaxiques de subordonnées. Ces battements marquent divers phénomènes discursifs. Sur le plan pragmatique, les battements marquent le contenu informationnel le plus pertinent de la séquence, mais aussi les récupérations après une recherche. Sur le plan des représentations partagées entre les participants, ils aident à faire appel aux connaissances partagées, en attirant l'attention du co-participant. Ces phénomènes discursifs sont tous compatibles avec les nombreux résultats de la littérature sur l'efficacité des battements pour faciliter le traitement des co-locuteurs. Puisque les battements sont de nature rythmique, nous pourrions dire que le seul indice de démarcation permettant d'isoler les relatives déterminatives possède un aspect prosodique. En revanche, nous avons vu que la relation temporelle des battements manuels aux frontières prosodiques n'est pas évidente et ne permet pas de conclure que ces gestes manuels délimitent la frontière entre SC et R. Les mouvements de tête, bien que peu nombreux, sont plus à même d'occuper cette fonction. Aucun autre articulateur ne crée de démarcation

avec cette construction syntaxique. Il n'est pas rare que les relatives déterminatives soient réalisées en co-occurrence avec des unités gestuelles en chevauchement sur le groupe intonatif gauche. L'intégration directe des déterminatives à leurs antécédents se produit alors également de manière gestuelle. Par ailleurs, un quart des relatives déterminatives sont produites avec des gestes manuels dont la configuration et la trajectoire sont formellement identiques à plusieurs autres gestes de la même séquence discursive. Plus qu'une intégration à un niveau local, la gestualité œuvre donc également à intégrer les relatives déterminatives à un niveau séquentiel dans le discours.

Compte-tenu du fait que les relatives déterminatives ne présentent qu'un seul indice significatif pouvant être assimilé à un phénomène de rupture, les interactions entre les différentes marques d'intégration observées sont davantage représentatives des séquences contenant ce type syntaxique. 20 % des relatives déterminatives sont par exemple produites en co-occurrence avec des gestes manuels qui font non seulement partie d'un *catchment*, mais sont aussi produits en chevauchement sur le groupe intonatif gauche (i.e. L) et/ou droit (i.e. R). La séquence (5) et la figure 6.5 ci-dessous illustrent cette configuration.

- | | | | |
|-----|-------|----|--------------------------------------|
| (5) | Joey | | the anthers and [(a) stuff # |
| | | | you know] [(b) the # bits] [(c) like |
| | | L | you know and the] [(d) little bits |
| | | SC | that stick out# |
| | Elena | | yeah] |
| | Joey | R | [(e) you have those |
| | | | like] with the pollen on # |

Joey explique à Elena comment obtenir un peu de suc des fuchsias. Elle a beaucoup de difficulté à récupérer l'item lexical approprié pour désigner la partie de ces fleurs où se trouve le pollen, et recourt à deux *catchments* afin d'activer le référent qu'elle s'efforce de verbaliser pour Elena. Le regard tourné vers la co-locutrice, Joey réalise d'abord un pointage bimanuel (a) en co-occurrence avec "stuff # you know" avant L. Ses deux mains sont positionnées en forme de pinces l'une à côté de l'autre, représentant les panicules à l'intérieur de la fleur. Une extraction conceptuelle est réalisée : en localisant les panicules dans les fuchsias, Joey ne désigne pas toute la fleur mais bien une partie au centre de celle-ci. Sans récupération lexicale ni réaction d'Elena, elle réalise un geste bimanuel très différent (b). Joey ouvre les pinces qu'elle avait formées en levant alternativement les index, partant à droite et à gauche. En restant dans le même espace de gestualisation, ce second geste (b) signale que le référent faisant l'objet d'une description est toujours le même. En revanche, ce der-

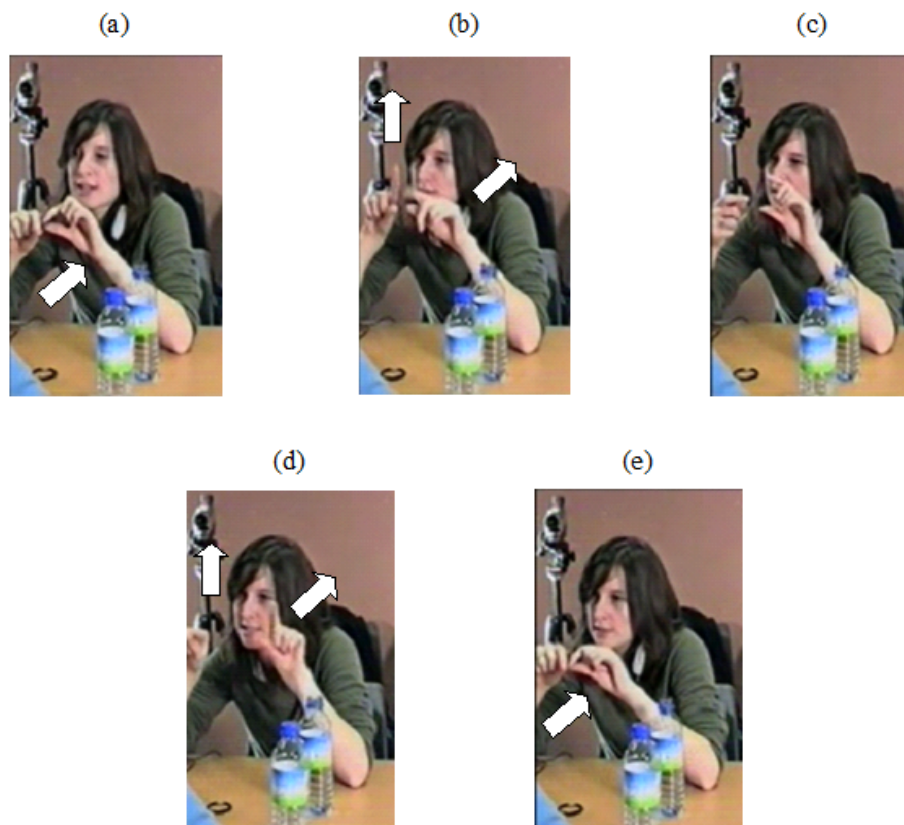


FIGURE 6.5 – Différents gestes manuels au cours de la séquence (5), où deux configurations sont récurrentes : les pinces visibles dans les vignettes (a) et (e), et les panicules dans les vignettes (b) et (d). Le geste manuel (d) est par ailleurs réalisé en chevauchement entre les groupes intonatifs de L et de la relative déterminative.

nier est moins statique et greffe un procès aux coordonnées établies par le premier geste (a). Ce mouvement est iconique : il représente les caractéristiques physiques des panicules, qui sortent de la fleur. La valeur déictique de ce mouvement est très secondaire : l’objectif n’est pas de désigner très précisément la direction dans laquelle partent les panicules. Ce geste apparaît en co-occurrence avec un item lexical extrêmement générique : “the # bits”. Cette association n’a toujours pas l’air de satisfaire le but communicatif de Joey, et ne déclenche pas de réaction de la part de la co-locutrice : l’unité communicationnelle est toujours considérée en cours de construction par les deux participantes. Joey a recours à un autre geste iconique (c) alors qu’elle fait à nouveau appel à la co-locutrice avec “you know”, l’invitant à tirer des inférences (Fox Tree et Schrock, 2002). Ce marqueur sert néanmoins davantage à atténuer le caractère fragmenté de son discours (Tannen, 1986; Schiffrin, 1987; Macaulay, 2002) et à centrer l’attention de la co-locutrice sur le discours à venir (Fox Tree et Schrock, 2002). Le pouce et l’index de Joey forment un croissant. L’espace

vide entre les deux doigts délimitent la taille d'un objet (Hassemer et Winter, 2016). Cet iconique anticipe la verbalisation de l'adjectif "little" qualifiant "bits", énoncé quelques instants plus tard. Le pointage simultané de la main droite de Joey vers cet iconique centre également l'attention de la co-locutrice sur ce geste. Toujours sans récupération lexicale ni collaboration de la co-locutrice, Joey reprend la configuration du geste iconique (b) en levant alternativement les index vers la gauche et la droite dans une nouvelle unité gestuelle (d). Ce geste (d) est d'une ampleur légèrement supérieure, et est réalisé en chevauchement sur L et sur SC. À l'aide de la relative déterminative, Joey réalise une description verbale un peu plus aboutie, qui possède un prédicat (i.e. "the little bits that stick out"). La répétition légèrement amplifiée de cette configuration manuelle, en co-occurrence avec la verbalisation d'un prédicat, déclenche un *backchannel* minimal de la part de la co-locutrice. Une fois assurée que le référent est activé pour Elena, Joey reprend ce référent à l'aide d'une forme lexicale réduite (i.e. le pronom distal "those"), corrélée à la répétition du pointage (a) du début de la séquence, les pinces un peu plus serrées et rapprochées (e). Deux *catchments* ont donc servi à la progression du discours dans cette séquence : la configuration bimanuelle des index levés alternativement à gauche et à droite utilisée dans les gestes (b) et (d), ainsi que la configuration des pinces utilisée dans les gestes (a) et (e).

6.2.2 Circonstanciellles

Présentant deux marqueurs de rupture prosodique, les circonstanciellles se démarquent par des ressources également liées aux gestes manuels. Notons d'abord que les circonstanciellles sont les subordonnées produites en co-occurrence avec le plus grand nombre de gestes manuels, aussi bien par rapport à leur co-texte (68 gestes manuels sur L; 97 sur SC; 72 sur R) que par rapport aux autres subordonnées (76 sur les appositives relatives; 90 sur les relatives déterminatives). Si cette observation ne constitue aucunement un indice de rupture, elle contextualise le premier d'entre eux. Les circonstanciellles sont les subordonnées présentant le plus grand nombre de gestes manuels dans les séquences où aucun autre geste manuel n'est produit ailleurs (i.e. sur L et/ou sur R). 14.5 % des circonstanciellles sont concernées. Si ce pourcentage paraît faible, il est à nuancer par le fait que les séquences sans aucun geste manuel ni sur L, ni sur R sont assez rares d'une manière générale dans le corpus. Cela représente par ailleurs significativement plus que les appositives relatives et les relatives déterminatives, qui comptent toutes deux le même nombre de séquences présentant ces caractéristiques ($F(54, 54) = 2.41, p < 0.001$). L'une des explications

réside dans le fait que L est significativement plus court que SC dans les séquences contenant une circonstancielle. L'exemple (6) associé à la figure 6.6 illustrent la production de gestes manuels en co-occurrence avec une circonstancielle, quand aucun autre geste n'est produit dans le co-texte (gauche comme droit).

- | | | | |
|-----|---------|----|---|
| (6) | Rhianna | | cause they do all those personality tests # |
| | Alex | | (h) [(a) # and # |
| | | | what |
| | | | to be an air hostess |
| | Rhianna | | yeah |
| | Alex | | really |
| | Rhianna | L | well |
| | | SC | if] [(b) you wanna do it on a long] [(c) # |
| | | | you know on long-term basis |
| | | R | then (h) have a number (d) of #] |



FIGURE 6.6 – Geste métaphorique produit en co-occurrence avec une circonstancielle conditionnelle au cours de la séquence (6), dans laquelle aucun autre geste manuel n'est produit.

Dans cette séquence de questions-réponses, Rhianna explique à Alex les différentes étapes du recrutement des hôtesse de l'air. L'image (a) ne représente pas de geste spécifique, mais illustre la position de repos des mains de Rhianna. Au début de cet extrait, Alex exprime sa surprise face à l'information qui vient d'être introduite (i.e. "personality tests") dans une insertion. Elle demande confirmation en liant explicitement cette information à "air hostess". Rhianna répond positivement à l'aide d'un grand hochement de tête, suivi par un *backchannel* minimal beaucoup plus atténué de la part d'Alex, énoncé avec peu d'intensité. Rhianna produit un marqueur de discours, "well", indiquant une insuffisance ou un décalage par rapport aux attentes de sa co-locutrice (Jucker, 1993). Ce décalage porte sur les informations présumées pertinentes à la séquence discursive. Par la simple insertion de ce marqueur, Rhianna indique que les connaissances d'arrière-plan utiles à la séquence vont être renégociées.

Cet arrière-plan contextuel est immédiatement explicité à l'aide de la circonstancielle, directement intégrée au groupe intonatif de L sur le plan prosodique. Si nous avons parlé de valeur prédictive (Dancygier et Sweetser, 2000) pour les circonstancielle conditionnelles dans le Chapitre 3, cette valeur prédictive reste ici secondaire. En surface, Rhianna met en place un raisonnement s'appuyant sur la création d'un cadre hypothétique, et sur la prédiction des conséquences de ce cadre, basée sur les connaissances du monde, et les relations typiques de cause à effet. L'information n'est en réalité aucunement hypothétique, mais bien du domaine de la *realis*. Le fait de présenter une information établie comme de l'arrière-plan présupposé sous la forme d'une circonstancielle est moins direct (Dancygier, 1998) et par conséquent plus poli qu'une clarification donnée à l'aide d'une proposition principale (Ford, 1993; Couper-Kuhlen et Thompson, 2009; Laury et Suzuki, 2011). Rhianna évite donc de remettre en cause les connaissances d'Alex, ou bien de présupposer un manque de connaissances de la part de cette dernière. La face conversationnelle de la co-locutrice est donc préservée (Goffman, 1967; Brown et Levinson, 1987), et les informations d'arrière-plan sont mises à jour. En revanche, Rhianna ne véhicule pas toutes les informations verbalement. En anticipant la recherche lexicale qui lui fait scinder sa proposition en deux groupes intonatifs, elle produit au début de SC un geste métaphorique de la main droite (b) qui complète la fin de son discours. Main ouverte et doigts tendus sur un plan vertical, Rhianna trace une ligne imaginaire droit devant elle sur laquelle elle définit deux points de repère par deux battements. Ce geste représente un axe temporel relativement conséquent, illustrant l'item lexical que Rhianna va récupérer quelques instants plus tard sur le second groupe intonatif de SC : "long-term basis". Au moment de la verbalisation de cette récupération, Rhianna ne gestualise déjà plus (c). Il est intéressant de remarquer que si Rhianna avait jusque-là les yeux fixés sur sa co-locutrice, elle la quitte des yeux dès le retour de ses mains à une position de repos. Elle énonce le second groupe intonatif de SC (présentant pourtant la récupération lexicale de l'information qu'elle a gestualisée) et celui de R avec le regard dirigé vers le côté gauche. La reprise de souffle audible en milieu d'unité intonative et l'abandon de la structure verbale de R vont également vers l'indication d'un désengagement de la part de Rhianna : elle souhaite abandonner son tour de parole, estimant que les informations d'arrière-plan nécessaires à la séquence ont été clarifiées (d).

Si cet exemple (6) présente un unique geste manuel dans la séquence en co-occurrence avec une circonstancielle conditionnelle, cette régularité concerne aussi bien les circonstancielle temporelles. De plus, contrairement à ce que nous attendions, elle concerne aussi bien les circonstancielle préfixées à la structure ver-

bale principale que les circonstancielles postposées. Les circonstancielles préfixées possèdent dans bien des cas un antécédent bref, au contenu sémantique moindre (i.e. série de marqueurs de discours), comme le montre l'exemple (6). Il aurait alors été possible de lier la densité plus importante de gestes manuels sur SC à la brièveté, mais aussi au manque de contenu sémantique de leur antécédent. L'exemple (7) illustre néanmoins la même régularité sur une circonstancielle temporelle insérée en position médiane par rapport à la structure syntaxique principale.

- (7) Tom
- | | |
|----|--|
| | [(a) the same with Plymouth |
| | (h) we went there # |
| L | last summer |
| SC | when we were doing] [(b) that little road trip further there] [(c) # |
| R | for a holiday instead of going abroad] |

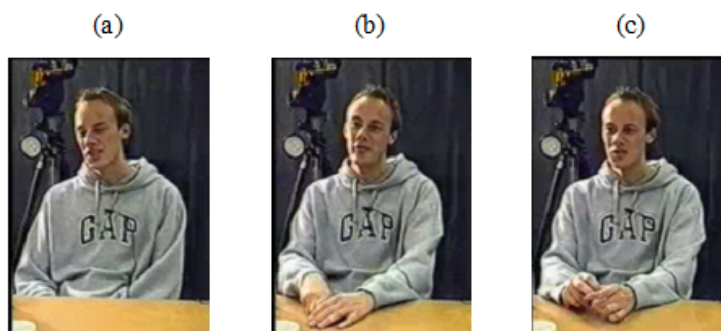


FIGURE 6.7 – Pointage réalisé en co-occurrence avec la circonstancielle temporelle au cours de la séquence (7), dans laquelle aucun autre geste manuel n'est produit.

Au cours d'une séquence argumentative co-construite, Tim et Tom comparent différentes villes du Royaume-Uni selon leur qualité de vie. Chacun avance les raisons qui les ont poussés à choisir l'université de Cardiff plutôt qu'une autre. Tim et Tom viennent d'entériner le fait que Sheffield est une ville très bétonnée. Alors que Tom établit un parallèle avec la ville de Plymouth, il précise les causes de sa visite à l'aide d'une circonstancielle temporelle. Tom lève les yeux vers le co-locuteur en introduisant l'élément référentiel "Plymouth", mais regarde aussitôt ailleurs et ne gestualise pas durant les deux prochains groupes intonatifs (a). Tom lève à nouveau les yeux vers Tim alors qu'il énonce "last summer" sur L, puis ne le quitte cette fois plus des yeux jusqu'à la fin de la séquence. À la différence de la circonstancielle conditionnelle de l'exemple (6), L et SC sont réalisés sous deux groupes intonatifs distincts sur le plan prosodique. Tom produit un pointage bimanuel au milieu de SC (b) en posant ses deux mains sur la table paumes vers le bas. L'index de sa

main droite localise “that little road trip” sur la table, en l’amenant sur la scène du discours (Cibulka, 2015), i.e. l’espace de gestualisation juste devant Tom, directement sous la vue de Tim, où les précédents gestes ont eu lieu durant l’interaction. Cette mise en focus d’un item référentiel se retrouve sur le plan verbal, avec l’introduction de “little road trip” par le pronom démonstratif distal “that”. SC élabore sur le topique discursif de “Plymouth”, présenté avant L, et ne contribue pas à la structure principale (i.e. l’argumentation). SC n’introduit pas moins de nouveaux éléments dans le discours par rapport à l’information nouvelle de L “last summer”, en livrant une expérience personnelle du locuteur. Si L peut être directement lié à R sans créer d’infélicité⁵ sémantique, conceptuelle, ou textuelle, c’est pourtant en co-occurrence avec SC que Tom choisit de produire un geste bimanuel. Les mains de Tom resteront sur la table pour le reste de la séquence (c), ce qui signale une implication de Tom dans la séquence un peu plus importante que L (où ses mains n’étaient pas visibles, son dos contre la chaise témoignant d’une position relativement relâchée), mais moins importante que SC (où ses mains sont directement en action et impliquées dans l’apport communicationnel; Cibulka, 2015). Par le simple fait de laisser ses mains sur la table à l’endroit où de précédents gestes ont été réalisés, Tom maintient son apport dans l’interaction.

Le fait que les circonstancielles soient produites avec significativement plus de gestes manuels est interprétable selon plusieurs traditions théoriques, qui sont ici complémentaires pour caractériser les circonstancielles. Si l’on considère le rôle interlocuteur des gestes co-verbaux (i.e. ces derniers remplissent une fonction communicative et sont également motivés par des facteurs sociaux; Bavelas et al., 1995), l’augmentation du dynamisme communicatif (Firbas, 1971; 1989) pendant la production des circonstancielles suggère que ces structures sont centrées sur le co-locuteur et constituent des signaux encodés comme importants pour ce dernier dans le discours. Le fait que les circonstancielles soient plus modulées intonativement que leur co-texte (cf. Chapitre 4 p. 149) constitue un indice en faveur de cette interprétation. En considérant le rôle intra-locuteur des gestes co-verbaux (i.e. la gestualité co-verbale aide à la production de la parole⁶ et joue un rôle cognitif essentiel), le recours à la gestualité plus important sur les circonstancielles que dans le co-texte suggère que

5. Une infélicité, dans le cadre d’une conversation, est la contradiction des attentes et/ou des présuppositions du co-locuteur (Goffman, 1981).

6. Nous ne faisons ici pas référence à la Lexical Retrieval Hypothesis de Rauscher et al. (1996) et de Krauss et al. (2000), qui propose que la gestualité co-verbale aide essentiellement à la mise en mots. Nous faisons ici appel à la position tenue entre autres par McNeill (1992; 2005) et Goldin-Meadow et al. (2001), reconnaissant l’important rôle cognitif joué par les gestes en plus de leur fonction communicative, en répartissant par exemple la charge cognitive sur plusieurs modalités et dimensions (Kita et Özyürek, 2003; Melinger et Levelt, 2004).

ces structures syntaxiques mobilisent une charge cognitive importante. Cette demande cognitive plus élevée peut être liée à la planification des unités (syntaxiques, discursives, et/ou prosodiques) et/ou à la charge et la complexité sémantiques du contenu à véhiculer (Goldin-Meadow et al., 2001; Goldin-Meadow, 2015). Le fait que les circonstancielles présentent la proportion la plus importante de marques vocales de travail de formulation (cf. Chapitre 4 p. 151) va dans le sens de cette supposition. La synthèse de ces deux interprétations concernant les rôles intra- et inter-locuteur des gestes accompagnant les circonstancielles montre le poids de ce type syntaxique sur le plan cognitif, à la fois pour le locuteur et le co-locuteur. Les caractéristiques verbales des circonstancielles vont également dans ce sens. La diversité de fonctionnement de leurs morphèmes introducteurs peu variés et leur mobilité dans la macrostructure que nous avons évoquées dans le Chapitre 3 expliquent la nécessité pour le locuteur de les rendre bien visibles et identifiables par le co-locuteur. Leur capacité à projeter et/ou clôturer un cadre d'interprétation pour plusieurs groupes intonatifs consécutifs justifie quant à elle leur important coût cognitif.

Selon le schéma *GSA*, i.e. *Gesture as Simulated Action* (Hostetter et Alibali, 2010), les gestes manuels sont des manifestations ouvertes des simulations perceptuelles et motrices qui sous-tendent la pensée et le langage (e.g. Barsalou, 2003). Selon ce schéma, les locuteurs sont capables d'agir sur leur potentiel à exprimer une situation particulière par un geste, en modifiant leur "seuil gestuel". Le seuil gestuel représente la quantité minimum d'action simulée requise pour que le système moteur produise un geste. Les locuteurs peuvent maintenir un seuil haut, empêchant alors la majorité de leurs simulations d'être produites sous la forme de gestes (par exemple dans des contextes où la gestualité est considérée comme inappropriée). De la même manière, les locuteurs peuvent maintenir un seuil bas, en augmentant alors le nombre de simulations qui seront exprimées sous la forme de gestes. Cette situation est particulièrement probable dans les contextes où les locuteurs estiment que les gestes sont particulièrement utiles ou dans des contextes où les locuteurs sont particulièrement motivés à communiquer clairement (Holler et Stevens, 2007; Jacobs et Garnham, 2007; Hostetter et al., 2011). Les raisons et phénomènes évoqués plus haut suggèrent que les locuteurs maintiennent un seuil bas pendant la production des circonstancielles⁷.

Le second marqueur de rupture des circonstancielles concerne la trajectoire des

7. Le fait que certains travaux sur ces phénomènes portent davantage sur les gestes référentiels que sur les gestes non-référentiels dans le discours n'enlève aucune pertinence à l'observation des circonstancielles à la lumière de ce schéma, en raison de l'augmentation significative de la production d'iconiques dans les circonstancielles par rapport à leur co-texte gauche.

gestes manuels. Ce type de subordonnée possède significativement plus de gestes manuels dont la direction est opposée à celle des gestes précédents que L et R (comptant le même nombre de gestes manuels montrant une trajectoire inverse aux gestes précédents; $F(54, 54) = 4.63, p < 0.0001$), et également significativement plus que les relatives déterminatives et les appositives relatives (comptant également le même nombre de gestes manuels montrant une trajectoire opposée aux gestes précédents; $F(54, 54) = 4.63, p < 0.0001$). Conformément à ce que nous attendions, cette régularité concerne surtout les circonstancielles temporelles (80 % des occurrences), plus enclines à être produites en co-occurrence avec l'exploitation gestuelle d'un axe horizontal. L'exemple (8) ainsi que la figure 6.8 illustrent cette régularité.

- (8) Tom L be[(a)cause #]
 SC [(b) when they] got married #
 R they [(c) bought a flat]
 for about ten thousand pounds #



FIGURE 6.8 – Plusieurs gestes manuels dans la séquence (8), parmi lesquels le geste métaphorique (b) réalisé en corrélation avec la circonstancielle présente une trajectoire opposée à celle des autres gestes manuels.

Tom explique à Tim comment ses parents sont revenus habiter à Cardiff après avoir vécu à Londres. Il produit d'abord un geste métaphorique non-référentiel ouvrant la séquence narrative en déplaçant la main droite vers la droite, paume semi-ouverte vers le haut, pouce et index en "L" (a). En produisant SC, Tom joint la main gauche en l'apposant de façon symétrique à la droite, les deux paumes se faisant face. Ce geste ressemble alors à une configuration de *conduit* (e.g. McNeill, 1992), où l'information véhiculée est modélisée en un objet délimité entre les mains du locuteur. En co-occurrence avec l'item verbal "when", Tom déplace cette configuration de

la droite vers le centre, puis revient à droite (b). Tom est encore dans l'organisation de son discours, son geste prenant cette fois-ci une dimension grammaticale en exploitant l'axe linéaire temporel. Dans cet extrait, l'importance des plans et des axes directionnels spatiaux en tant que marqueurs (Calbris, 2011) est primordiale. Sur R, Tom déplace cette même configuration un peu plus loin vers la droite, en localisant ses coordonnées sur la table par un battement corrélé avec "bought", marquant un événement important dans le récit (c). Tom utilise donc la progression sur l'axe horizontal gauche-droite comme une représentation métaphorique et linéaire du temps (Calbris, 2008), en effectuant un retour vers la gauche pour évoquer des événements antérieurs sur la circonstancielle⁸. Sur le plan prosodique, cet énoncé présente également un indice de rupture sur SC, sous la forme d'un intervalle non-neutre entre L et SC. La figure 6.9 montre le rehaussement initial entre la syllabe finale de L et la syllabe initiale de SC, où la F0 passe de 83 Hz à 103 Hz. Il est intéressant d'observer que la rupture créée sur SC est à la fois prosodique et gestuelle. Positionnée à l'intervalle entre L et SC, la rupture prosodique se manifeste néanmoins plus tôt dans le segment : les deux phénomènes de rupture ne sont pas exactement co-occurents.

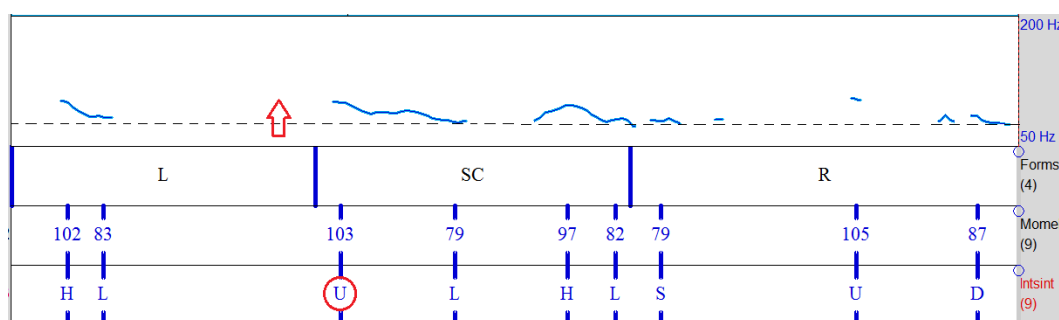


FIGURE 6.9 – Contour intonatif de l'exemple (8) illustré par la courbe intonative de Praat, la transcription des groupes intonatifs, et les valeurs Intsint (U = Upstep; D = Downstep; H = High; L = Low; S = Same). Le rehaussement initial observé sur SC est favorisé par la pause silencieuse qui sépare les segments L et SC. Si la syllabe en attaque de R est plus haute que la finale de SC, R présente un contour assez plat jusqu'à la tête du groupe intonatif.

Ce même exemple (8) permet de rappeler que les circonstanciels encodent verbalement une part considérable de procès ponctuels (58 % des occurrences; voir le Chapitre 3 p. 107). Or, les changements de direction gestuels sur les circonstan-

8. Cette représentation fait écho au système d'écriture occidentale, ainsi qu'à la structure informationnelle du langage écrit (Calbris, 2008), mais aussi oral dans une certaine mesure (i.e. la position des éléments connus, thématiques, à gauche du syntagme verbal pour l'écrit et en début de groupe intonatif pour l'oral, celle des éléments nouveaux, rhématiques, à droite du syntagme verbal pour l'écrit et en fin de groupe intonatif pour l'oral).

cielles sont tous corrélés à l'énonciation d'un procès plutôt que d'un verbe d'état. Si les changements de direction s'observent sur un axe temporel horizontal, à l'image de l'exemple (8), les changements de direction s'observent également sur d'autres coordonnées. La séquence argumentative de Rhianna dans l'exemple (9) permet d'illustrer ces cas de figure.

- | | | | |
|-----|---------|----|--|
| (9) | Rhianna | | the thing is |
| | | | my grand[(a) mother lives in Wimbledon]# |
| | Alex | | oh yeah |
| | Rhianna | L | (h) # and |
| | | SC | when she [(b) comes to get me] |
| | | R | [(c) it's #] kind of far # |

Rhianna donne ici l'une des raisons pour lesquelles elle trouve qu'atterrir à l'aéroport de Stansted n'est pas très pratique. La figure 6.10 permet d'observer les stratégies gestuelles dont elle fait preuve.

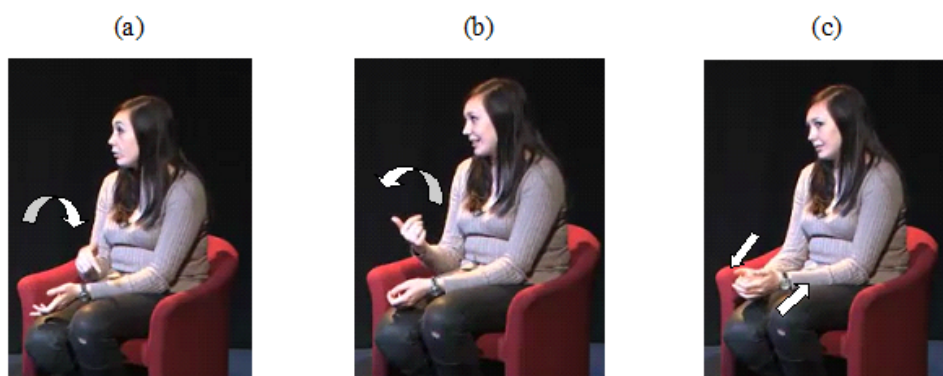


FIGURE 6.10 – Plusieurs gestes manuels dans la séquence (9), parmi lesquels le pointage (b) en corrélation avec la circonstancielle présente une trajectoire opposée à celle du pointage précédent (a).

En localisant le référent “Wimbledon” dans l'espace physique à l'aide d'un pointage abstrait dirigé vers ses genoux et légèrement à gauche (a), Rhianna tourne le regard vers Alex et hausse les sourcils : par cet appel au co-locuteur, elle vérifie que cet élément fait bien partie des connaissances partagées entre les deux participantes, qui sont toutes deux anglaises et connaissent bien les environs de Londres. Alex donne d'ailleurs un *backchannel* minimal d'alignement⁹, indiquant avec le marqueur de discours “oh” que la nouvelle information implique un changement d'interprétation de la séquence, même si ce nouvel élément lui est connu (“yeah”).

9. Le co-locuteur produit un signal encourageant la progression du locuteur (Stivers, 2008).

Rhianna détache ensuite le regard sur SC en réalisant un autre pointage abstrait référentiel (b), en corrélation avec l'énonciation du prédicat “comes to get me”. Ce pointage prend une trajectoire inverse au pointage précédent, partant de ses genoux vers l'extérieur, et désigne un espace moins précis à sa droite, repéré par rapport au premier. À l'image des choix lexicaux et verbaux de Rhianna, ce second pointage encode un mouvement dont sa grand-mère est à l'origine. En revanche, geste et parole expriment un point de vue différent. Alors que “comes” suggère uniquement la convergence vers Stansted, la direction du pointage permet de représenter une excursion de Wimbledon vers Stansted, dont la localisation est représentée comme lointaine. Rhianna produit un battement de tête en co-occurrence avec “comes”. Ce mouvement de tête suggère une étape importante du discours : Rhianna souligne non seulement que sa grand-mère se déplace pour aller la chercher, mais surtout que Stansted représente une distance considérable pour cette dernière, ce qui ennue Rhianna. Le battement fournit alors des informations à un niveau supérieur à celui des pointages qui implique directement les référents : celui de l'organisation pragmatique du discours. Elle produit sur R un commentaire conclusif en inclinant la tête sur le côté, ainsi qu'un geste métaphorique (c) qui incarne ce glissement dans les niveaux de discours. Rhianna ouvre ses deux mains paumes vers le haut, et les lève alternativement, pesant métaphoriquement différents paramètres comme dans l'élaboration d'un jugement. Elle prend ici position sur les événements qu'elle vient de décrire en tant qu'énonciatrice.

Trois autres régularités concernant les circonstancielles sont observables. Cependant, ces trois caractéristiques n'atteignent pas la significativité statistique et sont donc considérées comme des tendances. La première concerne la dynamicité des gestes manuels co-occurents à ce type syntaxique. Aucune circonstancielle de notre corpus ne présente de tenue de geste manuel (i.e. le maintien d'une configuration manuelle de manière statique, avant ou après la phase dynamique de réalisation du geste). Les gestes manuels réalisés dans ces constructions syntaxiques sont donc plutôt d'une nature brève et dynamique. Notons toutefois que la proportion de gestes réalisés en chevauchement est élevée (40 % des circonstancielles sont produites avec des gestes réalisés en chevauchement sur le co-texte gauche et/ou droit). Cette remarque ne s'applique donc pas au partage des unités gestuelles, assez commun, mais bien à la proportion de poses manuelles statiques. La tenue de gestes manuels s'observe par contraste sur 9.1 % des appositives relatives, et sur 3.6 % des relatives déterminatives. De même, la tenue de gestes concerne 5.5 % des co-textes gauches des circonstancielles, et 9.1 % des co-textes droits. L'exemple (10), associé à la figure 6.11, montre une série de gestes manuels différents en co-occurrence avec une

circonstancielle. Dans le co-texte droit (i.e. R), la configuration manuelle du dernier geste réalisé est tenue sur l'intégralité du groupe intonatif.

- (10) Joey cause if you go through # town #
 no #
 L [(a) yeah
 SC if you go #
 a bit] [# (b) like along Park Place]
 [(c) and up Newport Road #]
 R [(d) it's not that far]



FIGURE 6.11 – Série de pointages dans la séquence (10). La configuration du dernier pointage réalisé à la fin de SC est maintenue de manière statique sur l'intégralité de R.

Joey et Elena discutent du meilleur itinéraire en voiture pour aller au domicile de Joey en partant du centre-ville. Elena explique que l'itinéraire pris par l'un de leur amis a donné l'impression que le domicile de Joey était assez loin du centre. Joey produit une première circonstancielle conditionnelle, puis s'interrompt et se corrige par le signal "no" à la suite d'une longue pause silencieuse, suggérant que le début de son tour est un faux-départ. Elle effectue un retour sur son propre discours par le signal "yeah" sur L, en entérinant finalement les constituants du groupe intonatif qu'elle avait préalablement signalés comme une erreur. Au lieu de reformuler une partie de son discours précédent, Joey produit une seconde circonstancielle conditionnelle (SC), qu'elle scinde en trois groupes intonatifs, à l'aide entre autres d'une pause silencieuse démarcative. La seconde pause silencieuse peut s'apparenter à de l'hésitation due à la projection d'unités (le débit de parole est plus lent avant la pause qu'après (Ferré, 2004), et la structure "a bit" fait l'objet d'un abandon). Joey retrouve la fluidité après le marqueur "like", qui ressemble à un remplisseur au vu des difficultés de Joey, mais introduit aussi le focus du groupe intonatif, délivré sans

aucune disfluente. Cette circonstancielle conditionnelle présente un chemin alternatif à celui pris par l'ami des participantes, décrit précédemment par Elena. Les deux participantes ont co-construit la même conclusion concernant le chemin pris par leur ami : le domicile de Joey paraissait très loin de leur point de départ. Proposer un chemin alternatif sous la forme d'une conditionnelle permet d'orienter la séquence vers la conclusion inverse (i.e. "it's not that far" sur R), et de la présenter comme conséquence logique d'une série d'actions, à la manière d'une relation de cause à effet. La conditionnelle permet donc à Joey de présenter son itinéraire comme logique, et lui donne plus de pertinence tout en servant son argumentation. Alors qu'elle ne gestualise pas pendant la production de la première circonstancielle conditionnelle avant L, Joey effectue une série de pointages à l'aide de sa main droite, commençant sur L. Elle trace une série de trajectoires sur la table du bout des doigts, la table servant de plan de ville fictionnel. Les deux premiers pointages (a) et (b) marquent sensiblement la même trajectoire oblique, allant du haut de la table (vers Elena) à droite vers le bas de la table (vers Joey). En revanche, ces deux gestes ne sont pas réalisés avec la même configuration. Lors du premier pointage (a), tous les doigts de Joey sont en contact avec la table, mais sont groupés en forme de pince. Le second pointage est effectué seulement avec l'index. Ces deux pointages sont à l'image de sa production verbale. Alors que L et le premier groupe intonatif de SC témoignent d'une impulsion communicative sans donner d'information précise ("yeah if you go # a bit"), la fin du second groupe intonatif de SC donne des informations précises avec des référents définis. Le premier pointage définit le mouvement général alors le second représente un ré-encodage plus précis à l'image de la reformulation. Joey lève les yeux à la fin de ce second pointage alors qu'elle énonce le référent "Park Place", faisant appel aux connaissances partagées entre les participantes. Le troisième pointage (c) est différent en ce qu'il ne dessine plus de trajectoire directement sur la table, son index survolant la surface. Son index finit par se poser sur la table, et définit un point spécifique légèrement plus haut que les coordonnées des gestes précédents. Joey maintient le regard vers Elena, et marque "Newport Road" d'un grand battement de tête. Le maintien du regard prolonge l'appel aux connaissances partagées alors que le battement de tête annonce la fin de l'itinéraire. Joey produit la conclusion de cet itinéraire sur R, "it's not that far", en maintenant le regard vers Elena, mais aussi en tenant le dernier pointage réalisé sur SC. Son index est toujours sur la table, pointant vers les coordonnées assignées à "up Newport Road". L'immobilité de cette configuration suggère que les coordonnées spatiales établies précédemment ainsi que le cadre cognitif associé sont toujours valides pour le groupe intonatif qu'elle énonce (i.e. R). La configuration manuelle en elle-même indique que les référents précédemment activés dans le discours, à défaut d'être pleinement ac-

tifs, ne doivent pas être désactivés (Enfield, 2009).

La seconde tendance liée aux circonstancielle restant non-significative concerne la forme des gestes manuels. Seulement 9.10 % des séquences comprenant une circonstancielle comportent au moins une répétition formelle d'un geste manuel précédent (que ce dernier ait été produit sur L ou sur SC), alors que cette caractéristique concerne 18.2 % des relatives déterminatives et 12.7 % des appositives relatives. L'exemple (11) et la figure 6.12 présentent une séquence autour d'un même référent, dont les gestes manuels montrent des configurations bien distinctes.

- | | | | |
|------|-------|----|---|
| (11) | Elena | | [(a) cause she was saying |
| | | L | oh <u>it's</u> really] shiny |
| | | SC | when <u>it's</u> all [(b) straight |
| | Joey | | mh |
| | Elena | R | and] [(c) like <u>it</u> looks] really nice |



FIGURE 6.12 – Deux gestes manuels (b) et (c) aux configurations distinctes dans la séquence (11), malgré la reprise anaphorique d'un même référent en tant que topique sous une forme lexicale réduite dans la modalité verbale.

Elena livre un récit à propos du lissage de ses cheveux chez le coiffeur, et relate le point de vue de la personne l'ayant coiffée. Dans la modalité verbale, SC, comme L et R, présente le pronom "it" en élément thématique, recouvrant le topique de la séquence discursive, mentionné au début de celle-ci par Elena : "my hair". En plus d'impliquer ce référent sous la même forme réduite dans un rôle identique, cette circonstancielle s'intègre remarquablement dans la séquence, ne présentant aucune disjonction temporelle ou de rupture par rapport au type de la séquence. SC apporte un élément descriptif comme L et R, et l'intégralité de la séquence est réalisée au présent simple. L'espace de projection cognitive est le même pour toute la séquence,

then it's fine (laughs)

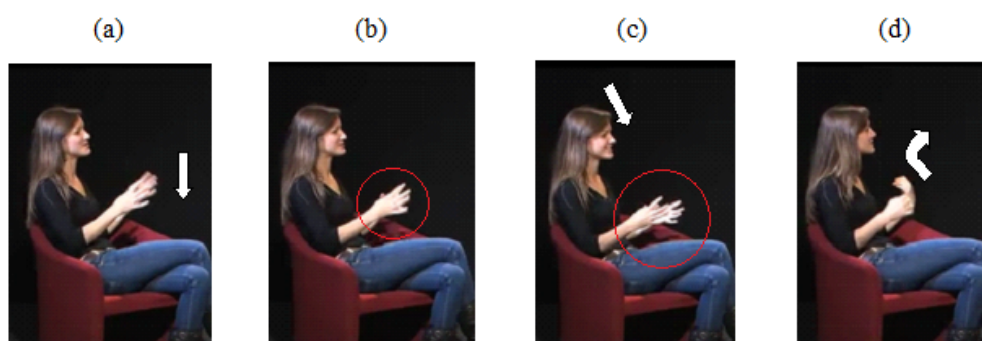


FIGURE 6.13 – Battement de tête (c) produit en co-occurrence avec une circonstancielle dans la séquence (12). La séquence est également caractérisée par une série de gestes manuels métaphoriques, et par un pointage.

Cette circonstancielle servant de cadre au discours d'Alex se trouve dans une description de l'angoisse qu'éprouve cette dernière à bord d'un avion. Alex produit un battement sur L à l'aide de ses deux mains paumes ouvertes vers la co-locutrice (a), pouvant s'apparenter à une technique de centrage d'attention. Alex joint ensuite ses mains du bout des doigts, atteignant une configuration en pyramide au début de SC (b). Cette configuration est d'ordinaire mobilisée pour atteindre la précision en délimitant un point discursif particulier (Calbris, 2011). Elle forme subséquentment un objet rond entre ses mains, paumes ouvertes se faisant face, les doigts écartés. Ce geste en *conduit* modélise la situation décrite en un petit objet délimité, donné à la co-locutrice. Il est accompagné de plusieurs battements de tête assez larges, en co-occurrence avec "turbulence" (c). Ces battements participent à la création d'un cadre de discours (McClave, 2000), et montrent que la structuration textuelle du discours dans les circonstanciels se réalise également gestuellement. À l'intérieur de ce cadre discursif, sur R, Alex spatialise et ancre des éléments référentiels (les "air hostesses" dont il est question plus tôt dans la séquence) dans l'espace physique par un pointage abstrait (d). Ce pointage sur R n'est pas au même niveau dans la structure du discours que L et que SC, d'une façon inverse à l'exemple (11) précédent. Alors qu'Alex gestualise en tant qu'énonciatrice à l'aide de battements et de métaphoriques qui donnent des informations sur les étapes de son discours, le pointage de R fait partie d'une description des items à l'intérieur de la situation discursive qui a été créée.

Si la capacité des battements de tête à marquer une frontière dans le discours est

beaucoup documentée par la littérature (Hadar et al., 1983; Maynard, 1987; Munnhall et al., 2004; Beskow et al., 2006; Krahmer et Swerts, 2007; Barkhuysen et al., 2008; Al Moubayed et Beskow, 2009; De Kok et Heylen, 2009) parmi la multiplicité de fonctions qu'ils remplissent dans le discours (McClave, 2000; Jantunen, 2017), Barkhuysen et al. (2008) vont plus loin en montrant que les battements de tête se produisent davantage en fin d'unité prosodique qu'en milieu. Ishi et al. (2014) montrent des résultats similaires sur du japonais, avec 80 % de battements de tête en co-occurrence avec la syllabe finale d'unité. En revanche, les battements de tête de notre corpus ne montrent pas cette caractéristique. Ils ne sont pas majoritairement localisés en fin d'unité dans notre corpus : seulement 47.9 % des battements sont produits en co-occurrence avec la syllabe finale du groupe intonatif en cours. La durée de ces battements n'est pas non plus significativement différente de ceux réalisés en milieu de groupe intonatif. La durée des battements réalisés en co-occurrence avec la syllabe finale du groupe intonatif de L n'est pas non plus significativement différente de celle des battements réalisés en co-occurrence avec la syllabe finale de SC.

Si les circonstancielles comptent un nombre raisonnable d'indices et de tendances de rupture, nous avons rappelé plus haut que la proportion de gestes manuels réalisés en chevauchement sur plusieurs groupes intonatifs est relativement élevée dans les séquences contenant les circonstancielles. Plus spécifiquement, les circonstancielles se différencient surtout des autres types syntaxiques par leur proportion de chevauchement entre SC et R (14.5 % vs. 3.6 % pour les relatives déterminatives et 1.8 % pour les appositives relatives), ce qui suggère un partage des unités gestuelles entre les circonstancielles et leur co-texte droit assez courant (relatives déterminatives : $F(54, 54) = 3.55$, $p < 0.0001$; appositives relatives : $F(54, 54) = 6.96$, $p < 0.0001$). Notons également que 9.1 % des séquences contenant une circonstancielle sont caractérisées par le chevauchement d'un geste manuel sur trois groupes intonatifs (i.e. L, SC, et R). Cette proportion est moins importante dans l'environnement des relatives déterminatives (7.3 %; en revanche $p > 0.05$) et dans celui des appositives relatives (3.6 %; $F(54, 54) = 2.36$, $p < 0.0001$). La séquence (13) ci-dessous, illustrée par la figure 6.14 à sa suite, comporte un geste manuel en chevauchement à la fois sur la fin de L, l'intégralité de SC, et le début de R :

- (13) Tim L [(a) al] [(b) though
 SC **when i passed**
 R (h) i didn't think] [(c) i was a very good driver #]

Alors que les deux participants discutent à propos du permis de conduire, Tim initie une séquence argumentative sur L. Ses mains sont en position de repos (a).

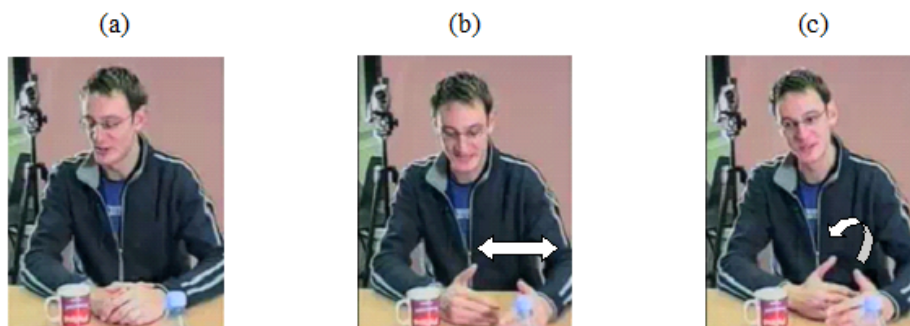


FIGURE 6.14 – Geste métaphorique (b) dans la séquence (13), produit en chevauchement sur les groupes intonatifs de L, de SC, et celui de R, incluant la circonstancielle temporelle dans une unité cognitive commune aux trois unités prosodiques.

Au milieu de l'énonciation de la conjonction servant de marqueur de discours sur L, Tim ouvre les mains en les éloignant l'une de l'autre (b), les deux pouces vers le haut. La main gauche glisse du centre vers la gauche alors que la main droite se rapproche du locuteur. Ce geste métaphorique réalisé à deux mains organise son discours : Tim se concentre sur l'ouverture d'une unité cognitive, plus large que ses unités propositionnelles verbales ou que ses groupes intonatifs : c'est un mouvement grammatical (ici un retour en arrière) que Tim manipule. Une fois ce cadre ouvert, Tim garde la configuration du geste métaphorique en y ajoutant un retournement de la main droite : il ouvre davantage la paume vers lui et réalise un battement de la main gauche en synchronie (c). Il lève également les yeux vers son interlocuteur : Tim marque son retour dans l'espace co-énonciatif et ajoute de la modalité à l'environnement gestuel qu'il a établi, en co-occurrence à la prise de position qualitative verbale "good driver".

En résumé, les circonstanciels sont caractérisés par une activité manuelle assez dense. Les gestes manuels sont d'abord plus susceptibles d'être produits en co-occurrence avec les circonstanciels qu'avec leur co-texte gauche et/ou droit. Il n'est également pas rare que les gestes manuels produits en co-occurrence avec les circonstanciels exploitent une trajectoire opposée à celle des gestes précédents. Cette régularité ne comporte toutefois pas une distribution équilibrée dans nos circonstanciels, concernant majoritairement les circonstanciels temporels. D'autres tendances sont observables, bien qu'elles manquent de robustesse statistique. La capacité des circonstanciels à exprimer des procès ponctuels est marquée à la fois verbalement et gestuellement, puisque ce type syntaxique n'est propice ni à la tenue de geste, ni aux répétitions formelles de gestes réalisés précédemment. En revanche,

il convient de noter que les gestes produits en co-occurrence avec les circonstancielles sont bien souvent en chevauchement avec le groupe intonatif gauche et/ou droit. L'espace ouvert par ces gestes a donc tendance à être plus grand que celui des unités prosodiques et syntaxiques. Cet espace plus large peut tout aussi bien concerner le co-texte gauche que le co-texte droit. Ce dernier phénomène représente une vraie spécificité chez les circonstancielles.

Malgré leur nombre raisonnable d'indices et de tendances de rupture observés, les circonstancielles font preuve d'une remarquable absence de combinaison de marques de rupture. Les gestes manuels isolés sur SC dans les séquences ne comportant aucun autre geste sur L et R n'exploitent pas spécialement une direction opposée à celle des gestes réalisés auparavant. Aucune de ces deux caractéristiques autour des gestes manuels ne montre une association avec la présence de battements de tête : seulement 3.6 % des circonstancielles de notre corpus présentent à la fois un geste manuel isolé (i.e. sans aucun autre geste sur L et R) et un battement de tête. De même, seulement 1.8 % des circonstancielles présentent à la fois un geste manuel exploitant une trajectoire opposée à celle utilisée par de précédents gestes manuels et un battement de tête. Cette absence d'association est relativement inattendue du fait de la grande fréquence de ces battements de tête, et de la grande fréquence des gestes manuels dans les circonstancielles. En ce qui concerne l'intégration des circonstancielles, nous avons mentionné la part importante de chevauchement des gestes manuels sur plusieurs groupes intonatifs. Or, aucun geste manuel exploitant une direction opposée à celle des gestes précédents sur SC ne se trouve également en chevauchement sur plusieurs groupes intonatifs. Les unités gestuelles des gestes manuels isolés initiés sur SC ne continuent pas sur R. Les battements de tête ne sont pas non plus co-occurents à des gestes manuels en chevauchement sur plusieurs groupes intonatifs (cette caractéristique concerne 7.3 % de nos circonstancielles). Nous concluons que la densité modale pour exprimer la rupture de manière gestuelle n'est pas importante, et que les circonstancielles ne cumulent pas plus d'un seul indice de rupture à la fois, tout comme elles ne mêlent pas indices de rupture et indices d'intégration dans une même séquence.

6.2.3 Appositives relatives

En rassemblant quatre indices significatifs de rupture, les appositives sont clairement autonomisées de manière gestuelle. Elles se distinguent réellement des relatives déterminatives et des circonstancielles par la variation de ces indices de rup-

ture. Alors que ces deux derniers types syntaxiques montrent majoritairement des caractéristiques liées aux gestes manuels, les indices de rupture des appositives impliquent une plus grande variété d’articulateurs.

En ce qui concerne les gestes manuels, les appositives relatives sont caractérisées par des unités gestuelles qui leur sont propres : ces unités gestuelles sont rarement en chevauchement sur plusieurs groupes intonatifs. Le chevauchement concerne seulement 14.5 % des séquences contenant une appositive relative, ce qui représente significativement moins que les circonstancielles et les relatives déterminatives, comptant la même proportion de chevauchement ($F(54, 54) = 1.93, p < 0.01$). Pourtant, nous avons vu dans le Chapitre 4 que les appositives sont significativement les subordinées les plus courtes et les plus rapides. L’exemple (14) ainsi que la figure 6.15 illustrent cette caractéristique.

- (14) Tim using a program called Author Catway
 L [(a) actually draw #] redevelopments #
 SC **which # [(b) we need to learn how to (c) use #]**
 R [(d) but it should be quite interesting] #

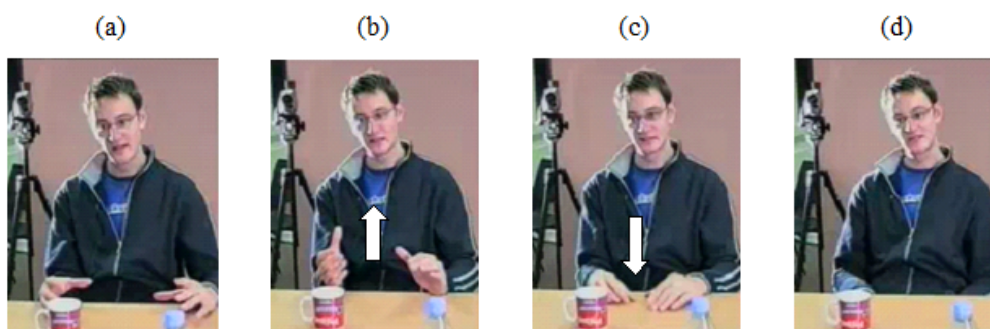


FIGURE 6.15 – Geste manuel iconique produit sur L (a), suivi d’un métaphorique (b) sur SC dans la séquence (14). Un battement (c) se greffe sur ce métaphorique.

Tim explique dans cette séquence descriptive la façon dont les étudiants urbanistes sont évalués. Il produit un iconique en éloignant ses deux mains l’une de l’autre sur un axe horizontal, paumes vers le bas. Les doigts écartés et semi-tendus, Tim représente une forme rectangulaire (a), en donnant une forme concrète au support sur lequel les “redevelopments” sont dessinés. Il maintient cette forme jusqu’à la pause silencieuse intra-constituant de SC, après laquelle ses deux mains sont légèrement levées (b). Sa main droite pivote par un flip très bref, pendant lequel sa paume ne fait plus face à la table mais à son autre main. Les deux mains sont ensuite

posées sur la table dans un battement (c). Ce mouvement donne des informations d'ordre pragmatique, en indiquant à la fois la fin des difficultés de planification de Tim (son discours a jusque-là été ponctué de pauses silencieuses, indiquant une recherche d'éléments lexicaux appropriés par le locuteur) et la complétion du segment discursif : Tim a donné toutes les informations nécessaires à la séquence. Une étape discursive est atteinte. Tom ne manifestant pas son intention de prendre le tour de parole, laissé libre par une longue pause silencieuse, Tim prend position qualitativement sur le contenu de L et de SC sur R sans produire aucun autre geste manuel (d). Ce segment s'apparente alors à une coda, i.e. un bilan conclusif de la séquence.

Notons en revanche que cet indice de faible proportion de chevauchement concerne plus les appositives relatives que les appositives canoniques. Si seulement 8.7 % des appositives relatives sont accompagnées de gestes manuels en chevauchement sur plusieurs groupes intonatifs, 27.3 % des appositives canoniques présentent cette caractéristique. Cela reste malgré tout une configuration minoritaire, et ce pourcentage demeure inférieur à celui des autres types syntaxiques. L'exemple (15) et la figure 6.16 présentent une appositive canonique accompagnée de gestes manuels en chevauchement sur plusieurs groupes intonatifs.

- | | | | |
|------|-------|----|---|
| (15) | Elena | | the best salads ever is Galician salads # |
| | Joey | | what are they |
| | | | what are they |
| | Elena | | they're from # |
| | | L | [(a) you know my d-] [(b) where my dad's from # |
| | Joey | | yeah |
| | Elena | SC | Galicia in] [(c) North West] # |
| | | | of Spain # |
| | Joey | | mh |
| | Elena | R | that's like # |
| | | | [(d) they just #] |

Elena a recours à une information accessible à Joey (“where’s my dad from”) pour expliquer à cette dernière ce que sont les salades galiciennes. L’élément de départ (i.e. l’antécédent L) est l’information la plus susceptible d’être connue par Joey, et implique indirectement Elena par sa relation filiale. Le début de L voit la réalisation du geste référentiel qui va caractériser les deux éléments de l’apposition : Elena forme un triangle avec ses doigts sur la table (a). Le regard d’Elena sur le triangle qu’elle forme prend une valeur déictique et centre l’attention de Joey sur la description. C’est aussi le travail du marqueur discursif “you know” dans la modalité verbale, qui demande l’attention de Joey tout en préservant la face conversationnelle de cette dernière (Elena mobilise une information qui a déjà été présentée à Joey, mais la fausse certitude établie par “you know” évacue toute mention de po-



FIGURE 6.16 – Geste déictique (b) produit en chevauchement sur les groupes intonatifs de L et de SC dans la séquence (15), suivi d’un second déictique (c) également produit sur SC. L’extrait se termine par un métaphorique (d) alors que la locutrice initie la description du contenu des salades galiciennes.

tentiel oublié de la part de Joey, à valeur plutôt négative). Le triangle formé par les deux mains d’Elena est déplacé plus haut, quittant la table pour être tenu au centre de l’espace interactionnel, les paumes et le regard d’Elena tournés vers Joey (b). La promotion spatiale de ce geste rassemble L et le début de SC en une seule unité cognitive (Enfield, 2009). Cette unité fait appel aux connaissances partagées avec “Galicia”, mentionnée dans le co-texte avec les “Galician salads”. En revanche, le début du complément informationnel de Galicia (“North West”) témoigne d’un changement de point de vue manuel : alors qu’Elena décrivait la localisation de cette région dans un coin de l’Espagne, Elena délimite désormais les contours de la région, les deux index quittant la configuration de triangle pour décrire deux arcs de cercles symétriques autour d’un espace où se trouve l’entité désormais non-matérialisée (c). Les deux gestes (b) et (c) sont semblables aux traits sémantiques verbaux des appositions canoniques en présentant deux réalisations possibles d’une même entité. Le geste métaphorique (d) à la suite de R annonce un mouvement textuel dans la description. L’étape de l’évocation de l’origine de ces salades est terminée et Elena commence à décrire leur contenu. Ce contenu se veut précis : les gestes de forme pyramidale dénotent la concentration d’Elena sur un autre élément constitutif des salades, afin de fournir une réponse pointue (Calbris, 2011) à la question de Joey. Après l’abandon d’une première structure verbale, Elena quitte Joey du regard alors qu’elle réfléchit à l’élément de départ le plus adéquat.

Un second indice de rupture concerne les gestes manuels produits en co-occurrence avec les appositives. Seulement 7.3 % de ces gestes manuels participent à la construction d’un *catchment* dans le discours. Cela représente significativement moins que les relatives déterminatives (25.5 %; $F(54, 54) = 2.81, p < 0.001$), et que les cir-

constanciennes (12.7 %; $F(54, 54) = 1.65, p < 0.05$). L'exemple (16)¹⁰ et la figure 6.17 présentent une appositive co-produite avec un geste manuel ne participant pas à la construction du *catchment* en cours. Les gestes (a), (b), (c), et (d) sont chacun documentés par deux images, afin de mieux décrire la trajectoire de leur mouvement.

- | | | |
|------|-----|---|
| (16) | Tom | but if you do a course
where you can [(a) # just # gradually go through #]
like # from the very beginning
the basics |
| | Tim | NODS |
| | Tom | [(b) right through to really kind of spi]ritual stuff #
it kind of # |
| | L | [(c) you know it's kind of a progressive st-] phase |
| | SC | which is what you need [(d) rather than #
like # just someone preaching at you#] |
| | R | [(e) and the chance to talk] |

Tom vante les mérites de son groupe de parole à Tim, qui vient d'assister à l'une des séances. Il développe l'argument selon lequel le groupe qu'il fréquente propose une progression graduelle dans la découverte. Avant la verbalisation de cet apprentissage progressif, Tom réalise un premier métaphorique de la main droite (a), d'abord posée paume contre table à gauche de son espace. Il déplace sa main dans une trajectoire oblique jusqu'au bord de la table à droite de son espace de gestualisation. Au long de cette trajectoire oblique, trois battements sur la table marquent trois étapes. Après que Tom a formulé un exemple plus précis à cette idée ("like # from the very beginning the basics"), Tim réalise un *backchannel* minimal gestuel, en hochant la tête (nod). Ce *backchannel* très peu intrusif est sûrement en relation avec le contour intonatif final montant de Tom sur "the basics", indiquant clairement une suite. Tim montre son alignement à la séquence en cours en encourageant sa progression (Stivers, 2008). Tom produit un second métaphorique de la même main, toujours paume contre table, reprenant la même trajectoire (i.e. de gauche à droite au bord de la table). Alors que le premier métaphorique est produit en co-occurrence avec la prédicat et l'adverbe exprimant l'action de progresser graduellement ("gradually go through"), le second métaphorique, pourtant semblable, est produit avec l'expression d'une borne à ce procès ("right through to really kind of spiritual stuff"). Le contour intonatif final sur "stuff" et la longue pause silencieuse qui s'ensuit suggère que l'endroit est choisi par Tom pour constituer une *TRP*. En revanche, Tim ne montre aucune intention de prendre le tour. Tom garde donc la parole et baisse le regard en se concentrant sur le prochain mouvement discursif de son argumentation. Après un abandon de structure ("it kind of"), un *butterworth* (ou geste de

10. Dans cet exemple, l'activité gestuelle du co-locuteur représente un *backchannel* et est décrite en majuscules.



FIGURE 6.17 – Série de métaphoriques aux coordonnées, configurations, et trajectoires récurrentes dans la séquence (16). Le métaphorique initié sur SC (d) établit un contraste parmi ces régularités.

recherche¹¹) l’aidant à organiser sa production et une autre pause silencieuse, Tom semble avoir eu le temps de projeter les prochaines unités de discours. Il centre l’attention de Tim sur le segment de discours à venir avec le marqueur de discours “you know” et produit un troisième métaphorique (c), en co-occurrence avec L. Ce métaphorique reprend la même trajectoire dans l’espace de gestualisation de Tom, servant de fil rouge à son argumentation. En revanche, le mouvement est effectué avec la tranche de la main de Tom, paume sur le côté et pouce vers le haut. SC représente un commentaire qualitatif, servant de manière plus directe l’argumentation de Tom. Le geste initié sur SC (d) constitue une rupture dans le sens où il

11. Notre traduction; geste désorganisé réalisé lors de recherches lexicales (portant le nom du psycholinguiste britannique Brian Butterworth).

n'exploite ni l'axe horizontal gauche-droite, ni la valeur référentielle mobilisés auparavant. Il se penche en avant en passant les mains autour de la bouteille d'eau sur la table, et garde les paumes ouvertes alors qu'il les oriente davantage vers le haut par un flip symétrique. La combinaison entre le changement de posture et l'ouverture des deux paumes permet à Tom d'entrer davantage dans l'espace interlocutif (Cibulka, 2015). Tom sort du premier niveau de son argumentation, i.e. la description du processus d'apprentissage voulue "objective", pour donner directement son avis. Ce métaphorique apporte de la modalité appréciative en même temps que de l'emphase avec le flip appuyant son propos. Le retour dans l'interlocution se voit aussi avec la hausse de F0, et avec le retour du pronom "you" (deux occurrences sur SC). Tom ne reprend ni la description du procédé d'apprentissage, ni les gestes référentiels de gauche à droite sur R. Il garde la même configuration manuelle et produit un second flip (e), annonçant cette fois une nouvelle étape de l'argumentation, avec la mention d'un argument différent ("the chance to talk"). En baissant les yeux sur le côté de la table, Tom est moins centré sur le co-locuteur.

Un phénomène supplémentaire concerne les gestes manuels. En revanche, il n'est pas considéré comme un indice de rupture en lui-même, concernant les propriétés temporelles d'un type de geste manuel dont l'emploi n'est pas caractéristique des appositives. Si les battements manuels sont caractéristiques des relatives déterminatives, nous avons vu qu'ils ne sont qu'en partie réalisés en co-occurrence avec une frontière de groupe intonatif. Ce n'est pas le cas des appositives, qui montrent 66.7 % de ces battements manuels en co-occurrence avec une frontière prosodique. De plus, 58.3 % de ces battements sont réalisés sur la syllabe finale de L, marquant une frontière relativement tôt dans la séquence, immédiatement avant l'appositive. Plus que cette distribution importante sur les frontières, la durée de ces battements manuels est significativement différente sur les frontières prosodiques qu'en milieu de groupe intonatif ($F(11, 11) = 14.5, p < 0.0001$), avec une moyenne de 0.71 secondes aux frontières, et 0.43 secondes ailleurs dans le groupe intonatif. En revanche, la différence de durée n'est pas significative entre les deux emplacements de frontière (i.e. syllabe finale de L, syllabe finale de SC). Cet allongement temporel des battements manuels fait écho aux résultats prosodiques des appositives, où l'allongement syllabique est conséquent sur les deux frontières (i.e. syllabe finale de L, syllabe finale de SC).

Si les indices de rupture concernent surtout les gestes manuels pour les relatives déterminatives et les circonstancielles, la démarcation est également exprimée par d'autres articulateurs lors de la production des appositives. Les locuteurs

montrent une mobilité du regard significativement plus grande lors de la production de séquences contenant une appositive que lors de la production de séquences contenant une relative déterminative $F(54, 54) = 2.94, p < 0.0001$) et que celle contenant une circonstancielle $F(54, 54) = 2.81, p < 0.001$). Seulement 7.3 % des séquences contenant une appositive sont en effet produites sans aucun changement du regard, alors que l'absence de mobilité du regard du locuteur concerne 27 % des relatives déterminatives et 26 % des circonstanciels. Dans l'exemple (17) accompagné de la figure 6.18, Joey décrit à Elena un hameau situé non-loin de Crickhowell, endroit qui s'avère connu des deux locutrices. SC est également un commentaire évaluatif.

- (17) Joey my mum's school
 it's the school i used to go to
 [(a) have this place
 L in #][(b) Gwyn Fechan]
 SC [(c) **which is**] [(d) # **just a place**
 it's] [(e) **not even a village** #
 R hem # which #] [(f) and the nearest town is Crickhowell]



FIGURE 6.18 – Changements de direction du regard de la locutrice Joey au cours de la séquence (17). La séquence est également caractérisée par une série de pointages manuels.

Cet extrait (17) est caractérisé par la mobilité du regard de Joey, chaque groupe intonatif comptant au moins un changement de direction. Joey réalise également

plusieurs gestes manuels, distincts en termes de configuration et de trajectoire. En revanche, ces gestes manuels sont tous des pointages et sont réalisés de la même main. Joey ne regarde pas la locutrice pendant la réalisation du groupe intonatif avant L (a). Elle pose sa main droite paume contre table d'une manière déictique quelques instants avant la mention de "this place". Sur L, elle marque "Gwyn Fechan" d'un battement de tête et lève les yeux vers la co-locutrice (b), indiquant une information majeure d'un point de vue pragmatique. Joey a d'ailleurs eu des difficultés à récupérer cette information lexicalement (précédée d'une pause silencieuse d'hésitation). Joey détourne le regard dès le début de SC (c). La paume droite de Joey, jusque là posée face contre la table, se lève : ses doigts sont encore en contact avec la table. Elle lève également les yeux vers Elena (d). Cette configuration n'est pas immobile : Joey réalise un mouvement circulaire sur la table, délimitant un espace plus large autour du point d'ancrage où elle a initié le mouvement. Ce pointage en corrélation avec "just a place" modifie la localisation d'un référent donné, puisque Joey délimite un espace plus large, aux contours plus flous. Ce geste prend des dimensions modales et métadiscursives. Il tourne autour d'une entité spécifique dont la référence lexicale est également commentée. En délimitant un espace approximatif pour le référent, Joey exprime une réserve quant à son exactitude. C'est un marqueur d'approximation. Elle produit également un geste de négation avec la tête (d) et fronce les sourcils en produisant le geste manuel circulaire. La valeur modale négative apportée par les mouvements de tête et de sourcils insiste sur l'assertion de la négation subséquente "not even a village". La co-occurrence du geste manuel et du mouvement de tête permet deux prises de position modales de nature différente. La première est à propos de la localisation de "Gwyn Fechan" en tant que référent, la seconde est à propos de sa pertinence dans la séquence discursive. Le groupe intonatif "it's not even a village" suivant le groupe intonatif de l'appositive ("which is just a place") est inclus dans le geste manuel circulaire, bien qu'il représente une proposition indépendante du point de vue grammatical. Joey a en revanche quitté du regard la co-locutrice pour produire ce second groupe intonatif (e). La tentative d'enclencher une nouvelle prédication sur R par le biais du pivot "which" est soldée d'un abandon et d'un recours au marqueur discursif "and". Le marqueur "and" marque la reprise d'un élément référentiel activé non pas sur L, mais dans les énoncés précédents : "Crickhowell" est le topique de la séquence. Joey lève à nouveau les yeux vers Elena (f), et produit l'intégralité de R en maintenant son regard vers la co-locutrice. Elle produit également un pointage manuel accompagné d'un mouvement de tête d'acquiescement. Bien qu'il soit produit avec la même main, le pointage est différent et beaucoup plus précis : un seul doigt est en contact avec la table. Par ce pointage de récupération (Joey vient de montrer des

difficultés de projection/calcul après SC par des pauses et des abandons lexicaux) et son mouvement de tête d’acquiescement, Joey met en focus Crickhowell. La succession d’un mouvement de tête de négation sur SC à un mouvement d’acquiescement sur R indique la coexistence de deux informations, différentes par la nature de leur pertinence au topique de la séquence. L’objectif de cette séquence était d’augmenter les connaissances partagées en évoquant un autre endroit situé non loin de ce lieu. Joey a procédé en spirale, partant de Crickhowell pour prédiquer à propos d’un lieu à côté, puis revenant à Crickhowell par un lien différent. À la fin de R, Joey a donc complété sa boucle tout en enrichissant les connaissances partagées.

Ce dernier exemple (17) montre que la mobilité du regard participe à la démarcation des appositives relatives et de leur contenu. En revanche, le fait que la locutrice lève les yeux vers la co-locutrice de nombreuses fois au cours de la séquence est indicateur de divers phénomènes, liés à la démarcation discursive, mais aussi à la structure informationnelle (e.g. emphase), et à la gestion de l’interaction (e.g. gestion du tour de parole). S’il a été montré que les locuteurs lèvent les yeux vers les co-locuteurs aux frontières prosodiques et/ou grammaticales pour un retour sur la façon dont les tones-units sont reçus (Argyle et Cook, 1976; Vertegaal et al., 2000; Cassell et al., 2001; Barkhuysen et al., 2008; Oben et Brône, 2016), les regards vers le co-locuteur sont également utilisés pour leur fonction emphatique (Argyle et Cook, 1976; Goodwin, 1981; Cassell et al., 1999; Rossano, 2012). Ces changements de la direction du regard dépendent aussi du type d’action discursive en cours (Cassell et al., 2001; Rossano, 2013). Une granularité plus fine est donc nécessaire pour distinguer les différentes fonctions du regard au cours des séquences contenant une appositive, en différenciant les changements de regard vers le co-locuteur en milieu de groupe intonatif et ceux en fin de groupe intonatif. Le tableau 6.1 permet de dresser le comportement prototypique d’un locuteur durant les séquences comprenant au moins un changement, et de déterminer les positions de ces différents changements de direction.

	milieu de groupe intonatif			fin de groupe intonatif		
	L	SC	R	L	SC	R
ailleurs que le co-p	2	11	8	2	9	5
vers le co-p	2	5	1	21	20	6

TABLE 6.1 – Distribution par nombre d’occurrences des changements de regard (ailleurs que le co-participant; vers le co-participant) dans les séquences comprenant une appositive.

Les changements de direction du regard les plus nombreux se produisent en fin

du groupe intonatif¹² de L. Durant le groupe intonatif de SC, il n'est pas rare que le locuteur détourne le regard pour l'élaboration de ce segment. Ce résultat fait écho aux travaux de Beattie (1978) et de Streeck (2014). Enfin, les locuteurs lèvent à nouveau les yeux vers les co-locuteurs en fin du groupe intonatif de SC dans presque autant d'occurrences qu'en fin de L. Si ces changements vers le co-locuteur en fin de groupe intonatif peuvent marquer les frontières (d'unités grammaticales et/ou d'unités prosodiques), ils peuvent également établir un focus informationnel, puisque que la plupart d'entre eux se produisent en co-occurrence avec les derniers items lexicaux des groupes intonatifs. Les derniers items lexicaux sont prototypiquement les plus informatifs (Halliday, 1985), contenant l'information nouvelle. Rien ne nous permet donc de différencier le marquage de ces deux phénomènes. La durée de ces unités de regard vers le co-locuteur n'a pas été établie comme significativement différente ou discriminante de l'un ou l'autre phénomène par la littérature. Toutefois, la démarcation et le focus ne sont ni incompatibles ni antithétiques. Nous considérons que le simple fait que les changements de regards soient co-occurents avec les frontières droites et gauches des groupes intonatifs participe déjà en lui-même à une rupture sur le plan linéaire.

Les appositives se distinguent également par leur proportion de haussements de sourcils, à la fois par rapport à leur co-texte (L : $p > 0.5$; R : $F(54, 54) = 2.13$, $p < 0.005$) et par rapport aux autres subordinées (circonstancielles et relatives déterminatives : $F(54, 54) = 1.59$, $p < 0.05$). En revanche, cette grande proportion de haussements concerne beaucoup plus les appositives relatives (66.7 %) que les canoniques (33.3 %), et est donc liée en grande partie à la prise de position modale. Nous avons en effet vu dans le Chapitre 3 que les appositives relatives sont propices à l'insertion de modalité et aux changements de position énonciative. Les séquences (18) et (19), illustrées par la figure 6.19 et la figure 6.20 comprennent respectivement une appositive relative et une appositive canonique, toutes deux caractérisées par un haussement de sourcils.

- (18) Beth L and then we went into # a place called Tropi[(a)cana]
 SC **which was [(b) horrible] (laughs) #**
 Kate where is [(c) that
 Beth R it's on Saint Mary Street
 near the castle]

Beth raconte ici ses activités du weekend précédent. Elle produit un battement

12. La mention "fin de groupe intonatif" désigne ici les changements de direction de regard se trouvant en chevauchement avec une frontière prosodique (i.e. la frontière entre syllabe finale de L/syllabe initiale de SC; syllabe finale de SC/syllabe initiale de R).

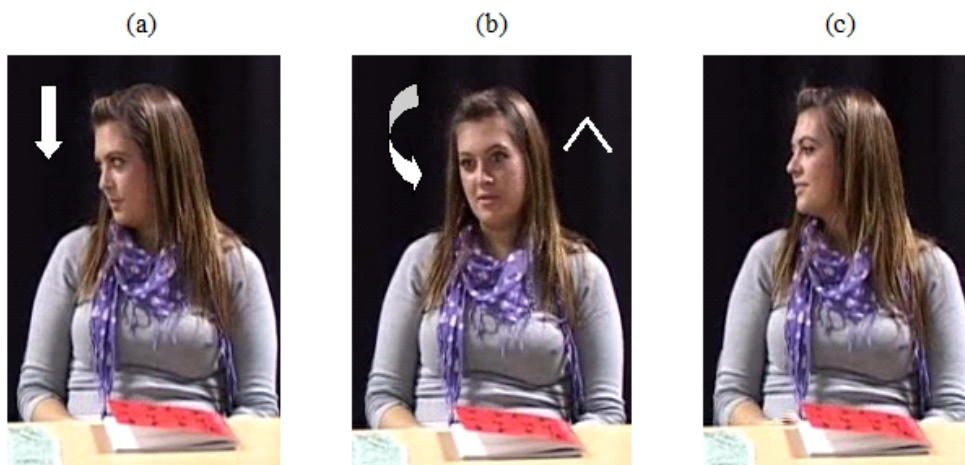


FIGURE 6.19 – Mouvements de tête et haussement de sourcils dans la séquence (18).

de tête sur l'élément rhématique de L (a). L est une proposition narrative qui ajoute un événement au récit de Beth. SC et R, à l'inverse, prédisent sur le même point de départ informationnel et n'ajoutent pas d'événement à la structure narrative. Avec l'ajout de modalité appréciative sur SC à l'élément rhématique de L (qualifié de "horrible"), Beth réalise un large mouvement de tête, ainsi qu'un haussement de sourcils (b). En tournant la tête sur le côté et en terminant par un battement, Beth exprime la totalité, qui intensifie son propos (McClave, 2000). Ce mouvement ajoute de la modalité assertive à l'intensité exprimée par l'emploi de l'adjectif hyperbolique "horrible", exprimant le degré le plus catégorique sur une échelle d'assertion. Beth lève à nouveau les yeux vers Kate après la demande de précision de cette dernière, et réalise R sans aucun mouvement de sourcil ni de tête (c).

La séquence (19) montre qu'un haussement de sourcils peut également être employé dans l'expression d'un autre type de modalité, à un degré d'engagement très différent.

- (19) Joey but the thing is
 [(a) the guy packed the bag
 L and hem] # [(b) he
 SC **like the Sainsbury's man**] [(c) (laughs)]
 R and hem # [(d) the pizza bases have got a crack in them]

Après avoir énoncé le premier événement de son récit (le caissier a mis les articles dans un sac) et réalisé un geste iconique (a) du point de vue du personnage (Joey prend le rôle du caissier en mimant l'action de ranger les articles dans un sac), Joey n'enchaîne pas directement avec l'événement majeur (en rangeant les articles,



FIGURE 6.20 – Haussement de sourcils (c) et différents pointages manuels dans la séquence (19).

il a fendu les pâtes à pizza toutes prêtes). Elle anticipe une question potentielle de son interlocutrice par un phénomène de *self-monitoring* (Clark et Krych, 2004, p. 64) : Elena pourrait trouver que la pertinence de “the guy” n’est pas optimale. Joey s’écarte alors de la réalisation du but principal de la séquence (Joey ne pourra peut-être pas utiliser les pâtes à pizza qu’elle a achetées) en repérant “the guy” par rapport à une autre forme lexicale de surface, “the Sainsbury’s man”. SC réalise alors un processus d’ancrage informationnel et de stabilisation référentielle pour l’élément thématique de la séquence nouvellement présenté sur L. “The Sainsbury’s man” est presque introduit comme une réparation par le marqueur “like”, corrélé à un geste métaphorique de la main droite (b). Paume vers la co-locutrice, Joey agite la main, suggérant à la fois l’approximation et le désengagement par rapport à son assertion précédente (Calbris, 1985), tout en haussant les sourcils. Ces traits indiquent que SC sert également à rendre la désignation plus polie, et à atténuer (*hedging*) le côté informel de la première réalisation lexicale. Elle fait suivre les termes apposés par un rire, contribuant lui aussi à faciliter la production de la séquence en lui donnant un ton léger. Bien que SC soit réalisé sous la forme d’une appositive canonique, Joey insère de la modalité dans son récit en prenant position sur les termes qu’elle vient d’utiliser, et sur leur potentiel effet dans l’interaction. Les sourcils toujours levés (c), elle réalise un adaptateur en se touchant le bout du nez. Cet adaptateur exprime lui aussi sa gêne toute relative, et signale en même temps une transition dans le discours : Joey garde son tour et s’apprête à continuer son récit. Après le marqueur de reprise “and” combiné à une pause remplie, Joey reprend son récit, et décrit les pizzas à l’aide d’un iconique, cette fois-ci d’un point de vue observateur¹³

13. McNeill (1992; 2005) évoque deux points de vue couramment utilisés par les locuteurs lorsqu’ils produisent des gestes iconiques. Le point de vue du personnage implique que le locuteur mime l’action ou l’entité décrite avec une partie ou l’intégralité de son corps, prenant alors un rôle de

en traçant les contours de la pâte (d).

Au premier abord, le fait que les haussements de sourcils apparaissent sur les appositives exprimant une position modale laisse penser que les mouvements de sourcils sont davantage liés à l'emphase (Cavé et al., 1996; House et al., 2001) plutôt qu'à la délimitation des unités prosodiques et/ou syntaxiques. Ce constat est fait par Granström et al. (1999), qui estiment que la relation entre les haussements de sourcils en tant qu'indices de frontière et ceux en tant qu'indices de proéminence est problématique. Il est en effet selon eux difficile de trancher entre les deux phénomènes dans les études de production, et l'étude de perception de ces derniers suggère qu'un haussement de sourcils en tant qu'indice de frontière n'est pas efficace en tant qu'indice indépendant dans le contexte d'une proposition ambiguë. Dans notre corpus, notons que les haussements de sourcils ne sont produits en corrélation avec un ton emphatique (i.e. montant-descendant) que dans seulement 1.8 % des appositives. Comme nous l'avons remarqué pour les changements de direction du regard, les phénomènes de démarcation et d'emphase ne sont à notre sens pas incompatibles, et sont chacun informatifs, sur deux modes d'organisation différents du discours. Le simple fait que les haussements de sourcils apparaissent beaucoup plus sur les appositives constitue une rupture avec le co-texte sur le plan linéaire. De plus, ces haussements de sourcils n'apparaissent pas n'importe où dans les groupes intonatifs. 68.8 % des haussements de sourcils dans les séquences contenant des appositives sont produits en chevauchement temporel avec une frontière prosodique (aussi bien sur la syllabe finale du groupes intonatifs de L que celle de SC). Ces haussements sont d'une durée moyenne de 0.83 secondes. Aucune différence de durée significative ne s'observe cependant entre les haussements dans les groupes intonatifs, ceux en chevauchement avec la syllabe finale de L, et ceux en chevauchement avec la syllabe finale de SC. Cavé et al. (1996) remarquent également une stabilité dans la durée des haussements de sourcils, alors que leur étude montre en même temps de grandes différences d'amplitude de mouvement. Ekman et Friesen (1969) et Ekman (1989) lient ce phénomène à celui de la dichotomie entre durée et amplitude des mouvements kinésiques.

Un dernier indice mérite d'être mentionné concernant les appositives. En revanche, cet indice n'atteint pas la significativité et représente donc une tendance.

personnage dans la situation qu'il décrit. Le point de vue de l'observateur implique que le locuteur ne mime pas directement l'action ou l'entité décrite, mais en représente ou décrit certains traits avec une partie de son corps. La recherche de McNeill suggère que les gestes réalisés du point de vue du personnage sont associés aux verbes transitifs, et ceux du point de vue de l'observateur aux verbes intransitifs (McNeill et al., 2007). Nous nous intéressons particulièrement aux changements de points de vue dans ce chapitre, comme dans l'exemple (19).

Comparé aux autres types syntaxiques, les séquences contenant une appositive sont les plus propices aux changements de main lors de la réalisation de gestes manuels. 5.5 % des séquences contenant une appositive comprennent des gestes manuels successifs réalisés avec une main différente, alors que cette caractéristique concerne 1.8 % des séquences contenant une circonstancielle et n'apparaît pas dans les séquences contenant une relative déterminative. Dans 90 % des cas, le changement de main s'effectue entre le dernier geste manuel produit sur l'appositive et le premier geste produit sur R. Notons aussi que cet indice s'observe exclusivement dans les séquences contenant une appositive relative. Les appositives canoniques ne sont donc pas concernées. Alex produit l'exemple (20)¹⁴ dans un récit à propos des trains Eurostar.

- (20) Alex i got my # hem return [(a) journey
 L for fifty pounds #]
 SC **which was [(b) really amazing] #**
 Rhianna [NODS yeah it's not bad]
 Alex R [(c) but then that was] #
 it was really fun-
 uh i- ironically
 it was the weekend
 when they had all the snow on the lines

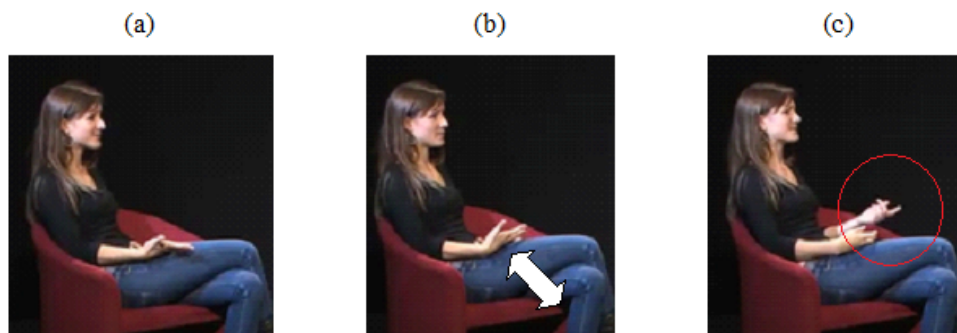


FIGURE 6.21 – Changement de main se produisant entre SC et R pour réaliser les gestes (b) et (c) dans la séquence (20).

Dans la figure 6.21, Alex réalise d'abord un geste métaphorique en posant la main droite paume ouverte sur ses cuisses (a). Ce geste est emphatique et ajoute de la modalité à son discours, la paume ouverte représentant un appel à la co-locutrice au même titre que le contour final montant sur le groupe intonatif de L. Alex greffe

14. Dans cet exemple, l'activité gestuelle de la co-locutrice représente un *backchannel* et est décrite en majuscules.

un battement sur cette configuration en fin de SC (b), en co-occurrence avec la prise de position énonciative très forte réalisée par “really amazing”. Bien que le contour intonatif de SC soit descendant, l’appel à la co-locutrice a bien été intensifié par le battement : Rhianna insère un *backchannel* affiliatif¹⁵ au point de vue exprimé par Alex. Alex change aussitôt de main en récupérant le tour de parole, et réalise sur R un pointage paume vers le haut, index vers la gauche (c). Ce pointage et ce changement de main sont très comparables au rôle de la combinaison des marqueurs de discours “but then”. “Then” fonctionne comme un déictique pour établir un point de référence temporel à la manière du pointage, alors que “but” lie les états de faits précédents à un segment imminent de manière adversative, à la manière du changement de main. Par des moyens lexico-discursifs et des moyens gestuels, Alex introduit la mise en parallèle d’un nouvel élément pivot dans le récit. Le changement a lieu au niveau du discours, où une nouvelle étape discursive a modifié la progression du récit (Enfield, 2004; 2009). Le pointage est également important au niveau de la gestion de l’interaction, en indiquant que le tour d’Alex n’est pas fini et que de nouvelles informations vont être apportées sur le topique auquel la co-locutrice vient de manifester son intérêt.

En résumé, les appositives présentent une rupture beaucoup plus perceptible que celle des autres types syntaxiques de subordinées, et sont encodées gestuellement comme des unités isolées. Plus que le fait qu’elles comportent quatre indices de rupture, leur vraie spécificité par rapport aux autres types syntaxiques réside dans le fait que ces indices soient répartis sur plusieurs articulateurs. En revanche, certains indices de rupture sont plus fréquents sur les appositives relatives que sur les appositives canoniques. Les indices concernés sont les changements de main et les haussements de sourcils. Aucune différence de longueur n’est significative entre ces deux sous-types sur le plan prosodique. Le temps de production des constructions ne rentre donc pas en compte. Ces deux constructions apparaissent également à proportion très égale dans les différents types de séquences discursives. À notre sens, la disparité relative aux changements de mains est due au positionnement séquentiel des constructions au sein du discours. Nous avons vu que les gestes co-verbaux successifs de main différente représentent l’une des ressources employées par le locuteur pour annoncer une transition structurelle dans la séquence discursive en cours. Ces changements de main apparaissent la plupart du temps immédiatement après la production de l’appositive relative. En tant que brefs résumés évaluatifs, les appositives relatives sont par ailleurs très utiles et fréquentes en fin d’unité discursive, à

15. Le co-participant montre son affiliation interactionnelle en adoptant la position modale attendue par le locuteur et/ou en reconnaissant l’importance de la nouvelle information présentée (Stivers, 2008).

la manière d'une coda positionnée avant le passage à une autre étape du discours. Les appositives canoniques montrent un placement séquentiel différent. Nous avons vu dans le Chapitre 3 qu'elles sont davantage utilisées en position macrosyntaxique médiane, dans la localité de l'item référentiel qu'elles modifient. Dans notre corpus, leur position au sein des séquences plus larges comme les récits ou les argumentations est très variable. Elles n'assument alors pas de rôle préférentiel quant à l'ouverture, le bon déroulement, ou la clôture des séquences discursives. En ce qui concerne les haussements de sourcils, la disparité semble davantage liée aux fonctions discursives de ces deux constructions, et à la gestion de l'espace d'échange interlocutif. Si l'ensemble des appositives introduisent de la variation dans une séquence homogène et permettent des glissements de point de vue, appositives relatives et appositives canoniques témoignent d'une orientation différente. Nous avons mis en lumière dans le Chapitre 3 la capacité des appositives relatives à véhiculer du contenu modal, où le locuteur prend position sur les relations exprimées par le co-texte. L'émphase est alors un phénomène relativement commun dans ce sous-type syntaxique. Les appositives canoniques, quant à elles, sont davantage représentatives du phénomène de *self-monitoring* de la part du locuteur, évaluant et modifiant une partie très restreinte de son discours en temps réel. Le fait que les haussements de sourcils ne soient pas autant utilisés dans les canoniques que dans les relatives suggère que les haussements sont davantage exploités pour des fonctions emphatiques et/ou modales (Ferré, 2014) que pour des fonctions d'auto-réparation et des prises de position métadiscursives. Bien qu'une appositive canonique puisse être chargée de contenu modal (par exemple en modifiant un référent par une autre forme lexicale de surface en décalage à des fins humoristiques), le locuteur est majoritairement centré sur son discours. Le co-locuteur montre d'ailleurs surtout son alignement à la séquence en cours plutôt qu'une affiliation au point de vue du locuteur. De même, bien que certaines appositives relatives de notre corpus représentent des auto-corrections métadiscursives ou pré-emptives, le locuteur est davantage centré sur le co-participant, et ce dernier montre davantage son affiliation au point de vue du locuteur que son alignement à la séquence en cours.

Les corrélations entre les différents indices à l'intérieur des appositives sont plutôt nombreuses. La plus forte d'entre elles concerne la caractéristique des appositives à être produites en co-occurrence avec des unités gestuelles distinctes (sans chevauchement sur les autres groupes intonatifs du co-texte), sans que ces gestes ne participent à la création ou au maintien d'un *catchment*. 58.2 % des appositives sont concernées, ce qui représente une proportion relativement importante compte-tenu du fait que 10 des appositives sont produites sans aucun geste manuel. Cette proportion passe à

71.1 % si l'on considère uniquement les appositives produites en co-occurrence avec un ou plusieurs gestes manuels. Dans l'exemple (21)¹⁶ illustré par la figure 6.22, Zoe décrit son année d'assistantat dans une petite ville de Bretagne, et plus particulièrement les inconvénients que cette situation présentait. Les deux seuls gestes manuels produits au sein de la séquence L-SC-R sont localisés sur l'appositive, et sont différents du battement réalisé quelques groupes intonatifs auparavant, à la fois dans leur trajectoire et dans leur configuration.

- | | |
|----------|--|
| (21) Zoe | cause you'd have to spend the
like the entirety the whole [(a) day there
it was just aw][(b) ful
and i went to families |
| | L every weekend] |
| | SC [(c) which was nice] [(d #)] |
| | R but i <u>just #</u> |
| Michelle | <u>did you feel</u> like a right charity case |

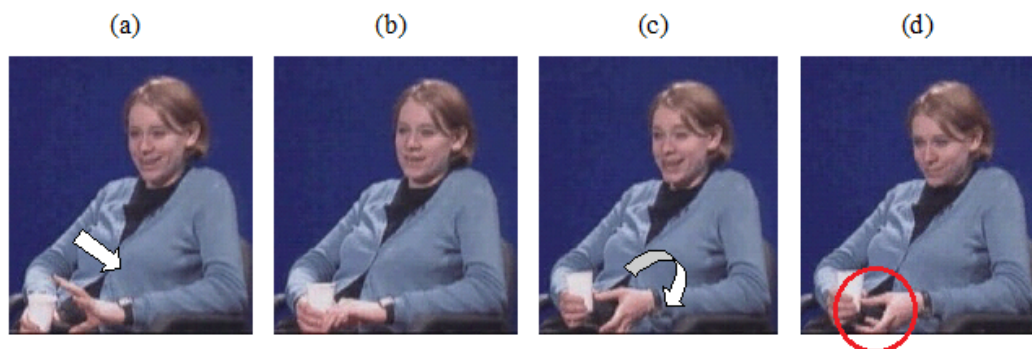


FIGURE 6.22 – Deux gestes métaphoriques (c) et (d) localisés sur SC dans la séquence (21), aux configurations bien distinctes du battement (a) réalisé quelques groupes intonatifs auparavant.

Zoe vient d'expliquer que seulement deux bus par jour faisaient la liaison entre Rennes et la ville où elle résidait. Le dernier battement de la séquence (a) permet d'insister sur l'idée de totalité, mise en avant dans la modalité verbale par la succession d'un adverbe ("entirely") à un adjectif ("whole"), et par un accent emphatique sur l'adverbe dans la modalité vocale. Après ce battement, Zoe retrouve une position de repos qu'elle gardera le long de la production des prochains groupes intonatifs, y compris l'intégralité de L (b). La paume gauche de Zoe est tournée vers le sol alors que ses doigts ne sont pas tendus. Au début de SC, alors que ses doigts s'écartent

16. Dans cette séquence, les éléments soulignés signalent un chevauchement de parole entre les participants.

et que sa main se crispe, Zoe tourne la paume à la verticale, en l’ouvrant vers le côté (c). Ce geste organise le discours précédent en modalisant son énoncé. Plus précisément, cette rotation du poignet marque l’opposition de deux points de vue dans son discours. Alors que L semble contribuer au point de vue principal de la séquence (i.e. l’expérience de vie dans une petite ville de Bretagne comportait beaucoup d’aspects très déplaisants), véhiculé auparavant par “awful”, SC en modifie le point de vue et marque une concession. La pause silencieuse suivante ne représente pas une *TRP* : Zoe crispe ses doigts comme pour saisir un objet (d), et baisse la tête en mimant son déglutissement. Ces mouvements œuvrent à la conservation de son tour de parole. Cependant, alors que Zoe débute R sans produire aucun autre geste manuel, Michelle prend le tour de parole, ignorant les signaux de conservation de tour précédemment réalisés par Zoe.

La seconde combinaison d’indices de rupture la plus importante concernant les appositives lie les deux caractéristiques précédentes (unités gestuelles distinctes pour les gestes manuels, pas de construction ou de contribution à un fil d’imagerie gestuelle pour ces gestes) à la mobilité du regard du locuteur. Cela représente 21.8 % des appositives de notre corpus. L’exemple (22)¹⁷ ainsi que la figure 6.23 illustrent cette combinaison.

- (22) Rhianna L i mean [(a) my mum’s pushing] me to [(b) get my license]
 SC **(h) uh which [(c) i guess i should] #**
 Alex it’s a good thing to have
 Rhianna R (h) [(d) but] #
 Alex if you ever need it i mean
 Rhianna well [(e) first of all
 for the moment
 i’m just not interested enough]

Rhianna explique ici les raisons pour lesquelles elle ne veut pas apprendre à conduire, et la séquence argumentative est centrée sur l’idée selon laquelle Rhianna n’est pas faite pour cette activité. Rhianna commence par évoquer des opinions adverses à la sienne, en leur reconnaissant une certaine valeur. L est un argument à part entière : sa mère voudrait qu’elle passe son permis. Rhianna marque cette information par un geste iconique (a), sa main droite balayant l’espace devant elle, correspondant à l’élément verbal “pushing”. Elle ne regarde pas sa co-locutrice. Ce segment prend une dimension hyperbolique, Rhianna donnant une expression littérale et concrète à la volonté de sa mère, dont les conseils sont matérialisés comme une forte pression. Le regard de Rhianna revient sur Alex alors que le groupe into-

17. Dans cette séquence, les éléments soulignés signalent un chevauchement de parole entre les participants.



FIGURE 6.23 – Changements de direction du regard au long de la séquence (22), également accompagnée de gestes manuels (a) et (c) divers en termes de types et configurations.

natif touche à sa fin (b). Le commentaire de SC n'est pas lié aux dires de sa mère, mais revient sur l'élément rhématique de L, i.e. "get my license". Par un geste d'acquiescement de la tête et un flip de la main droite beaucoup plus bas que le geste de L (c), Rhianna s'exprime à la fois en tant que personnage de la situation construite sur L (la mère de Rhianna l'exhorte, Rhianna acquiesce) et en tant qu'énonciatrice (elle reconnaît la validité des conseils de sa mère et marque cette concession par le flip). Elle hausse également les sourcils, le regard fixé dans le coin gauche. Ce haussement complète sa prise de position modale et marque le caractère contrastif de SC en tant que mouvement discursif. Rhianna revient sur la ligne principale de son argumentation sur R et se montre beaucoup plus catégorique dans ce segment : alors qu'elle acquiesce et semble plier sur SC, elle accompagne ce segment d'un geste de négation continu de la tête, commencé sur "but" (d). Son regard est revenu sur Alex, qui produit un *backchannel* substantiel et spécifique. Alors qu'elle reprend le tour et poursuit sa ligne d'argumentation, elle lève les yeux au ciel alors qu'elle organise ses unités de discours à l'aide de "first of all" (e). Cette séquence est alors

caractérisée par deux positions modales successives, qui ne sont pas de même intensité : l'énonciatrice de R est plus forte que celle de SC. Cette asymétrie reflète la structure discursive, puisque R continue l'agenda séquentiel de Rhianna alors que SC n'y contribue pas.

Dans l'extrait (22) précédent, nous avons mentionné le fait que la locutrice hausse également les sourcils durant la production de l'appositive. Les hausses de sourcils et les changements de direction du regard forment une autre corrélation également importante dans notre corpus, concernant 20 % des appositives.

La dernière combinaison notable d'indices dans notre corpus concerne le fait que 14.5 % des appositives comportent à la fois des gestes manuels distincts, ne participant pas à la construction d'un *catchment*, et présentant un changement d'articulateur (e.g. geste de la main gauche, puis geste de la main droite) au sein de la séquence. L'argumentation de l'exemple (23) menée par Tom à propos d'un groupe de parole comporte ces caractéristiques. Tom commence par citer directement l'une des participantes au groupe de cette année, qu'il lie sur SC à sa propre expérience du groupe. SC est donc ici un pivot discursif et énonciatif, et se compose de deux subordonnées successives.

- (23) Tom i [(a) really enjoy Alpha]
 L i look forward to it every week
 SC **which [(b) is how we were feeling last year**
 when we were on it]
 R we'd get to like # a saturday
 [(c) go and play football together]
 and go
 oh it's only a few days till Alpha

Dans la figure 6.24, Tom réalise d'abord un geste métaphorique emphatique en plaçant les deux mains de manière symétrique sur L (a), configurées en "L" (les paumes se faisant face, pouces et index tendus). Ce premier geste accompagne les propos de la participante de cette année, directement rapportés par Tom. Sur SC, cette configuration est modifiée : elle n'est plus symétrique, la main gauche se déplaçant vers ce côté, la paume un peu plus tournée vers le co-participant (b). La main droite, bien qu'elle tienne la configuration de L, est mise en retrait. Ce mouvement organise le discours de Tim, en désignant trois types de changement : le changement temporel, du présent vers le passé (l'année précédente) et le changement de participant (Tim et ses amis au lieu de la jeune fille de cette année). Si ces deux changements sont plutôt d'ordre grammatical, le troisième changement est d'ordre discursif, puisque Tom a quitté le rôle de personnage pour redevenir énonciateur.

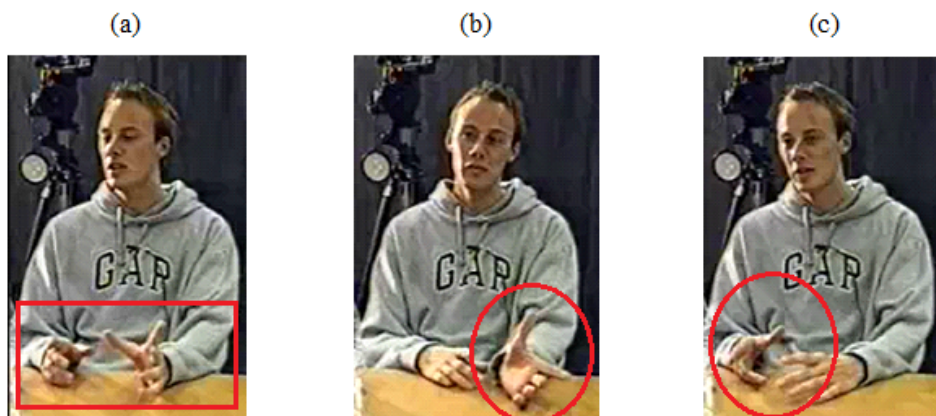


FIGURE 6.24 – Changement de main entre le geste localisé sur SC (b) et le geste suivant (c) au cours de la séquence (23).

La dominance manuelle change sur le prochain geste (c), qui a lieu sur le groupe intonatif suivant R. La main droite de Tom est remise en avant, pointant à droite, et la main gauche est un peu relâchée, la paume se rapprochant de la table. Ce changement a lieu bien que cette fois, ni les participants, ni le temps soient différents de ceux de SC. Le changement prend son sens au niveau séquentiel, où une nouvelle étape discursive a modifié la progression de l'argumentation. Plus que l'avancée d'un nouvel argument au thème principal (le groupe de parole a beaucoup de succès), R commence un nouveau récit, imbriqué dans l'argumentation, qui exemplifie l'argument avancé sur L et renforcé par SC.

6.3 Discussion et conclusion

Comme les ressources liées à la modalité verbale et celles à la modalité vocale, la gestualité permet de distinguer les différents types syntaxiques de subordonnées en fonction de leur degré d'autonomie. Si les relatives déterminatives sont clairement intégrées, le tableau est plus nuancé pour les circonstancielles. Quelques marques de rupture autour des gestes manuels empêchent de considérer les circonstancielles comme intégrées à leur co-texte, mais sont trop peu nombreuses et diversifiées pour réellement parler d'autonomie. Les appositives, quant à elles, comptent assez d'indices pour parler d'autonomie, dont l'expression est répartie sur plusieurs articulateurs.

Les relatives déterminatives sont les constructions les plus intégrées à leur co-

texte gestuellement. Les relatives déterminatives sont fréquemment intégrées à l'unité articulatoire produite avec le co-texte gauche, et ces constructions sont également produites en co-occurrence avec des unités gestuelles dont la configuration et la trajectoire sont très similaires aux gestes manuels environnant. En revanche, si les relatives déterminatives modifient des expressions nominales en affinant l'identification de leurs référents (Langacker, 2008), elles ne présentent pas les indices d'intégration et de continuité gestuels traditionnellement associés au maintien ou à la modification des référents dans le discours. Les configurations bimanuelles (symétriques comme asymétriques) décrites par Frederiksen (2016) liées à la hiérarchisation et à la modification des référents dans le discours ne sont pas courantes dans les relatives déterminatives. De même, les tenues de gestes manuels liées à l'intégration, l'extension, ou la modification d'un segment de discours décrites entre autres par Enfield (2009) sont rares. Le seul indice de rupture qui leur est associé est de nature prosodique. Les battements manuels qui caractérisent les relatives déterminatives sont à même de démarquer l'item lexical le plus pertinent pour une meilleure identification de ces référents. Seulement 50 % de ces battements sont réalisés en co-occurrence avec une frontière prosodique, et ne présentent pas d'allongement significatif par rapport aux battements réalisés en milieu d'unité. Nous avons vu que les battements servent une diversité de fonctions dans le discours et l'interaction, en indiquant l'état cognitif du locuteur et/ou en faisant appel aux connaissances partagées entre les locuteurs. Nous avons constaté dans le Chapitre 4 que si les relatives déterminatives ne présentent pas de frontière entre L et SC, une forte frontière rythmique est marquée entre SC et R (allongement syllabique final et pause silencieuse). Or, la relation des gestes à cette frontière n'est pas marquée avec l'indice que nous attendions. La durée des battements de tête fait écho à l'allongement significatif de la syllabe finale de SC, et participe donc spécifiquement au marquage de cette frontière. Les battements de tête ne sont cependant pas utilisés de manière caractéristique dans les relatives déterminatives.

À l'image des relatives déterminatives, les circonstancielles se démarquent uniquement par le biais des gestes manuels. En revanche, les circonstancielles ne sont pas propices aux répétitions formelles de gestes réalisés précédemment. Les gestes manuels les accompagnant sont plutôt divers en termes de trajectoire et de configuration. Il n'est cependant pas rare qu'un geste manuel soit produit en chevauchement sur plusieurs groupes intonatifs, dont celui correspondant à la circonstancielle. Le fait que le chevauchement soit beaucoup plus réparti entre le co-texte gauche et le co-texte droit exprime la capacité des circonstancielles à ouvrir ou élaborer des liens avec plusieurs unités du co-texte, facilitant la gestion d'espaces à une échelle

supérieure aux groupes intonatifs ou aux propositions syntaxiques. Nous avons vu que les battements de tête dans les séquences comprenant des circonstancielles ne sont pas spécifiquement produits en toute fin de groupe intonatif et ne présentent pas d'allongement temporel. L'absence d'indices gestuels de nature rythmique sur les circonstancielles fait écho aux caractéristiques prosodiques de ce type syntaxique, puisque les circonstancielles sont davantage démarquées par l'intonation que par le rythme. Il est très rare que les circonstancielles présentent plus d'un indice de rupture à la fois. En présentant des critères ne permettant pas leur catégorisation en cas d'intégration nette ni d'autonomie claire, nous concluons que le comportement gestuel des locuteurs lors de la production des circonstancielles œuvre davantage à la construction de zones de consensus et d'échange dans le discours, plutôt que de préciser l'organisation de ce type de subordonnée sur un plan linéaire.

Les appositives, quant à elles, offrent un contraste clair. Elles sont les seules constructions subordonnées clairement autonomisées par le biais de ressources gestuelles. Nous avons relevé une utilisation démarcative de plusieurs articulateurs, ainsi que la production d'unités gestuelles bien distinctes de celles du co-texte. L'autonomie de ces constructions s'observe à la fois sur la qualité et la forme des unités (gestes manuels formellement différents de ceux environnants) que sur leur temporalité (unités des gestes manuels ne franchissant pas les frontières des groupes intonatifs, positionnement d'unités réalisées avec d'autres articulateurs aux frontières prosodiques, allongement de certaines unités gestuelles positionnées aux frontières prosodiques). Comme pour les relatives déterminatives, nous avons constaté qu'un allongement significatif de la durée se produit bien sur certains gestes en chevauchement avec une syllabe finale de groupe intonatif. En revanche, cet allongement ne concerne pas le type de geste dont l'utilisation est spécifique aux séquences contenant une appositive. Par ailleurs, les appositives sont les seules subordonnées présentant l'articulation simultanée de plusieurs indices de rupture dans une seule et même séquence. La densité modale est alors une ressource mobilisée pour l'autonomisation de ce type syntaxique, et rend la rupture plus perceptible. Notons toutefois que certains indices ne montrent pas une distribution uniforme entre appositives relatives et appositives canoniques. La disparité entre les deux sous-types s'observe plus fortement sur la gestualité qu'à travers les caractéristiques prosodiques et verbales.

La démarcation gestuelle s'exprime majoritairement par des ressources correspondant aux gestes manuels. Cependant, les différences entre les trois types syntaxiques quant à la nature et à la distribution des indices suggèrent qu'aucun indice commun n'est systématiquement utilisé pour la démarcation des subordonnées. La

présence des battements (manuels dans les relatives déterminatives, de tête dans les circonstancielles) et des haussements de sourcils souligne en revanche l'importante contribution de l'aspect prosodique de la gestualité. Le rôle du rythme dans la démarcation gestuelle des subordonnées montre d'une part que la gestualité fait écho à la modalité vocale dans la production des subordonnées (nous avons vu dans le Chapitre 4 que les phénomènes rythmiques jouent un rôle essentiel dans la démarcation vocale des subordonnées; Lelandais et Ferré, 2016), et d'autre part que la prosodie n'est pas uniquement exprimée par les phénomènes vocaux (McNeill et al., 2001; Krahmer et Swerts, 2007; Mendoza-Denton et Jannedy, 2011; Esteve-Gibert et Prieto, 2013; Ferré, 2014; Krivokapic et al., 2017). Il est intéressant que ce constat puisse être fait non seulement avec le type syntaxique le plus en rupture, mais aussi avec celui le plus intégré. L'allongement de la durée de certaines unités gestuelles sur les frontières prosodiques va dans ce sens. Il a été prouvé que les effets de la structure prosodique s'étendent au-delà du conduit vocal pour inclure d'autres articulateurs, notamment les mains (Esteve-Gibert et Prieto, 2013; Krivokapic et al., 2017). Or, nous sommes en mesure d'affirmer ici que dans la modalité vocale comme dans la modalité gestuelle, certains mouvements articulatoires s'allongent immédiatement avant une frontière prosodique. De plus, cet allongement n'apparaît pas sur les mêmes frontières pour les relatives déterminatives (i.e. syllabe finale de SC) et pour les appositives (syllabe finale de L et syllabe finale de SC). Néanmoins, ces allongements s'observent à chaque fois sur l'utilisation d'un type de geste non caractéristique de la construction subordonnée en question (i.e. battements de tête pour les relatives déterminatives, alors que ces battements sont peu nombreux, battements manuels pour les appositives, alors que ces battements manuels sont significativement plus nombreux ailleurs).

La plupart des études travaillant sur l'expression vocale et gestuelle de la structure prosodique se focalisent sur le phénomène des proéminences (e.g. Beskow et al., 2006; Krahmer et Swerts, 2007; Al Moubayed et al., 2011), sur un axe davantage thématique que linéaire. Or, le fait que plusieurs types syntaxiques de subordonnées montrent des différences d'usage quant au partage/non-partage des unités gestuelles, offrent des contrastes quant à la forme et la direction des gestes d'une même séquence, et présentent des changements d'articulateurs indique qu'il est également possible de déterminer des indices jouant un rôle majeur dans la segmentation linéaire des unités de discours, sur plusieurs échelles. Notre analyse rejoint les études de Barkhuysen et al. (2008) et de De Kok et al. (2009), qui montrent l'importance de certains indices gestuels dans l'annonce d'une fin d'unité de parole. En revanche, notre analyse diffère de ces travaux en termes de granularité. Alors que

Barkhuysen et al. (2008) et De Kok et al. (2009) travaillent à l'échelle des tours de parole, ce travail a identifié des indices de frontière pour de plus petites unités, spécifiquement les groupes intonatifs contenant des constructions syntaxiques subordonnées.

Nous avons jusqu'ici constaté que l'aspect prosodique des indices de démarcation se reflète à la fois dans les phénomènes gestuels et dans les phénomènes vocaux, et que l'identification d'indices gestuels de segmentation du discours sur le plan linéaire est à la fois faisable et nécessaire. En revanche, la quantité et la distribution de ces indices gestuels sont moindres par rapport à celles des indices vocaux décrits dans le Chapitre 4. Dans les séquences contenant une subordonnée, la rupture est effectuée avec beaucoup plus de moyens vocaux que gestuels, aussi bien en nombre qu'en variété. La différence d'utilisation des modalités s'interprète en deux temps. Le faible nombre d'indices visuels de démarcation ainsi que le fort aspect prosodique parmi ceux qui sont utilisés suggèrent que les locuteurs préfèrent nettement indiquer une disjonction de manière vocale, et que la prosodie joue un rôle essentiel de manière générale dans l'expression multimodale de la rupture. Le nombre et la nature des indices gestuels utilisés pour la démarcation doivent être comparés à ceux des indices gestuels utilisés pour la proéminence et l'expression du premier plan informationnel pour confirmer cette observation. En revanche, la distribution de ces indices (i.e. l'absence de corrélation entre les indices gestuels dans la plupart des constructions) indique soit que les locuteurs accordent plus d'importance à la modalité vocale et attribuent d'emblée un faible rôle aux indices gestuels en complément aux phénomènes vocaux, soit que les locuteurs considèrent que le peu d'indices gestuels utilisés suffisent pour exprimer une démarcation, et que chacun d'eux possède un poids démarcatif important. Le fait que les battements de tête et les mouvements de sourcils n'apparaissent pas ensemble a par exemple été relevé dans d'autres situations de discours par House et al. (2001), qui considèrent que ces deux indices sont suffisamment forts (i.e. ont suffisamment d'impact sur la perception du co-locuteur) pour être employés seuls en tant que marqueurs. Or, rien ne nous permet d'invalidier l'une des deux interprétations, puisque notre analyse a défini des tendances observables pour la production des subordonnées. Les données perceptives recueillies grâce à la mise en place d'un test autour de la démarcation gestuelle nous permettra d'éclaircir cette question. Ce test de perception ne concernera en revanche que les gestes manuels. Il devra plus précisément confirmer que des frontières de nature gestuelle sont perceptibles dans les séquences contenant des subordonnées, que différents degrés de frontière peuvent être perçus, et que la rupture est plus perceptible dans les séquences contenant des appositives relatives. En

comprenant des constructions totalement intégrées à leur co-texte du point de vue vocal, ce test déterminera si les indices gestuels sont aussi fiables dans la perception des frontières que les indices vocaux. Les résultats de ce test de perception feront donc l'objet d'une comparaison avec le test mis en place autour de la rupture prosodique.

Lorsqu'ils sont employés à des fins démarcatives, la plupart des indices visuels donnent également beaucoup d'informations quant à l'organisation pragmatique du contenu (haussements de sourcils, battements, changement de la direction du regard, gestes paume ouverte vers le haut). Alors que ces indices sont utilisés à des fins démarcatives sur un mode d'organisation linéaire, ils remplissent simultanément des fonctions importantes et diverses sur un mode d'organisation thématique. Même en relevant des indices de rupture, nous constatons donc que les subordinées sont marquées par une gestualité d'ajustement (métadiscursif, discursif, représentationnel, interactionnel) et de modélisation pragmatique du discours. Nous sommes pour l'instant en mesure d'affirmer que les subordinées introduisent une rupture lorsqu'elles établissent une position assertive différente de celle du groupe intonatif précédent. Pour éviter un décalage entre l'apport du locuteur et les représentations du colocuteur, la gestualité marque cette rupture, mais est également en mesure d'indiquer la valeur informationnelle de cette rupture dans la séquence discursive en cours. L'analyse du marquage gestuel de l'apport informationnel est le seul moyen de déterminer si cette gestualité d'ajustement est marquée par davantage de ressources que la gestualité démarcative dont il était question dans ce chapitre.

Chapitre 7

Perception des frontières gestuelles

7.1 Introduction

Ce chapitre confronte les résultats de notre analyse gestuelle sur la production des subordonnées aux données perceptives, recueillies grâce à la mise en place d'un test autour de la rupture dans le discours exprimée par des moyens gestuels. Après avoir introduit les enjeux majeurs de notre test, nous proposons un bilan de l'état actuel des études sur le lien entre gestualité et prosodie, ainsi qu'entre gestualité et syntaxe. Nous abordons également les études perceptives sur les frontières discursives et prosodiques exprimées par des moyens gestuels. Nous exposons ensuite notre méthodologie, avant de donner nos résultats et leur interprétation.

À travers l'analyse de nos données sur une étude de production, nous avons vu dans le Chapitre 6 que les locuteurs sont capables d'exprimer une disjonction entre deux constituants de manière gestuelle. La démarcation gestuelle s'exprime majoritairement par des ressources correspondant aux gestes manuels, ainsi que par des changements de configuration formelle et de trajectoire. En revanche, aucun indice commun n'est systématiquement utilisé pour la démarcation des subordonnées en production. De même, la quantité et la distribution de ces indices gestuels sont moindres par rapport à celle des indices vocaux décrits dans le Chapitre 4. Des études de production (e.g. Özyürek et Kita, 1999; Kita et Özyürek, 2003; Özyürek et al., 2005; Vigliocco et Kita, 2006; Kita et al., 2007) ont montré que la gestualité co-verbale est couplée aux choix de découpage syntaxique, notamment celui des propositions. Des études expérimentales (Barkhuysen et al., 2008; De Kok et al., 2009) en gestualité ont également montré que les co-locuteurs utilisent les indices gestuels pour interpréter les fins d'unités de discours. En revanche, les études portant sur les

frontières gestuelles avec une granularité plus fine (par exemple pour marquer des frontières prosodiques ou syntaxiques; Krivokapic et al., 2015; 2017) ne portent pas sur la conversation spontanée et ne permettent pas de savoir si les indices gestuels les plus utilisés par les locuteurs dans un tel contexte sont également les plus utilisés par les co-locuteurs lors de la segmentation du discours.

La plupart des études travaillant sur l'expression vocale et gestuelle de la structure prosodique se focalisent sur le phénomène des proéminences (e.g. Beskow et al., 2006; Krahmer et Swerts, 2007; Al Moubayed et al., 2011), sur un axe davantage thématique que linéaire. Or, le fait que plusieurs types syntaxiques de subordinées montrent des différences d'usage quant au partage/non-partage des unités gestuelles, offrent des contrastes quant à la forme et la direction des gestes d'une même séquence, et présentent des changements d'articulateurs indique qu'il est également possible de déterminer des indices jouant un rôle majeur dans la segmentation linéaire des unités de discours, sur plusieurs échelles. Nos hypothèses sont les suivantes : (1) des frontières exprimées par des moyens gestuels sont perceptibles par des participants naïfs dans les séquences contenant des subordinées; (2) différents degrés de frontière peuvent être perçus; (3) la rupture est plus perceptible dans les séquences contenant des appositives relatives.

Notre test de perception porte uniquement sur les gestes manuels, en raison de la nature des stimuli utilisés. Nous avons vu en production dans le Chapitre 6 p. 240 qu'une unité de geste manuel formellement différente des autres unités gestuelles du co-texte représente un indice de rupture utilisé par les participants. Le test déterminera si cet indice est perceptible comme tel par les co-locuteurs. En comprenant des constructions totalement intégrées à leur co-texte du point de vue vocal, ce test déterminera si les indices gestuels sont aussi fiables dans la perception des frontières que les indices vocaux. Les résultats de ce test de perception font donc l'objet d'une comparaison avec le test mis en place autour de la rupture prosodique.

7.2 Etudes antérieures sur les frontières perçues en gestualité

7.2.1 La notion de frontière en gestualité

La notion de frontière en gestualité découle de l'étude des liens entre les unités gestuelles et les unités de la parole, qui a d'abord été menée sur des données de

production. Il a été montré que les gestes manuels produits en co-occurrence avec la parole sont synchronisés avec cette dernière sémantiquement (e.g. McNeill et Duncan, 2000; McNeill, 2005), temporellement (e.g. Chui, 2005), et structurellement (e.g. Kita et Özyürek, 2003; Lewandowski et Öskaliskan, 2018). En raison de cette coordination, les processus de production de la parole et de la gestualité sont liés à tous les niveaux (e.g. Iverson et Thelen, 1999; Bernardis et Gentilucci, 2006; Roustand, 2012; Ferré, 2014). La théorie du *Growth Point* (McNeill, 1992; McNeill et Duncan, 2010) se base par exemple sur l'observation selon laquelle parole et gestualité co-expriment des unités d'idée, utilisant chacun une forme différente. Alors que la gestualité emploie une forme de représentation globale et synthétique, la parole utilise une forme de représentation analytique et combinatoire (McNeill, 1992). Une interruption dans le processus de production de parole devrait donc avoir un effet sur les gestes, et vice-versa. Dans les sections ci-dessous, nous proposons un aperçu des travaux s'intéressant aux opérations gestuelles de hiérarchisation et de démarcation, en lien avec la prosodie et la syntaxe.

Gestualité et prosodie

Les liens entre gestualité et prosodie ont d'abord été étudiés au niveau de l'interaction. Schefflen (1968), Birdwhistell (1970), McQuown (1971), et Condon (1976) se sont intéressés aux liens entre gestualité et prosodie à un niveau interactionnel, et ont documenté une synchronie entre les pics d'énergie prosodique et gestuelle. Selon Bolinger (1983), la prosodie (particulièrement l'intonation) et la gestualité font partie d'une même forme sous deux apparences différentes, et expriment l'état émotionnel du locuteur d'une manière iconique. Il observe que les deux modalités produisent des mouvements ascendants dans des situations impliquant une tension émotionnelle, et des mouvements descendants dans des situations impliquant une relaxation émotionnelle. Comme Schefflen (1968) et Birdwhistell (1970), Bolinger (1986) suggère alors que la hauteur intonative et les parties du corps évoluent en parallèle.

Duncan (1972) enquête également à un niveau interactionnel sur les façons dont plusieurs moyens, notamment gestuels et intonatifs, sont utilisés dans la prise ou l'abandon de tour. L'étude remarque qu'un arrêt de la gestualisation peut être utilisé pour signaler un abandon de tour, de la même manière qu'un contour final montant ou descendant. Duncan et Fiske (1977) utilisent l'intonation et les gestes pour délimiter des frontières à des unités d'analyse mises au point pour l'analyse des interactions en face-à-face. Ces unités ressemblent aux syntagmes intonatifs (i.e.

les IP, cf. Chapitre 5 p. 176) : elles contiennent un accent nucléaire et une jonction (i.e. un mouvement intonatif) terminale. Les indices gestuels pour les frontières de ces unités sont la mise au repos des mains et des pieds, ainsi qu'une position de tête détournée du co-locuteur. En revanche, les annotations ne reposent sur aucun schéma moderne concernant l'intonation ou la gestualité : l'intonation est codée d'après le modèle de Trager et Smith (1957), comprenant quatre niveaux intonatifs et trois jonctions (i.e. types de tons) terminales. L'annotation gestuelle note les positions générales pour le corps et les mains ainsi que pour les mouvements, plutôt que des unités ou des types gestuels provenant de classifications établies auparavant par la littérature. De plus, l'unité de base sur laquelle les événements sont codés n'est pas le temps, mais la syllabe. Le temps linéaire n'est donc pas pris en compte : chaque événement est codé soit sur une syllabe, soit entre deux syllabes.

À l'échelle du discours, les travaux de Kendon (1972) et de McNeill et al. (2001) sur l'anglais attestent que le groupement d'unités prosodiques et discursives est indiqué par les gestes co-verbaux. Kendon (1980; 1983) propose un alignement entre les hiérarchies intonatives et gestuelles. Tuite (1993) remarque une récurrence rythmique impliquant les deux modalités. Comme d'autres études (Grosz et Hirschberg, 1992; Hirschberg et Nakatani, 1996), McNeill (1992; 2005) et Quek et al. (2000; 2002) notent des corrélations entre les segments de discours, l'intonation, et la gestualité. Afin d'identifier les segments de discours, les chercheurs ont annoté la prosodie à l'aide du schéma d'annotation ToBI (Silverman et al., 1992; cf. Chapitre 5 p. 182). McNeill et al. (2001) remarquent que chaque *catchment*, i.e. chaque groupe de gestes manuels présentant la même récurrence formelle (cf. Chapitre 6 p. 213), possède un ton de frontière distinct. McNeill (2005; 2017) montre par ailleurs que les battements et les déictiques marquent les différentes étapes structurelles d'une narration et fournissent des balises au co-locuteur quant à l'organisation et la structure du discours. Plus spécifiquement, Barkhuysen et al. (2008) exposent que les locuteurs du néerlandais ont tendance à détourner le regard et la tête du co-locuteur au milieu de la production d'un syntagme intonatif (i.e. les IP, cf. Chapitre 5 p. 176), et se tournent vers le co-locuteur en fin d'unité prosodique. Ces mêmes locuteurs montrent également davantage de clignements des yeux, de battements de tête, et d'acquiescements en fin de syntagme qu'ailleurs. De même, Ishi et al. (2014) trouvent dans une étude de parole spontanée en japonais que 80% des acquiescements sont réalisés pendant la syllabe finale du syntagme prosodique.

Les résultats de McClave (1991) précisent le rapport entre structure prosodique et gestualité. Contrairement à Bolinger, elle ne trouve pas de corrélation significa-

tive appuyant l'idée que gestualité et hauteur intonative évoluent dans des directions parallèles. En accord avec Kendon (1983), elle trouve cependant que les phrases gestuelles¹ sont alignées sur les frontières des groupes intonatifs. Dans presque tous les cas de son étude, les phrases gestuelles sont délimitées par les frontières des groupes intonatifs. Dans les rares cas où un geste s'étend sur plusieurs groupes intonatifs, les groupes intonatifs concernés ont des schémas accentuels et intonatifs parallèles.

En ce qui concerne les unités plus petites que les séries de gestes aux mêmes propriétés formelles ou que les syntagmes intonatifs, Kendon (1980) montre que la réalisation du geste (i.e. la partie dynamique du geste, après la phase de préparation et avant celle de rétraction; Kendon, 2004) est exécutée soit avant l'accent nucléaire du groupe intonatif avec lequel le geste apparaît en co-occurrence, soit juste à son attaque. Loehr (2004) reprend cette idée et montre plus précisément que les apogées des gestes (i.e. le déplacement maximum des articulateurs lors d'un geste par rapport à la position d'origine) sont alignées sur les accents nucléaires (également observé par Mendoza-Denton et Jannedy, 2011), et que les phrases gestuelles sont alignées sur les syntagmes intermédiaires. L'étude de Roustand (2012) sur les gestes manuels de pointage confirment ces résultats : la règle de synchronie phonologique, selon laquelle l'apogée d'un geste précède ou est synchrone avec le pic phonologique de l'unité de parole en co-occurrence, est majoritairement respectée.

Nous avons vu dans le Chapitre 4 qu'en anglais, les propriétés temporelles jouent un rôle important en plus des propriétés tonales dans le groupement des unités prosodiques en phrases (Barkhuysen et al., 2008, Breen et al., 2010; Krivokapic, 2014; Krivokapic et al., 2017), ainsi que dans leur proéminence (e.g. McNeill, 1992; Loehr, 2004; Mendoza-Denton et Jannedy, 2011 pour l'anglais; Swerts et Krahmer, 2010 en néerlandais; Rochet-Capellan et al., 2008 pour le portugais brésilien; Esteve-Gibert et Prieto, 2013 pour le catalan). Les travaux en phonétique articulatoire permettent de mettre en lumière l'action des gestes oraux (i.e. les mouvements des différents articulateurs du conduit vocal) dans la formation de la structure prosodique, et plus

1. Les phrases gestuelles, proposées par Kendon (2004), correspondent à l'unité de mouvement à partir du départ de la position de repos des articulateurs à la fin de la réalisation du geste (la phase dynamique du geste, qui lui apporte son sémantisme) et de la tenue de ce geste s'il y en a une (i.e. les mains restent en tension dans la même configuration de la réalisation du geste, mais ne bougent pas). Une phrase gestuelle est donc divisée en plusieurs phases. La phase de rétraction (i.e. le retour partiel ou complet des mains à une position de repos) et le rebond (i.e. un léger rebond des mains si elles sont posées sur la cuisse, par exemple) ne font pas partie de la phrase gestuelle. Les phrases gestuelles sont différentes des unités gestuelles, qui représentent l'unité supérieure dans la hiérarchie proposée par Kendon (2004). La segmentation des unités gestuelles repose sur la fin d'une phrase gestuelle qui n'est pas directement suivie par une autre phrase gestuelle, et sur un changement de mode de réalisation (e.g. passage d'un geste bimanuel symétrique à un geste bimanuel asymétrique).

particulièrement des frontières. Nous avons également vu dans le Chapitre 5 p. 181 qu'un certain nombre d'études comparent les syllabes sans accent aux syllabes accentuées, et trouvent que les segments acoustiques et articulatoires sont plus longs pour les syllabes accentuées que pour les syllabes qui ne sont pas accentuées (e.g. Edwards et al., 1991; de Jong et al., 1993; Fowler, 1995; Turk et Sawusch, 1997; Turk et White, 1999; Cho, 2006). Pour les syllabes accentuées aux frontières de constituants, les gestes articulatoires s'allongent temporellement à la fois en début et en fin d'unité (e.g. Edwards et al., 1991; Fougeron et Keating, 1997; Byrd et Saltzman, 1998; Byrd, 2000; Cho, 2006). Dans le signal acoustique, ce phénomène s'observe par l'allongement des segments adjacents aux frontières, i.e. l'allongement syllabique en fin de syntagme intonatif et parfois en début (e.g. Oller, 1973; Wightman et al., 1992; Shattuck-Hufnagel et Turk, 1998; Cho 2006). Ces effets temporels augmentent avec la force de la frontière : davantage d'allongement articulatoire et acoustique s'observe sur les frontières d'unités de hiérarchie supérieure (e.g. Wightman et al., 1992; Byrd et Saltzman, 1998; Shattuck-Hufnagel et Turk, 1998; Byrd, 2000; Cho, 2006 pour l'anglais). Les effets de cet allongement à la frontière sont locaux, i.e. ils sont produits autour de la frontière, et l'effet diminue lorsque ce phénomène s'éloigne (e.g. Byrd et al., 2006; Krivokapic, 2007).

Or, l'étude de Krivokapic (2014) sur de la parole spontanée en anglais montre que les fins d'unités prosodiques sont marquées par un allongement non seulement des gestes oraux, mais aussi des mouvements corporels, comme il avait déjà été démontré pour les tapotements de doigts aux syllabes proéminentes (Kelso et al., 1983; Parrell et al., 2014). Ce travail indique que l'influence de la structure prosodique s'étend au-delà des gestes du conduit vocal et inclut les gestes du corps (Wagner et al., 2014 pour un résumé), y compris les mouvements de sourcils et les hochements de tête sur les proéminences dans une moindre mesure (e.g. Hadar et al., 1983 pour l'anglais; Krahmer et Swerts, 2007 en néerlandais; Munhall et al., 2004 en japonais; Beskow et al., 2006; Al Moubayed et Beskow, 2009 pour le suédois). Dans le cas des proéminences, des études récentes ont indiqué que l'apogée (i.e. le déplacement maximum des articulateurs lors d'un geste par rapport à la position d'origine) des gestes est synchronisée avec la syllabe proéminente (e.g., Leonard et Cummins, 2011 pour l'anglais; Esteve-Gibert et Prieto, 2013 en catalan pour les pointages manuels; Esteve-Gibert et al., 2014 pour les acquiescements de tête). Un nombre de questions liées à cette coordination restent cependant ouvertes (Esteve-Gibert et Prieto, 2013, Wagner et al., 2014; Krivokapic, 2014), et le lien entre allongement des mouvements de tête et de sourcils aux frontières prosodiques n'est pas aussi documenté que celui aux proéminences prosodiques.

Gestualité et syntaxe

La relation entre gestualité et syntaxe a surtout été étudiée en psycholinguistique. McNeill (1992) observe un certain ratio entre les unités gestuelles manuelles et le format syntaxique : au cours de la tâche narrative imposée aux locuteurs, les locuteurs produisent environ un geste manuel par proposition. McNeill (*ibid.*) ne compte cependant que les gestes référentiels. D'autres études de production (e.g. Özyürek et Kita, 1999; Kita et Özyürek, 2003; Özyürek et al., 2005; Kita et al., 2007) ont également montré que la gestualité co-verbale est couplée aux choix de découpage syntaxique, notamment celui des propositions. En revanche, ces études ciblent surtout les gestes manuels référentiels, et s'intéressent à la description de mouvements lors de tâches narratives.

L'étude de Sekine et Kita (2017) repose sur le constat que les propositions dans la parole ont tendance à grouper et présenter l'information d'une manière similaire à celle des gestes co-verbaux en co-occurrence (Kita, 2000), et vice-versa. Ce constat s'inscrit dans l'hypothèse du groupement de l'information (Information Packaging Hypothesis) proposée par Kita (2000) : les gestes co-verbaux structurent de riches informations spatio-motrices en des groupements adaptés à la production de parole. Cette hypothèse prédit que la façon dont l'information est divisée en plusieurs gestes affecte la façon dont elle est divisée en plusieurs unités de parole (i.e. les propositions, d'après Bock et Cutting (1992)²). Cette similarité s'observe dans les représentations linguistiques (ici, de la parole) et gestuelle de la direction et de la manière lors de l'expression des événements de motion (Kita et Özyürek, 2003; Özyürek et al., 2005; Kita et al., 2007; Mol et Kita, 2012; Furman et al., 2014). Lorsque des locuteurs adultes expriment la direction et la manière dans une seule et même proposition finie, ces locuteurs ont tendance à produire un geste qui exprime direction et manière dans un seul et même mouvement. Par contraste, lorsque des locuteurs expriment la direction et la manière à l'aide de deux propositions finies

2. Bock et Cutting (1992) proposent le fait que les propositions finies représentent les unités de computation syntaxique. En utilisant une procédure pour éliciter des erreurs d'agrément entre les verbes, les auteurs trouvent que ces erreurs arrivent plus fréquemment lorsque la tête nominale et son verbe sont séparés par un syntagme (e.g. "The claim about the stolen babies was rejected", i.e. "l'affirmation à propos des bébés volés a été rejetée", p. 104) que lorsque qu'elles sont séparées par une proposition relative (e.g. "The claim that wolves were stealing babies was rejected", i.e. "L'affirmation selon laquelle les loups volaient des bébés a été rejetée"). En supposant une structure computationnelle hiérarchique, ils expliquent ce fait en statuant que plus l'information est introduite sous une seule unité computationnelle, plus les sources d'interférence sont nombreuses entre des éléments actifs similaires. Dans cet exemple, quand la tête nominale et le nom local font partie d'une seule proposition, le nombre de ces deux entités est plus susceptible d'interférer que lorsqu'elles sont réparties dans deux propositions différentes. En suivant Bock et Cutting (1992), Sekine et Kita (2017) utilisent la proposition finie en tant qu'unité pour estimer le groupement d'unités sémantiques dans la parole.

séparées, ces locuteurs ont tendance à les exprimer dans deux gestes séparés (Mol et Kita, 2012).

7.2.2 Les tests de perception en gestualité

Dans cette recherche comme dans celle portant sur les frontières prosodiques abordée dans le Chapitre 5, nous appréhendons la force des frontières en tant que notion perceptuelle en elle-même, comme un phénomène mesuré directement dans un test sur des participants naïfs, qui ne peuvent explicitement faire référence à des structures et phénomènes syntaxiques, phonologiques, ou prosodiques. Notre test de perception est construit sur le modèle de notre test portant sur les frontières prosodiques. Les deux tests peuvent ainsi faire l'objet de comparaisons en portant sur le même phénomène et en montrant les mêmes caractéristiques méthodologiques, en accord avec la perspective multimodale de ce travail. Le fait que davantage de travaux expérimentaux concernent la perception de l'apport informationnel dans la modalité gestuelle plutôt que la perception des frontières a également motivé cette orientation. Notre test a néanmoins été inspiré par divers travaux expérimentaux dans le domaine de la gestualité.

Les études perceptives sur la gestualité au niveau linguistique proviennent essentiellement du domaine des sciences cognitives ou de la psycholinguistique. Elles sont menées soit du point de vue des fonctions de la gestualité pouvant être bénéfiques au locuteur (avec la facilitation de la mise en mots et l'amélioration de la fluidité de la production vocale du locuteur (Krauss et al., 2000)), soit du point de vue des fonctions bénéfiques au co-locuteur (avec l'apport d'informations supplémentaires à la parole (e.g. McNeill, 1992; 2005; Goldin-Meadow, 2015), l'amélioration de la perception de descriptions (e.g. Driskell et Radtke, 2003), la désambiguïsation lexicale de certains mots (e.g. Holle et Gunter, 2007), l'amélioration de la compréhension dans des conditions de communication difficiles (e.g. Sueyoshi et Hardison, 2005), ou encore l'intérêt de la perception de gestes coverbaux sur la mémoire à long terme (e.g. Church et al., 2004)). En revanche, ces résultats sont en grande partie liés aux gestes référentiels, et répondent à des problématiques liées à la valeur communicative de ces types de gestes. Aucun travail, à notre connaissance, ne traite directement de la segmentation du discours à l'aide de gestes référentiels.

La relation entre les gestes non-référentiels (comme les battements) à la parole est encore moins étudiée en perception. Une ligne grandissante de travaux (Holle

et al., 2012; Biau et Soto-Faraco, 2013; Biau et al., 2016; Dimitrova et al., 2016) montre toutefois que les battements manuels jouent un grand rôle à la fois dans la segmentation syntaxique de la parole et dans la focalisation. Ces gestes manuels peuvent donc faciliter la segmentation du discours par les co-locuteurs. Toutefois, ces quelques études perceptives récentes sur la segmentation fine portent sur de la parole inventée et lue (Holle et al., 2012; Dimitrova et al., 2016; Biau et al., 2018), ou sur des extraits de discours politique (Biau et Soto-Faraco, 2013). Aucune étude sur les battements manuels, à notre connaissance, n'utilise de stimuli en partie issus de conversation spontanée.

Selon les variables observées, la tâche imposée aux participants, et le type de stimuli auxquels ces participants sont exposés, la performance des participants en perception peut être mesurée par des réponses à un questionnaire (e.g. Granström et al., 1999; House et al., 2001), par le temps de réaction à des stimuli (e.g. Barkuysen et al., 2008; Kelly et al., 2010), par l'oculométrie (e.g. Oben et Brône, 2015), par les PE (potentiels évoqués, i.e. la modification de l'activité électrique du système nerveux en réponse à une stimulation extérieure visuelle, auditive, sensitive ou motrice; e.g. Dimitrova et al., 2015), par les EEG (électro-encéphalogrammes, i.e. l'enregistrement de l'activité électrique produite par les neurones du cerveau; e.g. Meyer et al., 2012), qui permettent de surveiller certaines activités cognitives, ou par les IRM fonctionnels, qui mesurent l'activité des aires du cerveau en détectant les changements locaux de flux sanguin (Biau et al., 2016).

D'autres travaux expérimentaux sur la synchronisation entre indices prosodiques et gestuels dans la perception de la parole utilisent la modélisation, à des fins comparatives avec des extraits d'interaction spontanée (e.g. Schlangen, 2006; Atterer et al., 2008; Barkuysen et al., 2008). Ces études travaillent à l'implémentation de modèles sur des agents virtuels de conversation (e.g. Cassell et al., 2001). Dans ces travaux, l'acceptabilité des indices prosodiques et gestuels peut être jugée par des participants naïfs, directement par un questionnaire (Barkuysen et al., 2008) ou par leur aptitude à interagir avec les agents virtuels (Granstrom et House 2005; De Kok et Heylen, 2009). Ces travaux s'intéressent plus particulièrement aux indices multimodaux de fin de tour de parole. De Kok et Heylen (2009) montrent de cette manière que les mouvements de tête jouent un rôle primordial dans l'allocation et l'abandon des tours de parole (Duncan et Niederehe, 1974; Heylen, 2006; Rienks et al., 2010). Cette étude confirme également l'importance de la direction du regard pour ces mêmes tâches (Kendon, 1967; Duncan, 1972; Argyle et Cook, 1976; Goodwin, 1981). La relation des mouvements de sourcils à la perception de frontières

prosodiques dans la parole est également abordée, mais donne lieu à des résultats beaucoup plus mitigés (Granström et al., 1999).

L'utilité des indices gestuels dans la démarcation peut également être appuyée par des études plus générales sur l'attention portée à la gestualité par les co-locuteurs dans une interaction. Beaucoup de travaux montrent que les indices gestuels sont essentiels dans la construction et la gestion des connaissances communes dans l'interaction (Chartrand et Bargh, 1998; Parrill et Kimbara, 2006; Holler et al., 2011), et que locuteurs comme co-locuteurs s'appuient sur ces indices gestuels (Oben et Brône, 2016; Nakano et al., 2003; Gullberg, 2006; Hoetjes et al., 2015). Il a par exemple été montré que les locuteurs adaptent leur gestualité aux co-locuteurs, en s'appuyant sur les signaux que ces derniers communiquent (Galati et Brennan, 2010; 2014; Debreslioska et al., 2013; Perniss et Özyürek, 2015; Masson-Caro et al., 2016).

7.2.3 Les indices gestuels liés à la perception d'une frontière

Différentes études se sont d'abord concentrées sur les changements de posture en tant qu'indices de frontières dans le discours (Duncan, 1972; Beattie, 1978; Cassell et al., 2001), particulièrement au niveau de la prise et la fin de tour conversationnel. L'indice ayant cependant reçu le plus d'attention est la direction du regard. Alors que les locuteurs détournent le regard des co-locuteurs lors de l'élaboration du discours une fois le tour de parole pris (Kendon, 1967; Beattie, 1978; De Kok et Heylen, 2009), un changement dans la direction du regard vers le co-locuteur annonce une frontière dans le discours (Beattie, 1978; Novick et al., 1996; Vertegaal et al., 2000; Nakano et al., 2003; De Kok et Heylen, 2009) ou un appel au co-locuteur (Argyle et Cook, 1976; Goodwin, 1980; Holler et al., 2014). De même, lorsqu'une séquence de discours est achevée, les participants détournent le regard, alors que les partis maintiennent le regard l'un vers l'autre lorsque la séquence nécessite une expansion (Rossano, 2013). L'importance de cet indice est prêtée au fait que les yeux humains montrent une morphologie unique. Plusieurs études considèrent que le contraste entre l'iris noir et la large sclérotique blanche qui l'enveloppe a évolué au cours du temps pour faciliter la détection de la direction du regard par les humains (e.g. Kobayashi et Kohshima, 1997).

Les battements de tête isolés (i.e. un mouvement du menton vers le bas) et les acquiescements sont également considérés comme des indices de frontière, et sont reconnus comme tels par les co-locuteurs (Maynard, 1987; Heylen, 2006). Les mouve-

ments de sourcils, particulièrement les haussements, peuvent souligner et démarquer des entités particulières (Ekman, 1979; Cavé et al., 1996; Kraemer et Swerts, 2005). Les mouvements de tête et les haussements de sourcils sont liés à la prosodie, particulièrement à l’emphase et à la focalisation (Granström et House, 2005). Cependant, bien que le rôle des haussements de sourcils dans la démarcation de la parole soit attesté, la perception de ces indices en tant que tels donne lieu à des résultats moins probants (Granström et al., 1999; Barkhuysen et al., 2008).

Bien que la direction du regard, les mouvements de tête et les mouvements de sourcils jouent un rôle dans la démarcation de la parole, seuls les gestes manuels sont testés dans le cadre de cette étude. Ce choix résulte du fait que davantage d’études se sont intéressées aux mouvements du regard, de la tête, et des sourcils. De plus, dans les études de perception, ces dernières variables sont aisément manipulables dans le cadre de modélisation et d’implémentation sur des agents virtuels de conversation. Puisque nos stimuli audio sont issus de la conversation spontanée et synchronisés temporellement avec les gestes d’autres participants, la manipulation des gestes manuels uniquement est estimée plus pertinente dans le cadre de notre analyse de production du Chapitre 6, mettant clairement en avant le rôle des gestes manuels dans l’expression des frontières.

Peu d’études perceptives ciblent directement le rôle démarcatif des gestes manuels. Seule la capacité démarcative des battements manuels a été clairement mise en lumière (Holle et al., 2012; Biau et Soto-Faraco, 2013; Biau et al., 2016; Dimitrova et al., 2016). En revanche, il a été montré en production que deux gestes manuels consécutifs montrant des configurations et des coordonnées totalement différentes peuvent signaler une frontière dans le discours (Streeck, 2009; Enfield, 2009). Selon Calbris (2011, p. 46), “le changement crée une rupture dans le continuum et opère une fonction démarcative, alors que le type de changement fournit des informations et opère une fonction référentielle”. Ce phénomène s’explique par le fait que certains traits gestuels participent à la création de la cohésion discursive (Cassel et McNeill, 1990; Calbris, 2011; Hoetjes et al., 2015; Perniss et Özyürek, 2015). La représentation de référents par des gestes manuels est un processus cumulatif, souvent réalisé à travers une série de plusieurs unités gestuelles (Streeck, 2009). Les unités gestuelles répétées à travers une interaction créent de la cohésion discursive. McNeill (1992; 2005) observe et analyse ce phénomène par la notion de *catchment*, i.e. un groupe de gestes manuels présentant la même récurrence formelle dans leur configuration (cf. Chapitre 6 p. 213). De même, une série de gestes réalisés avec les mêmes coordonnées spatiales peut véhiculer la continuité. Il en va de même pour

l'utilisation des mains (main gauche, main droite, ou les deux), et pour la forme ou le style de geste (McNeill et Levy, 1993; Gullberg, 2006).

Dans cette même idée, l'adoption d'une position de repos pour les mains peut également signaler une frontière dans le discours si elle offre un contraste avec une séquence de gestes précédente (McCullough, 2005; Calbris, 2011). Au contraire, il a été montré que deux segments de parole peuvent également être liés par leur production dans une seule et même unité gestuelle (Enfield, 2009). En liant formellement un nouveau segment de discours à un précédent, la tenue d'un geste représente un autre moyen cohésif de représenter explicitement deux idées liées en même temps (Laursen, 2005; Enfield, 2009; Park-Doob, 2010; Frederiksen, 2016).

Les résultats de ces études montrent que les locuteurs signalent les frontières dans le discours par des moyens gestuels. En revanche, le constat du faible nombre d'études de perception explique le fait que l'on ne considère qu'un petit nombre de ces signaux gestuels de frontière comme manifestement interprétés en tant que tels par les co-locuteurs. Un grand nombre de questions subsistent également quant au poids de ces indices gestuels par rapport aux moyens prosodiques, et quant à l'effet combiné des deux modalités (Barkhuysen et al., 2008). Alors qu'il a été montré que les participants sont tout à fait capables de percevoir les frontières de manière auditive (cf. Chapitre 5), la capacité des participants à percevoir des frontières de manière visuelle de la même façon reste à documenter. La question de la manière dont les deux modalités sont liées dans la perception des frontières n'est pas résolue non plus. Barkhuysen et al. (2008) ont conduit une expérience basée sur des jugements de frontières en mesurant le temps de réaction aux stimuli, comptant trois conditions : une où les participants n'ont accès qu'aux informations visuelles, une où les participants n'ont accès qu'aux informations auditives, et une où les participants ont accès aux informations audio-visuelles. Les participants identifient les frontières plus vite dans la condition audiovisuelle. La différence n'est pas significative avec la condition où seules les informations auditives sont disponibles. La réponse la plus lente aux stimuli correspond à la condition où seules les informations visuelles sont disponibles. Les auteurs en concluent que la possibilité selon laquelle les participants ne se basent que sur les informations auditives pour la détection de frontières reste ouverte.

Dans notre étude, les participants ont accès aux informations audiovisuelles, mais aucun indice de frontière auditif n'est présent. Cette configuration nous permet de confirmer ou d'infirmer que les participants ne se basent que sur les informations au-

ditives pour la détection de frontières, et nous permet de mesurer la performance des participants lorsque des indices de démarcation uniquement visuels sont disponibles.

7.2.4 La force des frontières

Alors que les études perceptives en prosodie utilisent fréquemment des échelles graduelles permettant de mesurer la force des frontières (cf. Chapitre 5 p. 182), les études perceptives en gestualité ne font pas appel à ces échelles. La notion de force n'apparaît que chez Granström et al. (1999) dans les études s'intéressant à l'expression multimodale des frontières, et chez House et al. (2001) pour les proéminences, uniquement de manière indirecte. Elle est mesurée par la proportion de sujets indiquant avoir perçu une frontière ou une proéminence à un emplacement donné.

Construit en parallèle avec le test de perception des frontières prosodiques, notre test pose directement la question de la force des frontières aux participants naïfs. Il vise à évaluer la corrélation entre un nombre de variables gestuelles et les jugements quant à la présence de frontières dans le discours. Le terme de frontière est donc employé pour désigner une rupture perceptuelle produite par des moyens gestuels entre deux unités. L'échelle de Likert à cinq niveaux, utilisée dans notre test de perception des frontières prosodiques, est investie pour les frontières gestuelles. Cette échelle s'inspire des niveaux de perception définis par l'outil ToBI (Silverman et al., 1992) en prosodie et propose une notation relativement précise. Les degrés de force sont néanmoins adaptés aux auditeurs naïfs, en ne faisant référence à aucune unité ou structure verbale, prosodique, ou gestuelle. Elle présente les différents niveaux : pas de frontière (no boundary); frontière faible (weak boundary); incertain (uncertain); frontière (boundary); frontière forte (strong boundary).

Dans ce test, nos hypothèses sont les suivantes : (1) des frontières exprimées par des moyens gestuels sont perceptibles par des participants naïfs dans les séquences contenant des subordinées; (2) différents degrés de frontière peuvent être perçus; (3) la rupture est plus perceptible dans les séquences contenant des appositives relatives.

7.3 Corpus et stimuli

Un ensemble de vidéos ont été réalisées à partir d'extraits audio de notre corpus de conversation semi-spontanée en anglais britannique, ENVID. Les extraits vidéo sont d'une durée variable de 2 à 5 secondes (durée moyenne de 3.7 secondes), et proviennent de quatre locuteurs différents de notre corpus (deux femmes, deux hommes). La préparation des stimuli est détaillée ci-dessous.

7.3.1 Préparation des stimuli

En ce qui concerne l'audio, six extraits sans aucun trait prosodique assimilé à la création d'une rupture (4 extraits de locutrices, 2 extraits de locuteurs; les extraits comportant tout type de marqueur d'hésitation vocal ou verbal ont été exclus) ont été sélectionnés.

Dans chacun de ces extraits, la parole provient d'un seul locuteur, le signal est clair et ne présente aucun bruit. La séquence comprend trois unités : une proposition comportant une construction syntaxique subordonnée (deux extraits avec une appositive, deux extraits avec une circonstancielle, et deux extraits avec une relative déterminative), ainsi qu'une proposition précédant cette subordonnée (i. e. L, le co-texte gauche), et une proposition suivant cette subordonnée (i.e. R, le co-texte droit). Chaque extrait est donc analysable en tant que séquence de propositions L - SC - R.

Afin de s'assurer que ces stimuli soient clairement perçus comme sans rupture, seuls les stimuli ayant été validés par le pré-test auprès de six auditeurs naïfs anglophones natifs (mentionné dans le Chapitre 5 p. 184) ont été sélectionnés. Pour ce test, le taux d'accord entre participants est de 100 % : aucune rupture n'est perçue pour le set d'extraits.

En ce qui concerne la vidéo, nous avons filmé deux personnes (une femme pour les extraits audio provenant de locutrices, un homme pour les extraits audio provenant de locuteurs) produisant un, puis deux gestes (un seul geste dont la réalisation est en chevauchement avec l'énonciation de L et SC pour la première condition, un geste en co-occurrence avec l'énonciation de L + un geste en co-occurrence avec l'énonciation de SC pour les seconde et troisième conditions) dans chaque stimulus. La vidéo est réalisée avec une caméra d'appareil photo bridge, permettant un bon rapport entre haute définition, fluidité et taille des fichiers. La vidéo est ensuite

synchronisée temporellement avec le signal audio à l'aide du logiciel Movie Maker, donnant un stimulus audiovisuel. Une croix de fixation a été ajoutée à chaque stimulus afin de centrer l'attention et le regard des participants. Cette croix de fixation est visible dans la figure 7.1 ci-dessous.



FIGURE 7.1 – Croix de fixation apparaissant à l'écran pour une durée de deux secondes avant le début de chaque vidéo.

La nature et la configuration des gestes représentent les variables testées. Trois conditions différentes sont incluses, et nos stimuli comptent :

- 1) des extraits avec un seul geste³ produit en chevauchement sur les propositions de L et de SC, qui véhicule supposément la continuité dans le discours (Enfield, 2009)
- 2) des extraits avec deux gestes successifs identiques en termes de configuration (avec un geste sur L, un geste sur SC) séparés par un retour à la position de repos, qui participent à la cohésion du discours à travers la forme, l'espace, et la trajectoire (McNeill et Levy, 1993), mais sont séparés temporellement
- 3) des extraits avec deux gestes successifs différents en termes de configuration (avec un geste sur L, un geste sur SC) séparés par un retour à la position de repos, qui créent une forte frontière dans le discours (Calbris, 2011).

La réalisation des gestes a fait l'objet d'une synchronisation temporelle précise avec la parole. Pour chaque geste, l'apogée de la phase de réalisation du geste (i.e. l'apex) est alignée temporellement avec la réalisation de la syllabe proéminente accentuellement de l'affilié lexical. Tous les gestes montrent une durée, une amplitude, et une tension équivalentes selon les prérequis de chaque condition expérimentale.

3. Les types de gestes utilisés sont décrits plus loin dans la section p. 282.

Le visage des gestualisants est flouté, afin de filtrer les indices visuels disponibles (les mouvements de tête et de sourcils sont par exemple invisibilisés) et de fixer l'attention des spectateurs sur les mains en mouvement. Le filtre du visage permet également aux spectateurs naïfs de prêter facilement la parole entendue aux personnes gestualisantes, sans que la synchronisation des mouvements des lèvres avec l'audio soit nécessaire. Ce filtre est réalisé avec l'outil mis à disposition de la plateforme Youtube lors de la mise en ligne d'une vidéo. La posture de la personne gestualisante ainsi que le cadrage restent les mêmes à travers chaque vidéo : la personne est debout, jambes légèrement écartées, et son corps est visible de la tête aux genoux. La station debout permet de rendre les gestes manuels proéminents. Dans chaque vidéo quelles que soient la nature et la configuration des gestes testés, les mains sont en position de repos avant l'unité gestuelle. Un retour à cette position de repos s'observe après l'unité gestuelle. La position de repos est la même à travers les différents stimuli. La figure 7.2 ci-dessous montre les positions de repos de chaque personne gestualisante, ainsi que le cadrage des vidéos.

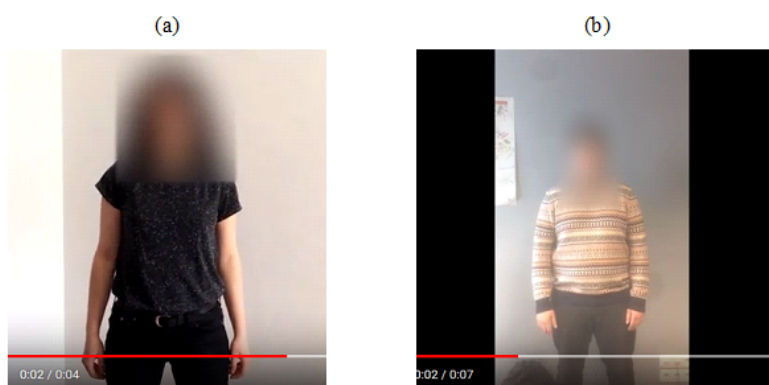


FIGURE 7.2 – Position de repos de la personne gestualisante dans les vidéos contenant des stimuli audio produits par des locutrices (a), et de la personne gestualisante dans les vidéos contenant des stimuli audio produits par des locuteurs (b). Ces positions de repos sont adoptées avant et après le geste.

En ce qui concerne les types de gestes exécutés dans les vidéos, nous avons choisi des gestes iconiques, métaphoriques, et des pointages à valeur référentielle, partageant un lien sémantique avec l'affilié lexical de chaque vidéo. Pour les vidéos contenant deux gestes répétés, nous avons privilégié des gestes aspectuels⁴ ou modaux⁵ car la répétition du même geste sans la répétition de l'affilié lexical auquel

4. Un geste manuel aspectuel véhicule des informations grammaticales sur le segment verbal avec lequel il est en co-occurrence. Il peut par exemple représenter un aspect duratif avec un geste circulaire.

5. Un geste manuel modal véhicule des informations à propos de la façon dont le locuteur prend position sur l'énoncé avec lequel le geste est en co-occurrence. Cette modalité peut être de nature

le geste est lié sémantiquement aurait donné un résultat incongru. L'exemple (1) ci-dessous illustre les choix quant aux types de gestes par rapport au sens de chaque stimulus audio.

- (1) Tim L you get an assessed [(a) piece of (b) work
 SC **which you do on a (c) computer]**
 R using a program called author catway

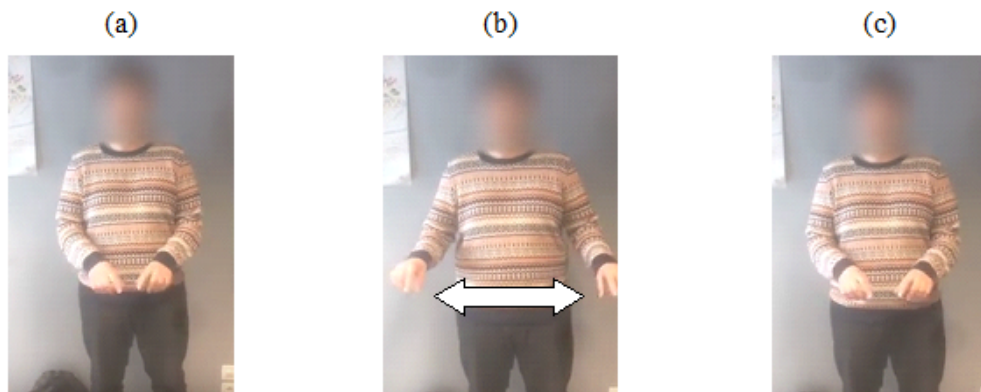


FIGURE 7.3 – Différents moments du geste iconique réalisé dans l'exemple (1). La préparation du geste est visible sur l'image (a), tandis que la réalisation et la rétraction sont respectivement visibles dans les images (b) et (c).

Dans le stimulus représenté par l'exemple (1) et la figure 7.3, le locuteur décrit la façon dont les étudiants urbanistes sont évalués ainsi que le type de travail qu'ils ont à rendre. L'item lexical affilié à l'apogée du geste iconique réalisé par la personne gestualisante est "work". Le geste trace un rectangle sur une surface plane, à la manière d'une feuille de papier sur un bureau. Ce stimulus fait partie de la condition 1 : il présente un seul geste, en chevauchement entre les propositions de L et de SC. Si la préparation et l'apogée du geste ont lieu sur L, la rétraction a lieu sur SC.

Les stimuli audiovisuels donnent 18 extraits et peuvent être listés comme tels :

- 6 extraits sans rupture audio avec réalisation d'un geste en chevauchement sur L et SC : condition 1 (avec 2 appositives relatives, 2 circonstancielles, 2 relatives déterminatives)
- 6 extraits sans rupture audio avec réalisation de deux gestes identiques en termes de configuration (un geste sur L, l'autre sur SC) : condition 2 (avec 2 appositives relatives, 2 circonstancielles, 2 relatives déterminatives)

assertive, épistémique, qualitative, et/ou déontique. Un geste paume ouverte vers le haut peut ainsi constituer un appel au co-locuteur, ou exprimer une incertitude par rapport à la parole en co-occurrence.

- 6 extraits sans rupture audio avec réalisation de deux gestes différents en termes de configuration (un geste sur L, l'autre sur SC) : condition 3 (avec 2 appositives relatives, 2 circonstancielles, 2 relatives déterminatives).

6 distracteurs viennent compléter les stimuli. Ces distracteurs ont été choisis selon les mêmes critères que les constructions authentiques en ce qui concerne l'audio (i.e. parole provenant d'un seul locuteur, signal clair et sans bruit), à la différence que les traits prosodiques et gestuels des distracteurs ne sont pas contrôlés. Ils se décomposent également en séquences L - SC - R.

Tous les stimuli en version finale ont été contrôlés par deux personnes étrangères à l'étude, afin de s'assurer que les stimuli audiovisuels ne donnent pas une impression d'incongruité. Pour chaque stimulus, la question posée était la suivante : la séquence de discours qui suit vous paraît-elle naturelle ? 100% des séquences ont été validées par ce processus.

Les 24 séquences (18 stimuli manipulés + 6 distracteurs) ont été randomisées dans un tableau Excel. Un numéro a été assigné à chaque vidéo (e.g. video 001, video 002). Les autres colonnes du tableau détaillent le type syntaxique auquel appartient la construction SC dans chaque extrait, son locuteur, ainsi que la condition pour les variables gestuelles et leur détail.

Les stimuli ont ensuite été déposés sur une playliste non-répertoriée sur Youtube, permettant quelques dernières finitions quant au traitement de l'image (amélioration de la fluidité, balance des couleurs).

7.4 Participants et procédure

7.4.1 Participants

Nos 24 stimuli ont été randomisés et présentés séquentiellement à 30 participants via une interface web (cf. infra p. 285). Une moitié des participants a été recrutée en ligne via des listes de diffusion mail ainsi que des réseaux sociaux. La seconde moitié des participants a été recrutée en personne à Bristol (Royaume-Uni). Pour cette seconde moitié, le test a été réalisé en présence de l'expérimentatrice. Des questions préliminaires sur l'interface web nous ont permis de nous assurer que tous les participants étaient anglophones natifs, qu'ils n'avaient aucun trouble auditif ou visuel à

leur connaissance, qu'ils n'avaient aucune expérience dans l'annotation prosodique et/ou gestuelle, et qu'ils avaient des écouteurs à disposition. Si au total 30 participants britanniques ont réalisé la tâche (âgés de 18 à 40 ans, moyenne : 24 ans), les réponses de 2 participants (recrutés en ligne) ont été exclues de l'analyse. Ces 2 participants ne remplissaient pas l'une ou plusieurs des conditions précédemment évoquées.

7.4.2 Le support de test

Le test a été réalisé sur eSurv (2017), une plateforme open-source de réalisation de questionnaires (cf. Chapitre 5 p. 189 pour la description de cette plateforme, ainsi que celle des avantages et des limites de la diffusion d'un test en ligne).

La diffusion de ce test en ligne possède les mêmes avantages et inconvénients que celle du test de perception des frontières prosodiques. En particulier, ce test nécessite une connexion internet de bonne qualité pour la fluidité de la vidéo, ainsi qu'un écran doté d'une bonne luminosité. Or, les participants disposent de matériel informatique dont la qualité peut varier. Si notre test nécessite également un casque audio pour des conditions de réalisation optimales, nous ne disposons d'aucun moyen de nous assurer que les auditeurs suivent à la lettre les instructions d'écoute et de visionnage. De même, la fatigue attentionnelle des participants n'est pas mesurée. Nous avons cependant évoqué p. 189 le fait qu'un nombre important d'études montre une absence de différences significatives entre les réponses de participants à un test réalisé en laboratoire et celles de participants à un test en ligne. Nous avons également opté pour un test relativement court à compléter (15 minutes).

7.4.3 Procédure

Les participants cliquant sur le lien du questionnaire sont dirigés sur la première page de test, où un texte explicatif expose brièvement les enjeux de l'étude. Ce texte explicatif est disponible en annexe p. 351.

Bien que le questionnaire soit anonyme, les participants répondent sur la seconde page à six questions d'informations personnelles (pays de résidence, langue maternelle, âge, sexe, port d'écouteurs, présence/absence de trouble auditif ou visuel). Ces questions ont pour but de filtrer les participations d'un public non-cible (e.g. un participant anglophone non-natif).

Les questions débutent sur la troisième page. La structure est identique pour chacune d’entre elles et les questions sont au nombre de dix par page. La figure 5.1 plus bas illustre l’interface de test. Chaque question présente la vidéo de l’extrait numéroté (e.g. video 001), ainsi qu’une transcription orthographique de l’extrait face à la vidéo. La transcription ne présente aucune ponctuation ni saut de ligne. L’endroit de réalisation potentielle d’une frontière gestuelle (i.e. entre L et SC) est marqué d’un astérisque rouge dans la transcription (e.g. “especially in class * when someone says something funny and I’m always dying”). Sous la vidéo (à gauche de la fenêtre de test) et la transcription (à droite de la fenêtre), une échelle de Likert à cinq niveaux présente les différents boutons : no boundary; weak boundary; uncertain; boundary; strong boundary. Cette échelle s’inspire des niveaux de perception définis par l’outil ToBI (Silverman et al., 1992) et ceux de Carlson et Swerts (2003) en proposant une notation relativement précise. La consigne figure en chapeau à chaque page du test : *Please rate the presence or absence of a boundary at **. La figure 7.4 illustre l’interface de test.

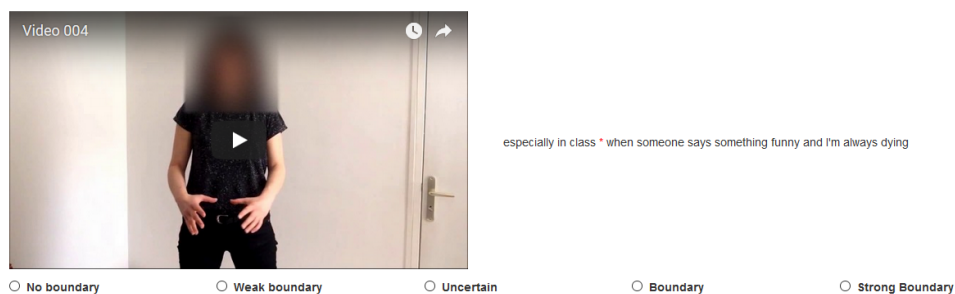


FIGURE 7.4 – Interface de la troisième page du test de perception réalisé sous eSurv, où figurent les premières questions. Chaque question présente la vidéo d’un stimulus ainsi que sa transcription, suivi de l’échelle graduelle de force de frontière perçue à cinq niveaux.

7.4.4 Analyse des données

Des ratios ont été calculés successivement pour chaque variable en respectant l’échelle à cinq niveaux. Une série de modèles linéaires mixtes généralisés (*GLMMs* estimés par la méthode du maximum de vraisemblance) a également été réalisée avec le programme R 3.4.0 (R Core Team, 2017) et le *package* lme4 (Bates et al., 2017), afin d’expliquer le degré d’influence de chaque variable sur la perception des frontières. De la même manière que pour le test de perception des frontières prosodiques, nous ne mesurons pas l’accord inter-annotateur à l’aide du Kappa de Fleiss,

en raison du nombre de catégories à choisir pour les participants (cinq) et du nombre relativement réduit de stimuli (cf. Chapitre 5 p. 191).

7.5 Résultats

Nous avons d’abord exploré les interactions potentielles entre les trois types de constructions syntaxiques (facteur fixe = Type; valeurs = appositive, circonstancielle, relative) et le jugement du poids des frontières (facteur fixe = Scores; valeurs = frontière forte (strong boundary), frontière (boundary), frontière faible (weak boundary), incertain (uncertain), pas de frontière (no boundary)). Des interactions ont été trouvées pour les scores “frontière forte”. L’effet principal des jugements “frontière forte” (facteur fixe = Strong; valeurs = oui, non) est significatif pour les appositives ($\beta = 2.30$, $SE = .35$, $p < 0.0001$) seulement. La différence avec les circonstancielle est également significative ($\beta = -1.34$, $SE = .61$, $p < 0.05$). De plus, l’effet principal des jugements “frontière forte” sur les appositives n’est pas significatif pour les stimuli faisant office de distracteurs. En revanche, la différence avec les relatives déterminatives n’est pas significative. La différence entre les appositives et les relatives déterminatives ne devient significative que lorsqu’on s’intéresse à l’effet principal des jugements “pas de frontière” (facteur fixe = No boundary; valeurs = oui, non), qui est significatif pour les appositives seulement ($\beta = -3.37$, $SE = .63$, $p < 0.0001$; dans le sens où les appositives comptent moins de jugements “pas de frontière” que les autres constructions), et dont la différence avec les relatives déterminatives est significative ($\beta = 0.90$, $SE = .37$, $p < 0.02$). Le tableau 7.1 résume ces résultats.

En revanche, aucun effet principal significatif n’est observable sur les circonstancielle, ni sur les relatives déterminatives.

	stimuli	distracteurs
appositives	frontière forte ***	n.s
circonstancielle	n.s	n.s
relatives déterminatives	n.s	n.s

TABLE 7.1 – Interactions significatives entre les types syntaxiques et les scores de frontière dans les stimuli et les distracteurs (***) = $p < 0.0001$; n.s = non significatif).

En résumé, alors que les appositives sont significativement associées à la perception de frontières fortes, les circonstancielle et les relatives déterminatives ne sont

pas associées à la perception de frontières. Les frontières gestuelles ne sont donc pas perçues avec la même intensité dans les trois constructions subordonnées. Un écart s’observe dans la perception de la force des frontières entre les appositives relatives, et les deux autres constructions subordonnées syntaxiques. Cet écart disparaît dans les distracteurs.

Nous avons ensuite exploré les interactions entre les indices gestuels et le jugement du poids des frontières dans nos trois types syntaxiques de constructions subordonnées.

En ce qui concerne les appositives relatives, l’effet principal de la condition 3 (deux gestes différents avec l’un sur L et l’autre sur SC) est significatif pour les frontières “fortes” ($\beta = 1.38$, $SE = .54$, $p < 0.01$). En revanche, la condition 1 (un seul geste produit en chevauchement sur L et sur SC) et la condition 2 (répétition du geste de L sur SC) ne montrent aucune interaction significative avec les jugements.

L’effet principal de la condition 3 (deux gestes différents avec l’un sur L et l’autre sur SC) est différent pour les circonstancielles. Il est seulement significatif pour les jugements “pas de frontière” ($\beta = -1.54$, $SE = .51$, $p < 0.005$), dans le sens où moins de stimuli de la condition 3 sont perçus comme n’ayant pas de frontière pour ce type syntaxique de subordonnée.

Les relatives déterminatives, elles, ne montrent pas d’interactions significatives avec les conditions gestuelles. Le tableau 7.2 offre un récapitulatif de ces résultats.

	appositives	circonstancielles	relatives déterminatives
interaction	condition 3 *	condition 3 **	n.s
score	“frontière forte”	“pas de frontière”	-

TABLE 7.2 – Interactions significatives entre les conditions gestuelles et les scores de frontière dans chaque type syntaxique (***) = $p < 0.0001$; ** = $p < 0.005$; * = $p < 0.05$).

En résumé, l’exploration des interactions entre les trois conditions gestuelles et le jugement du poids des frontières montre que la condition 3 est utilisée par les auditeurs naïfs pour percevoir des frontières gestuelles dans les constructions syntaxiques, mais à une force différente. Elle n’est également pas utilisée dans toutes les constructions syntaxiques, puisqu’aucune interaction ne concerne les relatives déterminatives. La production de deux gestes différents est donc un prédicteur dans la perception de frontières. La condition 1 et la condition 2 n’ont également aucun

effet significatif pour les jugements de frontière ou de non-frontière.

7.6 Discussion

En ce qui concerne les interactions potentielles entre les trois types de constructions syntaxiques (facteur fixe = Type; valeurs = appositive, circonstancielle, relative) et le jugement du poids des frontières (facteur fixe = Scores; valeurs = frontière forte (strong boundary), frontière (boundary), frontière faible (weak boundary), incertain (uncertain), pas de frontière (no boundary)), nous avons vu avec le tableau 7.1 qu'alors que les appositives sont significativement associées à la perception de frontières fortes, ni les circonstancielles, ni les relatives déterminatives ne sont associées à la perception de frontières. Ces résultats ne peuvent être expliqués en termes de longueur des stimuli, de position séquentielle, de statut discursif, ou d'indices prosodiques et/ou gestuels des constituants, puisque ces paramètres ont été contrôlés. La complexité et le poids des constituants restent toutefois des paramètres potentiels.

Nos résultats concernant les appositives sont en accord avec les résultats macrosyntaxiques, prosodiques, et gestuels en production, puisque les appositives expriment une grande rupture par rapport à la séquence hôte (cf. Chapitre 3 p. 117, Chapitre 4 p. 155, et Chapitre 6 p. 240). Nous avons particulièrement souligné le fait que les gestes manuels produits en co-occurrence avec les appositives sont très rarement en chevauchement sur plusieurs groupes intonatifs, et qu'ils montrent des configurations très différentes des autres gestes produits dans le co-texte. Du côté de la perception, les résultats concernant les appositives sont également en accord avec la perception des frontières prosodiques, puisque les appositives sont significativement associées à la perception de frontières fortes (cf. Chapitre 5 p. 193).

En revanche, les circonstancielles ne sont pas associées à la perception de frontières. Ce résultat n'est pas surprenant dans la mesure où ces constructions syntaxiques sont produites avec peu de marqueurs gestuels de rupture (cf. Chapitre 6 p. 223). Nous avons remarqué que quelques marques de rupture autour des gestes manuels empêchent de considérer les circonstancielles comme intégrées à leur co-texte, mais sont trop peu nombreuses et diversifiées pour réellement parler d'autonomie. Cette absence de frontières gestuelles n'est pas non plus surprenante en regard avec le plan macro-syntaxique, puisque nous avons vu dans le Chapitre 3 p. 104 que les circonstancielles montrent en majorité une autonomie macro-syntaxique partielle, où l'autonomie syntaxique ne correspond pas nécessairement à l'autonomie pragma-

tique. En termes de perception, les circonstancielles sont significativement associées aux frontières faibles en ce qui concerne la prosodie (cf. Chapitre 5 p. 193). Nous observons alors un écart dans les perceptions prosodiques et gestuelles des frontières dans le cas de ce type syntaxique.

Les relatives déterminatives ne sont pas non plus associées à la perception de frontières. Cette absence de frontières gestuelles perçues correspond à leur production gestuelle. Nous avons vu que les relatives déterminatives ne sont caractérisées que par un seul indice gestuel de frontière, les battements manuels (cf. Chapitre 6 p. 208). Nous avons également vu que l'usage de ces battements témoigne davantage de divers phénomènes discursifs que d'une réelle délimitation entre la relative déterminative et son co-texte gauche. De plus, les battements ne font pas partie des variables prises en compte dans ce test perceptif. Les résultats des relatives déterminatives correspondent également à la fois à leur perception prosodique (puisque les relatives déterminatives sont négativement corrélées à la perception de frontières prosodiques, cf. Chapitre 5 p. 193) et à leur production vocale (cf. Chapitre 4 p. 140). En termes de macro-syntaxe, nous avons vu dans le Chapitre 3 p. 89 que les relatives déterminatives sont majoritairement pleinement intégrées à leur co-texte gauche, régies syntaxiquement par la proposition du co-texte gauche.

En résumé, les frontières gestuelles ne sont pas perçues avec la même intensité dans les trois constructions subordonnées. Un écart s'observe dans la perception de la force des frontières entre d'une part les appositives relatives, et d'autre part les circonstancielles et relatives déterminatives. Cet écart ne suit pas le même schéma que les résultats de production et de perception prosodique, puisqu'une gradation entre les trois types syntaxiques s'observe pour la perception prosodique. En revanche, les appositives relatives restent les constructions perçues comme les plus en rupture. Le fait que les frontières prosodiques des circonstancielles soient perçues alors qu'aucune frontière gestuelle n'est perçue pour ces constructions syntaxiques suggère que les frontières exprimées de manière gestuelle dans notre test sont moins perceptibles que celles exprimées de manière prosodique. Cette différence de perception peut être liée à la différence de modalité, mais aussi à la nature de nos stimuli.

Nous avons ensuite exploré les interactions entre les indices gestuels et le jugement du poids des frontières, dans nos trois types syntaxiques de constructions subordonnées. Comme l'a montré le tableau 7.2, seule la condition 3 (production de deux gestes successifs différents en termes de configuration avec un geste sur L, un geste sur SC, séparés par un retour à la position de repos) a un effet significatif sur

la perception de frontières gestuelles pour deux de nos constructions syntaxiques sur trois. L'effet principal de la condition 3 est significatif pour la perception de frontières fortes dans le cas des appositives relatives. Ce même effet est également significatif dans le cas des circonstancielles, mais pour la perception de frontières à un moindre degré (perception de frontières faibles), et n'est pas significatif pour les relatives déterminatives. Ce résultat suit la tendance du précédent : les frontières gestuelles ne sont pas perçues dans les relatives déterminatives, même avec l'indice gestuel le plus fort. En revanche, à la différence de nos premiers résultats, les frontières sont perçues dans les circonstancielles. Ces seconds résultats sont alors plus proches des résultats du test de perception prosodique, puisqu'une rupture est perçue pour les circonstancielles et qu'une gradation s'observe dans la perception de rupture à travers les types syntaxiques pour la condition 3. La perception de la force des frontières gestuelles est donc également impactée par le jugement syntaxique en ce qui concerne la condition 3.

La condition 1 et la condition 2 n'ont aucun effet significatif pour les jugements de frontière, que ce soit dans le sens des frontières ou de l'absence de frontières. Si la condition 3 prouve que la gestualité a un impact sur la perception de frontières fortes, nos résultats ne montrent donc pas que la production d'un geste en chevauchement sur deux propositions (condition 1) ou que la répétition d'une même forme gestuelle sur deux propositions (condition 2) sont significativement associées à la perception d'absence de frontières, ou à la perception de la cohésion dans le discours. Les indices gestuels représentant ces deux conditions ne peuvent donc être considérés ici en tant qu'indices de continuité.

Ce dernier constat suit également les conclusions de l'observation des constructions syntaxiques du point de vue de leur production. Nous avons vu qu'elles ne présentent pas les indices d'intégration et de continuité gestuels traditionnellement associés au maintien ou à la modification des référents dans le discours. Les configurations bimanuelles (symétriques comme asymétriques) décrites par Frederiksen (2016) liées à la hiérarchisation et à la modification des référents dans le discours ne sont pas courantes dans nos types syntaxiques de subordonnées. De même, les tenues de gestes manuels liées à l'intégration, l'extension, ou la modification d'un segment de discours décrites entre autres par Enfield (2009) sont rares.

L'observation selon laquelle la condition 3 a un effet sur la perception des frontières dans les subordonnées est remarquable du point de vue des caractéristiques de production gestuelle des subordonnées. Nous avons souligné dans le Chapitre 6 que les

différences entre les trois types syntaxiques quant à la nature et à la distribution des indices suggèrent qu'aucun indice commun n'est systématiquement utilisé pour la démarcation des subordonnées. Or, malgré cette absence de systématisme dans l'emploi des indices gestuels, notre test montre que les participants sont sensibles à certains traits gestuels.

De plus, la production gestuelle des subordonnées est également marquée par la présence des battements (manuels dans les relatives déterminatives, de tête dans les circonstancielles) et des haussements de sourcils. Cette présence souligne l'importante contribution de l'aspect prosodique de la gestualité. Or, nous n'avons pas inclus de variable gestuelle qui fasse appel à l'aspect prosodique de la gestualité dans notre test de perception. Nous sommes donc en mesure d'affirmer que la segmentation du discours par la gestualité ne passe pas uniquement par la prosodie, mais exploite des paramètres propres à la modalité gestuelle, tels que la configuration des articulateurs ou la trajectoire. Notre étude est la première à montrer que des paramètres gestuels sans aspect prosodique influencent la perception de frontières dans le discours.

En production, la quantité et la distribution des indices gestuels sont moindres par rapport à celles des indices vocaux décrits dans le Chapitre 4. Nous avons suggéré des pistes d'explication en production dans le Chapitre 6, en envisageant deux possibilités. L'absence de corrélation entre les indices gestuels en production dans la plupart des constructions indique soit que les locuteurs accordent plus d'importance à la modalité vocale et attribuent d'emblée un faible rôle aux indices gestuels en complément aux phénomènes vocaux, soit que les locuteurs considèrent que le peu d'indices gestuels utilisés suffisent pour exprimer une démarcation, et que chacun d'eux possède un poids démarcatif important. Le fait que les battements de tête et les mouvements de sourcils n'apparaissent pas ensemble a par exemple été relevé dans d'autres situations de discours par House et al. (2001), qui considèrent que ces deux indices sont suffisamment forts (i.e. ont suffisamment d'impact sur la perception du co-locuteur) pour être employés seuls en tant que marqueurs. Ce test de perception nous permet de préciser ces possibilités. La production de deux gestes manuels successifs différents dans leur configuration (condition 3) suffit aux participants pour percevoir une frontière dans le discours, même en absence d'indice de rupture prosodique, mais le format syntaxique impacte la perception des frontières. Aussi, la production de deux gestes manuels successifs différents dans leur configuration suffit pour être considérée comme un indice de rupture dans les appositives relatives et les circonstancielles, mais pas dans le cas des relatives déterminatives.

Une suite possible du test de perception gestuel pourrait être d'agglomérer plusieurs indices de rupture gestuelle sur la production des relatives déterminatives, afin de déterminer si une corrélation d'indices contrecarre le jugement syntaxique et a un effet significatif sur la perception des frontières. Toujours est-il qu'un indice prosodique influence davantage les participants dans la perception de frontière qu'un indice gestuel, puisque dans le cas des indices vocaux, certains suffisent à influencer la perception de frontière dans nos trois types syntaxiques de subordinées. La force de la frontière perçue est la même entre les pauses silencieuses (ou l'allongement syllabique final) et deux gestes manuels différents dans les appositives relatives. En revanche, dans les circonstancielles et les relatives déterminatives, les pauses silencieuses en tant qu'indice isolé, l'allongement final en tant qu'indice isolé, et le rehaussement initial en tant qu'indice isolé conduisent à la perception de frontières d'une force supérieure à celle des frontières perçues à l'aide de deux gestes manuels différents en termes de configuration.

7.7 Conclusion

Nous avons présenté dans ce chapitre une expérience visant à étudier la perception des frontières gestuelles par des participants naïfs, dans des extraits de parole spontanée contenant une construction subordonnée syntaxiquement. Nous avons confirmé les hypothèses suivantes : (1) la perception des frontières gestuelles varie selon les différents types syntaxiques de subordination; (2) les différents types syntaxiques ne montrent pas les mêmes niveaux d'interaction avec les indices gestuels contenus dans les différentes conditions expérimentales; (3) la rupture est plus perceptible dans les séquences contenant des appositives relatives.

Nous avons vu que les frontières gestuelles ne sont pas perçues avec la même intensité en fonction des types syntaxiques : globalement, seules les appositives sont significativement associées à la perception de frontières. En explorant les interactions des types syntaxiques avec les indices gestuels représentés par les différentes conditions, la condition 3 (réalisation de deux gestes différents en termes de configuration, l'un sur L, l'autre sur SC) a une influence sur la perception de frontières. Ce résultat est conforme à ce que nous attendions, puisque les conditions 1 et 2 contiennent supposément des indices de continuité. Dans la condition 3, les appositives sont significativement associées à la perception de frontières fortes, les circonstancielles à la perception de frontières faibles, et les relatives déterminatives ne sont pas associées à la perception de frontières. Cette gradation dans la perception des

frontières correspond à la gradation de l'autonomie macro-syntaxique des différents types syntaxiques de constructions subordonnées traitée dans le Chapitre 3, et à la gradation de l'autonomie prosodique des différents types syntaxiques vue dans le Chapitre 4. Les correspondances entre la pertinence de l'indice représenté par la condition 3 (i.e. deux gestes différents en termes de configuration réalisés successivement) en perception et sa fréquence d'occurrence en production sont régulières, puisque la production de deux gestes manuels consécutifs formellement différents correspondent à la description de la production des appositives. Nous avons également relevé dans le Chapitre 6 qu'un très faible nombre de séquences comprenant une circonstancielle comportent une répétition formelle d'un geste manuel précédent (que ce dernier ait été produit sur L ou sur SC). Une correspondance s'établit donc en production, à la fois dans les appositives et dans les circonstancielles. L'absence de frontières perçues dans les relatives déterminatives en dépit de la condition 3 correspond également à la production gestuelle de ces subordonnées.

Nous avons également observé, à partir de nos résultats, que si la gestualité est bel et bien impliquée dans la perception des frontières, les frontières exprimées de manière gestuelle sont perçues à un moindre degré par rapport aux frontières exprimées de manière vocale, de la même manière qu'en production, les frontières sont exprimées avec davantage de moyens prosodiques que gestuels.

Le degré d'influence inférieur de la gestualité dans la perception des frontières par rapport à celui de la prosodie reste à nuancer, notamment en raison de la différence de poids de nos deux tests de perception. Ne comportant que deux indices de rupture (représentés par la condition 2 pour une rupture faible et par la condition 3 pour une rupture forte), ce test ne permet pas en revanche d'étudier l'influence des corrélations entre les indices dans la perception de frontières. La motivation de ce test, à valeur exploratoire puisqu'aucune étude ne documente la valeur de segmentation de la gestualité sans valeur prosodique, est d'abord de démontrer l'influence de la gestualité. De plus, nous sommes tenue à un faible nombre de stimuli, en raison de la nature du test (en ligne), et de la fatigue attentionnelle des participants. Nous tenons également à travailler à partir de vidéos de personnes gestualisantes qui ne sont pas des acteurs, en limitant les retouches et sans synthèse vidéo. Il sera intéressant, dans les suites potentielles de ce test, de mesurer l'intensité de la liaison entre les indices dans la perception d'une frontière gestuelle, et de tester si elle diffère en fonction du type syntaxique. Toutefois, nous savons d'ores et déjà que les correspondances entre les corrélations en production et perception ne concerneront que les appositives relatives, puisque rares sont les constructions produites avec plus d'un

indice de frontière à la fois. La pertinence des résultats concernant les corrélations d'indices de frontière en perception se trouvera donc davantage sur le plan de la modélisation.

En mettant à jour des différences dans la perception des frontières gestuelles en fonction du type syntaxique de subordination, ce test contribue à l'exploration de l'interface entre syntaxe et gestualité. Si notre test ne reflète pas les prises de décision spontanées et en temps réel des co-locuteurs en situation de conversation en raison de l'isolation des indices gestuels, il permet néanmoins d'exposer des participants naïfs à un stimulus ressemblant à des bribes de conversation spontanée. L'avantage de nos stimuli réside également dans le fait qu'ils ont été réalisés en présence d'un interlocuteur (Petroni et al., 2017). Toutefois, la présence de personnes gestualisantes suivant des instructions dans ce test représente également une limitation, puisque la partie vidéo de nos stimuli audiovisuels reste guidée, donc manipulée. Alors que les locuteurs des stimuli visuels s'ajoutent aux locuteurs des stimuli audio, les participants du test doivent associer les personnes gestualisantes aux voix des stimuli visuels. Si ce processus est utilisé dans la synthèse et la modélisation vidéo, le mécanisme d'association entre voix et personnes gestualisantes est différent, puisqu'il mêle stimuli authentiques, spontanés, à ceux qui résultent d'un suivi d'instructions. Cette influence est à prendre en compte dans la pertinence de nos résultats.

Tout comme notre test de perception prosodique, l'utilisation de modèles linéaires mixtes généralisés permet d'étudier l'influence de chaque variable sur la perception des frontières avec un traitement statistique robuste malgré le petit nombre de participants et de stimuli (accentuant notamment la différence entre les stimuli produits par des femmes et ceux produits par des hommes), ainsi que la potentielle hétérogénéité du groupe d'auditeurs. En revanche, les modèles linéaires ne prennent pas en compte certains paramètres situationnels relatifs aux conditions expérimentales. Si la fatigue attentionnelle pose beaucoup moins problème que pour le test prosodique en raison de la courte durée de notre test (20 minutes), adapté à des participants naïfs, nous supposons que la variation individuelle joue un grand rôle dans la perception des frontières gestuelles comme pour celle des frontières prosodiques. En revanche, la variation individuelle au niveau de l'assignation du poids des différents indices dans la perception de ces frontières (Roy et al., 2017) est davantage contrôlée, puisque notre test contient un seul indice de frontière, représenté par la condition expérimentale 3. Les développements potentiels de ce test devront diversifier la présence des indices gestuels de rupture tout en prêtant une attention particulière à cette variation individuelle.

Conclusion

S'appuyant sur un corpus vidéo de conversation spontanée en anglais britannique, le travail que nous présentons aborde la question de la hiérarchisation des constituants syntaxiques et discursifs en anglais oral spontané tel qu'il se manifeste dans un corpus conversationnel.

Ancré dans l'Analyse de Discours Multimodale, ce travail examine les choix des participants quant au déploiement des modalités communicatives dans la réalisation des constructions subordonnées.

En partant de l'hypothèse selon laquelle les constructions subordonnées recouvrent plus de diversité que ne leur prête la littérature, nous nous sommes attachée à décrire les différents types d'autonomie que présentent ces constructions par une analyse multimodale, du point de vue de leurs différents types syntaxiques (appositives, circonstancielles, relatives déterminatives).

A travers ces analyses, nous avons défini plusieurs unités communicationnelles pertinentes à l'étude dans les différentes modalités prises en compte, à différents niveaux. Nous avons exploré la notion de subordination des constituants du discours aussi bien au sens strict des relations syntaxiques qu'au sens plus large des relations discursives, 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 connaissances partagées. La proposition, le segment de discours, le tour de parole sont alors autant de ressources verbales simultanées et non-exclusives manipulées et projetées par les participants que nous avons prises en compte, au même titre que le groupe intonatif et le paragraphe à un niveau vocal, ainsi que l'unité gestuelle dans son contexte séquentiel au niveau de la gestualité.

Centrée autour du large concept de construction subordonnée, notre structu-

ration s'appuie sur un seul axe d'organisation discursive. Nous avons évalué le degré d'autonomie des différents types de subordinées sur un mode d'organisation linéaire (intégration vs. autonomie).

La multimodalité étudie les façons dont les contributions verbales, vocales, et visuelles s'articulent dans la construction du sens et de l'interaction. Au cours de la production des constructions subordinées, les locuteurs augmentent l'efficacité communicationnelle en véhiculant des informations différentes dans des modalités variées, qui prennent sens à plusieurs niveaux. Plus que le constat qu'une distribution du sens s'effectue dans les différentes modalités, nous sommes désormais en mesure d'observer que la production des différents types syntaxiques de subordinées n'implique pas les mêmes combinaisons : elles témoignent de stratégies multimodales bien différentes. Notre hypothèse de départ se trouve alors confirmée par les résultats de l'analyse.

Ils révèlent d'abord que l'identification et la description des subordinées sont basées sur des faisceaux d'indices bien distincts, et qu'elles sont loin de toutes montrer l'intégration au co-texte qui leur est prêtée.

Les relatives déterminatives sont les constructions les plus intégrées à leur environnement à la fois verbalement (par leur réaction grammaticale, leur anaphoricité), vocalement (soit par un phénomène direct de regroupement sous un seul et même groupe intonatif avec le co-texte gauche, soit par des contours montants finaux à valeur continuative), et gestuellement (par une intégration à l'unité articulatoire produite avec le co-texte gauche, ou par leur production en co-occurrence avec des unités gestuelles dont la configuration et la trajectoire sont très similaires aux gestes manuels environnants). Ce type syntaxique de subordinée se démarque seulement par le rythme (durée, pauses silencieuses intra-constituant, battements manuels isolés).

Les circonstancielles sont partiellement autonomes. Du point de vue verbal, leurs morphèmes introducteurs présentent une remarquable diversité de fonctionnement, et ces constructions subordinées sont mobiles dans la macrostructure. Leur intonation véhicule un changement, mais elle indique une altération des schémas interprétatifs et/ou informationnels (à l'image de leur capacité pragmatique à projeter un cadre interprétatif sur plusieurs groupes intonatifs consécutifs ou bien à clôturer une unité) plus qu'elle ne segmente les circonstancielles en unités isolées. En ce qui concerne la gestualité, les circonstancielles se démarquent uniquement par le biais des gestes manuels. Les gestes manuels les accompagnant sont plutôt divers

en termes de trajectoire et de configuration. L'absence d'indices gestuels de nature rythmique sur les circonstancielles fait écho aux caractéristiques prosodiques de ce type syntaxique, puisque les circonstancielles sont davantage démarquées par l'intonation que par le rythme. La gestualité des locuteurs pendant la production d'une séquence discursive contenant une circonstancielle travaille davantage à la construction de zones de consensus et d'échange dans le discours, qu'à la démarcation de ce type de subordonnée sur un plan linéaire.

Les appositives, quant à elles, sont clairement autonomes. Elles présentent un large faisceau d'indices macro-syntaxiques et discursifs d'autonomie. Les co-locuteurs sont sensibles à la rupture créée par les appositives, puisque la frontière entre celles-ci et leur co-texte droit représente un locus préférentiel pour la production de *back-channels*. Ce sont également les seules constructions à présenter un éventail aussi vaste d'indices de disjonction prosodique, tant au niveau segmental qu'au niveau suprasegmental. La rupture est donc plus perceptible, avec la plus grande différence significative de hauteur F0 non seulement entre la subordonnée et le co-texte droit, mais aussi entre la subordonnée et le co-texte gauche. Ce type de subordonnée est également réalisé avec la plus faible distribution de contours continuatifs. Au niveau gestuel, plusieurs articulateurs sont utilisés à des fins démarcatives. Les unités gestuelles produites en co-occurrence avec les appositives sont également bien distinctes de celles du co-texte. L'autonomie de ces constructions s'observe donc à la fois sur la qualité et la forme des unités (gestes manuels formellement différents de ceux environnants) que sur leur temporalité (unités des gestes manuels ne franchissant pas les frontières des groupes intonatifs, positionnement d'unités réalisées avec d'autres articulateurs aux frontières prosodiques, allongement de certaines unités gestuelles positionnées aux frontières prosodiques).

Nous pouvons donc placer nos trois constructions subordonnées sur un continuum représenté par la figure 7.5, allant d'une plus grande intégration à une plus grande autonomie multimodale.



FIGURE 7.5 – Placement des trois types de subordonnées sur un continuum d'autonomie.

Du côté de la perception, nous avons vu que les frontières prosodiques ne sont pas perçues avec la même intensité en fonction des types syntaxiques : les appositives sont significativement associées à la perception de frontières fortes, les circonstancielles à la perception de frontières faibles, et les relatives déterminatives ne sont pas associées à la perception de frontières. Cette gradation dans la perception des frontières prosodique confirme les résultats de production. En ce qui concerne les interactions des types syntaxiques avec les indices prosodiques de frontière en isolation, nous avons observé que les pauses silencieuses et l'allongement syllabique final sont utilisés par les auditeurs naïfs pour percevoir des frontières dans les trois constructions prosodiques, mais à un degré de force différent.

En ce qui concerne la perception des frontières gestuelles, nous avons vu que les frontières gestuelles ne sont pas non plus perçues avec la même intensité en fonction des types syntaxiques, mais que seules les appositives sont significativement associées à la perception de frontières. De même, seule la condition 3 (réalisation de deux gestes différents en termes de configuration, l'un sur L, l'autre sur SC) a une influence sur la perception de frontières. Les correspondances entre la pertinence de l'indice représenté par la condition 3 (i.e. deux gestes différents en termes de configuration réalisés successivement) en perception et sa fréquence d'occurrence en production sont régulières, et ce avec chaque type syntaxique. Nous avons également observé, à partir de nos résultats, que si la gestualité est bel et bien impliquée dans la perception des frontières, les frontières exprimées de manière gestuelle sont perçues à un moindre degré par rapport aux frontières exprimées de manière vocale.

La majorité des subordonnées permet d'observer que les locuteurs se montrent davantage préoccupés par l'économie et le bon déroulement du discours que par la présentation chronologique des événements : si la pertinence (plus spécifiquement l'idée du niveau de pertinence construite par le locuteur de son propre discours), pour mener à bien les divers objectifs discursifs, détermine le choix de l'ordre des énoncés (les subordonnées évoluent globalement dans la localité de l'élément qu'elles modifient), ces constructions tendent également à l'établissement d'un consensus entre locuteur et co-locuteur (qu'il soit de nature référentielle, métadiscursive, discursive, ou interactionnelle), par lequel progresse le discours.

Plus que l'observation selon laquelle les constructions subordonnées permettent d'effectuer des ajouts et des modifications en temps réel à des niveaux référentiels, métadiscursifs, discursifs, pragmatiques, et/ou interactionnels en conservant le tour de parole et l'activité interactionnelle en cours, le fait que ces structures puissent

à la fois porter sur le propositionnel, le textuel et les connaissances partagées des participants montre que le sens est négocié dans la co-construction du discours.

La forme courte prototypique des constructions subordonnées, plus ou moins séparées de leur entourage selon les types syntaxiques, est en adéquation avec les besoins et contraintes discursives dans l'espace et le contexte interactionnels. La progression du discours et de l'échange s'appuie sur des compromis aussi bien conceptuels que pragmatiques obtenus par des ajustements, inscrits dans l'économie du discours spontané grâce à cette forme. Ces compromis sont réalisés de manière émergente et multimodale dès lors que l'on considère le discours à la fois comme le produit et la base d'un processus de communication dynamique, modulaire, et composite.

Par notre relevé d'indices, une tendance se forme également en ce qui concerne le marquage d'une rupture : celle-ci s'exprime préférentiellement de manière vocale. Les locuteurs s'appuient sur une majorité d'indices prosodiques pour signaler une rupture de manière plutôt rétrospective (i.e. entre SC et R), et les co-participants s'appuient davantage sur des indices de rupture prosodiques pour percevoir des frontières que sur les indices de rupture gestuels. Les modalités verbales, vocales, et visuelles sont des ressources flexibles et dynamiques dans la construction des subordonnées, et fonctionnent en symbiose. La modalité vocale est préférentiellement utilisée pour ses fonctions démarcatives, alors que la modalité visuelle, où la linéarité est moins contrainte que dans le discours oral en s'inscrivant non seulement dans le temps mais aussi dans l'espace, est surtout exploitée pour ses capacités référentielles, en indiquant la nature et la portée des constructions subordonnées dans le discours.

Une approche en contexte a permis de mieux saisir le fonctionnement de ces types de constructions, dont la pertinence s'exprime au niveau de la structure du discours plutôt qu'au niveau phrastique. Cette observation suggère que la valeur des constructions subordonnées se saisit dès lors que l'on approche le discours en tant que construction dynamique et composite, se constituant dans l'interaction, à travers la mobilisation de ressources communicationnelles variées afin de construire le sens.

L'exploration des constructions subordonnées nécessite donc une préoccupation analytique multiple, concernant les processus de référence et de référenciation entendus comme l'élaboration de versions publiques du monde, la configuration émergente du discours et son organisation séquentielle, ainsi que la régulation de l'interaction.

L'étude de la subordination dans la conversation spontanée d'un point de vue multimodal fournit de nouvelles perspectives sur la production discursive, particulièrement sur la flexibilité de la planification et de la modélisation du discours. Davantage d'information sur la production discursive en temps réel avec un focus particulier sur les frontières contribue également aux champs d'études du traitement automatique du langage naturel et de l'acquisition du langage (Gilbert et al., 2016).

Nous avons évoqué notre ancrage dans la grammaire des constructions et la grammaire cognitive dans le Chapitre 1, tout en abordant quelques limites de ces approches, essentiellement liées à la représentation et à la modélisation de données multimodales de la conversation spontanée. Par son analyse multimodale de plusieurs types syntaxiques de constructions subordonnées, ce travail représente un apport théorique, dans le sens où nous proposons une approche multiparamétrique pour la représentation des subordonnées en conversation spontanée. Lorsqu'on analyse la façon dont des unités sémiotiques forment de plus grandes séquences d'action dans le discours et la conversation, la parole spontanée montre à la fois des chaînes complexes de structures emboîtées les unes dans les autres, et des interruptions où les parties du discours ne se suivent plus. Les constructions subordonnées introduisent une rupture lorsqu'elles ouvrent un cadre interprétatif différent de celle du groupe intonatif précédent. Alors que cette rupture peut être directement exprimée avec des moyens syntaxiques ou discursifs, la prosodie crée une rupture pendant la production de la subordonnée avec des traits rythmiques ou des rehaussements intonatifs, qui signalent que les éléments précédents doivent être recontextualisés. Afin d'éviter un décalage entre les représentations du co-participant et l'apport du locuteur, la gestualité donne des instructions d'ordre pragmatique à propos de la valeur informationnelle du contenu propositionnel (e.g. battements manuels, haussements de sourcils).

Nous proposons donc un modèle multiparamétrique des constructions subordonnées basé sur le fait que ces constructions sont des pratiques interactionnelles qui offrent une reconstruction interprétative du discours. Les différents types syntaxiques de subordonnées font référence à différents cadres interprétatifs dans l'expérience du locuteur (schémas 7.6, 7.7, et 7.8). La sélection de ces schémas interprétatifs est exprimée par la quantité différente d'indices de rupture prosodiques et gestuels produits en co-occurrence avec chaque type syntaxique de construction subordonnée.

Les relatives déterminatives marquent une identité des cadres interprétatifs entre

la proposition principale et la subordonnée (schéma 7.6 associé à l'exemple (1))⁶, sans rupture (à part quelques éléments rythmiques principalement à la frontière droite de la subordonnée) exprimée par la prosodie et la gestualité entre la subordonnée et le co-texte.

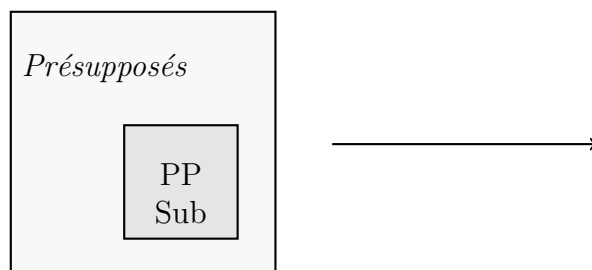


FIGURE 7.6 – Relatives déterminatives : identité des cadres interprétatifs pour la proposition principale (PP) et la construction subordonnée (Sub), avec une seule continuation possible (représentée par la flèche) après la production de la subordonnée.

- | | | | |
|-----|----------|----|--|
| (1) | Michelle | | who's she again # |
| | Zoe | | <u>Felicity's that woman that#</u> |
| | Michelle | L | oh [(a) she's that # woman |
| | | SC | that #] looks after [(b) the Nottingham crowd] (laughs) |
| | Zoe | | <u>looks after us</u> |
| | Michelle | R | [(c) that woman #] |

Dans l'exemple (1), Zoe vient de mentionner un référent ("Felicity"), qu'elle présuppose connu et facilement activable dans la mémoire de sa co-locutrice, Michelle. Michelle explicite le contraire, en demandant des informations sur l'identité de ce référent. Bien que SC permette une sélection plus précise du référent de L ("that woman"), et que l'apport de SC dans le discours ne puisse pas se résumer à une identité des présupposés, les deux propositions ("she's that woman" et "she looks after the Nottingham crowd") présupposent l'existence d'autres femmes à la connaissance des deux participantes. Michelle répète d'ailleurs "that woman" après la complétion de la relative déterminative, ce qui montre que la ré-analyse du segment de L provoquée par l'apport de la relative de SC lui permet d'accéder à une signification différente. L'information de SC peut alors être appréhendée comme une information différente de celle de L, tout en restant dans le même cadre interprétatif. Hormis les pauses silencieuses intra-constituant de Michelle, liées à la situation interactionnelle et plus précisément à la conservation de son tour de parole, aucun indice de rupture n'est utilisé par la locutrice avant la complétion de la

6. Les items verbaux soulignés représentent les chevauchements de parole, et les crochets représentent les unités gestuelles.

relative déterminative. Michelle produit un geste déictique sur L (a) qu'elle maintient, agrémenté d'un battement manuel en chevauchement sur L et SC. Ces deux segments sont donc marqués comme appartenant à la même unité cognitive.

Les circonstancielles, quant à elles, élargissent le cadre interprétatif de la proposition principale (schéma 7.7 associé à l'exemple (2)), avec un plus grand indice de rupture en prosodie (variations de hauteur F0 intra-constituant, davantage de modulations, pauses remplies produites à leur suite) et en gestualité (gestes manuels exploitant une trajectoire opposée à celle du co-texte précédent) que dans les relatives déterminatives. Toutefois, l'une de leurs spécificités concerne la propension des gestes produits en co-occurrence avec les circonstancielles à se trouver également en chevauchement avec les groupes intonatifs du co-texte. L'espace cognitif ouvert par ces gestes a donc tendance à être plus grand que celui des unités prosodiques et syntaxiques. Cet espace plus large peut tout aussi bien concerner le co-texte gauche que le co-texte droit.

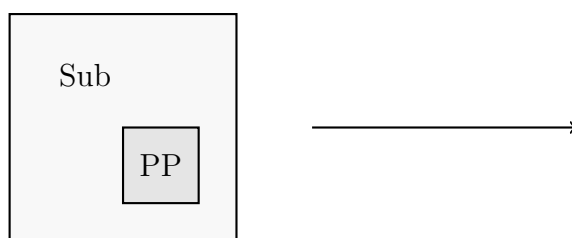


FIGURE 7.7 – Circonstancielles : la subordonnée (Sub) élargit le cadre interprétatif par rapport à celui de la proposition principale (PP), avec une seule continuation possible (représentée par la flèche) après la production de la subordonnée.

- (2) Rhianna L i tried [(a) driving once] in her car
 SC **when we were on a # [(b) little road in the countryside] #**
 R and hem (swallows) she said well turn left #

Dans l'exemple (2), Rhianna commence un récit. SC spécifie les circonstances dans lesquelles la relation prédicative “I tried driving once in her car” est réalisée, en localisant dans le temps la situation exprimée par le verbe et ses compléments. Ses éléments référentiels sont stabilisés dans le sens où leur portée est définie par SC. Le cadre interprétatif de L se trouve élargi par celui de SC. Au niveau prosodique, SC montre plus de modulations que L : la subordonnée est caractérisée par davantage de mouvements mélodiques. Le début de R est caractérisé par une combinaison de pauses remplies (“and” et “hem”), suggérant un coût cognitif plus élevé pour la projection de ce segment. Rhianna réalise un geste métaphorique emphatique (a) sur L, insistant sur le caractère exceptionnel de la situation en corrélation avec l’adverbe

“once”. Rhianna produit ensuite un geste iconique assez large (b) à l’aide de ses deux mains positionnées de façon symétrique en co-occurrence avec “little road on the countryside”. Elle trace deux droites parallèles devant elle, représentant la petite route de campagne sur laquelle sa voiture circulait. Rhianna se positionne donc à l’intérieur de son récit en tant que participante, contrairement à L où Rhianna se concentrait sur sa position d’énonciatrice, ne décrivant pas d’événement mais donnant des informations à propos de ces derniers. Le regard de Rhianna centre d’ailleurs l’attention sur ce geste, prenant une fonction déictique. Alors que R décrit le premier événement déclenchant la complication du récit (Rhianna doit changer de direction), ce segment n’est accompagné d’aucun geste de Rhianna; le cadre gestuel défini par SC, bien que toujours valable pour l’interprétation de R et des segments suivants, n’est pas maintenu.

Enfin, les appositives ouvrent un nouveau cadre interprétatif par rapport à celui de la principale (schéma 7.8 associé à l’exemple (3)), ce qui est marqué par un indice de rupture prosodique (pauses silencieuses inter-constituant, production sous des groupes intonatifs distincts, durée et débit distinctifs, allongement syllabique très significatif et contour descendant sur L, abaissement de registre) et gestuel (production d’unités gestuelles propres aux appositives sans chevauchement avec le co-texte, haussements de sourcils, changements de la direction du regard du locuteur) supérieur. Deux continuations sont possibles après la production d’une appositive : soit le cadre interprétatif de la proposition principale est repris, soit le cadre interprétatif ouvert dans la subordonnée est continué.

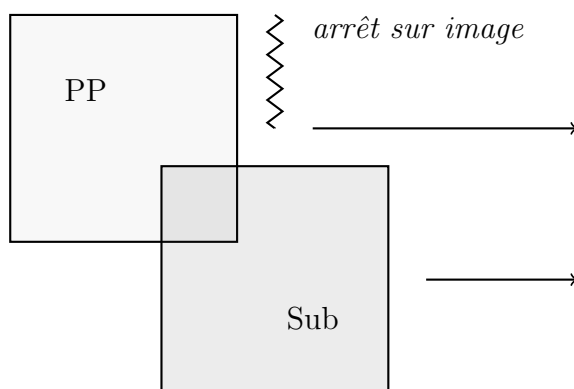


FIGURE 7.8 – Appositives relatives : la subordonnée (Sub) ouvre un nouveau cadre interprétatif par rapport à celui de la proposition principale (PP). Les deux cadres sont en léger chevauchement grâce au marqueur morphologique “which” permettant le basculement d’un cadre à l’autre. Deux continuations sont possibles après la production de l’appositive : soit le cadre interprétatif de la proposition principale est repris, soit le cadre interprétatif ouvert dans la subordonnée est continué.

- (3) Rhianna L i mean [(a) my mum's pushing] me to get my license
 SC **(h) uh which [(b) i guess i should] #**
 Alex it's a good thing to have
 Rhianna R [(c) (h) but #
 Alex if you ever need it i mean
 Rhianna well first of all #
 for the moment

Rhianna explique ici les raisons pour lesquelles elle ne veut pas apprendre à conduire. Rhianna commence par évoquer des opinions adverses à la sienne, en leur reconnaissant une certaine valeur. L est un argument à part entière : sa mère voudrait qu'elle passe son permis. Rhianna marque cette information par un geste iconique (a), sa main droite balayant l'espace devant elle, correspondant à l'élément verbal "pushing". Rhianna sépare SC de L par une reprise de souffe audible (h). Le commentaire de SC n'est pas lié à la volonté de sa mère, mais revient sur l'élément rhématique de L ("get my license"). Rhianna s'exprime en tant qu'énonciatrice : elle reconnaît la validité des conseils de sa mère et marque cette concession par un geste manuel métaphorique (b), paume vers le haut. Elle hausse également les sourcils. Ce haussement complète sa prise de position modale et marque le caractère contrastif de SC en tant que mouvement discursif. L et SC représentent deux cadres interprétatifs différents. Le groupe intonatif de SC est également plus court que le reste de la séquence : le commentaire concessif de Rhianna portant sur le prédicat de L est réalisé avec un changement de rythme. Rhianna revient sur la ligne principale de son argumentation sur R. L'effet du changement de rythme occasionné par SC est visible sur R, où Rhianna insère une pause silencieuse intra-constituant assimilable à une pause d'hésitation entre deux marqueurs de discours résumptifs ("but" et "well"), après la reprise de souffe audible. Elle se montre beaucoup plus catégorique dans ce segment : alors qu'elle acquiesce et semble plier sur SC, elle accompagne ce segment d'un geste de négation continu de la tête (c)). Cette séquence est alors caractérisée par deux positions modales successives, qui ne sont pas de même intensité : l'énonciatrice de R est plus forte que celle de SC. Cette asymétrie reflète la structure discursive, puisque R continue l'agenda séquentiel de Rhianna alors que SC n'y contribue pas. La cadre interprétatif ouvert sur SC, qui présente Rhianna en tant qu'énonciatrice prenant une position modale sur les arguments de sa mère, n'est pas continué. Rhianna est de retour dans l'argumentation en elle-même sur R, et présente ses arguments en opposition à ceux du parti adverse.

En ce qui concerne notre ancrage dans la grammaire des constructions, la problématique de cette thèse permet également d'avancer une représentation multimodale plus formelle de nos subordonnées, compatible avec le modèle ancré dans

la grammaire cognitive. La grammaire des constructions permet la représentation des différentes unités (i.e. constructions) verbales, vocales, et gestuelles à l'oeuvre dans une séquence contenant une subordonnée. Nous sommes désormais en mesure de montrer que le réel intérêt d'une telle représentation réside dans le fait que les interactions et chevauchements entre les constructions ne sont pas les mêmes en fonction du type syntaxique de subordonnée.

La représentation des relatives déterminatives pourrait être celle du graphe 7.9. Les graphes 7.9, 7.10, et 7.11 sont adaptés de ceux de Ferré (à paraître), qui ont été réalisés à partir de nos données. Dans les séquences contenant une relative déterminative, la prosodie et la gestualité lient le co-texte gauche à la subordonnée, car les deux segments partagent les mêmes unités (groupe intonatif GI, et unité gestuelle UG). En revanche, l'emplacement 3, qui marque le temps de fin de la subordonnée, constitue une frontière intonative et gestuelle. Le locuteur peut changer d'unité discursive (de SDa à SDb), ou peut continuer l'unité discursive en cours (SD).

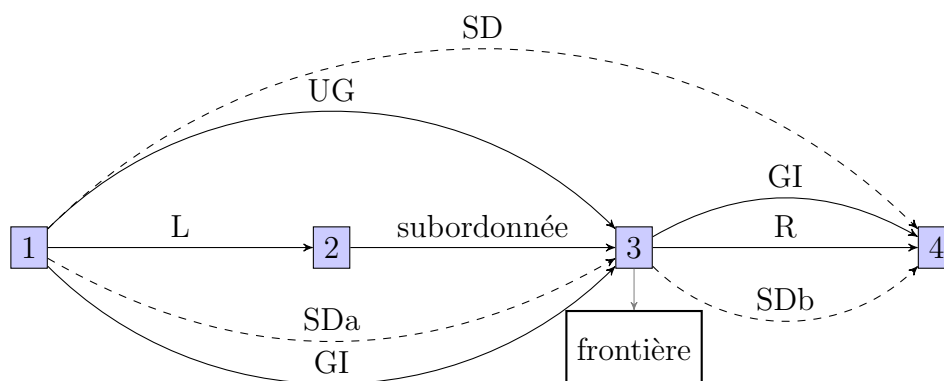


FIGURE 7.9 – Graphe représentant les unités prosodiques, gestuelles et discursives pour les propositions relatives (GI : groupe intonatif; UG : unité gestuelle; SD : segment de discours). Les emplacements 1, 2, 3, et 4 représentent des emplacements potentiels pour la production d'une frontière, qu'elle soit prosodique, syntaxique, discursive, et/ou gestuelle.

Les circonstancielles, quant à elles, montrent une organisation différente. Cette organisation est explicitée dans le modèle 7.10. Dans les séquences contenant une circonstancielle, le co-texte gauche et la subordonnée sont réalisées sous des unités intonatives distinctes. En revanche, il n'est pas rare que les unités gestuelles soient produites en chevauchement sur plusieurs propositions syntaxiques, soit avec le co-texte gauche (UGa), soit avec le co-texte droit (UGb). Comme dans les séquences contenant une relative déterminative, le locuteur peut changer d'unité discursive (de

SDa à SDb), ou peut continuer l'unité discursive en cours (SD).

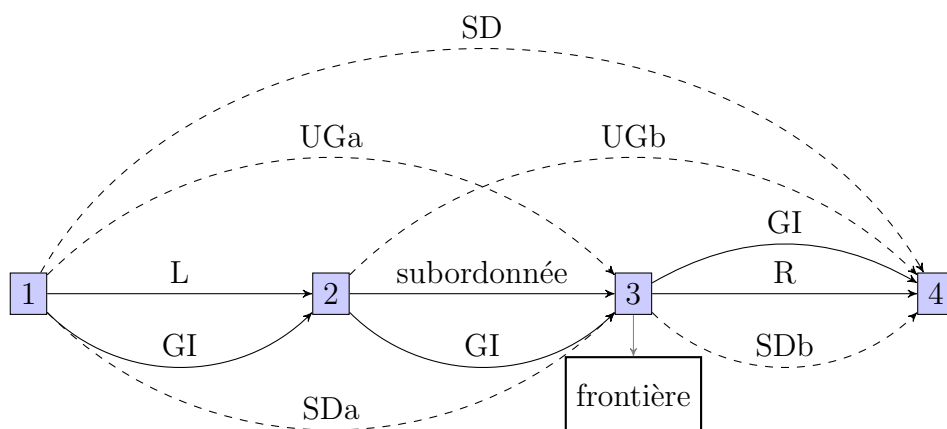


FIGURE 7.10 – Graphe représentant les unités prosodiques, gestuelles et discursives pour les circonstanciées (GI : groupe intonatif; UG : unité gestuelle; SD : segment de discours).

Enfin, les appositives peuvent être représentées par le modèle 7.11. Les emplacements 2 et 3 constituent des frontières prosodiques et gestuelles. Le fait que les subordonnées soient réalisées avec leur propre contour intonatif et leur propre unité gestuelle multiplie les possibilités de regroupements discursifs. Ces stratégies sont représentées par les lignes pointillées dans le graphe.

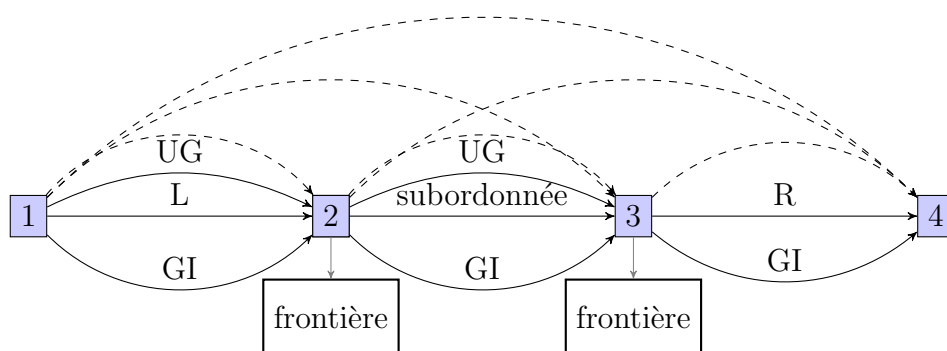


FIGURE 7.11 – Graphe représentant les unités prosodiques, gestuelles et discursives pour les appositives (GI : groupe intonatif; UG : unité gestuelle; SD : segment de discours).

Les modèles proposés ci-dessus sont basés sur l'étude des frontières produites dans les séquences contenant des subordonnées. Après avoir placé nos trois types de constructions subordonnées sur un continuum entre intégration et autonomie par

l'analyse des frontières dans les différentes modalités, il convient, dans les études suivantes, de dépasser la chaîne linéaire de l'organisation discursive pour nous intéresser à l'axe caractérisé par les choix et les contraintes thématiques à l'arrière de cette linéarité. Les constructions subordonnées doivent être analysées du point de vue de leur structure informationnelle et de leur poids communicatif, afin de déterminer si les subordonnées délivrent toutes, comme le stipule la littérature (e.g. Reinhart, 1984; Tomlin, 1985; Lambrecht, 1996) de l'information d'arrière-plan. Cette question de recherche devra être adressée à la lumière d'une liste d'indices multimodaux de premier plan informationnel. Nous émettons dès à présent l'hypothèse selon laquelle les différents types syntaxiques de subordonnées ne véhiculent pas tous de l'arrière-plan informationnel, et peuvent faire l'objet d'une gradation dans l'expression du premier plan informationnel. En outre, cette nouvelle ligne de recherche sur un mode d'organisation discursive thématique permettrait de montrer que l'apport informationnel n'est pas corrélé à l'autonomie d'une construction subordonnée, puisque nous émettons également l'hypothèse selon laquelle la gradation des types syntaxiques selon leur apport informationnel n'est pas la même que la gradation des types syntaxiques selon leur autonomie.

Bibliographie

- Aarts, B. et McMahon, A. 2008. *The handbook of English linguistics*. Wiley-Blackwell, Oxford.
- Adamou, E., Bergounioux, G., Bertrand, R., Bolly, C., Boutet, D., Chanard, C., Etienne, C., Jisa, H., et Parisse, C. 2014. IRCOM, Consortium Corpus Oraux & Multimodaux de l'IR-CORPUS.
- Aelbrecht, L., Haegeman, L., et Nye, R., eds. 2012. *Main clause phenomena. New horizons*. John Benjamins, Amsterdam.
- Al Moubayed, S., Beskow, J., Granström, B., et House, D. 2011. Audio-visual prosody: perception, detection, and synthesis of prominence. In Esposito, A., Esposito, M., Martone, R., Müller, V. C., et Scarpetta, G., eds, *Proceedings of the Third COST 2102 International Training School*, pages 55–71, Caserta, Italie. Springer.
- Al Moubayed, S. et Beskow, J. 2009. Effects of visual prominence cues on speech intelligibility. In *Proceedings of the International Conference on Auditory Visual Speech Processing AVSP*, pages 43–46, Norwich, Royaume-Uni.
- Amir, N., Silber-Varod, V., et Izre'el, S. 2004. Characteristics of Intonation Unit Boundaries in Spontaneous Spoken Hebrew – Perception and Acoustic Correlates. In *Proceedings of Speech Prosody 2004*, pages 1–4, Nara, Japon. ISCA.
- Argyle, M. et Cook, M. 1976. *Gaze and mutual gaze*. Cambridge University Press, Londres.
- Arnold, D. et Borsley, R. D. 2008. Non-restrictive Relative Clauses, Ellipsis and Anaphora. In *Proceedings of the HPSG08 Conference*, pages 5–25, Stanford, CA, USA. CSLI Publications.
- Asher, N. et Lascarides, A. 2003. *Logics of conversation*. Cambridge University Press, New York.
- Astruc-Aguilera, L. et Nolan, F. 2007. Variation in the intonation of extra-sentential elements. In Prieto, P., Mascaro, J., et Sole, M.-J., eds, *Segmental and Prosodic Issues in Romance Phonology*, pages 85–107. John Benjamins, Amsterdam.
- Astésano, C., Bertrand, R., Espesser, R., et Nguyen, N. 2012. Perception des frontières et des proéminences en français. In *Proceedings of JEP-TALN-RECITAL*, pages 1–5, Grenoble, France.
- Atterer, M., Baumann, T., et Schlangen, D. 2008. Towards incremental end-of-utterance detection in dialogue systems. In *Proceedings of the International*

- Conference on Computational Linguistics*, pages 11–14, Manchester, Royaume-Uni. ACM.
- Auer, P. 2005. Projection in interaction and projection in grammar. *Text*, 25(1):7–36.
- Auran, C., Bouzon, C., et Hirst, D. 2004. The AixMARSEC project: an evolutive database of spoken English. In *Proceedings of Speech Prosody 2004*, pages 561–564, Nara, Japon. ISCA.
- Auran, C., Colas, A., Portes, C., et Vion, M. 2005. Perception of breaks and discourse boundaries in spontaneous speech: developping an on-line technique. In *Proceedings of IDP05*, pages 1–7, Aix-en-Provence, France.
- Auran, C. et Loock, R. 2011. The prosody of discourse functions: the case of Appositive Relative Clauses in spoken British English. *Corpus Linguistics and Linguistic Theory*, 7(2):181–201.
- Azar, Z., Backus, A., et Ozyurek, A. 2016. Multimodal reference tracking in Dutch and Turkish discourse: Role of culture and typological differences. In *Proceedings of ISGS7*, pages 1–5, Paris, France.
- Bangerter, A. et Clark, H. H. 2003. Navigating joint projects with dialogue. *Cognitive Science*, 27(2):195–225.
- Barkhuysen, P., Kraemer, E., et Swerts, M. 2008. The interplay between the auditory and visual modality for end-of-utterance detection. *The journal of the Acoustical Society of America*, 123(1):354–365.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., et Wilson, C. D. 2003. Grounding conceptual knowledge in modality-specific systems. *TRENDS in Cognitive Sciences*, 7(2):84–91.
- Barsalou, L. W. 2008. Grounded cognition. *The Annual Review of Psychology*, 59:617–645.
- Barth-Weingarten, D. 2016. *Intonation Units Revisited. Cesuras in talk-in-interaction*. John Benjamins, Amsterdam.
- Bates, D., Maechler, M., Bolker, B., et Walker, S. 2017. Linear mixed-effects models using eigen and s4.
- Bavelas, J. B., Chovil, N., Coates, L., et Roe, L. 1995. Gestures Specialized for Dialogue. *Personality and Social Psychology Bulletin*, 21:394–405.
- Bavelas, J. B., Coates, L., et Johnson, T. 2000. Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79(6):941–952.
- Beattie, G. 1978. Sequential Temporal Patterns of Speech and Gaze in Dialogue. *Semiotica*, 23(1-2):29–52.
- Beaugendre, F. 1994. *Une Etude Perceptive de l'Intonation du Français. Développement d'un modèle et application à la génération automatique de l'intonation pour un système de synthèse à partir du texte*. Thèse de doctorat d'état, Université de Paris XI Orsay, France.
- Beckman, M. E., Hirschberg, J., et Shattuck-Hufnagel, S. 2005. The original ToBI system and the evolution of the ToBI framework. In Jun, S.-A., ed, *Prosodic Typology. The Phonology of Intonation and Phrasing*, pages 9–54. Oxford University Press, New York.

- Beckman, M. E. et Pierrehumbert, J. B. 1986. Intonational structure in Japanese and English. *Phonology*, 3(01):255–309.
- Benus, S., Enos, F., Hirschberg, J., et Shriberg, E. 2006. Pauses in deceptive speech. *Speech Prosody*, 18:2–5.
- Benus, S., Gravano, A., et Hirschberg, J. 2007. The prosody of backchannels in American English. In *Proceedings of ICPHS XVI*, pages 1065–1068, Sarrebruck, Allemagne.
- Benzitoun, C., Dister, A., Gerdes, K., Kahane, S., et Marlet, R. 2009. Annoter du des textes tu te demandes si c'est syntaxique tu vois. In *Proceedings of the 28th International Conference on Lexis and Grammar (LGC 2009)*, volume 4, pages 16–27, Bergen, Norvège. Presses de l'Université de Bergen.
- Benzitoun, C. 2013. Description des séquences introduites par quand en français parlé. In Debaisieux, J.-M., ed, *Analyses linguistiques sur corpus: subordination et insubordination en français*, pages 249–292. Lavoisier, Paris.
- Bernardis, P. et Gentilucci, M. 2006. Speech and gesture share the same communication system. *Neuropsychologia*, 44(2):178–190.
- Berrendonner, A. 2008. Pour une praxéologie des parenthèses. *Verbum Revue De Linguistique*, 30(1):5–23.
- Bertrand, R., Ferré, G., Blache, P., Espesser, R., et Rauzy, S. 2007. Backchannels revisited from a multimodal perspective. In *Proceedings of Auditory-visual Speech Processing*, pages 1–5, Hilvarenbeek, Pays-Bas.
- Bertrand, R., Ader, M., Blache, P., Ferré, G., Espesser, R., et Rauzy, S. 2009. Représentation, édition et exploitation de données multimodales: le cas des backchannels du corpus CID. *Cahiers de Linguistique*, 33(2):183–212.
- Bertrand, R., Ferré, G., et Guardiola, M. 2013. French face-to-face interaction: repetition as a multimodal resource. In Rojc, M. et Campbell, N., eds, *Coverbal Synchrony in Human-Machine Interaction*, pages 141–172. Taylor and Francis, Boca Raton, FL.
- Beskow, J., Granström, B., et House, D. 2006. Visual correlates to prominence in several expressive modes. In *Proceedings of INTERSPEECH 2006*, pages 1272–1275, Pittsburgh, USA.
- Bestgen, Y. 2009. The discourse functions of sentence-initial adverbials: studies in comprehension. In *Proceedings of Linguistic & Psycholinguistic Approaches to Text Structuring*, pages 7–13, Paris. Ecole Normale Supérieure.
- Biau, E., Fernandez, L. M., Holle, H., Avila, C., et Soto-Faraco, S. 2016. Hand gestures as visual prosody: BOLD responses to audio-visual alignment are modulated by the communicative nature of the stimuli. *NeuroImage*, 132:129–137.
- Biau, E., Fromont, L., et Soto-Faraco, S. 2018. Beat Gestures and Syntactic Parsing: An ERP Study. *Language Learning*, 68(1):102–126.
- Biau, E. et Soto-Faraco, S. 2013. Beat gestures modulate auditory integration in speech perception. *Brain and language*, 124(2):143–152.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., et Quirk, R. 1999. *Longman Grammar of Spoken and Written English*. Pearson Education, Londres.

- Bigi, B. et Hirst, D. 2012. SPeech Phonetization Aligment and Syllabification (SPPAS): a tool for the automatic analysis of speech prosody. In *Proceedings of Speech Prosody 2012*, pages 19–22, Shanghai, Chine. Tongji University Press.
- Bigi, B. 2012. SPPAS: a tool for the phonetic segmentation of Speech. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2012)*, pages 1748–1755, Istanbul, Turquie.
- Bird, S. et Liberman, M. 2001. A formal framework for linguistic annotation. *Speech Communication*, 33(1-2):23–60.
- Birdwhistell, R. R. 1970. *Kinesics and context*. University of Pennsylvania Press, Philadelphie.
- Blaauw, E. 1994. The Contribution of Prosodic Boundary Markers to the Perceptual Difference between Read and Spontaneous Speech. *Speech Communication*, 14(4):359–375.
- Blache, P. 2016. Representing syntax by means of properties: a Formal framework for descriptive approaches. *Journal of Language Modelling*, 4(2):183–224.
- Blakemore, D. 2006. Divisions of labour: The analysis of parentheticals. *Lingua*, 116(10):1670–1687.
- Blakemore, D. 2009. On the relevance of parentheticals. In *Proceedings of IDP 09*, pages 9–17, Paris, France.
- Blanche-Benveniste, C. 1990. Un modèle d’analyse syntaxique “en grilles” pour les productions orales. *Anuario de Psicología*, 47:11–28.
- Blanche-Benveniste, C. 1997. La notion de variation syntaxique dans la langue parlée. *Langue Française*, 115:19–29.
- Blühdorn, H. 2008. Subordination and coordination in syntax, semantics and discourse. In Fabricius-Hansen, C. et Ramm, W., eds, “*Subordination*” versus “*Coordination*” in *Sentence and Text. A Cross-Linguistic Perspective*, pages 59–85. John Benjamins, Amsterdam.
- Bock, K. et Cutting, J. C. 1992. Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, 31:99–127.
- Boersma, P. et Weenink, D. Praat: doing Phonetics by Computer. <http://www.fon.hum.uva.nl/praat/>.
- Bolinger, D. 1972. Around the Edge of Language: Intonation. In *Intonation: Selected Readings*, pages 19–29. Penguin, Londres.
- Bolinger, D. 1983. Gesture and intonation. *American Speech*, 58(2):156–174.
- Bolinger, D. 1984. Intonational signals of subordination. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, pages 401–413, Berkeley, CA, USA. eLanguage.
- Bolinger, D. 1989. *Intonation and Its Uses: Melody in Grammar and Discourse*. Stanford University Press, Stanford, CA, USA.
- Borsley, R. D. 1992. More on the Difference between English Restrictive and Non-Restrictive Relative Clauses. *Journal of Linguistics*, 28(1):139–148.
- Breen, M., Fedorenko, E., Wagner, M., et Gibson, E. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7-9):1044–1098.

- Bressem, J. et Ladewig, S. 2011. Rethinking gesture phases: Articulatory features of gestural movement? *Semiotica*, 184:53–91.
- Brinton, L. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. Cambridge University Press, Cambridge.
- Brône, G. et Sambre, P. 2015. Gestural specification and completion in multimodal construction grammar : A case study of instrumental causal actions involving Cut and Break. In *ICLC 13*, Newcastle-Upon-Tyne, Royaume-Uni.
- Brown, P. et Levinson, S. 1987. *Politeness: Some universals in language usage*, volume 4. Cambridge University Press, Cambridge, Royaume-Uni.
- Brown, G. et Yule, G. 1983. *Discourse Analysis*. Cambridge University Press, Cambridge.
- Brugos, A., Shattuck-Hufnagel, S., et Veilleux, N. 2006. Transcribing Prosodic Structure of Spoken Utterances with ToBI. Tutorial en ligne : <http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-with-tobi-january-iap-2006/>.
- Buchstaller, I. 2001. He goes and I'm like: The new quotatives re-visited. In *Proceedings of the 30th annual meeting on new ways of analyzing variation (NWA 30)*, pages 11–14, Raleigh, NC, USA.
- Buhmann, J., Caspers, J., van Heuven, V., Hoekstra, H., Martens, J.-P., et Swerts, M. 2002. Annotation of prominent words, prosodic boundaries and segmental lengthening by non-expert transcribers in the Spoken Dutch Corpus. In *Proceedings of LREC 2002*, pages 779–785, Las Palmas, Espagne. ELRA.
- Burton-Roberts, N. 1975. Nominal apposition. *Foundations of Language*, 13:391–419.
- Burton-Roberts, N. 1999. Language, linear precedence and parentheticals. In Collins, P. et Lee, D., eds, *The Clause in English: in honour of Rodney Huddleston*, pages 33–51. John Benjamins, Amsterdam.
- Bybee, J. 2008. Usage-based grammar and second language acquisition. In Robinson, P. et Ellis, N. C., eds, *A handbook of Cognitive Linguistics and SLA*, pages 216–236. Routledge, Londres.
- Byrd, D. et Saltzman, E. 1998. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, 26:173–199.
- Byrd, D. 2000. Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, 57:3–16.
- Calbris, G. 1985. Geste et parole. *Langue Française*, (68):66–84.
- Calbris, G. 2008. From left to right... : Covernal gestures and their symbolic use of space. In Cienki, A. J. et Müller, C., eds, *Metaphor and Gesture*, volume 3 of *Gesture Studies*, pages 27–53. John Benjamins, Amsterdam.
- Calbris, G. 2011. *Elements of meaning in gesture*, volume 5 of *Gesture Studies*. John Benjamins, Amsterdam.
- Campione, E. et Véronis, J. 2002. A large-scale multilingual study of silent pause duration. In *Proceedings of Speech prosody 2002*, pages 199–202, Aix-en-Provence, France. ISCA.

- Campione, E. et Véronis, J. 2004. Pauses et hésitations en français spontané. In *Actes des 25èmes Journées d'Études sur la Parole (JEP)*, pages 109–112, Fes, Maroc.
- Canac Marquis, R. et Tremblay, M. 1996. The wh- feature and the syntax of restrictive and non-restrictive relatives in French and English. In Lema, J. et Treviño, E., eds, *Theoretical Analyses of Romance Languages*, pages 127–141. John Benjamins, Amsterdam.
- Candea, M. 2000. *Contribution à l'étude des pauses silencieuses et des phénomènes dits "d'hésitation" en français oral spontané*. Thèse de doctorat d'état en sciences du langage : Linguistique et phonétique générale, Paris III - Sorbonne Nouvelle.
- Carlson, L., Marcu, D., et Okurowski, M. E. 2001. Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of the 2nd ACL Workshop on Discourse and Dialogue*, pages 30–39, Aalborg, Danemark.
- Carlson, R. et Swerts, M. 2003. Perceptually based prediction of upcoming prosodic breaks in spontaneous Swedish speech materials. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pages 507–510, Barcelone, Espagne. International Phonetic Association.
- Carston, R. 1993. Conjunction, explanation, and relevance. *Lingua*, 90:27–48.
- Cassell, J., Torres, O. E., et Prevost, S. 1999. Turn taking versus discourse structure. In Wilks, Y., ed, *Machine Conversations*, pages 143–153. Springer, New York.
- Cassell, J., Nakano, Y. I., Bickmore, T. W., Sidner, C. L., et Rich, C. 2001. Non-verbal cues for discourse structure. In *Proceedings of the 39th Annual Meeting on Association for Computational Linguistics*, pages 114–123, Toulouse, France.
- Cassell, J. et McNeill, D. 1990. Gesture and ground. In *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, volume 16, pages 57–68, Berkeley, CA, USA.
- Cavé, C., Guaitella, I., Bertrand, R., Santi, S., Harlay, F., et Espesser, R. 1996. About the relationship between eyebrow movements and Fo variations. In *Proceedings of the Fourth International Conference on Spoken Language (ICSLP 96)*, volume 4, pages 2175–2178, Philadelphia, PA, USA.
- Chafe, W. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View. In Li, C. N., ed, *Subject and Topic*, pages 25–55. Academic Press, New York.
- Chafe, W. 1984. How People Use Adverbial Clauses. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, pages 437–449, Berkeley, CA, USA. Linguistic Society of America.
- Chafe, W. 1988. Linking intonation units in spoken English. In Haiman, J. et Thompson, S. A., eds, *Clause Combining in Grammar and Discourse*, number 18 in *Typological Studies in Language*, pages 1–27. John Benjamins, Amsterdam.
- Chafe, W. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press, Chicago.
- Charolles, M. 2003. De la topicalité des adverbiaux détachés en tête de phrase. *Travaux de Linguistique*, 47(2):11–49.

- Chartrand, T. L. et Bargh, J. A. 1999. The chameleon effect: the perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6):893–910.
- Chemin-Dupontès. 1811. *Analyse grammaticale des phrases françaises graduées*. Gueffier, Paris.
- Chervel, A. 1979. Rhétorique et grammaire : petite histoire du circonstanciel. *Langue Française*, 41:5–19.
- Chevalier, J.-C. 2006. *Histoire de la syntaxe*. Champion, Paris.
- Cho, H. et Hirst, D. 2006. The contribution of silent pauses to the perception of prosodic boundaries in Korean read speech. In *Proceedings of Speech Prosody 2006*, Dresde, Allemagne. ISCA.
- Cho, T. 2006. Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. In Goldstein, L., ed, *Laboratory Phonology 8: Varieties of phonological competence*, pages 519–548. Walter de Gruyter, New York.
- Chomsky, N. 1965. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA.
- Chui, K. 2005. Temporal patterning of speech and iconic gestures in conversational discourse. *Journal of Pragmatics*, 37:871–887.
- Church, R. B., Ayman-Nolley, S., et Mahootian, S. 2004. The role of gesture in bilingual education : does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, 7(4):303–319.
- Cibulka, P. 2015. When the hands do not go home: A micro-study of the role of gesture phases in sequence suspension and closure. *Discourse Studies*, 17(1):3–24.
- Cienki, A. J. et Mittleberg, I. 2013. Creativity in the forms and functions of gestures with speech. In Veale, T., Feyaerts, C., et Forceville, C., eds, *Creativity and the agile mind: A multi-disciplinary study of a multi-faceted phenomenon*, pages 231–252. Mouton de Gruyter, New York, USA.
- Clark, H. H. et Fox Tree, J. E. 2002. Using “uh” and “um” in spontaneous speaking. *Cognition*, 84(1):73–111.
- Clark, H. et Gerrig, R. 1990. Quotations as demonstrations. *Language*, 66(4):764–805.
- Clark, H. H. et Krych, M. A. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1):62–81.
- Clark, H. H. 1996. *Using Language*. Cambridge University Press, Cambridge.
- Clifton, C., Carlson, K., et Frazier, L. 2002. Informative prosodic boundaries. *Language and Speech*, 45(2):87–114.
- Cohen, A. et 't Hart, J. 1967. On the anatomy of intonation. *Lingua*, 19(1-2):177–192.
- Cole, J., Mo, Y., et Baek, S. 2008. The role of syntactic structure in guiding prosody perception with spontaneous speech. In *Experimental and Theoretical Approach to Prosody (ETAP)*, volume 25, pages 2–42, Ithaca, New York. Language and Cognitive Processes.

- Cole, J., Mo, Y., et Baek, S. 2010. The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25(7-9):1141–1177.
- Cole, J., Mahrt, T., et Hualde, J. I. 2014. Listening for sound, listening for meaning: Task effects on prosodic transcription. In *Proceedings of Speech Prosody 2014*, pages 859–863, Dublin, Irlande. ISCA Archive.
- Collier, R., de Pijper, J. R., et Sanderman, A. 1993. Perceived prosodic boundaries and their phonetic correlates. In *Proceedings of the workshop on Human Language Technology*, pages 341–345, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Combettes, B. 2005. Les constructions détachées comme cadres de discours. *Langue Française*, (4):31–44.
- Condon, W. 1976. An analysis of behavioral organization. *Sign Language Studies*, 13.
- Corminboeuf, G. et Benzitoun, C. 2014. Evaluation critique des modèles graduels et non graduels de l'intégration syntaxique. *Corela [en ligne]*, 12(1).
- Corum, C., Smith-Stark, T. C., et Weiser, A., eds. 1973. *You take the high node and I'll take the low node: Papers from the Comparative Syntax Festival. The differences between main and subordinate clauses*. Chicago Linguistic Society, Chicago, USA.
- Cosnier, J. 1988. Grands tours et petits tours. In Cosnier, J., Gelas, N., et Kerbrat-Orecchioni, C., eds, *Echanges sur la Conversation*, pages 175–184. Editions du CNRS, Paris.
- Cotte, P. 1998. *L'Explication grammaticale de textes anglais*. Presses Universitaires de France, Paris.
- Cotte, P. 2008. Les propositions relatives et l'énonciation. *Cycnos [en ligne]*, 17 Spécial.
- Couper-Kuhlen, E. et Thompson, S. A. 2009. Conditionality and Conversational Complaining. Paper presented at the Department of Finnish Language and Literature. Helsinki, Finlande.
- Couper-Kuhlen, E. 1986a. Intonation and grammar. In *An Introduction to English Prosody*, pages 139–157. Max Niemeyer Verlag, Tübingen.
- Couper-Kuhlen, E. 1986b. Intonation and text/discourse. In *An Introduction to English Prosody*, pages 188–209. Max Niemeyer Verlag, Tübingen.
- Cristofaro, S. 2003. *Subordination*. Oxford University Press, Oxford.
- Croft, W. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press, Oxford.
- Crompton, P. 2009. Initial positioning of adverbials and discourse structure. In *Proceedings of Linguistic & Psycholinguistic Approaches to Text Structuring*, pages 18–24, Paris, France. Ecole Normale Supérieure.
- Cruttenden, A. 1986. *Intonation*. Cambridge University Press, Cambridge.
- Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge University Press, Cambridge.

- Culioli, A. 1999. *Pour une linguistique de l'énonciation: Formalisation et opérations de repérage*, volume 2 of *L'Homme dans la Langue*. Ophrys, Paris.
- Dancygier, B. et Sweetser, E. 2000. Constructions with if, since, and because: causality, epistemic stance, and clause order. In Couper-Kuhlen, E. et Kortmann, B., eds, *Cause, Condition, Concession, Contrast*, pages 111–142. Mouton de Gruyter, Berlin.
- Dancygier, B. 1998. *Conditionals and Prediction : Time, Knowledge, and Causation in Conditional Constructions*. Cambridge University Press, Cambridge.
- De Beaugrande, R.-A. et Dressler, W. U. 1981. *Introduction to text linguistics*, volume 1. Longman, Londres.
- de Fribourg, G. 2012. *Grammaire de la période*. Peter Lang, Berne.
- De Jong, K., Beckman, M. E., et Edwards, J. 1993. The interplay between prosodic structure and coarticulation. *Language and Speech*, 36:197–212.
- De Kok, I. et Heylen, D. 2009. Multimodal end-of-turn prediction in multi-party meetings. In *Proceedings of the 2009 International Conference on Multimodal interfaces*, pages 91–98, New York, USA. ACM.
- de Pijper, J. R. et Sanderman, A. A. 1994. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. *The Journal of the Acoustical Society of America*, 96(4):2037–2047.
- de Ruiter, J. P. 2007. Postcards from the mind: The relationship between speech, imagistic gesture, and thought. *Gesture*, 7(1):21–38.
- De Vries, M. 2002. *The syntax of relativization*. LOT, Utrecht.
- De Vries, M. 2006. The syntax of appositive relativization: On specifying coordination, false free relatives, and promotion. *Linguistic Inquiry*, 37(2):229–270.
- De Vries, M. 2007. Invisible constituents? Parentheses as B-merged adverbial phrases. In Dehé, N. et Kavalova, Y., eds, *Parentheticals*, pages 203–234. John Benjamins, Amsterdam.
- Debaisieux, J.-M. 2006. La distinction entre dépendance grammaticale et dépendance macrosyntaxique comme moyen de résoudre les paradoxes de la subordination. *Coordination et Subordination: typologie et modélisation*, (28):119–132.
- Debaisieux, J.-M. 2016. Toward a global approach to discourse uses of conjunctions in spoken French. *Language Sciences*, 58:79–94.
- Debras, C. 2017. The shrug : forms and meanings of a compound enactment. *Gesture*, 16(1):1–34.
- Debras, C. 2018. Complémentarité des corpus et des méthodes pour l'analyse linguistique des gestes. *Corpus*, 18:[en ligne].
- Debreslioska, S., Özyürek, A., Gullberg, M., et Perniss, P. 2013. Gestural viewpoint signals referent accessibility. *Discourse Processes*, 50(7):431–456.
- Dehé, N. 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics*, 45(3):569–615.
- Delais-Roussarie, E., Yoo, H., et Post, B. 2011. Quand frontières prosodiques et frontières syntaxiques se rencontrent. *Langue Française*, 170(2):29–44.

- Delais-Roussarie, E. et Choi-Jonin, I. 2004. Existent-ils des des indices intonatifs de segmentation en unités macro-syntaxiques? In *Actes des Journées d'Etude sur la Parole 2004*, pages 149–152, Fes, Maroc.
- Delais-Roussarie, E. et Post, B. 2008. Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche. In *Actes des Journées d'Etude sur la Parole JEP-TALN 8 [en ligne]*, Avignon, France. AFCP.
- Delais-Roussarie, E. 1995. *Pour une approche parallèle de la structure prosodique : Etude de l'organisation prosodique et rythmique de la phrase française*. Thèse de doctorat d'état, Université de Toulouse Le Mirail, Toulouse, France.
- Delais-Roussarie, E. 2000. Vers une nouvelle approche de la structure prosodique. *Langue Française*, 126:92–112.
- Delais-Roussarie, E. 2005. Vers une grammaire prosodique formelle: le cas des incidentes en français. In *Proceedings of IDP05*, pages 1–5, Aix-en-Provence, France.
- Delais-Roussarie, E. 2006. La prosodie des incidentes en français. *Cahiers de Grammaire*, 30:129–138.
- Delaveau, A. 2001. La subordination. In *Syntaxe: la phrase et la subordination*, pages 77–93. Armand Colin, Paris.
- Depraetere, I. 1996. Foregrounding in English relative clauses. *Linguistics*, 34(4):699–732.
- Deulofeu, J. 1999. *Recherches sur les formes de la prédication en français contemporain: le cas des énoncés introduits par que*. Ph.D. thesis, Thèse de doctorat d'état, Université Paris III Sorbonne Nouvelle, Paris, France.
- Deulofeu, J. 2010. La greffe d'un énoncé sur une construction: une combinaison originale de parataxe et de rection. In Béguelin, M. J., Avanzi, M., et Corminboeuf, G., eds, *La Parataxe*, volume 1, pages 175–208. Peter Lang, Berne.
- Deulofeu, J. 2013. Le rôle de l'élément que dans les phénomènes de subordination. In Debaisieux, J.-M., ed, *Analyses Linguistiques sur Corpus: subordination et insubordination en français*, pages 363–410. Lavoisier, Paris.
- Dik, S. 1989. *The Theory of Functional Grammar*. Foris, Dordrecht.
- Dimitrova, D., Chu, M., Wang, L., Özyürek, A., et Hagoort, P. 2016. Beat that Word: How Listeners Integrate Beat Gesture and Focus in Multimodal Speech Discourse. *Journal of Cognitive Neuroscience*, 28(9):1255–1269.
- Dohen, M. et Loevenbruck, H. 2009. Interaction of audition and vision for the perception of prosodic contrastive focus. *Language and Speech*, 52(2-3):177–206.
- Doron, E. 1994. The discourse function of appositives. In *Proceedings of IATL 9*, volume 9, pages 53–62, Jerusalem. Hebrew University.
- Driskell, J. E. et Radtke, P. H. 2003. The effect of gesture on speech production and comprehension. *The Journal of the Human Factors and Ergonomics Society*, 45(3):445–454.
- Ducrot, O. 1980. *Dire et ne pas dire: principes de sémantique linguistique*. Hermann, Paris.
- Duez, D. 1985. Perception of silent pauses in continuous speech. *Language and Speech*, 28:377–389.

- Duez, D. 1987. *Contribution à l'étude de la structuration temporelle de la parole en français*. Thèse de doctorat d'état, Université de Provence, Aix-en-Provence, France.
- Duez, D. 1991. *La pause dans la parole de l'homme politique*. Editions du CNRS, Paris, France.
- Duncan, S. et Fiske, D. 1977. *Face to Face Interaction : Research, Methods, and Theory*. Lawrence Erlbaum Associates, Hillsdale, USA.
- Duncan, S. et Niederehe, G. 1974. On signalling that it's your turn to speak. *Journal of Experimental Social Psychology*, 10(3):234–247.
- Duncan, S. 1972. Some signals and rules for taking speaking turns in conversation. *Journal of Personality and Social Psychology*, 23:283–292.
- Duvallon, O. et Routarinne, S. 2005. Parenthesis as a resource in the grammar of conversation. In Hakulinen, A. et Selting, M., eds, *Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction*, volume 17, pages 45–74. John Benjamins, Amsterdam.
- Edwards, J., Beckman, M. E., et Fletcher, J. 1991. The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89:369–382.
- Ehmer, O. 2016. Adverbial patterns in interaction. *Language Sciences*, 58:1–7.
- Eisenstein, J., Barzilay, R., et Davis, R. 2008. Gestural Cohesion for Topic Segmentation. In *Proceedings of ACL*, pages 852–860, Columbus, USA. Association for Computational Linguistics.
- Ekman, P. et Friesen, W. V. 1969. The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding. *Semiotica*, 1(1):49–98.
- Ekman, P. 1979. About brows : Emotional and conversational signals. In von Cranach, M., Foppa, K., Lepinies, W., et Ploog, D., eds, *Human Ethology : Claims and limits of a new discipline. Contributions to the Colloquium*, pages 169–248. Cambridge University Press, Cambridge, Royaume-Uni.
- Ekman, P. 1989. La mesure de l'expression faciale. *Science et Vie*, 168:24–31.
- Emonds, J. 1979. Appositive Relatives Have No Properties. *Linguistic Inquiry*, 10:211–243.
- Enfield, N. J. 2004. On linear segmentation and combinatorics in co-speech gesture: A symmetry-dominance construction in Lao fish trap descriptions. *Semiotica*, 149:57–124.
- Enfield, N. J. 2009. *The Anatomy of Meaning: Speech, Gesture and Composite Utterances*. Cambridge University Press, Cambridge.
- Enfield, N. J. 2017. *Dependencies in language*. Language Science Press, Berlin.
- Esteve-Gibert, N., Borràs-Comes, J., Swerts, M. G. J., et Prieto, P. 2014. Head gesture timing is constrained by prosodic structure. In *Proceedings of Speech Prosody 2014*, pages 356–360, Dublin, Irlande. ISCA.
- Esteve-Gibert, N. et Prieto, P. 2013. Prosodic structure shapes the temporal realization of intonation and manual gesture movements. *Journal of Speech, Language, and Hearing Research*, 56(3):850–864.
- eSurv. 2017. <http://esurv.org/>.

- Evans, N. 2007. Insubordination and its uses. In Nikolaeva, I., ed, *Finiteness: Theoretical and Empirical Foundations*, pages 366–431. Oxford University Press, Oxford.
- Everett, D. 2005. Cultural constraints on grammar and cognition in Pirahã: Another look at the design features of human language. *Current Anthropology*, 46:621–646.
- Fabb, N. 1990. The difference between English restrictive and nonrestrictive relative clauses. *Journal of Linguistics*, 26(1):57–77.
- Fauconnier, G. et Turner, M. 1996. Blending as a central concept in grammar. In Goldberg, A., ed, *Conceptual Structure, Discourse, and Language*. Cambridge University Press, Cambridge, MA.
- Fauconnier, G. 1984. *Espaces mentaux*. Minit, Paris.
- Fauconnier, G. 1991. Subdivision cognitive. *Communications*, 53:229–248.
- Ferré, G. 2004. *Relations entre discours, intonation et gestualité en anglais britannique*. Thèse de doctorat d'état, Université de la Sorbonne Nouvelle-Paris III, Paris, France.
- Ferré, G. 2014a. A Multimodal Approach to Markedness in Spoken French. *Speech Communication: Special Issue on Gesture and Speech in Interaction*, 57:268–282.
- Ferré, G. 2014b. Multimodal Hyperbole. *Multimodal Communication*, 3(1):25–50.
- Fillmore, C. 1985. Syntactic intrusions and the notion of grammatical construction. In *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society*, pages 73–86, Berkeley, CA, USA. eLanguage.
- Fillmore, C. 1990. Epistemic stance and grammatical form in English conditional sentences. In *Proceedings of CLS26*, pages 137–162, Chicago, USA. Chicago Linguistic Society.
- Firbas, J. 1971. On the concept of communicative dynamism in the theory of functional sentence perspective. *Brno Studies in English*, 7:12–47.
- Firbas, J. 1989. Degrees of communicative dynamism and degrees of prosodic prominence (weight). *Brno studies in English*, 18:21–66.
- Ford, C. E., Fox, B. A., et Thompson, S. A. 1996. Practices in the construction of turns: The TCU revisited. *Pragmatics*, 6:3–427.
- Ford, C. E. 1993. *Grammar in Interaction. Adverbial Clauses in American English Conversation*. Cambridge University Press, Cambridge.
- Forsgren, M. 1993. L'adjectif et la fonction d'apposition: observations syntaxiques, sémantiques et pragmatiques. *L'Information Grammaticale*, 58(1):15–22.
- Fougeron, C. et Keating, P. A. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America*, 101(6):3728–3740.
- Fowler, C. A. 1995. Acoustic and kinematic correlates of contrastive stress accent in spoken English. In Bell-Berti, F. et Raphael, L.-J., eds, *Producing speech: Contemporary issues*, pages 355–373. AIP Press, Woodbury, USA.
- Fox, B. A. et Thompson, S. A. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language*, 66(2):297–316.

- Fox, B. A. et Thompson, S. A. 2009. On formulating reference: An interactional approach to relative clauses in English conversation. *Papers in Pragmatics*, 4(1):183–196.
- Fox Tree, J. E. et Schrock, J. C. 1999. Discourse markers in spontaneous speech: Oh what a difference an oh makes. *Journal of Memory and Language*, 40(2):280–295.
- Fox Tree, J. et Schrock, J. 2002. Basic meanings of “You know” and “I mean”. *Journal of Pragmatics*, 34:727–747.
- Fraisse, P. 1974. *La Psychologie du Rythme*. PUF, Paris.
- Fraser, B. 1999. What are discourse markers? *Journal of Pragmatics*, 31:931–952.
- Frazier, L., Carlson, K., et Clifton Jr., C. 2006. Prosodic phrasing is central to language comprehension. *TRENDS in Cognitive Sciences*, 10(6):244–249.
- Frederiksen, A. T. 2016. Hold + Stroke Gesture Sequences as Cohesion Devices: Examples from Danish Narratives. *San Diego Linguistics Papers*, 6:2–13.
- Fromont, L., Biau, E., et Soto-Faraco, S. 2010. Taking prosody by the hand: The effect of pauses and gestures on sentence disambiguating. In *Proceedings of the 29th Conference on Generative Grammar*, Barcelone, Espagne.
- Fromont, L. A., Soto-Faraco, S., et Biau, E. 2017. Searching high and low: prosodic breaks disambiguate relative clauses. *Frontiers in Psychology [online]*, 8.
- Furman, R., Kuntay, A., et Özyürek, A. 2014. Early language-specificity of children’s event encoding in speech and gesture : Evidence from caused motion in Turkish. *Language, Cognition and Neuroscience*, 29:620–634.
- Furuwaka, N. 2005. Sémantique des propositions relatives : adnominale/prédicative. In *La syntaxe au coeur de la grammaire. Recueil offert en hommage pour le 60e anniversaire de Claude Muller*, pages 99–108. Presses Universitaires de Rennes, Rennes.
- Gaatonne, D. 2005. Relations de dépendance : direction et hiérarchie. In *La syntaxe au coeur de la grammaire. Recueil offert en hommage pour le 60e anniversaire de Claude Muller*, pages 109–120. Presses Universitaires de Rennes, Rennes.
- Galati, A. et Brennan, S. E. 2010. Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62(1):35–51.
- Galati, A. et Brennan, S. E. 2014. Speakers adapt gestures to addressees’ knowledge: implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, 29(4):435–451.
- Gilbert, A., Izhak, I., et Baum, S. 2016. A cross-language investigation of word segmentation by bilinguals with varying degrees of proficiency: Preliminary results. In *Proceedings of Speech Prosody 2016*, pages 36–39, Boston, MA, USA. ISCA.
- Givón, T. 1984. *Syntax: a functional-typological introduction*. John Benjamins, Amsterdam.
- Goffman, E. 1967. *Interaction rituals*. Anchor, Garden City, NY.
- Goffman, E. 1981. *Forms of talk*. University of Pennsylvania Press, Philadelphie.
- Goldberg, A. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press, Chicago, USA.

- Goldberg, A. 2006. *Construction at work : The nature of generalization in language*. Oxford University Press, Oxford.
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., et Wagner, S. 2001. Explaining math: Gesturing lightens the load. *Psychological Science*, 12(6):516–522.
- Goldin-Meadow, S. 2005. *Hearing gesture : How our hands help us think*. Harvard University Press, Cambridge, MA, USA.
- Goldin-Meadow, S. 2015. From action to abstraction : Gesture as a mechanism of change. *Developmental Review*, 38:167–184.
- Goodwin, C. et Goodwin, M. 1992. Context, activity and participation. In Auer, P. et Di Luzio, A., eds, *The Contextualization of Language*, pages 76–99. John Benjamins, Amsterdam.
- Goodwin, C. 1979. The interactive construction of a sentence in natural conversation. *Everyday Language: Studies in Ethnomethodology*, pages 97–121.
- Goodwin, C. 1980. Restarts, pauses, and the achievement of a state of mutual gaze at turn beginnings. *Sociological Inquiry*, 50(3-4):272–302.
- Goodwin, C. 1981. *Conversational organization: Interaction between speakers and hearers*. Academic Press, New York.
- Goodwin, C. 1986. Audience diversity, participation and interpretation. *Text*, 6(3):283–316.
- Goodwin, C. 1996. Transparent vision. In Ochs, E., Schegloff, E. A., et Thompson, S. A., eds, *Interaction and Grammar*, pages 370–405. Cambridge University Press, Cambridge, MA.
- Goodwin, C. 2012. The co-operative, transformative organization of human action and knowledge. *Journal of Pragmatics*, 46(1):8–23.
- Gosselin, L. 1990. Les circonstanciels: de la phrase au texte. *Langue Française*, (86):37–45.
- Granström, B., House, D., et Lundeberg, M. 1999. Prosodic cues in multimodal speech perception. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS99)*, pages 655–658, San Francisco, USA.
- Granström, B. et House, D. 2005. Audiovisual representation of prosody in expressive speech communication. *Speech Communication*, 46(3):473–484.
- Grimes, J. 1975. *The thread of discourse*, volume 207. Mouton de Gruyter, La Haye.
- Grosjean, F. et Deschamps, A. 1973. Analyse des variables temporelles du français spontané. *Phonetica*, 28(3-4):191–226.
- Grosjean, F. et Deschamps, A. 1975. Analyse contrastive des variables temporelles de l'anglais et du français: vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica*, 31(3-4):144–184.
- Gross, S. T. 1986. The kappa coefficient of agreement for multiple observers when the number of subjects is small. *Biometrics*, 42(4):883–893.
- Gross, G. 2005. Les circonstanciels sont des complétives. In Lambert, F. et Nolke, H., eds, *La Syntaxe au coeur de la Grammaire. Recueil offert en hommage pour le 60e anniversaire de Claude Muller*, pages 121–126. Presses Universitaires de Rennes, Rennes.

- Grosu, A. et Landman, F. 1998. Strange relatives of the third kind. *Natural Language Semantics*, 6(2):125–170.
- Grosz, B. J., Weinstein, S., et Joshi, A. K. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational linguistics*, 21(2):203–225.
- Grosz, B. J. et Hirschberg, J. 1992. Some intonational characteristics of discourse structure. In *Proceedings of the 2nd International Conference on Spoken Language Processing (ICSLP 92)*, pages 429–432, Banff, Alberta, Canada. ISCA.
- Grosz, B. J. et Sidner, C. L. 1985. Discourse structure and the proper treatment of interruptions. *ICJAI*, 85:832–839.
- Grosz, B. J. et Sidner, C. L. 1986. Attention, intentions, and the structure of discourse. *Computational linguistics*, 12(3):175–204.
- Gullberg, M. 2006. Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, 56(1):155–196.
- Gumperz, J. 1982. *Discourse strategies*. Number 1. Cambridge University Press, Cambridge.
- Gumperz, J. J. 1992. Contextualization and understanding. In Duranti, A. et Goodwin, C., eds, *Rethinking Context: Language as an interactive phenomenon*, number 11 in *Studies in the Social and Cultural Foundations of Language*, pages 229–252. Cambridge University Press, Cambridge.
- Gussenhoven, C. et Rietveld, A. C. M. 1992. Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, 20(3):283–303.
- Hadar, U., Steiner, T. J., Grant, E. C., et Clifford Rose, F. 1983. Head movement correlates of juncture and stress at sentence level. *Language and Speech*, 26(2):117–129.
- Haegeman, L. 2010. The internal syntax of adverbial clauses. *Lingua*, 120:628–648.
- Haiman, J. et Thompson, S. A. 1984. “Subordination” in Universal Grammar. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, pages 510–523, Berkeley, CA, USA. eLanguage.
- Hakulinen, A. et Selting, M. 2005. *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-In-Interaction*. John Benjamins, Amsterdam.
- Halitsky, D. 1974. Deep structure appositive and complement NPs. *Language*, 50:446–454.
- Halliday, M. et Hasan, R. 1976. *Cohesion in English*. Longman, Londres.
- Halliday, M. A. K. 1967. *Intonation and Grammar in British English*. Mouton de Gruyter, La Haye.
- Halliday, M. A. K. 1985. *An Introduction to Functional Grammar*. Edward Arnold, Londres.
- Harris, Z. 1982. *A Grammar of English on Mathematical Principles*. John Wiley, New York.
- Hassemer, J. et Winter, B. 2016. Producing and perceiving gestures conveying height or shape. *Gesture*, 15(3):404–424.

- Heringa, H. 2007. Appositional constructions: coordination and predication. In *Proceedings of the Fifth Semantics in the Netherlands Day*, pages 67–82, Groningue, Pays-Bas.
- Heritage, J. 1984. A change-of-state token and aspects of its sequential placement. In Atkinson, J. M. et Heritage, J., eds, *Structures of Social Action: Studies in conversation analysis*, pages 299–345. Cambridge University Press, Cambridge.
- Hermes, D. J. 2006. Stylization of Pitch Contours. In Sudhoff, S., Lenertova, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., et Schileßer, J., eds, *Methods in Empirical Prosody Research*, pages 29–61. de Gruyter, Berlin, Allemagne.
- Heylen, D. 2006. Head gestures, gaze, and the principles of conversational structure. *International Journal of Humanoid Robotics*, 3(3):241–267.
- Hirrel, L. 2018. *Cyclic gestures and multimodal symbolic assemblies : an argument for symbolic complexity in gesture*. Thèse de doctorat d'état, University of New Mexico, Albuquerque, NM, USA.
- Hirschberg, J., Nakatani, C., et Grosz, B. 1995. Discourse structure in spoken language: Studies on speech corpora. In *Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, pages 189–192, Palo Alto, CA, USA. ACM.
- Hirschberg, J. et Grosz, B. 1992. Intonational features of local and global discourse structure. In *Proceedings of the Workshop on Speech and Natural Language*, pages 441–446, Morristown, NJ, USA. Association for Computational Linguistics.
- Hirschberg, J. et Nakatani, C. 1996. A prosodic analysis of discourse segments in direction-giving monologues. In *Proceedings of the 34th annual meeting of the Association for Computational Linguistics*, pages 286–293, Santa Cruz, USA. ACM.
- Hirschberg, J. et Pierrehumbert, J. 1986. The intonational structuring of discourse. In *Proceedings of the 24th annual meeting of the Association for Computational Linguistics*, pages 136–144, New York, NY, USA. ACM.
- Hirst, D. J. 2007. A Praat plugin for Momel and INTSINT with improved algorithms for modelling and coding intonation. In *Proceedings of the XVIth International Conference of Phonetic Sciences*, pages 1233–1236, Sarrebruck, Allemagne. Univ. des Saarlandes.
- Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., et Swerts, M. 2015. Reduction in gesture during the production of repeated references. *Journal of Memory and Language*, 79:1–17.
- Hogrefe, K., Ziegler, W., et Goldenberg, G. 2011. Measuring the formal diversity of hand gestures by their hamming distance. In Stam, G. et Ishino, M., eds, *Integrating Gestures*, pages 75–88. John Benjamins, Amsterdam.
- Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., et Gunter, T. C. 2012. Gesture facilitates the syntactic analysis of speech. *Frontiers in Psychology*, 3(74):1–12.
- Holle, H. et Gunter, T. C. 2007. The role of iconic gestures in speech disambiguation : ERP evidence. *Journal of Cognitive Neuroscience*, 19:1175–1192.

- Holler, J., Tutton, M., et Wilkin, K. 2011. Co-speech gestures in the process of meaning coordination. In *Proceedings of GESPIN 2011*, Bielefeld, Allemagne. University of Bielefeld.
- Holler, J., Schubotz, L., Kelly, S., Hagoort, P., Schuetze, M., et Özyürek, A. 2014. Social eye gaze modulates processing of speech and co-speech gesture. *Cognition*, 133(3):692–697.
- Holler, J. et Stevens, R. 2007. The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26(1):4–27.
- Holler, J. et Wilkin, K. 2011. Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue. *Journal of Nonverbal Behavior*, 35(2):133–153.
- Holler, A. 2005. Expressing communicative-weight assignment discourse structurally. In *Proceedings of the Workshop on Constraints in Discourse*, volume 6, pages 88–94, Dortmund, Allemagne. Universität Dortmund.
- Horgues, C. 2010. *Prosodie de l'accent français en anglais et perception par des auditeurs anglophones*. Thèse de doctorat d'état, Université Paris Diderot Paris 7, Paris, France.
- Hostetter, A. B., Alibali, M. W., et Schrager, S. M. 2011. If you don't already know, I'm certainly not going to show you! Motivation to communicate affects gesture production. In Stam, G. et Ishino, M., eds, *Integrating Gestures*, pages 61–74. John Benjamins, Amsterdam.
- Hostetter, A. B. et Alibali, M. W. 2010. Language, gesture, action! A test of the Gesture as Simulated Action framework. *Journal of Memory and Language*, 63(2):245–257.
- House, D., Beskow, J., et Granström, B. 2001. Timing and interaction of visual cues for prominence in audiovisual speech perception. In *Proceedings of Eurospeech*, Aalborg, Danemark. Aalborg University.
- Hubbard, A. L., Wilson, S. M., Callan, D. E., et Dapretto, M. 2009. Giving speech a hand: Gesture modulates activity in auditory cortex during speech perception. *Human Brain Mapping*, 30(3):1028–1037.
- Huddleston, R. et Pullum, G. K. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge.
- Ishi, C. T., Ishiguro, H., et Hagita, N. 2014. Analysis of relationship between head motion events and speech in dialogue conversations. *Speech Communication*, 57:233–243.
- Ito, K., Turnbull, R., et Speer, S. 2017. Allophonic tunes of contrast : Lab and spontaneous speech lead to equivalent fixation. *Laboratory Phonology*, 8(1):1–29.
- Iverson, J. M. et Thelen, E. 1999. Hand, Mouth and Brain. The Dynamic Emergence of Speech and Gesture. *Journal of Consciousness Studies*, 6(11-12):19–40.
- Jackendoff, R. 1977. *X-bar Syntax: A Study of Phrase Structure*. The MIT Press, Cambridge, MA.

- Jacobs, N. et Garnham, A. 2007. The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, 56(2):291–303.
- Jantunen, T. 2017. Constructed Action, the Clause and the Nature of Syntax in Finnish Sign Language. *Open Linguistics*, 3(1):65–85.
- Jespersen, O. 1927. *A Modern English Grammar on Historical Principles*, volume 3. George Allen & Unwin, Londres.
- Jucker, A. H. 1993. The discourse marker well: A relevance-theoretical account. *Journal of Pragmatics*, 19(5):435–452.
- Jun, S.-A. et Fletcher, J. 2014. Methodology of studying intonation: From data collection to data analysis. In Jun, S.-A., ed, *Prosodic Typology II: The phonology of intonation and phrasing*, pages 493–519. Oxford University Press, New York.
- Jurafsky, D., Shriberg, E., Fox, B., et Curl, T. 1998. Lexical, prosodic, and syntactic cues for dialog acts. In *Proceedings of ACL/COLING-98 Workshop on Discourse Relations and Discourse Markers*, pages 114–120, Montreal, Canada. ACL.
- Kakouros, S. et Räsänen, O. 2016. Perception of sentence stress in speech correlates with the temporal unpredictability of prosodic features. *Cognitive Science*, 40(7):1739–1774.
- Katsika, A., Krivokapic, J., Mooshammer, C., Tiede, M. K., et Goldstein, L. 2014. The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics*, 44:62–82.
- Kay, P. et Fillmore, C. 1999. Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. *Language*, 75(1):1–33.
- Keenan, E. L. et Comrie, B. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry*, 8(1):63–99.
- Kelly, S. D., Manning, S. M., et Rodak, S. 2008. Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass*, 2(4):569–588.
- Kelly, S. D., Özyürek, A., et Maris, E. 2010. Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, 21(2):260–267.
- Kendon, A. 1967. Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26:22–63.
- Kendon, A. 1972. Some relationships between body motion and speech. In Siegman, A. W. et Pope, B., eds, *Studies in Dyadic Communication*, pages 177–210. Pergamon, New York.
- Kendon, A. 1980. Gesticulation and speech: Two aspects of the process of utterance. *The Relationship of Verbal and Nonverbal Communication*, 25:207–227.
- Kendon, A. 1983. Gesture and speech : How they interact. In Wiemann, J. M. et Harrison, R. P., eds, *Nonverbal interaction*, pages 13–45. Sage Publications, Beverly Hills, USA.
- Kendon, A. 1995. Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics*, 23(3):247–279.
- Kendon, A. 2004. *Gesture: Visible action as utterance*. Cambridge University Press, Cambridge.

- Khalifa, J.-C. 2004. *Syntaxe de l'anglais : théories et pratiques de l'énoncé complexe*. Ophrys, Paris.
- Kim, J.-B. 2012. Form and Function Mismatch in the English Appositional Construction. *Language Research*, 48(3):509–531.
- Kipp, M., Neff, M., et Albrecht, I. 2007. An annotation scheme for conversational gestures: how to economically capture timing and form. *Language Resources & Evaluation*, 41:325–339.
- Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., et Ishizuka, T. 2007. Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes*, 22(8):1212–1236.
- Kita, S. et Özyürek, A. 2003. What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48(1):16–32.
- Kita, S. 2000. How representational gestures help speaking. In McNeill, D., ed, *Language and Gesture*, pages 162–185. Cambridge University Press, Cambridge.
- Kleiber, G. 1980. Où en est-on de l'opposition relative restrictive / relative appositive? *L'Information Grammaticale*, 7:12–17.
- Knoeferle, P. et Kreysa, H. 2012. Can Speaker Gaze Modulate Syntactic Structuring and Thematic Role Assignment During Spoken Sentence Comprehension? *Frontiers in Psychology*, 3(538):1–15.
- Kobayashi, H. et Koshima, S. 1997. Unique morphology of the human eye. *Nature*, 387:767–768.
- Kok, K. et Cienki, A. J. 2016. Cognitive Grammar and gesture: Points of convergence, advances and challenges. *Cognitive Linguistics*, 27(1):67–100.
- Krahmer, E. et Swerts, M. 2005. More about brows. In Ruttkay, Z. et Pelachaud, C., eds, *From Brows to Trust: evaluating embodied conversational agents*, pages 191–216. Kluwer, Dordrecht.
- Krahmer, E. et Swerts, M. 2007a. The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, 57(3):396–414.
- Krahmer, E. et Swerts, M. 2007b. Perceiving focus. In Lee, C., Gordon, M., et Büring, D., eds, *Topic and Focus*, pages 121–137. Springer, Dordrecht.
- Kratzer, A. 2006. Decomposing attitude verbs. Talk delivered at the Hebrew University of Jerusalem. Jerusalem. The Hebrew University of Jerusalem.
- Krauss, R. M., Chen, Y., et Gotfexnum, R. F. 2000. Lexical gestures and lexical access: a process model. In McNeill, D., ed, *Language and Gesture*, volume 2, pages 261–282. Cambridge University Press, Cambridge.
- Kreiman, J. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics*, 10(2):163–175.
- Krifka, M. 2007. Basic Notions of Information Structure. *Interdisciplinary Studies on Information Structure*, 6:13–55.

- Krivokapic, J., Tiede, M. K., et Tyrone, M. E. 2015. A kinematic analysis of prosodic structure in speech and manual gestures. In *Proceeding of ICPHS 2015*, Glasgow, Royaume-Uni. International Phonetic Association.
- Krivokapic, J., Tiede, M. K., et Tyrone, M. E. 2017. A Kinematic Study of Prosodic Structure in Articulatory and Manual Gestures: Results from a Novel Method of Data Collection. *Laboratory Phonology*, 8(1):1–26.
- Krivokapic, J. 2014. Gestural coordination at prosodic boundaries and its role for prosodic structure and speech planning processes. *Philosophical Transactions of the Royal Society B.*, 369(20130397).
- Kuno, S. 1974. The position of relative clauses and conjunctions. *Linguistic Inquiry*, 5(1):117–136.
- Kuroda, S. Y. 1968. English relativization and certain related problems. *Language*, 44(2):244–266.
- Labov, W. et Waletzky, J. 1967. Narrative analysis: Oral versions of personal experience. In *Essays on the verbal and visual arts – Proceedings of the American Ethnological Society*, volume 7, pages 3–38, Seattle, Washington.
- Labov, W. 1972. Some principles of linguistic methodology. *Language in Society*, 1(01):97–120.
- Lacheret-Dujour, A. et Victorri, B. 2002. La période intonative comme unité d'analyse pour l'étude du français parlé : modélisation prosodique et enjeux linguistiques. *Verbum*, 24(1):55–72.
- Ladd, D. R. et Campbell, N. 1991. Theories of prosodic structure: evidence from syllable duration. In *Proceedings of the 12th International Congress of Phonetic Sciences*, volume 2, pages 290–293, Aix-en-Provence, France. IPA.
- Ladd, R. D. 2008. *Intonational phonology*. Cambridge University Press, Cambridge.
- Lago, J. 1994. L'apposition est-elle une fonction ou un mode de construction? *L'Information Grammaticale*, 63(1):12–17.
- Lakoff, G. et Johnson, M. 1980. *Metaphors we live by*. University of Chicago Press, Chicago.
- Lakoff, R. 1984. The Pragmatics of Subordination. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, pages 481–492, Berkeley, CA, USA. eLanguage.
- Lakoff, G. 1987. *Women fire and dangerous things : What categories reveal about the mind*. University of Chicago Press, Chicago, USA.
- Lambrecht, K. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press, New York.
- Lambrecht, K. 2000. Prédication seconde et structure informationnelle : la relative de perception comme construction présentative. *Langue Française*, 127:49–66.
- Langacker, R. W. 2001. Dynamicity in grammar. *Axiomathes*, 12(1):7–33.
- Langacker, R. W. 2008. *Cognitive Grammar. A Basic Introduction*. Oxford University Press, Oxford.
- Larreya, P. 1979. *Enoncés performatifs, présuppositions*. Nathan, Paris, France.

- Lascares, A. et Asher, N. 2007. Segmented discourse representation theory: Dynamic semantics with discourse structure. In *Computing Meaning*, pages 87–124. Springer, Amsterdam.
- Lascares, A. et Stone, M. 2009. Discourse coherence and gesture interpretation. *Gesture*, 9(2):147–180.
- Laursen, L. 2005. Towards an embodied Grammar: Gesture in tying practices Constructing obvious cohesion. In *Proceedings of ISGS2*, Lyon, France. ISGS.
- Laury, R. et Suzuki, R. 2011. *Subordination in conversation: A cross-linguistic perspective*, volume 24. John Benjamins, Amsterdam.
- Lazard, G. 1994. *L'Actance*. PUF, Paris.
- Le Goffic, P. 1979. Propositions relatives, identification et ambiguïté. *DRLAV*, 21:135–145.
- Lehiste, I. 1979. Perception of sentence and paragraph boundaries. In Lindblom, B. et Ohman, S., eds, *Frontiers of Speech Communication Research*, pages 191–201. Academic Press, New York.
- Lehmann, C. 1988. Towards a typology of clause linkage. In Haiman, J. et Thompson, S. A., eds, *Clause Combining in Grammar and Discourse*, pages 181–225. Academic Press, Amsterdam.
- Lelandais, M. et Ferré, G. 2014. Multimodal Analysis of Parentheticals in Conversational Speech. *Multimodal Communication*, 3(2):197–217.
- Lelandais, M. et Ferré, G. 2016. Prosodic boundaries in subordinate syntactic constructions. In *Proceedings of Speech Prosody 2016*, pages 183–187, Boston, MA, USA. ISCA.
- Lelandais, M. et Ferré, G. 2017. What do gestures in subordination tell us about (in)dependence? *Journal of Communication Studies*, 4(1-2):43–49.
- Lemke, J. L. 1991. Text Production and Dynamic Text Semantics. In Ventola, E., ed, *Functional and Systemic Linguistics: Approaches and Uses*, number 55 in Trends in Linguistics: Studies and Monographs, pages 23–38. Mouton de Gruyter, Berlin.
- Lemke, J. L. 2002. Travels in hypermodality. *Visual Communication*, 1(3):299–325.
- Lemke, J. L. 2005. Multimedia Genres and Transversals. *Folia Linguistica*, 39(1-2):45–56.
- Leonard, T. et Cummins, F. 2011. The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26(10):1457–1471.
- Levinson, S. C. 2003. Contextualizing “contextualization cues”. In Eerdmans, S., Prevignano, C. L., et Thibault, P., eds, *Language and Interaction: Discussions with John J. Gumperz*, pages 31–39. John Benjamins, Amsterdam.
- Lewandoski, W. et Özçaliskan, S. 2018. How event perspective influences speech and co-speech gestures about motion. *Journal of Pragmatics*, 128:22–29.
- Lindström, J., Lindholm, C., et Laury, R. 2016. The interactional emergence of conditional clauses as directives: constructions, trajectories and sequences of actions. *Language Sciences*, 58:8–21.

- Local, J. 1992. Continuing and restarting. In Auer, P. et Di Luzio, A., eds, *The Contextualization of Language*, pages 273–296. John Benjamins, Amsterdam.
- Local, J. 2007. Phonetic detail and the organisation of talk-in-interaction. In *Proceedings of the 16th ICPHS*, Sarrebruck, Allemagne.
- Loehr, D. P. 2004. *Gesture and intonation*. Thèse de doctorat d'état, Université de Georgetown.
- Longacre, R. E. 1996. *The Grammar of Discourse*. Springer, New York.
- Loock, R. et O'Connor, K. M. 2013. The Discourse Functions of Nonverbal Appositives. *Journal of English Linguistics*, 41(4):332–358.
- Loock, R. 2007. Appositive relative clauses and their functions in discourse. *Journal of Pragmatics*, 39(2):336–362.
- Loock, R. 2013. Pour (enfin?) en finir avec les deux types de relatives: la linguistique face aux limites de la catégorisation. *Cercles*, (29):21–45.
- Lytte, E. G. 1974. *A grammar of subordinate structures in English*. Mouton, La Haye.
- Lytvynova, M. et Dao, H. L. 2014. Les relatives appositives entre intégration syntaxique et intégration discursive. In *Actes du Congrès Mondial de Linguistique Française – CMLF 2014*, Berlin, Allemagne.
- Macaulay, R. 2002. You know, it depends. *Journal of Pragmatics*, 34:749–767.
- Mann, W. C. et Thompson, S. A. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.
- Martin, R. 1983. *La logique du sens*. PUF, Paris.
- Martinec, R. et Salway, A. 2005. A system for image-text relations in new (and old) media. *Visual Communication*, 4(3):337–371.
- Masson-Carro, I., Goudbeek, M., et Krahmer, E. 2016. Imposing Cognitive Constraints on Reference Production: The Interplay Between Speech and Gesture During Grounding. *Topics in Cognitive Science*, pages 1–18.
- Matthews, P. H. 1981. *Syntax*. Cambridge University Press, Cambridge.
- Matthiessen, C. et Thompson, S. A. 1988. The structure of discourse and “subordination”. In Haiman, J. et Thompson, S. A., eds, *Clause Combining in Grammar and Discourse*, pages 275–329. John Benjamins, Amsterdam.
- Maynard, S. K. 1987. Interactional functions of a nonverbal sign head movement in Japanese dyadic casual conversation. *Journal of Pragmatics*, 11(5):589–606.
- Mazeland, H. et Huiskes, M. 2001. Dutch “but” as a sequential conjunction. In Selting, M. et Couper-Kuhlen, E., eds, *Studies in Interactional Linguistics*, pages 141–69. John Benjamins, Amsterdam.
- McCawley, J. D. 1996. An overview of “appositive” constructions in English. In Przedziecki, M. et Whaley, L., eds, *Proceedings of ESCOL' 95*, pages 195–211, Ithaca, New York. Cornell University Press.
- McClave, E. Z. 1991. *Intonation and Gesture*. Thèse de doctorat d'état, Georgetown University, Washington D.C., USA.
- McClave, E. Z. 2000. Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32(7):855–878.

- McCullough, K. E. 2005. *Using gestures during speaking : Self-generating indexical fields*. Thèse de doctorat d'état, Université de Chicago, Chicago, USA.
- McNeill, D., Quek, F., McCullough, K. E., Duncan, S., Furuyama, N., Bryll, R., Ma, X. F., et Ansari, R. 2001. Catchments, prosody and discourse. *Gesture*, 1(1):9–33.
- McNeill, D., Duncan, S. D., Cassell, J., et Levy, E. T. 2007. *Gesture and the dynamic dimension of language: Essays in honor of David McNeill*. John Benjamins, Amsterdam.
- McNeill, D. et Duncan, S. 2000. Growth points in thinking-for-speaking. In McNeill, D., ed, *Language and Gesture*, volume 2, pages 141–161. Cambridge University Press, Cambridge.
- McNeill, D. et Levy, E. T. 1993. Cohesion and gesture. *Discourse Processes*, 16(4):363–386.
- McNeill, D. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago Press, Chicago.
- McNeill, D. 2000. *Language and Gesture*. Cambridge University Press, Cambridge.
- McNeill, D. 2005. *Gesture and thought*. University of Chicago Press, Chicago.
- McNeill, D. 2017. Gesture-speech unity - what it is, where it comes from. In Church, R. B., Alibali, M. W., et Kelly, S. D., eds, *Why Gesture? How the Hands Function in Speaking, Thinking, and Communicating*, pages 77–101. John Benjamins, Amsterdam.
- McQuown, N. A. 1971. *The natural history of an interview*. University of Chicago Press, Chicago.
- Melinger, A. et Levelt, W. M. 2004. Gesture and the communicative intention of the speaker. *Gesture*, 4(2):119–141.
- Melis, G. 2008. Relatives et types de qualification. *Cycnos [en ligne]*, 17.
- Mendoza-Denton, N. et Jannedy, S. 2011. Semiotic layering through gesture and intonation: a case study of complementary and supplementary multimodality in political speech. *Journal of English Linguistics*, 39(3):265–299.
- Mertens, P. et Simon, A.-C. 2013. Exploring acoustic and syntactic cues to prosodic boundaries in French. A multi-genre corpus study. In *Proceedings of ICPHS 2013*, pages 81–87, Glasgow, Royaume-Uni. IPA Archive.
- Mertens, P. 2008. Syntaxe, prosodie et structure informationnelle: une approche prédictive pour l'analyse de l'intonation dans le discours. *Travaux de Linguistique*, 56:97–124.
- Mettouchi, A., Lacheret-Dujour, A., Silber-Varod, V., et Izre'el, S. 2007. Only Prosody? Perception of speech segmentation in Kabyle and Hebrew. *Nouveaux Cahiers de Linguistique Française*, 28:207–218.
- Meyer, L., Obleser, J., Kiebel, S. J., et Friederici, A. D. 2012. Spatiotemporal dynamics of argument retrieval and reordering: an fMRI and EEG study on sentence processing. *Frontiers in Psychology*, 3:[en ligne].
- Miller, J. E., Weinert, R., et Miller, J. 1998. *Spontaneous Spoken Language: Syntax and Discourse*. Clarendon Press, Oxford.

- Mo, Y., Cole, J., et Lee, E.-K. 2008. Naïve listeners' prominence and boundary perception. In *Proceedings of Speech Prosody 2008*, pages 735–738, Campinas, Brésil. ISCA.
- Mo, Y. et Cole, J. 2010. Perception of prosodic boundaries in spontaneous speech with and without silent pauses. *Journal of the Acoustical Society of America*, 127(3):1956.
- Mo, Y. 2008. Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception. In *Proceedings of Speech Prosody 2008*, pages 739–742, Campinas, Brésil. ISCA.
- Mol, L. et Kita, S. 2012. Gesture structure affects syntactic structure in speech. In Miyake, N., Peebles, D., et Cooper, R. P., eds, *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pages 761–766, Austin, Texas, USA. Cognitive Science Society.
- Mondada, L. et Zay, F. 1999. Parenthèses et processus de configuration thématique: vers une redéfinition de la notion de topic. In *Pragmatics in 1998: Selected Papers from the 6th International Pragmatics Conference*, volume 2, pages 396–411, Anvers, Belgique. IprA.
- Mondada, L. 2008a. Contributions de la linguistique interactionnelle. In *Proceedings of Congrès Mondial de Linguistique Française 2008*, pages 1–17, Paris, France. EDP Sciences.
- Mondada, L. 2008b. Production du savoir et interactions multimodales. *Revue d'Anthropologie des Connaissances*, 2(2):219–266.
- Mondada, L. 2011a. Gestion du topic et organisation de la conversation. *Cadernos de Estudos Lingüísticos*, 41:7–35.
- Mondada, L. 2011b. Understanding as an embodied, situated and sequential achievement in interaction. *Journal of Pragmatics*, 43(2):542–552.
- Morel, M.-A. et Danon-Boileau, L. 1998. *Grammaire de l'Intonation*. Ophrys, Paris.
- Morel, M.-A. et Rialland, A. 1992. Emboîtements, autonomies, ruptures dans l'intonation française. In Chuquet, J. et Roulland, D., eds, *Subordination*, number 5 in Travaux linguistiques du CerLiCo, pages 221–243. Presses Universitaires de Rennes, Rennes.
- Morel, M.-A. 1992. Intonation et thématisation. *L'Information Grammaticale*, 54(1):26–35.
- Muller, C. 2006. Sur les propriétés des relatives. *Cahiers de Grammaire*, 30:319–337.
- Muller, C. 2008a. La relation au verbe principal dans les relatives prédicatives en français. *Faits de Langues*, 31(32):337–346.
- Muller, C. 2008b. Modes d'organisation syntaxique et dépendances multiples: micro et macro-syntaxe unifiées. In Van Raemdonck, D., ed, *Modèles Syntaxiques*, pages 213–229. Peter Lang, Berne.
- Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., et Vatikiotis-Bateson, E. 2004. Visual prosody and speech intelligibility head movement improves auditory speech perception. *Psychological science*, 15(2):133–137.

- Nakano, Y. I., Reinstein, G., Stocky, T., et Cassell, J. 2003. Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*, pages 553–561, Sapporo, Japon. Association for Computational Linguistics.
- Nespor, M. et Vogel, I. 1986. *Prosodic Phonology*. Foris, Dordrecht.
- Neveu, F. 2000. Quelle syntaxe pour l'apposition ? Les types d'appariement des appositions frontales et la continuité référentielle. *Langue Française*, 125(125):106–124.
- Nevins, A., Pesetsky, D., et Rodrigues, C. 2009. Pirahã exceptionality: A reassessment. *Language*, 85:355–404.
- Noël, Y. 2013. *Psychologie statistique avec R*. Springer, Berlin.
- Norris, S. 2004. *Analyzing multimodal interaction: A methodological framework*. Routledge, Londres.
- Novick, D. G., Hansen, B., et Ward, K. 1996. Coordinating turn-taking with gaze. In *Proceedings of ICSLP 96*, volume 3, pages 1888–1891. IEEE.
- Oben, B. et Brône, G. 2015. What you see is what you do: on the relationship between gaze and gesture in multimodal alignment. *Language and Cognition*, 7(04):546–562.
- Oben, B. et Brône, G. 2016. Explaining interactive alignment: A multimodal and multifactorial account. *Journal of Pragmatics*, 104:32–51.
- Ochs, E. 1979. Planned and unplanned discourse. In Givón, T., ed, *Syntax and Semantics (XII): Discourse and Syntax*, pages 51–80. Academic Press, New York.
- Oller, K. D. 1973. The effect of position in utterance on speech segment duration in English. *Journal of the Acoustical Society of America*, 54:1235–1247.
- Özyürek, A. How does linguistic framing of events influence co-speech gestures?: Insights from crosslinguistic variations and similarities. *Gesture*, 5(1-2).
- Özyürek, A. et Kita, S. 1999. Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In Hahn, M. et Stoness, S. C., eds, *Proceedings of the Twenty-First Annual Conference of the Cognitive Science Society*, pages 507–512, Mahwah, NJ. Lawrence Erlbaum Associates.
- Pagel, V., Carbonell, N., Laprie, Y., et Vaissière, J. 1995. Spotting prosodic boundaries in continuous speech in French. In *Proceedings of the XIIIth ICPPhS*, pages 308–311, Stockholm, Suède.
- Park-Doob, M. A. 2010. *Gesturing through time : Holds and intermodal timing in the stream of speech*. Thèse de doctorat d'état, Université de Chicago, Chicago, USA.
- Parrell, B., Goldstein, L., et Byrd, D. 2014. Spatiotemporal coupling between speech and manual motor actions. *Journal of Phonetics*, 42:1–11.
- Parrill, F. et Kimbara, I. 2006. Seeing and hearing double: The influence of mimicry in speech and gesture on observers. *Journal of Nonverbal Behavior*, 30(4):157–167.
- Payà, M. 2002. Incidental Clauses in Spoken Catalan: Prosodic Characteristics and Pragmatic Function. In *Proceedings of Speech Prosody 2002*, pages 559–562, Aix-en-Provence, France. ISCA.

- Perniss, P. et Özyürek, A. 2015. Visible Cohesion: A Comparison of Reference Tracking in Sign, Speech, and Co-Speech Gesture. *Topics in Cognitive Science*, 7(1):36–60.
- Peterson, P., Butt, M., et King, T. H. 2004. Non-restrictive relatives and other non-syntagmatic relations in a lexical-functional framework. In *Proceedings of LFG 2004 Conference*, pages 391–397, Stanford, CA, USA. CSLI Publications.
- Peterson, P. 1999. On the boundaries of syntax: non-syntagmatic relations. In Collins, P. et Lee, D., eds, *The Clause in English: in honour of Rodney Huddleston*, pages 229–250. John Benjamins, Amsterdam.
- Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., et Höhle, B. 2017. Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics*, 61:71–92.
- Pierrehumbert, J. et Hirschberg, J. 1990. The Meaning of Intonational contours in the Interpretation of Discourse. In Cohen, P., Morgan, J., et Pollack, M. E., eds, *Intentions in Communication*, pages 271–311. MIT Press, Cambridge.
- Pierrehumbert, J. B. 1980. *The phonology and phonetics of English intonation*. Thèse de doctorat d'état, Massachusetts Institute of Technology, Boston, MA.
- Polanyi, L. 1978. False starts can be true. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, volume 4, pages 629–639, Berkeley, CA, USA. eLanguage.
- Poole, M. E. et Field, T. W. 1976. A comparison of oral and written code elaboration. *Language and Speech*, 19(4):305–312.
- Portes, C. 2002. Approche instrumentale et cognitive de la prosodie du discours en français. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 21:101–119.
- Potts, C. 2005. *The Logic of Conventional Implicatures*. Oxford University Press, New York.
- Powers, D. M. 2012. The problem with kappa. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 345–355, Avignon, France. Association for Computational Linguistics.
- Priva, U. C. 2017. Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, 160:27–34.
- Péry-Woodley, M.-P. 2000. Cadrer ou centrer son discours? Introduteurs de cadres et centrage. *Verbum*, 22(1):59–78.
- Quazza, S. 1991. Modelling Italian intonation in a text-to-speech system. In *Proceedings of Eurospeech*, pages 1161–1164, Genova, Italie. ISCA Archive.
- Quek, F., McNeill, D., Bryll, R., Kirbas, C., Arslan, H., McCullough, K. E., Furuyama, N., et Ansari, R. 2000. Gesture, speech, and gaze cues for discourse segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 247–254, Hilton Head Island, USA. IEEE.
- Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., McCullough, K. E., et Ansari, R. 2002. Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(3):171–193.

- Quirk, R., Greenbaum, S., Leech, G., et Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. Longman, Londres.
- R Core Team. 2017. A language and environment for statistical computing. <http://www.r-project.org/>.
- Radford, A. 1981. *Transformational syntax: a student's guide to Chomsky's extended standard theory*. Cambridge University Press, Cambridge.
- Ramm, W. et Fabricius-Hansen, C. 2005. Coordination and discourse-structural salience from a cross-linguistic perspective. In *Proceedings of the 6th International Workshop on Multidisciplinary Approaches to Discourse (MAD'05) 'Salience in Discourse'*, pages 119–128, Chorin, Allemagne. Stichting/Nodus.
- Rauscher, F. H., Krauss, R. M., et Chen, Y. 1996. Gesture, speech, and lexical access: The role of lexical movements in speech production. *American Psychological Society*, 7(4):226–231.
- Redeker, G. 2006. Discourse markers as attentional cues at discourse transitions. In Fischer, K., ed, *Approaches to Discourse Particles*, volume 1, pages 339–358. Elsevier Science, Oxford.
- Reinhart, T. 1984. Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics*, 22:779–809.
- Rienks, R. J., Poppe, R., et Heylen, D. 2010. Differences in head orientation behavior for speakers and listeners : an experiment in a virtual environment. *Transactions on Applied Perception*, 7(1).
- Ritveld, A. C. et Gussenhoven, C. 1985. On the relation between pitch excursion size and prominence. *Journal of Phonetics*, 13:299–308.
- Romaine, S. et Lange, D. 1991. The use of like as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech*, 66(3):227–279.
- Rossano, F. 2012. *Gaze Behaviour in Face-to-face Interaction*. Thèse de doctorat d'état, Max-Planck Institut for Psycholinguistics, Nimègue, Pays-Bas.
- Rossano, F. 2013. Gaze in Conversation. In Sidnell, J. et Stivers, T., eds, *The Handbook of Conversation Analysis*, pages 308–329. Blackwell, Malden, MA, USA.
- Roustand, B. 2012. *Etude de la coordination gestes manuels / parole dans le cadre de la désignation*. Thèse de doctorat d'état, Université de Grenoble, Grenoble, France.
- Roux, G., Bertrand, R., Ghio, A., et Astésano, C. 2016. Naïve listeners' perception of prominence and and boundary in French spontaneous speech. In *Proceedings of Speech Prosody*, pages 912–916, Boston, MA, USA. ISCA.
- Roy, J., Cole, J., et Mahrt, T. 2017. Individual differences and patterns of convergence in prosody perception. *Laboratory Phonology*, 8(1):1–36.
- Ruth-Hirrel, L. et Wilcox, S. 2018. Speech-gesture constructions in cognitive grammar: The case of beats and points. *Cognitive Linguistics*, 29(3):[online].
- Sabio, F. 2013. Quelques aspects du clause linkage dans le français oral : l'annotation syntaxique des séquences “subordonnées”. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, 29:1–23.

- Sacks, H., Schegloff, E., et Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4):696–735.
- Sacks, H. 1995. *Lectures on Conversation*. Blackwell, Oxford.
- Schefflen, A. E. 1968. Human communication : behavioural programs and their integration in interaction. *Behavioral Sciences*, 13:44–55.
- Schegloff, E. A. 1982. Discourse as an interactional achievement: Some uses of “uh huh” and other things that come between sentences. In Tannen, D., ed, *Analyzing Discourse: Text and Talk*, pages 71–93. Georgetown University Press, Washington D.C.
- Schegloff, E. 1996. Turn organization: One intersection of grammar and interaction. *Studies in Interactional Sociolinguistics*, 13:52–133.
- Schiffrin, D. 1987. *Discourse Markers*. Number 5 in Studies in Interactional Sociolinguistics. Cambridge University Press, Cambridge.
- Schlangen, D. 2006. From reaction to prediction: Experiments with computational models of turn-taking. In *Proceedings of INTERSPEECH 2006*, Pittsburgh, USA. ISCA.
- Schoonjans, S. 2017. Multimodal Construction Grammar issues are Construction Grammar issues. *Linguistics Vanguard*, 3(1):[online].
- Sekine, K. et Kita, S. 2017. The listener automatically uses spatial story representations from the speaker’s cohesive gestures when processing subsequent sentences without gestures. *Acta Psychologica*, 179:89–95.
- Selkirk, E. 1978. On prosodic structure and its relation to syntactic structure. In *Proceedings of Nordic Prosody II*, pages 111–140, Trondheim, Tapir. Fretheim.
- Selkirk, E. O. 1984. *Phonology and Syntax: The Relation between Sound and Structure*. MIT Press, Cambridge, MA.
- Shattuck-Hufnagel, S. et Turk, A. 1998. The domain of phrase-final lengthening in English. In Kuhl, P. K. et Crum, L. A., eds, *The sound of the future: A global view of acoustics in the 21 st century. Proceedings of the 16th International Congress on Acoustics and 135th Meeting Acoustical Society of America*, pages 1235–1236, Seattle, USA. Acoustical Society of America.
- Shriberg, E. et Lickley, R. 1993. Intonation of clause-internal filled pauses. *Phonetica*, 50(3):172–179.
- Shyldkrot Hava, B.-Z. 1995. Subordonnées circonstancielles et dépendance sémantique. Comparaison, concession et condition : grammaticalisation et sens des connecteurs. *Faits de Langues*, 3(5):145–154.
- Silverman, K. E. A., Beckman, M. B., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., Pierrehumbert, J. B., et Hirschberg, J. 1992. TOBI: a standard for labeling English prosody. In *Proceedings of the 2nd International Conference on Spoken Language Processing (ICSLP 92)*, volume 2, pages 867–870, Banff, Alberta, Canada. ISCA Archive.
- Simon, A.-C. et Christodoulides, G. 2016. Perception of Prosodic Boundaries by Naïve Listeners in French. In *Proceedings of Speech Prosody*, pages 1158–1162, Boston, MA, USA. ISCA.

- Sloetjes, H. et Wittenburg, P. 2008. Annotation by Category: ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation*, Marrakech, Maroc. ELRA.
- Smessaert, H., Cornillie, B., Divjak, D., et Eynde, K. 2005. Degrees of clause integration: from endotactic to exotactic subordination in Dutch. *Linguistics*, 43(3):471–529.
- Smith, C. L. 2009. Naïve listeners' perceptions of French prosody compared to the predictions of theoretical models. In Yoo, H.-Y. et Delais-Roussarie, E., eds, *Proceedings of IDP09*, pages 335–349, Paris, France.
- Sprouse, J. 2011. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43:155–167.
- Stivers, T. et Robinson, J. 2006. A preference for progressivity in interaction. *Language in Society*, 35(03):367–392.
- Stivers, T. 2008. Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language & Social Interaction*, 41(1):31–57.
- Streeck, J. 2009. *Gesturecraft. The manufacture of meaning*. John Benjamins, Amsterdam.
- Streeck, J. 2014. Mutual gaze and recognition. Revisiting Kendon's "Gaze direction in two-person conversation". In Seyfeddinipur, M. et Gullberg, M., eds, *From Gesture in Conversation to Visible Action as Utterance*, pages 35–58. John Benjamins, Amsterdam.
- Streefkerk, B. M., Pols, L. C., et ten Bosch, L. F. 1997. Prominence In Read Aloud Sentences, As Marked By Listeners And Classified Automatically. In *Proceedings of IFA 21*, pages 101–116, Amsterdam, Pays-Bas. Institute of Phonetic Science.
- Sueyoshi, A. et Hardison, D. M. 2005. The role of gesture and facial cues in second language listening comprehension. *Language Learning*, 55(4):661–699.
- Svartvik, J. et Quirk, R. 1980. *A Corpus of English Conversation*. Lund University Press, Lund, Suède.
- Sweetser, E. 2006. Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In Gonzalez-Marquez, M., Mittleberg, I., Coulson, S., et Spivey, M., eds, *Methods in Cognitive Linguistics*, pages 203–226. John Benjamins, Amsterdam.
- Swerts, M. et Krahmer, E. 2005. Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, 53:81–94.
- Swerts, M. et Krahmer, E. 2008. Facial expression and prosodic prominence: Effects of modality and facial area. *Journal of Phonetics*, 36(2):219–238.
- Swerts, M. 1997. Prosodic features at discourse boundaries of different strength. *The Journal of the Acoustical Society of America*, 101(1):514–521.
- Swerts, M. 1998. Filled pauses as markers of discourse structure. *Journal of Pragmatics*, 30(4):485–496.
- 't Hart, J., Collier, R., et Cohen, A. 1990. *A Perceptual study of Intonation*.
Cambridge University Press, Cambridge.

- 't Hart, J. et Collier, R. 1975. Integrating Different Levels of Intonational Analysis. *Journal of Phonetics*, 3:235–255.
- 't Hart, J. 1981. Differential sensitivity to pitch distance, particularly in speech. *The Journal of the Acoustical Society of America*, 69:811–821.
- Taboada, M. et Mann, W. C. 2006. Rhetorical Structure Theory: looking back and moving ahead. *Discourse Studies*, 8(3):423–459.
- Tannen, D. 1986. Introducing constructed dialogue in Greek and American conversational and literary narrative. *Direct and Indirect Speech*, 3:11–32.
- Tao, H. et McCarthy, M. 2001. Understanding non-restrictive which-clauses in spoken English, which is not an easy thing. *Language Sciences*, 23:651–677.
- Thompson, S. A. et Couper-Kuhlen, E. 2005. The clause as a locus of grammar and interaction. *Discourse Studies*, 7(4-5):481–505.
- Thompson, S. A. et Longacre, R. E. 1985. Adverbial clauses. In Shopen, T., ed, *Language Typology and Syntactic Description: Complex Constructions*, pages 237–268. Cambridge University Press, Cambridge, UK.
- Thompson, S. A. 2002. “Object complements” and conversation: Towards a realistic account. *Studies in Language*, 26(1):125–163.
- Tolins, J. et Fox Tree, J. E. 2014. Addressee backchannels steer narrative development. *Journal of Pragmatics*, 70:152–164.
- Tomlin, R. S. 1985. Foreground-background information and the syntax of subordination. *Text*, 5(1-2):85–122.
- Trager, G. L. et Smith, H. L. 1957. *An Outline of English Structure*. American Council of Learned Societies, Washington D.C., USA.
- Traugott, E. 1999. The rhetoric of counter-expectation in semantic change: a study in subjectification. In Blank, A. et Koch, P., eds, *Historical Semantics and Condition*, pages 177–196. Mouton de Gruyter, Berlin.
- Traunmüller, H. et Eriksson, A. 1995. The perceptual evaluation of F0 excursions in speech as evidenced in liveliness estimations. *The Journal of the Acoustical Society of America*, 97(3):1905–1915.
- Tuite, K. 1993. The production of gesture. *Semiotica*, 93(1/2):83–105.
- Turk, A. et Sawusch, J. R. 1997. The domain of accentual lengthening in American English. *Journal of Phonetics*, 25:25–41.
- Turk, A. et White, L. 1999. Structural influences on accentual lengthening. *Journal of Phonetics*, 27:171–206.
- Turner, M. et Steen, F. 2013. Multimodal Construction Grammar. In Borkent, M., Barbara, D., et Hinnell, J., eds, *Language and the creative mind*, pages 255–274. CSLI Publications, Stanford, CA, USA.
- Underhill, R. 1988. Like is, like, focus. *American Speech*, 63(3):234–246.
- Van den Broeck, J. 1973. Determiners and relative clauses. *Leuvense Bijdragen*, 62(1):37–61.
- van Rijn, M. A. 2017. *The expression of modifiers and arguments in the noun phrase and beyond*. Thèse de doctorat d'état, Université d'Amsterdam, Amsterdam.

- Van Valin, R. D. et LaPolla, R. 1997. *Syntax. Structure, Meaning and Function*. Cambridge University Press, Cambridge.
- Van Valin, R. D. 1984. A Typology of Syntactic Relations in Clause Linkage. In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, pages 542–558, Berkeley, CA, USA. eLanguage.
- Vertegaal, R., Slagter, R., van der Veer, G., et Nijholt, A. 2000. Why conversational agents should catch the eye. In *Proceedings of CHI'00 Extended Abstracts on Human Factors in Computing Systems*, pages 257–258, La Haye, Pays-Bas. ACM.
- Vigliocco, G. et Kita, S. 2006. Language-specific properties of the lexicon: Implications for learning and processing. *Language and Cognitive Processes*, 21(7-8):790–816.
- Wagner, P., Malisz, Z., et Kopp, S. 2014. Gesture and speech in interaction: An overview. *Speech Communication*, 57:209–232.
- Wagner, M. et Watson, D. G. 2010. Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, 25(7-9):905–945.
- Wang, L. et Chu, M. 2013. The role of beat gesture and pitch accent in semantic processing: an ERP study. *Neuropsychologia*, 51(13):2847–2855.
- Ward, G. et Hirschberg, J. 1985. Implicating uncertainty: The pragmatics of fall-rise intonation. *Language*, 61(4):747–776.
- Watson, D. et Gibson, E. 2005. Intonational phrasing and constituency in language production and comprehension. *Studia Linguistica*, 59(2-3):279–300.
- Wells, J. 1997. SAMPA computer readable phonetic alphabet. In Gibbon, D., Moore, R., et Winski, R., eds, *Handbook of Standards and Resources for Spoken Language Systems*. Mouton de Gruyter, Part IV, section B. Berlin.
- Wells, J. 2006. *English Intonation: An Introduction*. Cambridge University Press, Cambridge.
- Wennerstrom, A. 2001. *The Music of Everyday Speech. Prosody and Discourse Analysis*. Oxford University Press, Oxford.
- Wichmann, A. 2000. *Intonation in Text and Discourse*. Longman, Londres.
- Wichmann, A. 2001. Spoken parentheticals. *Gothenburg studies in English*, (81):177–193.
- Wightman, C., Shattuck-Hufnagel, S., Otsendorf, M., et Price, P. J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of Acoustical Society of America*, 91:1707–1717.
- Willems, N., Collier, R., et 't Hart, J. 1988. A synthesis scheme for British English intonation. *The Journal of the Acoustical Society of America*, 84(4):1250–1261.
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., et Spence, C. 2015. Conducting perception research over the internet: a tutorial review. *PeerJ [online]*, 3(1058).
- Wyld, H. 2003. Adverbial clauses: an enunciative approach. In Celle, A. et Gresset, S., eds, *La Subordination en Anglais: une approche énonciative*, pages 15–38. Presses Universitaires du Mirail, Toulouse.
- Yngve, V. H. 1970. On getting a word in edgewise. In *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, pages 567–578, Berkeley, CA, USA. eLanguage.

- Yoon, T.-J., Chavarria, S., Cole, J., et Hasegawa-Johnson, M. 2004. Intertranscriber Reliability of Prosodic Labeling on Telephone Conversation using ToBI. In *Proceedings of INTERSPEECH 2004*, pages 2722–2732, Jeju, Corée du Sud.
- Yoon, T.-J., Cole, J., et Hasegawa-Johnson, M. 2007. On the edge : acoustic cues to layered prosodic domains. In *Proceedings of ICPhS XVI*, pages 1264–1267, Sarrebruck, Allemagne.
- Yule, G. 1980. Speakers' topics and major paratones. *Lingua*, 52(1):33–47.
- Zima, E. 2014. English multimodal motion constructions. A construction grammar perspective. *Studies van de BKL - Travaux du CBL - Papers of the LSB*, 8:[en ligne].

Index

- affiliation interactionnelle, 109, 127, 129, 131, 133, 135, 136, 152, 162, 253, 255
- alignement interactionnel, 95, 109, 115, 127, 131, 133, 135, 136, 162, 231, 243, 255
- Analyse Conversationnelle, 23, 40
- Analyse de Discours Multimodale, 1, 11–14, 24, 34, 40, 299
- Analyse du Discours, 1, 4, 23, 39
- ancrage informationnel, 31, 251
- appositives, 2, 5–9, 15–18, 44, 76, 301, 310
 - canoniques, 8, 76
 - relatives, 7, 8, 46, 47, 76
- approche IPO, 177
- arrière-plan informationnel, 5, 11, 43, 50, 56, 224, 311
- assemblage symbolique, 32, 33
- backchannel, 67, 95, 106, 127–131, 133, 143, 152, 154, 213, 216, 243, 253, 301
 - minimal, 15, 95–97, 101, 102, 127, 133, 135, 136, 152, 155, 161, 222, 223, 231, 243
 - substantiel, 126–130, 134, 259
- cadre interprétatif, 2–4, 9, 48, 53, 54, 56, 103, 104, 109, 110, 112, 115–117, 134, 153–155, 170, 200, 224, 227, 234, 236, 238, 300, 304–308
- charge cognitive, 30, 227
- circonstanciennes, 2–4, 9, 15, 17, 44, 47, 76, 300, 309
 - conditionnelles, 2, 3, 76
 - temporelles, 2, 3, 76
- cognition incarnée, 25
- connaissances partagées, 25, 31, 73, 109, 154, 155, 159, 163, 210, 211, 217–219, 230, 233, 242, 248, 261, 299, 303
- coordination syntaxique, 9, 46, 48, 51, 53, 92, 106, 135
- délexicalisation, 185, 188, 191, 192, 198
- densité modale, 209, 215, 239, 262
- discours écrit, 10, 36, 63, 64, 229
- disfluente, 97, 98, 110, 121, 122, 182, 209, 210, 233, 235
- Distributed Little Red Hen Lab, 28
- ELAN, 66, 75, 81
- emphase, 33, 34, 38, 120, 128, 211, 215, 216, 245, 248, 252, 253, 255, 257, 259, 277, 306
- énoncés composites, 28
- ENVID, 13, 15, 16, 18, 61, 183, 280
- espace interactionnel, 11, 73, 163, 217, 242
- eSurv, 189, 191, 192, 285, 286
- face conversationnelle, 100, 224, 242
- focus, 11, 35, 37, 53, 111, 116, 121, 132, 135, 210, 211, 218, 226, 233, 248, 249, 275, 277

- fonction non-référentielle, 68, 72, 73, 228, 235, 274
 fonction référentielle, 68, 70, 72, 213, 218, 227, 231, 241, 245, 273, 274, 282
 Gesture as Simulated Action, 227
 gesture catchments, 213, 214, 219, 220, 222, 243, 256, 259, 270, 277
 grammaire cognitive, 14, 29, 30, 32, 34, 48, 53, 304, 309
 grammaire des constructions, 14, 25–29, 32, 34, 304, 308, 309
 grammaire générative, 26, 30
 histoire conversationnelle, 25, 109, 154, 159
 insubordination, 105, 106
 IPA, 64
 linguistique interactionnelle, 12, 14, 34
 marqueur de discours, 97, 101, 106, 108, 111, 123, 125, 129, 130, 132, 145, 147, 154, 159, 161, 162, 169, 223, 231, 244, 308
 modèle autosegmental de phonologie intonative, 39, 40, 175–177, 182
 modèle linéaire mixte généralisé, 82, 191, 205, 286, 295
 modification syntaxique, 1, 8, 9, 44–46, 49, 53, 98, 121, 302
 Momel Intsint, 38–40, 65
 objets de discours, 24, 97, 116
 parenthétiques, 10, 135, 164, 170, 171
 pause remplie, 16, 132, 151, 153, 158, 168, 171, 180, 202, 252, 306
 pause silencieuse, 55, 75, 76, 141, 166, 173, 174, 180, 181, 184, 186, 187, 194, 200, 230, 241
 d'hésitation, 74, 142, 151, 158, 168, 169, 180, 233, 246
 de démarcation, 94, 113, 140–142, 144, 157, 232
 de focalisation, 140, 141, 147, 158, 163
 phases gestuelles, 27, 231, 271
 phrases gestuelles, 66, 271
 Praat, 15, 63, 66, 75, 79, 184–188
 premier plan informationnel, 9, 11, 43, 56, 264, 311
 proéminence, 39, 42, 212, 252, 263, 264, 268, 271, 272, 279
 propositions commentatives, 10
 recherche lexicale, 73, 74, 142, 209, 224
 rection grammaticale, 3, 49, 59, 87, 88, 90, 101–104, 108, 114, 118, 119, 133, 300
 récursivité, 10, 44
 relatives déterminatives, 2, 5, 6, 9, 15, 17, 18, 46, 76, 300, 309
 réorientation interprétative, 9, 15, 18, 136, 170, 304
 repère constitutif, 30, 99
 reprise de souffle audible, 111, 125, 130, 151, 152, 155, 156, 169, 224
 Rhetorical Structure Theory, 50
 SAMPA, 64
 segmental, 24, 155, 156, 158, 163, 170, 301
 Segmented Discourse Representation Theory, 50
 signaux de contextualisation, 52
 SPPAS, 64
 structure attentionnelle, 30
 structure informationnelle, 30, 33, 35, 40, 229, 248, 311

- structure présentative, 92, 129, 133, 145,
160
- subordonnées complétives, 4, 9, 10, 44, 46
- suprasegmental, 16, 24, 148, 155, 170, 171,
301
- synthèse vocale, 186
- système d'annotation ToBI, 179, 182, 191,
270, 279, 286
- technique de “rush-through”, 142, 143,
151, 152, 209
- théorie du Growth Point, 42, 269
- tradition britannique de l'intonation, 14,
37–39, 63, 175
- Transition Relevance Place, 142, 143, 167,
211, 244, 257
- travail de formulation, 36, 64, 142, 145,
151, 152, 158, 163, 227

Annexe A

Formulaire de consentement éclairé

Le modèle de formulaire de consentement éclairé utilisé pour notre contribution au corpus ENVID (un dialogue, 30 min) se trouve ci-dessous.

Constitution d'un corpus en vue d'une recherche en linguistique

Vous êtes invité à participer à la constitution d'un corpus en vue d'une étude de linguistique. Le but de cette étude est une analyse de l'anglais oral.

Si vous décidez d'y participer, il sera demandé d'entretenir une conversation d'une demi-heure, qui sera enregistrée et filmée.

Les données seront seulement utilisées dans un contexte de recherche ou d'enseignement, sous forme textuelle (transcriptions) ou sous forme d'images. La transcription pourra être diffusée en intégralité (dans une thèse) ou partiellement (sous forme de courts extraits dans des articles).

Concernant les données vidéo, l'image du participant est également susceptible d'être utilisée et publiée dans une thèse, ainsi que dans de possibles articles de conférence scientifique.

L'investigatrice principale s'engage à mener cette recherche selon les dispositions éthiques et déontologiques, et à assurer la confidentialité des informations recueillies

en dehors du cadre universitaire décrit plus haut.

Il n'y a aucun risque connu à la participation à cette recherche.

Le participant ne sera identifié que par son prénom. Le participant a la possibilité d'obtenir des informations supplémentaires concernant cette étude auprès de l'investigateur, et ce dans les limites des contraintes du plan de recherche. Un résumé des résultats de la recherche pourra être délivré sous format électronique à la demande du participant après leur validation.

La participation à cette étude est entièrement volontaire.

Je soussigné, X, déclare avoir lu les informations ci-dessus, ainsi qu'avoir obtenu une réponse satisfaisante à toute question posée. J'accepte de participer à cette recherche, je possède une copie de ce formulaire.

Creation of a corpus in linguistic studies

You are invited to participate in a study of linguistics. The purpose of the study is an analysis of spoken English.

If you decide to participate, you will be asked to have a 30-min conversation that will be recorded and filmed.

The collected data will only be used in an academic context, either as a text (transcriptions) or as images. The transcription might be published as a whole (in a thesis) or in short extracts (in papers).

The main investigator is committed to leading this study in an ethical way, as well as to preserving the confidentiality of the data in any other context than university.

There is no particular risk linked to the participation in this study.

Participants will only be identified by their first names. Extra information can be given to participants by the main investigator, provided it does not affect the study in itself. A summary of the results can be made available to participants on

request after the validation of the study.

Participation in this study is entirely voluntary.

I, X, have read the information above and any questions I have asked have been answered to my satisfaction. I agree to participate in this research. I have been given a copy of this form to keep.

Annexe B

Indications des tests de perception

Indications aux participants du test des frontières prosodiques

“Welcome to the study!

In this experiment, I'm interested to know what sort of boundary native speakers of English perceive when they listen to a complex utterance. A boundary is any cue which acts as a separator between parts of speech. If you agree to take part in this study, you'll be asked to complete a short questionnaire that gives me general information on who you are although the survey will remain anonymous. You'll also be asked to complete a questionnaire measuring if you perceive any boundaries and how strongly you may perceive them in 52 utterances. These utterances are in British English.

The questionnaire should take approximately 30-40 minutes to complete.

You will be required to listen to some sound files and I strongly recommend that you use headphones for better hearing quality. The pages containing sound files may take a few seconds to load. Each page presents a block of ten questions with their associated sound files. The sound excerpts have been modified by speech synthesis so they may sometimes sound a bit unnatural and you may as well not be able to distinguish the words said, but I would like you to judge from your intuition the strength of the boundary or break at the place in the utterance marked with a red *. You are allowed to listen to the sound file two or three times if that helps you make a decision.

Your answers won't be saved until you click the finish button at the end of the experiment so make sure you have a long enough time span to complete the questionnaire. You may exit the survey at any time by simply closing the window in your browser in which case none of your answers will be recorded. At the end of the experiment, you'll have the opportunity to leave feedback when you have completed the questionnaire but, if you do have any questions, you can contact me by using the details at the bottom of this page.

By clicking "Next", you agree to participate in the experiment.

Thank you for your participation."

Indications aux participants du test des frontières gestuelles

"Welcome to the study!

In this experiment, I'm interested to know what sort of boundary native speakers of English perceive when they watch short extracts of speech. A boundary is any cue which acts as a separator between parts of speech. If you agree to take part in this study, you'll be asked to complete a short questionnaire that gives me general information on who you are although the survey will remain anonymous. You'll also be asked to complete a questionnaire measuring if you perceive any boundaries and how strongly you may perceive them in 24 video clips. These video clips contain speech in British English.

The questionnaire should take approximately 15 minutes to complete.

You will be required to watch some video files containing sound and I strongly recommend that you use headphones for better hearing quality. The pages containing video files may take a few seconds to load. Each page presents a block of ten questions with their associated video files. You may not be able to distinguish the words said, as some extracts contain fast speech, but I would like you to judge from your intuition the strength of the boundary or break at the place in the utterance marked with a red *. You are allowed to watch the video files two or three times if that helps you make a decision.

Your answers won't be saved until you click the finish button at the end of the experiment so make sure you have a long enough time span to complete the questionnaire. You may exit the survey at any time by simply closing the window in

your browser in which case none of your answers will be recorded. At the end of the experiment, you'll have the opportunity to leave feedback when you have completed the questionnaire but, if you do have any questions, you can contact me by using the details at the bottom of this page.

By clicking "Next", you agree to participate in the experiment.

Thank you for your participation."

Annexe C

Liste des subordonnées sélectionnées

Relatives déterminatives

1. she would be the one **that i would have thought would be most # kind of # problematic** but she hasn't been
2. the first # year **my brother was in uni #** think we went there just for lunch
3. cause some people **that # haven't been #** the Spanish girls that were there #
4. the Spanish girls **that were there #** on our second one
5. that's an Ikea one **they've bought themselves #** like like little bits and pieces of it
6. live in a place **where you can be # like self-sufficient #** and have your own # garden #
7. you know these pots **that have like # holes in them # like funnelly bits yeah** they're good for strawberry growing apparently #
8. the ones **that dangle down #** yeah
9. you know and the little bits **that stick out #** you have those
10. and stuff **that's gonna like crawl along #** so it's like shallow
11. the guy **that lives there #** said he'd help us
12. (h) and so she # disowned everything **that # she could associate my nana #** for example her accent # and #
13. do you remember that first night **when everyone turned up** and i # insulted

14. oh she's that # woman **that # looks after the Nottingham crowd (laughs)** that woman #
15. i met people **who like # (h) haven't been to #** well they've probably been to first lesson
16. it was better than that other place **i was at #** Rochefort # (laughs)
17. ten # girls **who were # in their final year #** so they must have been like eighteen year olds
18. but i put it on the bit **where hem (cough) (h) they were uh # in the garden** and they were talking
19. everything **that was going on #** ooh would you like another s' # would you like another biscuit
20. she said it's an American **she met from New York #** and he was on a business trip or something
21. only friends **that haven't met yet #** (laughs) probably really cheesy
22. we'd get a present **that he'd # had like # two years ago** or something really silly
23. and my friends **that were here** they didn't go # at all
24. i hate speaking to people **that i don't know that well #** yeah but every si- it's nice
25. but how do we explain to people **that obviously use computers a lot more than me #** (h) it's not very good #
26. and the daughter **that she's had all her life #** the # cos there's the black one
27. i've never seen the Irish bar then **that you're talking about** i don't know what you're talking about
28. i'm coming back for someone **that's not even there (laughs)** (h) it was awful #
29. to pick up anything **that was there #** and i just had my hand there
30. it's the guy # from **who was in the # hem # the show** but he didn't win
31. and everything else **that's sweet** like chocolate and anything #
32. you know the one **where they dress up in underwear #** there was like hem
33. uh # for (h) people **who speak good English** so hopefully #

34. but the only # place **where you can do that** is in Paris
35. all the people i know **who have a job** is because # mum works there or
36. uh the school # **which is right # below** that uh rents # the flat out
37. Paris is the only place **you could ever live in #** any other town
38. (h) in the # test **they gave us** there were actually m- # grammar mistakes
39. but everyone else **i've spoken to** is having trouble finding something #
40. and navigate to somewhere **i don't know (h) #** and like # big roundabouts
and #
41. cause it's the one thing **that you don't do** b- hem # you're not allowed to
do
42. the reasons **they gave #** i mean Blair was also #
43. had to rely on false # information **that they thought was correct #** i
mean #
44. around the corner **where a slip road joins the motorway #** and i think
someone must have gone down the slip road
45. he ended up having lessons from everyone **he knew** you know you can #
give lessons
46. who lived down in the old estate **where i lived #** yeah
47. with this guy **who does it #** he does like a Top Gear programme
48. looking at anything **you pick #** i'm not sure of his name
49. the things **he looks at #** but they're very
50. recently **that they published this #** er surv- oh not survey like a # a
piece of work
51. went and found this salmon farmer in Scotland **who's # lost forty percent
of his business literally overnight** and had to scale down and #
52. they got the guy **who had written the report** and he was saying
53. like lead to the cells **that develop cancer** but # and again # why not
54. check the fact **that it says # eighty percent # good meat** kind of i
don't know what it would say
55. i always wonder how the people # **who are there as # participants #**
see it

Circonstanciennes

1. oh it's really shiny **when it's all straight** and like it looks really nice
2. that with the honeysuckle # **if you take it off** and just squeeze down
3. so that **if it grows out** it's not a problem #
4. cos someone # **when they had that party** # do you remember that party on #
5. and that's happened before as well **when # Holly and Beth were really drunk** # and Gemini drove them home
6. i hope so # **if i go to dance tonight** i'll ask her
7. but **if i don't** i'll ask Lydia
8. yeah **if you go # a bit # like along Park Place and up Newport Road** # it's not that far
9. because sometimes # **(h) when everybody went home** i'd stay for a weekend # and #
10. my dad used to teach in Hebburn **when he was first starting teaching** and he was getting harassed by all the pupils
11. but **if you # ring me #** you can buy it today
12. speaking French **when i wasn't used to it** all the time #
13. well you know i mean i **if i was had a computer in front of me** i'm sure i could work it out #
14. it's really kind of like # pathos **when you saw her mother's face** because she just treated her so badly
15. i'd panic **if we got stuck under the tunnel** i'm hyper claustrophobic #
16. you didn't wear a c- # a corset **when you were thirteen** did you (laughs)
17. and **if it's some ten pounds a head** then you get like #
18. like you know **when you imagine it's gonna be amazing** # but apparently it wasn't
19. (h) # and **when she comes to get me** it's # kind of far #
20. so that for example **if there's a terrorist attack** (h) hem # there are loads of solutions
21. well **if you wanna do it on a long # you know on long-term basis** then (h) have a number of #
22. and **when i see # something coming towards me** i kind of go hhh

23. i tried driving once in her car **when we were on a # little road in the countryside #** and hem (swallows) she said well turn left #
24. so **# if i ever want to # to do a turn-o** it's kind of covered
25. (h) but hem (swallows) **# when you kind of go # (h) around #** hem it's a bit # yeah
26. i actually uh **when they said no** i actually called the guy
27. and then **when i arrived #** i heard of that # French and British programme
28. (h) you know **when you're not allowed to laugh #** (h) and then there's like a massive silence #
29. especially in class **when someone says something funny** and i'm always # dying
30. so what would you do **if you did that** would you be like a lawyer or #
31. i suppose **when you do law** you kind of do have to speak a minimum # amount of English
32. you know **when you do languages #** (h) you have to do
33. you know **when there's turbulence #** (h) i always look at them i'm like
34. (h) i always look at them i'm like **(h) if they're standing there smiling** then it's fine (laughs)
35. and **if there's one of them that goes wrong** your whole journey is completely messed up #
36. it's like **(h) if i do something stupid** then it's other people's lives
37. it's a good thing to have **if you ever need it** i mean
38. (h) i'm okay **when i know where i'm going** it's # (h) # like if i know exactly the
39. whatever like **if i know the route** it's fine
40. well i guess **if the if the exams are written by French # people and French speakers** no matter how good their English is
41. in London **if you're paying London rent** it's gonna be like what #
42. although **when i passed** i didn't think i was a very good driver #
43. so you can # go and to **when you're finished** you get a job in
44. i suppose **if they'd had Cardiff City** then they wouldn't be that good #
45. sometimes **if you have close calls** you know you can't stop
46. you can't turn your neck # **if you try to drive** you're like #

47. yeah definitely **if it'd been a much more # bigger car behind me** then i'd probably've had some
48. and Lynn's came back from uni **when i was twenty-one** kind of hanging round with them again
49. and hanging round with me **when i was eighteen #** and so i got to know them
50. last summer **when we were doing that little road trip further there #** for a holiday instead of going abroad
51. because # **when they got married #** they bought a flat
52. because they did not **if you had a burglar alarm box #** you didn't get burgled #
53. you know **if you go into most # leisure centres** you've got hard floor
54. i'm gonna # get it **when it comes out on video #** was it good
55. make sure # **if you're buying sausages #** check the fact

Appositives

1. cos i was gonna go with Lowri # **my friend from home #** hem # on # like like she said
2. you know my d- where my dad's from # **Galicia in North West # of Spain #** that's like #
3. hem cebolla # **which is # onion #** and # hem # lettuce
4. Woodville Road # uh uh **which is more the way** is that more the way he went
5. Holly went out # **which is a bit unsocial really but #** hem # and # Beth had locked her room
6. but someone # **which is a bit silly** had broken into Beth's room #
7. she's good friends with Jenna # and # and Cath # **who she does live with #** but Cath always goes out with the other optometrists
8. you know Jamie's brother **Tom #** he's had like a thing with her
9. and hem # he **like the Sainsbury's man (laughs)** and hem # the pizza bases have got a crack in them
10. Keith # **the guy that lives there #** said he'd help us
11. in Gwyn Fechan # **which is # just a place it's not even a village #** hem # which # and the nearest town is Crickhowell

12. they've got some really nice crafty shops **gifty shops** # yeah #
13. and they have a really nice garden **you can sit in** # yeah #
14. i never see any of my French people **my French hem flatmates** they never talk to me you know
15. i was living with this Scottish girl called Nicola # **Nicola (laughs)** (h) and just upstairs
16. and it's got this plaque on the wall **that i like reading** it says (h)
17. and i'm from Durham **which is in the north-east of England** # (h) have you ever been there
18. one's in # Tyne and Wear **which is Newcastle** # (h) # and
19. every weekend **which was nice** # but i just #
20. well for the train **that i hope # you're gonna get and i've got** is the nine o'clock #
21. and then we went into # a place called Tropicana **which was horrible (laughs)** # it's on Saint Mary Street
22. a big this one # **where you can drive to in** # one of the plac-
23. and it's right at the bottom of Park Street **which is the best street to go shopping in ever** # so yeah
24. in Nantes **which i might do** # (h) and #
25. even my mum **who's an English teacher** didn't get a hundred percent so
26. very well paid # **which is why** # (h) **i was really looking forward to it and** # they called me back and said
27. even compared to Easyjet **which is another low cost company** # yeah i hate Ryanair #
28. i kind of go hhh # **which is really not** # n- w- when it's a car it's
29. i mean my mum's pushing me to get my license (h) **uh which i guess i should** # (h) but # well first of all #
30. cos my # French grandmother # **my # dad's mum** # is #
31. basically she had this deal with us # **with # every single one of her grandchildren** is (h) if you pass your baccalaureate
32. i don't know if you've heard of it **UCL** # (snuffles) and hem
33. so i did the second year of hem # like English # here (h) **(laughs) which was interesting** (h) hem yeah and then i just graduated now #
34. for fifty pounds # **which was really amazing** # but then that was #

35. and (h) it took us to Limoges # **which was like # what like # four hours away from here #** (h) so yeah that was pretty grim
36. but he said **the guy** said that like
37. it's only one shop for the whole thing now **which is quite bad #** yeah #
38. you get an assessed piece of work **which you do on a computer** using a program called author catway
39. redevelopments # **which # we need to learn how to use #** but it should be quite interesting
40. did you see # like Michael Moore's # film **Fahrenheit nine eleven #** yeah #
41. oh we went to hem # music pub **One Four Seven #** open mic
42. hem well just the regulars **the locals #** the band called The Misfits #
43. the locals # **the band called The Misfits #** and
44. and they played this Irish # tune **which was awesome #** this guy with hem # a whistle
45. hem Hannah # **my friend #** you know Hannah
46. but some of them # **a couple of my friends** just won't drive on a motorway
47. there's North Pentwyn **which is not North of Pentwyn (giggles) #** so why d'they call it North Pentwyn
48. because it was all # fenced off **the whole road #** we used to play street hockey in the street
49. and football and that **which was great for us** but #
50. you know Glenwood Church **the big hall #** it's got blue carpet down now
51. it's got blue carpet down now **which is supposed to be all-s # all-purpose** but # you can't skate on it
52. well i looked on the website # **on BBC website** earlier on today
53. he had a pig's head # **this is a butcher #** and he like cut
54. i'll happily eat black pudding **which i know is disgusting #** blood and guts and #
55. i look forward to it every week **which is how we were feeling last year when we were on it** we'd get to like # a Saturday

Résumé

À partir d'un corpus vidéo de conversation spontanée en anglais, notre travail de thèse s'attache à déterminer si plusieurs types syntaxiques de constructions subordonnées expriment le même degré d'intégration à leur environnement co-textuel, d'une perspective multimodale. La littérature syntaxique décrit les subordonnées comme des formes dépendantes, qui spécifient ou élaborent le contenu d'une autre proposition. En montrant que les constructions sous étude n'expriment pas une dépendance uniforme à leur environnement selon la façon dont les locuteurs utilisent les modalités prosodique et gestuelle pour exprimer plus ou moins de démarcation, les résultats en production comme en perception suggèrent d'une part que les appositives sont produites avec davantage de rupture que les autres types syntaxiques, et d'autre part que la création d'une rupture s'appuie majoritairement sur des moyens davantage prosodiques que gestuels.

Mots clés : Syntaxe; frontières; subordination; prosodie; gestualité; Analyse de Discours Multimodale.

Abstract

Based on a video recording of conversational British English and within the framework of Multimodal Discourse Analysis, this study tests whether three different syntactic types of subordinate structures are evenly integrated to their environment. Subordinate constructions have been described in syntax as dependent forms elaborating on primary elements of discourse. Beyond showing that subordinate constructions are not evenly dependent on their environment depending on how speakers use the prosodic and kinetic modalities to express greater (in)dependency, our results in production as in perception suggest on the one hand that appositive clauses show more break than the other syntactic types, and on the other hand that the creation of a break mainly relies on prosodic cues.

Keywords : Syntax; boundaries; subordination; prosody; gesture; Multimodal Discourse Analysis.

Titre : Expression multimodale de la subordination en anglais

Mots clés : Syntaxe; frontières; subordination; prosodie; gestualité; Analyse de Discours Multimodale

Résumé : À partir d'un corpus vidéo de conversation spontanée en anglais, notre travail de thèse s'attache à déterminer si plusieurs types syntaxiques de constructions subordonnées expriment le même degré d'intégration à leur environnement co-textuel, d'une perspective multimodale. La littérature syntaxique décrit les subordonnées comme des formes dépendantes, qui spécifient ou élaborent le contenu d'une autre proposition.

En montrant que les constructions sous étude n'expriment pas une dépendance uniforme à leur environnement selon la façon dont les locuteurs utilisent les modalités prosodique et gestuelle pour exprimer plus ou moins de démarcation, les résultats en production comme en perception suggèrent d'une part que les appositives sont produites avec davantage de rupture que les autres types syntaxiques, et d'autre part que la création d'une rupture s'appuie majoritairement sur des moyens davantage prosodiques que gestuels.

Title : The multimodal expression of subordination in English

Keywords : Syntax; boundaries; subordination; prosody; gesture; Multimodal Discourse Analysis.

Abstract : Based on a video recording of conversational British English and within the framework of Multimodal Discourse Analysis, this study tests whether three different syntactic types of subordinate structures are evenly integrated to their environment. Subordinate constructions have been described in syntax as dependent forms elaborating on primary elements of discourse.

Beyond showing that subordinate constructions are not evenly dependent on their environment depending on how speakers use the prosodic and kinetic modalities to express greater (in)dependency, our results in production as in perception suggest on the one hand that appositive clauses show more break than the other syntactic types, and on the other hand that the creation of a break mainly relies on prosodic cues.