



LUND UNIVERSITY
Joint Faculties of Humanities and Theology

Centre de langues et de littérature

HT 2013

L'intonation montante en français spontané

Continuation vs. Question

Mémoire de Master II

EMMA VALTERSSON

Directeurs de mémoire:

Paul Touati (Université de Lund)

Francisco Torreira (Max Planck Institute for Psycholinguistics)

TABLE DES MATIÈRES

| | |
|---|----|
| TABLE DES MATIÈRES..... | 1 |
| 0. INTRODUCTION | 2 |
| 1. PARTIE THÉORIQUE..... | 5 |
| 1.1. INTONATION | 5 |
| 1.2. CONTINUATIONS..... | 7 |
| 1.3. QUESTIONS..... | 8 |
| 1.4. DIFFÉRENCE ENTRE LA CONTINUATION ET LA QUESTION MONTANTE..... | 9 |
| 2. HYPOTHÈSE ET PRÉVISIONS..... | 13 |
| 3. MÉTHODE | 15 |
| 3.1. « SEMI-MANUAL-EFFORT METHODOLOGY » | 15 |
| 4. MATÉRIEL..... | 27 |
| 4.1. THE NIJMEGEN CORPUS OF CASUAL FRENCH | 27 |
| 4.2. ORGANISATION SPATIALE | 28 |
| 4.3. AUDIO – DEUX CANAUX..... | 29 |
| 5. RÉSULTATS | 31 |
| 5.1. EXEMPLES TIRÉS DU CORPUS | 31 |
| 5.2. SIGNAUX FINAUX..... | 33 |
| 5.3. SIGNAUX NON-FINAUX | 41 |
| 5.4. LE RÔLE DE LA SYNTAXE..... | 44 |
| 5.5. LE POIDS DES VARIABLES DANS LA CLASSIFICATION DES FONCTIONS..... | 45 |
| 5.6. CONCLUSIONS DES RÉSULTATS ET TRACES DE SYNTHÈSE DE LA F0 | 48 |
| 6. DISCUSSION | 53 |
| RÉFÉRENCES BIBLIOGRAPHIQUES | 57 |

0. INTRODUCTION

Phénomène linguistique multifonctionnel, l'intonation favorise l'empaquetage du message linguistique, sa production, par le locuteur aussi bien que sa compréhension par l'interlocuteur et ceci grâce à une mise en hiérarchie de l'information. Alors que cette multifonctionnalité facilite en direct la communication entre les locuteurs en véhiculant des informations en ce qui concerne les émotions, de l'organisation séquentielle de la conversation, elle en complique *l'analyse du sens* dans la mesure où une et même intonation a plusieurs portées. Cette multifonctionnalité se reflète dans l'analyse linguistique qui est effectuée. En effet, certaines approches empiriques ou théoriques adoptent une perspective analytique qui met en relief soit son rôle dans la réalisation d'un énoncé ou d'un paragraphe isolés, ou ultérieurement la manière dont elle est supposée signaler à l'interlocuteur de l'information liée au changement du locuteur.

La recherche sur l'intonation du français a gagné du terrain dans les années soixante lorsque deux grands laboratoires au sud de la France ont commencé à focaliser leur recherche sur ce phénomène. Grâce aux avancées de la technologie, ces laboratoires ont pu effectuer une recherche empirique systématique des faits suprasegmentaux, dépassant ainsi le stade des analyses auditives et des impressions subjectives. Ceci a eu pour conséquence une explosion des publications consacrées à la recherche sur les faits prosodiques. Ces avancées technologiques, maintenant mises au service de chacun qui dispose d'un ordinateur, ont rendu possible notre propre recherche qui se veut empirique et quantitative. C'est ainsi que nous avons décidé de nous attaquer à une problématique classique de l'intonation du français, à savoir la description quantitative et qualitative des différences entre les montées intonatives dites de continuation et de question.

Un rapide état de lieux de la recherche sur l'intonation du français a mis en évidence un désaccord concernant la description et l'interprétation de ces deux montées de la fréquence fondamentale. Certains chercheurs (cf. Delattre, 1966 ; Rossi et al., 1981) considèrent qu'elles sont différentes au sujet de sa composition intonative (par exemple une différence au niveau de la fréquence fondamentale atténuée). Alors que d'autres chercheurs encore les considèrent comme identiques, elles seraient en quelque sorte des versions différentes d'un même contour (Di Cristo, 1978). Les travaux plus récentes de Portes et Bertrand (cf. Portes et Bertrand, 2007a ; entre autres) acceptent une différence structurelle entre les deux montées qui sont censées présenter différentes conséquences en ce qui concerne

la structure de la conversation (la montée de la continuation implique une suite par le même locuteur, tandis que la montée de la question contient l'attente d'une réponse par l'interlocuteur). D'autres chercheurs encore (cf. le travail de Marandin et al., 2004) les considèrent également différentes, leur interprétation est plutôt liée à la réaction anticipée par le locuteur sur le contenu du message qu'elles accompagne. Cependant, ces travaux plus récentes qui semblent accepter une distinction entre les deux contours ne font aucune référence à sa configuration intonative. Il nous semble par conséquent primordial d'étudier la différence intonative proposée ou réfutée par les recherches antérieures, qui ne disposaient pas des mêmes possibilités techniques, à l'aide d'une étude de corpus des conversations spontanées.

Nous avons remarqué (conformément à Léon et Bhatt, 1987) que le contexte interactionnel est informatif lors de l'identification d'une question. D'après nos observations initiales, il nous semble que la distinction entre la continuation et la question montante n'est pas problématique quant à son intégralité dans son contexte, alors qu'exclu de celui-ci elles ne sont pas aussi facilement distinguables. L'organisation syntaxique ne permet pas non plus une identification correcte puisque les contours présentent toutes les deux le même ordre (SVO) la plupart du temps. Étant donné qu'elles ont été décrites comme différentes ou similaires au sujet de ses caractéristiques intonatives (cf. Delattre, 1966 ; Rossi et al., 1981 ; Di Cristo, 1978), notre but avec cette étude est simplement d'examiner si une identification basée sur l'intonation seule est envisageable. Autrement dit, nous voulons voir si les continuations sont discernables des questions au cas où toutes les informations supplémentaires à l'intonation (par exemple le contexte, la syntaxe, l'autorité épistémique) sont exclues et les contours sont considérés en isolation. Un but secondaire est de discuter autour le rôle potentiel des mesures au sujet de l'alternance des tours de parole.

Nous nous proposons de donner un rapide état de lieux sur la recherche en ce domaine, un chapitre qui inclut la description des continuations et des questions ainsi que la différence entre ces deux groupes. D'après les informations évoquées dans la partie théorique nous avons émis une hypothèse et des prévisions au sujet des différences entre les continuations et les questions manifestées par l'intonation dans le Chapitre 1. Cependant, avant de centrer notre attention sur le thème de notre propre travail, nous offrons une courte introduction au phénomène de l'intonation (Ch. 1.). Le premier chapitre est alors focalisé sur les deux contours mélodiques étudiés dans ce travail, ainsi que leur rôle potentiel dans

l'alternance des tours de parole. Dans le troisième chapitre, nous décrivons en détail notre propre méthodologie avec une focalisation sur les avantages d'une approche empirique et quantitative basée sur un corpus de données spontanées (voir. Ch. 4 pour une description des données). Finalement, les résultats statistiques (Ch. 5) et les conclusions que nous en avons tirés (Ch. 6) sont présentés.

1. PARTIE THÉORIQUE

Ce chapitre nous offre la description du phénomène – l’intonation – qui est central dans ce mémoire et des classifications différentes qui ont été proposées en ce qui concerne ses composantes et leurs interrelations. Nous évoquons par la suite les descriptions des deux contours montants étudiés dans ce travail, afin de finalement les mettre en parallèle et discuter de ses similitudes et de ses différences à partir des propositions avancées dans la littérature.

1.1. INTONATION

Avant de centrer notre attention sur le sujet de ce mémoire, c’est-à-dire du rôle distinctif de l’intonation dans des séquences conversationnelles (dépendant du contour intonatif), il nous semble nécessaire d’effectuer un survol des phénomènes suprasegmentaux. Aussi bien de leur interprétation au sein du système intonatif du français ainsi que de la terminologie qui lui est propre. Un terme générique est fréquemment utilisé lors du regroupement des phénomènes accentuels, rythmiques et intonatifs, c’est celui de prosodie. On affirme ainsi que la prosodie sert à l’interprétation des énoncés en permettant « la reconstruction des structures syntaxique, sémantique et pragmatique » (Avanzi et Delais-Roussaries, 2011 : 2) ; elle se situe donc « au cœur de la communication langagière » (Lacheret, 2011 : 87). Rossi (1985 ; cité par Touati, 1987 : 23) a suggéré que la prosodie assume une fonction modale, une fonction expressive et une fonction d’organisation, ce qui montre à l’évidence la multifonctionnalité de ce phénomène. Lacheret (2013 : 1) quant à elle limite cette multifonctionnalité à seulement deux fonctions, celle du groupement des unités de sens conditionné par l’intonation et celle de la mise en relief associée à l’accentuation. Bien que les chercheurs ne soient pas d’accord sur le nombre de fonctions assumées par la prosodie, il est néanmoins raisonnable de constater que sa propriété multifonctionnelle complique l’interprétation d’un fait prosodique en empêchant la reconnaissance d’un lien direct entre lui et ce qu’il communique. Avanzi et Delais-Roussarie (2011 : 2) ont remarqué cette problématique en constatant qu’une « manifestation prosodique et une fonction linguistique ou discursive » ne sont pas reliées d’une manière univoque. La relation directe entre un mouvement de la fréquence fondamentale et un signifié étant impossible, la recherche des faits prosodiques se retrouve face à un problème puisque une montée fréquentielle ne signifie pas toujours une question, étant donné qu’elle peut également être accompagnée d’une chute de la f_0 (Martin, 2004).

La Figure 1 ci-dessous illustre la conception de la prosodie appliquée dans ce travail (empruntée à Lacheret-Dujour, 2011) et indique les paramètres acoustiques qui nous serviront de manifestation empirique dans notre analyse (l'intonation – c'est-à-dire la fréquence fondamentale, l'intensité et la durée). Il convient de noter que les phénomènes acoustiques présentés et décrits dans le reste de ce chapitre n'illustrent que la description empirique de ces paramètres, sans qu'il soit possible de dire avec certitude si les locuteurs les perçoivent de la même manière que l'analyste. Il s'agit de permettre un point de vue analytique modélisé et susceptible de prédire des manifestations pertinentes de la variation intonative et non pas d'une « spéculation » argumentée tendant à retrouver des « traces » du travail de production interactif des locuteurs. L'effort analytique est ainsi centré sur la structuration intonative du message.

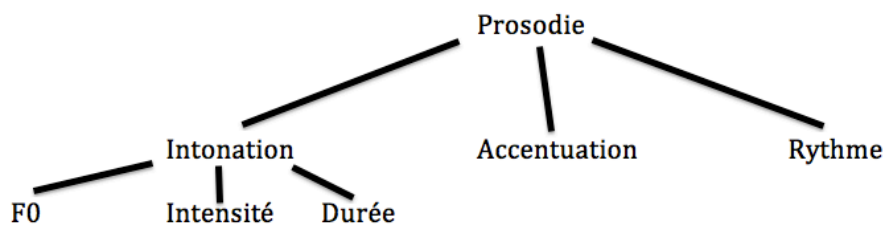


Figure 1. Cette figure illustre la conception de la prosodie (construite d'après la proposition de Lacheret-Dujour (2011)) divisée en trois parties principales.

Le locuteur dispose de l'intonation afin de signaler l'information importante du message et de procéder à la segmentation appropriée du flot de parole (Caspers, 1998 : 377). La problématique qui se pose en étudiant la fonction de l'intonation, et notamment les contours mélodiques, provient du fait que des éléments qui relèvent du contexte se manifestent au même endroit, dérangeant ainsi l'identification des traits qui sont propres au contour. La tâche de discernement entre le sens d'un contour mélodique pur et le sens signalé par le contexte, c'est-à-dire la distinction entre les phénomènes linguistiques et paralinguistiques, est alors compliquée (Caspers, 1998 : 379). Dans les langues post-lexicales, telle que le français, il est possible d'observer que le sens de l'énoncé « Marie corrige encore les copies » (Smith, 2002 : 173) dépend du patron intonatif avec lequel il est prononcé. S'il est présenté avec un patron intonatif descendant, on a tendance à l'interpréter comme une affirmation, alors qu'accompagné d'une montée fréquentielle il est plus probable que l'on le comprend comme une question de confirmation (cf. Grundstrom, 1973 : 19). Finalement il

faut noter que nous considérons deux montées intonatives qui signalent les deux une différence quant à l'implication interactionnelle (maintien ou abandon du tour de parole). Autrement dit, nous nous intéressons ici qu'à la différence intonative potentielle entre deux montées et les impacts que cela peut manifester lors d'une conversation, sans prêter aucune attention aux autres mouvements intonatifs qui peuvent aider à signaler l'abandon du tour (par exemple une descente de la f_0 – cf. Portes et Bertrand, 2007a). Dans ce qui suit nous essayons de décrire les deux contours (la continuation et la question) séparément, afin d'ensuite les comparer au sujet de leurs caractéristiques intonatives.

1.2. CONTINUATIONS

Le contour de continuation est souvent décrit à l'aide de sa fonction structurelle, c'est-à-dire l'implication du fait que le locuteur continuera à parler (Portes et Bertrand, 2007a : 1 ; Martin, 2007). Certains chercheurs établissent une distinction entre la continuation majeure et la continuation mineure, toutes deux montantes, (cf. Portes et Bertrand, 2007b ; Post, 2002) alors que cette dichotomie n'est pas basée sur des caractéristiques solides, ce qui complique l'identification. Portes et Bertrand (2007a : 8) avancent que la différence entre les deux est basée sur le fait que la continuation majeure coïncide avec des points de transition possible (un point où un changement de locuteur peut s'effectuer), ce qui ne nous empêche pas à conclure qu'elle ne provoque pas un changement de locuteur, vu que tous les points de transition possible ne résultent pas en une alternance de tour (cf. Selting, 2000). D'après les travaux de Portes et Bertrand (2007a) il s'avère alors qu'elle peut se manifester à plusieurs endroits dans le flux de parole, alors qu'il n'existe pas des résultats démontrant sa coïncidence avec un changement de tour de parole (d'après nos connaissances). Il est ainsi important de remarquer que la continuation majeure sera (d'après les travaux cités) toujours suivi par le même locuteur qui l'a produit. Pourtant, la différence entre les deux contours de continuation ne peut pas être établie à l'aide de la fréquence fondamentale (c'est-à-dire le maximum qu'ils atteignent) d'après une étude empirique (Portes et Bertrand, 2007b). Cependant, ils ont le même mouvement intonatif (montant) et impliquent un maintien de la parole, ce qui nous mène à nous servir du nom regroupant de « continuation », sans faire la différence entre mineure et majeure.

Les travaux de Portes, Bertrand et leurs collègues qui étudient la différence entre les deux contours de continuation se servent d'un codage intuitif (d'après une correspondance personnelle avec Roxane Bertrand) ce qui est circulaire dans la manière où une impression est

basée sur des informations intonatives présentes dans le signal. D'après une recherche qui emploie un tel codage, Portes et Bertrand (2007b) ont remarqué que la continuation majeure a une version qui est employée par les locuteurs lors de la production d'une énumération. Cette version a une différence phonétique audible par rapport à la continuation majeure dite normale, étant ainsi classée comme une sorte spécifique d'une continuation majeure. Ce contour peut terminer un paragraphe ou une séquence par l'initiation d'une liste « virtuelle » que l'interlocuteur a charge de compléter en imagination (Bertrand et al. 2007 : 69). Cette version de la continuation majeure ne sera pas incluse dans notre recherche à cause du nombre limité des exemples (seulement dix) qui se sont manifestés avant un silence (un nombre insuffisant pour une étude quantitative). Il faut néanmoins souligner que, comme une différence ne semble pas pouvoir être établie entre une continuation mineure et une continuation majeure (toutes les deux montantes) sur des bases objectifs (autorité épistémique, réaction par le locuteur), nous nous servons du terme « continuation » étant donné que la séparation de ces deux contours est l'objectif pour une autre étude – alors que la nôtre ne s'intéresse qu'à la comparaison entre deux montées fréquentielles qui ont des implications différentes lors d'une interaction.

1.3. QUESTIONS

Une inversion syntaxique et l'emploi de l'expression « est-ce que » au début de phrase sont toutes deux des stratégies disponibles lors de la demande de l'information, ainsi l'intonation n'est qu'un moyen parmi d'autres auquel un locuteur peut recourir lorsqu'il pose une question. Notre propre recherche, qui porte de l'intérêt aux montées intonatives, n'inclut que les questions qui sont accompagnées d'une fréquence fondamentale montante, excluant de cette manière les questions qui ne présentent pas cette caractéristique (p. ex. l'emploi d' « est-ce que », cf. Martin, 2004). La hauteur de la f_0 a été suggérée être un indice universel dont la plupart des langues se servent lors de la signalisation d'une question (Bolinger, 1989) qui posséderait d'une fréquence fondamentale en moyenne supérieure à celle des non-questions. Au sujet du français, Grundstrom (1973) a remarqué, dans son étude de corpus, que la majorité des questions ont des courbes montantes. Il prétend ainsi qu'il n'existe que trois courbes, toutes montantes et réalisées sous la dernière syllabe, qui signalent à elles seules l'interrogation (1973 : 32). Van Heuven et Van Zanten. (2005 : 96) concluent leur recherche empirique en adoptant les idées d'Ohala (1984) qui propose qu'une montée mélodique typique de la question est liée aux vestiges de notre comportement animal, car les créatures

petites et inoffensives ont une fréquence fondamentale supérieure aux créatures grandes et dangereuses.

La suggestion au sujet d'un maximum supérieur pour les questions est le plus souvent correcte étant donné que des questions à inversion et des questions écho sont accompagnées d'une montée finale, alors qu'une question partielle est (la plupart du temps) exprimée avec une descente de la f_0 (Smith, 2002 : 145). Les questions totales qui présentent une montée intonative en français ne se servent pas d'un mot interrogatif tel que « pourquoi », au contraire une questions avec une inversion combine intonation et inversion du pronom personnel et du verbe (« as-tu vu le chien ? » (Smith, 2002)). Une question écho (« Marie corrige encore des copies ? ») peut également être une phrase déclarative en remplaçant la montée par une descente de la f_0 , étant donné que l'ordre des mots est identique pour les deux types de phrases (Jun et Fougeron, 2000 : 222) (le mot type, ainsi que fonction est utilisé dans notre propre travail comme synonyme au mot contour dans nos discussions). Martin (2007 : 111) propose qu'« en [l'] absence d'une autre marque de la modalité interrogative, le contour mélodique final montant de l'exemple « tu viens » est pertinent et donc non neutralisé. La présence d'un autre marqueur (inversion syntaxique, mot interrogatif, indice pragmatique) pourra entraîner la neutralisation du contour qui pourra ou non être effectivement réalisé ». Nous allons remarquer par la suite que les deux contours étudiés dans ce travail présentent tous les deux une montée de la f_0 , ainsi qu'un même ordre des mots, ce qui sont les raisons pour la difficulté de la séparation. Le chapitre suivant traite des informations intonatives qui dans la littérature ont été proposées responsables afin d'établir une différence entre les deux contours.

1.4. DIFFÉRENCE ENTRE LA CONTINUATION ET LA QUESTION MONTANTE

En 1966, Delattre propose un inventaire de dix contours intonatifs stylisés du français, ceux-ci reflétant différentes fonctions dans le discours et se manifestent sous la syllabe finale. Trois entre eux sont de nature montante où deux projettent une suite par le même locuteur (la continuation mineure et la continuation majeure), tandis que le dernier est celui de question (totale) qui dans sa nature indique la demande d'une réponse par l'interlocuteur (Delattre, 1966 : 14). Dans l'inventaire de Delattre (1966) (voir Figure 2 ci-dessous), le registre du locuteur est divisé en quatre niveaux sur lesquels il situe les contours. Comme nous pouvons le voir dans la Figure 2, tous les contours montants commencent au même niveau (le deuxième) afin d'atteindre des niveaux différents lorsqu'ils se terminent (continuation

mineure = 3, continuation majeure = 4, question = 4+) sous la dernière voyelle (Delattre, 1967).

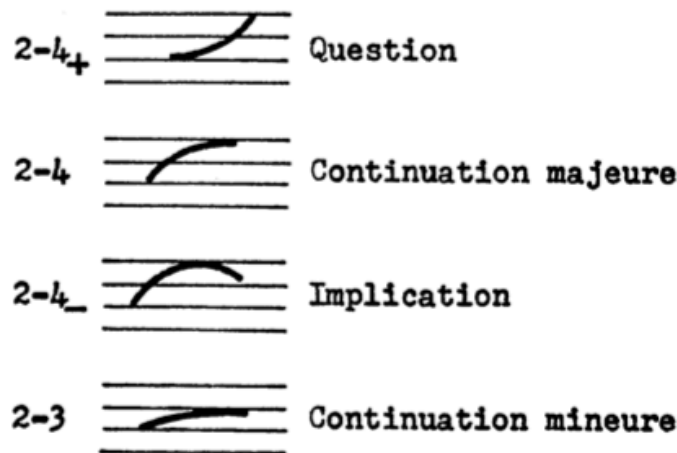


Figure 2. Cette figure, empruntée à Delattre (1966 : 4), démontre les différences en maximum, gamme (range en anglais) et forme entre les montées de question, de continuation majeure et de continuation mineure.

L'inventaire proposé par Delattre (1966) n'a pas fait l'unanimité ; d'autres chercheurs de l'intonation du français ont introduit d'autres traits distinctifs afin de séparer la continuation de la question montante. Rossi et al. (1981 : 158-159) ont effectué une étude de corpus expérimental qui confirme qu'une montée interrogative et une montée de continuation (non-interrogative) diffèrent quant aux niveaux maximum qu'elles atteignent à la fin de l'énoncé. À l'aide de leur étude, ils ont aussi pu observer que la continuation semble avoir un ton statique, alors que celui de la question est montant pendant le contour entier. En ce qui concerne les deux autres paramètres acoustiques de l'intonation, la continuation se manifeste par un glissement d'intensité entre 3 et 8 dB ainsi qu'une légère tendance à prendre une durée inférieure de la voyelle tonique (finale) par rapport à celle de la montée interrogative (Rossi et al. 1981 : 159). Une autre chose qu'il est possible de constater concernant la Figure 5 ci-dessus est que la forme de la courbe montante de la continuation se distingue de celle de la question. La courbe de continuation monte plus tôt (déjà vers un tiers de sa durée) afin de s'aplanir pendant le reste de sa durée, alors que la courbe de question manifeste plutôt sa montée forte vers la fin de la courbe (après le deuxième tiers). Rossi et al. (1981 : 153) émettent l'hypothèse que les deux montées ne sont en vérité que deux variantes contextuelles (fin/non-fin d'énoncé) d'un même contour, cependant ils concluent qu'elles peuvent se différencier par l'étendu du glissando (c'est-à-dire le nombre de niveaux traversés pendant la

montée (1981 : 162-163 et 176). Ces résultats confirment les conclusions de Di Cristo (1978) qui proposent que la réalisation phonétique différente entre la continuation et la question totale est due au contexte, et que le contour sous-jacent est en effet le même.

Contrairement à ce que Delattre (1966) a proposé quant à la forme, Rossi et al. (1981 : 153), Di Cristo (1976) et Rossi (1978) ont tous démontré que la forme du contour ne sert pas à établir la distinction entre les deux montées étudiées. Rossi (1978) suggère que la perception de la forme qui ne se manifeste pas dans le signal est due au changement d'intensité qui transforme la perception de la courbe. Cela a également été manifesté par la recherche empirique de Grundstrom (1973 : 29). Néanmoins, son test de perception a démontré qu'il existe une hiérarchie entre les contours montants où leur forme et leur degré d'interrogation sont liés. Selon ses résultats la montée en croissant (concave) est toujours le contour le plus interrogatif, suivi par une montée constante et finalement d'une montée en décroissant (convexe) qui n'est pas interrogative dans sa nature (1973 : 45). Il admet cependant que « la forme de la montée y est pour quelque chose » (1973 : 45), alors que le maximum de la f_0 joue un rôle plus grand dans la distinction entre les deux montées.

L'idée centrale est que les deux fonctions (c'est-à-dire continuation et question) résultent en différentes réactions interactionnelles par l'interlocuteur (cf. Stivers et Rossano, 2010) où une question demande une réponse et ainsi une réaction, alors que la continuation implique structurellement que le locuteur en cours maintient la parole. Basée sur cette idée, il est également possible à s'attendre que les deux fonctions coïncident avec des signaux de garde ou de maintien de tour de parole. D'après les études sur les turn-yielding cues (signaux d'abandon du tour de parole) en anglais (Duncan, 1972 ; Gravano et Hirschberg, 2011), il a été possible à constater que les énoncés finaux de tour présentent une intensité inférieure relative aux énoncés au milieu de tour. L'étude la plus récente admet également que la durée du mot final d'énoncé est inférieure à celui placé à l'intérieur du tour, ce qui contredit les idées sur l'« allongement final » en fin d'énoncé (cf.; Duncan, 1972 ; Cutler et Pearson, 1986 ; Stephens et Beattie, 1986). Une étude comparative entre trois langues (le malais de Manado, le dialecte d'anglais parlé aux Orcades et le néerlandais) effectuée par Van Heuven et Van Zanten (2005) confirme que les questions ont également un débit de parole supérieur en comparaison avec une phrase déclarative avec la même structure linguistique. Ce chapitre comparatif nous permet ainsi de noter plusieurs caractéristiques liées à l'intonation qui sont avancées comme établissant la distinction entre les deux contours ou les actes

verbales (speech act en anglais) avec lesquels ils coïncident.

2. HYPOTHÈSE ET PRÉVISIONS

Selon la littérature sur l'intonation du français, la continuation et la question sont différentes quant aux traits phonétiques qu'elles manifestent, ainsi qu'à leurs placements dans les séquences conversationnelles (Delattre, 1966 ; Delattre, 1967 ; Portes et Bertrand, 2007a). Certains phénomènes inclus dans nos prévisions ne découlent pas seulement de la recherche focalisée sur l'intonation du français, mais aussi d'études consacrées aux signaux de maintien et d'abandon du tour de parole. L'identification de la fonction des exemples a été déterminée à l'aide du paramètre de changement de tour de parole puisqu'une continuation ne résulte pas en une alternance (cf. Bertrand et al. 2007), contrairement à une question (cf. Stivers et Rossano, 2010). Ces prédictions servent de base pour les tests statistiques effectués et présentés dans le chapitre 5, justifiés presque entièrement par des affirmations dans la littérature. La distinction est faite entre variables finales (sous la syllabe finale) et non-finales (le syntagme intonatif final entier), afin de pouvoir examiner si les exemples des deux montées peuvent être catégorisés avant que la syllabe finale (la montée) arrive. Notre hypothèse au sujet des deux contours est :

« Les continuations sont différentes en comparaison aux questions totales en ce qui concerne leurs compositions intonatives, aussi bien en ce qui concerne les mesures finales que non-finales ».

Les mesures finales comprennent le minimum et le maximum standardisés de la f_0 , la forme du contour, le maximum d'intensité standardisé et la durée de la voyelle finale. Toutes ces mesures ont été prises sous la syllabe finale et parfois même sous la voyelle finale. En revanche, les mesures non-finales ont été prises quelque part dans le syntagme intonatif précédant à la syllabe finale dans le but de vérifier si un locuteur signale, à l'aide des paramètres intonatifs, le type de montée qu'il est en train de produire. Le débit de parole et le registre de l'énoncé, calculé des mesures de la f_0 jusqu'à la syllabe finale, dans le groupe des mesures non-finales, étant donné que ces mesures comprennent des informations qui émergent avant la syllabe finale.

Au sujet des mesures finales, nous prévisions que les questions ont un maximum supérieur à celui des continuations selon les observations de Delattre (1966 ; 1967). Il a également évoqué une différence concernant la forme du contour, de ce fait nous nous attendons à ce que les questions manifestent une courbe concave sous la voyelle finale

(Delattre, 1966) alors que les continuations sont accompagnées d'une courbe convexe. L'intensité maximale standardisée de la voyelle finale présente, selon nos prévisions, une valeur inférieure pour les questions (Rossi et al. 1981 ; Cutler et Pearson, 1986 ; Gravano et Hirschberg, 2011), puisqu'elle est un signal d'abandon de tour. En revanche, les minima standardisés de la f_0 sont égaux, étant donné que les deux contours commencent au même niveau dans le registre du locuteur. Quant à la durée de la voyelle finale, nous nous attendons à ce qu'une question et une continuation n'ont de durées différentes, d'après l'observation de Gravano et Hirschberg (2011 : 614) qui propose un allongement final plus important avant les bornes de phrase intra-tour qu'inter-tours (*turn-medial* et *turn-final*).

Les mesures non-finales sont peu nombreuses en comparaison avec les mesures finales, une conséquence naturelle au fait que la syllabe finale reçoit souvent l'attention lors de ce genre d'études. Le débit de parole est censé être plus rapide pour un énoncé final, ce qui dans cette étude est représentée par la question, que celui d'un énoncé non-final (Gravano et Hirschberg, 2011 ; Van Heuven et Van Zanten, 2005). La dernière mesure choisie dans cette recherche, qui n'a pas été discutée auparavant dans les recherches concernant les deux montées, est le registre de l'énoncé entier jusqu'à la syllabe finale. Cependant, l'inventaire de Delattre (1966) indique que la montée finale des deux contours commence au même niveau, ce qui nous mène à nous attendre à ce que cette mesure ne distingue pas les deux fonctions entre elles. Le terme registre fait désormais dans notre propre travail référence à un certain niveau (cf. Delattre, 1966) dont les locuteurs se servent à un certain moment. Le registre se place à l'intérieur du champ tonal qui est l'étendue entre le plancher et le plafond du locuteur.

3. MÉTHODE

Ce chapitre introduit notre propre méthodologie (« semi-manual-effort methodology ») inspirée par celle de Heldner et Edlund (2010). Elle essaie de rassembler deux approches distinctes qui font de la recherche sur le même objet – la langue parlée. D’une part, la méthodologie d’analyse conversationnelle représente les études qualitatives qui effectuent la recherche sur une collection restreinte d’exemples tirés des conversations naturelles, alors que d’autre part l’approche phonétique se sert plutôt de la quantification. Les deux approches méritent d’être louées quant à leurs qualités différentes ; l’analyse conversationnelle prête beaucoup d’attention aux exemples en examinant ce qui les précède dans la conversation, alors que la quantification présente la possibilité d’examiner tout un corpus et la manière dont un phénomène spécifique en est représenté. L’approche statistique permet également d’effectuer des tests statistiques où la taille d’échantillon est suffisamment grande pour tirer des conclusions générales. En nous inspirant des travaux d’Ogden (2006) et de Heldner et Edlund (2010), un compromis entre les deux approches nous semble être une manière nécessaire afin d’examiner les différences et les similitudes potentielles des deux montées. Nous essayons le plus possible d’inclure des informations sur le contexte interactionnel (l’autorité épistémique, la réaction par le locuteur, etc.), afin d’en même temps nous servir d’une approche qui nous permet d’inclure plus d’exemples que dans une étude de collection. Ce chapitre contient également des présentations des logiciels employés dans ce travail, ainsi qu’une justification du codage créé et des mesures intonatives prises dans chaque exemple d’une manière automatique.

3.1. « SEMI-MANUAL-EFFORT METHODOLOGY »

À l’aide d’une méthodologie complètement objective sans l’influence d’un chercheur, Heldner et Edlund (2010) ont employé un logiciel qui leur a permis de couper le signal audio en morceaux cohérents ne pas séparés par un silence. Comme cette méthodologie ne prend en compte que les traits intonatifs sans faire référence à la syntaxe, au contexte communicatif entre les participants, etc., il est ainsi impossible de dire qu’elle intègre toutes les informations qui peuvent avoir une importance dans la communication. Nous combinons leur méthodologie avec une analyse semi-qualitative (seules les informations linguistiques dans l’énoncé sont incluses) avec l’espoir de l’améliorer afin qu’elle soit mieux ajustée à notre propre approche qui se veut plus proche à celle de l’analyse conversationnelle qu’était celle de Heldner et Edlund (2010).

En pratique, notre méthodologie implique que les exemples identifiés sont codés par rapport à la complétude syntaxique, l'autorité épistémique et la réaction interactionnelle de l'interlocuteur (un régulateur, une réponse, etc.). Ces exemples sont ensuite annotés dans Praat, où des bornes autour de différents endroits dans l'énoncé (par exemples la voyelle finale) ont été mises manuellement dans le TextGrid (voir *infra*). Ces indications manuelles sont faites afin de faciliter les mesures intonatives (voir Figure 3), ainsi qu'assurer la systématisme de leurs prises au même endroit dans tous les exemples. Outre la borne autour de la voyelle finale, un point qui indique le silence suivant la montée a été marqué. Après que ces indices ont été mis dans le TextGrid, les mesures intonatives ont été prises automatiquement sans qu'il n'y ait d'influence par le chercheur sur les valeurs obtenues sauf leurs points de repère.

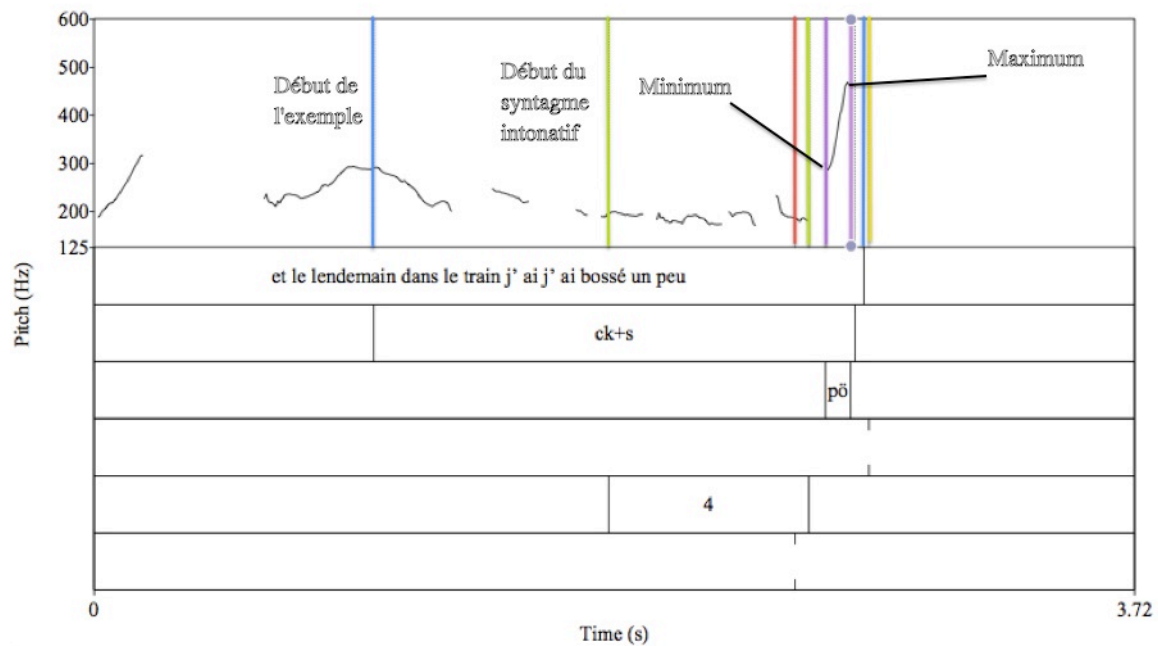


Figure 3. Cette figure présente un exemple d'un token avec des bornes et des points indiqués à l'aide de la couleur. L'exemple en entier est indiqué par les bornes bleues, alors que les bornes vertes illustrent le syntagme intonatif dont la syllabe finale fait partie (la syllabe finale exclue). La ligne jaune, qui indique où le silence qui suit la montée se situe, est le dernier point de repère pour le logiciel. Ces indices permettent au script de savoir entre quels points une mesure spécifique doit être prise.

Il nous semble difficile d'étudier les différences des montées intonatives du français ainsi que leurs rôles potentiels dans l'interaction sans faire référence au contexte conversationnel (cf. la méthodologie employée par Heldner et Edlund, 2010). Cependant, il nous semble crucial de

prendre les mesures intonatives aux mêmes endroits dans chaque exemple afin d'obtenir une sorte de systématique pour l'analyse. Même s'il est évidemment problématique de tenter de combiner ces deux méthodologies, néanmoins nous avons fait de notre mieux pour considérer des revers possibles en intégrant les deux approches dans une méthodologie en ce qui concerne l'objectivité du classement des contours et l'automatisme. D'abord le fait de décider en avance où il faut prendre la mesure d'un maximum d'intensité ne convient pas aux analystes conversationnels, qui préfèrent étudier l'énoncé entier pour expliquer certains phénomènes conversationnels (cf. Curl, 2004 ; Stivers et Rossano, 2010). Ensuite, la différence inévitable entre les exemples en ce qui concerne leur longueur, les voyelles différentes qui se trouvent à la position finale, le débit de parole, le sexe du locuteur, etc. sont tous des paramètres qui introduisent une variété d'exemples. Nous avons fait de notre mieux afin de compenser pour ces revers en essayant de standardiser les mesures le plus que possible. Les valeurs obtenues après cette standardisation peuvent être généralisées à travers des exemples distincts.

Nous concevons bien qu'une méthodologie adoptée du domaine de l'analyse conversationnelle propose l'avantage de prêter une attention spéciale à chaque exemple, sachant qu'aucune conversation ne ressemble à une autre. Il nous semble cependant important d'examiner une quantité d'exemples afin de pouvoir en tirer des conclusions généralisées sur les différences intonatives des deux contours et sur leur fonction discursive. C'est la raison pour laquelle nous avons introduit cette méthodologie qui (d'après nos connaissances) n'a pas été utilisée dans une étude similaire sur l'intonation du français dans une conversation spontanée. Nous sommes conscients du fait que cette méthodologie peut en effet être améliorée, ce qui est notre but pour l'avenir, mais elle nous semble convenir pour ce travail.

3.1.1. IDENTIFICATION AUDITIVE

La première sélection des exemples a été faite à l'aide du logiciel ELAN (Brugman et Russel, 2004) développé à Max Planck Institute for Psycholinguistics à Nimègue. En premier lieu, ce logiciel sert à annoter la multimodalité dans les conversations et un de ses buts est de faciliter la recherche des analyses sur les gestes dans la communication. Ce logiciel présente également la possibilité de simultanément consulter la vidéo, écouter le son de bon qualité (pas celui enregistré par la caméra) aussi bien qu'annoter à quels instants un phénomène linguistique se manifeste (dans notre cas des montées intonatives avant un silence) ; une qualité admirable pour un logiciel d'annotation. Deux annotateurs (les deux compétents en

français ; l'auteur de ce mémoire et le chercheur Francisco Torreira) ont séparément identifié les exemples en écoutant les dialogues dans leur totalité pour ensuite se mettre d'accord sur quels exemples inclure dans l'analyse finale. Grâce à cette manière de travailler, il nous semble qu'aucun exemple n'a échappé l'identification. En cas de désaccord sur si un contour était montant ou non (un contour d'implication réalisé avec une f0 très haute a des fois été confondu avec des montées intonatives étudiées dans ce travail), le logiciel Praat a été consultée comme estimation objective.

Au début du travail nous avons l'intention d'inclure davantage de types de montées intonatives, telles que des questions tags (avec l'ajout d'un « hein » ou d'un « non ») et des régulateurs montants. Nous nous sommes ensuite rendus compte qu'il y avait trop d'exemples de ces deux types pour une étude de cette étendue. C'est pourquoi nous nous sommes concentrés exclusivement sur les montées de continuation et de question, limitant ainsi le nombre d'exemples à 323. Ces exemples ont ensuite été codés pour leur complétude syntaxique, leur autorité épistémique et la réaction interactionnelle manifestée par l'interlocuteur. ELAN a eu un avantage sur Praat puisqu'il permet de consulter l'enregistrement vidéo, étant donné que certaines réactions par l'interlocuteur ont été du type gestuel.

3.1.2. CODAGE

En vue d'être objectif lors de la classification des questions montantes et des continuations, nous avons créé un système de codage qui permet une objectivité supérieure à l'intuition. La division en catégories est ainsi un peu plus robuste. En plus de ce codage objectif, nous avons également classifié les catégories 'c' (continuation) et 'q' (question) afin d'avoir une idée générale du nombre d'exemples de chaque catégorie dont l'on possédait. Dans les chapitres qui suivent nous allons présenter chaque type de codage ainsi que la manière dont il reflète les affirmations faites par la littérature. La correspondance entre les codages 'intuitifs' et les codages 'objectifs' est également calculée et commentée. Le codage 'j' (joke) ne comprend pas beaucoup d'exemples (3), mais nous l'avons inclus tout de même afin de pouvoir faire référence au fait qu'un exemple particulier peut se comporter différemment des autres puisque le locuteur fait une sorte de plaisanterie.

| | | | | |
|-----------------------------|----------------------|---------------------------|----------------|--------------------|
| Codage | | | | |
| Syntaxe | s (syntaxe complète) | i (syntaxe incomplète) | | |
| Autorité épistémique | + (locuteur) | - (interlocuteur) | ~ (partagée) | |
| Réaction | f (réponse pleine) | k (maintien de la parole) | b (régulateur) | n (aucune réponse) |
| Plaisanterie | j (plaisanterie) | | | |

Tableau 1. Notre propre système de codage développé pour ce travail.

3.1.2.1 CODAGE INTUITIF

Ce codage n'est pas employé officiellement dans nos analyses, puisqu'il est basé sur des impressions intuitives que nous avons eues de ce qu'est une question ou une continuation. Il est possible de classer certaines questions (inversion) à partir de la syntaxe, alors que d'autres ne sont classées qu'avec l'impression que l'on obtient en ayant écouté la conversation dans sa totalité. Le motif pour ce codage est la possibilité d'une comparaison avec notre propre codage (voir Ch. 3.1.2.) et ainsi observer leur correspondance (voir Ch. 3.1.2.5.). Ce codage peut être comparé à celui de Bertrand et al. (2007) qui ont utilisé un codage intuitif dans leur article. Après une communication personnelle avec Roxane Bertrand, nous avons confirmé que le codage dans leur étude (Bertrand et al. 2007) a été basé sur l'intuition d'un chercheur en phonétique qui reconnaît bien les patrons intonatifs du français.

3.1.2.2. CODAGE SYNTAXIQUE

Cette partie du codage est centrée sur la complétude syntaxique et la projection d'une suite potentielle qu'elle peut signaler. Comme nous l'avons déjà mentionné, les locuteurs se servent de la syntaxe quand ils cherchent à effectuer un changement de parole (cf. Sacks et al., 1974 ; Selting, 2000). C'est la raison pour laquelle nous avons décidé d'introduire la complétude syntaxique dans la classification des deux montées. Il est bien possible que la syntaxe puisse être incomplète pour les continuations qui projettent une suite, puisqu'elles signalent à l'aide d'une collaboration entre la syntaxe et l'intonation que le locuteur en cours n'a pas fini son tour de parole (cf. Selting, 2000). Les termes *syntactiquement 'complet'* et *incomplet'* ont été jugés sur des bases de finitude possible ou non. Lorsqu'un énoncé est syntactiquement complet, cela indique que le locuteur peut parfaitement bien s'arrêter là sans qu'il ne doive nécessairement continuer (la pragmatique n'a pas été prise en compte). En revanche, lors d'un énoncé syntactiquement incomplet, le locuteur suspend ce qu'il est en train de dire au milieu d'un syntagme qui ne peut pas être considéré complet (manque d'un élément nécessaire). Un exemple d'un énoncé syntactiquement incomplet tiré du corpus est « *c'est vrai que si on s'y consacre pas vraiment* », qui projette clairement une suite est ensuite suivie de « *je trouve*

que c'est un truc hyper dur quoi ». Aussi bien la continuation prototypique que la question prototypique peuvent contenir le codage 's (complet). Alors que seules les continuations peuvent également être incomplètes (i), étant donné qu'elles sont susceptibles de renforcer l'information de ne pas avoir terminé le tour de parole avec une syntaxe incomplète.

3.1.2.3. CODAGE ÉPISTÉMIQUE

Le codage épistémique, basé sur les idées avancées par Heritage (2012), illustre l'autorité épistémique par un des interlocuteurs du thème conversationnel traité à un moment donné. Il existe trois possibilités de classification : '+', '-' ou '~'. Lorsque le locuteur parle de quelque chose qui lui appartient, le signe '+' est employé, par exemple en utilisant le pronom personnel 'je', ou en racontant une histoire que le locuteur a vécu. En revanche, quand l'interlocuteur possède le plus d'autorité épistémique concernant un sujet, nous avons annoté cela à l'aide du signe '-'. Cette catégorie peut par exemple être illustrées avec l'emploi du pronom personnel 'tu'. Un problème a émergé quand personne n'avait l'autorité épistémique (les locuteurs ont discuté d'un match de football que les deux avaient regardé). Dans ces cas peu nombreux nous nous sommes servis d'un tilde '~', ce qui implique qu'aucun d'entre eux ne manifesterait une autorité épistémique.

Une continuation prototypique est signalée par le fait que le locuteur en cours sait le plus concernant ce que les interlocuteurs discutent (il possède alors l'autorité épistémique). Elle peut également être employée lorsque le locuteur parle de quelque chose de partagé. Quant aux questions prototypiques, elles demandent de l'information à l'interlocuteur (sauf dans les cas des questions rhétoriques, mais nous n'en avons pas trouvé dans nos exemples) et il est alors évident que l'interlocuteur connaît le mieux le sujet. Les questions sont souvent marquées ou par l'usage du pronom personnel 'tu', ou par le fait que le thème traité appartient clairement à l'interlocuteur. La force du pronom personnel 'tu' comme un signal (indépendamment de l'intonation et de la forme grammaticale) de question a été proposé par Geluykens (1987, cité par Bolinger, 1989 : 101) d'après un test de perception où les participants ont identifié les phrases déclaratives avec un 'tu' comme des questions plus de la moitié du temps.

3.1.2.4. CODAGE DE RÉACTION

La réaction interactionnelle de l'interlocuteur est censée être différente selon le type de contour émis par le locuteur, c'est-à-dire qu'une question provoque une réponse par l'interlocuteur (cf. Stivers et Rossano, 2010). En revanche, une continuation projette une suite

par le locuteur en cours. En suivant ce raisonnement, nous avons identifié quatre réactions possibles aux deux montées intonatives. Il est possible que le locuteur en cours maintienne la parole et continue à parler, ce qui est marqué avec un 'k'. L'interlocuteur peut également réagir et produire une réponse pleine 'f' à ce que le locuteur vient de dire, ou répondre avec un régulateur ('b'). Tous ces codages impliquent que la réaction est cohérente avec ce qui vient d'être dit, c'est-à-dire qu'une question reçoit une réponse qui lui est liée. Toutes ces réactions se manifestent après un silence qui ne dépasse pas 2 secondes, lorsque cela n'a pas été le cas nous sommes servis du signe 'n'. Les continuations prototypiques sont suivies ou d'un maintien de la parole ou d'un régulateur, alors que les questions prototypiques peuvent apparaître avant une réponse pleine ou un régulateur (un régulateur peut aussi être composé de gestes (Portes et Bertrand, 2007a : 6)). Quant au codage 'n', il peut être observé après les deux types de contours dû aux chevauchements qui peuvent diriger la conversation dans une autre direction lié à autre chose que le thème de l'énoncé étudié.

3.1.2.5. L'ACCORD ENTRE CODAGE INTUITIF ET CODAGE OBJECTIF

Nous avons créé de nouvelles variables pour les analyses statistiques que nous avons nommé 'continuation prototypique' et 'question prototypique'. Ces variables comprennent les valeurs qui correspondent à un type prototypique afin d'obtenir une classification sans l'influence subjective sur ce qui est une continuation ou une question. Afin d'être le plus objectif possible, nous avons par conséquent combiné les valeurs suivantes pour une continuation : 's'/'i', '+/'~', 'k'/'b'/'n'. Une continuation a ainsi la possibilité d'être syntaxiquement complète ou incomplète, qu'elle doit traiter un thème qui est maîtrisé par le locuteur, ou partagé par les deux locuteurs. Les réactions, qui peuvent être provoquées par un contour de continuation sont une suite par le locuteur ou un régulateur émis par l'interlocuteur, ou en cas de chevauchement une réaction sans rapport avec ce qui précède.

Quant aux questions prototypiques, les traits suivants sont inclus : 's', '–', 'f'/'b'/'n'. Une question doit être syntaxiquement complète puisqu'un interlocuteur ne peut pas répondre si l'énoncé interrogatif n'est pas terminé. Il est clair que l'autorité épistémique est possédée par l'interlocuteur, car il n'est pas nécessaire de poser une question si on connaît déjà la réponse (sauf dans des cas d'une question rhétorique). Une question produit normalement une réponse pleine de l'interlocuteur, alors qu'un régulateur ou un manque de réponse dans le cas d'un chevauchement ou une plaisanterie (par exemple « T'es allumée ? »

adressé à la caméra) sont également possibles. Le tableau 2 ci-dessous montre l'accord entre les codages intuitifs et les codages objectifs qui est supérieur à 95 %.

| | Type intuitif (continuation) | Type intuitif (question) |
|---|-------------------------------------|---------------------------------|
| Type prototypique (continuation) | 189 | 6 |
| Type prototypique (question) | 6 | 122 |

Tableau 2. Ce tableau montre l'accord entre les codages objectifs et les codages intuitifs.

3.1.3. MESURES INTONATIVES AUTOMATIQUES

Le slogan du logiciel Praat, développé à l'Université d'Amsterdam par Boersma et Weenink (2013), « doing phonetics by computer » reflète bien l'idée fondamentale de ce travail, c'est-à-dire le désir d'inclure un ordinateur qui complète l'ouïe par des visualisations et des reproductions des données sans l'influence d'autres traits dans l'énoncé. Ce logiciel a eu le rôle d'instrument de mesure, qui d'après nos indications a pris des mesures systématiques aux mêmes endroits dans tous les exemples, à l'aide des scripts écrits par l'auteur et le chercheur Francisco Torreira. Ce chapitre présente les mesures choisies, ainsi que le placement des bornes de la voyelle finale, de la fin de l'énoncé, du début du syntagme intonatif, etc. (voir Figure 3).

3.1.3.1. SCRIPTS AVANT LES MESURES

La règle fondamentale dans la programmation est que l'ordinateur fait exactement ce que l'on lui a demandé de faire, c'est pourquoi il faut bien indiquer dans quel endroit du fichier qu'il doit effectuer la commande. Les indications du début et de la fin des voyelles finales sont faites (puisque c'est sous la voyelle finale que le contour se manifeste, cf. Mertens, 2009) en consultant des informations visuelles dans l'oscillogramme et des formants dans le spectrogramme. Des bornes ont également été mises autour du syntagme intonatif final, en excluant la dernière syllabe, ce qui nous permet de mesurer le débit de parole ainsi que la médiane de l'énoncé jusqu'à la syllabe finale (la montée de la f0). Finalement, un point dans le silence qui suit le contour montant permet de mesurer la trace de la f0 après la voyelle finale, jusqu'à ce que l'énoncé soit terminé.

3.1.3.2. MESURES DE F0

Plusieurs mesures de la fréquence fondamentale ont été prises vers la fin de l'énoncé, aussi bien sous la voyelle finale qu'avant et après celle-ci. La plupart entre elles sont tirées des propositions faites dans la littérature au sujet des différences en forme et en maximum des deux montées. Dans un morceau de 10 minutes au début de la conversation, Praat a calculé la médiane (un point de référence) de chaque locuteur, qui nous permet ensuite de standardiser chaque mesure de la f0. En soustrayant la médiane globale de la valeur exacte, nous obtenons des valeurs standardisées (toutes en demi-tons). Cela permet la comparaison des exemples prononcés par des locuteurs et ceux par des locutrices, sans faire de calculs supplémentaires (les voix des hommes et des femmes se diffèrent en hauteur, ce qui complique la comparaison absolue entre elles).

Il nous semble naturel de commencer la présentation par la mesure prise au début du contour montant – le minimum de la montée. Cette mesure représente le point le plus bas de la montée sous la voyelle finale. Les exemples avec une attaque se composant d'une consonne sourde posent des problèmes lorsque la montée de la f0 est censée commencer pendant cette consonne. Étant donné que la f0 n'est pas visible (c'est-à-dire pas mesurable) dans le spectrogramme, la décision a été prise de n'inclure que le minimum standardisé pris sous la voyelle (qui est voisée). Delattre (1966) indique que les deux types de contours commencent au même niveau dans le registre (c'est-à-dire champ tonal dans nos termes) des locuteurs, ce qui implique que les deux montées n'ont pas de minima se distinguant au niveau de leur hauteur. La mesure qui s'oppose à la précédente est le maximum de la f0 à la fin de la montée. Après une examen des visualisations des montées, nous avons observé que certaines codas suivant la voyelle finale peut causer une descente légère de la f0 vers la fin de la voyelle, alors que d'autres codas manifestent une montée continue. Cela nous a amené à mesurer le point maximum pendant la syllabe finale avant le silence (remarquons que nous avons seulement choisi des montées qui se trouvent avant un silence afin de faciliter la prise des mesures automatiques). En considérant les propositions dans la partie théorique, il est possible que ces mesures distinguent les deux montées (Delattre, 1966 ; entre autres).

3.1.3.3. MESURES D'INTENSITÉ

En menant une étude qui s'inscrit dans la tradition exploratrice, nous avons l'intention d'examiner si les deux types de montées se distinguent en ce qui concerne l'intensité qu'ils présentent. Étant donné que chaque voyelle a une intensité inhérente liée au type de voyelle

([a] a l'intensité la plus haute, alors que [i] a l'intensité la plus basse), il n'est pas possible de simplement comparer les mesures d'intensité totales entre elles. La standardisation est notre solution à ce problème, en supposant qu'il existe au moins une voyelle (le noyau de la syllabe pénultième) dans la seconde précédant la voyelle finale (avec un débit de parole à une syllabe par seconde, il existe plusieurs voyelles dans cette seconde). Nous avons mesuré le maximum d'intensité dans cet extrait. Cette valeur est soustraite du maximum de la voyelle finale. Ainsi, nous possédons une mesure de l'intensité (mesurée en dB), qui nous indique si la voyelle finale est plus faible ou plus forte en comparaison avec ce qui la précède.

3.1.3.4. MESURES DE DURÉE

Comparer la durée de plusieurs exemples pose des problèmes liés à une comparaison sans standardisation. Afin d'éviter cette problématique, nous incluons le débit de parole du syntagme intonatif final comme variable exogène dans le test statistique. En nous servant du TextGrid, l'indication du nombre de syllabes qui se trouvent dans le même syntagme intonatif que la syllabe finale et de leur durée a été possible. En divisant le nombre de syllabes par leur durée, le résultat désigne le débit de parole (syllabes/seconde) pendant le dernier syntagme intonatif de l'énoncé. Ces deux mesures de durée (la durée de la voyelle finale et le débit de parole) permettent la comparaison des différences intonatives potentielles entre la continuation et la question. Comme une voyelle dans une syllabe fermée est différente dans sa durée qu'une voyelle dans une syllabe ouverte (Léon, 1996), nous avons contrôlé cette variable exogène à l'aide de l'inclusion de l'information sur la présence ou l'absence d'une coda syllabique dans le test statistique. Néanmoins, il nous semble que notre propre tentative est une manière acceptable d'inclure la durée comme variable endogène dans une étude quantitative semi-automatique.

3.1.3.5. EXCLUSION D'EXEMPLES PROBLÉMATIQUES

Le choix d'une étude exploratrice, en essayant d'inclure le plus de mesures possibles qui pourraient avoir une influence sur la distinction entre les deux montées, nous amène à l'exclusion de certains exemples qui présentent une difficulté quant à leur analyse automatique. Lorsque les participants ont rigolé lors de la production d'une montée intonative, nous avons choisi de ne pas l'inclure à cause de l'impossibilité d'une analyse dans Praat. Des exemples avec des voyelles contiguës n'ont pas été analysés non plus en raison d'une difficulté de l'indication du début de la voyelle finale (p. ex. le mot cible *théâtre* [teatr] ou la difficulté de séparation s'est présentée).

3.1.4. LE RAISONNEMENT DES TESTS STATISTIQUES

Les analyses statistiques dans cette étude ont été effectuées avec le logiciel statistique R (R Core Team, 2013) en examinant l'influence d'une ou plusieurs variables exogènes sur une variable endogène à l'aide des régressions linéaires. Ce chapitre sert à présenter l'essentiel de ces tests statistiques, ainsi que le polynôme employé dans le calcul du rayon de courbure des contours montants. Les variables endogènes qui, d'après les résultats statistiques significatifs sont liées au type prototypique, sont ultérieurement incluses dans une simulation d'une validation croisée qui rend possible l'examen des variables qui prédisent le mieux le type de contour.

3.1.4.1. TESTS STATISTIQUES

Cette partie du chapitre sur la méthodologie appliquée dans cette recherche présente les régressions linéaires dont nous nous sommes servies, ainsi que le motif derrière ce choix. Il faut se rappeler que la statistique sert à collecter des données dans le monde réel qui sont ensuite introduites dans des modèles statistiques afin de pouvoir faire des inférences sur comment ce phénomène se comporte dans la population entière (Field et al. 2012 : 33-34). Cette stratégie est incorporée dans la recherche dans la mesure où nous essayons de conclure s'il semble exister un patron intonatif dans la manière dont les locuteurs signalent la continuation ou la question dans la conversation. La logique derrière les régressions linéaires est de prédire une variable endogène d'après des variables exogènes avec une droite (Field et al. 2012 : 246). Cette ligne nous permet de prédire une valeur endogène d'après une ou plusieurs variables exogènes (régression simple ou multiple) en superposant la ligne qui représente le mieux la tendance générale des données (Field et al. 2012 : 246-247). Il est important de n'inclure que les variables qui semblent avoir une relation avec la variable endogène, puisque les résultats obtenus sont liés aux variables dans le modèle, c'est-à-dire qu'un modèle avec trop de variables exogènes sans une relation théorique avec la variable endogène ne résulte pas en un modèle fiable (Field et al. 2012 : 263-264).

La différence entre le modèle linéaire simple et le modèle linéaire multiple provient du nombre de prédicteurs qui sont inclus dans la régression (un ou plusieurs), alors qu'ils produisent tous les deux des résultats avec des coefficients à interpréter. Ces coefficients nous offrent de l'information sur l'ordonnée à l'origine sur l'axe y (c'est-à-dire la valeur où la ligne croise l'axe), ainsi que la pente de la ligne (Field et al. 2012 : 247). Dans une régression multiple avec une entrée forcée (toutes les variables introduites au modèle

simultanément) une décision sur quelle variable a le plus d'importance n'existe pas (Field et al. 2012 : 264). Les coefficients présentés par le modèle pour chaque variable exogène (au sujet de sa relation avec la variable endogène) indiquent l'influence qu'une variable a sur la pente si toutes les autres variables sont tenues constantes. De cette manière il est possible d'observer si, toutes autres variables tenues constantes, une certaine variable peut expliquer quelque chose dans la variabilité des données.

3.1.4.3. LE RAYON DE COURBURE DES CONTOURS

Afin de pouvoir comparer la forme d'un contour à un autre, en faisant référence à autre chose qu'une droite, une parabole a été employée en calculant le rayon de courbure pour chaque contour. Elle permet ensuite une comparaison entre la forme et la fonction. Il est ainsi possible d'examiner si, comme l'a proposé Delattre (1966), une montée de continuation se distingue au sujet de sa forme d'une question montante. Andruski et Costello (2004) ont démontré que des coefficients quadratiques représentent mieux les contours de la f_0 , par rapport aux coefficients linéaires, en améliorant la classification des contours dans leur étude du hmong vert (une langue à tons parlée en Asie du Sud-Est). Ils ont pris une mesure de la f_0 dans la courbe toutes les 10 ms (2004 : 128) afin d'introduire ces mesures dans un polynôme du second degré (c'est-à-dire que le troisième coefficient est au carré). Cette méthodologie a ensuite été adoptée par Grabe et al. (2007) et Torreira (2007), qui ont tous les deux employé ces équations quadratiques dans leurs illustrations de la concavité et la convexité des courbes de la fréquence fondamentale.

L'équation qui nous aide à obtenir le coefficient qui nous intéresse peut être exprimée de la manière suivante : $y = a + bx + cx^2$. a fait référence à la fréquence fondamentale de la courbe, b représente la pente initiale et c représente le rayon de courbure de la montée. En examinant les coefficients obtenus, l'intérêt est centré sur le coefficient c , qui lorsqu'elle est négative ou basse indique une montée précoce. En revanche, une montée tardive est représentée par une valeur positive ou haute (Torreira, 2007 : 1074). Ces coefficients sont ensuite étudiés concernant leur relation avec la fonction.

4. MATÉRIEL

4.1. THE NIJMEGEN CORPUS OF CASUAL FRENCH

Les recherches concernant la relation entre l'intonation et l'alternance des tours de parole ont deux choix au sujet du matériel ; aussi bien le lieu d'enregistrement (soit intra-, soit extra-muros) ainsi que le degré de spontanéité de la conversation doivent être considérés. Selon la description auparavant (voir Ch. 3), les chercheurs du domaine de l'analyse conversationnelle préfèrent travailler avec des interactions spontanées qui reflètent au mieux la communication humaine, alors que les phonéticiens optent pour du matériel où la qualité de l'audio est excellente. Cette opposition conduit à une problématique au sujet des conclusions que les deux approches peuvent tirer de leurs matériaux. L'analyse conversationnelle se retrouve face à un problème où les différences intonatives ne sont pas toujours mesurables, tandis que les phonéticiens examinent des données non-spontanées qui ne peuvent pas servir à tirer de conclusions sur les conversations dans leur milieu naturel.

En conséquence (de ces observations), nous avons l'intention de nous placer entre ces deux approches en cherchant à combiner spontanéité et une bonne qualité de l'audio. Il nous semble par conséquent fondamental de disposer d'un corpus qui possède, dans la mesure du possible, ces deux traits. Nous nous sommes donc servis d'un corpus de 23 conversations spontanées du français compilées par Torreira et al. (2010) : « The Nijmegen Corpus of Casual French ». Ce corpus comprend 23 enregistrements audio de 46 participants (24 locutrices et 22 locuteurs) qui permettent des analyses intonatives détaillées sans pour autant perdre le trait de spontanéité. Le corpus contient aussi bien des enregistrements audio que vidéo, ce qui nous a permis de prendre en compte les signaux d'écoute (régulateurs) émis par le canal visuel (gestes). Les enregistrements dans leur totalité durent aux alentours de 90 minutes (Torreira et al. 2010 : 203), alors que nous n'employons que la partie où la conversation est de nature dialogique, ce qui a résulté en 323 exemples.

Chaque conversation comprend trois parties thématiquement différentes et trois participants, avec un confédéré (un locuteur conscient de la procédure de l'enregistrement qui recrute deux amis). Deux minutes après le début de la conversation, le technicien entre dans la pièce d'enregistrement et indique que le micro du confédéré ne marche pas. Ils sortent ensemble sous prétexte que le micro doit être remplacé. Le confédéré parti, il ne reste alors que deux participants dans la pièce d'enregistrement qui ne sont pas conscients du fait qu'ils

sont toujours enregistrés. Il est cependant clair que la plupart d'entre eux soupçonnent que l'enregistrement n'est pas arrêté lorsqu'ils disent par exemple « tu penses qu'on est enregistré ? ». Après que le confédéré est sorti avec le technicien, le dialogue qui suit est de nature semi-spontanée et les participants paraissent oublier la situation dans laquelle ils se trouvent. Cette partie de la conversation (sans le confédéré) est de différente longueur (de quatre à 40 minutes) et elle continue jusqu'à ce que le dialogue perde sa vivacité. Nous avons cessé de consulter la conversation lorsque le confédéré rentre dans la pièce d'enregistrement et le dialogue devient à nouveau un triangle.

4.2. ORGANISATION SPATIALE

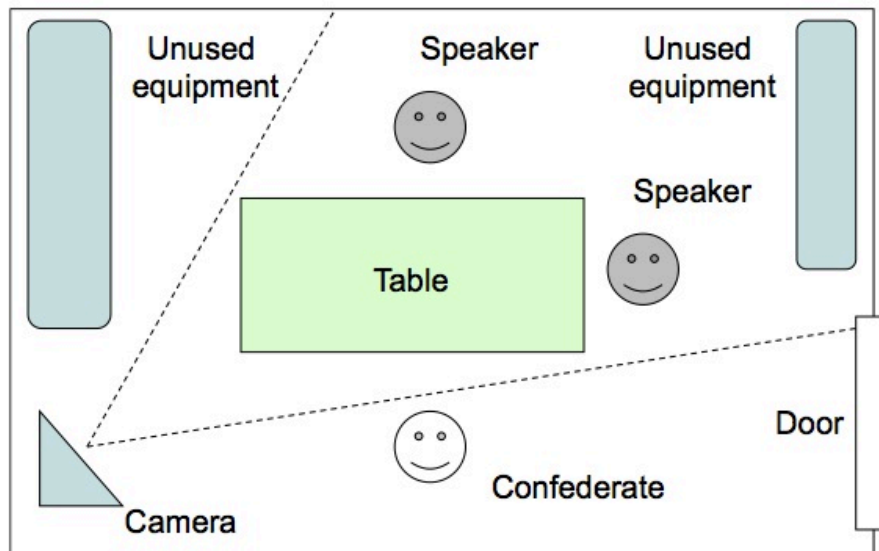


Fig. 1. Layout of the recording room.

Figure 4. Cette figure est une capture d'écran d'une figure présentée dans Torreira et al. (2010 : 23). Elle illustre l'organisation spatiale des participants dans la pièce d'enregistrement.

L'enregistrement n'a pas seulement concerné des canaux audio, mais du matériel vidéo a également été tourné sans que les participants en soient conscients. Pour que leur comportement ne soit pas affecté de la conscience d'être filmé, la lumière rouge de la caméra était couverte avec du ruban adhésif et de l'équipement électronique (ordinateur, écran, etc.) a été placé sur des tables dans la pièce d'enregistrement (voir Figure 4). Comme nous avons pu le remarquer au sujet des soupçons de l'enregistrement audio, certains participants ont également fait référence à la possibilité d'être filmés en commentant la caméra (« T'es

allumé ? » adressé à la caméra). Avec un trépied, la caméra a été placée dans un angle afin de bien capter l'interaction entre les deux participants (qui restent pendant le dialogue). Il est ainsi possible de suivre avec aisance les regards et les gestes des interlocuteurs (voir Figure 5). La caméra utilisée est un Canon XM2 Mini-DV, fonctionnant sur batterie, avec une pendule débranchée pour donner l'impression qu'elle ne tournait pas.



Figure 5. Cette figure (adaptée à Torreira et al. (2010 : 23)) est une capture d'écran d'un des vidéos du corpus.

4.3. AUDIO – DEUX CANAUX

Même si trois locuteurs étaient présents durant la majeure partie de la conversation (sauf la partie dialogique), les paroles ne sont enregistrées que pour deux participants (ceux qui se trouvent dans le cadre de la caméra). Afin de permettre une analyse phonétique de chaque locuteur sans que des sons émis par l'interlocuteur se chevauchent, deux pistes séparées ont été utilisées. L'enregistrement a été effectué à l'aide d'un enregistreur portable stéréo Edirol R-09, des micros Samson QV head-mounted unidirectionnel et d'un préamplificateur microphone stéréo. La fréquence d'échantillonnage était mise à 48 kHz avec une résolution de quantifications de 32 bits. Pour éviter des phrases occlusives des consonnes obstruantes, d'autres bruits non-linguistiques, etc., les micros ont été placés au côté gauche à 5 cm de la bouche. Les participants ont également reçu l'instruction de ne pas taper sur la table, de ne pas jouer avec des clés, ou de faire autre chose qui gêne l'enregistrement. Avec ces deux

canaux, l'analyse phonétique de la fréquence fondamentale a manifestée de belles traces de la fréquence fondamentale.

5. RÉSULTATS

Ce chapitre présente les différents résultats qui essaient d'établir la distinction entre les deux montées de la f_0 , à l'aide des traits intonatifs sous la syllabe finale (finaux) et des traits qui se manifestent plus tôt dans l'énoncé (non-finaux). Selon cette division des traits, il est possible de déterminer si les deux fonctions sont également différentes avant la syllabe finale qui est généralement censée être l'endroit privilégié de distinction des deux montées. Le but primaire de cette étude est alors de déterminer si les deux montées peuvent être séparées à l'aide de l'intonation, un but secondaire étant de discuter autour le rôle potentiel des mesures au sujet de l'alternance des tours de parole. Puisque cette étude n'est pas centrée sur la perception, notre analyse conclut uniquement sur la présence des signaux émis par les locuteurs, sans proposer pour autant d'explication sur la manière dont les interlocuteurs peuvent s'en servir.

Field et al. (2012 : 266) argumentent que le premier modèle linéaire doit inclure toutes les variables exogènes qui sont censées avoir une influence significative sur la variable endogène, afin d'ensuite répéter la régression en excluant les variables qui ne manifestent pas un effet signification (voir Ch. 3.1.4.1. pour un exposé plus détaillé sur la régression linéaire et ses éléments). Le modèle est ainsi optimisé en n'incluant que des prédicteurs qui ont une influence significative sur le résultat du modèle. À la fin de ce chapitre une estimation du degré de la capacité de prédiction procuré par chaque variable endogène sur laquelle la fonction manifeste une influence significative est calculée à l'aide d'une simulation « leave-one-out cross-validation ». Les résultats des tests statistiques sont illustrés à l'aide des diagrammes en boîte dont nous nous servons afin d'illustrer les tendances des modèles linéaires qui indiquent les médianes avec la barre au milieu de la boîte. Les boîtes comprennent 50 % des données, où 25 % sont placés sur chaque côté de la barre au milieu, alors que le reste des données est résumé par les deux lignes à chaque côté de la boîte. Les données aberrantes (le terme pour des valeurs extrêmes) sont indiquées par des points.

5.1. EXEMPLES TIRÉS DU CORPUS

Ce chapitre présente des exemples tirés du corpus, qui illustrent de quelle manière une locutrice se sert de la f_0 afin de signaler une continuation ou une question. L'observation des figures permet de remarquer que les deux fonctions diffèrent peu en ce qui concerne les traces de la f_0 lors de l'énoncé final. La décision a été prise de décrire une continuation et une question pour illustrer les différences qui ne concernent que les variables liées à la fréquence

fondamentale, puisqu'elle est le seul paramètre visible. La similitude au sujet du maximum de la f0 de ces exemples démontre bien la nécessité d'une étude empirique.

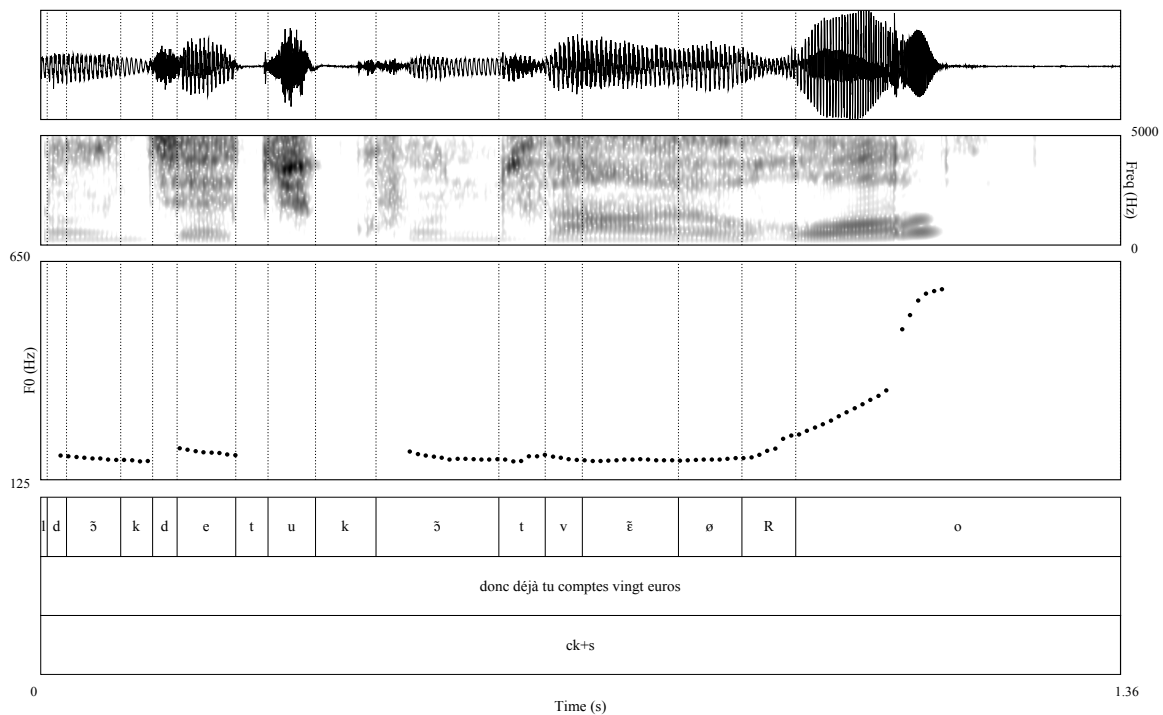


Figure 6. Cette figure montre une continuation produite par une femme. (22-11-07_1_p1_F06R_c_1012205).

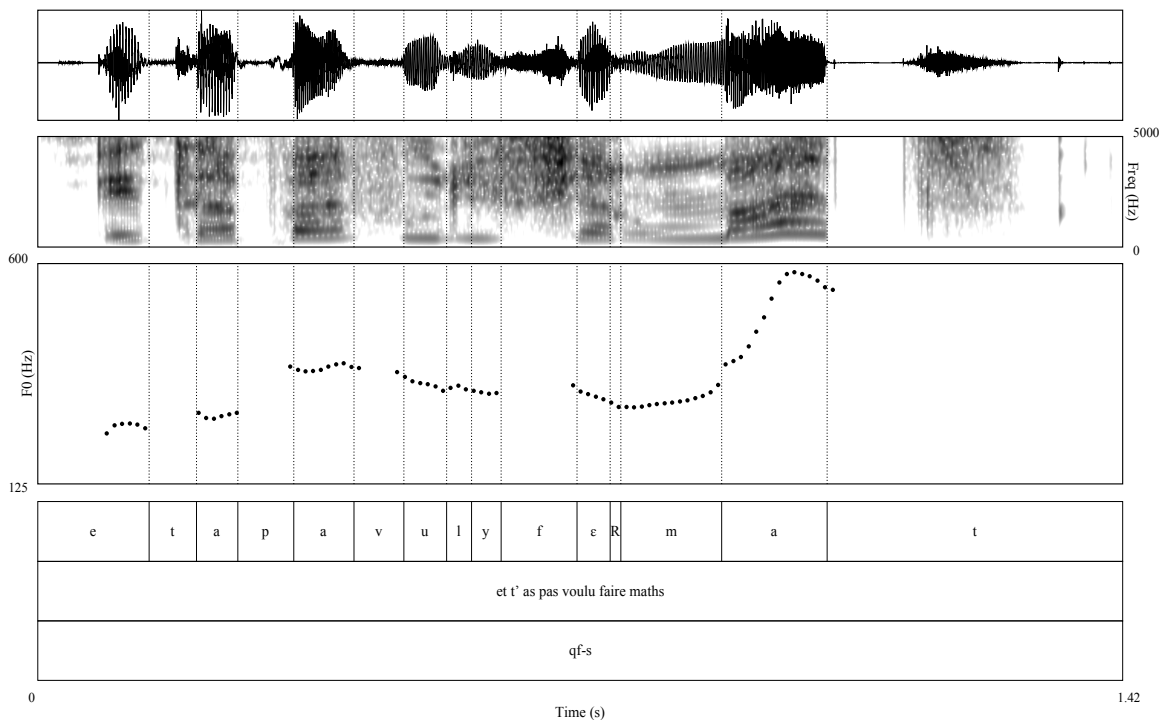


Figure 7. Cette figure montre question produite par une femme. (22-11-07_2_p1_F07R_q_857418).

5.2. SIGNAUX FINAUX

Le but de ce chapitre est d'examiner l'influence d'une fonction spécifique sur des différents signaux intonatifs situés sous la syllabe finale des énoncés. En ce qui concerne la variable prédite (le maximum standardisé de la f_0 , le minimum standardisé de la f_0 , la forme de la courbe montante, le maximum standardisé de l'intensité et la durée de la voyelle finale), nous incluons toujours la variable du sexe du locuteur dans les tests statistiques liés à la f_0 , étant donné que les hommes et les femmes ne se servent pas également de ce paramètre d'après nos observations initiales. En revanche, les variables de la présence ou l'absence d'une attaque (une consonne avant le noyau vocalique de la syllabe) ou d'une coda (une consonne après le noyau vocalique de la syllabe), ainsi que du degré d'aperture de la voyelle ne sont incluses que dans les modèles où elles peuvent avoir une influence. La raison derrière cette décision est qu'une coda voisée permet à la f_0 de continuer à monter après la fin de la voyelle, tout comme la présence d'une attaque peut influencer la f_0 de la voyelle (cf. la proposition de Ladefoged, 2003). Les propositions de Delattre (1966) indiquant une différence sous la syllabe finale des deux fonctions (la f_0), ainsi que des différences en intensité et de durée (Rossi et al. 1981 ; entre autres) entre les deux fonctions peuvent être évalués après ce chapitre.

5.2.1. LE MAXIMUM DE LA F_0

Cette partie est focalisée sur les mesures du maximum de la fréquence fondamentale qui se manifeste sous la syllabe finale de l'énoncé. Aussi bien Delattre (1966) que Rossi et al. (1981) remarquent que ce point sert à distinguer les continuations des questions, car ces dernières sont censées atteindre un niveau supérieur comparé aux premières.

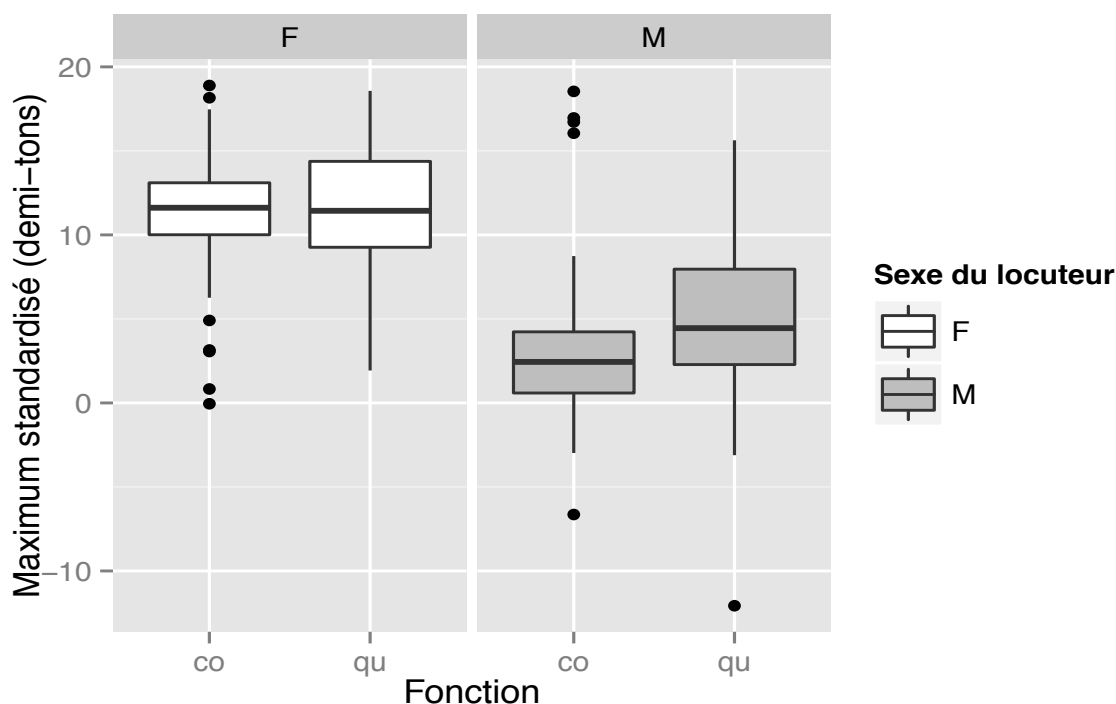


Figure 8. Cette figure présente le maximum standardisé de la f_0 (indiqué en demi-tons) qui est partagé d'après les fonctions (co = continuation, qu = question). Une division est également faite entre les sexes des locuteurs (F = locutrices (blanc), M = locuteurs (gris)).

Une tendance pour les locutrices de se servir d'un maximum de la f_0 similaire pour les deux fonctions est illustrée par la Figure 8 ci-dessus. Cette tendance se manifeste par la proximité des médianes des deux fonctions (indiquées par la barre au milieu de la boîte) et le chevauchement des boîtes. Elle est confirmée par le test statistique qui ne démontre pas une différence significative. Les deux boîtes des locuteurs ne se chevauchent pas autant que les locutrices, ce qui indique que les locuteurs font plus de distinction entre les deux fonctions à l'aide du maximum de la fréquence fondamentale. En excluant les données aberrantes, la régression linéaire démontre cette différence puisque les questions ont un maximum standardisé supérieur aux continuations ($\beta = 1.92$, $t = 2.13$, $p < 0.05$). Les locutrices se servent en général d'un maximum plus haut par rapport à leur médiane globale (mesurée dans un extrait de dix minutes et indiquée avec la valeur 0 dans le diagramme en boîte), ce qui confirme l'observation de Bolinger (1989) et bien d'autres qui ont remarqué que les locutrices emploient un champ tonal (*register* en anglais – nous concevons avec ce terme l'étendue entre son plancher et son plafond) plus étendu que les locuteurs.

5.2.2. LE MINIMUM DE LA F0

Quant au minimum de la f_0 (le point minimal de la voyelle finale), la description de Delattre (1966) illustre un même niveau pour les deux fonctions en relation au registre (c'est-à-dire champ tonal dans nos termes) du locuteur. Comme cette recherche se veut exploratrice, nous voulons examiner si cette proposition se confirme lors d'une étude empirique. Une autre possibilité est que la distinction entre un contour de continuation et un contour de question peut être établie déjà au début de la voyelle finale. Puisqu'une attaque sourde peut influencer la f_0 de la voyelle suivante (cf. Ladefoged, 2003), l'information sur sa présence est incluse comme variable exogène dans la régression linéaire.

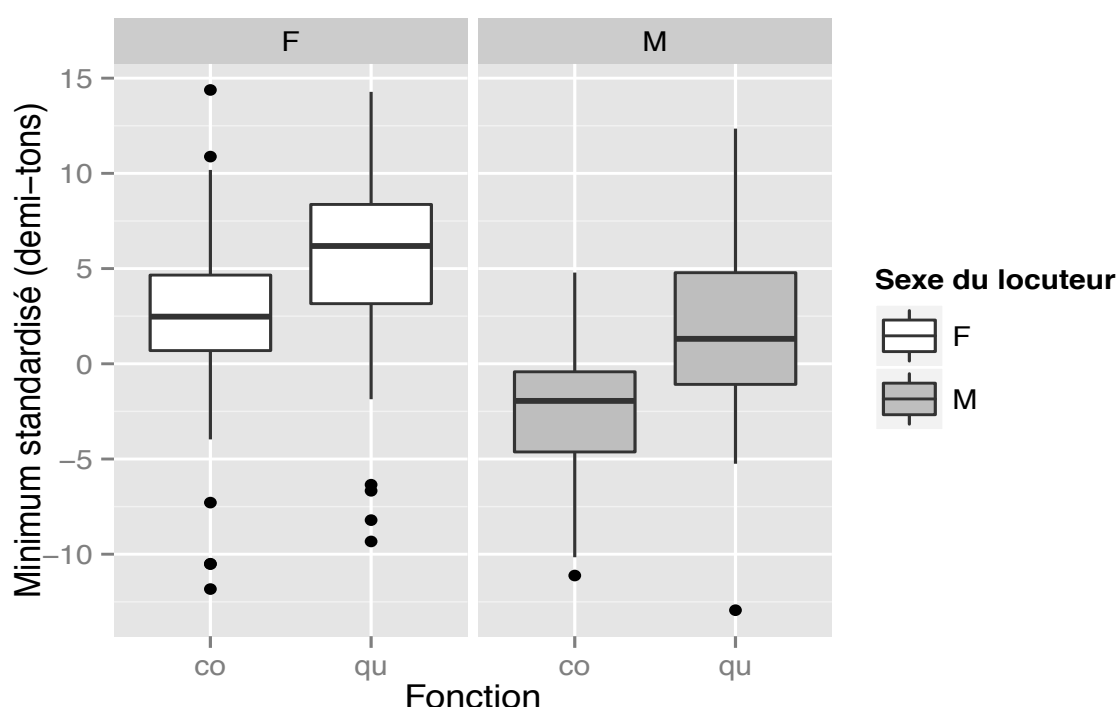


Figure 9. Ce diagramme en boîte illustre le minimum standardisé (par rapport à la médiane globale du locuteur) en demi-tons, partagé selon la fonction (co = continuation, qu = question) ainsi que selon le sexe du locuteur (F = locutrices (blanc), M = locuteurs (gris)).

Les locuteurs et les locutrices semblent signaler la différence entre les continuations et les questions de la même manière, c'est-à-dire en manifestant une f_0 supérieure déjà au début de la voyelle pour les questions. Cette tendance est visible par le peu de chevauchement entre les boîtes, ainsi que dans la régression linéaire où les questions démontrent une f_0 supérieure aussi bien pour les locuteurs que pour les locutrices ($\beta = 3.66$, $t = 7.70$, $p < 0.001$). La différence entre les locuteurs et les locutrices observée au sujet du maximum de la f_0 émerge

également pour le minimum standardisé de la f_0 . Les résultats liés au minimum de la f_0 contredisent la proposition de Delattre (1966) puisque les types démontrent une différence significative. De plus, les boîtes ne se chevauchent pas autant pour ce diagramme en comparaison avec celui du maximum de la f_0 (voir Figure 8), ainsi le minimum de la f_0 paraît un meilleur indice pour faire la distinction entre les deux fonctions aussi bien pour les locuteurs que pour les locutrices.

5.2.3. LA FORME DE LA COURBE MÉLODIQUE MONTANTE

Plusieurs propositions existent concernant la forme de la courbe mélodique des continuations et des questions. Certaines entre elles suggèrent que les deux fonctions peuvent être séparées à l'aide de leur forme. Delattre (1966) illustre les deux fonctions avec des formes opposées : les continuations avec une montée convexe (une montée précoce) et les questions avec une montée concave (une montée tardive). Grundstrom (1973) argumente cependant que la différence nette entre les deux types ne peut pas être fait seulement par la forme, admettant quand même qu'une montée en croissant (concave) est plus interrogative qu'un ton statique, et que la montée en décroissant (convexe) n'est pas interrogative du tout. Rossi (1978), ainsi que Rossi et al. (1981) s'opposent à cette affirmation en démontrant que la forme ne peut pas distinguer les deux fonctions.

Afin d'examiner quelle proposition qui est valable dans nos données, nous avons prêté attention au rayon de la courbure des deux types de montées, en employant un polynôme qui nous permet de calculer le degré de courbure pour une montée. D'après les coefficients obtenus, il est possible de déterminer si une certaine montée est ou concave ou convexe. La méthodologie est empruntée à Grabe et al. (2007), Andruski et Costello (2004) et Torreira (2007). Elle utilise une régression quadratique qui nous signale le coefficient du rayon de la courbure. Un coefficient négatif indique une montée précoce dans la courbe (convexe), alors qu'une valeur positive indique une montée tardive (concave).

La distribution fait preuve d'une grande dispersion entre les données en ce qui concerne leur forme de la f_0 sous la voyelle finale. Il s'avère que certains exemples manifestent une brève chute de la f_0 lors des dernières ms de la montée de la syllabe identifiée à partir de la forme d'onde, ce qui est probablement due au fait que la glotte des locuteurs retourne à une position neutre pendant un intervalle pendant laquelle la phonation n'a pas terminé. Ainsi, la variable de la forme ici ne comprend que la montée jusqu'à son point

maximum. La distribution montre que les questions ont une légère tendance à des coefficients plus négatifs que les continuations (une montée plus précoce), bien que les deux fonctions se chevauchent beaucoup. Néanmoins, les résultats de la régression linéaire sont non-significatifs ($\beta = -60, t = -0.70, p < 0.5$), tout comme l'analyse suivante qui ne comprend que les exemples se trouvant à l'intérieur de la portée de deux écarts types aux deux côtés de la moyenne ($\beta = -103, t = -1.76, p < 0.08$). Cela nous amène à conclure que la proposition de Delattre (1966) n'est pas reproductible dans une étude de corpus, ce qui a également été remarqué par Rossi et al. (1981). Même en éliminant les données aberrantes une différence significative ne peut pas être obtenue, un résultat illustré par les figures ci-dessous qui démontrent que les deux fonctions peuvent être accompagnées des deux formes.

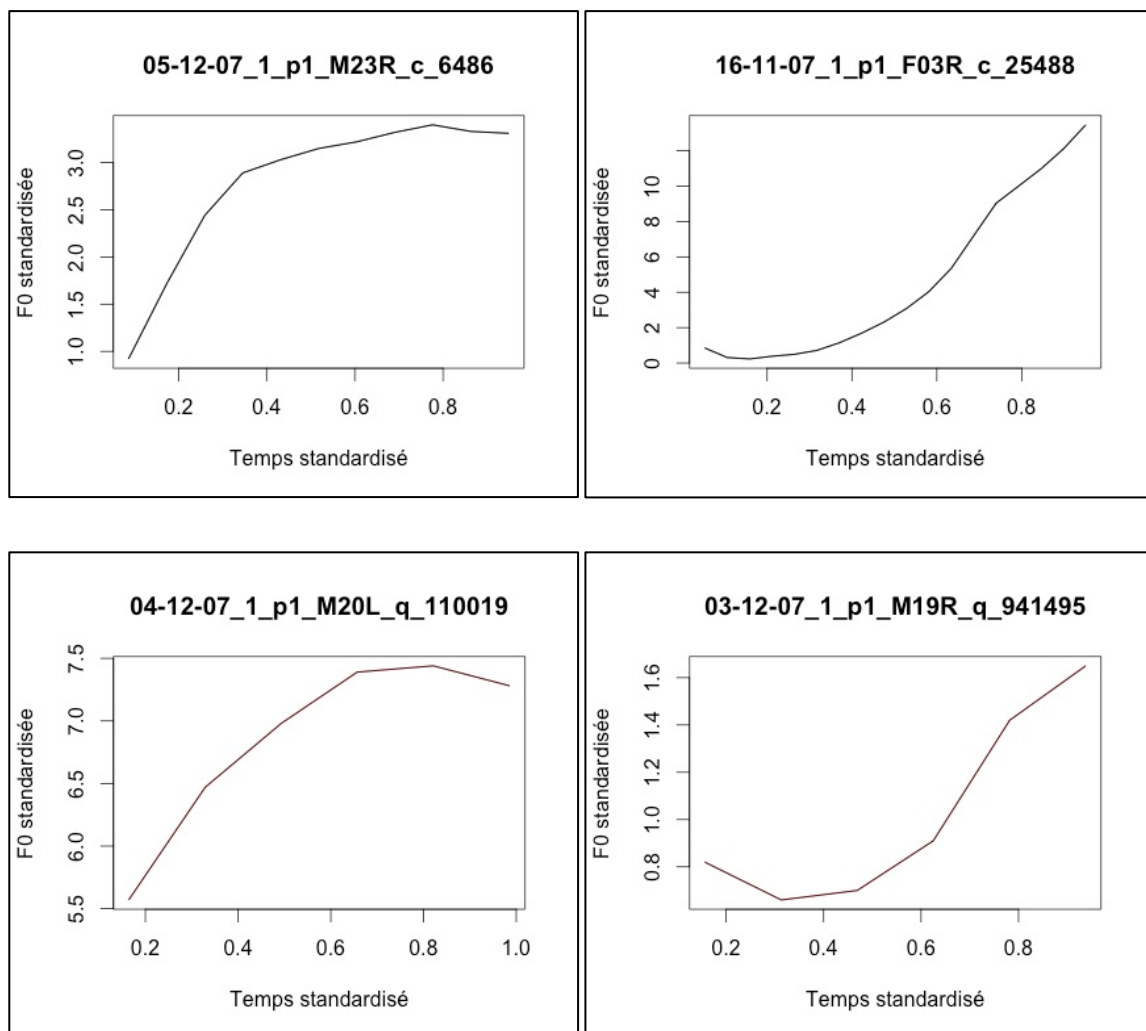


Figure 10. Ces diagrammes illustrent deux images de chaque fonction (continuation et question). Il est ainsi possible d'observer que les formes de la f_0 des deux fonctions peuvent être convexes aussi bien que concaves. Les contours en haut indiquent des continuations et les contours en bas indiquent des questions.

5.2.4. L'INTENSITÉ

Selon Duncan (1972), une intensité en position finale inférieure comparé à ce qui la précède a été classée comme un turn-yielding cue d'après son étude qualitative en anglais de deux enregistrements vidéo. Cette proposition a récemment été validée par une étude de corpus, toujours pour l'anglais, par Gravano et Hirschberg (2011). Comme l'alternance des tours de parole est un trait incorporé dans le type prototypique des questions, nous nous attendons à ce que ce groupe manifeste une intensité inférieure comparée à celles des continuations. Étant donné que l'intensité est inhérente au type de voyelle, une classification concernant le degré d'aperture est introduite dans le test statistique (fermées, moyennes et ouvertes). D'après la classification de Léon (1996 : 82) nous avons classé les voyelles d'après leur degré d'aperture qui varie entre ouvert et fermé. Le groupe dit *moyennes* contient ce que Léon nomme ouverte et fermée, alors que *très fermées* et *très ouvertes* dans sa classification correspondent à *fermées* et *ouvertes* dans la nôtre.

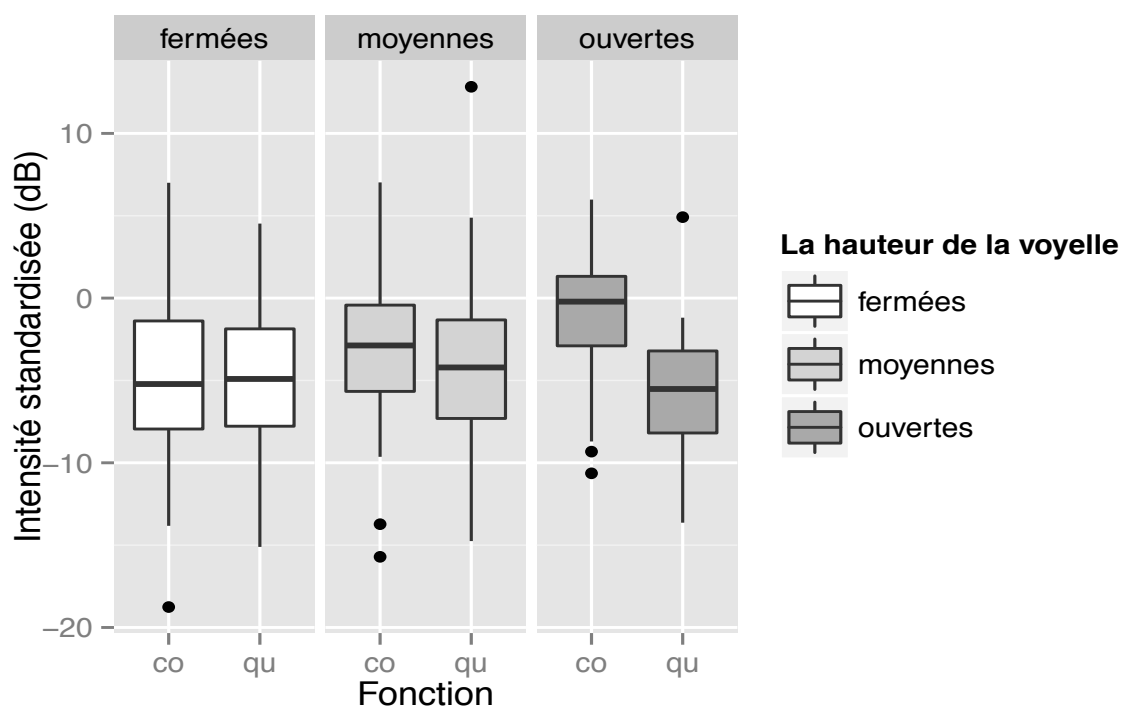


Figure 11. Ce diagramme présente le maximum de l'intensité standardisée (par rapport au maximum dans la seconde précédente) indiquée en dB, avec une division par le degré d'aperture de la voyelle. La fonction est indiquée avec 'co' (continuation) et 'qu' (question). Les voyelles fermées sont indiquées avec le blanc, les voyelles moyennes sont grises, et les voyelles ouvertes ont le couleur gris foncé.

La différence significative entre les deux fonctions dans les voyelles ouvertes pour lesquelles les boîtes sont distinctes ($\beta = -5.09$, $t = -3.34$, $p < 0.001$) est illustrée par le peu de chevauchement dans le diagramme ci-dessus. En revanche, les voyelles fermées démontrent un chevauchement entre les boîtes, alors que les voyelles moyennes manifestent une légère tendance d'absence de chevauchement. Au contraire, cette différence n'est pas représentée dans les statistiques même en excluant les données aberrantes. Cela pourrait indiquer que seules les voyelles ouvertes, qui exhibent une intensité supérieure aux autres voyelles (plus la bouche est ouverte, plus l'intensité augmente (Ladefoged, 2003)), présentent une différence suffisamment grande entre le maximum sous la voyelle finale et le maximum dans la seconde précédente. Elles seraient donc les seules à révéler une différence significative entre les deux fonctions. En ce qui concerne les voyelles ouvertes, les résultats avec une intensité inférieure pour les questions confirment les résultats de Duncan (1972) et Gravano et Hirschberg (2011) avec la relation entre une intensité inférieure et un abandon de tour.

5.2.5. LA DURÉE DE LA VOYELLE FINALE

Ce chapitre traite de la durée de la voyelle finale dans le but d'examiner si la variabilité au sujet de cet indice intonatif peut être expliqué par la fonction. La présence ou l'absence d'une coda syllabique est introduite comme variable exogène dans le test statistique, puisqu'il est bien connu que cette variable peut influencer la durée de la voyelle (Léon, 1996 ; Di Cristo, 1978). Di Cristo a également établi une distinction entre la durée inhérente des voyelles orales et nasales, un facteur qui ne sera cependant pas inclus dans notre test statistique, étant donné que l'échantillon ne nous semble pas suffisamment grande pour faire cette distinction. Dans la littérature qui concerne la distinction entre les deux fonctions, Rossi et al. (1981) ont remarqué qu'une question à laquelle on peut répondre oui ou non (« C'est l'ami de Luc ? ») manifeste une durée vocalique significativement plus longue que celle d'une continuation, alors que d'autres types de questions ne présentent pas cette même tendance. D'après Duncan (1972) et Cutler et Pearson (1986) entre autres, un mot dans une position finale d'énoncé tend à être plus long qu'un mot en position non-finale. Toutefois, Gravano et Hirschberg (2011) rapportent que l'allongement peut apparaître aussi bien en position finale d'énoncé au milieu et à la fin de tour, au contraire il semble plutôt avoir plus d'importance dans une position non-finale de tour au cause de la différence au sujet de la différence de débit de parole.

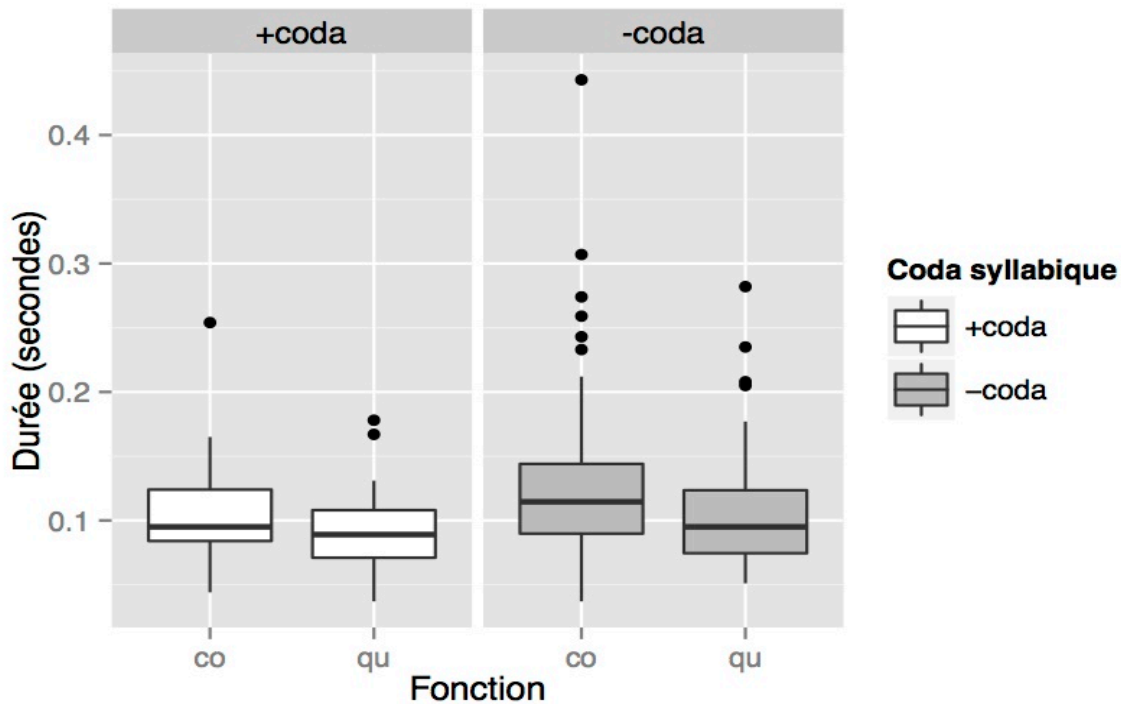


Figure 22. La durée de la voyelle finale (en secondes) est divisée par fonction (co = continuation, qu = question) ainsi que par la présence (blanc) et l'absence d'une coda (gris).

Le diagramme en boîte de la durée en secondes semble indiquer que la distinction entre les deux fonctions est établie à l'aide de l'allongement de la voyelle finale dans un énoncé final. Cette distinction est visible aussi bien pour les syllabes ouvertes que pour les voyelles fermées, or la distinction pour les syllabes ouvertes semble être un peu plus importante. La régression linéaire comprend les variables du débit de parole, ainsi que de la présence d'une coda, qui toutes deux peuvent influencer la durée d'une voyelle. Un débit de parole supérieur raccourcit la durée d'une voyelle, alors que la présence d'une coda pourrait également l'influencer. Les tendances signalées dans le diagramme sont démontrées significatives d'après notre modèle ($\beta = -0.02$, $t = -3.23$, $p < 0.01$) avec une durée supérieure pour les continuations. Ces résultats sont néanmoins conformes à ce que Gravano et Hirschberg (2011) ont proposé en indiquant que l'allongement du mot final est plus marqué pour un énoncé non final de tour (2011 : 614). Il faut cependant remarquer que la différence n'est que de 20 ms. Ensuite, leur différence de 400 ms entre les mots finaux est calculée sur le mot entier, ce qui est difficile à comparer à notre étude qui n'a pris que la voyelle finale en compte.

5.3. SIGNAUX NON-FINAUX

Cette partie du chapitre sur les résultats présente les données obtenues par mesures prises avant la syllabe finale. Comme certains phénomènes intonatifs se manifestent avant la syllabe finale (le débit de parole, le registre de l'énoncé), ils sont traités dans une partie séparée de l'analyse. Dans le cas où ces mesures pourraient établir une distinction entre les deux fonctions, il serait possible de dire que le signal pour une fonction se fait déjà avant la syllabe finale qui est souvent l'endroit censé être responsable pour la distinction.

5.3.1. LE REGISTRE

D'après les observations précédentes (cf. Ch. 5.2.2.), il se peut que la fonction soit signalée différemment déjà avant la syllabe finale. Étant donné que les questions ont déjà une f_0 plus haut dans le champ tonal du locuteur au début de la montée, il serait intéressant de regarder si cette différence est déjà signalée par le registre (un niveau à l'intérieur du champ tonal) employé par le locuteur avant que la syllabe finale ne commence. Afin d'étudier cette possibilité, nous avons mesuré la médiane dans chaque énoncé (la syllabe finale exclue) qui est ensuite standardisée par rapport à la médiane globale obtenue dans l'extrait de dix minutes au début de chaque conversation.

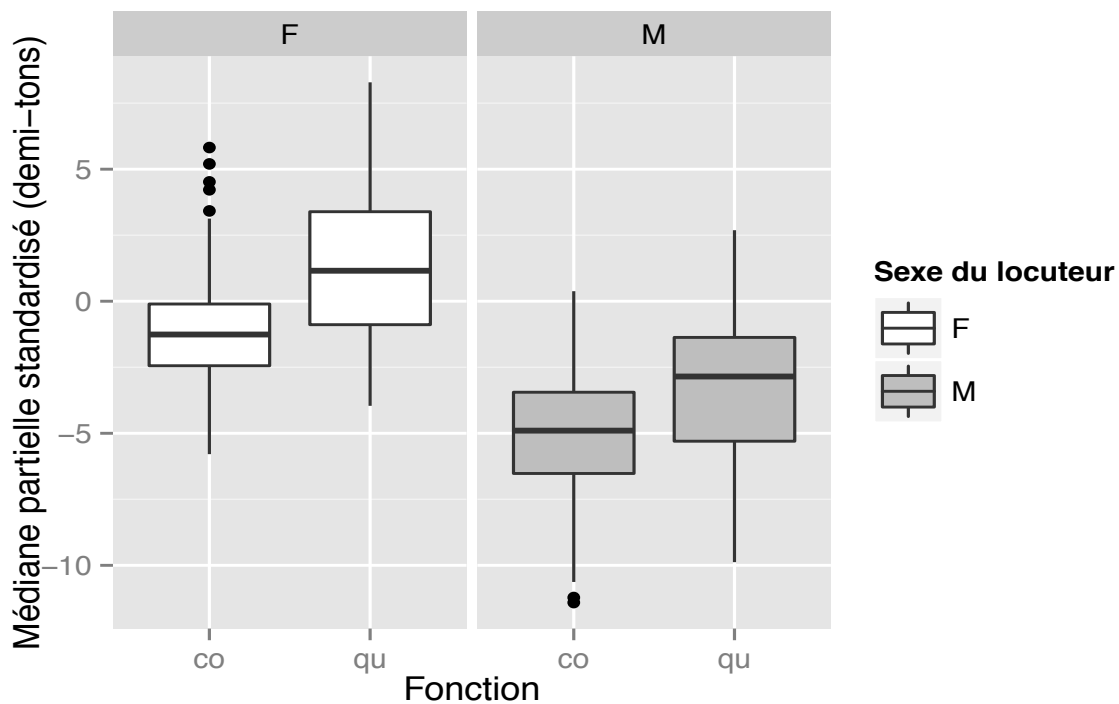


Figure 13. Ce diagramme en boîte illustre la médiane de l'énoncé en excluant la syllabe finale où la montée se réalise (mesurée en demi-tons), divisée par fonction ('co' = continuation, 'qu' = question) et par le sexe du locuteur (locutrices = blanc, locuteurs = gris).

Il semble que la médiane de l'énoncé (jusqu'à la syllabe finale) peut aider à établir une différence entre les deux montées. Les locuteurs et les locutrices manifestent la différence de la même manière en présentant les questions avec une f_0 en moyenne supérieure comparé à celle des continuations ($\beta = 2.15$, $t = 6.99$, $p < 0.001$). Ces résultats peuvent indiquer que ce n'est pas seulement la syllabe finale qui est différente au sujet des mesures de la f_0 , mais que l'énoncé entier peut signaler cette tendance (la syllabe finale exclue). Après avoir consulté les données de nouveau, nous pouvons constater que la différence concernant le registre n'est pas due à un accent haut pendant l'énoncé, étant donné que les traces de la f_0 sont assez plates avant que la montée finale ne commence dans la plupart des cas. Les locutrices démontrent une gamme (range en anglais – c'est-à-dire la différence entre les points minimum et maximum) inférieure pour les questions (un minimum supérieur mais un maximum égal aux continuations), ce qui peut être expliqué par le fait qu'elles parlent dès le début à un niveau supérieur dans ces énoncés qui ne leur permet pas de produire un maximum supérieur pour les questions. Cela manifeste un effet plafond (c'est-à-dire qu'il n'est pas aisé pour elles de monter davantage puisqu'elles parlent déjà à une fréquence plus haute) en ce qui concerne cette fonction.

Cette tendance peut également expliquer les différences entre les locuteurs et les locutrices observées pour les autres mesures de la f_0 . Une explication possible est que les locutrices parlent dès le début d'un énoncé interrogatif à un niveau supérieur comparé à leur médiane globale (mesurée dans un extrait de dix minutes). Cependant, en ce qui concerne le maximum de la f_0 , puisqu'elles parlent déjà à un niveau supérieur, elles ne peuvent pas faire la distinction entre les questions et les continuations.

5.3.2. LE DÉBIT DE PAROLE

Gravano et Hirschberg (2011) ont, à l'aide d'une étude de corpus map task, observé que le débit de parole augmente avant une borne de tour de parole. Leurs résultats indiquent ainsi que les questions seront prononcées avec un débit de parole supérieur comparé à celui des continuations, puisque les premières sont liées à une alternance de tour de parole et ainsi de locuteur. Le débit de parole est calculé pendant le dernier syntagme intonatif de l'énoncé en

indiquant son début et sa fin (avant la syllabe finale de l'énoncé) et en divisant le nombre de syllabes prononcées par la durée. Cette mesure est compliquée dans les exemples qui ne comprennent qu'une syllabe précédant la syllabe finale, un facteur qui ne peut pas être contrôlé dans des données spontanées. La mesure du débit de parole reflète le nombre de syllabes par seconde (Gravano et Hirschberg (2011) ; Cangemi et D'Imperio (2011) ; Van Heuven et Van Zanten (2005)).

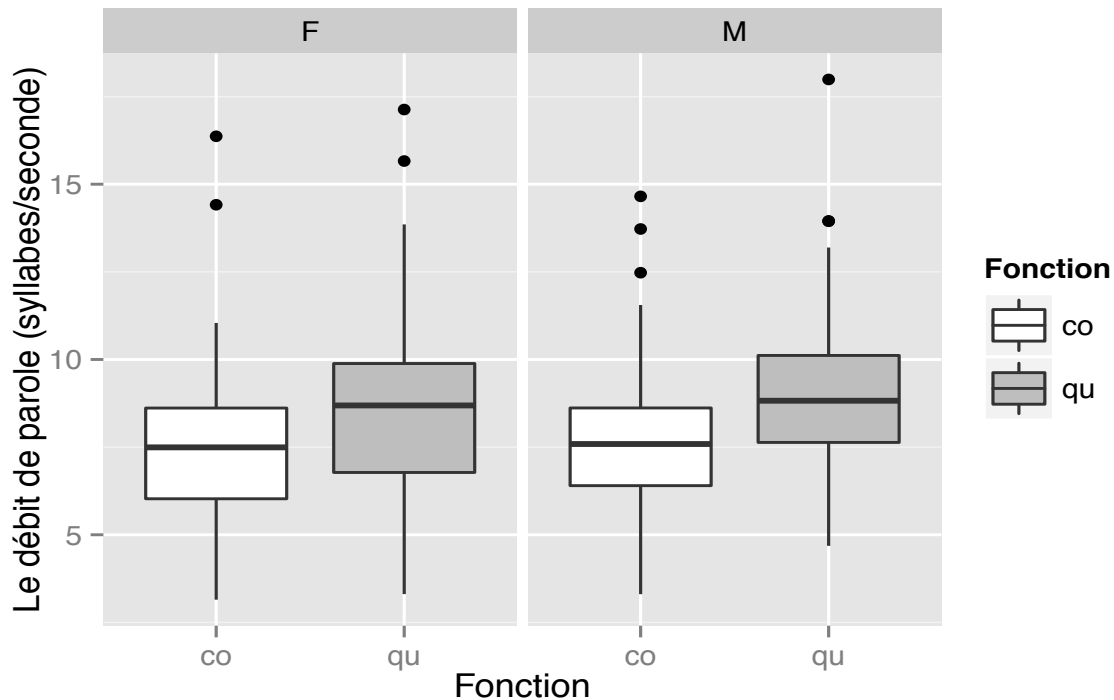


Figure 14. Le débit de parole en syllabes par seconde est indiqué dans ce diagramme en boîte, divisé par fonction ('co' = continuation, 'qu' = question). Les boîtes blanches indiquent les locutrices et les boîtes grises les locuteurs.

La tendance observée par Gravano et Hirschberg (2011) semble également se manifester dans nos données où la boîte des questions indique un débit de parole supérieur à celui des continuations ($\beta = 1.15$, $t = 4.45$, $p < 0.001$). Comme les questions signalent une alternance de tour de parole, le débit de parole semble jouer le même rôle dans le corpus utilisé par Gravano et Hirschberg (2011) que dans le nôtre. Ce trait intonatif est ainsi encore un qui aide à établir la différence entre les deux fonctions avant la syllabe finale.

5.4. LE RÔLE DE LA SYNTAXE

Comme les alternances de locuteurs sont en partie liées à la complétude de la syntaxe, nous voulons examiner son influence sur les mesures intonatives. Seules les continuations peuvent être soit complètes, soit incomplètes syntaxiquement, c'est la raison pour laquelle nous n'avons incluse que cette fonction dans les premières analyses. Il s'avère que la complétude syntaxique n'a de l'influence que sur les mesures qui ne sont pas liées à la fréquence fondamentale. Cette influence manifeste la même tendance que les questions en faisant descendre l'intensité ($\beta = -1.69, t = -2.01, p < 0.05$), ainsi que la durée de la voyelle finale ($\beta = -0.03, t = -3.02, p < 0.01$), alors qu'elle fait augmenter le débit de parole ($\beta = 1.4, t = 3.49, p < 0.001$). Ces résultats démontrent qu'une continuation qui indique le maintien du tour de parole à l'aide de la syntaxe (incomplète) renforce ce signal étant donné qu'elle est accompagnée des autres traits qui indiquent un maintien. Cela veut dire que les continuations syntaxiquement incomplètes signalent encore plus le maintien aussi bien avec l'intonation que la syntaxe que les continuations syntaxiquement complètes.

La différence entre les continuations complètes et les questions concerne aussi bien la fréquence fondamentale que les autres variables intonatives. Bien que les deux soient syntaxiquement complètes, ces deux groupes se distinguent concernant les mesures intonatives. Les variables qui ne reflètent pas la f_0 dépeignent une trichotomie où les continuations incomplètes se distinguent des continuations complètes qui à leur tour sont différentes des questions. En centrant d'abord l'attention sur la différence de la f_0 entre les continuations complètes et les questions, ces dernières ont un maximum standardisé supérieur ($\beta = 1.56, t = 3.36, p < 0.001$), un minimum standardisé de la voyelle finale supérieur ($\beta = 3.46, t = 6.78, p < 0.001$) et un registre supérieur avant la syllabe finale ($\beta = 2.26, t = 7.13, p < 0.001$) par rapport aux premières (après avoir contrôlé l'effet du sexe du locuteur). En ce qui concerne les autres variables intonatives, les questions manifestent une intensité inférieure ($\beta = -1.54, t = -2.96, p < 0.01$) et un débit de parole supérieur ($\beta = 0.94, t = 3.56, p < 0.001$) alors que la différence en durée n'est pas significative.

D'après les résultats dans ce chapitre, il est possible de constater que la syntaxe joue bien un rôle à l'intérieur des continuations. Les variables qui sont censées jouer un rôle dans l'alternance des tours de parole (cf. Duncan, 1972 ; Gravano et Hirschberg, 2011 ; entre autres) semblent être liées à la complétude syntaxique, bien que les questions sont toujours différentes des continuations. Au sujet de la différence de la durée rapportée par Gravano et

Hirschberg (2011) l'on pourrait s'imaginer qu'elle était potentiellement confondue avec la complétude syntaxique. Il est intéressant de noter que non seulement la syntaxe aide à garder la parole, mais que les autres variables de l'intonation (sauf la durée) font aussi une distinction qui est additionnelle à la syntaxe. L'effet exhibé par la durée de la continuation peut être interprété comme une sorte de moment pour la planification de ce qui suit. En produisant une continuation qui n'est pas syntaxiquement liée à ce qui suit (complétude syntaxique) les locuteurs se servent d'une durée supérieure, alors qu'ils n'ont pas ce besoin aux cas où la planification est plus facile. En ce qui concerne les continuations complètes, elles se trouvent au centre de la trichotomie. D'après la proposition de Selting (2000) qui suggère que l'importance de la syntaxe peut dépasser la prosodie, l'on pourrait imaginer qu'une montée qui signale un maintien de parole avec la syntaxe ne doit pas émettre plus de signaux qui indiquent que ce contour empêcherait une alternance, cependant cette hypothèse n'est pas vérifiée par nos données. Les deux types de variables (syntaxe et intonation) semblent collaborer en manifestant le maintien du tour, cela est rendu possible car les continuations complètes se distinguent aussi bien par la f_0 que par d'autres variables de l'intonation des questions. Afin d'interpréter la différence entre les continuation complètes et les questions, le locuteur émet alors plus d'indices (la f_0 étant supplémentaire) pour faciliter la compréhension de l'interlocuteur.

5.5. LE POIDS DES VARIABLES DANS LA CLASSIFICATION DES FONCTIONS

Ce chapitre présente notre recherche sur le poids qu'une ou plusieurs variables expriment dans la détermination de la fonction. Une simulation « leave-one-out cross-validation » (Picard et Cook, 1984) nous permet de créer une boucle capable de prédire un dernier exemple à partir des informations d'un modèle estimé à partir du reste des exemples (322 sur 323). Il est possible de la considérer sous la forme d'un cercle où toute cercle sauf une partie est inclus dans le modèle. Après que ce modèle ait prédit la partie exclue, le résultat sur son exactitude est sauvegardé. Le modèle est effacé et un nouveau modèle qui contient tout le cercle sauf une nouvelle partie exclue (un autre exemple) est construit. La simulation fait cette procédure jusqu'à ce que tous les exemples aient été prédits. Le raisonnement derrière cette simulation consiste à évaluer la contribution d'une variable concernant la classification des types prototypiques, ce qui permet d'observer leur influence dans la distinction entre les deux fonctions dans de nouvelles données jamais rencontrées. Un modèle linéaire généralisé est entraîné sur les exemples qui prédisent le type du dernier exemple, c'est-à-dire que 323

prédictions sont faites, incluant toujours les informations des coefficients obtenus à partir du reste des exemples. Comme variables exogènes, nous incluons toutes les variables endogènes dans les régressions linéaires décrites dans le chapitre des résultats sur lesquelles le type prototypique a eu une influence significative. Ces variables sont alors divisées en deux groupes : signaux finaux et non-finaux. Les signaux finaux comprennent le maximum standardisé de la f_0 , le minimum standardisé de la f_0 , la chute d'intensité dans la voyelle finale et la durée de la voyelle finale. Le registre de l'énoncé partiel et le débit de parole avant la syllabe finale font partis des signaux non-finaux. Ensuite la variable du sexe du locuteur (ainsi que des interactions entre les mesures de la f_0 et le sexe du locuteur dans les cas où cette dernière variable a joué un rôle) est également incluse.

Notre intérêt est d'étudier comment ces variables permettent de faire la distinction entre les deux types sans avoir accès à l'information de la complétude syntaxique, de l'autorité épistémique et d'autres signaux communicatifs que l'interlocuteur possède lors de la conversation (gestes, etc.). Une classification supérieure à 50 % pour un choix binaire est jugée comme bonne, ce qui signifie que la simulation fait une classification qui est au-dessus du niveau de la chance (c'est-à-dire que le modèle ne prend pas des décisions au sujet du type par hasard). Outre, une différence significative, si la classification est bonne cela indique encore plus que les mesures peuvent être employées par les interlocuteurs afin d'identifier les fonctions.

| Les variables incluses dans la simulation | Le pourcentage des choix corrects |
|---|--|
| Le maximum et le minimum de la f_0 , l'intensité, la durée, le registre, le débit de parole | 75 % |
| Le maximum et le minimum de la f_0 , le registre | 74 % |
| Le maximum et le minimum de la f_0 | 75 % |
| L'intensité, la durée, le débit de parole | 67 % |
| Le minimum de la f_0 , le registre | 73 % |
| Le registre | 68 % |

Tableau 3. Ce tableau indique le pourcentage de choix corrects que les différentes simulations manifestent.

Tous les résultats d'après les différentes simulations « leave-one-out cross-validation » sont résumés dans le tableau 3 ci-dessus auquel nous faisons référence tout au long du reste de ce chapitre. La première simulation, qui inclut comme variables exogènes toutes les variables endogènes significatives mentionnées auparavant, fait une classification qui est bien au-dessus du niveau de la chance. 75 % des exemples sont classés correctement d'après le modèle puisque 242 sur 323 exemples ont eu la bonne étiquette. Comme plusieurs

variables exogènes sont incluses dans cette simulation qui atteint un pourcentage des choix corrects supérieur à la chance, nous voulons également étudier comment des modèles plus simples sont comparables au modèle comprenant toutes les variables. Il est alors intéressant d'étudier si seules les mesures de la fréquence fondamentale peuvent établir une distinction comparable entre les deux fonctions. Il s'avère que cette simulation (n'incluant que des mesures de la f_0) ne fait descendre le pourcentage d'exactitude que d'un pourcent à 74 %. En revanche, la simulation qui n'inclut que les variables de la f_0 qui se manifestent sous la syllabe finale fait remonter à nouveau le pourcentage d'exactitude à 75 %. Cependant, la simulation qui exclue les mesures de la f_0 , n'incluant que les autres mesures de l'intonation (l'intensité, la durée, le débit de parole), fait descendre l'exactitude à 67 % (toujours au-dessus du niveau de la chance). Il nous semble pourtant que les valeurs de la f_0 prises avant la montée (le minimum standardisé et la médiane partielle de l'énoncé) sont les mesures individuelles qui nous indiquent le mieux la fonction. Le minimum standardisé atteint un taux de 73 % des prédictions correctes, alors que la médiane avant la syllabe finale atteint 68 %. Ces résultats indiquent clairement que les informations émises précédemment à la montée de la f_0 peuvent également jouer un rôle lors de la classification des deux fonctions.

Notre étude établit une distinction entre des variables finales et non-finales, puisque nous voulons savoir si une question est signalée comme telle bien avant la syllabe finale de l'énoncé. Il est alors souhaitable d'examiner quel groupe de variables permet le mieux de séparer les fonctions à l'aide de cette simulation. Il s'avère que les deux groupes atteignent un taux similaire en ce qui concerne la séparation où les mesures non-finales placent 72 % des exemples dans la bonne catégorie, alors que les mesures finales jugent 73 % d'entre eux correctement. Ces résultats nous permettent de conclure qu'aussi bien les mesures finales que non-finales servent à différencier les deux fonctions entre elles, ce qui illustre que les deux groupes de mesures ont une importance à peu près égale à propos de la séparation.

D'après nos résultats, il faut alors reconsidérer la description des deux exemples et leurs contours stylisés proposés par Delattre (1966 ; 1967). Avec les figures dans le chapitre suivant nous tentons de décrire la manière dont les locuteurs et les locutrices se servent des variables de la f_0 en signalant une certaine fonction. Il faut insister sur le fait que les deux groupes de variables (finales vs. non-finales) atteignent un niveau similaire de distinction entre les deux fonctions, indiquant ainsi que l'attention qui a précédemment été centrée sur la syllabe finale doit également inclure ce qui la précède dans l'énoncé.

Il faut noter que le poids de chaque groupe de variable est différent puisque l'ajout des variables de la f0 aux autres variables de l'intonation améliore le choix entre les deux fonctions, alors que l'ajout des variables intonatives aux variables de la f0 ne démontre pas ce même effet. Cela veut dire que la f0 est le facteur le plus informatif quand elle est étudiée seule, et que les autres dimensions de l'intonation sont redondantes dans cet aspect. Par cela, nous ne voulons pas dire que cela est vraiment la manière dont les interlocuteurs interprètent ces informations, puisqu'il faudrait faire une étude de perception afin de déterminer de plus près si les interlocuteurs se servent de ces signaux. Nous pouvons néanmoins affirmer que les locuteurs émettent certaines informations intonatives en produisant un certain type de montée qui sont également disponibles aux interlocuteurs, même si nous ne savons pas si et comment ils s'en servent.

5.6. CONCLUSIONS DES RÉSULTATS ET TRACES DE SYNTHÈSE DE LA F0

Avant de résumer les contours d'après leurs tendances générales en ce qui concerne le paramètre de la fréquence fondamentale, nous allons brièvement résumer les résultats en illustrant les types prototypiques au sujet de toutes les caractéristiques intonatives étudiées en revenant à nos prédictions. Nous constatons d'abord que les deux fonctions se distinguent aussi bien pour les mesures finales que non-finales. Le maximum est inférieur pour les continuations (pour les locuteurs), ce qui confirme les propositions de Delattre (1966) et Rossi et al. (1981). Les fonctions se distinguent plus fortement pour les mesures avant le début de la montée, étant donné que les questions commencent également déjà à un niveau supérieur aussi bien au début de la voyelle finale que dans le syntagme intonatif précédant celle-ci. Ces résultats s'opposent à nos prévisions d'après Delattre (1966) qui a proposé un début au même niveau pour les deux contours. Les questions confirment cependant les prédictions liées aux turn-yielding cues observés pour l'anglais par Duncan (1972) et Gravano et Hirschberg (2011) en présentant une intensité inférieure (cependant que pour les voyelles ouvertes) ainsi qu'une durée inférieure pour la voyelle finale. La dernière mesure qui permet de les distinguer est le débit de la parole qui est supérieur pour les questions, conformément aux résultats de Gravano et Hirschberg (2011) et Van Heuven et Van Zanten (2005). Le seul paramètre qui ne semble pas servir à une séparation intonative entre les deux contours est la forme de la montée, qui s'avère aussi bien convexe que concave pour les deux types. Il s'avère alors que plusieurs mesures servent à établir la distinction entre les deux contours, alors qu'il faut pourtant noter que ces distinctions ne sont pas nettes (cf. la discussion au Chapitre 6).

Les variables de la fréquence fondamentale ont été les meilleurs prédicteurs dans les simulations décrites dans le chapitre précédent. Étant donné que nous avons observé des variables de la f_0 qui servent à séparer les deux fonctions entre elles (le maximum standardisé, le minimum standardisé, la médiane), ces indices servent ici à illustrer les tendances générales (basées sur les médianes des diagrammes en boîte) que les locutrices et les locuteurs emploient. Cette partie du chapitre sur les résultats résume les stratégies des deux sexes, présentées sous la forme des traces de la f_0 stylisées. La ligne horizontale indique la médiane globale d'un extrait de dix minutes qui sert à standardiser les mesures de la f_0 , alors que la ligne verticale indique le début de la voyelle. Étant donné que les figures sont des versions stylisées, il faut garder en tête qu'elles ne sont qu'une somme de tous les exemples, ce qui n'implique pas que tous les exemples présentent cette tendance.

Continuation (femmes)

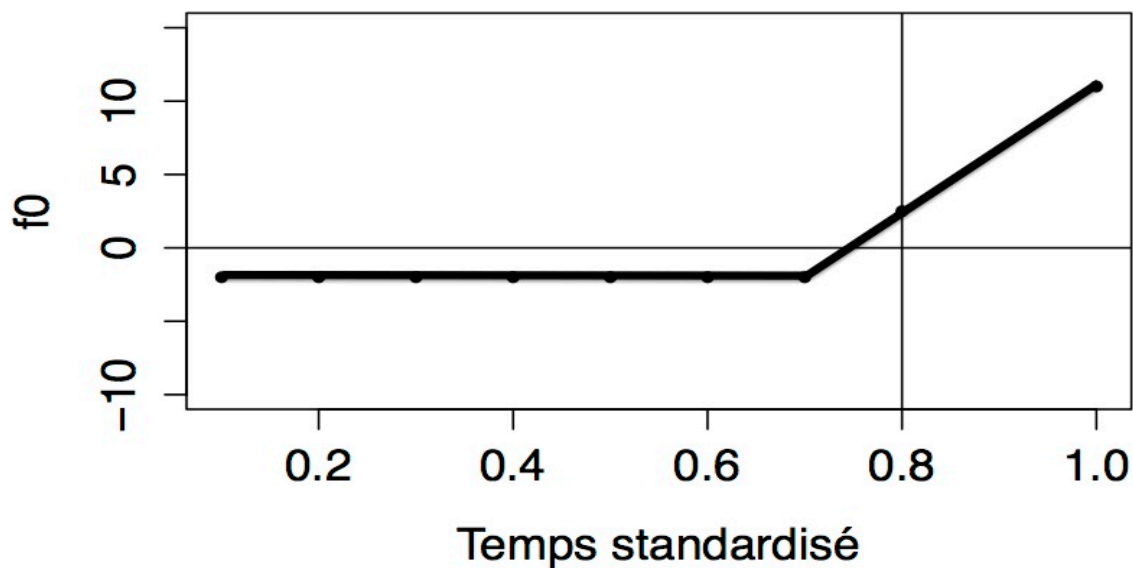


Figure 15. Cette figure illustre la tendance générale (basée sur la médiane des diagrammes en boîtes à trouver dans le reste de ce chapitre) des locutrices en produisant une continuation à l'aide de la fréquence fondamentale. La ligne horizontale indique la médiane globale (tirée d'un extrait de dix minutes) et la ligne verticale le début de la voyelle finale.

Question (femmes)

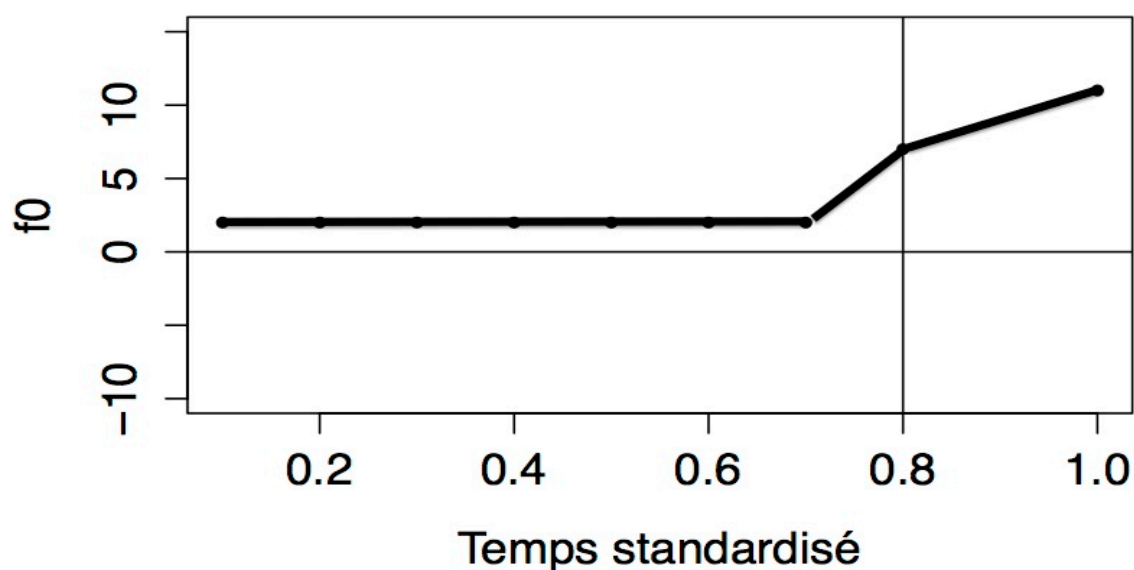


Figure 16. Cette figure illustre la tendance générale (basée sur la médiane des diagrammes en boîtes à trouver dans le reste de ce chapitre) des locutrices en produisant une question montante à l'aide de la fréquence fondamentale. La ligne horizontale indique la médiane globale (tirée d'un extrait de dix minutes) et la ligne verticale le début de la voyelle finale.

Les continuations des locutrices sont en général signalées dès le début de l'énoncé cible à l'aide d'un registre qui est inférieur à la médiane globale de la locutrice. Au moment où la voyelle finale commence, la f_0 est légèrement supérieure à la médiane et elle monte beaucoup pendant la voyelle. Quant à la question montante, les locutrices l'indiquent déjà au début de l'énoncé à l'aide d'un registre supérieur à leur médiane globale et un minimum supérieur comparé à celui de la continuation. Cela illustre bien l'effet plafond mentionné auparavant, c'est-à-dire que les locutrices ne montent pas au-dessus d'un certain niveau, peut-être en raison du fait qu'il n'est pas aisé pour elles de parler à un registre supérieur à celui atteint par le maximum.

Question (hommes)

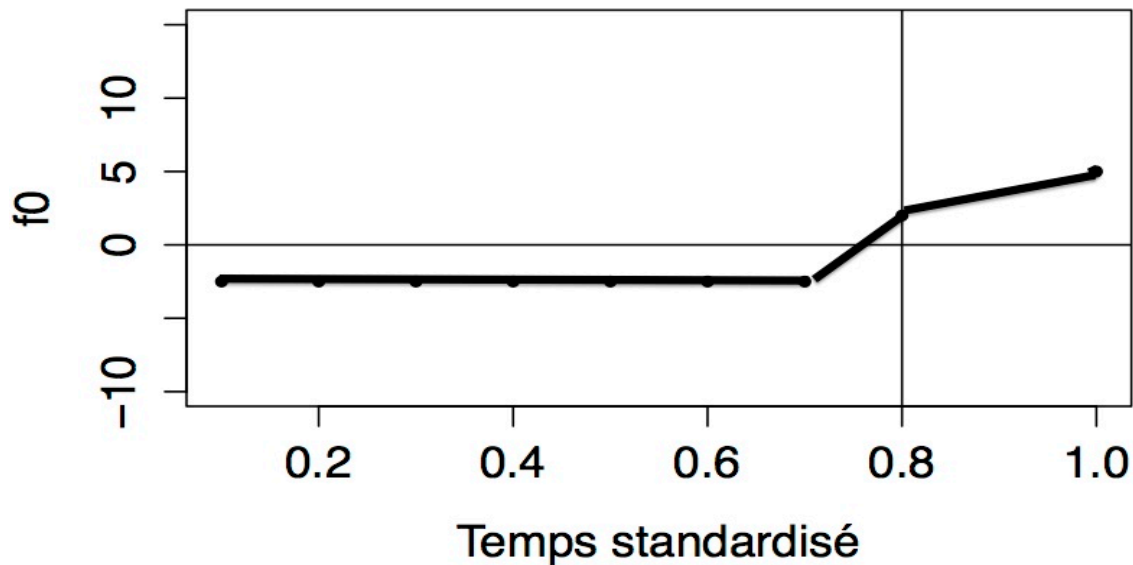


Figure 17. Cette figure illustre la tendance générale (basée sur la médiane des diagrammes en boîtes à trouver dans le reste de ce chapitre) des locuteurs en produisant une continuation à l'aide de la fréquence fondamentale. La ligne horizontale indique la médiane globale (tirée d'un extrait de dix minutes) et la ligne verticale le début de la voyelle finale.

Le registre des locuteurs au début de l'exemple de la continuation est plus inférieur à leur médiane globale que celui des questions. Ainsi les locuteurs et les locutrices signalent le registre de la même manière en accompagnant la question d'un registre supérieur. Les autres deux points de repère de la f_0 se distinguent également avec un minimum et un maximum supérieur pour les questions.

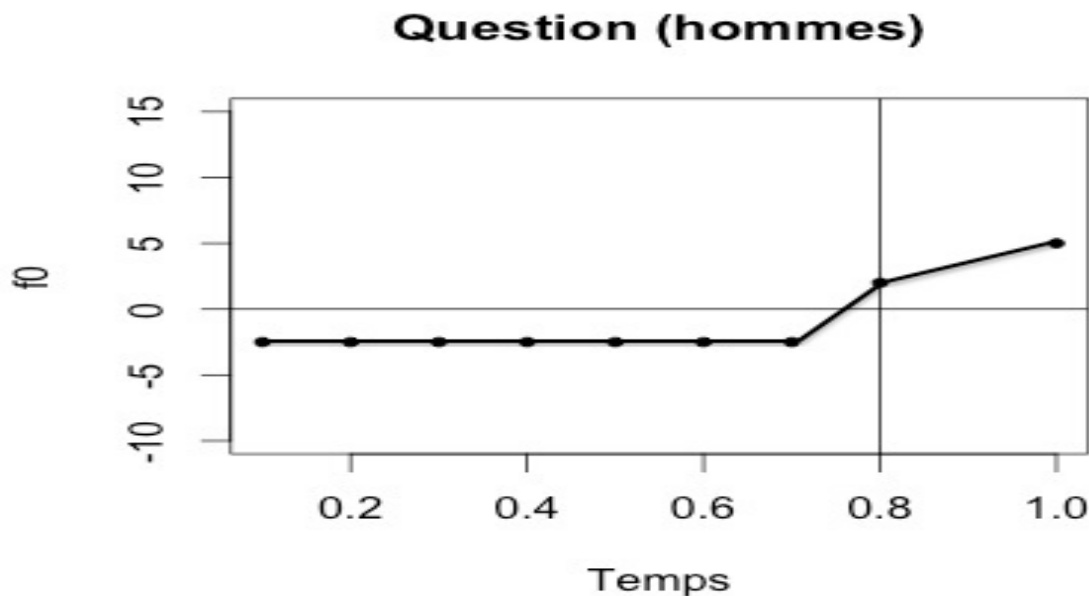


Figure 18. Cette figure illustre la tendance générale (basée sur la médiane des diagrammes en boîtes à trouver dans le reste de ce chapitre) des locuteurs en produisant une question montante à l'aide de la fréquence fondamentale. La ligne horizontale indique la médiane globale (tirée d'un extrait de dix minutes) et la ligne verticale le début de la voyelle finale.

D'après les résultats de notre travail empirique, il est intéressant de remarquer la différence du registre qui se manifeste déjà avant la fin de l'énoncé cible émis par le locuteur afin de signaler les différentes fonctions montantes. Il faut ainsi reconsidérer le fait de centrer toute l'attention sur la syllabe finale de la montée, prenant également en compte que des différences peuvent s'affirmer avant cet endroit. Ainsi la médiane de l'énoncé partiel (le registre) est la mesure la moins conservatrice et celle qui se prête le mieux pour effectuer une comparaison. De toute façon, la plupart des mesures de la f_0 proposées par Delattre (1966) et des mesures intonatives liées à l'alternance des tours de parole (Gravano et Hirschberg (2011) entre autres) apparaissent également comme importantes. Même si les fonctions se distinguent à plusieurs points, il nous ne semble pas possible de les classer comme des contours différents à cause de nombreux chevauchements des fonctions et le peu de distinction nette (voir la discussion, Ch.6.).

6. DISCUSSION

Le but de ce mémoire a été double dans le sens où nous avons voulu examiner aussi bien les différences intonatives des deux montées ainsi que leur implication dans l'alternance des tours de parole. Quant au premier objectif, il provient du fait que le statut des contours de la continuation et de la question montante est débattu dans la littérature (cf. Delattre, 1966 ; Rossi et al. 1981 ; Di Cristo, 1978 ; entre autres). Étant donné qu'ils sont souvent classés comme deux contours séparés qui se distinguent au sujet de leur maximum de la f_0 , leur forme, ainsi que par d'autres traits intonatifs, nous avons eu l'intention de vérifier ces propositions lors d'une étude empirique. La plupart des études menées autour de la séparation de ces contours ont été effectuées avant la coutume d'un emploi des corpus des conversations spontanées. Les traits distinctifs provenant de l'intonation méritent alors d'être examinés sur du matériel autre que celui qui est dirigé quant à son contenu (les phrases lues). La première partie de la discussion porte ainsi sur une séparation potentielle entre les deux contours.

La deuxième partie de cette discussion considère les conséquences que les informations intonatives peuvent avoir lors d'une alternance de tour de parole. Il faut cependant nous réserver concernant les implications de ces informations, puisque nous ne commentons que sur leur présence. En revanche, nous ne savons rien sur l'intention par le locuteur en les produisant ou sur la manière dont l'interlocuteur les perçoit. Comme il a été constaté dans le chapitre sur les résultats (Ch. 5.), certaines informations intonatives établissent la distinction entre les deux fonctions, ce qui mérite d'être mis en relief. Le système des alternances du tour est évidemment plus complexe que les informations intonatives proposées lors de nos analyses, cependant elles peuvent potentiellement ajouter de l'information supplémentaire à la syntaxe (voir. Ch. 5.4.). Il faut toutefois d'abord centrer l'attention sur la description des contours avant de décrire leurs rôles potentiels lors une alternance de tour.

En ce qui concerne le statut de la continuation et de la question montante qui ne permettent pas une distinction nette entre elles au sujet des informations intonatives, nous concluons que les manifestations des différentes fonctions proviennent d'un seul contour. Nous marchons sur les traces de Di Cristo (1978) puisque nous considérons qu'il s'agit d'un seul contour qui est réalisé de manière quelque peu différente. Toutes les informations intonatives examinées dans le chapitre précédent servent à établir une distinction significative

entre les deux fonctions, alors que ces différences ne sont pas aussi nettes afin de conclure qu'il s'agit bien des contours différents. Le chevauchement entre les boîtes dans les diagrammes diffère quant à leur degré, tandis qu'il n'y a jamais une différence aussi grande qui permet de constater que les mesures des continuations se distinguent suffisamment des questions pour en conclure que certains exemples ne sont jamais ambigus. Dû au fait qu'il y a des chevauchements au sujet de la plupart des mesures, nous considérons que les questions sont souvent réalisées avec un maximum supérieur, un minimum supérieur, etc., alors qu'il n'est pas possible d'identifier à 100 % la catégorie d'un exemple à partir des traits intonatifs.

Il nous semble aussi étrange que les locutrices et les locuteurs se servent des contours différents, une raison de plus afin de classifier les fonctions comme des versions d'un même contour. Étant donné que les locutrices se servent de certains indices (p. ex. le maximum) d'une manière différente par rapport aux locuteurs, cela n'indique pas que les deux groupes se servent des contours différents. Notre interprétation de l'absence d'une différence concernant cette mesure est que c'est un effet plafond qui se manifeste, avec un choix des locutrices de ne pas produire cette différence puisqu'elle n'est pas aisée. Notre décision de ne pas attribuer des contours différents aux locuteurs et aux locutrices va ensemble avec la proposition de Bolinger (1989) qui mentionne que les deux sexes ne se servent pas des contours différents, alors qu'il peut y avoir une variation quant à leur emploi. Il n'est alors pas surprenant que certains phénomènes intonatifs sont utilisés différemment par les locuteurs et les locutrices, sans que pour autant conclure qu'ils se servent des formes sous-jacentes de la langue d'une manière différente.

D'après nos connaissances, notre propre étude est la première étude qui adresse cette problématique à l'aide d'un corpus des données spontanées. Il nous semble cependant nécessaire d'étudier les réalisations de ce contour dans lorsqu'il est employé naturellement par les locuteurs. La proposition mise en avant par Di Cristo (1978) qui concerne la réalisation différente par le même contour lors de son placement (final et non-final) semble ainsi être confirmé par nos données. En supposant que le lieu d'apparition est la raison pour laquelle une réalisation particulière se fait du contour montant, il est possible d'expliquer la variabilité dans nos données avec un maximum et un minimum supérieur pour les questions, etc. Selon l'inventaire de Marandin et al. (2004), le contour de la continuation mérite ainsi d'être décrit avec les symboles L-H%. Comme le diacritique (%) indique une borne majeure (Gussenhoven et Jacobs, 2005) avec par exemple une durée supérieure, il nous semble que

cette annotation est possible, puisque la différence en durée n'est pas immense entre les deux catégories (continuation et question).

D'après avoir conclu que les deux fonctions sont deux versions différentes d'un même contour, la discussion peut se tourner vers le rôle que les différentes réalisations peuvent manifester au sujet de l'alternance des tours de parole. Une différence entre les deux fonctions peut être établie aussi bien au sujet des informations de la f_0 qu'avec d'autres informations intonatives (intensité, durée, débit de parole). La séparation des questions par rapport aux continuations reflète les rôles différents que ces deux fonctions jouent dans la structuration d'une conversation. La question, qui indique un abandon de tour d'après notre classification, comprend des turn-yielding cues proposés par Gravano et Hirschberg (2011), Cutler et Pearson (1986), Duncan (1972). Étant donné qu'elle est accompagnée d'une intensité inférieure à ce qui précède, d'une durée inférieure et un débit de parole supérieur aux continuations, l'on peut constater que cette fonction se comporte comme un signal d'abandon de tour exemplaire.

Il ne nous semble pas exister une étude précédente à la nôtre qui examine la différence entre la continuation et la question en français à l'aide d'un corpus en ce qui concerne leurs rôles lors de l'alternance des tours de parole. C'est la raison pour laquelle il est intéressant de noter que les locuteurs émettent des indices liés à l'abandon du tour (turn-yielding) lorsqu'ils produisent des questions (qui signalent l'abandon du tour avec bien d'autres indices – la syntaxe, l'autorité épistémique). Nous avons également noté que les questions syntaxiquement complètes sont un groupe séparé des continuations syntaxiquement complètes, ce qui indique que dans les cas où la syntaxe ne peut pas être considérée comme l'indice de garde de la parole, l'intonation peut toujours apporter de l'information. Les différences intonatives ne peuvent pas être considérées employées par les interlocuteurs lorsqu'ils prennent une décision sur une intervention potentielle à un point particulier de la conversation. Afin de pouvoir dire quelque chose concernant cette utilité il faut faire une étude perceptive où l'intonation est le seul trait distinctif entre les deux fonctions. Néanmoins, les résultats, que nous avons obtenus au sujet des deux fonctions, nous semblent importants à considérer puisqu'ils semblent révéler un système sous-jacent qui n'est peut-être pas spécifique à la langue française en ce qui concerne l'alternance des tours de parole.

La dernière découverte sur laquelle nous allons discuter dans ce chapitre est la complétude syntaxique et sa relation à la durée de la voyelle finale. Gravano et Hirschberg

(2011) ont trouvé une différence en ce qui concerne la durée du mot en position non-finale et finale d'énoncé. D'après nos résultats, cette différence peut potentiellement être confondue avec la complétude syntaxique. Nous avons observé que les continuations ne diffèrent pas beaucoup par rapport aux questions (20 ms, voir Ch. 5.2.4.), alors que les continuations incomplètes se distinguent des continuations complètes avec une différence légèrement supérieure (30 ms, voir Ch. 5.4.). Nous interprétons cette différence en durée comme un moment de planification de ce qui suit, étant donné que les continuations incomplètes demandent potentiellement une organisation de l'information linguistique dans ce qui suit. Même si nous ne possédons pas de l'information sur la complétude syntaxique des données de Gravano et Hirschberg (2011), il nous semble que la position du mot (turn-medial position) implique que ce qui suit est apparenté à ce qui le précède.

Finalement nous voulons récapituler sur l'importance de ne pas classifier les deux fonctions comme des contours différents, puisqu'elles ne démontrent pas des distinctions nettes au sujet des mesures intonatives différentes. Leur rôle dans le système de l'alternance des tours de parole n'est pas résolu d'après cette étude, alors qu'une indication sur les manifestations divergentes de l'intonation est démontrée. Notre étude n'a pas du tout pris en compte le rôle de la multimodalité d'une manière systématique, un facteur additionnel qui doit être incorporé lors des recherches ultérieures. La relation entre les gestes et les fonctions est quelque chose qui mérite d'être examinée, alors que la portée de cette étude ne l'a pas permis, mais c'est une possibilité ouverte pour de la recherche suivante.

RÉFÉRENCES BIBLIOGRAPHIQUES

- ANDRUSKI, J. E., COSTELLO, J. (2004). « Using polynomial equations to model pitch contour shape in lexical tones : an example from Green Mong, *Journal of the International Phonetic Association*, Vol. 34, No. 2, pp. 125-140.
- AVANZI, M., DELAIS-ROUSSARIE, E. (2011). « Introduction : Regards croisés sur la prosodie du français – des données à la modélisation », *Journal of French Language Studies*, Vol. 21, No. 1, pp. 1-12.
- BERTRAND, R., PORTES, C., SABIO, F. (2007). « Distribution syntaxique, discursive et interactionnelle des contours intonatifs du français dans un corpus de conversation », *Travaux neuchâtelois de linguistique*, 47, pp. 59-77.
- BOERSMA, P., & WEENINK, D. (2013). Praat: doing phonetics by computer (Version 5.3.51). <http://www.praat.org>
- BOLINGER, D. L. M. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford University Press.
- BRUGMAN, H., RUSSEL, A. (2004). Annotating Multimedia/ Multi-modal resources with ELAN. In Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation. <http://tla.mpi.nl/tools/tla-tools/elan/>
- CANGEMI, F., D'IMPERIO, M. (2011). « Local speech rate differences between questions and statements in Italian », *International Congress of Phonetic Sciences XVII, Hong Kong*, pp. 392-395.
- CASPER, J. (1998). « Who's next? The Melodic Marking of Question vs. Continuation in Dutch », *Language and Speech*, Vol. 41, No. 3-4, pp. 375-398.
- CURL, T. S. (2004). The relationship of phonetic structure and sequence organization. *Sound Patterns in Interaction: Cross-Linguistic Studies from Conversation*, 62, 273.
- CUTLER, A., & PEARSON, M. (1986). On the analysis of prosodic turn-taking cues. *Intonation in discourse*, 139-156.
- DELATTRE, P. (1966). « Les dix intonations de base du français », *French Review*, Vol. 40, No. 1, pp. 1-14.
- DELATTRE, P. (1967). « La Nuance de sens par l'intonation », *French Review*, Vol. 41, No. 3, pp. 326-339.
- DI CRISTO, A. (1976). « Des indices prosodiques aux traits perceptuels. Application d'un modèle d'analyse à l'étude du vocatif », *Travaux de l'Institut de Phonétique d'Aix*, 3, p.213-358.
- DI CRISTO, A. (1978). De la Microprosodie à l'Intonosyntaxe. Thèse de Doctorat d'Etat (Univ. de Provence), Publications de l'Université de Provence, 2 vol. 1985.
- DUNCAN, S. (1972). « Some signals and rules for taking speaking turns in conversations », *Journal of personality and social psychology*, Vol. 23, No. 2, 283-292.
- FIELD, A., MILES, J., FIELD, Z. (2012). *Discovering Statistics using R*, Sage Publications, London
- GELUYKENS, R. (1987). « Intonation and speech act type : an experimental approach to rising intonation in declaratives », *Journal of pragmatics*, Vol. 11, pp. 483-494.
- GRABE, E., KOCHANSKI, G., COLEMAN, J. (2007). « Connecting intonation labels to mathematical descriptions of fundamental frequency », *Language and speech*, Vol. 50, No. 3, pp. 281-310.
- GRAVANO, A., HIRSCHBERG, J. (2011). « Turn-taking cues in task-oriented dialogue », *Computer Speech & Language*, Vol. 25, No. 3, pp. 601-634.

- GRUNDSTROM, A. (1973). « L'intonation des questions en français standard », in *Interrogation et intonation, Studia Phonetica*, 8, pp. 19-49.
- GUSSENHOVEN, C., JACOBS, H. (2005). *Understanding Phonology*, Hodder Education, London
- HELDNER, M., EDLUND, J. (2010). « Pauses, gaps and overlaps in conversations », *Journal of Phonetics*, Vol. 38, No. 4, pp. 555-568.
- HERITAGE, J. (2012). « Epistemics in action: Action formation and territories of knowledge », *Research on Language & Social Interaction*, Vol. 45, No. 1, pp. 1-29.
- JUN, S. A., FOUGERON, C. (2000). « A Phonological model of French intonation ». In A. Botinis (éd.) *Intonation : Analysis, Modeling and Technology*. Dordrecht, Kluwer Academic Publishers, pp. 209-242.
- LACHERET, A. (2013). « Les développements récents en prosodie du français parlé: approches fonctionnelles », *Le Français moderne*, Vol. 81, No. 2, pp. 230-241.
- LACHERET-DUJOUR, A. (2011). « Petit parcours prosodique de Grenoble à Aix-en-Provence : pionniers, dissidents et fédérateurs », *Un siècle de phonétique expérimentale : Histoire et développement de Théodore Rosset à John Ohala*, (Abry Boë et al éd.), ENS Editions, Lyon, pp. 277-297.
- LADEFOGED, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.
- LÉON, P. (1996). *Phonétisme et prononciations du Français avec des travaux pratiques d'application et leurs corrigés*. Editions Nathan, Paris
- LÉON, P., BHATT, P. (1987). « Structures prosodiques du questionnement radiophonique ». *Études de linguistique appliquée*, Vol. 66, pp. 88-105.
- MARANDIN, J.-M., BEYSSADE, C., DELAIS-ROUSSARIE, E., RIALLAND, A., DE FORNEL, M. (2004). « The meaning of final contours in French », Ms available at <http://www.llf.cnrs/Gens/Marandin>.
- MARTIN, P. (2004). « Intonation de la phrase dans les langues romanes: l'exception du français », *Langue française*, Vol. 1, pp. 36-55.
- MARTIN, P. (2007). « Structure prosodique, structure de contrastes ». *TRANEL (Travaux neuchâtelois de linguistique)*, 103.
- MERTENS, P. (2009). « Prosodie, syntaxe, discours: autour d'une approche prédictive ». *Interface Discours & Prosodie (IDP)*, pp. 19-32.
- OGDEN, R. (2006). « Phonetics and social action in agreements and disagreements », *Journal of Pragmatics*, Vol. 38, No. 10, pp. 1752-1775.
- OHALA, J. J. (1984). « An ethological perspective on common cross-language utilization of F0 of voice ». *Phonetica*, Vol. 41, No. 1, pp. 1-16.
- PICARD, R. R., COOK, R. D. (1984). « Cross-validation of regression models ». *Journal of the American Statistical Association*, Vol. 79, No. 387, pp. 575-583.
- PORTES, C., BERTRAND, R. (2007a). « Some cues about the interactional value of 'continuation' contour in French », *Actes, Discours et Prosodie comme Interface Complexe (IDP)*, pp. 1-14.
- PORTES, C., BERTRAND, R. (2007b). « Contribution to a grammar of intonation in French. Form and function of three rising patterns », *Nouveaux cahiers de linguistique française*, No. 28, pp. 155-162.
- POST, B. (2002). « French tonal structures ». *Speech Prosody 2002, International Conference*.
- R CORE TEAM. (2013). « R: A Language and Environment for Statistical Computing », R

- Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>
- ROSSI, M. (1978). « Interactions of intensity glides and frequency glissandos », *Language and speech*, Vol. 21, No. 4, pp. 384-396.
- ROSSI, M. (1985). « L'intonation et l'organisation de l'énoncé ». *Phonetica*, Vol. 42, No. 2-3, pp. 135-153.
- ROSSI, M., DI CRISTO, A., HIRST, D., MARTIN, P., & NISHINUMA, Y. (1981). *L'intonation: de l'acoustique à la sémantique*.
- SACKS, H., SCHEGLOFF, E. A., & JEFFERSON, G. (1974). « A simplest systematics for the organization of turn-taking for conversation », *Language*, pp. 696-735.
- SELTING, M. (2000). « The construction of units in conversational talk », *Language in society*, Vol. 29, No. 4, pp. 477-517.
- SMITH, C. L. (2002). « Prosodic Finality and Sentence Type in French », Vol. 45, No. 2, pp. 141-178.
- STEPHENS, J., BEATTIE, G. (1986). « On judging the ends of speaker turns in conversation ». *Journal of Language and Social Psychology*, Vol. 5, No. 2, pp. 119-134.
- STIVERS, T., ROSSANO, F. (2010). « Mobilizing Response », *Research on Language and Social Interaction*, Vol. 43, No. 1, pp. 3-31.
- TORREIRA, F. (2007). « Tonal realization of syllabic affiliation in Spanish », *ICPhS XVI, Saarbrücken*, pp. 6-10.
- TORREIRA, F., ADDA-DECKER, M., & ERNESTUS, M. (2010). « The Nijmegen corpus of casual French », *Speech Communication*, Vol. 52, No. 3, pp. 201-212.
- TOUATI, P. (1987). *Structures prosodiques du suédois et du français: profils temporels et configurations tonales*. Lund University Press.
- VAN HEUVEN, V. J., VAN ZANTEN, E. (2005). « Speech rate as a secondary prosodic characteristic of polarity questions in three languages, *Speech Communication*, Vol. 47, pp. 87-99.