



HAL
open science

Projection de la valence émotionnelle par la prosodie : aspects phonétiques, phonologiques et perceptifs dans les préfaces aux annonces de nouvelles

Emilie Marty

► To cite this version:

Emilie Marty. Projection de la valence émotionnelle par la prosodie : aspects phonétiques, phonologiques et perceptifs dans les préfaces aux annonces de nouvelles. Linguistique. Aix Marseille Université, 2023. Français. NNT : . tel-04501618

HAL Id: tel-04501618

<https://theses.hal.science/tel-04501618>

Submitted on 12 Mar 2024

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

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

Copyright

.....

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 15 décembre 2023 par

Emilie MARTY

Projection de la valence émotionnelle par la prosodie : aspects phonétiques, phonologiques et perceptifs dans les préfaces aux annonces de nouvelles

Discipline

Sciences du langage

École doctorale

ED 356 – Cognition, Langage, Éducation

Laboratoire/Partenaires de recherche

Laboratoire Parole et Langage

– CNRS/UMR 7309



Composition du jury

Pilar PRIETO

Rapporteuse

Universitat Pompeu Fabra

Marc SWERTS

Rapporteur

Tilburg University

Štefan BEŇUŠ

Examineur

Constantine the Philosopher University,

Slovak Academy of Sciences.

Roxane BERTRAND

Présidente du jury

Aix-Marseille Université

James GERMAN

Directeur de thèse

Aix-Marseille Université

Caterina PETRONE

Co-directrice de thèse

Aix-Marseille Université

Affidavit

Je soussignée, Emilie Marty, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de James German et Caterina Petrone, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Aix-en-Provence, le 19 octobre 2023



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Affidavit

I, undersigned, Emilie Marty, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of James German and Caterina Petrone, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Place Aix-en-Provence, date 19 octobre 2023



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Liste de publications et/ou brevets et participation aux conférences

1) Liste des publications réalisées dans le cadre du projet de thèse :

Marty, É., Bertrand, R., Petrone, C., & German, J. S. (2022). Prosodic Correlates of Discourse Structure and Emotion in Discourse Markers that Preface Announcements of News. In *Proceedings. 12th International Conference on Speech Prosody 2022. 23-27 May 2022, Lisbon, Portugal*, 605-609.

2) Participation aux conférences et écoles d'été au cours de la période de thèse :

1. Marty, E., Petrone, C., & German, J. (2019, December). On the impact of emotional prosody in announcements: The case of the pre-sequence. In *Journée Annuelle des Doctorants*. (hal-03383155)
2. Marty, É., Bertrand, R., Petrone, C., & German, J. S. (2022). Prosodic Correlates of Discourse Structure and Emotion in Discourse Markers that Preface Announcements of News. *12th International Conference on Speech Prosody 2022. 23-27 May 2022, Lisbon, Portugal*.
3. ILCB summer school (3rd edition): Intermediate courses in Language and Emotion, Early Development of Communication, Roles of neural oscillations in cognitive processes, Mixed Models and Linear Regression, Basic Course in Machine Learning. 31 August – 4 September 2020. Held at Centre International de Recherche en Mathématiques (CIRM), Luminy, Marseille (France).
4. ILCB summer school (4th edition): Advanced courses in Signal Processing, Intermediate courses in Linear Regression, Mixed Models. 30 August – 3 September 2021. Held at Centre International de Recherche en Mathématiques (CIRM), Luminy, Marseille (France).

Résumé

Cette thèse examine le rôle projectif de la prosodie affective en français. Nous nous sommes concentrés sur le cas particulier des préfaces aux annonces de nouvelles, lieu privilégié pour la projection d'indices en interaction. La complexité de la notion d'affect, la multitude de paramètres acoustiques et de moyens de les évaluer, ainsi que les positions théoriques traditionnelles de l'intonation rendent difficile l'appréhension que l'on peut avoir de cette projection affective au travers de la prosodie. Cette thèse soutient qu'il existe bel et bien des indices de projection affective, et que ceux-ci peuvent se trouver non seulement dans les caractéristiques acoustiques continues de la voix, mais aussi dans la phonologie de l'intonation, de façon plus catégorielle. De plus, cette thèse soutient que ces caractéristiques, produites par le locuteur lorsqu'il s'apprête à annoncer une nouvelle, sont perçues et utilisées par les interlocuteurs pour se préparer à la valence affective, négative, positive ou neutre, de l'annonce.

Les résultats corroborent l'existence d'un tel phénomène. En production, la prosodie produite avant l'annonce d'une nouvelle révèle des différences, principalement au niveau de la fréquence fondamentale et de la qualité de la voix, selon la valence de la nouvelle. De même, l'intonation semble indiquer que l'affect pourrait être porté au travers de catégories phonologiques, bien que nous n'ayons que peu de différences à ce niveau. De plus, des variations phonétiques au niveau de l'intonation sont aussi présentes et varient en fonction de la valence affective de la nouvelle. Enfin, cette thèse révèle que les auditeurs ont une réaction affective différente en fonction des valences affectives portées par la préface, sur la base uniquement de changements dans la phonétique et la phonologie de l'intonation.

Mots clés : prosodie affective, intonation, indices de projection, anticipation, annonces de nouvelles

Abstract

This thesis examines the projective role of affective prosody in French. We focus on the case of prefaces to news announcements, a privileged site for the projection of interactional cues. The complexity of the notion of affect, the multitude of acoustic parameters and ways of assessing them, and the traditional theoretical positions of intonation studies make it difficult to apprehend this affective projection through prosody. This thesis argues that cues to affective projection do exist, and that these can be found not only in the continuous acoustic features of the voice, but also in the phonology of intonation, in a more categorical way. Furthermore, this thesis argues that these features, produced by the speaker when about to announce news, are perceived and used by interlocutors to prepare for the affective valence, whether it be negative, positive, or neutral, of the announcement.

The results corroborate the existence of such a phenomenon. In production, the prosody produced before a news announcement reveals differences, mainly in fundamental frequency and voice quality, depending on the valence of the news. Similarly, intonation seems to indicate that affect could be carried through phonological categories, although we have few differences at this level. Moreover, phonetic variations in intonation are also present, and vary according to the affective valence of the news item. Finally, this thesis reveals that listeners have different affective reactions based on affective valence carried solely on changes in the phonetics and phonology of intonation.

Keywords: affective prosody, intonation, projection cues, anticipation, news announcements

Remerciements

Je tiens tout d'abord à remercier mes directeurs Caterina Petrone et James German. Merci à Caterina pour le temps qu'elle a pris avec moi pour m'aider à comprendre ce sujet dans lequel je me suis lancée, pour m'avoir conseillée et apporté des méthodes de travail qui me serviront longtemps. Merci pour toutes ses relectures attentives et sa disponibilité. Merci à James pour ses retours également, sa disponibilité, et son approche du travail et de la santé. Merci d'avoir cru en mes capacités et d'avoir accepté cette co-direction il y a 4 ans. Merci, à tous les deux pour votre patience, votre gentillesse, votre compréhension et votre suivi. Merci d'avoir cru en mon travail et surtout, d'avoir été des modèles en tant que chercheurs mais aussi en tant que personnes, par votre rigueur de travail et votre patience.

Je remercie les membres du jury, Marc Swerts, Pilar Prieto, Roxane Bertrand, et Štefan Beňuš d'avoir accepté de lire et d'évaluer ce travail.

Je remercie à nouveau Roxane, qui fut ma directrice de mémoire dans un premier temps. Merci de m'avoir suivie dans ma thèse, conseillée sur ce qui touchait son domaine mais aussi sur l'approche qu'il faut avoir du travail. Merci de m'avoir, la première, donné goût à la recherche, merci pour ta gentillesse et ton soutien.

Merci à Cristel Portes et Štefan Beňuš, d'avoir accepté de monitorer le déroulé de la thèse en tant que comité de suivi. Merci pour vos conseils et votre écoute. Merci Cristel pour ces rendez-vous sur l'annotation prosodique qui m'ont permis de progresser, merci d'avoir accepté la double annotation. Merci d'avoir été, lors du master et de la licence, une professeure passionnée et passionante.

Merci aux membres des équipes POP, SystUs, S2S, et du Groupe Transversal Prosodie pour l'opportunité qui m'a été fournie de présenter mon travail, en discuter, et obtenir des retours toujours constructifs. Merci particulièrement à toutes les personnes qui, à un moment ou un autre, m'ont apporté leur aide pour comprendre un sujet, choisir une voie, et prendre confiance en moi. Merci notamment à Alain, Amandine, Brigitte, Cristel, Gilles, Maud et Muriel.

Merci à nos merveilleuses documentalistes, Joëlle et Claudia. Merci pour votre bonne humeur, et votre patience avec une doctorante qui rend toujours les livres un peu en retard !

Merci à Rémi et Axel, pour ces verres pris en terrasse avec vous au début de ma thèse, qui m'ont permis de me sentir à l'aise et à ma place dans cet open space tout nouveau. Merci pour vos conseils, votre aide sur la compréhension des statistiques et de la création de scripts Praat. Je suis ravie que nous ayons partagé un directeur de thèse pendant quelques années. Merci à mes collègues, Airelle, Auriane, Axel, Bingie, Clara, Clothaire, Eiber, Estelle, Estelle, Francesca, Giulia, Giusy, Joana, Léna, Lola, Lydia, Leela, Marine, Marie, Marie, Mary, Outhmane, Rémi, Salomé, Sneža, et Lydia. Merci Estelle, ma collègue devenue mon amie. Pour avoir relu les mails d'une grande stressée sans jamais râler, pour ta présence, nos fous rires et nos discussions. Merci de m'avoir accompagnée au yoga.

Merci Airelle, mon amie devenue ma collègue. Ma partenaire de thèse sans qui je ne suis pas sûre que je serais arrivée au bout de ce travail. Merci pour toutes ces heures zoomées et en présentiel, partagées à travailler, rire et pleurer. Merci de m'avoir accueillie les bras ouverts dans ton monde de musique. Ta présence et ton amitié, à eux seuls, ont illuminé ce parcours de quatre ans.

Merci au Coco Bohème, dont le souvenir tendre aura accompagné ma philosophie de vie et de travail. Merci à ce lieu au travers duquel j'ai rencontré la plupart de mes amis dans la région. Merci à ma deuxième maison, mon chez moi à jamais. Merci Ben, d'avoir créé ce beau lieu, de l'avoir tenu pendant plusieurs années et d'avoir, après sa fermeture, ponctué mes années de petits cafés rares et épars, mais toujours bienvenus. Merci à l'Antre de Calliopée, au Sweet Panda, et à l'Orienthé, de m'avoir fourni de petits coins cosy pour, parfois travailler, parfois me détendre, toujours avec un bon petit thé parfumé.

Merci à mes amis, qui ont été là tout au long de cette thèse ou que j'ai découvert plus récemment : Cédric, Pauline, Apolline, Made, Lenn, Cindy, Alexandre, Sara, Georgia, Nikita, Isa, Daniel, Maël, Henry, Olivier, Léa... Merci à Constance, mon amie d'enfance. Merci à Elliott pour toutes ces virées en ville, ces cafés, ces sessions boulot chez toi, le partage de notre passion crochet et tricot, ton amitié. Merci à Dr. Aloïs, pour les multiples sessions boulot et papotage en ville. Partager avec toi les émotions par lesquelles nous ont fait passer nos thèses respectives m'aura permis de relativiser un peu sur celles-ci. Merci aux parents de Nicolas, Mihaela et Patrick, de m'avoir accueillie chez eux à bras ouverts pendant le confinement au cours de ma thèse.

Merci à ma meilleure amie Eva, Vava, mon Castor d'amûr, et à mes deux grands amis, Alex et Théo. Mes limougeauds adorés qui ont fait de mes années lycée un régal, puis de mes Nouvel Ans LE moment de l'année que j'attend toujours avec impatience. Soumeix 4 ever. Merci d'être là depuis toutes ces années.

Merci à ma famille. Papa, Maman, Marion, et Barbara. Merci à tous les 4 d'avoir été présents pour moi au cours de ces dernières années et des précédentes. De m'avoir coucounée à souhait, écoutée, et de m'avoir conseillée chaque fois que j'en avais besoin. Merci tonton Fabi pour tes conseils et ton soutien. Merci Hadrien. Merci Anita pour ton écoute et le partage de tes passions. Merci Michel. Merci tatie Christiane, merci tonton Gilles. Merci Loïc. Merci Laure, ma cousine adorée et merci à ta fille Louna d'avoir embelli cette année par ses radieux sourires de nourrisson. Merci à mes grands-parents, ceux qui sont toujours là et ceux qui sont partis. Merci Papi Raymond, Mamie Odile, Papi René, et Mamie Coco. Vos vies ont toujours été une inspiration pour moi, votre amour, une force. Je vous aime du fond du cœur. Merci Pandala.

Merci, enfin, à mon compagnon Nicolas. Pour chaque instant de ces quelques années. Merci d'avoir été présent pour moi, de m'avoir aidée à traverser les passages compliqués. De t'être assuré que je n'oubliais pas de manger en ces derniers mois de rédaction. Merci pour ton calme, ta joie, et ton amour.

« Le téléphone sonne. [...] « Allô ?... Quoi ? » C'est D. qui a dormi à côté. [...] « Quelles nouvelles ? » Silence. On parle de l'autre côté de Paris. J'essaye d'arracher le téléphone, c'est dur, c'est impossible. « Et alors ? Des camarades ? » D. lâche le téléphone et il me dit : « Ce sont des camarades de Robert qui sont arrivés au Gaumont. » Elle hurle : « Ce n'est pas vrai. » D. a repris l'appareil. « Et Robert ? » [...] D. se tourne vers elle : « Ils l'ont quitté il y a deux jours, il était vivant. » Elle n'essaye plus d'arracher le téléphone. Elle est par terre, tombée. Quelque chose a crevé avec les mots disant qu'il était vivant il y a deux jours. »

La douleur, Marguerite Duras

A mes grands-parents.

Table des matières

Affidavit	2
Affidavit	3
Liste de publications et/ou brevets et participation aux conférences	4
Résumé	5
Abstract	6
Remerciements	7
Table des matières	15
Introduction	21
Chapitre 1 : Etat de l'art	25
1. La projection dans l'interaction	27
1.1. Les indices de projection	27
1.1. Dispositifs de structuration du discours	33
1.2. Préfaces et marqueurs discursifs	42
1.3. Projection d'indices sur la valeur affective d'une séquence	49
2. Prosodie affective : phonétique, phonologie, ou les deux ?	53
2.1. Sens linguistique <i>versus</i> sens paralinguistique	53
2.2. Distinction entre phonétique et phonologie	56
2.3. Théorie AM et Phonologie Intonationnelle	61
2.4. Annotation phonologique de l'intonation du français	65
2.5. Dualité de structure	69
3. Corrélat de la valence affective	73
3.1. Définir les types d'affects	73
3.2. Enregistrer la voix affective : quel corpus choisir ?	80
3.3. Quels indices de projection attendre ? Etat de la recherche sur la parole expressive	86
3.4. Projection de structure et projection affective	94
Chapitre 2 – Etude de production : indices de projection phonétiques de la valence affective	101
	15

1. Problématique	103
1.1. Séquences interactionnelles	104
1.2. Paramètres acoustiques	105
1.3. Hypothèses	107
2. Enregistrement du corpus	109
2.1. Matériel	109
2.1.1. Phase de construction 1	109
2.1.2. Phase de construction 2	111
2.1. Participants	114
2.2. Procédure	115
2.2.1. Enregistrements	115
2.2.2. Mesures phonétiques	116
2.2.3. Tests statistiques	120
3. Résultats	125
3.1. Mesures globales de la f0	125
3.1.1. Au sein de la préface	125
3.1.2. Au sein des marqueurs discursifs « voilà donc »	129
3.1.3. Au sein de la séquence d'annonce de la nouvelle	132
3.1. Prise en compte du temps pour la f0	135
3.1.1. Au début de la préface	135
3.1.2. Au sein des marqueurs discursifs	138
3.2. Mesures des données temporelles	140
3.2.1. Au sein de la préface	140
3.2.2. Au sein des marqueurs discursifs « voilà donc »	142
3.2.3. Au sein des séquences d'annonces	143
3.3. Caractéristiques spectrales	145
3.3.1. Au sein de la préface	145
3.3.2. Au sein des marqueurs discursifs « voilà donc »	155
4. Discussion	164
Chapitre 3 – Etude phonologique de la parole en annonce de nouvelle	175
1. Problématique	177
1.1. Les configurations de fin d'IP en français	178

1.2.	Description phonologique de « voilà » et « donc »	181
1.3.	Hypothèses	185
2.	Méthodologie	187
2.1.	Annotation phonologique	187
2.2.	Accord inter-annotateurs	188
2.3.	Critères pour l’alignement des cibles tonales	191
2.3.1.	Critères pour la préface	191
2.3.2.	Critères pour « voilà donc »	193
2.4.	Analyses statistiques	197
3.	Résultats	199
3.1.	Accord inter-annotateur	199
3.1.1.	Préface	199
3.1.2.	« voilà donc »	200
3.2.	Analyses des patrons phonologiques	202
3.2.1.	Préface	202
3.2.2.	« voilà donc »	202
3.2.3.	Valeur mélodique et alignement	205
3.2.4.	Préface	206
3.2.5.	« voilà donc »	210
4.	Discussion	222
	Chapitre 4 – Perception des préfaces aux annonces de nouvelles	231
1.	Problématique	233
1.1.	Notion de feedbacks	235
1.2.	Perception des affects : facteurs de variabilité	237
1.3.	Hypothèses	241
2.	Tâche de perception des indices de projection affective	244
2.1.	Matériel	244
2.1.1.	Phase de construction 1	244
2.1.2.	Phase de construction 2	246
2.1.	Participants	248
2.2.	Procédure	249
2.3.	Analyses statistiques	251

3. Résultats	255
3.1. Tâche 1 : Positif – Négatif	256
3.2. Tâche 2 : Négatif – Neutre	260
3.3. Tâche 3 : Positif – Neutre	264
3.4. Observation de l’empathie à travers les trois tâches	268
4. Discussion	272
Discussion générale	283
Bibliographie	297
Table des figures	344
Liste des tableaux	350
Annexes	353
A. Etude 1 – Analyses phonétiques.	355
Annexe 1. Scénarios initiaux (21) – étude préliminaire 1	355
Annexe 2. Résultats – tâche préliminaire 1	362
Annexe 3. Productions libres – tâche préliminaire 2	364
Annexe 4. Scripts pour enregistrement du corpus – étude principale	380
Annexe 5. Graphiques des résultats non significatifs – étude principale	387
Annexe 6. Résumés des modèles LMER réalisés sur les données de la f0	389
Annexe 7. Résumés des modèles LMER réalisés sur les données temporelles	392
Annexe 8. Segments significatifs et résumés des GAMMs	394
B. Etude 1 – Analyses phonologiques.	400
Annexe 9. Ensemble des patrons produits dans la préface	400
Annexe 10. Ensemble des patrons produits dans les marqueurs discursifs	401
Annexe 11. Graphiques des résultats non significatifs dans la préface	403
Annexe 12. Résumés des modèles réalisés sur la valeur mélodique et l’alignement dans la préface	404
Annexe 13. Graphiques des résultats non significatifs dans les marqueurs discursifs	406
Annexe 14. Résumés des modèles réalisés sur la valeur mélodique et l’alignement dans les marqueurs discursifs	409
C. Etude de perception	416
Annexe 15. Liste des stimuli de la tâche préliminaire	416

Annexe 16. Questionnaire d'empathie BES-A	418
Annexe 17. Plots de densité pour les temps de réaction dans l'étude principale	419
Annexe 18. Résumés des emmeans réalisés sur le RT selon les différences tâches	420
Annexe 19. Taux de réponses par feedback selon la valence, résultats de la tâche préliminaire (phase de construction 1)	423

Introduction

« Il faut qu'on parle... je pense que cette relation ne va nulle part »

L'énoncé « il faut qu'on parle », ancré dans l'imaginaire collectif, laisse rarement indifférent celui qui le reçoit. Il est communément compris comme précédant l'annonce d'une nouvelle à minima importante, au pire mauvaise. Pourtant, selon la manière dont nous prononçons cette phrase, notre interlocuteur peut parfois s'attendre à recevoir plutôt une bonne nouvelle : « Il faut qu'on parle... je pense savoir où partir en vacances ». Nous pouvons alors nous demander où se situe la différence. Un élément de réponse réside dans des affirmations que l'on entend parfois : « ce ton ne me dit rien qui vaille », « ce ton n'annonce rien de bon ». Lorsque notre interlocuteur réagit de cette manière à un énoncé tel que « Il faut qu'on parle », cela laisse entendre que notre voix, notre « ton », comporte des indices sur la valeur affective de la nouvelle que l'on s'apprête à nous donner. Cela indique aussi que notre interlocuteur les a repérés, et se prépare à recevoir la nouvelle en conséquence. Les caractéristiques présentes dans la voix, qui varient selon un affect donné (ex. la joie, la peur, l'ennui, le positif), font l'objet d'études depuis longtemps de la part des spécialistes de la prosodie. De grandes questions sont étudiées : quels sont les indices de l'affect présents dans la prosodie ? Comment sont-ils produits ? Comment sont-ils perçus (Audibert, 2008; Bänziger et al., 2001; Laukkanen et al., 1996; Scherer, 2003) ?

Malgré le nombre très important d'articles et d'ouvrages proposés sur ce vaste champ d'étude, ces questions sont toujours pertinentes. Régulièrement, nous découvrons de nouveaux moyens d'observer la prosodie affective, et de nouveaux angles pour questionner son fonctionnement. De nombreuses questions demeurent ainsi ouvertes, et contribuent à une meilleure définition de l'objet. Certains indices prosodiques sont-ils plus proéminents que d'autres ? Lesquels facilitent le plus la perception d'un affect donné ? Par quel canal le sens affectif peut-il être porté, est-ce uniquement à travers de la variation continue, ou peut-il être aussi porté par des catégories de l'intonation ?

L'étude de la prosodie affective se concentre principalement sur les caractéristiques continues présentes dans la voix, comme véhicules d'indices de l'affect. Ces

caractéristiques phonétiques peuvent porter sur des variations de la hauteur, de la force, de la temporalité, et du type de voix. L'idée que le sens affectif puisse être porté par des catégories dans l'intonation a longtemps été écartée, notamment dans les cadres théoriques auxquels nous nous rattachons pour ce travail : la théorie Métrique Autosegmentale (J. B. Pierrehumbert, 1980), et plus généralement la phonologie de l'intonation (D. R. Ladd, 2008). Ce n'est que très récemment que cela a été remis en question, en particulier par trois études clés (Cao et al., 2014; Liscombe et al., 2003; Wang et al., 2020).

Au-delà de la question des caractéristiques vocales de l'affect, continues ou catégorielles, nous nous intéressons à un autre aspect de la parole affective, que nous avons mis en évidence dans l'exemple proposé au début de cette introduction. Si avec seulement l'énoncé « il faut qu'on parle », nous pouvons indiquer la valeur affective de la nouvelle que nous allons annoncer, alors nous devrions considérer les indices portés par la prosodie comme « projectifs » de cette valeur affective de la nouvelle. Les indices ainsi projetés dans l'énoncé pourraient être repris par l'interlocuteur pour anticiper la valeur affective de la nouvelle : sera-t-elle positive, ou négative ? Le terme de projection nous vient des domaines de l'Analyse Conversationnelle ou CA (Sacks et al., 1974) et de la Linguistique Interactionnelle ou LI (Couper-Kuhlen & Selting, 1996, 2017). Il fait référence au fait que les locuteurs en interaction présagent toujours des éléments de ce qu'ils vont dire ensuite, préparant ainsi leur interlocuteur, lui permettant même parfois d'anticiper et de répondre à ce qui va être dit, avant que cela n'ait été dit.

« Speakers consistently foreshadow their next moves, and while this foreshadowing does not specify all features of this move but only projects a silhouette, it nevertheless prepares the audience and often enables them to initiate a response before the act to which they respond is actually completed. » (Streeck, 1995, p. 105)

C'est à cette fonction projective de la prosodie affective que nous nous sommes intéressés tout au long de ce travail. Suivant l'exemple de Swerts et Hirschberg (2010), nous avons étudié le rôle projectif de la prosodie affective au moyen d'analyses phonétiques continues de la parole produite. A cela, nous avons ajouté l'analyse de caractéristiques phonologiques intonationnelles, et leur variation phonétique. Nous avons ainsi observé le comportement de la prosodie affective au sein d'énoncés clés

décrits dans les domaines de la CA et de la LI, qui précèdent les annonces de nouvelles. Enfin, nous avons observé la perception de ces énoncés, afin de vérifier qu'ils sont bien perçus, et qu'ils permettent une préparation interactionnelle et affective de celui qui reçoit la nouvelle. Cette préparation affective permet à l'interlocuteur d'anticiper la valeur affective positive ou négative de la nouvelle qui va lui être annoncée, et de s'y préparer pour réagir émotionnellement le mieux possible à l'annonce.

Problématique : La question de recherche qui a guidé les travaux de cette thèse est donc celle de la fonction projective de la prosodie affective. Dans quelle mesure projetons-nous des indices prosodiques de la valence affective d'une nouvelle que nous nous apprêtons à annoncer ? Par quels canaux cette projection prosodique passe-t-elle ? Cette projection est-elle perceptible pour l'interlocuteur, et utilisée pour anticiper et se préparer à l'annonce sur un plan affectif ?

Nos résultats principaux montrent que les énoncés précédant une annonce de nouvelle négative sont produits avec des valeurs élevées de la hauteur de la voix, et qu'ils sont marqués par une voix plus « soufflée », par rapport à lorsque la nouvelle est positive, ce qui est cohérent avec la littérature (Freese & Maynard, 1998; Swerts & Hirschberg, 2010). Lorsque la nouvelle est neutre, les caractéristiques acoustiques sont assez proches de celles de la nouvelle positive, bien que nous retrouvons des différences significatives. Nous avons trouvé des différences au niveau de la phonologie, bien que nos résultats soient moins probants que ceux que nous avons évoqués sur l'intonation affective (Cao et al., 2014; Liscombe et al., 2003; Wang et al., 2020). Les études sur la variation phonétique au sein des catégories phonologiques observées indiquent des différences significatives selon la valeur affective de la nouvelle à annoncer. Cela indique que nous devrions nous pencher sur ces caractéristiques pour mieux comprendre comment la parole affective est produite. Enfin, nous avons trouvé que les auditeurs qui écoutent des énoncés produits juste avant l'annonce d'une nouvelle semblent faire une différence selon la valeur affective de la nouvelle qui suit, sans l'avoir entendue. Les énoncés et les informations qu'ils portent ont influencé la réaction affective des auditeurs. Les effets de ces résultats sont cependant assez variables notamment lorsque l'on compare les nouvelles positives aux nouvelles neutres, qui pourraient être perceptivement confondues. Nous n'avons pu déterminer avec assurance si les indices de projection ont influencé uniquement le

ressenti affectif de l'interlocuteur, ou si des enjeux de règles interactionnelles ont aussi été pris en compte.

Le Chapitre 1 de ce manuscrit est consacré au cadre théorique que nous avons suivi, présenté en trois parties. La première partie (1.1.) introduit la notion d'indices de projection, tels que décrits dans les cadres de l'Analyse Conversationnelle et de la Linguistique Interactionnelle. Nous y explorons des notions connexes telles que la progressivité, la préférence, la préface, les marqueurs discursifs et les séquences d'annonces de nouvelles. La deuxième partie (1.2.) présente les fondements théoriques qui entourent la prosodie affective. Nous y voyons les distinctions entre sens linguistique et paralinguistique, phonétique et phonologie, les théories de l'intonation et son annotation, ainsi que le lien entre intonation et dualité de structure. La troisième partie (1.3.) présente les corrélats de la valence affective. Nous y explorons l'étendue des différences de catégorisation des affects, le type de corpus qui existent et permettent d'explorer ces notions, les corrélats phonétiques et phonologiques de l'affect que nous pourrions attendre dans ce travail, et les travaux réalisés sur les corrélats prosodiques des annonces de nouvelles, et valeur projective de la prosodie dans ce cadre. Le Chapitre 2 présente la construction du corpus que nous avons créé pour ce travail, ainsi que les analyses phonétiques qui ont été réalisées dessus. Le Chapitre 3 présente le travail d'annotation phonologique, ainsi que les analyses réalisées sur ces annotations, et l'étude de la variation phonétique rattachée à l'intonation. Le Chapitre 4 présente l'étude de perception réalisée sur les préfaces aux annonces de nouvelles. Un dernier chapitre, le Chapitre 5, est consacré à la discussion générale des résultats de nos études, en fonction des objectifs, et des apports de ce travail.

Chapitre 1 : Etat de l'art

1. La projection dans l'interaction

1.1. Les indices de projection

En interaction, il est fréquent d'avoir l'impression que l'on peut anticiper ce que le locuteur s'apprête à nous dire, ou une partie de l'information, et ce avant même qu'il ne l'exprime verbalement. Par exemple, lorsque nous recevons un appel d'un ami qui déclare, « Il faut trop que je te raconte un truc ! », et que l'on anticipe déjà une interaction où l'on va nous raconter une histoire, qui peut être amusante, ou au moins étonnante. Quand la maîtresse d'école au cours d'un rendez-vous parent-professeur commence par dire « Écoutez ce semestre est bien mieux » et on imagine immédiatement qu'elle va nous dire que l'enfant concerné a eu un bon comportement, de bonnes notes ou bien, qu'il a moins manqué l'école. Ou encore, lorsqu'un proche nous appelle et nous dit d'un ton grave « Bon. Écoute. C'est à propos de mamie... » avec un ton de la voix grave, et que l'on se prépare instinctivement au pire, ou au moins à une situation difficile. Les locuteurs semblent donc transmettre à travers plusieurs sources (tels que le choix précis de leurs mots, le ton de la voix) des informations sur ce qu'ils vont dire juste après.

Le phénomène par lequel le locuteur convoie des informations sur ce qu'il va dire après découle de ce que l'on appelle la « projection » (Auer, 2005; Ford & Thompson, 1996; Goodwin, 2002; Schegloff, 1996) : Goffman (1969) explore la manière dont la projection d'informations se déroule au cours des interactions. Il affirme que le sens d'un message dépend par exemple de la manière dont il est dit. Il est possible de moduler la voix pour faire un énoncé de manière sérieuse, joyeuse, sarcastique, ou encore moqueuse. Si l'on reprend la phrase proposée plus tôt, « Il faut trop que je te raconte un truc ! », et que celle-ci est prononcée d'une voix sérieuse ou d'une voix joyeuse, l'énoncé ne portera alors plus tout à fait le même sens. Ce type d'information sur la nature du message, porté par des « indices » dans la voix, va alors cadrer le sens du message. L'interlocuteur à qui l'on s'adresse va utiliser les « indices » tels que l'intonation et dériver le sens de l'énoncé en fonction de ceux-ci. Le locuteur projette donc des indices sur le sens de son message lorsqu'il fait un énoncé. On parle alors « d'indices de projection ». Dans notre exemple précédent lorsque la maîtresse dit « Écoutez ce semestre est bien mieux », elle produit

plusieurs indices de projection. En effet, la projection peut être appliquée à différents niveaux. Certains indices sont lexicaux : « bien mieux » est un terme qui indique déjà la nature positive de son message. La projection peut aussi être syntaxique, par exemple lorsque l'on utilise des questions courtes (« Qui ? ») et que l'interlocuteur sait qu'il peut immédiatement répondre après cet unique mot, sans avoir à attendre d'autres mots (Sacks et al., 1974). Elle peut être discursive, avec l'utilisation de mots tels que « bon », « en fait » ou « tiens » (Dargnat, 2020). Ces mots, sur lesquels nous reviendrons en section 1.2 projettent alors des informations sur la structure du discours, comme le fait de passer à un autre thème, ou d'introduire un nouveau sujet dans la conversation. La projection peut être aussi gestuelle. Par exemple, si l'on dit « Il ne fera pas ça » en pointant du doigt un manège (Tieu et al., 2018), on signifie à notre interlocuteur qu'une personne tierce (« il ») ne fera pas le manège, sans avoir explicitement verbalisé le mot « manège ». Enfin, la projection peut être prosodique. Si la maîtresse de notre exemple parle de façon sérieuse, cela constituera un indice additionnel de projection sur la nature positive du message, en plus du lexique. Si elle parle de façon sarcastique en revanche, ou avec une expression faciale de mécontentement, cela sera en conflit avec « bien mieux », laissant penser qu'au contraire, le semestre s'est mal passé, les indices de projection laissant plutôt imaginer une nature négative du message. La prosodie peut aussi projeter la fin de la prise de parole d'un individu. Par exemple, un locuteur peut utiliser des changements dans la hauteur de la voix, la qualité, ou dans la vitesse de parole ou la longueur des voyelles pour indiquer la fin d'un tour de parole ou d'un sujet (Brusco et al., 2020; Geluykens & Swerts, 1994; Wells & Macfarlane, 1998).

Nous allons nous intéresser à la projection au travers de deux domaines d'étude de la linguistique que sont l'Analyse Conversationnelle et l'Analyse Interactionnelle. L'Analyse Conversationnelle (Sacks et al., 1978) est une approche linguistique dans laquelle on s'intéresse aux interactions verbales entre individus dans les conversations. Son but est de comprendre comment les personnes échangent des informations, co-construisent le sens, et quelles pratiques conversationnelles elles utilisent pour créer une communication fluide. La Linguistique Interactionnelle (Couper-Kuhlen & Selting, 1996; E. Ochs et al., 1996) est une autre approche linguistique dans laquelle on s'intéresse aussi aux interactions verbales, dans les conversations, mais pas uniquement. Cette approche s'intéresse à d'autres types d'interactions sociales (entretiens, communication en ligne). Elle examine aussi la co-construction du sens, en prenant en compte l'influence que

peuvent avoir les contextes sociaux, culturels et discursifs. Ces deux approches nous permettent de comprendre les mécanismes conversationnels qui entourent la projection, et les multiples utilités qu'elle peut avoir en interaction.

La projection porte sur ce qui va suivre dans l'immédiat, mais peut aussi permettre que d'autres « actions » ou « actions conversationnelles », c'est-à-dire tout acte ou intention communicative accompli par les participants dans une interaction verbale, aient lieu. Les actions conversationnelles peuvent être, par exemple, le fait de poser une question, d'y répondre, exprimer un désaccord, offrir, contester, demander, taquiner, affiner, se conformer, exécuter, remarquer ou encore le fait de promettre quelque chose (Schegloff, 2007). Pour Auer (2005), la projection est justement une action individuelle qui en prépare une autre. La projection peut, par ailleurs, être déroulée en plusieurs temps et suivre ainsi une « trajectoire de projection », le déroulement temporel sur lequel la projection se développe et parvient à sa clôture ou à sa résolution. Elle porte sur plusieurs niveaux de la parole. Le locuteur peut vouloir projeter une action future, comme dans la définition d'Auer : un locuteur pourrait par exemple dire à son ou sa partenaire « j'ai quelque chose à te demander » pour projeter l'action de demande, dans laquelle le locuteur va demander à adopter un chat. Le locuteur, par cet énoncé, va aussi projeter l'énoncé suivant qui correspond à une partie de cette action, dans lequel le locuteur va explicitement dire « Est-ce que tu es d'accord pour qu'on adopte un chat ? », et dans laquelle l'interlocuteur va répondre si oui ou non. Par la même occasion, le locuteur projette son tour de parole, c'est-à-dire le moment où il prend la parole pour verbaliser la demande (Mondada, 1999).

Mais la projection ne peut fonctionner uniquement avec le locuteur. Pour qu'elle marche, il faut non seulement un locuteur qui fasse des indices de projection, mais aussi un interlocuteur qui écoute, récupère et utilise ces indices pour le message, la structure de l'énoncé, savoir s'il peut prendre la parole à son tour, ou encore anticiper le prochain énoncé du locuteur. Mais cela ne s'arrête pas à la seule compréhension de l'interlocuteur. Le locuteur va s'adapter à la réaction de celui-ci, alors même qu'il est encore en train de parler, modifiant ses propos pour s'assurer de la compréhension de l'autre. Si l'interlocuteur prend la parole, il devient alors lui-même locuteur et peut projeter des indices, récupérés par le locuteur devenu interlocuteur. Ce système de va-et-vient constant entre les interactants est inhérent aux conversations. La conversation est, en

effet, fondamentalement collaborative. On dit qu'elle est « co-construite » par les interactants.

Ce terme de co-construction (Goodwin, 1979; Jacoby & Ochs, 1995) est utilisé pour désigner le processus par lequel les participants à une interaction collaborent, en construisant conjointement le sens et la structure de la conversation. Cela s'oppose à l'idée qu'un échange serait linéaire, avec un locuteur-émetteur qui transmet un message, et un récepteur qui écoute (Shannon & Weaver, 1949). La co-construction implique une interdépendance étroite entre les participants pour élaborer ensemble les énoncés et les différentes parties du discours, d'où l'utilisation des termes « locuteur », « interlocuteur » et « interactants », qui permettent de rendre compte du caractère co-construit de « l'interaction ». La projection est justement l'un des dispositifs mis en place pour cette co-construction, puisqu'elle permet de proposer une « trajectoire interactionnelle » à l'interlocuteur, qui peut accepter ou non cette proposition. Streeck (1995) définit le terme de trajectoire tel qu'il est utilisé ici comme un moyen de décrire le déroulé des unités d'interaction, de la première indication sur le fait qu'il se passe quelque chose, à la complétion de l'unité. Le terme de trajectoire conversationnelle permet pour Streeck (1995) de répondre à un besoin de décrire le langage et les comportements comme des phénomènes en progression se déroulant vers leur complétion.

Dans une analyse du corpus du CID (Bertrand et al., 2006), Priego-Valverde et al. (2022) illustrent l'initiation d'une nouvelle trajectoire interactionnelle avec l'exemple suivant extrait d'une discussion où l'un des deux interactants s'apprête à aborder un sujet lié au temps qu'il a passé à travailler en Belgique :

- | | | |
|---|-----------|---|
| 1 | EB gpd_62 | ouais ça c'était ap ob assez particulier c'était euh c'était quand je |
| 2 | | bossais en Belgique euh |
| 3 | EB gpd_63 | et euh |
| 4 | SR gpd_49 | ah oui c'est vrai tu as bossé en Belgique je euh retire tout ce que |
| 5 | | j'ai dit euh |

Extrait de Priego-Valverde et al. (2022, p. 5)

Dans cet extrait, une trajectoire est proposée par EB (lignes 1 et 2) pour parler de quand EB travaillait en Belgique. Une trajectoire peut être acceptée, refusée ou même modifiée par l'interlocuteur. Dans cet exemple, la trajectoire initiale d'EB est modifiée par l'interlocuteur SR (ligne 4). SR transforme la modifie la trajectoire d'une conversation

sérieuse sur la Belgique vers une trajectoire où la conversation devient humoristique (« je euh retire tout ce que j'ai dit euh »). Cela montre qu'un interlocuteur B, en réponse à la projection d'une trajectoire par A, peut à son tour proposer une trajectoire qui sera en accord ou en désaccord, ou déviée de celle proposée par A. Il est possible qu'une trajectoire puisse être délaissée pendant quelque temps avant que les interactants ne la reprennent. Pour retourner à l'exemple du CID, il est possible que les interactants restent sur une trajectoire humoristique quelque temps, puis reprenne la trajectoire sérieuse proposée en ligne 1. Il est aussi possible qu'ils en dévient jusqu'à la fin de l'interaction, soit en restant sur le ton humoristique que SR a proposée, soit en enchaînant sur une autre trajectoire qui n'aura rien à voir avec le travail d'EB en Belgique (« En parlant de travail, ta sœur a eu son poste ? »), nouvelle trajectoire qui déviara alors complètement la conversation.

L'ordre de la conversation est ainsi l'objet d'une « négociation » conjointe par les interactants, processus où les participants à une conversation ajustent leurs tours de parole, clarifient le sens de leurs énoncés, co-construisent le sens, le tout afin de maintenir une communication fluide et inter-compréhensible (Goodwin, 1995; Lerner, 1993). On peut voir, au cours d'une interaction, de multiples trajectoires conversationnelles émerger, être séparées en plusieurs trajectoires (avec par exemple une trajectoire humoristique et une sérieuse sur un même thème), ou être suspendues puis reprises (Priego-Valverde et al., 2022).

Tout au long de la conversation, les interactants vont s'efforcer au travers de négociations d'assurer en priorité la « progressivité » de l'interaction (Stivers & Robinson, 2006), qui permet d'accomplir une activité interactionnelle. C'est-à-dire qu'ils lui permettent de se dérouler et d'avancer, petit à petit, vers la fin de l'interaction. Schegloff (1979) précise que la progressivité de l'interaction est faite et maintenue au moyen de matériaux linguistiques déjà utilisés dans l'interaction. Ce n'est donc pas un outil qui est utilisé, mais la conséquence de multiples procédés utilisés en interaction pour permettre à celle-ci d'arriver à son terme. Le progrès d'une activité donnée est donc au centre des préoccupations des interactants qui, en projetant et négociant les trajectoires conversationnelles, évitent la rupture complète du progrès de l'activité. Ils vont, par exemple, tout faire pour que la narration d'une histoire initiée par le locuteur progresse, malgré quelques potentielles digressions, vers l'accomplissement complet de la narration.

Lorsque l'on fait des indices de projection sur ce que l'on va dire, on projette aussi d'autres informations comme ce que l'on pense sur ce que l'on va dire, ce que l'on pense sur ce que l'on dit, ou encore ce que l'on voudrait entendre comme réponse de la part de l'interlocuteur. Par exemple, si l'on dit « J'ai une question un peu naïve... » avec une voix hésitante, on projette plusieurs informations. Tout d'abord, on projette le fait que l'on va poser une question. De plus, on projette aussi que l'on n'est pas confiant sur notre question (elle est « naïve », voix hésitante). Mais aussi, on projette le fait que l'on aimerait obtenir une réponse de l'interlocuteur (car on va poser une question), et on indique que l'on souhaite éviter une moquerie en prévenant d'avance que la question est « naïve ». Ce dernier point, le fait de vouloir éviter une réponse de l'interlocuteur dans laquelle il se moquerait, est ce que l'on appelle la « préférence » (Sacks, 1973; Schegloff, 1972).

Face à un énoncé de la part d'un locuteur, l'interlocuteur peut répondre de multiples façons. Si par exemple, un locuteur évoque un sujet et demande un assentiment à l'interlocuteur, celui-ci pourra être d'accord avec lui, et lui accorder l'assentiment, ou ne pas être d'accord et le refuser. Il pourrait aussi ne pas répondre du tout. Il y a entre ces trois options des différences de préférence, que nous illustrons dans la Figure 1. Le locuteur qui a demandé l'assentiment préférera, en priorité, un accord. C'est la réponse préférée. Si l'interlocuteur exprime un désaccord, il fournira alors une réponse non-préférée. Mais il peut aussi ne pas répondre du tout. Or, un locuteur préférera toujours, entre un désaccord et une absence de réponse, un désaccord. En effet le désaccord, bien qu'il puisse mettre dans une situation inconfortable les interlocuteurs, et qu'il perturbe la progressivité de l'interaction, n'arrêtera pas complètement l'échange. En revanche, une absence complète de réponse rompt totalement la progressivité, car l'action entamée par le locuteur, dans laquelle il demande l'assentiment de son interlocuteur, n'est pas complétée, et l'interaction échoue. Il y a donc dans cet exemple trois niveaux de préférence, respectivement : une réponse préférée, et deux non-préférées, hiérarchisées dans cette non-préférence.

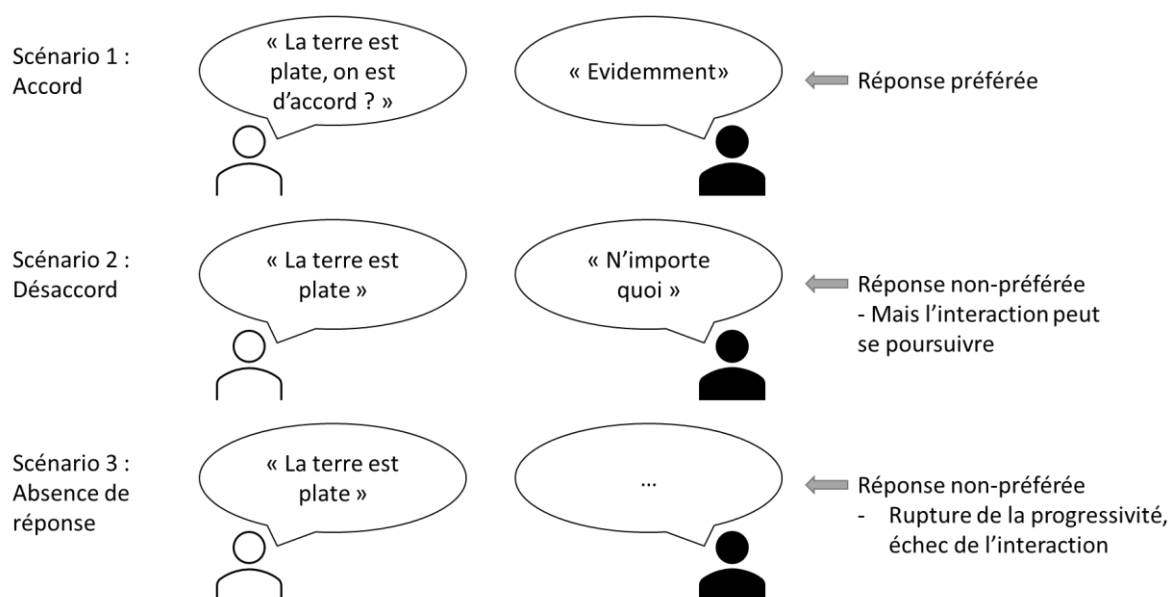


Figure 1 – Exemple de hiérarchie de préférence entre les réponses

Schegloff (2007) qualifie la structure de préférence de « caractéristique clé » dans l'agencement des unités interactionnelles, terme générique qui regroupe tout segment de discours qui forme une unité cohérente dans le contexte d'une conversation. En produisant un énoncé, on peut projeter la préférence que l'on a entre les différentes réponses possibles. Pour les participants à une interaction, il est primordial de pouvoir déceler les indices de préférence du locuteur, afin de l'anticiper et ainsi pouvoir répondre la réponse préférée, ou celle la plus haute possible sur la structure de préférence. La préférence entre donc dans les éléments à prendre en compte lorsque l'on co-construit le discours, car elle permet d'en assurer la progressivité en évitant des réponses non préférées qui pourraient mettre fin à une activité initiée par le locuteur.

1.1. Dispositifs de structuration du discours

Les interactants cherchent à assurer la progressivité de l'interaction en déterminant une trajectoire, en la négociant, et en essayant de fournir un maximum de réponses préférées. Mais comment cette progressivité est-elle déroulée et structurée ? En Analyse Conversationnelle et en Linguistique Interactionnelle, les phrases (entités linguistiques abstraites) et les énoncés (flux de parole réellement produits) sont comprises comme des formes d'action situées dans des contextes spécifiques et créées avec une attention

spécifique au contexte (Goodwin & Heritage, 1990), et c'est à travers ces phrases et énoncés, à travers ces « actions conversationnelles », que l'interaction est déroulée et découpée en plus petites parties.

Ces actions sont organisées au travers de « séquences », définies par Schegloff, (2007) comme des lignes de conduite (« *courses of action* ») qui sont mises en place à travers la parole, et composées d'au minimum une paire adjacente. Par exemple, si l'on parle d'une séquence de requête, on parle d'un moment de l'interaction pendant lequel les interactants suivent la ligne de conduite correspondante au fait de faire une requête. Cette requête sera constituée d'au moins une paire adjacente avec une première partie dans laquelle le locuteur va faire sa requête, et une seconde dans laquelle l'interlocuteur devra répondre à cette requête en l'acceptant ou non. Les paires adjacentes sont elles-mêmes constituées justement d'au moins deux énoncés, positionnés de façon adjacente, et prononcés par deux locuteurs différents, le locuteur et son interlocuteur (Sacks et al., 1974; Schegloff & Sacks, 1973). Selon Selting (2010), la première partie de la paire adjacente initie la séquence et rend pertinente, ou appropriée, la seconde. Les paires adjacentes peuvent donc être des séquences de requête (« Je peux m'asseoir ? » - « Oui »), d'annonce (« J'ai accouché » - « C'est merveilleux ! », ou de récit (« J'ai vu un chat l'autre jour, il dormait dans un vase, c'était adorable » - « Trop mignon »). Les séquences permettent d'accomplir une activité (Schegloff, 2007) et la progressivité cette activité sera assurée par les interactants, séquence après séquence.

Pour que cette structuration du discours fonctionne, il faut cependant aussi que lorsqu'un interlocuteur propose une trajectoire en entamant une séquence, l'interlocuteur comprenne quelle est la trajectoire proposée et l'accepte. S'il le fait, on considère alors qu'il s'est « aligné » à son interlocuteur (Stivers, 2008). Il va par exemple, répondre à une question posée. En revanche s'il répond à une question par une autre question, on pourrait considérer qu'il ne s'est pas aligné et refuse la trajectoire. Par exemple si la locutrice A demande « Je suis trop contente, je t'ai dit pour l'accouchement de ma cousine ? » et que le locuteur B répond « On n'irait pas manger ? », il refuse une annonce de nouvelle et propose une autre trajectoire conversationnelle. Se faisant, il rompt complètement la progressivité de l'activité d'annonce de nouvelle d'un accouchement, et se désaligne au niveau structurel de la conversation. Mais il pourrait aussi rester aligné et pour autant ne pas donner une réponse à laquelle l'on pourrait s'attendre. Par exemple, il pourrait répondre « Oh ma pauvre, tu vas plus la voir ta

cousine ». Se faisant, il est en rupture avec le point de vue affectif exprimé par la locutrice, appelé en anglais « *stance* », proposé par le locuteur A qui disait être « contente ». On considère alors que le locuteur B s'est « désaffilié ». S'il avait répondu en gardant dans sa réponse une valeur affective positive (« C'est super ! »), il se serait affilié au locuteur A. Hors, répondre en étant aligné et affilié est souvent la meilleure voie pour arriver à atteindre le but commun à toutes les interactions au cours d'une conversation : la « réussite de l'interaction » (Clark, 1996; Grice et al., 1975). Réussir une interaction correspond à être arrivé au bout de l'activité interactionnelle, en maintenant une interaction fluide, qui a du sens, et qui est mutuellement compréhensible.

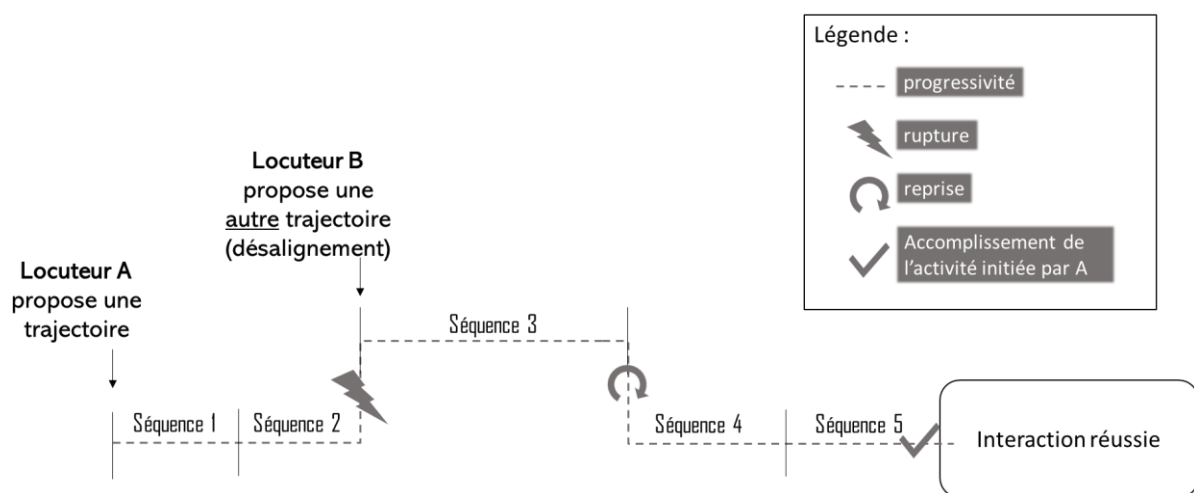


Figure 2 – Représentation schématique du cours de l'interaction

Nous pensons donc que l'utilisation d'indices de projection est un dispositif interactionnel qui permet de donner un maximum de chances à l'interlocuteur de rester aligné et affilié avec le locuteur. En récupérant ces indices, et donc en anticipant au mieux la trajectoire interactionnelle proposée par le locuteur, l'interlocuteur pourrait se préparer à répondre en gardant affiliation, et alignement. La projection au fil des séquences permettrait donc d'optimiser les chances de réussite de l'interaction. Certains types de séquences sont plus propices que d'autres à la projection.

Nous allons maintenant nous intéresser à ces espaces conversationnels particulièrement propices à la projection. Il existe un type particulier d'unités structurelles définies en Analyse Conversationnelle (Schegloff, 1980) qui répondent à cette demande : les « pré-séquences ». Ces actions de discours indirect, différentes des séquences, sont définies par Heritage (1989) comme des dispositifs de défrichage

destinés à établir le caractère approprié ou la pertinence de l'action ultérieure projetée par le locuteur. Les pré-séquences permettent donc de « préparer le terrain » pour lancer une séquence et s'assurer qu'elle soit bien reçue par l'interlocuteur, qui validera sa pertinence en acceptant que cette séquence projetée démarre. Il existe plusieurs types de pré-séquences, correspondant à chacune des actions. Par exemple, une pré-offre comme « Tu as l'air d'avoir du mal à porter cette valise » permet au locuteur d'établir la pertinence de la séquence d'offre, dans laquelle il pourrait produire un énoncé comme « Est-ce que tu veux mon aide pour la porter ? ». Une pré-invitation « J'ai envie de faire une pause » établira la pertinence de « Je t'offre un café ? », et une pré-annonce comme « Je suis désolée pour le repas » établira la pertinence de « J'ai un peu brûlé les carottes ».

Ainsi, il est possible avant d'annoncer les résultats d'un examen de préparer le terrain en commençant par un tour de parole pour introduire le sujet. Par exemple, "J'ai reçu mes résultats" projette une seconde séquence, celle où l'on donne les résultats en question. La pré-séquence correspond donc à une séquence qui en projette une autre, et projette donc une ligne de conduite, un « *course of action* » (dans notre exemple, l'annonce). Les pré-séquences peuvent avoir plusieurs utilités. Elles permettent d'arrêter en amont une séquence projetée dans laquelle il y a un risque de désaccord, de conflit, ou de rejet (Goodwin & Heritage, 1990), ou au contraire elles permettent une affiliation en laissant l'opportunité à l'interlocuteur d'anticiper la trajectoire conversationnelle. Les pré-séquences peuvent aussi induire une réponse préférée : éliciter une offre de la part de l'interlocuteur, en réponse à une pré-requête ou, dans le cas d'une annonce de mauvaise nouvelle, pousser le destinataire de la nouvelle à la deviner, afin que le locuteur qui devait l'annoncer n'ait finalement pas à le faire (Goodwin & Heritage, 1990). Pour Goodwin & Heritage (1990), la notion de pré-séquence de l'Analyse Conversationnelle serait un complément de la notion « d'acte indirect » de parole de la pragmatique. C'est-à-dire que par la pré-séquence, le locuteur fait l'acte d'introduire une nouvelle séquence sans nécessairement formuler le fait qu'il va faire une annonce, une requête ou une blague. Etant donné ces utilisations de la pré-séquence que nous avons décrites jusqu'ici, nous pensons que la pré-séquence est en effet, par essence, un espace conversationnel entièrement dédié à la projection. En permettant aux interactants d'anticiper le cours de la conversation, elle pourrait leur permettre de se préparer et protéger l'image d'eux-mêmes qu'ils souhaitent montrer.

Cette idée que les individus souhaitent protéger et conserver une image publique de soi a été théorisée dans les années 50 avec les travaux de Goffman (1955, 1967), et plus tard de Brown et Levinson (1987) soit la notion de « face ». Les participants à une interaction chercheraient à protéger leur propre face, mais aussi celle des autres interlocuteurs, afin qu'elles ne soient pas abîmées et afin que ces faces soient appréciées par autrui (c'est-à-dire, les interlocuteurs, mais aussi toute personne auprès de laquelle ils souhaitent conserver une certaine image). La présence de cet enjeu peut expliquer certains choix interactionnels faits par les interlocuteurs. Notamment, elle peut expliquer le fait que l'on cherche à rester alignés, affiliés, et à donner des réponses préférées. Heritage (1989) affirme en effet que l'organisation des préférences, dans un grand nombre de cas, est fortement reliée avec le fait d'éviter le mettre en danger la face (la sienne ou celle d'autrui). Si l'on reprend l'exemple en Figure 1, lorsque le locuteur demande un assentiment « La terre est plate, on est d'accord ? », si l'interlocuteur lui donne la réponse préférée en étant d'accord, il ne met pas en danger la face de l'autre, ni la sienne. S'il exprime un désaccord comme dans le scénario 2, il remet en question les connaissances du locuteur et peut démarrer une dispute. S'il ne répond pas, il rend la question non-pertinente dans la conversation et fait échouer la séquence, qui s'arrête. Les scénarios 2 et 3, aux réponses non-préférées, met donc en danger les faces du locuteur et de l'interlocuteur.

Cependant il n'est pas toujours possible de donner son assentiment simplement pour protéger la face de l'interlocuteur et sa propre face, notamment si l'on n'est pas d'accord avec l'affirmation faite par le locuteur. A défaut de pouvoir le faire, les interactants vont donc utiliser la projection dans la pré-séquence d'indices sur ce qui va être dit pour l'empêcher en amont si besoin. Dans notre exemple, le locuteur aurait pu faire une pré-séquence avant de dire « La terre est plate, on est d'accord ? », en commençant par dire « Tu sais, tous ces débats scientifiques sur la terre qui serait ronde, ou plate... ». Il aurait alors préparé le terrain pour une séquence interactionnelle avec une demande, une affirmation ou une question, sur la forme de la terre. Son interlocuteur aurait pu répondre, s'il pense que la terre est plate « Ah oui, j'ai toujours eu du mal avec la science, je n'y crois pas trop », et le locuteur aurait alors pu demander son assentiment sans trop risquer de mettre en danger leurs faces, ou de recevoir un désaccord. En revanche si l'interlocuteur ne croit pas aux théories complotistes, il pourrait répondre à la pré-séquence « Tu sais, tous ces débats scientifiques sur la terre qui serait ronde, ou plate... » par « Je ne sais pas

si on peut appeler ça un débat scientifique, la recherche a un avis clair sur la question » ou bien « haha, oui, les théories complotistes ». Il aura alors déjà donné des indications sur son avis sur la question, en montrant son adhérence avec la recherche, ou en utilisant les termes « théories complotistes », péjoratifs, pour désigner l'idée que la terre est plate. Le locuteur qui s'apprêtait à demander un assentiment pourra alors, s'il souhaite protéger sa face et celle de son interlocuteur, annuler la demande qu'il allait faire, et changer de sujet. Pour prendre un autre exemple, moins propice au conflit, un interlocuteur n'aura plus à refuser une invitation à une sortie si après une pré-séquence telle que « j'ai envie de sortir ce soir », il a répondu « et moi j'ai très envie de dormir ». En effet le locuteur aura alors probablement annulé l'invitation qu'il allait faire.

Le terme « pré-séquence » regroupe plusieurs variantes, plusieurs types de pré-séquences qui sont adaptées à la séquence que le locuteur cherche à introduire. La séquence d'invitation sera introduite par une pré-invitation, la séquence de demande sera introduite par une pré-demande, et la séquence d'annonce sera introduite par une pré-annonce. Toutes ces « pré » ont fait l'objet d'études en Analyse Conversationnelle et Linguistique Interactionnelle, et nous allons maintenant nous attarder quelque peu sur ce qui les caractérise. Les études qui se sont attachées à décrire ces pré-séquences sont faites sur des corpus naturels. Par exemple il y a la pré-invitation, dans laquelle on peut imaginer le locuteur dire « il y a un nouveau café en bas de la rue » avant de proposer d'aller y boire. Drew (1984) les décrit comme le fait d'indiquer au destinataire de l'invitation future qu'il y a une opportunité de faire une sortie, un évènement social. Cela laisse au destinataire de l'invitation la possibilité de proposer, de lui-même, un arrangement explicite de sortie qui correspond en fait à l'invitation que le locuteur souhaitait faire. Cela peut être fait en rapportant seulement l'activité qui est possible, et en retenant l'implication de l'activité sociale, c'est-à-dire l'invitation. Dans notre exemple, on mentionne l'ouverture d'un café, qui devient une opportunité pour une activité sociale, et on retient l'invitation même à aller y boire.

Heritage (1989), décrit la pré-invitation comme étant l'un des cas les plus clairs dans sa revue sur l'état d'avancée de l'Analyse Conversationnelle. Pour lui comme pour Drew (1984), elle permet de ne pas inviter une personne directement. Il voit un bel avantage à cela : puisque pré-inviter permet d'éviter au locuteur qui veut faire l'invitation d'avoir à la faire, si le destinataire de l'invitation projetée refuse l'invitation, alors cela ne sera pas perçu comme un réel refus. Si après que le locuteur qui invite a mentionné l'existence d'un

nouveau café, l'interlocuteur invité refuse en disant « Je n'aime pas trop le café », alors le locuteur peut changer la séquence d'invitation projetée (« Tu veux y aller ? ») et la remplacer par une autre (« Ah bon ? Moi j'adore. » ou bien « Ah, tu préfères le thé ? »). En effet, l'invitation n'a pas encore été faite, donc elle peut facilement être annulée. La pré-invitation aura donc permis d'éviter une situation de refus d'une activité sociale proposée, un événement interactionnel qui aurait pu mettre en danger la face du locuteur qui invite et celle de l'interlocuteur destinataire. Un refus aurait été une réponse non préférée qui aurait fragilisé la progressivité de l'interaction et mis en péril sa réussite.

Un deuxième type de pré-séquence est la pré-demande, comme lorsque l'on dit « je vois qu'il te reste des petits gâteaux » avant d'en demander un. Elles peuvent être opérées, comme dans notre exemple, en vérifiant la disponibilité d'un objet avant de le demander (Fox, 2015; Levinson, 1983; Schegloff, 2007). Là encore, on peut imaginer que l'interlocuteur qui reçoit la pré-demande sur la disponibilité des petits gâteaux pourrait, de lui-même, les proposer (« Il m'en reste oui. Tu en veux ? »). Cela évitera au locuteur d'avoir à réellement demander les gâteaux, et risquer un refus. Les pré-demandes ont été analysées dans la vie quotidienne ainsi que dans des contextes institutionnels, sur des corpus naturels dont un corpus vidéo, permettant notamment l'analyse de quelques éléments gestuels qui peuvent rentrer en compte dans la pré-demande (Lindström et al., 2005). Un troisième type de pré-séquence est la pré-blague, lorsque l'on s'exclame « elle va te faire rire celle-là » avant de raconter la blague. Sacks (1989) compare son fonctionnement à celui des pré-histoires, que l'on peut faire avec le même type de phrases : « Tu veux savoir », avant de raconter une blague ou une histoire. Il y a, enfin, la pré-annonce, lorsque l'on appelle un proche et que l'on dit « J'ai eu les résultats du concours » avant d'annoncer qu'on l'a réussi. Ce cas, analysé en détails par Terasaki (1976, 2004) et repris par Levinson (1983) puis Heritage (1989), nous retient tout particulièrement. Ce travail est en effet dédié à l'étude de ce qui se passe avant l'annonce d'une nouvelle. Nous nous y intéressons donc dans la prochaine partie de ce chapitre.

Pour bien comprendre le fonctionnement de la pré-annonce, il est nécessaire de s'intéresser aux enjeux d'une annonce de nouvelle. Pour ce faire nous allons prendre un exemple très direct : l'annonce d'une mauvaise nouvelle. Cette action interactionnelle comporte de multiples risques car elle est considérée comme une « action non-préférée » (Levinson, 1983), pour le locuteur comme pour l'interlocuteur. Cette action est non - préférée car elle initie une activité non désirée, faire l'annonce de la nouvelle et de ses

conséquences pour le locuteur, et la recevoir, ainsi que ses conséquences, pour l'interlocuteur destinataire de la nouvelle. C'est une activité qui peut mettre en danger la face de chacun. Pour l'interlocuteur destinataire, recevoir une mauvaise nouvelle peut être difficile car il peut réagir fort, en termes affectifs et interactionnels, et potentiellement perdre la face. Par exemple lorsqu'on lui annonce une maladie, il pourrait pleurer, se mettre en colère, ou rentrer dans le déni. Pour le locuteur, faire l'annonce d'une mauvaise nouvelle est une activité difficile justement parce qu'elle fera réagir de façon imprévisible l'interlocuteur qui ne voudrait ou ne saurait gérer des pleurs (en particulier s'ils ne sont pas considérés comme appropriés, comme c'est souvent le cas dans le monde du travail). De plus cela peut compliquer le déroulement de l'interaction, sa progressivité, son succès. La réticence du locuteur à donner la mauvaise nouvelle se vérifie interactionnellement par la présence de pauses longues ou d'excuses.

Levinson (1983) insiste sur un autre point important lorsqu'on annonce une nouvelle : comme dans tout autre type d'activité, le locuteur essaie de ne pas communiquer des informations à autrui si la personne les connaît déjà, donc de respecter la maxime de quantité de Grice (1979) : on évite de dire une chose à quelqu'un si la personne est déjà au courant. Dans le cas de l'annonce d'une nouvelle, on évite de l'annoncer si notre interlocuteur l'a déjà entendue. Pour pouvoir éviter de l'annoncer, il faut s'être renseigné au préalable sur l'état des connaissances de l'interlocuteur destinataire de la nouvelle. Le locuteur peut justement se renseigner au travers d'une pré-séquence. En conclusion, la pré-séquence d'annonce (aussi appelée pré-annonce) sert à la fois à répondre à la maxime de quantité, et à protéger les faces des interactants. Ci-dessous, nous présentons un exemple d'annonce de bonne nouvelle, et un exemple d'annonce de mauvaise nouvelle (D représente « *deliverer* », le locuteur qui annonce ; R représente « *recipient* », l'interlocuteur qui reçoit, « *Didju* » est une transcription de la prononciation orale de « *Did you* ») :

- 1 D : I forgot to tell you the two best things that happen' to me today.
- 2 R : Oh super = what where they.
- 3 D : I got a B + on my math test ... and I got an athletic award.

Bonne nouvelle : Extrait de Levinson (1983, p. 349), citant Terasaki (1976, p. 53)

- 1 D : Didju hear the terrible news?
2 R : No. What.
3 D : Y'know your Grandpa Bill's brother Dan?
4 R : He died.
5 D : Yeah.

Mauvaise nouvelle : Extrait de Levinson (1983, p. 350), citant Terasaki (1976, p. 28)

En produisant un énoncé de pré-annonce comme dans les deux exemples ci-dessous (en ligne 1), le locuteur va projeter plusieurs informations. Dans la préface à une bonne nouvelle, il indique qu'il y a deux éléments positifs, il donne une indication sur leur temporalité (« *today* »), et il montre la valeur positive de la nouvelle (« *best things* »). L'interlocuteur lui demande alors quelles sont ces nouvelles, ce qui rend pertinente l'annonce de la nouvelle en ligne 2, et en ligne 3. Dans l'extrait sur la mauvaise nouvelle, le déroulé est légèrement différent. La pré-annonce en ligne 1 comporte le terme « terrible », qui donne une indication sur la valeur affective négative de la nouvelle qu'il va annoncer, et qui donne une gradation de cette valence très négative (il aurait simplement pu dire « bad news »). L'interlocuteur destinataire de la nouvelle répond en ligne 2 qu'il ne la connaît pas. Le locuteur qui annonce pose alors une seconde question en ligne 3 au lieu d'annoncer la mauvaise nouvelle, donnant ainsi plus d'indications sur cette nouvelle (elle est au sujet d'un grand oncle). Cela laisse une nouvelle chance à l'interlocuteur pour deviner la mauvaise nouvelle, et évite au locuteur d'avoir à l'annoncer et donc à performer une action non désirée. En ligne 4, l'interlocuteur fait alors une supposition : « *He died* ». La nouvelle ayant été correctement anticipée, le locuteur qui voulait annoncer la nouvelle n'a alors plus qu'à confirmer la supposition, en ligne 5. Ainsi, tous les indices projetés par le locuteur peuvent être récupérés et interprétés par l'interlocuteur afin d'anticiper des informations sur la nature de la nouvelle.

Anticiper ces informations aura plusieurs utilités, pour l'interlocuteur et pour le locuteur. L'interlocuteur pourrait refuser que la nouvelle lui soit donnée en amont de l'annonce en montrant au locuteur qui projette l'action d'annonce qu'elle ne serait pas appropriée, pertinente, possible ou désirée, ce qui permettra d'éviter de faire un rejet de l'activité qui mettrait en danger la face des interactants (Heritage, 1989). Ce refus peut être dû au fait que l'interlocuteur connaît déjà la nouvelle, ou bien qu'il ne la veuille pas entendre. Pour le locuteur qui a projeté des indices sur une nouvelle négative, l'objectif serait ainsi d'annuler complètement la nécessité de faire l'action non-préférée d'annonce

en préfigurant l'action de manière que l'interlocuteur puisse supposer la nouvelle (Levinson, 1983). La pré-séquence est aussi pour le locuteur une opportunité de s'assurer qu'il ne va pas violer la maxime de quantité de Grice (1979) en donnant une nouvelle déjà connue. En conclusion, faire une pré-annonce donne à l'interlocuteur une opportunité d'anticiper la nouvelle, se préparer, optimiser ses chances de garder la face. Elle donne au locuteur une chance d'éviter une action non préférée, de respecter la maxime de quantité, et de garder la face. Pour les deux interactants, elle augmente les chances d'alignement et d'affiliation, protège donc la progressivité de l'interaction, et ses chances de réussite.

1.2. Préfaces et marqueurs discursifs

Il convient de discuter maintenant d'un point important de terminologie. Jusqu'à présent, nous avons parlé de cette partie qui précède l'annonce d'une nouvelle avec le terme « pré-séquence » (Heritage, 1989; Levinson, 1983; Terasaki, 1976; Terasaki, 2004). Mais il en existe un autre, proche ; celui de « préface ». Ces deux termes semblent avoir souvent été utilisés de façon interchangeable dans la littérature (Drew, 1984; Heritage, 1989; Schegloff, 2007), le terme de préface étant, semble-t-il, utilisé pour parler du même type de phénomène (Sacks, 1989). La citation suivante est, par exemple, une définition qui colle totalement à ce que nous avons nommé jusqu'à maintenant « pré-séquence ». Pourtant, le terme utilisé est celui de « préface » :

« The role of prefaces - or pre's (as conversation analysts have fondly nicknamed these pet phenomena) - is to 'foreshadow' or 'project' (Sacks et al. 1974) something that comes after them, to bring it into play and 'prepare the scene' (Schegloff, 1984b). They allow other participants a certain premonition as to what this actor might be up to next »

(Streeck, 1995, p. 87)

Streeck (1995) définit ici les préfaces comme des éléments qui projettent ce qui vient après. Les préfaces préparent le terrain, laissant l'interlocuteur anticiper en partie de ce que le locuteur va dire ensuite. Pourtant dans la littérature, on trouve des études sur des préfaces d'un ou deux mots tout au plus, qui ne ressemblent pas tout à fait aux pré-séquences. En anglais on trouve par exemple une étude sur « *well* » (Heritage, 2015), que

l'on pourrait traduire par « eh bien » ou « bien » et une autre sur « *I mean* » (Maynard, 2011), que l'on traduit par « je veux dire ». En finnois, on trouve une étude sur « *siis* » (Hakulinen & Couper-Kuhlen, 2015), qui veut dire « donc », et qui est utilisé pour faire la préface d'une séquence avec un tour de parole très émotif. Mais Cashion (1986) utilise aussi le terme de préface pour parler de petites phrases précédant des blagues, telles que « *You'll love this one* » (« Tu vas adorer celle-là »), ce qui nous empêche à nouveau de différencier préfaces et pré-séquences simplement par la longueur. Cashion (1986) souligne dans son article que les préfaces servent à conserver la face de l'interlocuteur et celle du locuteur qui raconte la blague, exactement comme la pré-séquence. Pour Heritage (2015), la pré-séquence comme la préface servent à alerter de l'arrivée d'une action dans laquelle le locuteur y privilégie sa propre perspective, puisqu'il va démarrer une nouvelle partie de l'interaction sur un sujet qui le concerne et qu'il voit de son point de vue (c'est son invitation, sa demande, sa blague, ou sa nouvelle). Si la différence entre pré-séquences et préfaces ne se fait ni par la longueur, ni par la fonction générale de préparation de l'arrivée d'une séquence, il faut donc regarder si l'on peut les distinguer à un niveau plus subtil.

Seulement deux articles mentionnent les deux termes et font une différence entre pré-séquence et préface, et ce dans le cadre d'annonces de nouvelles. Freese et Maynard (1998) et Maynard (2017) se sont intéressés à ces phénomènes précisément dans le cas d'annonces de mauvaises nouvelles en milieu médical. Maynard définit la pré-séquence d'annonce (qu'il nomme « pré-annonce ») comme un dispositif par lequel un locuteur qui va annoncer peut découvrir si l'interlocuteur destinataire connaît déjà la nouvelle, et donc lui laisser l'opportunité de la dire, montrer qu'il la connaît, ou signaler au locuteur qu'il peut annoncer la nouvelle. Cet accord peut être donné avec un simple silence comme dans l'exemple suivant où un docteur appelle un patient pour annoncer une maladie et où « *A chilling pause* » est interprété comme le feu vert donné par le destinataire :

« *That evening, the doctor called. "It's bad," he said.*

A chilling pause.

"You have acute myelogenous leukemia—you have to go to the hospital tonight." »

Extrait de Maynard (2017, p.6)

La préface au contraire, est définie comme un énoncé qui empêche l’interlocuteur destinataire de la nouvelle de donner son feu vert, par la parole ou par un silence. Cela permet du même coup au locuteur qui va annoncer la nouvelle de faire la préface et d’enchaîner immédiatement avec l’annonce. Elle serait préférée lorsque celui qui annonce la nouvelle est à peu près sûr que le destinataire ne la connaît pas déjà. Maynard prend l’exemple suivant dans lequel un médecin annonce à son patient :

« *This combination of cerebellar dysfunction in one arm and corticospinal tract dysfunction in the other (shakes head, raises eyebrows, looks at patient). I’m sorry you know it’s stronger than any other laboratory test we have. It’s... there’s no other disease but multiple sclerosis that will do it* »

Extrait de Maynard (2017, p.6)

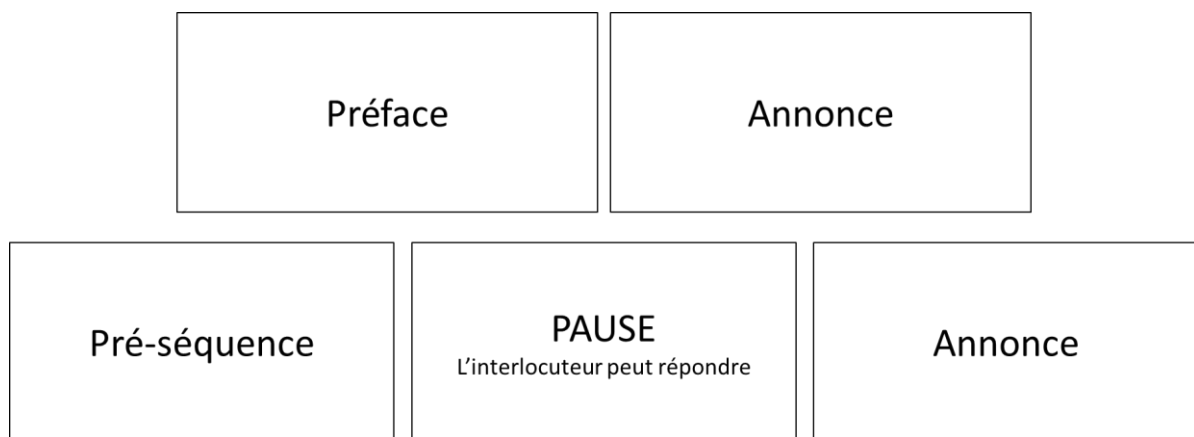


Figure 3 – Représentation schématique de la différence entre pré-annonce et préface à l’annonce, proposée d’après la distinction de Maynard (2017)

Pour Freese et Maynard (1998), la différence entre pré-séquence et préface réside donc dans l’avis du locuteur qui annonce une nouvelle sur les connaissances de l’interlocuteur à ce sujet : connaît-il déjà potentiellement la nouvelle ? La pré-séquence est faite lorsque l’on pense que le destinataire connaît ou peut deviner la nouvelle, et on attend son feu vert pour l’annoncer. La préface est faite lorsque l’on pense que le destinataire ne connaît pas la nouvelle et ne la devinera pas, et n’appelle donc pas à une réponse de sa part : elle sert simplement à communiquer l’intention du locuteur en début de séquence. Elle serait favorisée aussi dans les contextes où l’interlocuteur destinataire de la nouvelle n’a pas de

possibilité de répondre dans l'immédiat au locuteur qui annonce, comme ce serait le cas dans un message vocal sur répondeur téléphonique. Suivant la distinction de Maynard (2017), nous préférons donc utiliser le terme « préface » à partir de maintenant, plus adapté à notre étude sur répondeur téléphonique.

Nous avons brièvement évoqué plus tôt qu'il est possible de faire des préfaces de seulement un mot ou deux avec notamment le cas des tours de parole, c'est-à-dire des portions de l'interaction dans lesquelles un participant parle ou communique sans interruption, préfacés par « *well* » (« bien ») en anglais (Heritage, 2015), mais on retrouve aussi d'autres mots sur lesquels Heritage a travaillé comme « *and* » (« et ») (Heritage & Sorjonen, 1994) ou « *ei* » (« non ») en estonien (Keevallik, 2012, 2016). Il se trouve qu'au-delà de son utilisation en tant que préface, « *well* » est avant tout défini comme un « marqueur discursif », terme parfois abrégé en « DM », de l'anglais « *discourse markers* ». On retrouve dans les marqueurs du discours des mots comme « *well* », « *so* », « *oh* », « *y'know* » en anglais, ou « voilà », « donc », « ok », en français. Si un marqueur discursif peut constituer à lui seul une préface, il apparaît alors pertinent de s'y intéresser pour notre objet d'étude. Nous allons donc maintenant définir ce qu'ils sont. Les marqueurs discursifs sont devenus un champ d'étude productif à partir des années 1980. Comme le souligne Fraser (1988) dans sa revue des types de DMs en anglais, Levinson (1983) fut la première personne à souligner l'intérêt de ces petits mots, à l'époque nommés « particules de discours » qui avaient jusque-là été considérés comme peu importants. Puis le thème a été repris par plusieurs linguistes en pragmatique ou dans le cadre de la théorie de la pertinence (Sperber & Wilson, 2004), théorie selon laquelle les individus visent à obtenir un équilibre optimal entre effort cognitif et compréhension.

Les marqueurs du discours contribuent à la cohérence du discours, ayant une fonction d'intégration dans le discours : ils coordonnent les énoncés en contexte, indexent les énoncés adjacents pour les interlocuteurs, et indexent un énoncé précédent ou suivant. Ils sont décrits comme de la colle à discours qui connecte les énoncés, séquentiellement dépendants (Fraser, 1988; Schiffrin, 1987; Schourup, 1999). Ils sont des indicateurs de la cohérence du discours (Ochs & Schieffelin, 1989). Du point de vue de la pertinence, les marqueurs du discours seraient un moyen de montrer la pertinence d'un segment de discours par rapport à un autre en « inférant une connexion » entre les deux (Blakemore, 1987; Fraser, 1988; Jucker, 1993). Dans l'énoncé « je pense donc je suis », le marqueur « donc » infère une connexion entre ce qui est avant et ce qui est après, « je pense »

rendant pertinent l'énoncé « je suis ». Dans l'énoncé « J'ai pas envie de parler, je suis fatiguée, voilà ! », le marqueur « voilà » infère une connexion entre « J'ai pas envie de parler » et « je suis fatiguée », établissant la pertinence de la première partie de l'énoncé par rapport à la seconde.

Les marqueurs discursifs peuvent avoir de multiples fonctions en tant que « pivots structurels ». Il y a par exemple, la fonction d'introduction : « Alors, et si on décidait où on part en vacances ? » où le marqueur « alors » introduit un nouveau sujet. Un peu plus tard dans la conversation ainsi entamée, on pourrait avoir un marqueur discursif de résomption d'un sujet : « Bon, donc on part à a mer ? », où « bon » permet, après un long débat sur le meilleur moyen de se ressourcer, de retourner au sujet initial qui serait de choisir un lieu de vacances. Une autre fonction est celle de conclusion : « Enfin voilà pourquoi je préfère la montagne », où « enfin » marque la fin d'un argumentaire sur les bienfaits de la montagne par rapport à la mer. Ces fonctions peuvent aussi être plus précises. Par exemple « *so* » en anglais peut servir à introduire un sujet de façon à faire comprendre que ce qui va être abordé était dans la tête du locuteur depuis un moment, bien que ça n'arrive sur le tapis qu'au moment où le marqueur du discours est dit (Bolden, 2006).

Les marqueurs du discours sont donc une aide à la structuration du discours, où ils ajoutent du lien entre les énoncés pour fluidifier l'interaction. Ils partagent avec les préfaces une fonction d'introduction au prochain énoncé, bien qu'ils n'agissent pas tout à fait sur le même plan : les marqueurs du discours ayant avant tout une fonction de structuration du discours, et les préfaces une fonction de projection du discours. Cependant, ces points communs rendent idéale l'utilisation de DMs en préface. Depuis les années 2000, les marqueurs du discours semblent justement avoir été l'objet d'étude dans des cadres où ils sont produits en préface d'un énoncé, dans des études sur corpus naturels en Analyse Conversationnelle (Bolden, 2006, 2008, 2009; Maschler & Schiffrin, 2015). Bolden (2006) s'est par exemple attachée à décrire l'utilisation du marqueur discursifs « *so* » anglais en dehors de sa fonction de structuration du discours, lorsqu'il est utilisé en préface d'énoncés pour lancer de nouveaux sujets de conversation (tels que des discussions sur un sujet, des demandes, des invitations ou des offres). Les marqueurs utilisés de cette manière permettent d'initier des séquences d'action (Sacks et al., 1974; Schegloff, 2006). Bolden (2009) conclue que « *so* », lorsqu'il est utilisé en contexte institutionnel, sert à prévoir le tour de parole qui va suivre et qui montre un changement

de programme. Il agit alors en marqueur projetant ce changement sur la trajectoire de l'action.

En français, les marqueurs du discours sont aussi multiples : bien, alors, voilà, donc, ok. Nous nous intéressons ici à deux marqueurs du discours en français, « voilà » et « donc », qui sont tous deux présents dans le corpus que nous avons enregistré et analysé pour cette thèse, à la jonction entre préface et annonce d'une nouvelle. Le marqueur « voilà » a été étudié en linguistique par de nombreux auteurs (Cesare, 2011; Delahaie, 2013; Petit, 2020). En interaction notamment, il est observé pour son statut d'élément de discours pouvant servir plusieurs usages, séparément ou en même temps. Parfois décrit comme « tic de langage » (Col et al., 2020; Pourquery, 2014) de notre époque au vu de son utilisation généralisée dans beaucoup de contextes différents, c'est un mot extrêmement polyvalent. Col, Danino et Rault (2015), puis Col et Danino (2020) ont dressé des revues de la littérature à son sujet. Dans l'étude de 2015, le corpus utilisé pour analyser « voilà » est composé de multiples sources (des transcriptions de discours oral, d'un roman, et de textes issus de la presse écrite). Le corpus est annoté et transcrit en fonction de multiples critères. Le marqueur discursif « voilà » peut avoir différents statuts syntaxiques : adverbe (« eh bien voilà »), interjection (« assieds-toi, voilà »), pivot verbal lorsqu'il se comporte presque comme un verbe (« me voilà propre » qui se rapproche de « je suis propre »), préposition lorsqu'il introduit une information liée à des circonstances (« voilà deux décennies »). Lorsqu'il est utilisé en interjection, il est en position détachée du reste de l'énoncé, voire holophrastique, c'est-à-dire qu'il peut être utilisé entièrement seul. Col et al. (2015 ; 2020) distinguent deux grandes fonctions discursives pour « voilà », qui peut avoir une fonction introductive (« voilà une nouvelle robe »), pour entamer une nouvelle partie du discours ou introduire un nouvel élément, ou une fonction conclusive (« alors c'est décidé on fait comme ça, voilà »), pour fermer une partie du discours. Ce marqueur discursif a, enfin, deux valeurs possibles :

- Valeur de balisage, pour délimiter des régions dans la structure du discours, notamment le début ou la fin d'une nouvelle séquence : « ça y est, il mange, *voilà*, ça lui fait du bien »
- Valeur prédicative, servant à structurer cette fois les informations contenues dans le discours, pour associer une cause à un élément, ou introduire un nouvel élément : « voilà la raison de mon silence », « me voilà à Marseille »

Nous nous intéressons également à un second marqueur discursif : « donc ». Plusieurs études ont cherché à comprendre l'utilisation de « donc » en français (Ferrari & Rossari, 1994; Hansen, 1997; Pellet, 2005; Vlemings, 2003; Zénone, 1981). Vlemings (2003) précise que lorsqu'il est utilisé, « donc » infère que le sujet abordé n'est pas entièrement nouveau dans la conversation mais fait au contraire référence à un antécédent (linguistique ou non), car il a en général une fonction de résomption ou de conclusion : par exemple on peut dire à un collègue « donc tu n'as pas mangé ce midi » en référence au fait qu'on l'a vu ne pas manger au repas, ou en référence à une discussion au cours de laquelle il a mentionné ne pas avoir mangé. Le marqueur « donc » permettrait aussi, par exemple, de passer d'une préface à une séquence qui devrait être inférée de ladite préface : « j'ai quelque chose à te dire... donc... j'ai accepté le boulot à Marseille ».

Pellet (2005) liste dans sa thèse sur l'apprentissage du français L2 les fonctions de « donc » attribuées au fil des études faite sur ce marqueur discursif. Comme pour « voilà », on retrouve la fonction de résomption proposé par Zénone (1981). Pellet (2005) considère que « donc » fonctionne de façon métalinguistique, c'est-à-dire qu'il est utilisé par le locuteur lorsqu'il réfléchit à la structure de son énoncé. « Donc » peut également permettre d'imposer sa propre perspective à l'interlocuteur (Pellet, 2005), comme dans l'exemple proposé ci-dessous où le locuteur A utilise « donc » en ligne 3 pour créer un lien entre ce que B a dit en lignes 1 et 2, et ce que lui dit en ligne 3, imposant alors sa perspective sur ce qui a été dit en lignes 1 et 2 :

- | | | |
|---|-----|--|
| 1 | B : | non mais c'était une position encore une fois fondée sur des |
| 2 | | principes, je les ai rappelés tout à l'heure |
| 3 | A : | donc vous avez changé de principes |
| 4 | B : | non, nous n'avons pas changé de principes |
| 5 | A : | Ah ben si puisque maintenant vous voulez une liste unique |

Extrait de Pellet (2005, p. 85), citant Hansen (1997, p. 167)

En considérant les moments où les marqueurs discursifs sont parfois utilisés comme préfaces, il serait pertinent d'examiner comment les marqueurs tels que « voilà » et « donc » se comportent lorsqu'ils sont positionnés de manière à susciter cette utilisation. Si, par exemple, ils sont placés entre une préface et une annonce de nouvelle, comme dans la phrase « En ce qui concerne ton chat, voilà donc, il est malade », nous pensons qu'ils peuvent alors porter des indices de projection sur la nouvelle qui va être annoncée, et

pour cela nous porterons notre attention non seulement sur les préfaces mais aussi sur les marqueurs discursifs.

1.3. Projection d'indices sur la valeur affective d'une séquence

Nous avons à présent établi plusieurs éléments : tout d'abord l'existence des préfaces, espaces dédiés à la projection en interaction ; mais aussi, nous avons vu ce que sont les marqueurs du discours, leur rôle de projection structurelle, et nous avons remarqué qu'ils peuvent parfois être utilisés en préface. La possibilité d'une projection de valence affective au sein de la préface a rapidement été évoquée, et nous allons maintenant développer cette idée. Nous pensons que cette projection d'affect est possible au vu de la nature de la préface, qui peut être utilisée pour négocier la trajectoire conversationnelle, car comme le souligne Selting (2010) l'engagement émotionnel et l'affect est montré, répondu et négocié en interaction. On peut alors penser que la préface pourrait permettre de s'accorder sur les affects à montrer à son interlocuteur dans le cadre de la négociation de la trajectoire conversationnelle, mais aussi d'une forme de « trajectoire affective ». Celle-ci serait projetée par l'interlocuteur pour signaler la valeur affective qui sera rattachée à la séquence projetée : une préface à une annonce de mauvaise nouvelle serait produite avec des indices de projection sur la valeur négative de la nouvelle, indiquant vers quelles affectivités le locuteur se dirige (une séquence affective positive, ou une séquence affective négative). Bien que peu de recherche semble avoir pour focus la valeur affective de projection dans la préface, nous retrouvons tout de même deux études qui mentionnent ces utilisations.

Dans une étude récente, Stranovska et al. (2016) évoquent le fait que la production d'une pré-séquence peut permettre la réalisation d'une séquence de politesse. L'exemple qui est pris est celui d'une étudiante qui demande à emprunter le téléphone de son professeur pour passer un appel :

*« Hello, professor XY. I have a request on you.
I forgot my phone at home and I need to make an urgent call.*

Can I use your phone, please? »

Extrait de Stranovská et al. (2016, p.924)

Dans cet exemple, la pré-séquence "*I have a request on you*" est une pré-demande. Elle projette la séquence de demande, mais elle permet aussi de faire cette demande de façon polie. Mais il y a une seconde partie à cette pré-séquence, dans laquelle l'étudiante précise l'objet et la raison de sa demande avant de la faire. Cette seconde partie contient le mot "*urgent*" et Stranovska et al. (2016) soulignent que ce terme est un intensificateur qui apporte un facteur émotionnel (l'urgence de l'appel). Un tel type de terme intensificateur contribue à persuader le professeur d'accepter la demande.

Dans une étude sur les annonces de mauvaises nouvelles dans le domaine médical, Maynard (1996) propose l'exemple d'une pré-séquence (et spécifiquement une pré-annonce) dans laquelle on retrouve un élément lexical qui projette la valeur affective de la nouvelle qui va être donnée (le décès d'un proche) :

« Janey's mother said, "I've got some rather sad news" [...]

Janey asked, "What?" [...]

Janey's mother answered, "Grandpa died this morning" »

Extrait de Maynard (1996, p.115)

Dans cet exemple, la pré-séquence « *I've got some rather sad news* » est une pré-annonce. Elle projette la séquence d'annonce et sa nature de façon évidente (« *rather sad news* » où « *sad* » indique que la nouvelle est négative, et « *rather* » est un moyen d'atténuer le terme, comme lorsque l'on dit en français « j'ai des nouvelles assez mauvaises »). Puis, « *Janey* », l'interlocuteur destinataire de la nouvelle, complète la seconde partie de la paire de la pré-séquence et autorise sa mère à faire l'annonce (« *What?* ») en rendant l'annonce pertinente. Sa mère fait alors l'annonce du décès du grand-père.

Maynard (1996) oppose ces pré-séquences où la valeur affective de la nouvelle est dite explicitement à des pré-séquences où il n'y a aucun indice de projection affective lexical (potentiellement car la nouvelle n'est ni positive, ni négative). Il souligne que, pourtant, ces pré-séquences plus vagues permettent quand même de porter la nature de la nouvelle (positive, négative, sérieuse, importante, drôle), mentionnant le rôle potentiel

du ton « sérieux » et « préoccupé ». Il donne pour exemple la pré-annonce suivante, faite par un homme qui s'apprête à annoncer à sa famille qu'il est gay :

« *At dinner one night during Christmas break, I said, "There is something that I want to tell you two, and I want you to be an active part of my life, and feel that without telling, you won't be."* (Edgar 1994:10) »

Extrait de Maynard (1996, p.115)

Cette nouvelle, bien que neutre, peut susciter des réactions négatives de la part de sa famille, si elle est homophobe. On peut alors imaginer qu'il y a effectivement, dans la voix du locuteur qui s'apprête à annoncer la nouvelle, quelque chose qui indique la nature importante et sérieuse de la nouvelle malgré le fait que celle-ci n'est à aucun moment directement qualifiée d'importante (nous trouvons tout de même une formulation qui laisse présager du sérieux de l'annonce : « *There is something I need to tell you* »). C'est sur cet aspect du « ton » qui laisse présager de la nature de la nouvelle que ce sont penchés Freese et Maynard (1998). Ils considèrent en effet que les pré-séquences et les préfaces de nouvelles peuvent attacher une valence à leur annonce de façon lexicale, comme dans les exemples proposés précédemment, ou bien avec la prosodie. Freese et Maynard (1998) affirment notamment que le fait que l'on puisse avoir une façon de parler particulière qui « sonne » comme des bonnes nouvelles ou des mauvaises nouvelles est un fait reconnu par beaucoup.

Tout au long de cette section, nous avons établi plusieurs faits et plusieurs concepts importants autour de la notion de projection en interaction. Dans la sous-section 1.1 sur les indices de projection, nous avons abordé les notions d'indices et de projection, de co-construction de l'interaction, de trajectoire, de progressivité et de préférence. Nous avons vu qu'au travers d'indices de projection, les locuteurs co-construisent le discours. Ils proposent, refusent, modifient ou acceptent des trajectoires interactionnelles qui permettent de naviguer l'interaction en assurant sa progressivité ; et en respectant la hiérarchisation des préférences, ils collaborent pour le succès de l'interaction. Dans la sous-section 1.1 **Erreur ! Source du renvoi introuvable.** sur les dispositifs de structuration du discours, nous avons expliqué le découpage en séquences, paires-minimales, et tours de parole de l'interaction. Nous avons défini les pré-séquences et ses sous-types (pré-invitation, pré-demande, pré-annonce) et expliqué leur utilité dans les mécanismes de protection de la face des interactants, d'alignement et d'affiliation. Dans la sous-section 1.2 sur les préfaces et marqueurs discursifs, nous avons fait une différence entre pré-séquences et préfaces, et défini les marqueurs discursifs comme des unités polyvalentes de structuration du discours, pouvant aussi faire office de préface. Dans la sous-section 1.3 sur la projection d'indices de la valeur affective d'une séquence, nous avons étudié l'utilité spécifique des pré-séquences et préfaces dans le cadre d'annonces de nouvelles et montré que ces espaces interactionnels servent à projeter des indices sur la nouvelle à annoncer, mais aussi des indices sur sa valeur affective. Nous avons évoqué le fait que cela peut être opéré au moyen du vocabulaire employé, mais aussi au niveau du « ton de la voix ». Nous allons donc maintenant passer à la seconde section de cet état de l'art, dédiée à la notion de prosodie, à ses fonctions, et aux canaux qui permettent de la produire.

2. Prosodie affective : phonétique, phonologie, ou les deux ?

2.1. Sens linguistique *versus* sens paralinguistique

Arvaniti (2020) propose une définition complète de la « prosodie » comme un terme « parapluie » englobant divers phénomènes linguistiques tels que le stress, le rythme, le phrasé et l'intonation, qui tous contribuent à modeler le sens et la perception d'une énonciation. Le « *stress* » peut être appliqué au niveau du mot, engendrant des propriétés lexicales spécifiques, ou au niveau de l'énoncé dans son ensemble. Il se réfère à la proéminence perçue de certaines syllabes par rapport aux autres. Cette proéminence des syllabes peut être créée à travers diverses caractéristiques acoustiques, lesquelles varient d'une langue à l'autre. Arvaniti (2020) définit le « rythme » (« *rhythm* ») comme un terme controversé qui renvoie à la perception de régularités et de schémas rythmiques dans la parole, souvent confondu avec le concept de « *timing* », qui correspond à la régularité temporelle effective d'une unité donnée. Pour elle, le rythme, comme le stress, est un phénomène perceptuel qui est porté par différentes caractéristiques acoustiques. Le terme « phrasé » (« *phrasing* ») est utilisé pour décrire la segmentation naturelle de la parole en unités distinctes qui s'agencent ensemble, facilitant ainsi l'expression pour le locuteur et la compréhension pour l'auditeur. En effet, le discours oral ne se présente pas comme un flux ininterrompu d'entités linguistiques isolées les unes des autres, mais plutôt comme une série d'éléments interconnectés. « Intonation », enfin, est un terme utilisé pour faire référence à des modulations systématiques de la fréquence fondamentale (F0). Ces modulations couvrent l'ensemble de l'énoncé et possèdent des fonctions grammaticales cruciales, telles que l'encodage de l'information pragmatique et le marquage des frontières entre les phrases. En somme, la prosodie, constituée de ses composantes multiples et interdépendantes, contribue fortement à guider la manière dont nous comprenons et interagissons les uns avec les autres en interaction.

La prosodie occupe une place cruciale dans la parole, remplissant deux fonctions majeures : celle de véhiculer le « sens linguistique » et le « sens paralinguistique ». Avant de décrire les corrélats de la prosodie dans la parole, il est nécessaire de faire un point sur ces deux concepts fondamentaux. Nous allons donc commencer par les définir. Pour Laver (1994), le « sens linguistique » se manifeste à travers divers aspects de la parole, tels que la phonologie segmentale, la phonologie suprasegmentale, la morphologie et la syntaxe. Au travers de la communication verbale, le sens linguistique porte ce que le locuteur souhaite communiquer à son interlocuteur. Il se réfère à la signification des mots, des phrases et des discours. Le sens linguistique dans la prosodie, lorsqu'il se déploie au niveau suprasegmental, englobe les indices que les locuteurs emploient pour structurer leur discours, préciser des nuances de sens des formes verbales, en ajouter ou même le modifier.

Au-delà des indices verbaux du sens linguistique, il y a des indices gestuels, faciaux et prosodiques du sens linguistique, comme l'illustre les recherches sur la prosodie audiovisuelle. Plusieurs chercheurs parlent donc de prosodie « multimodale ». Ces études explorent comment les éléments prosodiques sont transmis et perçus à travers à la fois l'ouïe et la vue, en exploitant les indices visuels tels que les gestes, les mouvements des lèvres et les expressions faciales pour enrichir la compréhension des aspects prosodiques de la parole (Barkhuysen et al., 2005; Brown & Prieto, 2021; Krahmer & Swerts, 2009). La prosodie linguistique joue un rôle pour la segmentation des mots (Cutler et al., 1986; Cutler & Norris, 1988; Mattys et al., 1999), la structuration du discours (Grosz & Sidner, 1986; Nakatani et al., 1995), la désambiguïsation syntaxique (Klatt, 1975; Lehiste, 1973; P. J. Price et al., 1991), la focalisation contrastive (Hamblin, 1973; Rooth, 1985, 1992), et elle joue aussi un rôle dans la construction de la hiérarchie la syntaxe et de l'information (Jackendoff, 1997; Nespors & Vogel, 1986; Selkirk, 1984).

Laver (1999) définit le sens paralinguistique comme ayant des rôles de support, et de complétion ou contradiction du langage verbal. La communication paralinguistique se manifeste au sein du discours de multiples façons : en tant que substitut à la communication verbale, en renforçant certains éléments verbaux, en véhiculant des informations affectives telles que les attitudes et les émotions, en consolidant la nature des relations sociales entre les interlocuteurs, en orchestrant la répartition du temps alloué à chaque tour de parole et en gérant les différentes phases de l'interaction, que ce soient les séquences d'ouverture, les séquences intermédiaires ou les séquences de

clôture. Les informations paralinguistiques incluent aussi des informations sur le sexe ou le genre de la personne (Pernet & Belin, 2012), le statut social de la personne qui s'exprime ou de celle à laquelle on s'adresse (Brown & Prieto, 2017; Elordieta & Prieto, 2012), ou bien sa provenance géographique régionale ou nationale (E. Grabe et al., 2004). Ces informations peuvent être portées, entre autres, par la prosodie, en parallèle de l'utilisation d'autres marqueurs paralinguistiques, par exemple lexicaux, segmentaux, ou syntaxiques. Le sens paralinguistique fournit donc une dimension supplémentaire à la communication en reflétant les émotions, les rapports sociaux et les nuances culturelles des locuteurs, tout en contribuant activement à la gestion des interactions pour créer un échange fluide.

Le sens paralinguistique tel qu'il est porté par la prosodie a fait l'objet de multiples études. Certaines études se sont attardées sur le thème des émotions (Audibert, 2008; Banse & Scherer, 1996; Bänziger et al., 2001; Scherer, 2003), de l'autorité et de la politesse (Borràs-Comes et al., 2015; Jeong & Potts, 2016; Nadeu & Prieto, 2011; Petrone et al., 2019), du comportement social et de l'alignement (Levitan et al., 2012), de « *l'interpersonal stance* » ou posture interpersonnelle (cf. étude sur les relations de pouvoir et de solidarité dans la prosodie de Wichmann (2011)). D'autres études se sont concentrées sur l'ironie (Bryant, 2010; Deliens et al., 2018; Mauchand et al., 2020; Mazzocconi et al., 2022; Rivière et al., 2018) ou la politesse (Brown & Prieto, 2017; Caballero et al., 2018). Nous développerons les différences entre les domaines d'études liés à des notions d'affect comme les émotions, les attitudes, ou la posture interpersonnelle en section 3.1. Au travers de ces définitions, il apparaît que la prosodie, puisqu'elle peut porter des indications sur la valeur affective d'un message grâce à sa fonction paralinguistique, pourrait aussi être utilisée pour projeter la valeur affective d'un message à venir. Nous allons maintenant nous intéresser aux canaux qui permettent de véhiculer les informations dans la prosodie : la phonétique, et la phonologie.

2.2. Distinction entre phonétique et phonologie

Au cours des dernières décennies, les chercheurs en prosodie ont étudié comment les significations linguistiques et paralinguistiques sont portées par la prosodie, ainsi que la façon dont elle est perçue par les auditeurs. Ces études ont observé la réalisation des nuances prosodiques à travers deux canaux qui peuvent porter la prosodie, et qui sont fortement corrélés : la « phonétique » et la « phonologie ». La phonétique est la branche de la linguistique qui étudie les sons de la parole, à la fois dans leur production (comment les sons sont produits par le locuteur) et dans leur perception (comment les sons sont perçus par l'auditeur). La phonétique examine les caractéristiques acoustiques, articulatoires et auditives des sons du langage. La phonologie est la branche de la linguistique qui étudie la manière dont les sons du langage fonctionnent à l'intérieur de systèmes linguistiques spécifiques (les langues) pour créer des distinctions de sens. Ces distinctions seraient portées par des unités minimales de sens phonologiques.

Dans un premier temps, nous allons présenter les corrélats phonétiques de la prosodie, qui correspondent à des variables continues mesurées dans la voix. Ces variables font varier la perception de quatre composantes de la voix :

- La fréquence fondamentale (f_0), qui influence la perception du pitch
- L'intensité, qui influence la perception du niveau sonore
- La durée, qui influence la perception du tempo
- Les caractéristiques spectrales, qui influencent notamment la perception de la qualité de la voix, des voyelles, de la proéminence

La « fréquence fondamentale », qui est une mesure de la hauteur de la voix, est généralement abrégée en « f_0 ». Au niveau acoustique, elle correspond à la fréquence la plus basse d'un son périodique, c'est-à-dire la fréquence à laquelle les cycles de vibration des plis vocaux se répètent régulièrement. Dans la tradition britannique, le terme de « pitch » ou de f_0 sont parfois utilisés indifféremment pour parler de la fréquence fondamentale. Cependant, le terme pitch est en général utilisé pour désigner la hauteur de la voix perçue par l'oreille humaine (Cruttenden, 1997). Bien que cette perception soit

en grande partie influencée par les variations de la f_0 , d'autres caractéristiques acoustiques entrent en compte, comme la variation de l'intensité (Thompson et al., 2012). La f_0 est mesurée en Hertz (Hz). Nous utiliserons la dénomination f_0 lorsque nous parlerons des analyses acoustiques réalisées, et le terme « pitch » lorsque nous évoquerons la perception et pour la notion d'intonation décrite plus bas.

Nous avons ensuite « l'intensité », qui correspond à la force mise dans la voix, à son énergie. Elle est mesurée en décibels (dB). L'intensité est déterminée par l'amplitude des vibrations sonores. Plus l'amplitude est grande, plus le son est fort, et cela signifie que l'intensité dépend de la quantité d'énergie sonore transmise. L'intensité a été étudiée pour les liens qu'elle pourrait avoir avec la notion de proéminence, les études montrant en général qu'elle n'a que peu d'influence sur la proéminence ou le *stress* (Fry, 1955; Ladd & Arvaniti, 2023; Mol & Uhlenbeck, 1955).

Les mesures de « durées » correspondent à la durée de différents segments de la parole (durée du phonème, de la syllabe, du mot ou de l'énoncé, ou même durée des pauses dans la parole). A partir de la durée, il est possible de faire une mesure de la vitesse à laquelle on parle, et elle se mesure généralement en nombre de syllabes par seconde (syll/sec) : c'est la vitesse de parole, si les pauses sont incluses, ou le débit articulatoire, si les pauses sont exclues. Il est aussi possible, dans la mesure des pauses, de différencier pauses vides, composées de silence, et pauses pleines, composées de voix (lorsque le locuteur prononce des « hm » ou des « euh », par exemple). Les mesures de durées peuvent parfois être intégrées dans la notion de « rythme » ou d'isochronie qui veut que certaines langues aient un rythme basé sur la syllabe, le *stress*, ou la *more* (Arvaniti, 2020; D. R. Ladd & Arvaniti, 2023; Lancia et al., 2019; Tilsen, 2019; Turk & Shattuck-Hufnagel, 2013). Suivant Turk et Shattuck-Hufnagel (2013), nous les englobons donc plutôt dans les notions de durée sous la dénomination de « caractéristiques temporelles apparentes de la parole » (« *surface speech timing patterns* »).

Nous avons, enfin, les « caractéristiques spectrales » de la voix, qui influencent, entre autres, la façon dont l'on perçoit la « qualité » de la voix de quelqu'un, la proéminence, et la qualité des voyelles. Les caractéristiques spectrales sont liées, acoustiquement, à la manière dont l'énergie sonore est distribuée sur différentes fréquences vocales, au sein des harmoniques, des formants, et du spectre sonore. Campbell et Mokhtari (2004) décrivent la qualité de la voix, liée aux caractéristiques spectrales, comme la quatrième dimension prosodique qui pourrait elle aussi porter des informations sociales et

paralinguistiques. De nombreuses études se sont intéressées à cette notion de qualité de la voix et à ce qu'elle apporte pour la perception des affects (Beňuš & Rusko, 2009). Pour cette raison, nous nous intéresserons aux caractéristiques spectrales dans le cadre de la modification de qualité de la voix. Ces caractéristiques permettent de différencier le type de voix, qui peut être modal (voix normale), soufflé (« *breathy* », lorsque l'on émet de l'air de façon audible en même temps que la vibration des plis vocaux), craqué (« *creaky* », lorsque les plis vocaux vibrent lentement et de façon irrégulière), nasal (lorsque l'air passe aussi par la cavité nasale), clair (lorsque aucun son produit par l'air sortant ne vient à s'ajouter à celui des plis vocaux) ou encore, il est possible de faire une « voix de crécelle » (« *vocal fry* », si les plis vocaux vibrent plus lentement que pour la voix soufflée et que le registre de la voix est très bas). De multiples paramètres ont été proposés pour calculer la qualité de la voix, et aucun consensus total n'a été atteint jusqu'à aujourd'hui. En effet, selon les raisons qui nous poussent à nous intéresser à la qualité de la voix (voix pathologiques, étude de l'émotion), et selon la longueur de l'enregistrement ou bien sa qualité, il faudra choisir le paramètre en fonction. Nous reviendrons sur ces paramètres de la qualité de la voix plus tard dans cet état de l'art.

La prosodie comprend, enfin, l'étude de l'intonation. La f_0 est l'indice acoustique principal de l'intonation, et celui sur lequel porte majoritairement notre attention dans ce document. L'intonation est liée à la manière dont les sons sont organisés en termes de hauteur, de montées et de descentes dans la mélodie. Elle peut être étudiée du point de vue de la phonétique, mais aussi de la phonologie. C'est sur ce point que nous porterons notre attention dans les lignes qui suivent. La phonologie de l'intonation correspond à des variables discrètes que nous décrivons comme des mouvements tonals, ou des contours prosodiques. On considère que ces contours pourraient porter du sens dans le discours. Au cours des dernières décennies, les chercheurs en phonologie intonationnelle ont œuvré en direction d'un but commun : parvenir à un accord sur la manière la plus adaptée de décrire le phénomène. Ainsi, de nombreuses théories ont vu le jour depuis que la recherche s'est penchée sur la question de la phonologie de la prosodie.

Portes (2004) fait un état des lieux de deux approches de la phonologie de l'intonation, en se focalisant sur le français. D'une part, il y a les approches « holistiques », dans lesquelles on considère que les contours intonatifs décrits ne peuvent être découpés en unités porteuses de sens plus petites, et où ils sont souvent appelés « intonèmes », de la même façon que l'on parle de phonèmes ou de morphèmes. En français, ces intonèmes

ont été décrits par Delattre (1966) et Rossi (2001). Ils sont représentés sur des portées similaires à celles que l'on trouve dans la musique. On trouve par exemple des intonèmes qui commencent avec une montée mélodique puis descendent, et qui seraient souvent utilisés pour indiquer une question simple (question "oui/non"), ou encore des intonèmes descendant qui signifient que l'énoncé produit est déclaratif. D'autre part, il y a les approches « autosegmentales » dans lesquelles on considère qu'il est possible de découper les contours prosodiques en unités encore plus petites, c'est-à-dire en tons, qui peuvent être produits sur les syllabes. C'est à ce deuxième type d'approche que nous nous intéressons ici.

C'est Goldsmith (1976) qui dans sa thèse propose la théorie autosegmentale. Cette théorie vise à expliquer le fait que certains traits linguistiques (comme le ton ou l'accentuation) ne sont pas liés de manière linéaire aux segments (lexicaux), mais sont représentés en tant « qu'autosegments » distincts. Ces autosegments peuvent alors être associés entre les niveaux segmental (lexical) et suprasegmental (intonationnel) à l'aide de traits : visuellement, on créera par exemple une tire (un niveau visuel de représentation) pour le mot « *balloon* », et une tire pour l'intonation. Sur la tire de l'intonation, on pourra annoter des tons hauts H ou bas L. Les tons hauts correspondent à la cible haute d'une montée intonative, et les tons bas correspondent à la cible basse d'une descente intonative. Entre les deux tires, on ajoutera des traits pour montrer à quels segments ils sont associés entre eux sur les deux niveaux. Deux règles régissent ces associations : toutes les syllabes sont associées à au moins un ton, et *vice versa*, et les traits d'association ne peuvent pas se croiser.

Lieberman (1975) et Liberman et Prince (1977), furent les premiers à proposer l'existence d'une organisation hiérarchique « métrique » au sein de la phonologie pour l'intonation. Liberman (1975) propose dans sa thèse de s'intéresser à ce qu'est l'intonation exactement, et notamment à son statut dans la grammaire. Il veut savoir comment représenter l'intonation, et où la représenter. Pour cela il l'étudie de la même façon que les autres phénomènes linguistiques ont été étudiés jusqu'ici : il cherche à montrer qu'il existe des formes sous-jacentes et des formes de surface de l'intonation. Il veut observer la « dérivation », le lien entre forme sous-jacente et forme de surface.

Son idée est donc que, comme pour les phonèmes où nous avons une forme sous-jacente qui porte du sens (mettons, le phonème /i/ en français) et des formes de surface qui ne changeront pas le sens (/i/ peut être produit [i], ou [i:] par exemple), nous devrions

pouvoir trouver l'équivalent pour l'intonation. Il observe pour cela le cas des « *tune differences* » ou différences de réglages (variations dans le contour de la f0), qui pourraient modifier le sens d'une même phrase : selon le contour intonatif, produire une affirmation, une question, ou démontrer de l'incrédulité. Il aborde également le cas des « *phrasing differences* » ou différences de phrasé qu'il lie aux pauses, virgules, ou « breaks » intonationnels, et qui modifient là encore le sens d'une phrase. Pour prendre un exemple commun donné pour faire comprendre l'importance de la virgule à l'écrit, les « breaks » ou pauses font la différence entre « On mange les enfants ? » et « On mange, les enfants ? ». Dans le second exemple, à l'oral, il y aurait un break intonational. Ces éléments lui permettent de montrer qu'il existe des caractéristiques tonales (en anglais) qui peuvent porter du sens indépendamment du texte auquel elles sont associées, de façon ordonnée. Ces tons peuvent être hauts ou bas et sont « associés » au niveau lexical en suivant des règles linguistiques précises.

Lieberman et Prince (1977) proposent l'utilisation d'une grille métrique sur laquelle on projetterait la proéminence de syllabes qui portent des poids différents les unes par rapport aux autres (syllabe forte « s », syllabe faible « w »), permettant ainsi l'organisation rythmique des syllabes, mots, et phrases syntaxiques. Le concept de grilles métriques est alors un moyen d'analyser la structure de l'intonation. Elles sont des représentations visuelles utilisées pour capturer les schémas rythmiques et temporels de la parole, et permettent de décrire l'organisation des syllabes et des tons dans un énoncé, et se présentent comme dans l'exemple ci-dessous. Sur le mot « *balloon* », la première syllabe est faible et la seconde, forte.

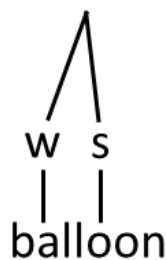


Figure 4 – Représentation du poids des syllabes, extrait de Liberman et Prince (1977)

Par ces représentations en grilles métriques, Liberman et Prince (1977) cherchent à proposer une théorie dans laquelle on considère que le niveau phonétique et phonologique n'est pas la concaténation pure et simple de multiples éléments de base. Au

contraire, il y a une hiérarchie qui régit le *stress* dans la parole avec un système de proéminence relative des syllabes les unes par rapport aux autres. Selkirk (1980, 1984), puis Nespor et Vogel (1986, 2007) reprennent aussi ces travaux pour la théorie de la phonologie prosodique afin d'analyser la manière dont le flux de la parole est organisé en unités phonologiques mises ensembles. Elles proposent de considérer pour étudier ces unités les interactions entre la phonologie et les autres composantes de la parole sous forme de règles qui permettent de construire une « structure phonologique ». Elles proposent notamment de hiérarchiser la prosodie en trois niveaux : un niveau pour la syllabe, le pied et le mot phonologique, un pour les groupes clitiques et la phrase phonologique, et un pour le syntagme intonational et l'énoncé phonologique.

La théorie autosegmentale et de la théorie métrique proposent donc une approche non linéaire, hiérarchique, de l'étude de l'intonation. D'une part, le niveau de lexical et le niveau intonational sont considérés distincts : segmental, et suprasegmental. Il existe entre ces niveaux des liens d'associations régis par des règles précises. D'autre part, il existe un système de proéminence relative des syllabes les unes par rapport aux autres. Ces proéminences sont liées aux *stress* lexicaux, mais aussi aux *breaks* ou « pauses » que l'on peut percevoir dans la parole. La proéminence est changeante, et peut modifier le sens de l'énoncé. Enfin, il semble y avoir une hiérarchie, elle aussi sur plusieurs niveaux, allant du niveau de la phrase phonologique à celui de la syllabe. Là encore, ces niveaux peuvent être associés entre eux différemment, créant des distinctions de sens dans l'énoncé. De ces différentes théories est née la proposition d'une nouvelle théorie : la théorie métrique autosegmentale, que nous décrivons dans la prochaine section.

2.3. Théorie AM et Phonologie Intonationale

Il y a 40 ans environ, la « théorie métrique autosegmentale » (abrégée en théorie AM, de l'anglais *autosegmental metrical theory*) a été développée. Plusieurs chercheurs ont contribué au développement de cette théorie, en travaillant sur diverses langues (Bruce, 1977; Gussenhoven, 1984; D. R. Ladd, 2008; M. Y. Liberman, 1975; J. B. Pierrehumbert, 1980; J. Pierrehumbert & Beckman, 1988). La théorie AM a été développée pour proposer un modèle de la structure de l'intonation, qui prend en considération le principe de

compositionalité de la mélodie de l'intonation, afin de rendre compte du rythme de la parole, de notions liées au *stress*, des mouvements de la courbe de la f0, et afin de pouvoir comparer l'intonation entre différentes langues. De même que dans les théories métrique et segmentale, l'AM ne considère pas l'intonation comme un simple flux continu que l'on ne peut découper en unités plus petites et que l'on peut représenter de façon impressionniste. Au contraire, l'intonation est « composée » d'évènements qui, liés ensemble, forment des contours mélodiques.

Ces « évènements » sont des tons, qui comme dans la théorie autosegmentale peuvent être hauts ou bas (*high* H ou *low* L) et sont considérés comme des autosegments (indépendants du segmental). Ils sont annotés en série et peuvent être combinés ensembles (Arvaniti, 2022; Ladd, 2008; Pierrehumbert, 1980; Pierrehumbert & Beckman, 1988). Arvaniti (2022) décrit les tons comme phonologiquement associés à des positions structurelles dans la représentation métrique d'un énoncé. Elle précise que cette association entre des propriétés métriques et autosegmentales a justement donné son nom à la théorie, et permet de représenter l'intonation avec des tons autosegmentaux qui s'associent métriquement à des constituants spécifiques de la phrase intonative (« *pitch accent* », noté T*) ainsi qu'aux frontières de celle-ci (« *edge tone* », noté T- ou T% selon la force de la frontière). Dans ce travail, nous utiliserons « accent mélodique » pour désigner les « *pitch accents* » (Bertrand et al., 2022). Les tons transcrits peuvent être associés entre eux et former un groupement de tons (signifié visuellement par un +), appelé accent mélodique bitonal. Un accent bitonal sera par exemple marqué L+H. Pour indiquer un « *downstep* » ou abaissement progressif de la f0 sur un ton, la diacritique « ! » est utilisée. Par exemple, si nous trouvons deux tons H consécutifs et que le second est abaissé, cela sera noté H !H comme dans l'exemple de Ladd (2008) ci-dessous, qui présente une transcription phonologique d'un exemple en français de contour phonologique portant un sens d'incrédulité ou de déception, et propose une transcription suivant les règles AM.

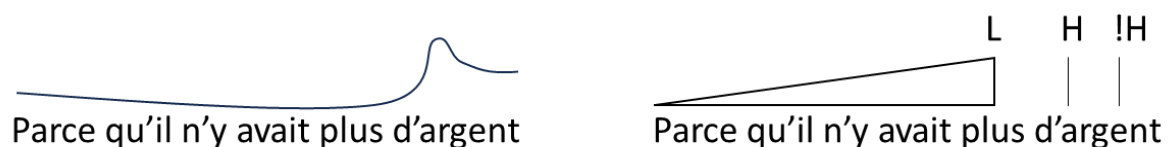


Figure 5 – Exemple de représentation des tons dans la théorie AM (à droite), d'après Ladd (2008)

La théorie AM de l'intonation repose sur quatre principes (Ladd, 2008). Tout d'abord, il y a l'acceptation de l'existence d'une structure tonale séquentielle dans l'intonation. Cela signifie que les contours mélodiques (variations de la f_0) sont organisés de manière séquentielle : il y a, le long du contour de la f_0 , des moments clés localisés qui sont appelés « cibles tonales », et peuvent être vus comme la réalisation d'objets phonologiques appelés « tons » (Arvaniti et al., 1998). Ensembles, les cibles tonales permettent de réaliser la structure tonale d'un patron phonologique. Entre les tons, le contour est considéré comme « non spécifié », et l'espace entre deux tons est considéré comme une transition. Les tons sont associés à certains points spécifiques du niveau segmental de l'énoncé.

Ensuite, il y a la distinction entre les accents mélodiques et le *stress* (ou accentuation). D'une part on trouve les accents mélodiques, qui sont des caractéristiques intonationnelles associées avec certaines syllabes selon divers principes de l'organisation prosodique. D'autre part on trouve le *stress* qui serait lié à la force métrique et créerait une perception de certaines syllabes accentuées comme étant plus proéminentes de d'autres (« PERmit » vs « perMIT »).

Puis, il y a l'analyse des accents mélodiques comme cibles de hauteur. Ces cibles peuvent être hautes (ton *high*, abrégé H), ou elles peuvent être basses (ton *low*, abrégé L). Elles sont jugées comme telles en fonction notamment de différences relatives de hauteur entre elles et vis-à-vis de l'empan vocal du locuteur.

Enfin, il y a les notions de « valeur mélodique » (« *scaling* ») et « d'alignement », qui ensemble définissent les cibles tonales (H ou L). La valeur mélodique correspond à l'implémentation d'un ton, H ou L, à différentes hauteurs : sa réalisation phonétique varie en fonction de cette hauteur. L'implication de la valeur mélodique dans la représentation phonologique a notamment été l'objet de plusieurs études (Lieberman & Pierrehumbert, 2005; Prieto, 1998, 2003; Vanrell, 2007). L'alignement est lié à la notion d'association, qui désigne le placement dans le temps d'une cible tonale au sein de l'unité porteuse de ton à laquelle il appartient : par exemple, au sein de l'énoncé « les pots de fleurs », s'il y a un accent tonal H, il pourra être associé à différentes syllabes, voyelles, où frontières de l'énoncé. Pour Pierrehumbert (1980), l'association du ton est phonologique et peut créer des différences entre plusieurs accents bitonaux qui ne porteront pas le même sens (en anglais américain : H^*+L , $H+L^*$ pour les accents bitonaux descendants, et L^*+H , et $L+H^*$

pour les accents bitonaux montants). Le placement de l'étoile sur le ton associé à la syllabe accentuée dans l'accent fait la différence entre descente ou montée « précoce » (« *early fall* », « *early rise* ») ou « tardive » (« *late fall* », « *late rise* ») (Pierrehumbert & Beckman, 1988). A cette « association primaire » s'ajoute une « association secondaire » (Grice, 1995; Pierrehumbert & Beckman, 1988) ou « ancrage phonologique » (Prieto et al., 2005). Elle intervient lorsqu'un ton est associé non seulement à l'unité porteuse de ton, mais aussi à une frontière de more, de syllabe, ou de mot (ou unité) prosodique. Le ton est alors sujet à une double association. La réalisation du ton en cible tonale peut être alignée différemment à l'intérieur ou en décalage avec la syllabe à laquelle le ton est associé. L'alignement peut être influencé par de multiples facteurs tels que la vitesse de parole, et la longueur des voyelles, ainsi que la structure de la syllabe à laquelle est rattaché le ton. L'implication de l'alignement dans la représentation phonologique a, comme la valeur mélodique, été l'objet de plusieurs études (Kohler, 1987; Niebuhr, 2007; Niebuhr & Kohler, 2004; Pierrehumbert, 1980; Pierrehumbert & Steele, 2009).

Ainsi, un objectif de la théorie AM est de fournir un cadre adapté à la description et à la comparaison de l'intonation dans plusieurs langues, afin de mieux rendre compte de sa signification linguistique. Ladd (2008) propose le terme de phonologie intonationnelle, qui désigne donc l'étude d'une grammaire intonationnelle dont le but est de décrire des modèles prosodiques, à la recherche de significations systématisées attachées à ces modèles (Gussenhoven, 2002). Suivant les différentes théories présentées jusqu'ici, nous considérerons l'intonation comme un véhicule du sens dans la parole dans notre travail. L'intonation est un paramètre sécable en unités de sens que la recherche en phonologie intonationnelle cherche à décrire et comprendre. Nous nous intéresserons également aux variations de valeur mélodique et d'alignement que nous pourrions trouver dans la réalisation des cibles tonales. Nous allons maintenant porter notre attention sur la façon dont la phonologie intonationnelle est annotée en français, en évoquant notamment la convention d'annotation « ToBI ».

2.4. Annotation phonologique de l'intonation du français

L'intonation peut être découpée en unités intonationnelles hiérarchisées entre elles. En fonction des langues, le nombre d'unités proposées diffère (Ladd, 2008). En français, selon les articles, on considère qu'il existe deux ou trois niveaux d'unités prosodiques. Deux de ces unités sont majoritairement acceptées dans la littérature (cf Petrone et al., 2021 pour un état de l'art) : le syntagme accentuel, et le syntagme intonational.

Le syntagme accentuel (AP, « *accentual phrase* ») est considéré comme l'unité de base du français. Au niveau lexical, l'AP est constituée d'au moins un mot. Elle est le domaine où l'on retrouve l'accent mélodique (Beckman & Pierrehumbert, 1986). Cet accent mélodique fait partie en français de la montée finale obligatoire, aussi appelée accent primaire ou accent final. La montée finale marque la frontière droite de l'AP. Le deuxième constituant de cette unité de base qu'est l'AP est la montée initiale, aussi appelée accent secondaire ou accent initial. La montée initiale marque la frontière gauche du constituant (German & D'Imperio, 2015; Jun & Fougeron, 2000, 2002; Petrone et al., 2021; Welby, 2006).

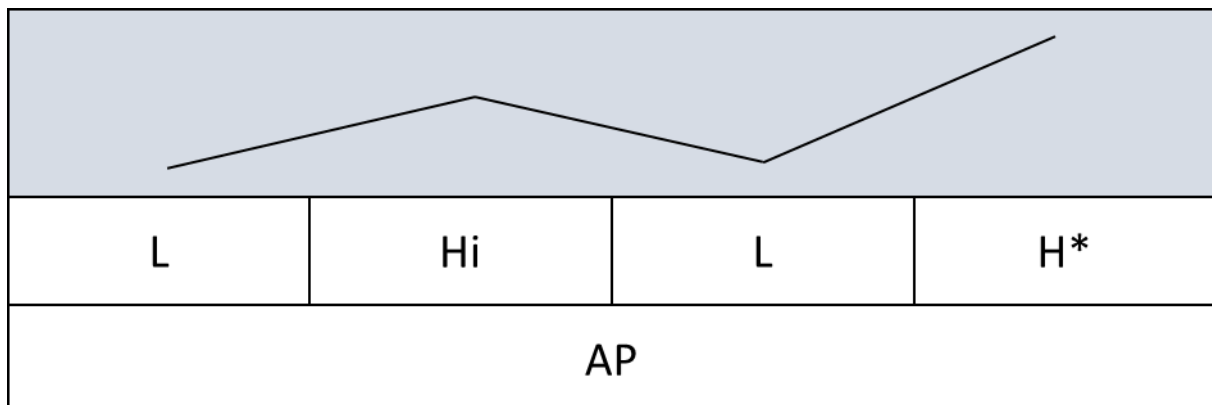


Figure 6 – Patron tonal sous-jacent du français

L'AP est donc constituée de deux montées intonatives : une montée initiale optionnelle, et une montée finale obligatoire qui contient l'accent mélodique. Nous allons maintenant exposer la transcription de ces montées selon Jun et Fougeron (2000, 2002), qui ont décrit

la phonologie intonationnelle du français en suivant la méthodologie de la théorie AM, et pour qui le patron intonational sous-jacent pour le français est /LHiLH*/.

La montée finale est constituée d'un accent bitonal LH*, qui est composé d'un ton L bas réalisé dans l'une des deux dernières syllabes de l'AP, et de l'accent tonal. L'accent tonal est réalisé, dans ce patron sous-jacent, en ton haut H*. Il est associé à la frontière droite de l'AP. En français, l'accent tonal est lié à la dernière syllabe d'un mot, en général. Il est cependant réalisé uniquement si ce mot est en fin de syntagme accentuel. Le ton H* est donc non seulement associé à la frontière droite, mais aussi à la syllabe accentuée du dernier mot de l'AP. La montée initiale est constituée d'un accent bitonal LHi. Il est constitué d'un ton L bas, associé à la frontière gauche de l'AP, et à la première syllabe du premier mot de contenu de celle-ci. Puis il y a le ton Hi, qui peut être sur la première ou la deuxième syllabe de ce mot de contenu. La présence de la montée initiale est conditionnée à plusieurs facteurs dont la longueur de l'AP qui le contient, la composition segmentale et la vitesse d'articulation du locuteur (German & D'Imperio, 2015; Jun & Fougeron, 2002; Welby, 2006).

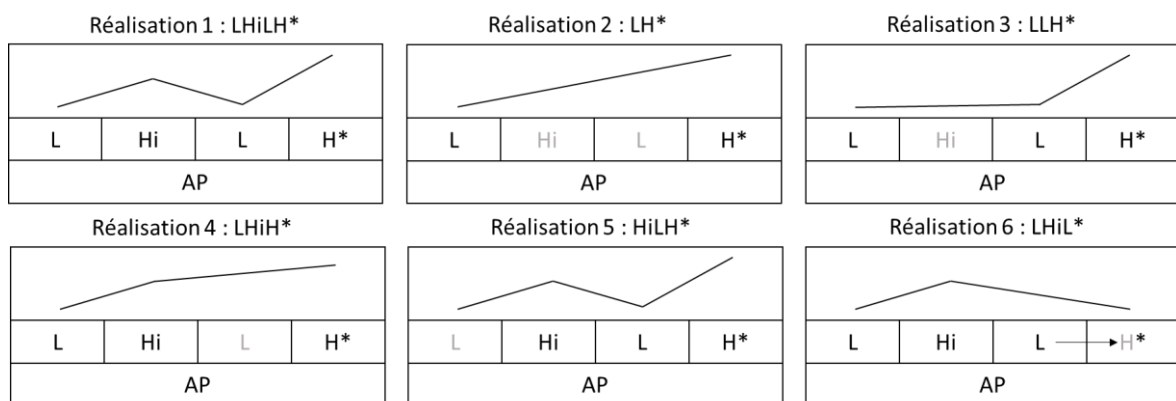


Figure 7 – Les cinq réalisations possibles du patron tonal sous-jacent, d'après Jun et Fougeron (2000, 2002)

Le patron sous-jacent /LHiLH*/ peut être réalisé de multiples manières : LHiLH* (tel quel), LH* (les tons du milieu de l'AP ne sont pas réalisés), comme LLH* (Hi n'est pas réalisé), LHiH* (le L du milieu de l'AP n'est pas réalisé), ou encore HiLH* (sans le L initial). Ces différents patrons sont proposés dans la Figure 7 : les tons non réalisés sont grisés, la réalisation d'un ton L* à la place de H* est signifiée par une flèche. Jun et Fougeron (2002) estiment que si H* ne peut pas être réalisé en raison de certaines contraintes, il pourra

alors exceptionnellement être réalisé en tant que L*, créant ainsi un patron LHiL*. Cela arrive lorsque l'accent tonal est immédiatement précédé de Hi, et suivi d'un autre Hi dans l'AP suivante, ce qui démontre que dans l'intonation, le fait de produire trois tons H consécutifs est évité. Jun et Fougeron (2002) nomment cette contrainte « *avoid HHH* ».

Les différents patrons présentés peuvent être bordés par des tons de frontières. Ces tons sont associés aux deux autres niveaux d'unités prosodiques que nous n'avons pas encore abordées. Le syntagme intonational (IP, de l'anglais *intonational phrase*) est l'unité la plus large, composée d'au minimum une AP. Selkirk (1984) la décrit comme une unité qui correspond à une portion de phrase associée à une mélodie ou un contour intonational caractéristique. Elle est généralement caractérisée par un allongement de la dernière voyelle qui la compose (Jun & Fougeron, 2002). Le ton de frontière qui marque sa frontière droite est situé sur la dernière syllabe de l'IP, et est distingué comme suit : L% ou H%. La dernière syllabe de l'IP est aussi la dernière syllabe de la dernière des AP qui la compose. En conséquence, l'accent tonal de la dernière AP est situé sur la même syllabe que le ton de frontière. Le dernier niveau hiérarchique est un niveau intermédiaire entre l'AP et l'IP, appelé justement le syntagme intermédiaire (ip, de l'anglais *intermediate phrase*). Elle contient au moins une syllabe accentuée, et au moins une AP. Son existence en français, bien qu'encore controversée, a été argumentée et montrée notamment par des indices de durée (un allongement de la dernière voyelle dans la dernière syllabe) significativement plus longs que dans l'AP, mais plus courts que dans l'IP (Michelas & D'Imperio, 2010, 2009). La structure hiérarchique des constituants de la phrase prosodique dans l'étude de l'intonation du français peut être présenté sous forme d'un arbre :

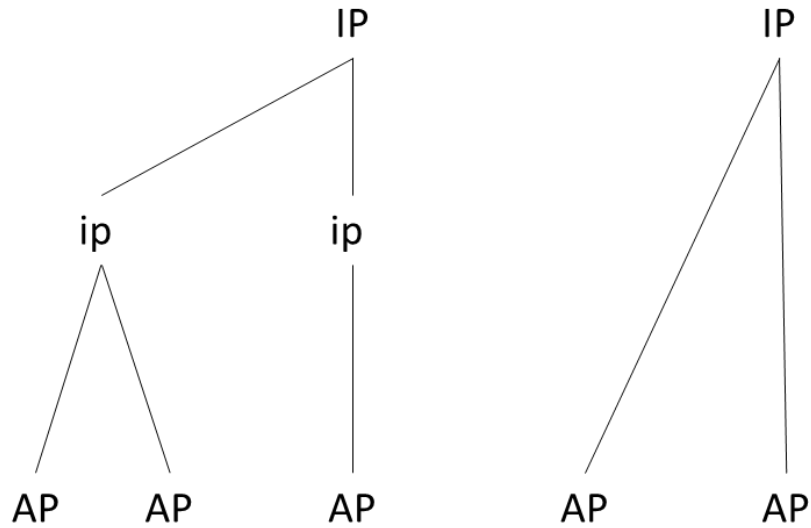


Figure 8 – Représentation schématique de la hiérarchie entre les niveaux de la phrase prosodique

A partir des idées de la phonologie intonationnelle décrite dans la section précédente (2.3), la convention d'annotation « Tones and Breaks Indices », ou ToBI, a été développée (Silverman et al., 1992). Le but de ToBI est de fournir un système de transcription pratique, adapté à partir des concepts introduits par la théorie AM. Cela permet de faciliter l'analyse et la comparaison des schémas d'intonation dans les langues parlées. Cette convention, originellement développée pour l'anglais, et spécifiquement pour une seule variété de l'anglais, est devenue un cadre général pour le développement de systèmes d'annotations prosodiques dans d'autres variétés d'anglais et dans d'autres langues (Beckman et al., 2004). Les transcriptions ToBI sont constituées de plusieurs « tires » parallèles qui permettent d'annoter plusieurs informations à un temps *T*. Tout d'abord, il y a la tire des tons, sur laquelle sont annotés les tons et groupements de tons, toujours basés sur les tons hauts annotés « H » pour *high*, et les tons bas annotés « L » pour *low*. Ensuite, il y a la tire des *break index*, où l'on annote la force des frontières entre les phrases intonationnelles et autres catégories. Enfin, il y a la tire *miscellaneous*, où l'on annote les remarques potentielles, les hésitations sur l'annotation, les disfluences. Depuis sa création, ToBI a beaucoup évolué et il existe maintenant une convention de transcription ToBI adaptée pour le français : F-ToBI (Delais-Roussarie et al., 2015). Dans cette Thèse, nous avons suivi dans nos annotations une adaptation de F-ToBI suivant les propositions de Jun et Fougeron (2000, 2002) et Welby (2006). Les détails de cette annotation seront donnés dans le chapitre 3.

2.5. Dualité de structure

Dans la cadre de la Phonologie Intonationnelle et des théories qui l'ont précédée (autosegmentale, segmentale, métrique), toute la place est donnée à l'étude du sens linguistique. Le sens paralinguistique, bien que reconnu comme existant et ne devant pas être complètement écarté, n'est que très peu mentionné : au début, puis à la fin de sa démonstration sur la phonologie intonationnelle. A l'origine de ce choix, on retrouve la notion de « dualité de structure » (Gussenhoven, 2004; Ladd, 2008; Prieto, 2015). C'est l'idée qu'il existe un fossé fondamental au sein de la prosodie concernant la cartographie du son au sens. Gussenhoven (2004) et Ladd (2008) font référence à cette division en marquant la différence entre la « moitié apprivoisée » (« *the tamed half* ») de l'intonation, qui est le sens linguistique, et la « moitié indomptée » (« *the untamed half* » ou « *savage half* ») qui est la signification paralinguistique. Dans le modèle de superposition ou *overlay model* qui a précédé la phonologie intonationnelle, Bolinger (1964), rapporté par Ladd (1988), décrit même la prosodie à travers une métaphore marine. Dans cette métaphore, les fonctions de la prosodie sont affectées à un type de vague différente en fonction de leur importance et de leur fonctionnement dans le langage. Le but de cette métaphore est notamment de démontrer la nature fluide et graduelle de la prosodie.

« The surface of the ocean responds to the forces that act upon it in movements resembling the ups and downs of the human voice. If our vision could take it all in at once, we would discern several types of motion, involving a greater and greater expanse of sea and volume of water: ripples, waves, swells, and tides. It would be more accurate to say ripples on waves on swells on tides, because each larger movement carries the smaller ones on its back [...] The ripples are the accidental changes in pitch, the irrelevant quavers. The waves are the peaks and valleys that we call accent. The swells are the separations of our discourse into its larger segments. The tides are the tides of emotion [...] The extremes—ripples and tides—are the easiest to describe and the least significant » (Bolinger, 1964, p. 282-283)

Il décrit ainsi d'une part les *ripples*, infimes changements que nous rapprochons de ce qui est en général appelé la microprosodie, et d'autres part les *tides*, marées d'émotion affectant de façon globale la hauteur du pitch ou bien l'empan vocal. Ce seraient donc des

mouvements trop graduels pour être découpés en unité de sens. Bolinger (1964) qualifie ces affirmations de faits évidents qu'il ne mentionne que pour les écarter de l'analyse, et il considère qu'ils seront probablement les mêmes dans toutes les langues. Dans cette métaphore, les affects créent donc dans la prosodie des mouvements continus qui influeraient sur l'ensemble des autres vagues, sans pouvoir être analysés comme des catégories discrètes. Les émotions ne devraient alors être abordées que par des mesures elles aussi continues. En conclusion, alors que nous transmettrions le sens linguistique à la fois par la phonétique avec des variables continues et par la phonologie avec des variables discrètes distinctes, le sens paralinguistique créerait une variation qui serait trop graduelle pour être capturée par la phonologie. Nous ne pourrions donc la porter et l'analyser que par l'observation des indices phonétiques acoustiques.

La dualité de structure est donc bien ancrée aux racines même de l'étude de l'intonation. Cette dualité bien reconnue et acceptée par la communauté des chercheurs en prosodie a durablement affecté l'étude de celle-ci. Plus précisément, elle a affecté l'étude de la « prosodie affective », l'un des champs très étudiés du sens paralinguistique. Il est possible que celle-ci ait été particulièrement affectée en raison de plusieurs facteurs. D'une part, comme nous le verrons en section 3.1, il est difficile d'étudier les « affects » car il existe de multiples catégorisations de ceux-ci, de multiples définitions et sous-catégories.

Pourtant, malgré le peu d'attention porté au sens paralinguistique dans la recherche en intonation, un fil rouge semble se détacher : toujours, les théories proposées pour rendre compte de la phonologie de l'intonation se retrouvent face à certains éléments qui ne rentrent pas tout à fait dans le cadre de la phonologie, ni tout à fait dans le cadre de la phonétique. C'est un problème que met en avant Ladd (2008) à la fin du livre *Intonational Phonology*, soulignant la difficulté constante à réellement démêler ce qui est linguistique, et ce qui est paralinguistique. La possibilité d'un intermédiaire entre linguistique et paralinguistique offre alors de nouveaux questionnements. Si la solution pour comprendre les catégories à mi-chemin entre les deux est métrique, donc hiérarchisée et phonologique, où s'arrêter ? Nous pouvons nous demander si entre linguistique et paralinguistique, il n'y a vraiment qu'un seul intermédiaire ou si, au lieu de cela, nous ne ferions pas plutôt face à un continuum entre deux extrémités, comme cela est souvent le cas dès que l'on tente de classifier un phénomène, quelle que soit la discipline. Comme nous le verrons plus tard, l'affect peut être ressenti mais aussi montré intentionnellement

en interaction. De plus, nous l'avons vu dans la première section de ce chapitre : les notions d'alignement, d'affiliation et de projection interactionnels indiquent un certain niveau de contrôle sur ce que l'on montre en termes affectifs en interaction. Si l'affect peut être intentionnellement montré en interaction pour ajouter du sens à un énoncé, et si cela peut être fait à travers la prosodie, est-il si aberrant d'imaginer que l'intonation puisse porter ce sens à travers des catégories phonologiques ? Le flou de la frontière entre sens linguistique et sens paralinguistique, de même que le flou parfois pointé entre phonétique et phonologie nous font formuler cette hypothèse : il est possible que la phonologie de l'intonation ne porte non pas seulement du sens linguistique, mais aussi du sens paralinguistique et spécifiquement, affectif.

Au cours des dernières années, de nombreuses études ont été réalisées sur les affects. La conséquence de cette notion de dualité de structure est que ces études ont toujours observé uniquement la réalisation phonétique des affects, en examinant les indices de débit de la parole, de hauteur ou bien de qualité de la voix (Batliner et al., 2000; Juslin & Laukka, 2001, 2003; Laukka et al., 2011; Scherer, 2003). Ces études, en ne s'intéressant pas à la variation phonologique, négligent potentiellement des variables qui pourraient expliquer les différences perçues de sens affectif dans la parole. Suivant Pierrehumbert (1990), nous pensons que la phonétique doit être étudiée en parallèle de la phonologie, si l'on souhaite découvrir comment le sens est porté dans la parole au travers de la prosodie.

Au cours de cette section 2. , nous avons abordé plusieurs thématiques tournant autour de la prosodie. Dans la sous-section 2.1 introductive sur le sens linguistique et le sens paralinguistique, nous avons abordé la notion de prosodie, un terme parapluie qui recouvre des notions de *stress*, rythme, phrasé, et intonation. Nous avons ensuite défini le sens linguistique et le sens paralinguistique. Dans la sous-section 0 sur la distinction entre phonétique et phonologie, nous avons listé et définit les corrélats acoustiques phonétiques couramment reconnus dans la littérature (f_0 , intensité, durée, caractéristiques spectrales), puis nous avons introduit le thème de l'intonation. Dans la sous-section 2.3 sur la théorie AM et la phonologie intonationnelle, nous avons commencé par relater le développement de la théorie autosegmentale et de la théorie métrique. Nous avons montré qu'elles ont introduit la notion d'un découpage de l'intonation en unités de sens plus petites que le « contour intonatif », proposé la description par tons hauts ou bas, démontré l'existence d'une hiérarchie entre les constituants de la phrase qui interagit avec le système de proéminence relative des syllabes ; créant des variations rythmiques dans la parole. Puis, nous avons décrit la théorie AM et ses propositions pour l'analyse de l'intonation, de mêmes que celles de la phonologie intonationnelle. Dans la sous-section 2.4 sur l'annotation phonologique du français, nous avons apporté des précisions sur la manière dont la théorie AM a été utilisée et implémentée pour l'analyse du français et de ses spécificités. Nous avons expliqué la composition du patron phonologique sous-jacent et ses réalisations et, enfin, évoqué le système de transcription F-ToBI dont une adaptation est utilisée dans ce travail. Dans la sous-section 2.5 sur la dualité de structure, nous nous sommes attardés sur cette séparation historique entre le sens linguistique qui pourrait être porté par des variations phonétiques et par la phonologie, et sur le sens paralinguistique qui lui ne pourrait être porté que par les variations continues phonétiques. Nous avons expliqué que cette dualité est remise en question, et évoqué l'intérêt que pourrait porter l'étude de la phonologie de l'intonation telle qu'elle pourrait porter le sens affectif. Nous allons donc maintenant passer à la troisième section de cet état de l'art, dédiée à la notion d'affect, de ses corrélats phonétiques et phonologiques, et à l'étude qui a inspiré ce travail.

3. Corrélats de la valence affective

3.1. Définir les types d'affects

Il existe une multitude de termes de catégorisation de ce que l'être humain ressent : émotions, humeur (*moods*), point de vue interpersonnel (*interpersonal stance*), attitudes, ou encore affects (Karunaratne et al., 2011). Tous ces termes se réfèrent à différents types d'expériences affectives subjectives et personnelles (Russell, 2003). En effet, un affect intense pour une personne pourrait ne pas l'être pour une autre. Entre ces différents types d'affects, certains phénomènes de contagion existent. Par exemple, une personne éprouvant de la joie peut aussi ressentir une humeur plus positive. Le phénomène de contagion affective existe aussi entre les individus. Par exemple, Bono et Ilies (2006) ont montré l'effet que peut avoir des démonstrations d'émotions positives de la part d'un leader sur l'humeur générale des personnes qui le suivent. Les différents termes utilisés pour désigner les affects sont utilisés au sein de multiples disciplines telles que la linguistique, la psychologie et la psycholinguistique. La terminologie exacte utilisée peut varier au sein de chaque discipline. Toutes ces variations, de subjectivité, de contagion et de champs d'études créent un flou général sur la compréhension de ce que sont les différents types d'affects. En conséquence, ces termes sont souvent intervertis, bien qu'ils désignent des concepts différents. Nous allons donc prendre le temps dans cette section de présenter les différents termes associés aux affects. Cela nous permettra de faciliter la compréhension des études sur le sujet, et définir notre propre objet d'intérêt. Nous suivons dans les paragraphes qui suivent la classification proposée par (Karunaratne et al., 2011), inspirée de la classification de Scherer (2003).

Design Features	Response Characteristics			Antecedents		Consequences	
	Intensity	Duration	Synchro- nization	Event focus	Appraisal elicitation	Rapidity of change	Behavior impact
Emotions: <i>angry, sad, joyful, fearful, ashamed, proud, elated, desperate</i>	●	●	●	●	●	●	●
Moods: <i>cheerful, gloomy, irritable, listless, depressed, buoyant</i>	●	●	●	●	●	●	●
Interpersonal stances: <i>distant, cold, warm, supportive, contemptuous</i>	●	●	●	●	●	●	●
Preferences/ Attitudes: <i>liking, loving, hating, valuing, desiring</i>	●	●	●	●	●	●	●
Affect dispositions: <i>nervous, anxious, reckless, morose, hostile</i>	●	●	●	●	●	●	●

Figure3. Defining different types of affect: A design feature approach
Importance of the design feature in identifying the affect type is indicated by the size of the dot. More important features are marked with bolder dots while the less important features are marked with smaller dots. (i.e. duration and behavior impact are more important in defining affect dispositions than other four design features)

Figure 9 – Résumé des types d’affects, exemples, et caractéristiques, extrait de Karunaratne et al. (2011)

Tout d’abord, nous avons « l’affect ». Ce terme, large, est utilisé dans l’étude pour englober tous les autres comme des types d’affects particuliers. De même, Ochs et Schieffelin (1989) utilisent aussi le terme d’affect pour regrouper sous son sens tout ce qui a trait à ce que l’on peut ressentir physiologiquement, mentalement, démontrer en interaction, ou dégager vis-à-vis d’autrui : les ressentis, l’humeur, les dispositions, les attitudes.

Nous avons ensuite les « émotions » comme la peur, la joie ou la tristesse. Ce terme revient très régulièrement tant dans la littérature que dans la vie quotidienne. Elles sont définies par Lange (1922) comme l’interprétation que nous faisons des changements physiologiques perçus dans notre corps : par exemple, on ferait l’expérience de la peur lorsque l’on réalise que notre cœur bat vite. Les émotions concerneraient donc les états instinctifs ressentis par une personne, physiologiquement et « affectivement » de manière non contrôlée. En Linguistique Interactionnelle et en Analyse Conversationnelle, les émotions sont considérées comme des démonstrations ou expositions (*displays*), construites socialement dans un contexte précis, et ayant pour but la communication (Harré, 1986; Selting, 2010). Cette seconde description des émotions semble en contraste

avec la première, car on parle ici de quelque chose qui serait construit, exposé, et qui ne peut alors pas être instinctif. Harré (1984), reconnaît néanmoins que les émotions sont quelque chose qui est ressenti mais aussi démontré socialement en interaction. Lorsque l'on parle des émotions, on pense rapidement aux émotions fondamentales (Ekman, 1992, 1993) : peur, colère, joie, tristesse, dégoût et surprise. Ces émotions ont en effet très souvent été analysées mais l'idée qu'il y ait seulement 6 émotions fondamentales ne fait pas consensus et le nombre exact d'émotions varie. Dans le tableau de Karunaratne et al. (2011) (Figure 9), on voit que c'est un affect qui est lié à un évènement précis, qui peut changer rapidement et affecter le comportement.

Ensuite, nous avons les humeurs comme la gaîté ou la mélancolie. Karunaratne et al. (2011) présentent les humeurs comme quelque chose qui ne serait pas relatif à un évènement et serait donc un ressenti plus général, n'affectant pas nécessairement le comportement de la personne. Comme les émotions, l'humeur peut changer assez rapidement bien que cette rapidité soit moindre. Les auteurs ne s'étendent pas plus sur cette définition. Cependant, nous retrouvons une étude détaillée sur les émotions et les humeurs proposée par Beedie et al. (2005). Ils comparent la distinction qui est faite entre ces termes au sein de la littérature scientifique à travers une large comparaison d'articles scientifiques, et au sein de la population non-académique à travers plusieurs questionnaires d'évaluations et de notations. Ils concluent que l'émotion est définie comme 1) venant du cœur et ressentie, 2) intense mais brève, 3) ayant une cause connue et identifiée par celui qui la ressent, 4) démontrée en interaction. Ces quatre points sont en adhérence avec la définition de l'émotion proposée par Karunaratne et al. (2011) L'humeur au contraire, est définie comme 1) venant de l'esprit et serait pensée, 2) étant d'intensité moyenne, 3) ayant une cause floue mal identifiée par la personne qui la ressent, 4) n'étant pas démontrée en interaction. Là encore, les définitions concordent. Ainsi, l'humeur est un affect moyennement intense, sans cause définie, et n'est pas démontré en interaction.

Le « stance » ou point de vue interpersonnel (« *interpersonal stance* ») correspond au fait d'être froid, distant, méprisant, ou de soutenir autrui. C'est un affect qui est lié à un évènement précis. Selon Karunaratne et al (2011) le comportement d'une personne serait peu affecté par son propre point de vue interpersonnel, et qui peut changer aussi rapidement qu'une émotion. C'est un terme de l'Analyse Conversationnelle. Pour Stivers (2008), il est le point de vue émotionnel et attitudinal que nous exprimons par rapport à

nos propres paroles. Par exemple, le fait d'annoncer une bonne nouvelle en montrant du mépris peut véhiculer des informations sur ce que l'on ressent par rapport à la nouvelle : « Anna a validé son diplôme » dit avec un *stance* méprisant peut ajouter comme information que cette nouvelle ne ravie pas la personne qui l'annonce, qu'elle n'apprécie pas Anna, ou encore qu'elle estime qu'elle n'aurait pas dû avoir son diplôme. Pour Selting (2010), ce terme est régulièrement utilisé pour adresser n'importe quel affect dans le cadre de la CA, comme les émotions, les ressentis, les humeurs, les dispositions et les attitudes. Le *stance* pourrait donc être considéré comme un moyen de désigner les affects tels qu'ils sont compris dans l'interaction : liés à ce que Goodwin et al. (2012) nomment les flux changeants de la conversation (« *the back-and-forth flow of conversation* »). Il est possible d'accorder son *stance* à celui de l'interlocuteur auquel on s'adresse, ce qui peut être utile dans le fonctionnement de l'affiliation en interaction : pour reprendre notre exemple, si celui qui annonce « Anna a validé son diplôme » le fait avec un *stance* méprisant et que son interlocuteur est un ami, l'interlocuteur pourrait répondre « ah bon ? Ça m'étonne » en s'accordant au *stance* méprisant. Locuteur et interlocuteur seront alors affiliés sur du mépris.

Les « attitudes » correspondent au fait d'aimer, de détester, d'apprécier. Ce sont des affects qui, contrairement au *stance*, ne sont pas liés à un événement précis, et ne changent pas rapidement dans le temps. Comme le *stance*, en revanche, ils peuvent avoir un impact moyen sur le comportement. Pour Rilliard et al. (2009), les attitudes sont un affect lié à la dimension culturelle du langage et des affects. Dans leur étude conduite sur le japonais et le français, ils définissent les expressions attitudinales comme encodées de façon conventionnelle par une culture et un langage. Ils différencient les attitudes des émotions, qui ne seraient pas autant construites socialement que les attitudes. Pour Rilliard et al. (2009), l'étude des expressions attitudinales serait par leur dimension culturelle moins impactée que l'étude des émotions lorsqu'elles sont enregistrées en laboratoire. En effet, si les expressions attitudinales sont déjà mises en scène de manière plus consciente que les émotions par les locuteurs, alors il devrait être moins difficile de les reproduire en laboratoire. Cependant, les auteurs ne cherchent pas à complètement les distinguer au sein de l'étude de la prosodie. Pour eux, la prosodie serait en effet porteuse d'informations à la fois attitudinales et émotionnelles, et il ne serait alors peut-être pas nécessaire de toujours les distinguer. Dans une étude sur les effets attitudinaux de la prosodie, Wichmann (2000) observe le lien avec les émotions. Pour elle, les émotions (la colère, la

joie, la tristesse) seraient ce que l'on ressent, et correspondraient à ce qu'elle appelle l'intonation expressive. Les attitudes en revanche, seraient une sorte de comportement de la part du locuteur, quelque chose de montré en interaction donc. Elle prend comme exemple le fait d'être condescendant, amical, ou impoli. Ces exemples nous amènent à penser que ce que Wichmann (2002) décrit comme les attitudes se rapproche plus de ce que nous avons défini comme *stance*. Cet exemple nous montre, encore, à quel point il peut être complexe lorsque nous souhaitons dresser un inventaire des caractéristiques prosodiques d'un affect en particulier, de démêler les informations données dans la littérature et de ne pas confondre des affects qui sont pourtant bien différents entre deux études.

Les dispositions affectives, enfin, correspondent au fait d'être nerveux, anxieux, ou encore hostile. Ce sont des affects qui durent dans le temps et affectent le comportement de la personne. Pour Karunaratne et al. (2011), ce sont les dispositions affectives qui peuvent induire des désordres affectifs tels que le stress. Pour Mühlhoff (2019), les dispositions affectives sont des traces des relations, événements et rencontres passées, et qui affectent le présent. Elles seraient donc ancrées dans le vécu de la personne et liées à la façon dont le corps réagit pour se préserver de situations. Mühlhoff (2019) soulève le fait qu'en psychologie, le terme de « disposition » est même utilisé comme un trait de personnalité, et qu'en sociologie, il est relié à la notion d'*habitus*, qui désigne l'ensemble des schémas de perception, de réaction et de comportement d'un individu liés au contexte de socialisation dans lequel il se trouve. Cette dernière définition nous permet de comprendre un autre aspect de la complexité de l'étude des affects. Les affects peuvent être vécus sur des périodes variées, peuvent plus ou moins affecter le comportement de l'individu, ont des causes parfois inexplicables et parfois bien définies, mais aussi, ils sont liés à la personnalité et au vécu même des personnes qui vivent ces affects. Tous ces types d'affects sont entremêlés et il peut être difficile de toujours être sûr que ce que l'on étudie est exactement ce que l'on souhaite étudier : si l'on cherche à observer les effets de la joie sur la prosodie, mais que le participant qui se présente est d'humeur mélancolique, ou qu'il a une attitude négative vis-à-vis de l'étude, cela pourrait énormément impacter sa production.

Nous avons montré ici qu'il existe de multiples façons de catégoriser les affects. Nous avons jusqu'ici présenté des définitions de l'affect où tous sont séparés les uns des autres, c'est ce que l'on appelle les émotions discrètes (*discrete emotions*), ou bien les classer

selon des axes ou catégories, et ce sont les émotions catégorielles (*categorical emotions*). Nous allons maintenant nous intéresser à la catégorisation par axe. Russel (1980) a développé un modèle de catégorisation de l'affect qui comporte deux grands axes. Il place, en abscisses, l'axe de la « valence ». Cet axe oppose deux extrémités, le plaisir et le déplaisir, aussi différencié par positif et négatif : on retrouve par exemple en valence positive des émotions comme le plaisir ou la sérénité, et en valence négative la frustration et le sentiment d'être misérable. En ordonnées, il place l'axe de l'activation ou de l'« excitation » (*activation* ou *arousal*) qui oppose, selon les versions, l'activation à la désactivation, ou l'excitation forte vs excitation faible : on retrouve par exemple en excitation haute des émotions comme l'étonnement ou le sentiment d'alarme, et en excitation basse le sommeil et l'ennui. Ces axes sont des échelles sur lesquelles les affects peuvent être placés. Cette catégorisation des affects appelée le modèle circumplexe (Russell, 1980, 2003) visible ci-dessous tel que représenté par Seo et Huh (2019).

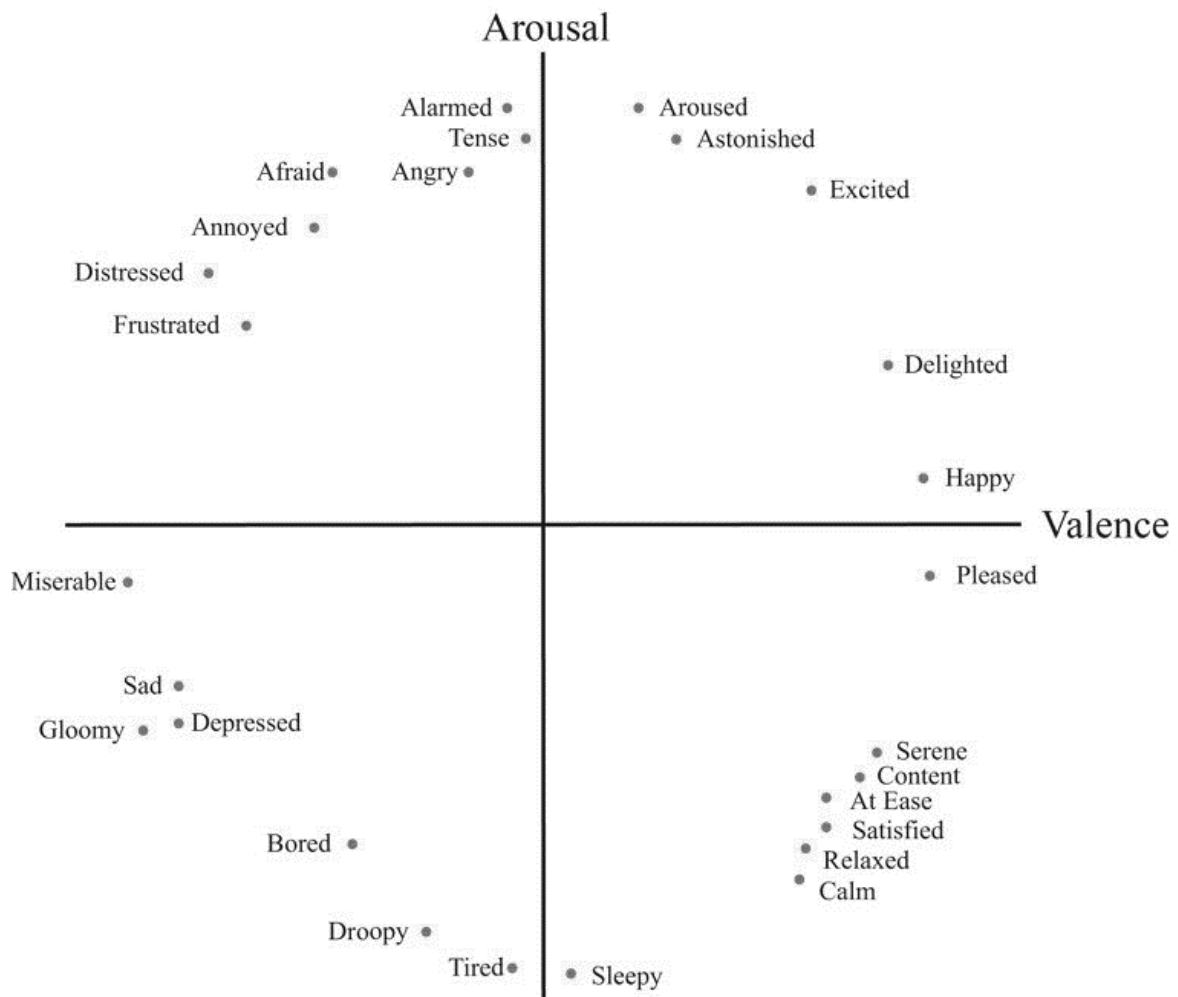


Figure 10 – Modèle circomplexe (Russel, 2003, p.143)

Dans ce modèle, nous avons donc d’une part la valence positive ou négative de l’affect, et de l’autre l’excitation haute ou basse. Une troisième dimension de la dominance existe, proposée par Mehrabian et Russell (1974). Cependant, le modèle bidimensionnel est généralement le plus utilisé (Bakker et al., 2014). Si nous nous reportons à la Figure 10 du modèle à deux axes, nous remarquons que certaines des émotions fondamentales d’Ekman sont représentées à des points clés sur l’axe de la valence : la joie (*happy*) est placée à l’extrémité positive de la valence, et sur l’axe de l’excitation haute bien que plus bas que l’excitation ou la colère. A l’extrémité opposée dans la valence négative, on retrouve la tristesse. Elle est située sur l’axe de l’excitation basse, mais plus haut que la fatigue ou l’ennui.

Dans cette Thèse, nous ferons une différence entre affect à valence positive, et affect à valence négative. Nous avons choisi de nous concentrer sur cette distinction car elle est

cohérente avec notre objet d'étude : les annonces de nouvelles. En effet, les nouvelles peuvent être, comme cela a été précédemment décrit, des « bonnes nouvelles », à valence positive, ou bien des « mauvaises nouvelles », à valence négative. Bien que l'axe de l'excitation ait aussi un certain intérêt, il n'est pas le plus adapté pour nous et nous ne contrôlerons donc pas l'arousal, car nous pensons que les nouvelles négatives ou positives éliciteront des affects sur l'axe de la valence plus que sur celui de l'excitation. En revanche, les nouvelles que nous allons étudier seront positives et négatives et devraient donc éliciter des affects que nous pensons proches de la joie et de la tristesse, respectivement.

Nous avons maintenant établi les types d'affects auxquels il est possible de s'intéresser, et lesquels nous regarderons. Il faut maintenant soulever la question de la façon dont ces affects peuvent être étudiés. En effet, nous souhaitons observer la prosodie affective liée à ces valences positives et négatives. Nous allons maintenant nous intéresser à la manière dont la prosodie affective est étudiée, avec quels types de corpus, et faire un état de la littérature dans le domaine des affects véhiculés par la prosodie.

3.2. Enregistrer la voix affective : quel corpus choisir ?

Il existe plusieurs types de corpus que l'on peut utiliser pour analyser l'expression vocale. En fonction de la manière dont le corpus sera enregistré, on pourra obtenir des affects plus ou moins typiques. Laukka et al. (2012) se sont penchés sur cette notion de « typicalité » d'un affect, ou plus précisément d'une émotion, dans une étude sur l'expression des émotions vocales. Une émotion dite « typique », serait une émotion dont les corrélats prosodiques et gestuels seraient au plus proche d'une émotion catégorielle déterminée, une sorte d'exemple « parfait » (ex, de joie, tristesse) qui sera aisément reconnu s'il fait l'objet d'une étude de perception. Deux types de corpus sont opposés : les corpus « actés », dans lesquelles les émotions proposées sont généralement bien reconnues car elles auraient des caractéristiques acoustiques uniques, et les corpus « spontanés » où les expressions émotionnelles discrètes telles que les émotions de base (ex., joie, colère) sont bien moins fréquentes, et souvent moins reconnaissables. Nous

allons maintenant développer les avantages et désavantages des deux méthodes d'enregistrement de l'affect.

Les expressions vocales « simulées », ou « actées » sont comme leur nom l'indique, une reproduction d'affects. Les « expressions actées » peuvent être obtenues de plusieurs façons. Scherer (2003) les présente comme une des méthodes préférées dans le domaine de l'étude des émotions vocales : on enregistre, en laboratoire, des acteurs professionnels ou semi-professionnels et on leur demande de jouer les émotions. Cette demande est explicite, par exemple : « produisez l'énoncé X de façon triste », ou « dites ces énoncés comme si vous étiez très heureux ». Cette méthode d'enregistrement de corpus facilite énormément l'obtention de données affectives de plusieurs façons. Tout d'abord, cela permet de « choisir » les émotions que l'on va étudier, mais aussi de les enregistrer avec une haute qualité audio, en chambre-sourde ou dans une pièce adaptée. Cependant, il est parfois reproché aux études utilisant les acteurs d'obtenir des résultats sur une version exagérée des affects, qui serait donc moins réaliste. Scherer (2003) précise que bien qu'il soit très probable que la représentation simulée des affects soit beaucoup plus intense et prototypique que lorsqu'ils sont enregistrés en contexte naturel, on ne peut pas ignorer la possibilité que les acteurs exagèrent certains indices évidents des affects et en oublient d'autres, plus subtils.

Cette possibilité que certains indices soit exagérés et d'autres négligés a depuis quelques années fait l'objet d'études qui se sont attachées à la comparaison des deux types de corpus, naturels et actés (Batliner et al., 2000; Fernandez & Picard, 2011). Elles ont montré que les acteurs vont en effet produire les affects d'une façon différente de lorsque ces affects sont produits en contexte naturel, ou comparés à un locuteur naïf auquel on demanderait de produire la même émotion. Batliner et al. (2000) traitent de la reconnaissance des émotions dans la parole et de son application dans les systèmes de dialogue. Ils cherchent à mettre en évidence les points positifs et négatifs des corpus actés par des acteurs, et des corpus enregistrés avec des locuteurs naïfs, tous deux utilisés pour rendre possible des interactions homme-machine comprenant des affects dans la vie réelle. Plus précisément, ils se concentrent sur la reconnaissance des émotions dans la communication homme-machine lors de dialogues de prise de rendez-vous au téléphone. L'étude montre que les corpus sur acteurs obtiennent des meilleurs résultats dans la discrimination des émotions par leurs caractéristiques acoustiques. Il attribue cette différence avec les autres corpus au fait que les acteurs proposent des émotions plus

« manifestes » car c'est ce qui leur est demandé de faire. Par ailleurs, l'étude met aussi en avant la possibilité que les acteurs puissent avoir une gamme limitée de variations émotionnelles par rapport aux locuteurs de la vie réelle, ne permettant donc pas d'entraîner les systèmes de dialogues correctement pour les situations réelles.

En accord avec l'idée que les acteurs ne produisent pas tout à fait les mêmes affects que ce que l'on trouve dans des corpus naturels, Berry et Brown (2019) rappellent que l'on apprend aux acteurs à utiliser gestes, expressions faciales et ton de la voix pour faire le portrait de leur personnage (Benedetti, 2012; Berry & Brown, 2019; Brestoff, 1995; Goldstein & Bloom, 2011; Kemp, 2012; Konijn, 2000; Schechner et al., 2013; Zarrilli & Hulton, 2009). Ils mettent en avant le fait que les acteurs peuvent avoir différentes méthodes de jeu et pourraient donc produire des énoncés différents. Ils nomment l'étude de Juslin et al. (2018) qui s'est attachée à observer la différence entre les affects en corpus actés et naturels. Ils ont trouvé que les énoncés actés étaient jugés comme ayant une intensité affective plus forte. Les énoncés naturels, eux, étaient jugés comme plus authentiques. Ainsi, les corpus actés ont pour avantage des expressions vocales très définies et claires, qui rendent possible de dresser des inventaires de caractéristiques prosodiques récurrentes dans certaines émotions, établissant ainsi un point de départ stable pour explorer les affects. Il semble cependant que ces corpus ne permettent pas à eux seuls de déterminer toutes les caractéristiques prosodiques qui entrent en compte dans la voix affective. D'une part, les logiciels de reconnaissance vocale des affects entraînés sur les affects actés performant moins bien en situation naturelle. D'autre part, les voix actées sont jugées moins authentiques par les auditeurs. Il n'est donc pas exclu que cela puisse influencer la perception que les auditeurs auront de l'affect dans les études.

Certaines méthodes « d'induction d'affect », et généralement plutôt d'humeurs sont parfois utilisées pour rendre les corpus plus « authentiques » et se rapprocher de la parole naturelle (Alpert et al., 1963; Altarriba, 2012; Bachorowski & Owren, 1995; Bonner, 1943; Havrdová & Morávek, 1979; Hicks, 1979; Karlsson et al., 2000; Markel et al., 1973; Roessler & Lester, 1979; Scherer, 1979, 1985, 2003; Skinner, 1935; Tolkmitt & Scherer, 1986). Scherer (2003) reconnaît que l'induction d'humeur présente l'avantage d'un fort degré de contrôle pour le chercheur, mais il lui reproche un certain degré d'incertitude : pour lui, on ne peut être sûr que l'induction a fonctionné de la même façon sur tous les participants, et il avance que les effets de l'induction sont en général assez faibles.

L'expression vocale « naturelle » des affects serait plus proche de la réalité, mais plus difficile à obtenir, avec des résultats montrant des effets moins forts que dans les études en émotions simulées. En effet, afin d'obtenir des corpus d'affects vocaux naturels, il faut trouver un moyen d'enregistrer dans un tel contexte. Pour se faire, les études se tournent en général vers des conversations en centres d'appels téléphoniques, qui peuvent aisément être enregistrées. Dans leur étude sur les émotions vocales actées et naturelles, Fernandez et Picard (2011) ont justement travaillé sur ces types de corpus, en prenant deux sortes de corpus non actés. D'une part, un corpus qui est composé d'une collecte d'interactions de service téléphonique entre une personne et son opérateur téléphonique. Ce corpus n'est aucunement guidé par les chercheurs, mais simplement récupéré à posteriori. Ce n'est donc pas un corpus créé spécifiquement pour l'étude des affects mais il peut néanmoins contenir des interactions avec de la production d'affects vocaux, et ce en contexte entièrement naturel. Un second corpus est utilisé. Celui-ci est semi-dirigé. Il contient des enregistrements d'appels passés par un participant enregistré en laboratoire à un ami ou un proche.

D'autres propositions de corpus d'émotions naturelles existent, dont trois projets de constitution de corpus naturels que Douglas-Cowie et al. (2003) s'attachent à décrire et comparer dans un article qui traite des progrès de la recherche sur la parole et l'affect, en particulier dans le contexte de l'interaction homme-machine. L'accent est mis sur l'importance de développer des bases de données appropriées pour l'étude de la parole affective et aborde quatre questions principales : la portée du corpus, l'aspect naturel des productions, le contexte et les descripteurs utilisés pour classifier les affects. Les auteurs préconisent de prendre la création de corpus naturels contenant beaucoup de démonstrations d'affects comme une tâche à part entière/ En effet, la variabilité des affects nous oblige à créer de grandes bases de données. Ils conseillent de ne pas s'attacher à l'étude d'émotions trop spécifiques, mais plutôt aux états émotionnels reliés à cette émotion (ne pas regarder la colère ou la tristesse, mais les états émotionnels « fortement négatifs »), et mettent en avant le besoin d'un consensus sur ces catégories. Ces états émotionnels généraux en grandes catégories, bien qu'ils ne correspondent pas tout à fait à la notion d'affects négatifs et positifs qui nous intéresse, s'en rapprochent : nous avons par exemple des affects « pas très émotionnels », (neutre, confiant), positifs non orientés (joyeux, amusé), positifs orientés vers un objet (aimant, affectueux).

Laukka et al. (2012) résument eux aussi les différences entre les types de corpus et comment les choisir. Dans leur étude, ils observent des corpus et comparent trois types d'expression : actées par des professionnels, actées par des personnes naïves, non-actées. Ils réalisent sur une étude de perception dans laquelle ils demandent aux participants de juger de la typicalité de l'expression (joie, colère, peur), de donner la fréquence à laquelle ils s'attendent à entendre ce type de voix pour l'émotion jugée dans la vie réelle, et si la représentation vocale qu'ils entendent est la plus idéale ou non pour véhiculer l'émotion étudiée. Il ressort de leur étude que les émotions actées, surtout celles jouées par les acteurs, sont reconnues comme étant les plus « prototypiques » par les auditeurs. Elles correspondent à l'idée de l'expression idéale d'une émotion qu'un individu peut avoir, plutôt qu'à la fréquence à laquelle on peut l'utiliser dans la vie de tous les jours. Les émotions prototypiques, proches de « l'idéal » que les auditeurs imaginent, seraient reconnues comme la bonne expression de l'émotion car elles seraient plus intenses. Les expressions vocales des émotions seraient moins intenses et moins prototypiques, mais elles seraient cependant beaucoup plus courantes dans la vie de tous les jours que les expressions très prototypiques / plus intenses. Pour Laukka et al. (2012), il convient d'utiliser des émotions moins prototypiques si l'on souhaite comprendre « l'essence » de l'expression des émotions en interaction spontanée. En effet, les expressions spontanées sont souvent masquées ou placées stratégiquement. Elles portent donc des états affectifs complexes non prototypiques (Laukka et al., 2008; Scherer, 1989) qui ne pourraient être entièrement compris par l'étude uniquement des expressions prototypiques.

A la lumière de ces travaux, il apparaît que les corpus de données naturelles permettent l'étude d'interactions affectives en contexte naturel et donc potentiellement plus « authentique ». Les affects véhiculés seraient assez complexes et peu prototypiques car masqués ou contrôlés en partie par le locuteur. L'étude de ce type de corpus est donc très importante si l'on souhaite comprendre en profondeur le fonctionnement de la parole affective dans les contextes naturels, à la fois pour la recherche fondamentale, mais aussi pour des domaines appliqués comme l'interaction homme-machine. Cependant, ces corpus restent plus complexes à analyser que les corpus actés par des acteurs professionnels ou naïfs. En effet comme nous l'avons pointé plus tôt, les affects enregistrés en corpus naturels sont plus difficiles à classer, particulièrement si l'on souhaite les classer en catégories bien définies comme les émotions de base (joie, tristesse). Ils

nécessitent donc en général une évaluation perceptive par des participants avant de pouvoir être étudiés pour leurs caractéristiques prosodiques.

Douglas-Cowie et al. (2003) évoquent justement le fait que le prix de l'authenticité des affects est le manque de contrôle sur les données. Par ailleurs, il est beaucoup plus difficile d'extraire des énoncés suffisamment proches pour être comparables si l'on cherche à observer la phonologie intonationnelle au sein du corpus. Pour trouver des phrases identiques afin d'obtenir un set de données bien équilibré, il faudrait un corpus très grand et l'annotation serait extrêmement longue. De plus, ce sont des corpus difficiles à obtenir, à cause de leur format, et les règles éthiques et autorisations à demander pour la protection des participants rendent le processus de création de corpus plus long que lorsqu'il est simplement enregistré en laboratoire. Les corpus naturels sont donc pertinents pour l'étude des affects mais ne sont pas toujours la solution la plus adaptée selon l'objet de l'étude. Il nous apparaît donc qu'il convient de continuer à utiliser de multiples méthodologies d'enregistrement de corpus afin de pouvoir explorer tous les aspects de la voix affective.

Afin de trouver un compromis entre expression naturelle et expression actée, nous nous intéresserons dans notre propre corpus d'annonces de nouvelles à l'expression vocale des affects telle qu'elle peut être produite par des non-acteurs. En demandant à nos locuteurs d'énoncer des nouvelles qui s'avèrent être neutres, bonnes ou mauvaises, nous ne leur demandons pas d'imiter une émotion particulière dans leur appel (comme de la tristesse pour les mauvaises nouvelles ou de la joie pour les bonnes nouvelles). Nous leur demanderons simplement d'annoncer des nouvelles positives, négatives ou neutres, par leur contenu. En effet, notre intérêt se porte sur l'étude de la valence de nouvelles annoncées telle qu'elle est projetée au travers de la prosodie de la préface. Nous souhaitons regarder la production d'un locuteur en contexte conversationnel, lorsqu'il se laisse guider par la trajectoire interactionnelle proposée par le texte à lire. Si le texte est une annonce de nouvelle au sujet d'un vol annulé, nous nous attendons à ce que, guidé par la valence négative de la nouvelle, donc la trajectoire d'une annonce de nouvelle négative, le locuteur produise des corrélats prosodiques en conséquence. Nous pensons que la valence positive ou négative de la nouvelle à annoncer conduira le locuteur à produire des indices de projection prosodique sur cette valence. Afin de pouvoir établir par la suite des hypothèses quant aux indices acoustiques et phonologiques que nous pourrions trouver,

il nous faut maintenant nous pencher sur ce que la littérature en prosodie émotionnelle ou affective a découvert au cours des ans.

3.3. Quels indices de projection attendre ?

Etat de la recherche sur la parole expressive

Comme nous l'avons évoqué précédemment, l'idée d'une dualité de structure inhérente à la prosodie en a profondément affecté l'étude (voir section 2.5). En conséquence, la plupart des études sur la production et la perception de la prosodie affective se sont, depuis 50 ans, concentrées sur sa réalisation acoustique phonétique, donc les corrélats sont décrits en section 3.3 : f_0 , intensité, rythme, et qualité de la voix (Batliner et al., 2000; Juslin & Laukka, 2001; Laukka et al., 2011; Scherer, 2003). Ces études ont permis de montrer quels sont les paramètres acoustiques utilisés dans la prosodie affective, mais aussi de montrer qu'ils peuvent être utilisés indépendamment du lexique.

La recherche qui a été faite sur ses corrélats phonétiques l'a été sur les « émotions fondamentales » (Ekman, 1992). Comme le soulignent Berry et Brown (2019), une grande partie de cette littérature porte spécifiquement sur la joie et la tristesse, les deux affects qui nous intéressent tout particulièrement (Banse & Scherer, 1996; Goudbeek & Scherer, 2010; Juslin & Laukka, 2003; Juslin & Scherer, 2005; Konijn, 2000; Laukka, 2005; Laukka et al., 2005; Murray & Arnott, 1993; Scherer, 2003). Nous faisons l'hypothèse que les corrélats prosodiques d'une annonce de nouvelle positive seront proches des corrélats prosodiques de la joie, et que les corrélats prosodiques d'une annonce de nouvelle négative seront proches des corrélats prosodiques de la tristesse. Ces études, et celles qui portent sur d'autres émotions fondamentales ont mis en évidence plusieurs corrélats acoustiques. Berry et Brown (2019) rapportent que la colère et la joie sont marqués par des augmentations du débit de parole, de l'intensité et sa variabilité, de l'énergie dans les hautes fréquences (donc la qualité de la voix), de la f_0 moyenne et sa variabilité et pour la colère spécifiquement par moins de pauses. A l'inverse, la tristesse et la tendresse sont marquées par des diminutions du débit de parole, de l'intensité et sa variabilité, de l'énergie dans les hautes fréquences, de la F_0 moyenne et sa variabilité et pour la tristesse spécifiquement d'un plus grand nombre de pauses.

Synthetic compilation of the review of empirical data on acoustic patterning of basic emotions (based on Johnstone and Scherer, 2000)

	Stress	Anger/rage	Fear/panic	Sadness	Joy/elation	Boredom
Intensity	↗	↗	↗	↘	↗	
F0 floor/mean	↗	↗	↗	↘	↗	
F0 variability		↗		↘	↗	↘
F0 range		↗	↗(↘)	↘	↗	↘
Sentence contours		↘		↘		
High frequency energy		↗	↗	↘	(↗)	
Speech and articulation rate		↗	↗	↘	(↗)	↘

Figure 11 – Tableau des différences acoustiques entre des émotions discrètes, extrait de Scherer (2003)

Scherer (2003) présente aussi un résumé des caractéristiques acoustiques d'émotions basiques (stress, colère, peur, tristesse, joie, ennui) que nous présentons ci-dessus (Figure 11). Ce résumé présente des résultats cross-linguistiques, car il est élaboré à partir de plusieurs études faites sur différentes langues et des résultats de Banse et Scherer (1996) sur l'allemand. Si l'on regarde les caractéristiques acoustiques de la tristesse, nous pouvons remarquer qu'elles sont toutes abaissées par rapport à ce que l'on peut trouver pour les autres affects, notamment en comparaison avec la colère et la joie, deux affects plus actifs sur l'axe de l'excitation que la tristesse.

Cependant il est possible de trouver des différences dans les paramètres acoustiques selon la langue, et nous allons donc maintenant présenter ce qui a été trouvé dans les études réalisées sur le français. Bänziger et al. (2001) présentent justement une étude qui utilise les corrélats phonétiques proposés par Banse et Scherer (1996) et par Scherer (2003) pour faire de la synthèse vocale des émotions, et évaluer si les émotions vocales de synthèse sont bien reconnues. Cette étude utilise les caractéristiques cross-linguistiques proposées par Scherer (2003) sur des phrases du français, pour recréer la joie et la tristesse. L'étude montre que les émotions de synthèse ont été bien reconnues par les participants à l'étude de perception.

Dans une étude récente, Audibert et al. (2005) ont observé le poids relatif des paramètres acoustiques dans la reconnaissance des émotions. Ils se sont intéressés à plusieurs scénarios de perception : resynthèse complète de la voix contenant tous les stimuli, intensité seule, f0 seule, f0 + contour de l'intensité et qualité de la voix + durée. La qualité de la voix est synthétisée à partir d'une forme d'onde du flux glottal (« *glottal flow* ») de référence, modifié pour exprimer différentes qualités vocales (ex., voix craquée ou soufflée), suivant la méthode utilisée par Gobl et Chasaide (2000). Avec ces différentes conditions d'écoute, ils ont pu remarquer que lorsque la voix est entièrement synthétisée,

la joie est identifiée 80% du temps, alors que la tristesse ne l'est qu'environ 60% du temps, et que le neutre est confondu 48% du temps avec elle. Dans la condition avec pour seul indice l'intensité, la joie était très mal reconnue, et la tristesse était reconnue environ 35% du temps. En revanche, elle était confondue 83% du temps avec le neutre. Dans les conditions avec f_0 et f_0 + intensité, la joie était la mieux reconnue (60%), et la tristesse était toujours confondue avec le neutre environ 38% du temps. Dans la condition avec la qualité de la voix et la durée, enfin la tristesse était mieux reconnue que la joie (40% vs 20%), et la joie était plus souvent confondue avec le neutre (36%) que la tristesse ne l'était (28%).

Il ressort de ces deux articles que les caractéristiques acoustiques trouvées dans les autres langues pourraient aussi être trouvées en français. La fréquence fondamentale apparaît alors comme le paramètre le plus pertinent pour distinguer la joie des autres affects en français, alors que la tristesse semble plus facilement confondue avec le neutre. Lorsque l'on prend en compte la qualité de la voix, en revanche, on remarque que la tristesse se détache mieux du reste. Le neutre, enfin, ne peut être réellement considéré comme un affect à part entière et il est parfois considéré en dehors des classes d'émotions (Vidrascu & Devillers, 2007).

Comme nous l'avons montré, beaucoup d'études ont été réalisées sur les corrélats acoustiques phonétiques des émotions discrètes et autres affects. Cependant, certaines études se sont aussi concentrées sur les corrélats liés à des catégories plus larges, celles des axes d'excitation et de valence du modèle circomplexe de Russel. A partir de cette classification des émotions, Berry et Brown (2019) listent les caractéristiques acoustiques liées aux axes de l'excitation et de la valence en résumant les travaux de plusieurs chercheurs (Goudbeek & Scherer, 2010; Juslin & Scherer, 2005; Laukka et al., 2005; Owren & Bachorowski, 2007; Schröder, 2004; Schröder et al., 2001).

Ils trouvent que, pour l'axe de l'excitation, l'excitation est positivement corrélée avec, pour la f_0 , des augmentations de la moyenne, de l'écart-type (variabilité), et de l'empan vocal, pour l'intensité des augmentations de la moyenne et de l'écart-type, pour les valeurs de durée des augmentations de la parole, de la longueur des phrases, et une diminution de la longueur des pauses, et enfin pour les caractéristiques spectrales, un timbre vocal qualifié de « tonitruant » (« *blaring* », qui correspond à des pentes spectrales plus plates dues à un plus grand effort vocal), plus d'énergie des hautes fréquences, et du *shimmer* (mesure des perturbations de l'amplitude vocale). Ainsi, les valeurs de la f_0 , de

l'intensité, des durées (des segments et des pauses), et des caractéristiques spectrales sont autant de facteurs que l'on peut observer pour différencier les émotions à haute activation et les émotions à faible activation. Bien qu'au cours de notre travail, nous nous intéresserons plutôt à l'axe de la valence, nous avons précédemment noté que la joie, que nous relierons à notre valence positive, est un peu plus haute sur l'axe de l'excitation que la tristesse, que nous relierons à notre valence négative. Il est donc possible que nous retrouvions certaines des caractéristiques des affectss à haute excitation dans les annonces de nouvelles positives. De même nous avons précédemment noté que la tristesse, que nous relierons à notre valence négative, est à l'inverse un peu plus basse sur l'axe de l'excitation. Il est donc possible que nous retrouvions des caractéristiques opposées à celle des affects de haute activation, donc plus proches des affects à basse excitation dans les annonces de nouvelles négatives. Berry et Brown (2019) font ensuite un résumé des corrélats acoustiques de l'axe de la valence. Ils trouvent que lorsque la valence est positive, la prosodie est réalisée avec un élargissement de l'empan vocal, une diminution de l'énergie donc du niveau sonore, une augmentation du débit de parole et pauses plus courtes, et enfin moins d'énergie dans les hautes fréquences, un signal plus clair (donc moins de bruit dans la voix par rapport aux harmoniques), et une qualité de voix « chaude » (qui correspondrait à des pentes spectrales plus plates, un effort vocal plus grand et une voix basse et plus soufflée). Ces paramètres semblent donc pertinents à observer pour notre étude.

Dans une étude comparant la valence et l'activation, Liscombe et al. (2003) extraient différents paramètres prosodiques, automatiquement et manuellement. Ils analysent notamment deux mesures acoustiques des pentes spectrales (qui évaluent la façon dont l'énergie est répartie, dans les fréquences hautes ou basses) : les pentes spectrales des voyelles avec un stress nucléaire, et les pentes spectrales des voyelles avec la plus haute amplitude. Les corrélations faites entre ces mesures et la valence montrent une corrélation positive entre les pentes spectrales (surtout dans les voyelles à haute amplitude), et les émotions positives, donc plus d'énergie dans les fréquences hautes pour les émotions positives. Liscombe et al. (2003) concluent que le paramètre du spectral tilt, qui est lié à la qualité de la voix, semble une mesure pertinente à analyser pour l'étude de prosodie en fonction de la valence émotionnelle qui lui est attachée particulièrement pour des émotions comme la colère, l'anxiété, l'ennui et l'amicalité. En revanche pour la colère, la confiance, l'intérêt, la frustration, mais aussi pour la joie et la tristesse, le paramètre le

plus pertinent est la f_0 . Pour la joie, c'est la f_0 minimum qui ressort le plus et qui est positivement corrélée, donc plus haute en moyenne, et pour la tristesse, la f_0 maximum qui est négativement corrélée, donc plus basse en moyenne. Ces deux critères nous apparaissent donc comme pertinent pour l'observation des différences entre valence positive et valence négative.

La littérature a montré que les énoncés à valence positive tels que la joie sont caractérisés par une moyenne et une variabilité de la F_0 élevées, et un large empan vocal par rapport aux énoncés à valeur négative tels que la tristesse. La vitesse d'articulation sera plus rapide dans les énoncés à valence positive que dans les émotions à valence négative. D'autres études ont également examiné les caractéristiques phonétiques liées à la qualité de la voix, autres que les pentes spectrales. Par exemple, il y a le ratio harmonique sur bruit (HNR, *harmonics-to-noise ratio*), et la différence entre la première et la deuxième harmonique (H1-H2). Ces deux données sont largement utilisées dans la recherche linguistique pour rendre compte des types de phonation : voix modale, voix craquée, voix soufflée (Chai & Garellek, 2022; Jiang & Pell, 2017; Kuo, 2013; Yang et al., 2020). Une autre mesure de la qualité de la voix est celle de la proéminence du pic cepstral lissée (CPPs, *cepstral peak prominence smoothed*). C'est une mesure plus récente qui rend également compte de la phonation modale vs soufflée. Elle est également utilisée en prosodie, pour observer le rôle de la qualité de la voix dans la distinction de fonctions pragmatiques (Heldner et al., 2019). Plusieurs études montrent qu'il y a plus d'énergie à haute fréquence dans les émotions à valence positive (Johnstone & Scherer, 2000; Scherer, 2003) que dans les émotions à valence négative, et que la voix sera donc plus soufflée lorsque l'on produit un énoncé triste, et plutôt modales lorsque l'on produit un énoncé joyeux (Laukkanen et al., 1996; Scherer, 1986; Yuan et al., 2002). En français, nous retrouvons également de nombreuses études sur la perception au cours des deux dernières décennies, dont Audibert et al. (2005), Bänziger et al. (2001), et Juslin et Laukka (2001) dont nous avons explorés les résultats plus tôt dans cette section. Goudbeek et Scherer (2010), enfin, ont réalisé une étude sur plusieurs paramètres et observé notamment les mesures de la qualité de la voix (HNR) en comparant deux corpus, en français et en allemand. La HNR avait des valeurs plus élevées pour la joie que pour la tristesse.

Ainsi, beaucoup d'études ont été faites sur les corrélats acoustiques des affects et autres catégorisations affectives. Malgré cette focalisation sur la phonétique pour l'étude

de la signification paralinguistique, il a été reconnu à plusieurs reprises (Ladd, 2008; Prieto, 2015) que la différence entre sens linguistique et paralinguistique n'est pas entièrement claire et qu'elle devrait, à un moment donné, être étudiée. Depuis quelques années, les chercheurs ont justement commencé à remettre en question la notion de dualité de structure, et à s'intéresser à la possibilité d'une phonologie de la prosodie affective. Avec la remise en question de la dualité de structure, une nouvelle vision de la discrétisation du sens affectif a vu le jour. On y considère que la phonologie pourrait être un deuxième canal pour transmettre le sens affectif. Ainsi, de nouvelles études en phonologie intonationnelle sur les caractéristiques phonologiques du sens affectif sont menées, bien qu'encore rares. Ces études se concentrent sur le type de tons, ou sur l'addition de certains tons qui pourraient être associés à un sens affectif (Beňuš et al., 2007; Cao et al., 2014; Liscombe et al., 2003; Stibbard, 2001; Wang et al., 2020). Nous rapportons ici trois études principales, toutes réalisées sur l'anglais américain.

Dans le premier article Liscombe et al. (2003), explorent en effet les corrélats phonétiques, mais aussi phonologiques, de la prosodie affective au sein d'un corpus enregistré par des acteurs professionnels. Ils cherchent à différencier valence positive et valence négative en observant ces corrélats. Ils regardent, ainsi, des données comme la vitesse de parole, les pentes spectrales ou l'amplitude, mais aussi la forme du contour intonatif, le placement de l'accent mélodique et des tons de frontière, annotés avec la convention ToBI (élaboré pour l'anglais américain). Ils ont trouvé que les affects négatifs sont corrélés positivement avec des contours de plateau H-L% qui sont donc plus présents, et corrélés négativement ou non corrélés avec les patrons L-L% qui sont donc moins présents. En revanche les affects positifs sont corrélés négativement avec H-L%, mais corrélés positivement avec L-L%. Liscombe et al. (2003) concluent que le type d'accent mélodique, ainsi que les tons de frontière, pourraient être liés à la valence négative ou positive d'un affect, et donc utilisés pour les différencier.

Cependant cette étude est réalisée sur l'anglais américain, et il se pourrait fortement que nous ne retrouvions pas les mêmes accents mélodiques et tons de frontières en français, dont l'intonation diffère. Bien que cet article regarde l'influence des contours intonatifs et de certaines parties des contours, il ne s'étend pas sur des considérations théoriques où sur les implications que les résultats pourraient avoir sur l'étude de la voix affective. Nous considérons en effet que l'étude présentée par Liscombe et al. (2003) représente la première pierre d'un chemin qui ouvre la voie à la remise en question de la

notion de dualité de structure et à l'étude approfondie et extensive des contours intonatifs, tons de frontières, et accents mélodiques comme véhicules potentiels de sens paralinguistique autant que linguistique. Car si l'intonation n'a pas toujours été complètement écartée de l'étude du sens affectif (D. R. Ladd et al., 1985; Paeschke et al., 1999; Scherer, 2003), elle n'avait pas encore été étudiée du point de vue de la phonologie intonationnelle. Enfin, le corpus est constitué uniquement d'énoncés très courts de quatre syllabes contenant des dates et des nombres (« *two thousand four* ») : on peut donc s'attendre à des résultats moins évidents et des corrélations moins fortes si l'on observe des énoncés plus longs et portant plus de sens linguistique et affectif.

Le second article, de Cao et al. (2014) examine les indices phonologiques de l'émotion, combinés à une analyse acoustique de la f_0 et des coefficients cepstraux de fréquence mel lissés (MFCC) sur corpus acté constitué d'énoncés prononcés avec une émotion donnée. Les indices phonologiques potentiels explorés étaient la force des ruptures (*break index*), les tons de syntagmes intermédiaires, les tons de frontière en fin d'énoncés, les accents mélodiques et les bigrammes, annotés suivant les conventions ToBI. Ils observent ces données en comparant six affects : neutre, colère, dégoût, peur, joie et tristesse. Ils ont notamment trouvé des effets significatifs dans la partie nucléaire des patrons intonatifs, c'est-à-dire à la fin. Ils ont constaté par exemple que H^* et $H-L\%$ sont plus fréquents dans la tristesse, tandis que la joie est plutôt associée à $L + H^*$. Les deux affects sont moins souvent sujets à des tons de frontière finaux $L-L\%$, et plutôt à $!H-L\%$ et $H-L\%$. Le neutre avait pour unique ton phrasal intermédiaire le $H-$. Ainsi, ils ont constaté que les caractéristiques de la phonologie intonationnelle annotés avec ToBI et les caractéristiques acoustiques continues montrent des performances comparables sur les prédictions d'émotions dans l'énoncé. Cao et al. (2014) soulignent que l'état émotionnel neutre était celui pour lequel les patrons phonologiques annotés avec ToBI étaient le plus robuste, par rapport aux autres émotions.

Cet article met en avant la nouveauté de l'étude de l'implémentation phonologique de l'émotion, et met l'accent sur l'intérêt de ces représentations catégorielles, qui ont par le passé plusieurs fois été montrées comme pertinentes (Ladd et al., 1985; Mozziconacci, 2002). Elle démontre l'importance de l'étude des catégories phonologiques ToBI comme prédicteurs potentiels de l'émotion. Ils sont à prendre en compte au même titre que les prédicteurs phonétiques continus, pour continuer à développer notre compréhension du rôle de la prosodie dans le port d'informations paralinguistiques affectives.

La troisième étude que nous présentons est la plus récente. Wang et al. (2020) ont, eux aussi, également examiné les indices phonologiques et phonétiques des émotions dans la parole. Cet article porte, spécifiquement, sur le domaine de la SER (« *Speech Emotion Recognition* »). Les auteurs de l'article soulignent que peu de travaux dans leur domaine proposent des méthodes pour améliorer la reconnaissance automatique des affects dans la parole, et qu'ils ont été inspirés par les travaux de Cao et al. (2014). Le but de leur étude est de proposer une nouvelle méthode de reconnaissance automatisée des affects dans la parole. Comme dans les travaux précédents, ils suivent les conventions ToBI, en utilisant cette fois un outil d'annotation ToBI automatique. Wang et al. (2020) observent les corrélats prosodiques de quatre affects : le neutre, la colère, la joie, et la tristesse. Ils les classent par valence, activation et domination. Les résultats de cette étude montrent que les caractéristiques phonologiques suivantes sont fortement corrélées à la dimension de la valence : les accents mélodiques ! H*, H, !H* et L*+ H, l'accent mélodique en frontière L*+H, et le *break index* 1.

Là encore, il est démontré que le sens affectif peut aussi être porté par des catégories discrètes phonologiques qui s'avèrent être pertinentes pour améliorer les systèmes de classification de la voix affective. Cette étude est réalisée sur la base d'un corpus d'enregistrement d'interactions émotionnelles en dyades d'acteurs et actrices d'une durée de 5 minutes. Tout comme pour l'étude précédente, nous pouvons nous attendre à ce que les acteurs aient produit des caractéristiques phonologiques plus régulièrement que ce que nous pourrions trouver dans un corpus naturel. Nous notons cependant que la nature du corpus permet d'obtenir une approche différente de l'étude des caractéristiques phonologiques des affects, dans un cadre interactionnel, et en observant des portions de parole beaucoup plus longs qu'un énoncé seul.

Ces trois études ont donc observé les corrélations qui peuvent exister entre des catégories phonologiques et les affects au sein de la prosodie, et ont démontré l'intérêt de cette étude. Ce changement étant encore récent, peu de recherches explorent les aspects phonologiques de l'affect, seuls ou en combinaison avec les aspects phonétiques. De plus comme nous l'avons mentionné, ces études ont toutes été réalisées sur l'anglais américain, dont la phonologie est différente de celle du français ou d'autres langues. Ainsi, nous pensons qu'il est possible que nous trouvions aussi des liens entre patrons phonologiques dans la prosodie et valence affective en français, bien que nous ne trouvions probablement pas les mêmes résultats que dans les études citées ci-dessus.

3.4. Projection de structure et projection affective

Peu de travaux ont fait l'étude de la prosodie affective liée à la valence d'annonces de nouvelles. Nous en décrivons deux dans cette section, et nous nous attarderons particulièrement sur la seconde, donc notre première étude est inspirée. La première étude est celle de Freese et Maynard (1998) sur l'anglais, réalisée dans le cadre de l'Analyse Conversationnelle. Ils ont réalisé leur étude sur 100 occurrences d'annonces de nouvelles dans un corpus naturel de conversations téléphoniques, en anglais américain et anglais britannique. Ils s'attardent sur le fait que la signification affective d'une nouvelle ne soit pas intrinsèquement déterminée par la nature bonne ou mauvaise de la nouvelle, mais plutôt réalisée par les interactants. Attribuer la valence à la nouvelle pourrait, selon eux, être réalisé lexicalement ou prosodiquement. Ils s'intéressent à la valence attribuée par la prosodie, portées par des variations continues phonétiques. Pour les annonces de bonnes nouvelles, l'étude a démontré l'utilisation d'une f_0 plus haute avec un empan plus large, des contours phonologiques fréquents et abrupts, une qualité de voix normale, un niveau sonore plus fort sur les mots clés et une vitesse de parole plus rapide. Pour les mauvaises nouvelles, ils ont trouvé une f_0 basse à l'empan étroit, des voyelles allongées avec un ton descendant, une voix souvent soufflée ou craquée, des mots clés dits parfois plus bas et enfin, une vitesse de parole plus basse. Ces directions de ces différentes caractéristiques est similaire à celle trouvée pour certains affects dans les études sur la prosodie émotionnelle : la joie, et la tristesse. Les caractéristiques retrouvées pour les bonnes nouvelles sont proches de celles de la joie, alors que les caractéristiques retrouvées pour la tristesse sont proches de celles de la tristesse. L'étude se concentre cependant sur la partie de l'annonce de la nouvelle, et non sur ce qui vient avant l'annonce, comme la préface ou les marqueurs discursifs. Elle permet donc d'observer le rôle de la prosodie au moment où l'annonce est faite, qui serait ainsi de proposer une valence pour l'entièreté de la séquence d'annonce, et qui permettraient de structurer l'interprétation de la nouvelle et de ses implications pour l'interlocuteur destinataire de la nouvelle. Freese et Maynard (1998) mentionnent aussi le rôle de cette information pour permettre

à l'interlocuteur de s'aligner au prochain tour de parole en prenant en compte la valence proposée pour la nouvelle. L'étude ne permet cependant pas d'observer le rôle projectif de la prosodie en amont de l'annonce de nouvelle.

La seconde étude que nous allons décrire est celle de Swerts et Hirschberg (2010). Ils ont conduit une étude en néerlandais sur les « prédicteurs prosodiques » du contenu positif ou négatif à venir, dans un message vocal. Ils s'intéressent dans cette étude aux prédicteurs prosodiques de la parole émotionnelle lorsque l'on annonce de bonnes ou mauvaises nouvelles. Pour étudier ces prédicteurs, ils ont créé un corpus semi-dirigé dans lesquels des participants devaient appeler une personne fictive et laisser un message sur un répondeur en leur annonçant les résultats d'un entretien d'embauche.

Pour chacun des 40 locuteurs enregistrés, l'humeur était évaluée au moyen d'un questionnaire court, où sur 6 questions les locuteurs choisissaient entre deux possibilités sur une échelle de Lickert à 7 points (ex. « satisfait » vs « insatisfait »). Puis, une nouvelle humeur positive ou négative était élicitée avant l'enregistrement au moyen d'un extrait vidéo de 7 minutes : *Friends* (série humoristique) pour l'induction d'une humeur positive, et *La liste de Schindler* (film sur la Shoah) pour l'induction d'une humeur négative. Les personnes pour lesquelles l'humeur était positive devaient laisser un message de bonne nouvelle (obtention du travail), et les personnes pour lesquelles elle était négative devaient laisser un message de mauvaise nouvelle (refus d'embauche). Chaque message existait en deux versions, une dans laquelle la nouvelle était annoncée, et une dans laquelle elle ne l'était pas et on demandait à l'interlocuteur de rappeler pour la connaître. Dans les nouvelles avec décision, on peut distinguer une première partie introductrice de contextualisation de l'appel, et une partie avec la décision pour l'embauche.

TABLE I. Examples of original Dutch utterances (with English translations within brackets) elicited in positive and negative mood conditions of the 'with decision' and 'without decision' experiments.

Experiment	Mood	Example
With decision	Positive	<p>“Goeie morgen, Mirjam, u spreekt met Mieke Peeters van L&L. We hebben vorige week sollicitatiegesprekken gevoerd, en ik wil je mededelen dat [je bent aangenomen, en ik hoop dat je nog steeds geïnteresseerd bent in deze functie. Zou je mij zo snel terug kunnen bellen op 03/22234476. Alvast bedankt. Groetjes.]”</p> <p><i>(“Good morning, Mirjam, this is Mieke Peeters talking from L&L. We have had interviews last week, and I would like to inform you that [we want to offer you the job, and I hope that you are still interested in this position. Would you be so kind to call me back at 03/22234476 as soon as possible. Thanks a lot. Cheers.]”)</i></p>
	Negative	<p>“Goeie middag, Mirjam, je spreekt met Michael Dooren van L&L. We hebben laatst een gesprek gehad, een sollicitatiegesprek gehad. En ik wou heel even doorgeven dat [je voor ons niet geschikt genoeg bent bevonden voor deze functie. Als je vragen hebt, kun je nog contact met ons opnemen. Hartstikke bedankt. Tot ziens.]”</p> <p><i>(“Good afternoon, Mirjam, this is Michael Dooren from L&L. We have recently had a talk, we had an interview. And I would like to let you know that [we feel you are not suitable enough for this job. If you have any questions, then you can contact us. Thanks a lot. See you.]”)</i></p>

Figure 12 – Exemples de messages vocaux d’annonce avec décision, extrait de Swerts et Hirschberg (2010)

Après enregistrement, le corpus a fait l’objet d’une étude de perception. Pour les messages avec décision, seule la partie avant l’annonce a été conservée. Pour les messages sans décision, l’entièreté a été conservée. Afin de s’assurer qu’aucun indice lexical, ou autre que prosodique, ne leur ait échappé et influence les résultats, les annonces tronquées et les annonces sans décision étaient proposées aux participants avec l’audio, ou en version texte. Dans une tâche à choix forcé, les participants devaient dire s’ils pensaient que le message qui suivrait allait être une nouvelle positive ou négative. Les résultats montrent une différence significative entre stimuli audios et écrits. Cette partie de l’étude a permis de montrer que lorsque les stimuli n’étaient qu’écrits, aucune différence significative n’est trouvée : aucun indice lexical n’est donc présent dans les stimuli. Lorsque les stimuli étaient audios, et enregistrés sans annoncer la décision dans le message vocal, aucune différence significative n’a été trouvée, ce qui semble indiquer que les locuteurs n’ont pas projeté d’indices de la valence du message dans ce contexte.

Les résultats principaux sont trouvés dans la dernière condition : lorsque les stimuli étaient audios, enregistrés avec l’annonce de la décision, et que les auditeurs entendaient la partie tronquée avant décision. Dans cette condition, il y avait une différence significative dans le jugement des participants à l’étude. La dernière partie de l’étude, est une analyse des corrélats acoustiques (f_0 , énergie, temporalité) et de leurs corrélations avec des jugements sur l’humeur (amicale, plaisante, intéressée...) et les jugements sur la valence (bonne nouvelle, mauvaise nouvelle) perçus dans le message vocal. Nous ne nous étendrons pas les résultats de l’humeur qui n’est pas évaluée dans notre étude. En

revanche, dans la condition « avec décision » le pitch maximum et minimum, sont négativement corrélés avec la valence perçue par l'auditeur. Ainsi, les annonces de nouvelles perçues comme mauvaises sont produites avec un empan vocal plus bas. La vitesse de parole est positivement corrélée avec la valence perçue, c'est-à-dire que les annonces de mauvaises nouvelles seraient produites plus rapidement. Swerts et Hirschberg (2010) précisent que bien que les corrélats acoustiques trouvés dans la condition « sans décision » diffèrent de ceux que nous venons de présenter, le consensus entre les participants pour la perception d'un message comme positif ou négatif était plus grand dans la condition « avec décision ». Ils font donc l'hypothèse que les corrélats trouvés dans la condition « avec décision » seraient ceux utiles pour porter la bonne nouvelle ou la mauvaise. Ils ajoutent dans leur conclusion que les indices précoces du contenu affectif d'un message sont probablement pertinents en ce qu'ils guident l'interlocuteur destinataire de la nouvelle pour se préparer mentalement à la bonne nouvelle ou à la mauvaise nouvelle.

Cette étude démontre la présence de corrélats acoustiques prosodiques en amont d'une annonce de nouvelle. Elle indique aussi que les interlocuteurs qui écoutent l'annonce semblent utiliser les « indices de prédiction » pour guider leur décision dans la tâche d'évaluation de la valence. Swerts et Hirschberg (2010) soulignent, enfin, que leur étude permet de démontrer que les locuteurs, lorsqu'ils doivent annoncer une nouvelle, sont influencés par la valence du message qu'ils doivent transmettre. Cela révèle le fait que le locuteur qui annonce se sent mal à l'aise à l'idée de devoir annoncer des nouvelles déplaisantes lorsqu'il doit aussi annoncer la décision.

Il semble alors crucial de s'intéresser à la présence de ces indices. Ce que Swerts et Hirschberg (2010) appellent « indices de prédiction » sont très proches des « indices de projection » de l'Analyse Conversationnelle. Cependant, la notion de prédiction sera à notre avis moins adaptée pour notre cadre théorique. En effet, l'idée de prédiction ne semble pas prendre en compte la dimension interactionnelle à laquelle nous nous attachons. Elle est généralement plutôt utilisée pour parler des mécanismes cognitifs en jeu dans ce procès. La projection, en revanche, implique en retour l'anticipation de l'interlocuteur. En suivant Swerts et Hirschberg (2010), on peut alors se demander si nous retrouverons les indices de projection de la valence en observant non pas toute l'introduction de l'appel pour annoncer une nouvelle, mais plutôt en observant uniquement la préface à l'annonce de nouvelles. Nous savons que la préface est une partie

de discours pertinente pour la projection d'indices sur ce que l'on va dire, puisqu'elle est dédiée à la projection. Le but de la préface est, pour le locuteur, de projeter des indices sur ce qu'il va dire, et pour l'interlocuteur, de récupérer les indices et anticiper ce qui va être dit. Cela lui permet de se préparer interactionnellement.

Nous pensons comme Swerts et Hirschberg (2010) que les indices de projection de la valence de la nouvelle qui va être annoncée permettent à l'interlocuteur de se préparer « mentalement », et tout particulièrement en termes affectifs. L'interlocuteur se prépare au mieux afin de pouvoir anticiper la valence de la nouvelle et réagir de façon alignée. Pour que l'interlocuteur anticipe cela, il faut que le locuteur projette des indices de la valence affective dans la préface. Pour ce faire, le locuteur pourrait utiliser dans la préface un lexique qui tende vers une valence ou une autre, comme nous l'avons précédemment établi. Cependant, il pourrait aussi le faire à travers la prosodie affective, en produisant une préface à l'annonce de nouvelle avec des corrélats phonétiques et phonologiques qui seraient liés à une valence positive, ou des corrélats liés à une valence négative. Dans cette Thèse, nous allons observer les corrélats phonétiques mais aussi phonologiques des indices de projection de la valence affective dans la prosodie. Nous nous intéresserons à la f_0 , la vitesse d'articulation et la durée des énoncés, ainsi qu'aux caractéristiques spectrales de la voix. Nous observerons également les patrons intonationnels réalisés avant l'annonce d'une nouvelle et leur implémentation phonétique. Enfin, nous regarderons si les indices de projection présents avant l'annonce dans les préfaces permettent aux interlocuteurs d'anticiper la valence de la nouvelle et de s'y préparer affectivement.

Au cours de cette section 3. , nous avons abordé plusieurs thématiques tournant autour de la valence affective. Dans la sous-section 3.1 sur les différents types d'affects, nous avons évoqué la multitude de termes utilisés pour parler de cette notion : émotion, les attitudes, le *stance*, les humeurs. Nous avons expliqué que cette multitude de noms, de termes dans la littérature, parce qu'ils sont parfois utilisés de façon interchangeable, mènent à une confusion générale dans leur définition. Dans la sous-section 3.2 sur l'enregistrement de la voix affective, nous avons abordé les types de corpus qui existent (corpus actés, corpus naturels, corpus enregistrés par des participants naïfs). Nous en avons ensuite listé et discuté des avantages et des inconvénients. Dans la sous-section 3.3 sur les indices de projection qui peuvent être attendus lorsque nous étudions la projection d'affects en interaction, nous avons évoqué les corrélats acoustiques phonétiques qui ont été trouvés dans la littérature sur la prosodie affective. Nous avons vu que les émotions basiques telles que la joie, la colère ou la tristesse ont été très souvent étudiées, mais que la recherche s'est étendue à d'autres types d'affects, et d'autres classifications des affects comme les axes de la valence et de l'excitation de Russel (1980). Nous avons expliqué l'intérêt que nous avons pour les corrélats de la tristesse, que nous rapprochons de ce que nous pourrions trouver dans les annonces de nouvelles à valence négative, et pour les corrélats de la joie, que nous rapprochons de ce que nous pourrions trouver dans les annonces de nouvelles à valence positive. Puis, nous avons discuté de trois études récentes qui se sont intéressées aux corrélats phonologiques de l'affect en anglais américain et ont trouvé que certains patrons intonatifs ou types d'accents revenaient plus souvent selon l'affect. Dans la sous-section 3.4 sur la projection de structure et la projection affective, nous avons fait la description de l'étude qui a fortement inspiré ce travail : celle de Swerts et Hirschberg (2010). Ils se sont intéressés, dans un corpus d'annonces de nouvelles données sur répondeur téléphonique, aux informations portées dans la prosodie sur la valence affective de la nouvelle. Précisément, ils ont regardé quels étaient les corrélats phonétiques qui étaient faits en fonction de cette valence, positive ou négative, dans la préface de l'annonce. Ils ont ensuite regardé si les annonces étaient bien reconnues comme positives ou négatives, et ont regardé les corrélations entre les facteurs phonétiques et le jugement des auditeurs sur la valence de la nouvelle.

C'est ici que se termine ce chapitre d'état de l'art. Nous avons évoqué trois grands thèmes importants pour cette thèse. En premier, la notion de projection dans la préface et tout ce qui y est relié, dans le cadre de l'Analyse Conversationnelle et de la Linguistique Interactionnelle. En deuxième, les notions de prosodie affective et d'intonation, à travers le prisme de la phonétique et celui de la phonologie. En troisième, la notion d'affect et celle de valence, les corrélats phonétiques et phonologiques trouvés dans la littérature jusqu'ici. Ces trois grandes sections nous ont permis d'expliquer ce que sont la projection, la prosodie, et la valence affective séparément, pour présenter le cœur de notre sujet d'étude : la projection prosodique de valence affective, dont nous souhaitons évaluer si et dans quelle mesure elle existe dans les préfaces à l'annonce de nouvelles, et comment elles se manifestent au travers de la phonétique et la phonologie en français.

Chapitre 2 – Etude de production : indices de projection phonétiques de la valence affective

1. Problématique

Dans ce chapitre, nous présentons le processus de création du corpus qui nous a permis d'observer la production et la perception d'indices de projection, ainsi que l'analyse phonétique des productions des participants. Ces analyses doivent nous permettre de répondre à l'une de nos questions de recherche : voir si la valence affective est portée dans la préface par des indices de projection phonétiques, au travers de caractéristiques acoustiques continues. Ainsi, nous souhaitons voir si le locuteur qui annonce une nouvelle projette des indices phonétiques de la valence affective de l'annonce. Si c'est le cas, nous souhaitons découvrir quels indices phonétiques entrent en jeu dans cette projection, et où ils sont utilisés.

Nous avons donc créé un corpus, inspiré de celui de Swerts et Hirschberg (2010), pour étudier le rôle projectif de la prosodie, et répondre à une partie de nos questions de recherche. Il s'agit d'un corpus acoustique, créé à partir d'une tâche de lecture. Le choix d'une tâche de lecture plutôt qu'une tâche de production spontanée nous permet de contrôler certains facteurs qui peuvent compliquer la comparaison de la prosodie entre les enregistrements, comme la variation segmentale et syntaxique. Ce corpus oral porte sur des nouvelles qui pourraient être annoncées dans un milieu institutionnel ou professionnel, ce qui nous permet d'observer les corrélats phonétiques de la valence de la nouvelle dans une multitude de contextes. Trois valences affectives peuvent être portées par la nouvelle, suivant l'axe de la valence : valence positive, valence négative, valence neutre. Chaque annonce de nouvelle est composée de plusieurs parties suivant globalement la structure décrite pour les annonces de nouvelles par Terasaki (1976), et celle des conversations téléphoniques décrites par Schegloff (1986) : ouverture, préface, marqueurs discursifs, séquence d'annonce, fermeture. Nous allons maintenant décrire plus en profondeur les différents éléments qui nous ont conduits à choisir la préface et les marqueurs discursifs « voilà donc », ainsi que la séquence d'annonce, comme objets d'étude plutôt que les séquences d'ouverture ou de fermeture. Ensuite, nous expliquerons pourquoi nous avons choisi d'étudier certains paramètres acoustiques plutôt que d'autres.

1.1. Séquences interactionnelles

Dans l'état de l'art général de ce document, nous avons rapporté plusieurs éléments essentiels pour donner le cadre théorique de ce chapitre. Nous avons rapporté ce que la littérature décrit du fonctionnement d'une annonce de nouvelle, nous avons discuté de la notion de préface, et des marqueurs discursifs. Nous avons défini le type d'affects auxquels nous nous intéressons. Enfin, nous avons présenté les corrélats phonétiques de la prosodie, les effets de la valence affective sur ceux-ci, et ceux trouvés spécifiquement dans le cadre des annonces de nouvelles. Ce que nous avons présenté montre qu'il existe des éléments interactionnels qui permettent au locuteur qui annonce une nouvelle de projeter des indices concernant la valence de cette nouvelle, et que certains pourraient se trouver dans la prosodie. Avant de déterminer quels corrélats exactement nous souhaitons étudier, nous avons dû définir les « lieux » interactionnels sur lesquels porter notre intérêt.

Nous avons pris des mesures pour la préface, les marqueurs discursifs « voilà donc », et la séquence d'annonce séparément. La préface étant l'objet principal de notre étude, nous y avons fait une analyse pour chacune des caractéristiques acoustiques observées. Cependant pour certaines mesures que nous détaillons dans la section suivante, nous n'avons pu étudier qu'une seule partie de la préface (« Je vous appelle au sujet ») du fait de la variation des items en fonction du scénario de la nouvelle dans la seconde portion de la préface (« Je vous appelle au sujet de votre chat Minouche » vs « Je vous appelle au sujet de votre enfant Léo »). Les marqueurs du discours, du fait de leur importance structurelle, de leur placement stratégique entre la préface et la séquence d'annonce, et de leur potentielle utilisation en tant que préface (1.2), ont aussi été analysés. Leur constance verbale (« voilà donc ») nous permet une analyse des indices de projection phonétique pour la plupart des mesures que nous prendrons. Enfin, nous avons choisi d'observer la projection prosodique au sein de l'annonce, afin de pouvoir comparer les corrélats phonétiques entre ce qu'il se passe avant que la nouvelle n'ait été dite, et ce qui se passe pendant. Nous pourrions ainsi voir si la direction des effets reste la même ou si elle change, une fois que l'on est plus en train de préparer la nouvelle en amont, mais plutôt en train de l'annoncer. Cependant, la forte variation inter-item et intra-item (ex. « il est toujours malade, et le traitement va coûter cher » pour la nouvelle négative ; « il est guéri, et vous pourrez venir le chercher dès demain » pour la nouvelle positive) ne nous permet de prendre que peu de mesures phonétiques comparables. Nous détaillons

maintenant les mesures phonétiques que nous avons choisies ainsi que le domaine temporel exact sur lequel elles sont extraites.

1.2. Paramètres acoustiques

L'étude des corrélats phonétiques est essentielle pour comprendre le fonctionnement de l'annonce mais surtout, pour observer la fonction projective de la prosodie affective en interaction. Cependant, il existe une infinité de mesures acoustiques qui peuvent être prises sur une production orale, et nous avons donc dû faire des choix et sélectionner certaines mesures parmi les quatre composantes de la voix : la fréquence fondamentale, l'intensité, les mesures temporelles, et les caractéristiques spectrales. Pour l'intensité tout d'abord, nous avons choisi de ne pas prendre de mesures. Les mesures d'intensité sont sensibles à la distance entre le micro et le participant. Or, c'est un paramètre que nous n'avons pas eu la capacité de réguler. Les enregistrements du corpus ayant été réalisés pendant la crise sanitaire du COVID-19, nous avons choisi des micros sur pied plus respectueux des mesures sanitaires, au lieu des micro-casques habituellement utilisés pour contrôler la distance bouche – micro. Nous nous sommes donc concentrés sur la f_0 , les données temporelles, et les caractéristiques spectrales.

Pour la fréquence fondamentale, nous avons choisi de prendre quatre mesures : la f_0 minimum, la f_0 maximum, l'empan vocal, et la moyenne de la f_0 . Ces mesures ont été prises en faisant une moyenne sur l'ensemble de l'énoncé étudié. De telles mesures, globales, peuvent aisément être sélectionnées pour comparer des énoncés différents malgré la variation verbale inter ou intra items (comme les préfaces, ou les séquences d'annonce). Nous nous sommes intéressés à la variation de la f_0 car c'est un des paramètres qui permet une bonne reconnaissance de la joie particulièrement mais aussi de la tristesse (Audibert et al., 2005). Il pourrait donc être utilisé pour produire des indices de projection dans le cadre d'annonces de nouvelles positives. Nous pensons donc que ce paramètre pourrait présenter des variations significatives entre la valence positive et les deux autres valences. C'est par ailleurs un paramètre acoustique très souvent analysé dans les études sur les affects, présentant des tendances opposées dans les deux affects que l'on compare à nos valences, avec des valeurs élevées pour la joie, et abaissées pour la tristesse. Ce résultat est trouvé pour la moyenne de la f_0 et l'empan vocal (Berry & Brown, 2019; Liscombe et al., 2003; Scherer, 2003) ; ainsi que pour la f_0 maximum et la

f0 minimum (Liscombe et al., 2003). Les deux études s'étant intéressées à la prosodie dans les annonces de nouvelles ont-elles aussi proposé une analyse de la f0 (Freese & Maynard, 1998; Swerts & Hirschberg, 2010).

Pour les données temporelles, des mesures comme la vitesse de parole ont été démontrées comme faisant une différence entre les affects neutres et d'autres affects ; avec une vitesse plus lente pour le neutre (Vidrascu & Devillers, 2007) et plus rapide pour le négatif par rapport au positif (Scherer, 2003; Swerts & Hirschberg, 2010). Le terme « vitesse de parole » est en général employé pour parler du ratio nombre de syllabes par seconde, avec un calcul incluant les pauses. Cependant, il arrive qu'il désigne un ensemble de mesures telles que la durée du signal (longueur de l'énoncé), la durée des pauses, et le nombre de pauses. Berry et Brown (2019) utilisent justement ce terme en ce sens-là, et trouvent un débit plus rapide et un raccourcissement des pauses lorsque la valence d'un affect est positive. La vitesse d'articulation est, elle aussi, un paramètre qui présente de la variation dans la littérature sur les affects (Scherer, 2003), et qui suit la même direction que la vitesse de parole. Nous avons choisi de nous intéresser aux données temporelles en les observant sous plusieurs angles : vitesse d'articulation, durée de l'énoncé, durée des pauses. La vitesse d'articulation, étant une mesure globale, peut être prise malgré la variation inter et intra item, ce qui a influencé notre choix de l'étudier. Elle a donc été prise dans les trois domaines : préfaces, marqueurs discursifs, annonce. Il en est de même pour la durée des énoncés. La longueur des pauses, quant à elle, n'a pu être évaluée qu'au sein de la préface et des séquences d'annonce. Effectivement, marqueurs discursifs ne contenaient en effet pas assez souvent de pauses pour que cette mesure puisse être réalisée.

Les mesures spectrales enfin, sont étudiées de façons extrêmement variées dans la littérature pour analyser la « qualité de la voix ». Notre choix s'est porté sur trois mesures : le ratio harmoniques – bruit (*Harmonics-to-Noise Ratio* ou HNR, en anglais), la différence entre la première et la deuxième harmonique (H1-H2), et le Pic de Proéminence Cepstrale lissé (*Cepstral Peak Prominence smoothed* ou CPPs, en anglais). La HNR ainsi que H1-H2 sont des mesures très répandues dans la recherche en linguistique pour rendre compte du caractère « soufflé », « modal » ou « craqué » de la phonation d'un individu (Jiang & Pell, 2017; Kuo, 2013; Yang et al., 2020). En effet plus la HNR est basse, plus la voix est perçue comme soufflée ou craquée, alors que lorsqu'elle est haute, la voix est considérée comme saine ou modale (Eskenazi et al., 1990; Krom, 1995; Yumoto, 1983; Yumoto et al.,

1982). La combinaison de cette mesure avec H1-H2 nous permet de déterminer si la voix est soufflée ou bien craquée (Garellek, 2019). Si la HNR est basse et H1-H2 bas (pente spectrale faible) alors la voix est plutôt craquée. Si la HNR est basse et H1-H2 haut, alors la voix est plutôt soufflée. Le CPPs est une mesure plus récente qui rend également compte du caractère soufflé de la phonation. Un CPPs significativement moins prononcé indique une voix plus soufflée (Stoller & Dixon, 2016). L'étude de cette mesure nous apparaît pertinente, de telles analyses ayant démontré de la variation en fonction de la valence émotionnelle dans la littérature. En effet, il semble qu'il y ait plus d'énergie dans les hautes fréquences lorsque l'affect est à valence positive par rapport à lorsqu'il est à valence négative (Johnstone & Scherer, 2000; Scherer, 2003), les voix étant plus soufflées pour la tristesse et modales pour la joie (Laukkanen et al., 1996; Scherer, 1986, p. 198; Yuan et al., 2002). En français, nous trouvons de multiples études en perception qui se sont aussi intéressées à de telles mesures (Audibert et al., 2006; Juslin & Laukka, 2001; Laukka et al., 2012). Goudbeek & Scherer (2010) spécifiquement, ont montré que les affects ont un effet significatif sur les mesures spectrales comme la HNR, aux valeurs plus hautes pour la joie que pour la tristesse, donc une voix plus « saine » dans la joie, et plus « rauque » dans la tristesse. Les mesures des caractéristiques spectrales étant comparées au sein du phonème, et non d'énoncés entiers comme valeur moyenne, nous ne les analysons pas tout à fait de la même façon que les autres mesures. Pour être comparable, il est nécessaire de les observer au sein de phonèmes identiques, et spécifiquement des phonèmes voisés. Le voisement, donc la présence d'harmoniques, est essentiel à ces mesures qui sont calculées avec les harmoniques. Ainsi, nous ne pouvons comparer que des énoncés au contenu verbal identique, phonème par phonème. Les domaines sur lesquels nous pouvons réaliser cela est le début de la préface « je vous appelle au sujet », et les marqueurs discursifs « voilà donc ». Ces deux portions de discours sont identiques pour tous les items.

1.3. Hypothèses

Nous faisons les hypothèses suivantes. Nous pensons qu'un locuteur qui annonce une nouvelle va projeter des indices phonétiques de la valence affective. Ces indices seront similaires à ceux trouvés dans les études sur la joie et la tristesse (Bänziger & Scherer, 2005; Berry & Brown, 2019; Goudbeek & Scherer, 2010; Scherer, 2003), et dans les études

sur la prosodie dans les annonces de nouvelles (Freese & Maynard, 1998; Swerts & Hirschberg, 2010), que nous avons décrit dans l'état de l'art. Nous pensons trouver des annonces de nouvelles positives produites avec des caractéristiques phonétiques proches de la joie, et des nouvelles négatives produites avec des caractéristiques proches de la tristesse. Nous nous attendons à retrouver ces caractéristiques même lorsque l'on observe une partie de discours qui a déjà d'autres fonctions, et à ce qu'elles soient similaires quelle que soit la partie de l'annonce étudiée (préface, marqueurs discursifs, séquence d'annonce). Enfin, nous faisons le postulat que la projection d'indices phonétiques sera perceptible malgré la variété de contextes sociaux d'annonces de nouvelles, et l'absence de nouvelles graves.

Plus précisément pour chaque mesure, nous nous attendons à trouver des valeurs de la f_0 élevées avec un empan plus large, une vitesse d'articulation accrue et des durées écourtées, et plus d'énergie dans les hautes fréquences donc la voix plus modale (caractérisée par un pic cepstral plus haut, une haute HNR ainsi qu'une petite différence entre H1 et H2) lorsque la valence de la nouvelle est positive. A l'inverse lorsqu'elle est négative, nous nous attendons à trouver des valeurs de la f_0 abaissées et un empan plus étroit, une vitesse d'articulation ralentie et des durées allongées, et moins d'énergie dans les hautes fréquences donc une voix plus craquée ou soufflée. Nous nous attendons à ce que les valeurs de ces paramètres, lorsque la valence de la nouvelle est neutre (condition contrôle), soient situées entre celles du positif et du négatif. Nous faisons également l'hypothèse que les différences significatives dans la courbe de la f_0 pourraient être trouvées à certains points-clés de l'énoncé, par exemple sur la fin des syntagmes intonationnels ou accentuels, où l'on retrouve les accents mélodiques et les tons de frontières, ou bien en début d'énoncé.

2. Enregistrement du corpus

2.1. Matériel

2.1.1. Phase de construction 1

Afin de créer le corpus le plus optimal possible pour l'étude de la projection prosodique à travers la phonétique, nous avons utilisé douze scénarios d'annonces de nouvelles, et douze scripts d'annonces de nouvelles. Les scénarios présentent le contexte dans lequel une annonce de nouvelle est faite, et les scripts sont ce que les participants ont lu pour enregistrer les annonces. Ceux-ci résultent de deux tâches préliminaires faites dans une première phase de construction du corpus. Nous allons commencer par décrire cette première phase. Les deux tâches préliminaires nous ont permis de nous assurer que la valence affective des scénarios créés était bien reconnue comme positive, négative, ou neutre (tâche préliminaire 1), et de créer des scripts d'annonces de nouvelles inspirés de productions libres réalisées sur la base des scénarios validés (tâche préliminaire 2).

L'objectif de la tâche préliminaire 1 est de créer 21 scénarios d'annonces de nouvelles (tels que ceux présentés en Tableau 1) et la validation de 12 d'entre eux comme positifs, négatifs ou neutres. Trente participants de langue maternelle française ont pris part à cette étude préliminaire (28 femmes et 2 hommes, âge moyen de 24 ans, SD = 2,5 ans). Le niveau moyen d'études supérieures (niveau 4 de la Cife) était de 1,9 ans (SD = 1,2 ans). Les participants n'avaient pas de troubles de la lecture. Le recrutement a été réalisé via les réseaux sociaux. Les participants n'ont pas été rémunérés. Tous ont validé un formulaire de consentement en ligne sur Qualtrics. Cet outil nous permet de proposer aux participants de répondre à la question qui leur est posée au moyen d'une échelle visuelle analogue (EVA).

L'EVA est une échelle continue horizontale ou verticale, avec un ancrage verbal (par exemple, « positif » à -100 - « négatif » à +100) à chaque extrémité de la ligne continue (Gorrall et al., 2016). Ce type d'échelle vient de la recherche médicale et est préféré aux échelles ordonnées (Laerhoven et al., 2004; D. D. Price et al., 2012; Wewers & Lowe, 1990), mais son utilisation s'est répandue dans la recherche en linguistique (Bänziger & Scherer, 2005; A. Chen et al., 2004; Kreiman et al., 1993). L'échelle visuelle analogue serait

à préférer pour les études en ligne comme la nôtre, car elle donnerait des résultats plus exacts du ressenti des participants et demanderait moins d'explications préalables (Reips & Funke, 2008). Les participants étaient assignés à trois listes au contenu randomisé, contrebalancées pour être courtes mais équilibrées sur le nombre d'items par valence affective. Dans chaque liste nous trouvons, 7 scénarios positifs, 7 négatifs et 7 neutres. Les participants ont jugé chaque scénario, dans une seule version de valence affective. Chaque liste a été évaluée par 10 participants. La question posée était « Est-ce une mauvaise nouvelle / bonne nouvelle ? ». Les participants devaient répondre en utilisant l'EVA à leur disposition. Les extrémités étaient labélisées MAUVAISE NOUVELLE (à gauche, -100) et BONNE NOUVELLE (à droite, 100). Le neutre correspondait donc au milieu, soit 0. Pour les scénarios négatifs et positifs nous avons conservé ceux respectivement notés, en moyenne, en dessous de -75, ou au-dessus de 75. Pour le neutre, nous avons conservé les scénarios dont le score était entre -25 et 50. En effet, la valence « neutre » va plus difficilement être notée par des scores négatifs (en dessous du zéro), et au contraire tendre vers des scores bas, mais au-dessus de zéro. Interactionnellement, les nouvelles neutres et positives sont plus proches par rapport au négatif, puisqu'elles n'engagent pas une action non préférée. La validation de la valence des scénarios nous a permis de conserver les 12 scénarios de l'étude principale (voir Annexe 1 pour le détail des 21 scénarios initiaux).

L'objectif de la tâche 2 était d'obtenir un canevas à partir duquel nous avons créé les annonces de nouvelles entièrement scriptées utilisées pour la tâche de lecture de l'étude principale. Cette production libre d'annonces de nouvelles réalisée en deuxième phase de construction du corpus a été faite sur les 12 scénarios sélectionnés à partir de l'étude préliminaire 1. Quatre participants de langue maternelle française (3 femmes, 1 homme, âge moyen de 33 ans, SD = 15 ans) ont pris part à l'étude. Le niveau moyen d'études après baccalauréat était de 4 ans. Les participants n'avaient pas de troubles de la lecture ou de l'audition. Le recrutement a été réalisé via le bouche-à-oreille. Les participants n'ont pas été rémunérés. Ils avaient pour consigne de produire une annonce de nouvelle comme s'ils laissaient un message sur un répondeur vocal à partir des scénarios validés. Suivant les conditions sanitaires du COVID-19, deux des enregistrements ont pu être réalisés en chambre sourde, mais les deux autres ont dû être réalisés sur la plateforme de conférences en ligne zoom. Les 144 productions d'annonces de nouvelles enregistrées (4 participants X 3 valences X 12 items) ont ensuite été transcrites manuellement (voir Annexe 3). Puis, les transcriptions ont été utilisées comme bases pour créer des scripts

d'annonces de nouvelles au déroulé plus strict et plus court qui ont par la suite été utilisés dans la tâche de production de l'étude principale.

Exemple de production libre, locuteur 2 :

« Bonjour Amélie ici Camille Cumont euh à l'appareil *je vous appelle au sujet de votre demande pour passer en télétravail le vendredi euh voilà donc euh* malheureusement euh vous ne pourrez pas passer en télétravail puisque Monique est déjà en télétravail ce jour-là donc nous avons besoin de vous et je suis vraiment navrée de devoir vous faire venir le vendredi [...] bonne journée au revoir »

Nous pouvons noter plusieurs éléments dans cet exemple. Tout d'abord nous avons la présence d'une préface : « je vous appelle au sujet de votre demande pour passer en télétravail le vendredi ». De plus, nous avons la présence de marqueurs discursifs entre la préface (marquée en italique) et l'annonce : « voilà donc ». C'est la présence de tels marqueurs dans les annonces libres de cette seconde phase de construction du corpus qui nous a poussé à en inclure dans les scripts utilisés pour les enregistrements de l'étude principale.

2.1.2. Phase de construction 2

Pour l'enregistrement du corpus, nous avons réutilisé plusieurs éléments des tâches préliminaires 1 et 2 que nous venons de décrire. Les scénarios (Tableau 1) validés par la tâche 1 décrivaient des contextes dans lesquels une annonce de nouvelle allait être faite. Nous les avons donc utilisés comme supports contextuels proposés aux locutrices pour l'enregistrement, afin qu'elles puissent comprendre la raison pour laquelle elles allaient enregistrer un message. Ces scénarios étaient lus mentalement avant la lecture orale enregistrée des scripts. Les lectures orales se faisaient comme si les locutrices laissaient un message sur un répondeur téléphonique.

VALENCE DE LA NOUVELLE	SCENARIO D'ANNONCE EN CONTEXTE MEDICAL	SCENARIO D'ANNONCE EN CONTEXTE DE SERVICE
POSITIVE	Vous êtes médecin et vous avez reçu une patiente. Vous lui avez fait faire un examen récemment : ses analyses sont meilleures que la dernière fois.	Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol est un long courrier et elle a un siège « classe économique ». Vous pouvez

	Vous l'appellez et laissez un message sur son répondeur pour l'en informer.	la relocaliser en première classe pour le même prix. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.
NEGATIVE	Vous êtes médecin et vous recevez une patiente. Vous lui avez fait faire un examen récemment : ses analyses sont mauvaises et vous devez en discuter. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.	Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol a été annulé et vous n'avez pas trouvé de solution : elle ne pourra prendre l'avion que le lendemain. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.
NEUTRE	Vous êtes médecin et vous recevez une patiente régulière. Vous allez lui faire un examen. Aujourd'hui, vous vous êtes installé dans le cabinet en face du vôtre car ce dernier est en rénovation. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.	Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol a été déplacé du terminal 1 au terminal 4. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Tableau 1 – Exemple de scénario d'annonce de nouvelle

Chaque scénario propose au lecteur de se mettre dans la peau d'une personne avec un rôle social particulier (médecin, professeur), devant annoncer une nouvelle à autrui (patient, étudiant). Une première phrase permet de donner le rôle du lecteur, et de celui qui reçoit l'annonce : « Vous êtes médecin et vous recevez une patiente ». Une deuxième phrase exprime un problème ou une demande précédente de l'interlocuteur, et la solution ou réponse que le rôle du lecteur doit apporter : « Vous lui avez fait faire un examen récemment : ses analyses sont mauvaises et vous devez en discuter ». Enfin, une instruction est donnée au lecteur pour indiquer qu'il va falloir laisser un message vocal : « Vous l'appellez et laissez un message sur son répondeur pour l'en informer ». Selon la nouvelle à annoncer, il y a des contextes juridiques, médicaux, de service ou encore d'enseignement.

Les nouvelles sont déclinées de trois manières possibles : une annonce à valence positive (bonne nouvelle), une annonce à valence négative (mauvaise nouvelle), et une annonce à valence neutre (c'est-à-dire sans connotation affective). Les nouvelles négatives ne sont jamais liées à des événements graves. Ce choix a été fait pour des raisons éthiques : ne pas risquer de déclencher des réactions affectives fortes chez les

participants, liées à leur vécu personnel ou leur sensibilité. De plus, nous avons préféré éviter l'enregistrement de nouvelles graves qui auraient pu créer de fortes variations prosodiques entre les items. Les nouvelles dont la valence est modérément forte risquaient moins de créer cette variation inter-item, ou de déclencher une réaction affective trop forte pour les participants. Pour créer les bonnes nouvelles en revanche, il n'y avait pas autant de contraintes liées à des potentielles réactions négatives fortes chez les participants. Par souci de cohérence, et afin de pouvoir comparer les indices de projection produits sur des affects de même type de valence, nous n'avons pas utilisé d'annonces de nouvelles positives extrêmes (ex., mariage, naissance). Les nouvelles neutres décrivent une information ou un fait. Elles se rapportent à la confirmation d'une information déjà sue ou attendue (un rendez-vous, la réception d'une demande), ou à un délai ou une légère modification d'un élément déjà convenu (venir 30 minutes plus tard, changer un point de rendez-vous).

Dans la rédaction des scripts d'annonces de nouvelles, nous avons respecté plusieurs règles afin que ceux-ci soient cohérents et réalistes. Nous reprenons certains éléments de l'étude préliminaire 2 où les annonces étaient produites librement. Notamment, nous avons inclus des marqueurs discursifs pour le réalisme de l'annonce et la structuration du discours. Nous reprenons aussi certains éléments de la littérature, et avons donc inclus quatre sections ou « phases » de l'annonce de nouvelle. Chaque message contient donc ces quatre « phases », toujours présentées dans le même ordre : ouverture, préface (se terminant par des marqueurs discursifs, ou MDs), annonce, fermeture. Dans la phase d'ouverture, la locutrice se présente. Ensuite, dans la phase de préface, elle donne le sujet général de l'appel sans révéler, au moyen de ses mots, la valence des nouvelles à venir. La préface se termine par le marqueur discursif « voilà donc » qui aide à organiser et structurer le discours en servant de pivot entre les parties avant, et pendant / après l'annonce, et a été ajouté au vu de sa présence dans les productions de l'étude préliminaire 2. Dans la phase d'annonce, la locutrice annonce le message qu'elle devait délivrer, c'est-à-dire une nouvelle positive, négative ou neutre. Enfin, il y a la séquence de clôture, où l'appelant prononce quelques formalités de fin d'appel avant de raccrocher.

Nous avons besoin que les scripts d'annonces soient les plus similaires possible de l'un à l'autre pour que la prosodie y soit comparable : les scripts qui seraient lus pour le corpus devaient être relativement courts, avoir une structure syntaxique identique, et avoir une variation segmentale réduite au minimum. Pour chaque scénario, le contenu lexical de

l'ouverture, de début de la préface, du marqueur discursif à la fin de la préface, et de la phase de clôture est donc constant. En revanche, le contenu lexical de la fin de la préface, et de la séquence d'annonce changent en fonction des spécificités de la nouvelle à annoncer. Nous présentons un exemple de script de message vocal dans le tableau 2 (voir Annexe 4 pour la liste complète des annonces).

Ouverture	Préface	MDS	ANNONCE		Fermeture
Bonjour Madame Monat, c'est Camille Dumont à l'appareil votre chef d'escale	Je vous appelle au sujet de votre vol de ce soir	Voilà donc	L'embarquement aura lieu en porte 1 et non en porte 4	<i>Nouvelle neutre</i>	Donc je reste à disposition pour toute question, bonne journée, à bientôt
			Nous vous offrons une place en première classe, au même prix	<i>Nouvelle positive</i>	
			Il a été annulé et nous ne pouvons pas vous en proposer un autre aujourd'hui	<i>Nouvelle négative</i>	

Tableau 2 – Présentation de la structure des appels d'annonce de nouvelle

2.1. Participants

Trente locutrices natives françaises (âge moyen = 38 ans, SD = 6,5), sans troubles de la vue, de l'audition, ou dyslexie, ont été recrutées pour l'étude. Les nouvelles à annoncer étant dans des contextes professionnels, nous avons défini un âge minimum de 28 ans, et maximum de 47 ans, afin d'augmenter les chances que les personnes soient habituées à laisser un message vocal lié dans ce type de situation. Nous avons calculé les années d'éducation effectuées à partir des années effectuées après le niveau 3 d'étude, d'après la Classification internationale type de l'éducation (ou Cite). Le niveau 3 correspond en France au niveau du baccalauréat. Le niveau 4 correspond de façon internationale à l'enseignement supérieur. Le nombre d'années moyen d'études supérieures était de 4,7

ans (SD = 1,2). Le recrutement a été réalisé via les réseaux sociaux et distributions de flyers. Les participantes ont été rémunérées 10 euros chacune. Toutes ont signé un formulaire de consentement avant de démarrer l'enregistrement. L'étude de production présentée ici ainsi que celle de perception présentée au Chapitre 4 ont été validées par le comité éthique d'Aix-Marseille Université (référence : 2021-01-07-02).

2.2. Procédure

2.2.1. Enregistrements

Les consignes données aux participants étaient de lire les annonces en se mettant à la place du locuteur qui va appeler et laisser un message sur répondeur pour annoncer une nouvelle. Chaque participant voyait : les 12 scénarios de nouvelles x les 3 variations selon la valence, soit 36 items. Les scénarios de contextualisation et les scripts à lire et enregistrer leur étaient présentés sur un PowerPoint à la suite des consignes, elles aussi écrites. Ils étaient divisés en 3 blocs suivant un ordre strict : les nouvelles neutres, les bonnes, puis les mauvaises. Nous les avons placés dans cet ordre afin de s'assurer que la production des annonces neutres ne serait pas influencée par une annonce de mauvaise nouvelle juste avant, par exemple. En effet, nous avons observé de façon informelle que lorsque l'on enchaîne plusieurs productions avec des valences différentes, on peut se laisser influencer par les valences affectives précédentes. Nous avons donc séparé par valence, et ordonné de façon que le neutre, à faible activation, soit fait en premier. Au début de chaque bloc, deux essais d'entraînement étaient proposés. Ils permettaient de s'assurer que le participant avait bien compris le changement de bloc et de faire tampon entre les valences des blocs.

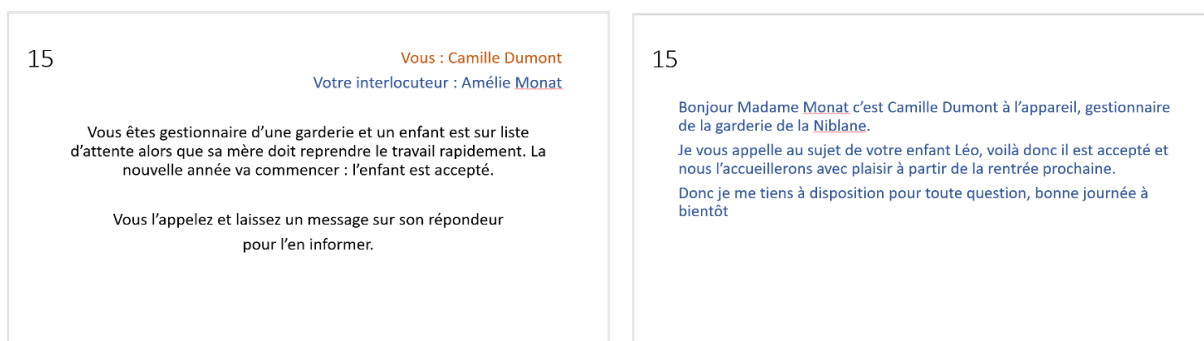


Figure 13 – Présentation des stimuli pour l'enregistrement des annonces de nouvelles

Les enregistrements ont été faits en chambre sourde via un micro à pied et étaient sauvegardés sous le logiciel Audacity. En respectant les règles sanitaires mises en place pour le COVID-19, le participant n'était invité à retirer son masque qu'une fois installé seul dans la chambre sourde. La durée totale de l'enregistrement étant d'environ 40 minutes, une pause était imposée entre chaque bloc pour que le participant puisse boire, bouger et reposer sa voix afin d'éviter tout risque de fatigue vocale.

2.2.2. Mesures phonétiques

Afin de pouvoir réaliser les analyses phonétiques telles que faire des mesures de la vitesse de parole ou plus tard, des annotations phonologiques, nous avons fait des transcriptions automatiques du corpus. Les transcriptions ont été réalisées semi-automatiquement en utilisant la fonction d'alignement forcée de SPPAS 3.7 (Bigi, 2015; Bigi & Meunier, 2018). L'alignement a été réalisé au niveau du mot, de la syllabe et du phonème. Le logiciel SPASS permet en effet, en fournissant un texte prédécoupé au format texte, de créer automatiquement un fichier TextGrid adapté pour le logiciel Praat (Boersma & Weenink, 2023) afin de pouvoir l'afficher aligné au fichier son, avec plusieurs tires à choisir en amont. Les deux tires qui nous intéressaient étaient la tire d'intervalles « TokensAlign », qui fournit le découpage par mots, et la tier SyllAlign, qui fournit le découpage par syllabes. Ces tires nous permettent le calcul de la vitesse d'articulation, et nous serviront par la suite à l'annotation phonologique de l'intonation des préfaces et marqueurs discursifs. Pour les analyses des caractéristiques spectrales, nous avons par la suite ajouté la tier PhonAlign, qui fournit le découpage par phonème, que nous n'avons utilisé que pour certaines analyses décrites plus bas.

La valeur de détection des pauses était mise à 200 ms minimum, le paramètre recommandé par défaut pour le français (Bigi & Priego-Valverde, 2019). Les pauses sont automatiquement annotées avec un dièse « # ». La première étape de l'annotation automatique est une segmentation en unités inter-pausales (IPU, tire appelée « Transcription »), qui découpe la transcription donnée et l'aligne à l'audio en fonction des pauses proposées dans le fichier texte (signalées manuellement par un retour à la ligne) et des pauses détectées dans l'audio (Bigi & Hirst, 2012). Cette tire permet l'annotation des suivantes, et met en avant les pauses produites par le locuteur. Bigi et al. (2010) précisent qu'après le découpage en IPU, vient la phonétisation : découpage au niveau du phonème. Dans SPPAS, la phonétisation est basée sur un système mélangeant l'utilisation d'un dictionnaire de phonèmes existants pour les langues, et sur des règles d'inférence établies sur la base d'annotation de linguistes experts d'une langue donnée, afin de repérer au mieux les frontières entre les phonèmes et de les reconnaître. La phonétisation est alignée à l'audio puis utilisée pour la syllabification qui est elle aussi basée sur un système de règles élaboré pour le français. Nous avons ensuite vérifié et corrigé à la main les grandes erreurs de détection (absence de détection d'une pause de plus de 200 ms, décalage dans l'annotation des mots ou syllabes pouvant être dû à mauvaise détection d'IPU) sous Praat.

Un total de 1080 observations (12 scénarios x 3 valences x 30 locutrices) a été obtenu à l'enregistrement et 1078 ont été conservées. En effet deux éléments ont été exclus en raison d'une mauvaise lecture des marqueurs du discours par les locutrices (inversion ou omission d'un des deux mots, « donc euh voilà » au lieu de « voilà donc »).

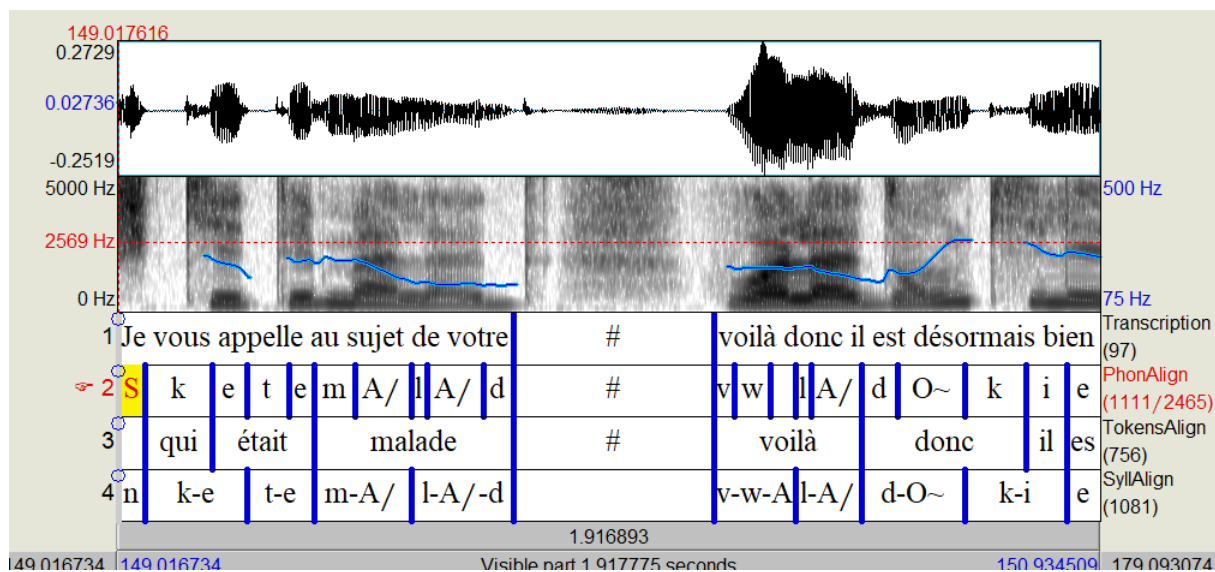


Figure 14 – Fichiers Praat et audio avec les tires d’annotation automatique : Transcription, PhonAlign, TokensAlign et SyllAlign

Après ces corrections, nous avons extrait les données phonétiques à l’aide de scripts PRAAT dans les phases du discours étudiées, les préfaces et les marqueurs du discours, ainsi que dans l’annonce de nouvelles afin de pouvoir comparer le comportement des indices phonétiques dans cette partie du discours. En ce qui concerne les analyses de la f_0 , nous avons mesuré la f_0 maximum, minimum, les valeurs moyennes et l’empan de la f_0 . Pour les données temporelles, nous avons observé la vitesse d’articulation, donc le nombre de syllabes par seconde (ce qui exclut les pauses). Nous avons aussi mesuré la durée de l’énoncé, pour les préfaces et marqueurs discursifs seulement, les séquences d’annonce étant différentes au sein d’un seul item, donc de longueurs inégales non comparables. Enfin, nous avons mesuré la durée des pauses lorsqu’elles étaient présentes, au sein des préfaces et séquences d’annonces, mais pas au sein des marqueurs discursifs (trop peu d’occurrences comportaient des pauses). Enfin, pour les caractéristiques spectrales, nous avons pris des mesures du rapport harmoniques/bruit (HNR), de la différence entre la première et la deuxième harmonique (H1-H2) et de la proéminence du pic cepstral lissé (CPPs). Les valeurs de la HNR ont été prises sur plusieurs niveaux : HNR05 pour une mesure dans l’intervalle 0- 500Hz, HNR15 (0-1500Hz), HNR25 (0-2500Hz) and HNR35 (0-3500Hz) (Kyaw & Sagisaka, 2017). Cela nous permet de nous concentrer sur différentes plages de fréquences du signal vocal. Par exemple, HNR05 peut être plus sensible aux bruits ou aux irrégularités de basse fréquence, tandis que HNR35 peut capturer des informations dans les composantes de haute fréquence. En analysant la

HNR à différentes bandes de fréquences, nous obtenons des mesures plus précises que si l'on prenait une mesure globale de la HNR. Ces analyses des caractéristiques spectrales ont pu être réalisées sur le début des préfaces, et les marqueurs discursifs. Les séquences d'annonces, elles, ne permettent pas l'analyse des caractéristiques spectrales car comme la seconde partie des préfaces, elles varient en termes de contenu verbal.

Nous allons maintenant décrire les méthodes d'extraction de données que nous avons utilisées. Pour les valeurs de la f_0 , extraites en Hertz (Hz), nous avons adapté un script existant (Lennes, 2017) pour automatiser l'extraction de f_0 maximum dans plusieurs fichiers sonores à un intervalle donné. Nous avons ajouté l'extraction de la f_0 minimum et moyenne à l'aide des fonctions intégrées de Praat. Ensuite, nous avons calculé l'empan de la f_0 en utilisant les valeurs maximale et minimale (empan vocal = f_0 maximum - f_0 minimum). Après une première extraction des données, nous avons découvert un effet « plafond » et un effet « sol » pour les valeurs maximum et minimum à l'observation de graphiques *boxplots* réalisés sous R avec le package *ggplot2* (Wickham, 2016), c'est-à-dire que les valeurs étaient biaisées vers une valeur maximum et une valeur minimum qu'elles ne pouvaient pas dépasser. Pour remédier à cela et éliminer les valeurs extrêmes présentes en raison d'erreurs de détection de la f_0 , nous avons pris les valeurs maximale et minimale respectivement au 90e quantile et au 10e quantile. Les mesures étaient faites en Hertz (Hz), et la fenêtre de calcul a été réglée sur 75Hz à 350 Hz.

Pour les paramètres de données temporelles (vitesse d'articulation, durée de l'énoncé, durée des pauses), nous avons créé un script sur Praat permettant de calculer le nombre de syllabes dans la préface ou dans les marqueurs discursifs, ainsi que la durée totale en excluant ou en incluant la durée des pauses du calcul. La vitesse de parole est donc calculée en syllabes par seconde. La durée de l'énoncé et la durée des pauses est calculée en secondes.

En ce qui concerne les mesures des caractéristiques spectrales, nous avons utilisé *PraatSauce* (Kirby, 2018), un ensemble de scripts Praat développés d'après *voiceSauce* (Shue et al., 2009). Il peut extraire au niveau du phonème, sur le début de la préface et les marqueurs discursifs « voilà donc », entre autres, plusieurs mesures de la HNR (HNR 15, 20, 30 et 35), les deux premières harmoniques H1 et H2 et le CPPs. Il offre la possibilité d'extraire plusieurs points équidistants au sein du segment. Nous avons opté pour la seconde option, qui nous permettra par la suite un alignement temporel entre les préfaces pour pouvoir les comparer. Pour déterminer combien de points équidistants prendre,

nous avons procédé à une recherche au sein de la littérature où aucun consensus ne ressort. Certains chercheurs travaillant sur les syllabes ont pris 7 points sur seulement le deuxième tiers du segment pour éviter des effets de coarticulation (Kaiser & Bořil, 2018), certains supprimaient les points en bordure de consonnes (Rhee et al., 2019). Cependant, les effets de coarticulation devraient être déjà pris en compte dans nos données, car nous ne prenons pas de valeurs dès le début du segment ou à la toute fin et nous nous intéressons à l'évolution des caractéristiques spectrales sur une portion de discours plus longue que la syllabe. Deux nombres d'extraction de points assez communs sont 5 (Gao & Hallé, 2015; Hussain, 2021; Martínez Garcia, 2020), et 10 (Icht et al., 2021). Nous avons donc extrait 5 points équidistants sur chaque segment de la partie identique dans la préface (« Je vous appelle au sujet »), ainsi que sur chaque segment des marqueurs discursifs (« voilà donc »).

2.2.3. Tests statistiques

Afin d'analyser les données récoltées, dans la préface, dans les marqueurs du discours, et dans l'annonce séparément, nous avons utilisé des modèles linéaires mixtes (Linear Mixed Effect Model, LMER, Pinheiro & Bates, 2000) car les variables observées sont continues. Les modèles linéaires mixtes permettent de prendre en compte les facteurs fixes et aléatoires dans une analyse. En utilisant ce type de modèle, nous avons testé l'effet de la VALENCE (positive, négative, neutre) sur chacune des variables, soit les corrélats de la f_0 (empan vocal, f_0 moyenne, f_0 maximum, f_0 minimum), et les données temporelles (vitesse d'articulation, durée des énoncés, durée des pauses). Pour la f_0 , nous avons choisi de prendre les valeurs en log (ex., $\log(\text{Maximum}.f_0)$) afin de nous assurer d'avoir une distribution normale des données. Néanmoins dans les graphiques, les données de la f_0 restent représentées en données brutes en Hertz pour faciliter leur exploration. Notre modèle maximal comportait quatre éléments spécifiés aléatoires. Un intercepte aléatoire ITEM pour rendre compte de la variation inter-items, un intercepte aléatoire LOCUTEUR pour rendre compte de la variation inter-individuelle, et une pente aléatoire VALENCE par LOCUTEUR. Nous avons comparé ce modèle maximal à d'autres modèles emboîtés, qui différaient par une seule variable (par exemple, absence de la pente VALENCE par ITEM). La comparaison entre les modèles est faite au moyen des tests de vraisemblance avec la fonction *anova()* de R. Plus précisément, nous avons évalué si l'ajout de variables explicatives supplémentaires (les interceptes et pente) dans un modèle améliore la

capacité du modèle à expliquer la variance dans la variable dépendante. Le test anova donne plusieurs valeurs pour comparer les modèles emboîtés. Parmi celles-ci, les valeurs BIC (Critère d'Information Bayésien) et AIC (Critère d'Information d'Akaike), ainsi que la valeur p. Pour choisir entre les modèles nous avons observé la valeur p et, suivant Pinheiro & Bates (2000), nous avons choisi le modèle avec les plus petites valeurs BIC et AIC, qui représente le mieux nos données. Le modèle final inclut uniquement la pente aléatoire VALENCE par LOCUTEUR, ainsi que LOCUTEUR et ITEM comme interceptes aléatoires. Nous avons, pour chaque modèle, opéré un changement de l'intercepte sur les données avec la fonction *relevel*, avec dans un premier temps la condition « négatif » comme intercepte, puis « positif » après *relevel*. Cela nous permet de comparer directement les nouvelles positives aux nouvelles neutres. Pour cette raison, nous avons corrigé le niveau alpha de la *p-value* de 0.05 à 0.025 (0.05/2 modèles exécutés par chaque variable dépendante).

Modèle LMER utilisé pour les analyses globales :

$$\text{lmer}(\log(\text{Maximum.f0}) \sim \text{Valence} + (1|\text{Locuteur}) + (1|\text{Item}) + (1|\text{Valence:Locuteur}), \text{data})$$

Pour l'analyse des données des caractéristiques spectrales, nous avons opté pour un autre modèle statistique. Comme nous l'avons établi dans le paragraphe de problématique, ce type de mesures étant très variable en fonction du type de phonèmes dans lesquelles elles sont prises, elles doivent être prises au niveau du phonème entre des énoncés identiques phonologiquement afin de pouvoir faire des comparaisons directes : « je vous appelle au sujet », et « voilà donc ». Communément prises sur des items très courts, ces mesures ne peuvent être analysées de la même façon que la f0, en mesures globales moyennées sur une grande portion de discours. Nous avons donc extrait 5 points par phonème, et la comparaison doit se faire point par point. Ainsi donc, un modèle lmer standard ne rendrait pas compte correctement de la variation spectrale car il ne peut analyser chaque groupe de points séparément. Afin de pouvoir comparer nos énoncés, nous avons opté pour un modèle qui compare les items en prenant en compte l'évolution dans le temps du paramètre étudié : le Modèle Mixte Additif Généralisé (« Generalized Additive Mixed Model » en anglais, ou GAMM). Les GAMM fonctionnent de la même façon que les modèles lmer (modèles mixtes linéaires). Ils permettent d'observer les effets de

variables indépendantes sur des variables dépendantes, en incluant les pentes aléatoires et des intercepts aléatoires : ce sont les termes paramétriques (Sóskuthy, 2017). Cependant, ils permettent en plus de prendre en compte les variables non linéaires, c'est-à-dire des variables sur lesquelles des mesures répétées dans le temps ont été faites, et pour lesquelles on observe une certaine fluctuation (la « non-linéarité »). Dans nos données, les caractéristiques spectrales ont de grandes chances d'entretenir une relation non-linéaire avec la valence de la nouvelle. En effet, nous nous attendons à ce qu'elles varient beaucoup en fonction des caractéristiques acoustiques du phonème qui les contient. Ainsi, il est possible que de des différences ressortent uniquement en comparant au sein de chaque phonème plutôt que sur l'ensemble de l'énoncé (ex., « Je vous appelle au sujet »). La conséquence de la non-linéarité d'une variable est justement que, si l'on se contente d'un modèle linéaire qui rendra compte de la variation globale, aucune différence significative ne ressorte. Pourtant, cela ne signifie pas obligatoirement qu'aucune portion du signal ne présente de différences significatives : elles pourraient se situer sur un seul mot, quelques phonèmes, ou une syllabe particulière. En prenant en compte la non-linéarité des données, les modèles GAMM rendent compte de l'évolution dans le temps de la variable observée et sont donc plus adaptés (Roy et al., 2017).

La prise en compte de la non-linéarité est réalisée par l'inclusion de termes lissés, en plus des termes paramétriques. Le terme lissé est celui du temps. Pour que le modèle puisse rendre compte au mieux de la variation du temps, il faut inclure la spécification du paramètre de régression spline (« bs »). Ce paramètre permet au modèle de faire une interpolation entre plusieurs points fixes appelés « nœuds » (*knots*, en anglais), qui doivent être manuellement choisis, pour créer une courbe qui, n'étant pas linéaire, représentera mieux des données non-linéaires. Le paramètre de régression spline recommandé par défaut pour les modèles GAMM est le « *thin plate regression spline* » (Wood, 2017), car contrairement aux autres, il ne nécessite pas la sélection de nœuds. Le choix des *thin plate* peut être représenté comme suit : `bs="tp"`. Un dernier paramètre à spécifier pour le terme lissé est le choix de la dimension (*basis dimension*, « k »). La dimension permet de définir le nombre maximal de degrés de liberté possibles pour le terme, et ainsi de contrôler le lissage de la courbe afin qu'elle ne soit ni surspécifiée, ni sous-spécifiée. Ainsi, la spécification des paramètres de régression contient trois informations : le terme utilisé pour le lissage, le type de spline de régression, et la dimension. En plus du terme lissé il est possible d'inclure des intercepts ou pentes

aléatoires, comme dans les modèles glmer. Ces intercepts et pentes aléatoires sont eux aussi lissés pour rendre compte de la non-linéarité des données. Ils sont appelés « lissage aléatoire », ou « *random smooth* » en anglais, et sont spécifiés par « bs = re ».

Nos modèles GAMM sont réalisés avec la fonction bam du *package mgcv* sur R (Wood, 2017). Ils testent l'effet de la VALENCE sur les différentes mesures prises : f0, HNR (05, 15, 25, 35), H1-H2 et CPPs. Dans chaque modèle, nous observons l'effet de la VALENCE (négatif/positif/neutre) sur les différentes mesures des caractéristiques spectrales. Nous avons sélectionné comme facteur de lissage (*smooth factor*) le TEMPS (*Time*) par la VALENCE, avec le paramètre de régression spline réglé par défaut « bs= « tp » (Wood, 2017). Le paramètre k de la dimension est réglé à 30, car il doit être réglé au-dessus de 10 au moins lorsque l'on s'attend à une courbe comportant de fortes ondulations, comme les mesures de la parole (Sóskuthy, 2017b). Nous avons sélectionné la valeur k= 30 plutôt qu'une autre en la comparant à d'autres valeurs avec la fonction *gam.check*. Elle nous permet d'obtenir un index k. Si l'index est plus proche de 1 pour k = 30 que pour les autres valeurs, alors k = 30 est considéré comme plus adapté que les autres valeurs pour notre modèle. La valeur k=30 présente un meilleur index k (0.9) que k=10 (0.83). Enfin, le modèle comporte une pente aléatoire LOCUTEUR par VALENCE.

Modèle LMER utilisé pour les analyses de caractéristiques spectrales :

$$\text{bam}(\text{CPPs} \sim \text{Valence} + s(\text{Temps}, \text{by}=\text{Valence}, \text{bs}=\text{"tp"}, k=30) + s(\text{Locuteur}, \text{Valence}, \text{bs}=\text{"re"}), \text{data})$$

Dans nos modèles GAMM, le facteur de lissage TEMPS correspond à un temps normalisé en fonction des segments qui composent l'énoncé analysé. Dans la préface, sur « *Je vous appelle au sujet* » (/ʒøvuzapɛlosyʒe/) le temps normalisé va de 0 à 70 : 1-5 = ʒ, 6-10 = ø, 11-15 = v, 16-20 = u, 21-25 = z, 26-30 = a, 31-35 = p, 36-40 = ε, 41-45 = l, 46-50 = o, 51-55 = s, 56-60 = y, 61-65 = ʒ, 66-70 = e. Dans les marqueurs discursifs « voilà donc » (/vwaladøk/), le temps normalisé va de 1 à 40 : 1-5 = v, 6-10 = w, 11-15 = a (a1), 16-20 = l, 21-25 = a (a2), 26-30 = d, 31-35 = ø, 36-40 = k. Cette normalisation du temps nous permet de représenter l'évolution des mesures prises au cours de la préface et des marqueurs discursifs sans être gêné par les différences de vitesse de parole, en comparant chaque phonème dans un énoncé à celui qui lui correspond dans un autre.

Les modèles GAMM rendent compte visuellement de cette variation temporelle, au moyen d'un graphique de données variant dans le temps (Xiao et al., 2023). C'est sur ces graphiques que nous pouvons voir apparaître les portions du signal où la valence affective a créé des différences significatives sur les caractéristiques spectrales (mis en évidence en rouge dans les graphiques présentés dans la section résultats ci-dessous). Par ailleurs, les modèles GAMM fournissent un pourcentage de déviance, qui indique la proportion de données dont la variance est expliquée par les prédicteurs du modèle (la valence et la pente aléatoire). Cela permet de mieux se rendre compte de leur rôle dans les variations significatives perçues.

Pour présenter nos résultats, nous devons donc prendre en considération plusieurs éléments afin de comprendre l'influence de la valence sur le paramètre observé, généralement toujours rapportée dans la littérature utilisant les GAMM : le pourcentage de déviance (Arnold et al., 2013), les termes lisses et paramétriques (Arnhold & Kyröläinen, 2017), et les segments temporels où l'on trouve des différences *significatives* (Kaland et al., 2023). Tout d'abord, nous rapporterons le pourcentage de déviance, puisqu'il donne une indication sur le degré de variance expliqué par le modèle. Ensuite, nous rapportons les p-values du modèle pour les termes lissés, avant de discuter des graphiques rattachés à ces informations. Les graphiques montrent en effet les zones où les différences significatives peuvent être trouvées en fonction du temps normalisé (la temporalité normalisée exacte pour les zones de significativité est rapportée en Annexe 8). Enfin, nous rapporterons les p-values des coefficients paramétriques. Cela nous permet de voir si les différences significatives localisées dans le temps se retrouvent sur la globalité de l'énoncé ou si elles sont rattachées uniquement à certains points précis de celui-ci. Par exemple, la f_0 pourrait être plus haute sur l'ensemble « je vous appelle au sujet » dans la préface, ou au contraire seulement sur les phonèmes /a/ et /y/, ou encore sur un mot particulier comme « appelle » uniquement. Le premier modèle GAMM sera présenté avec un tableau des coefficients paramétriques et termes lisses, afin de présenter le fonctionnement du modèle. Par la suite, ces tableaux seront proposés en annexes (Annexe 8).

3. Résultats

Nous présentons dans cette section les résultats des analyses phonétiques faites sur le corpus. Nous commençons ce compte-rendu par la présentation des analyses globales de données de f_0 (empan vocal, moyenne, maximum, minimum), suivies des analyses GAMM de la f_0 . Puis, nous présentons les résultats d'analyses des données temporelles (vitesse d'articulation, durée de l'énoncé, durée des pauses) réalisées avec les modèles LMER. Enfin, nous présentons les résultats des analyses des caractéristiques spectrales (HNR, H1-H2, CPPs) avec les modèles GAMM. Dans chacune de ces sous-parties, nous parlons d'abord des préfaces, puis nous continuons par les marqueurs discursifs, et nous terminons par la séquence d'annonce. Par souci de clarté et pour faciliter la lecture, nous donnerons les résultats statistiques non significatifs complets en Annexe 6 pour les modèles réalisés sur la f_0 , et en Annexe 7 pour les modèles réalisés sur les données temporelles.

3.1. Mesures globales de la f_0

3.1.1. Au sein de la préface

Comme nous nous y attendions, nous avons trouvé des différences significatives de hauteur de la f_0 et largeur d'empan dans les mesures globales prises sur la préface. Ces différences apparaissent comme très marquées entre les valences positive et négative, marquées entre les valences négative et neutre, et peu marquées entre les valences positive et neutre. Le Tableau 3 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

Valence	Empan de la f0		f0 moyenne		f0 maximum		f0 minimum	
	Moy.	SD	Moy.	SD	Moy.	SD	Moy.	SD
Négative	67,1	16,8	214,5	23,4	250,6	27,0	184,9	21,9
Positive	93,3	22,5	230,7	22,7	280,4	25,4	188,4	22,2
Neutre	78,8	19,2	224,4	22,3	265,8	26,0	188,2	21,9

Tableau 3 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans la préface

Nous voyons que les valeurs de l’empan vocal présentées en Figure 15 présentent des différences significatives entre toutes les conditions et suivent les mêmes directions. L’empan est plus étroit dans les préfaces lorsque la valence de la nouvelle est négative par rapport à lorsqu’elle est positive [$\beta = -0.16$, SE = 0.021, $t = -7.53$, $p < 0.0001$]. Il est également plus étroit lorsque la valence de la nouvelle est négative par rapport à lorsqu’elle est neutre [$\beta = -0.11$, SE = 0.021, $t = -5.20$, $p < 0.0001$]. Enfin, l’empan apparaît plus large lorsque la valence de la nouvelle est positive par rapport à lorsqu’elle est neutre [$\beta = 0.05$, SE = 0.02, $t = 2.31$, $p < 0.025$].

Nous remarquons les mêmes tendances pour les valeurs moyennes (Figure 16), maximum (Figure 17) et minimum (Figure 18) de la f0. Les valeurs sont abaissées lorsque la valence de la nouvelle est négative par rapport à lorsqu’elle est positive : pour la f0 moyenne [$\beta = -0.04$, SE = 0.006, $t = -7.33$, $p < 0.0001$], maximum [$\beta = -0.06$, SE = 0.007, $t = -8.34$, $p < 0.0001$], et minimum [$\beta = -0.02$, SE = 0.007, $t = -2.88$, $p < 0.01$]. Les valeurs sont également abaissées lorsque la valence de la nouvelle est négative par rapport à lorsqu’elle est neutre : pour la f0 moyenne [$\beta = -0.04$, SE = 0.006, $t = -3.92$, $p < 0.001$] et maximum [$\beta = -0.03$, SE = 0.007, $t = -4.97$, $p < 0.0001$]. Elles sont élevées lorsque la valence de la nouvelle est positive par rapport à lorsqu’elle est neutre : pour la f0 moyenne [$\beta = 0.02$, SE = 0.007, $t = 3.4$, $p < 0.01$] et maximum [$\beta = 0.02$, SE = 0.007, $t = 3.36$, $p < 0.01$].

En revanche pour la f0 minimum (Figure 18), les différences ne ressortent pas statistiquement entre valences négative et neutre [$p = 0.09$], et valences positive et neutre [$p = 0.2$]. Pour la valeur minimum de la f0, nous avons calculé sur un sous-groupe de données réalisé sur l’ensemble des données, en excluant les anomalies de détection de la f0, repérées grâce à une première visualisation par *boxplots*. Nous avons exclu les *outliers* présents dans ces données (ces *outliers* étaient en général sous 120 Hertz), ce qui représentait 3% des données.

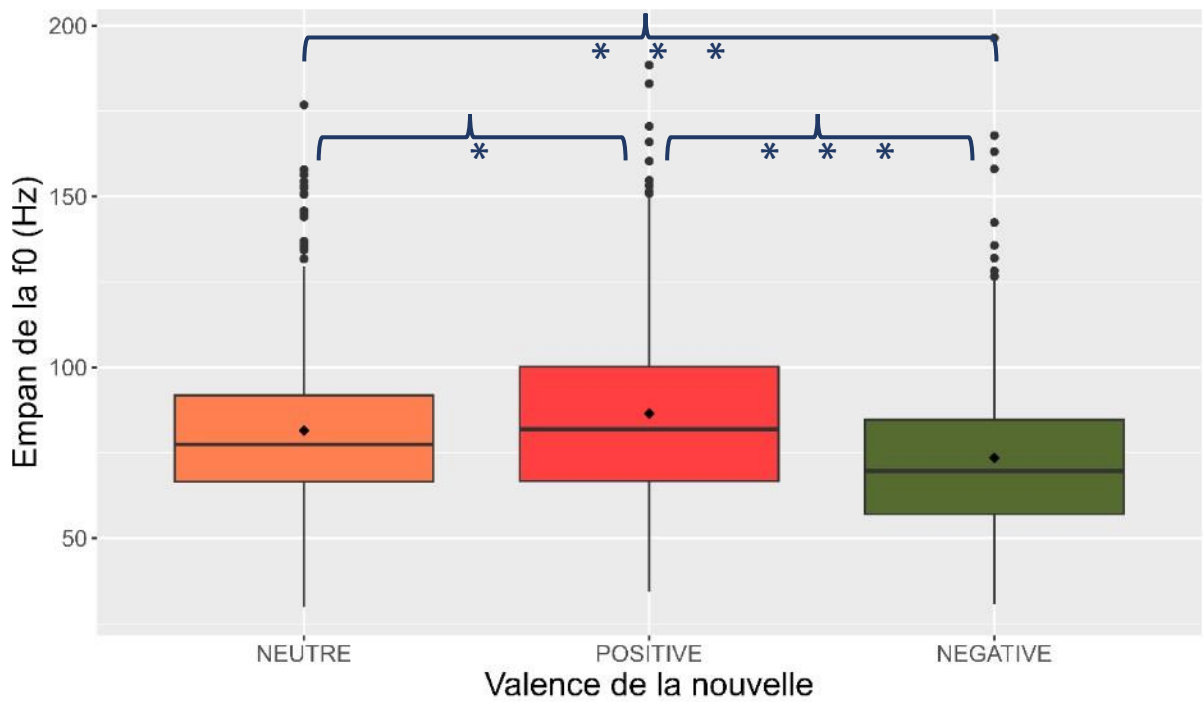


Figure 15 – Empan de la f0 (Hz) dans les préfaces par valence de la nouvelle

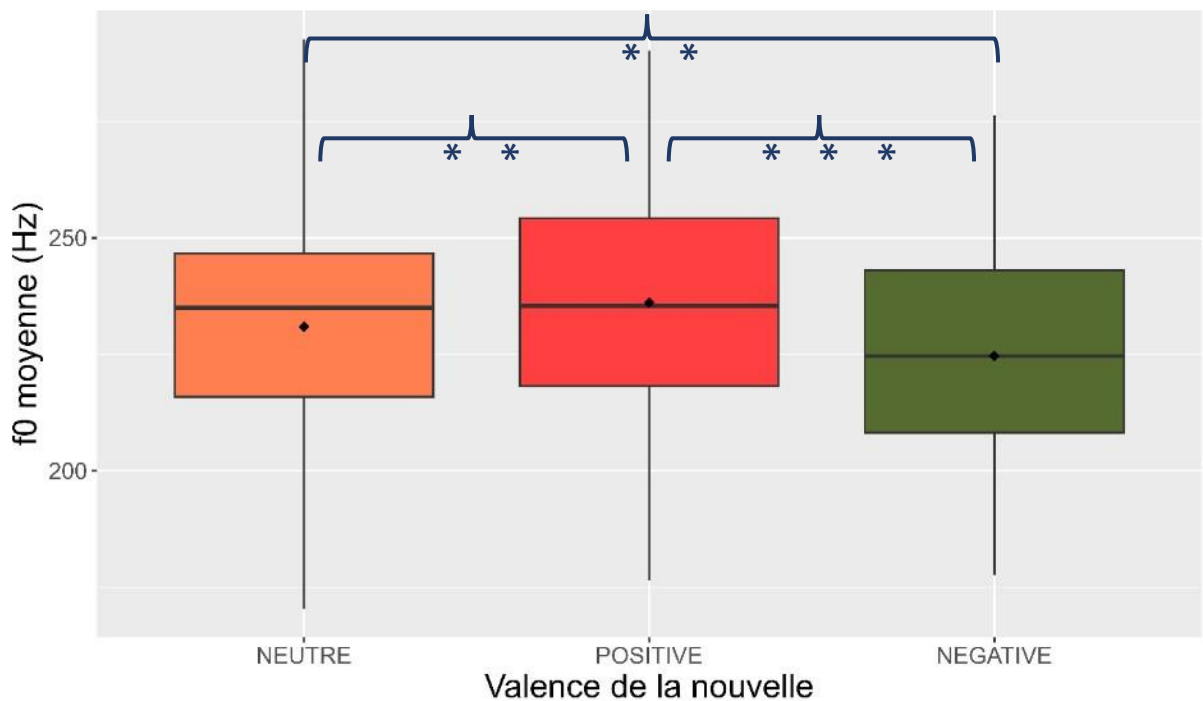


Figure 16 – Moyenne de f0 (Hz) dans les préfaces par valence de la nouvelle

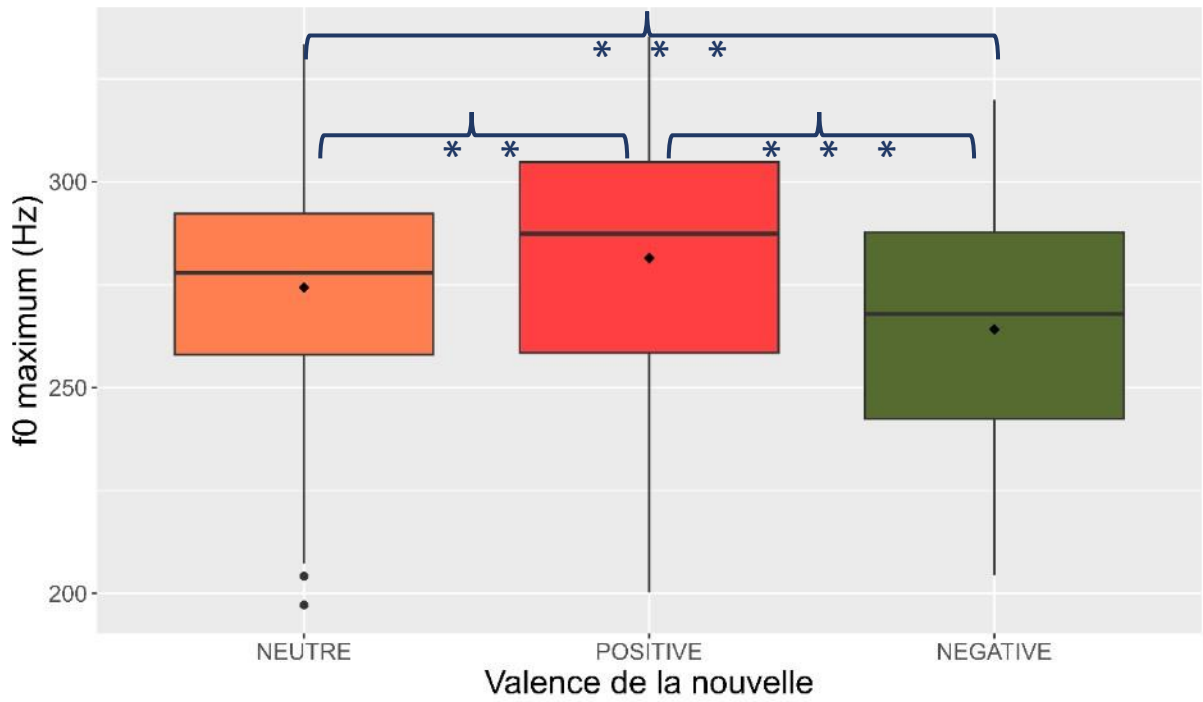


Figure 17 – f0 maximum (Hz) dans les préfaces par valence de la nouvelle

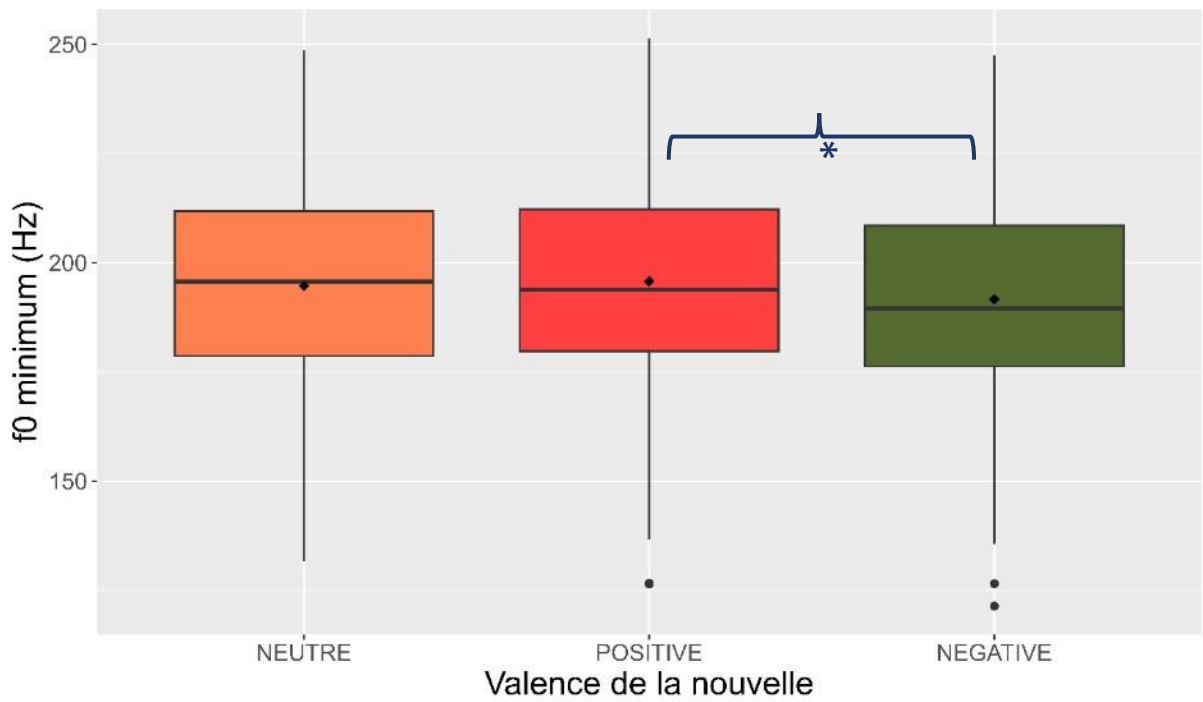


Figure 18 – f0 minimum (Hz) dans les préfaces par valence de la nouvelle

3.1.2. Au sein des marqueurs discursifs « voilà donc »

Nous avons trouvé des différences significatives de hauteur de la f0 et largeur d'empan dans les mesures globales de la f0 prises sur les marqueurs discursifs, bien que ces différences ressortent moins que dans la préface. Elles apparaissent comme très marquées entre les valences positive et négative, parfois marquées entre les valences négative et neutre, et n'apparaissent presque jamais entre les valences positive et neutre. Le Tableau 4 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

Valence	Empan de la f0		f0 moyenne		f0 maximum		f0 minimum	
	Moy.	SD	Moy.	SD	Moy.	SD	Moy.	SD
Négative	51,8	30,5	203,4	24,4	229,2	32,1	184,4	22,1
Positive	66,9	37,2	215,2	24,1	249,4	33,6	187,3	23,3
Neutre	61,2	35,1	210,6	25,0	241,5	31,2	187,4	24,0

Tableau 4 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans les marqueurs discursifs

Nous voyons que l'empan vocal est plus étroit (Figure 19) et les valeurs moyennes de f0 abaissées (Figure 20) dans les marqueurs discursifs lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive : pour l'empan [$\beta = -0.27$, SE = 0.04, $t = -6.22$, $p < 0.0001$] et la f0 moyenne [$\beta = -0.05$, SE = 0.01, $t = -5.73$, $p < 0.0001$]. L'empan est également plus étroit et la f0 moyenne abaissée lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre : pour l'empan [$\beta = -0.18$, SE = 0.04, $t = -4.21$, $p < 0.0001$], et pour la f0 moyenne [$\beta = -0.03$, SE = 0.01, $t = -3.45$, $p < 0.01$]. Aucune différence n'apparaît entre les valences positive et neutre, ni pour l'empan [$p=0.04$], ni pour la f0 moyenne [$p=0.02$].

Pour la valeur maximum de la f0 (Figure 21), des différences semblent apparaître entre chaque comparaison de valence. Nous trouvons qu'elle est significativement plus basse lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive [$\beta = -0.08$, SE = 0.01, $t = -7$, $p < 0.0001$]. Elle l'est également lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre [$\beta = -0.05$, SE = 0.08, $t = -4.43$, $p < 0.0001$]. De même, elle est significativement plus haute lorsque la valence est positive par rapport à lorsqu'elle est neutre [$\beta = 0.03$, SE = 0.01, $t = 2.57$, $p < 0.025$].

Pour la valeur minimum de la f0 (Annexe 5), aucune différence ne ressort et nous ne trouvons en effet aucune différence significative de la valence de la nouvelle, quelles que soient les valences comparées [$p > 0.025$].

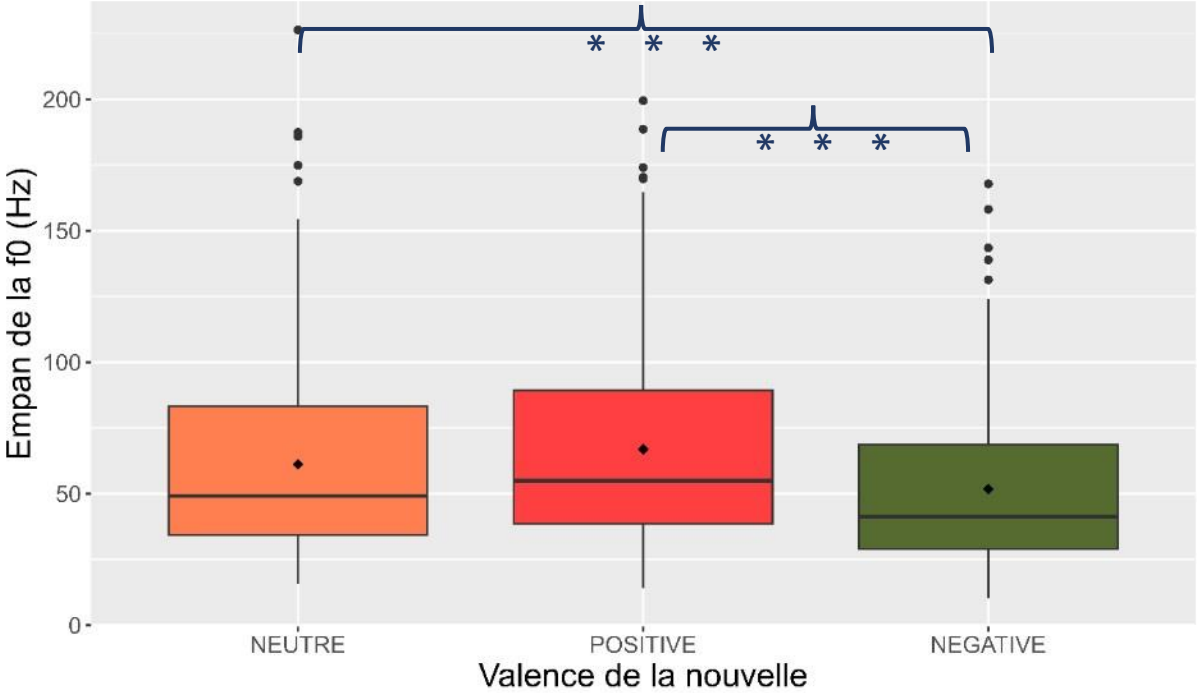


Figure 19 – Empan de la f0 (Hz) dans les marqueurs discursifs par valence de la nouvelle

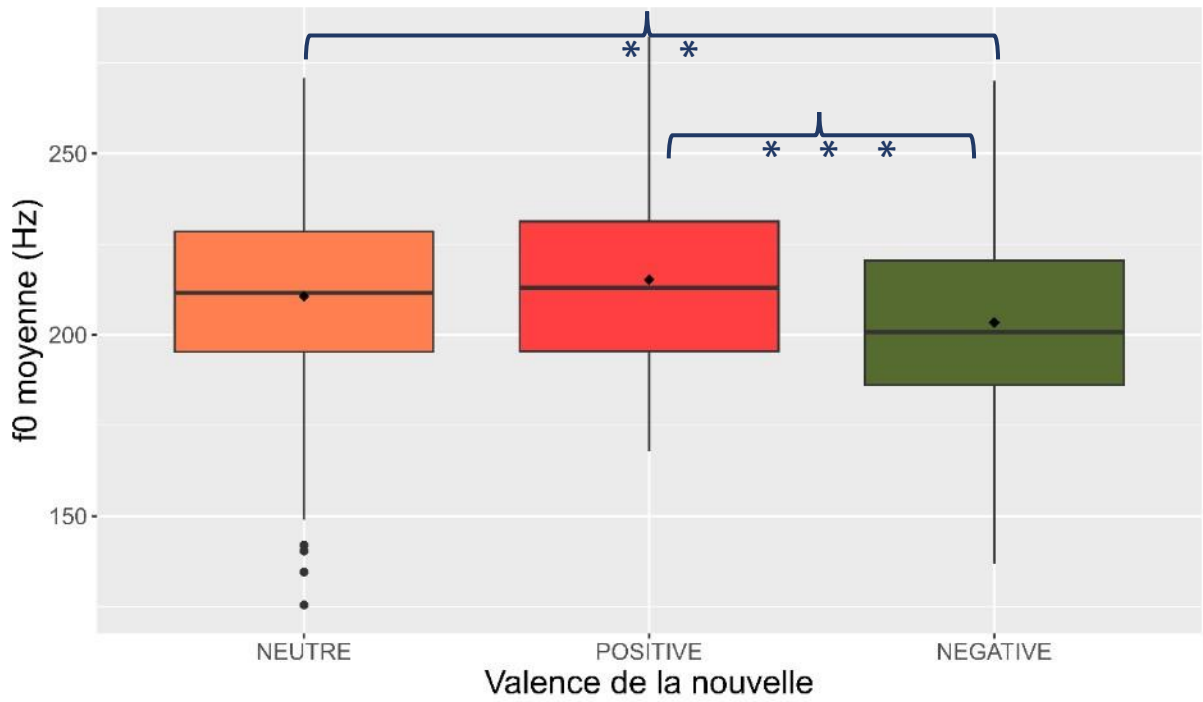


Figure 20 – Moyenne de f0 (Hz) dans les marqueurs discursifs par valence de la nouvelle

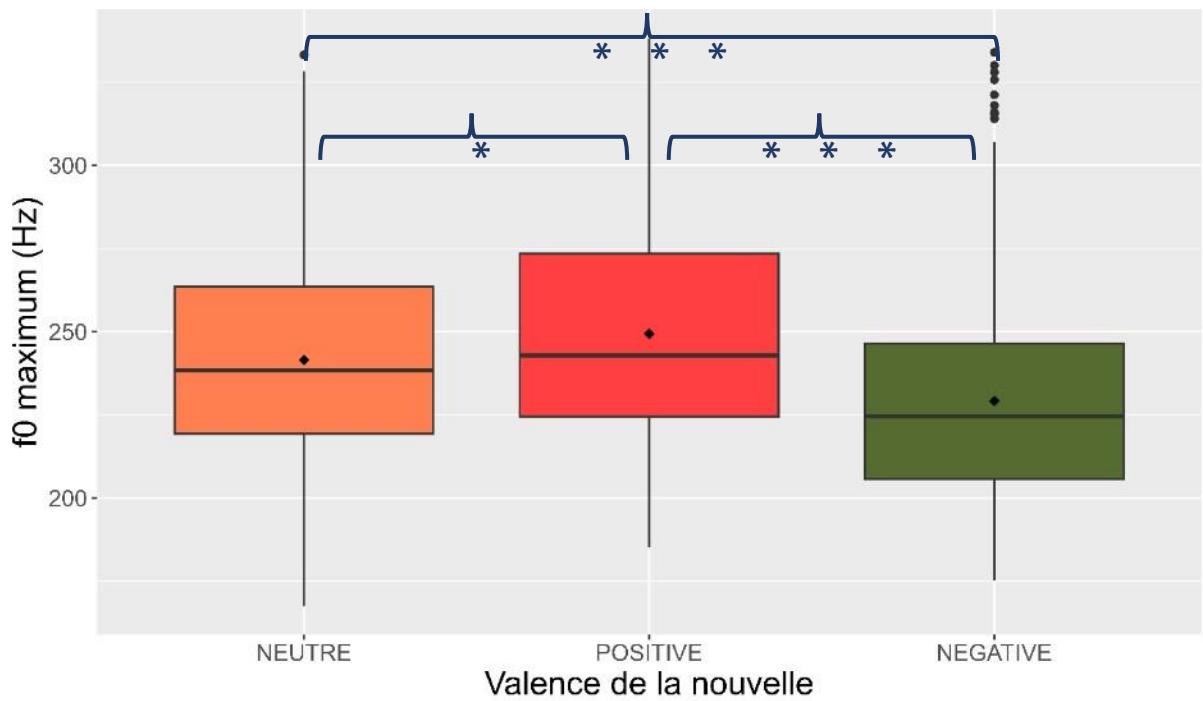


Figure 21 – f0 maximum (Hz) dans les marqueurs discursifs par valence de la nouvelle

3.1.3. Au sein de la séquence d'annonce de la nouvelle

Nous avons trouvé des différences significatives de hauteur de la f0 et largeur d'empan dans les mesures globales de la f0 prises sur la séquence d'annonce, très similaires à celles trouvées dans la préface. Ces différences apparaissent en effet comme très marquées entre les valences positive et négative, marquées entre les valences négative et neutre, et peu marquées entre les valences positive et neutre. Le neutre semble généralement situé à mi-chemin entre les valences positive et neutre.

Le Tableau 5 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

Valence	Empan de la f0		f0 moyenne		f0 maximum		f0 minimum	
	Moy.	SD	Moy.	SD	Moy.	SD	Moy.	SD
Négative	73,5	23,5	224,6	22,7	264,2	28,6	191,6	22,5
Positive	86,5	26,6	236,1	23,4	281,5	29,2	195,8	23,3
Neutre	81,5	22,8	230,9	22,3	274,3	26,5	194,7	22,5

Tableau 5 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans les séquences d'annonce

Nous trouvons que l'empan vocal (Figure 22) paraît plus étroit et les valeurs de la f0 moyenne (Figure 23), maximum (Figure 24) et minimum (Figure 25) abaissées lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive : pour l'empan [$\beta = 0$, SE = 0, t = -25.30, p < 0.0001], la f0 moyenne [$\beta = -0.07$, SE = 0.007, t = -10.22, p < 0.0001], la f0 maximum [$\beta = -0.11$, SE = 0.008, t = -13.02, p < 0.0001], et la f0 minimum [$\beta = -0.02$, SE = 0.005, t = -3.38, p < 0.01]. L'empan est également plus étroit et les valeurs de la f0 abaissées lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre : pour l'empan [$\beta = 0$, SE = 0, t = -12.44, p < 0.0001], la f0 moyenne [$\beta = -0.04$, SE = 0.007, t = -6.15, p < 0.0001], la f0 maximum [$\beta = -0.05$, SE = 0.008, t = -6.76, p < 0.0001], et la f0 minimum [$\beta = -0.01$, SE = 0.005, t = -2.69, p < 0.01]. De même, les valeurs sont significativement plus larges (empan) et plus hautes lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre : pour l'empan [$\beta = 0.167$, SE = 0,013, t = 12.77, p < 0.0001], la f0 moyenne [$\beta = 0.02$, SE = 0.007, t = 4.06, p < 0.001], et la f0

maximum [$\beta = 0.05$, $SE = 0.008$, $t = 6.25$, $p < 0.0001$]. Cette dernière différence ne ressort pas (Figure 25) pour la f_0 minimum [$p = 0.4$].

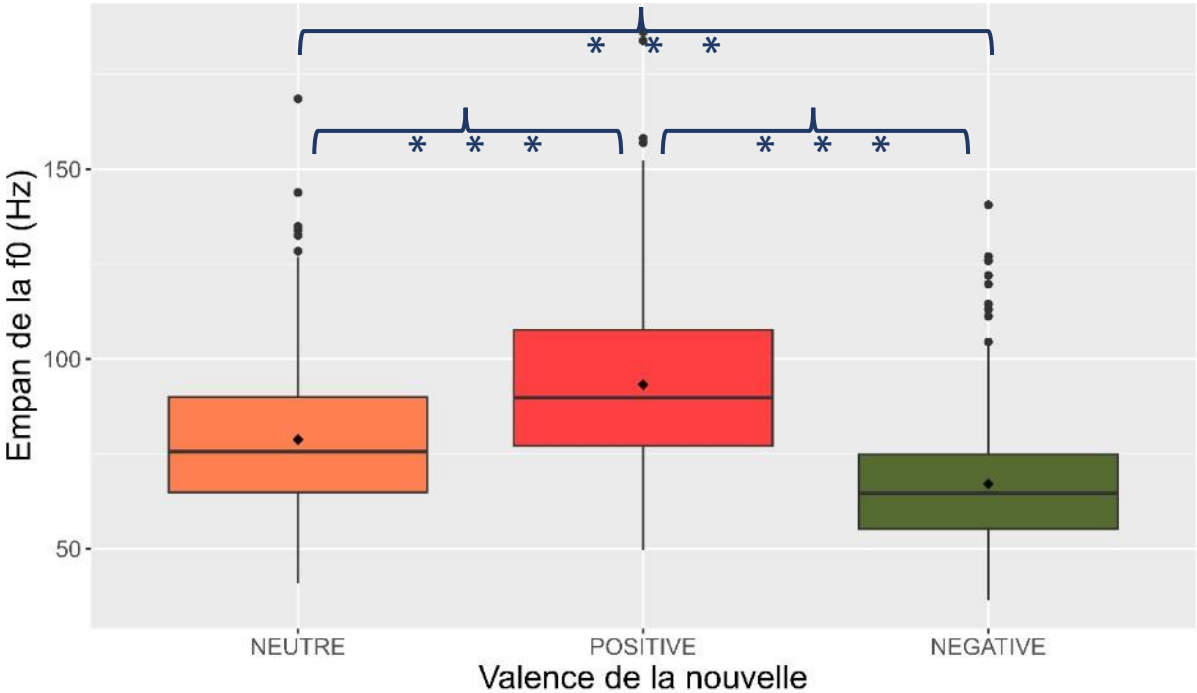


Figure 22 – Empan de la f_0 (Hz) dans les annonces par valence de la nouvelle

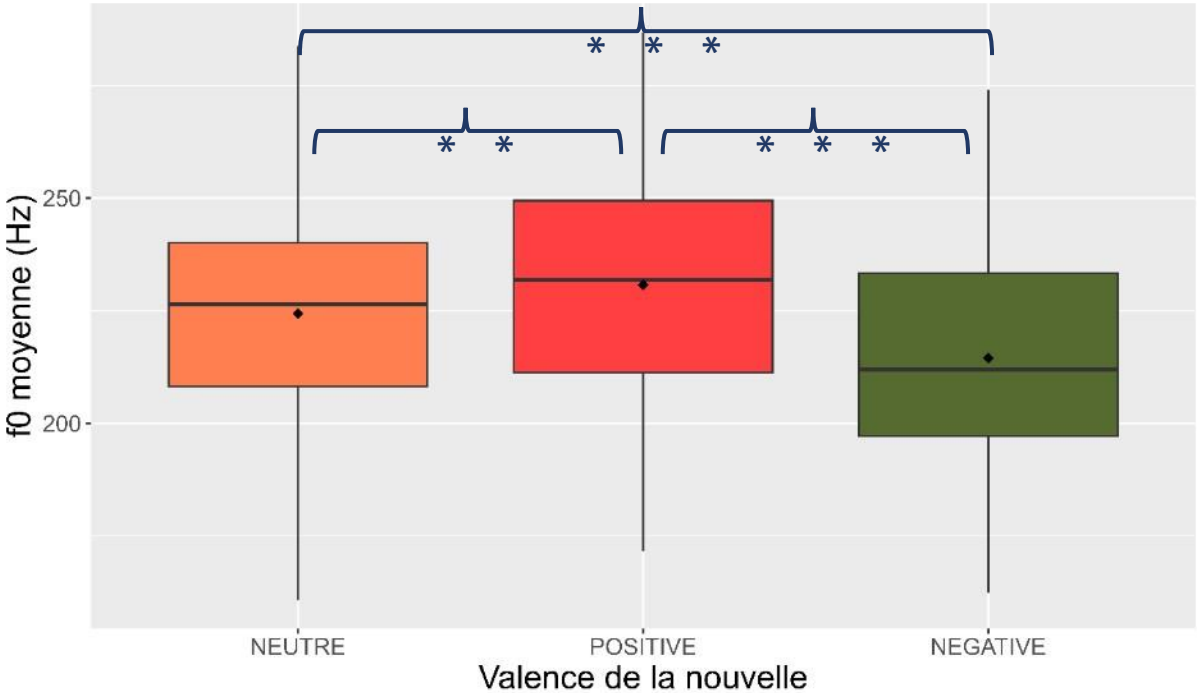


Figure 23 – Moyenne de f_0 (Hz) dans les annonces par valence de la nouvelle

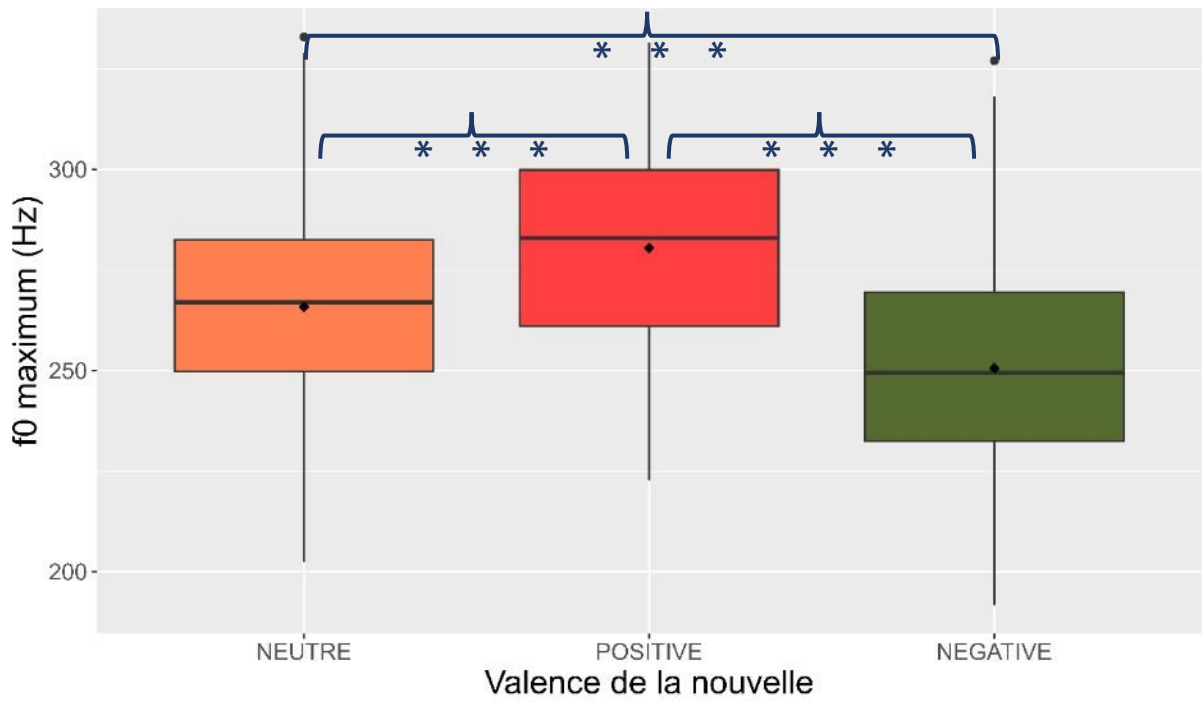


Figure 24 – f0 maximum (Hz) dans les annonces par valence de la nouvelle

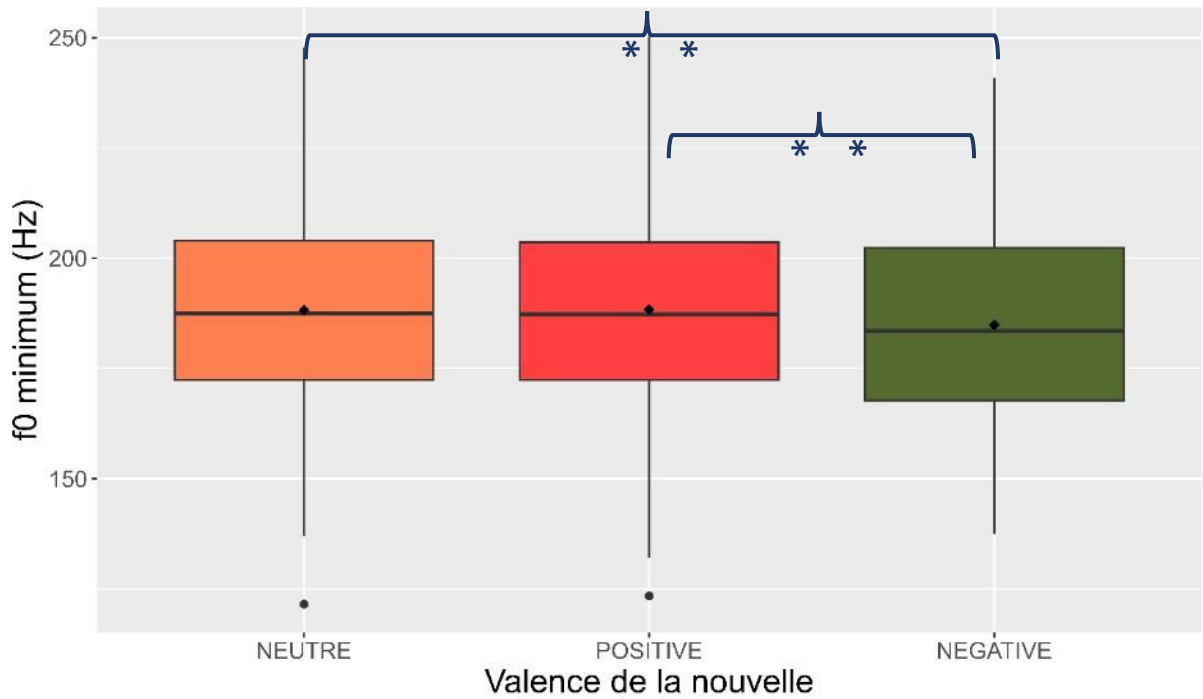


Figure 25 – f0 minimum (Hz) dans les annonces par valence de la nouvelle

3.1. Prise en compte du temps pour la f0

En plus des analyses globales de la f0, nous avons souhaité faire une observation de l'évolution dans le temps de la courbe de f0. Cela nous permet de voir si les différences significatives que l'on peut trouver, et que l'on a trouvées, sur les valeurs globales de la f0, se retrouvent sur l'ensemble de la portion observée, ou si elles sont localisées à certains points-clés de la préface et des marqueurs discursifs, par exemple sur la fin des syntagmes intonationnels ou accentuels, où l'on retrouve les accents mélodiques et les tons de frontières, ou bien en début d'énoncé.

3.1.1. Au début de la préface

Pour la courbe de f0 dans les préfaces, le Tableau 6 présente les coefficients pour les prédicteurs paramétriques ainsi que ceux pour les termes lissés. Nous allons commencer par observer les effets de la valence pour les termes lissés (quand le temps est donc pris en compte), avant de s'attarder sur les termes paramétriques. Les termes lissés, qui prennent en compte l'effet du temps sur chaque valence, ainsi que la pente aléatoire valence par locuteur, présentent des différences significatives [$p < 0,0001$], et des degrés de liberté (edf) largement supérieurs à 1, ce qui indique que les prédicteurs sont bien non-linéaires (Arnhold & Kyröläinen, 2017).

Coefficients paramétriques	β	SE	t	p
Intercept	205.51	3.92	52.41	<0.0001
Valence neutre	6.26	5.54	1.12	0.258
Valence positif	13.31	5.54	2.40	<0.05
Termes lissés	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.86	29	1130.6	<0.0001
s(Temps):Valence neutre	28.88	29	1227.9	<0.0001
s(Temps):Valence positif	28.89	29	1309.5	<0.0001
s(Participant, Valence)	86.14	87	100.5	<0.0001

Tableau 6 – Coefficients paramétriques et termes lissés du modèle GAMM réalisé sur la courbe de f0 au sein des préfaces

Nous pouvons donc nous attendre à ce qu'une inspection visuelle des données grâce au graphique lissé et aux graphiques de différence (Figure 26, Figure 27) mette en lumière des zones dans l'axe du temps, sur lesquels on retrouve des différences localisées entre les valences. Nous trouvons une courbe de la f0 au sein de la préface significativement

plus basse lorsque la valence de la nouvelle (dans l'annonce) est négative par rapport à lorsqu'elle est positive, sur l'ensemble de la fenêtre temporelle choisie « je vous appelle au sujet » (/ʒəvuzzaapɛlosyʒɛ/, le temps normalisé correspondant à chaque zone de différences significatives est proposé en Annexe 8). La courbe de la f_0 dans la préface est significativement plus basse lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre, sur une partie seulement des segments : /ʒ/ (uniquement au point 3.7), /ap/, et /ɛ/. Enfin, la courbe de la f_0 est significativement plus basse lorsque la valence de la nouvelle est neutre par rapport à lorsqu'elle est positive, sur une partie seulement des segments : /p/ (au point 31.6), /o/. Les coefficients paramétriques présentés dans le Tableau 6 démontrent des différences significatives uniquement entre la valence de la nouvelle lorsqu'elle est négative et lorsqu'elle est positive [$p < 0,025$]. Cela signifie que les différences trouvées le long de la courbe avec les termes lissés se retrouvent aussi de façon globale, lorsque l'on compare les courbes en moyenne, sans prendre en compte le facteur temps. En revanche, cette différence globale ne se retrouve pas lorsque l'on compare les autres valences, alors que des différences ressortaient dans les mesures globales de la f_0 rapportées précédemment. Cela pourrait indiquer que les différences significatives sont situées après la portion de signal étudiée par les GAMM (« Je vous appelle au sujet *de vos examens médicaux* »). Il se pourrait également que les différences soient trop faibles au sein de chaque segment pour ressortir, alors que les mesures globales captureraient mieux ces différences.

négatif
 neutre
 positif

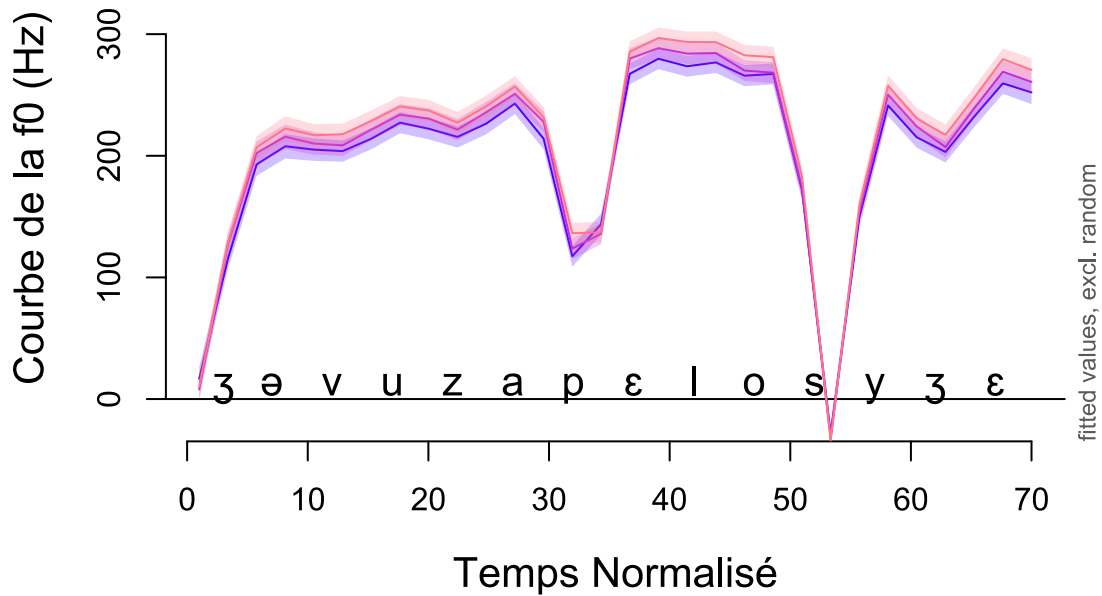


Figure 26 – Courbes lissées de la f0 au sein de la préface par valence

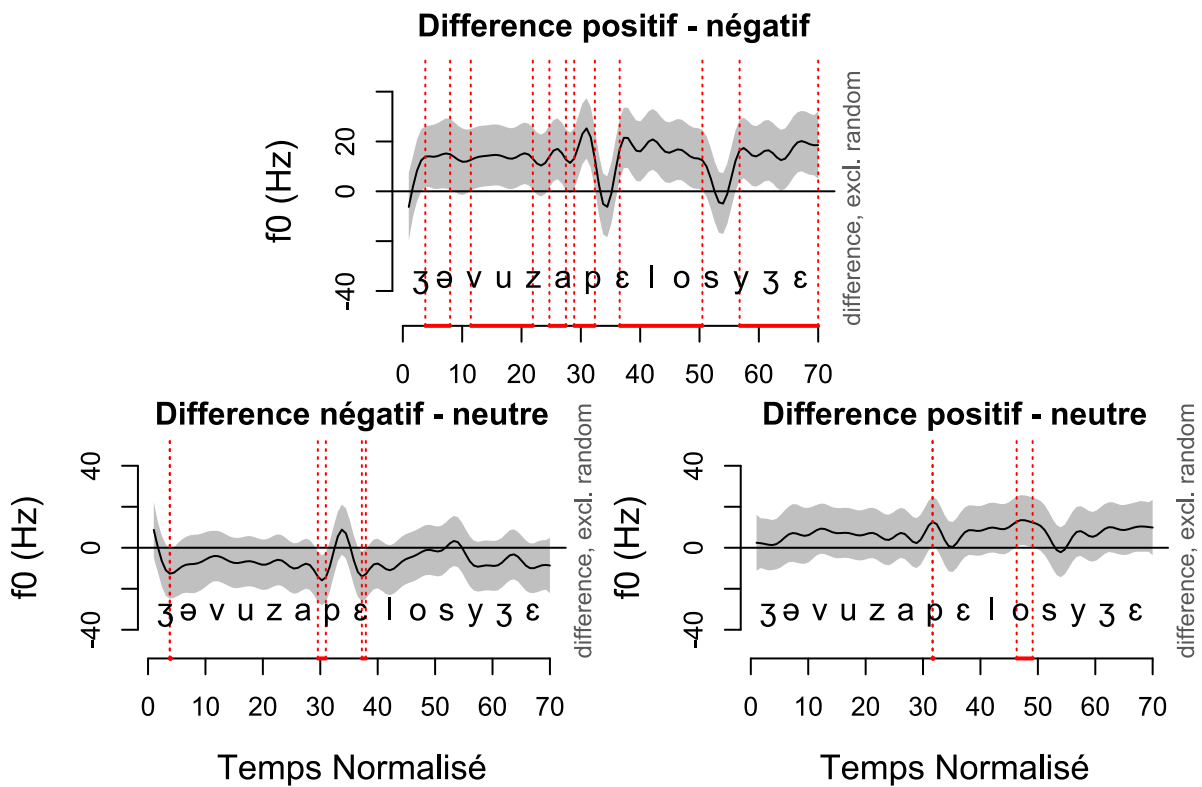


Figure 27 – Différences significatives de la f0 au sein de la préface

3.1.2. Au sein des marqueurs discursifs

Pour la courbe de f_0 dans les marqueurs discursifs, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Une inspection visuelle des données (Figure 28, Figure 29) devrait donc mettre en lumière des différences localisées entre les valences.

Nous trouvons une courbe de la f_0 significativement plus basse lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive, uniquement sur le mot « voilà ». La courbe de la f_0 est significativement plus basse lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre, sur une partie seulement des segments, au début de « voilà » et à la jonction entre « voilà » et « donc ». Aucune différence n'est trouvée entre le positif et le neutre ce qui est cohérent avec les analyses globales de la f_0 pour les marqueurs discursifs. En effet, la différence trouvée sur /k/ ne peut être prise en compte puisque c'est une consonne non voisée sur laquelle la f_0 n'est normalement pas produite. Les coefficients paramétriques présentés en Annexe 8, qui ne prennent pas en compte le facteur temps, ne démontrent aucune différence significative quelle que soit la valence [$p > 0,025$]. Cela signifie que les différences trouvées le long de la courbe avec les termes lissés ne se retrouvent pas lorsque l'on compare les marqueurs discursifs de façon globale, alors qu'ils étaient retrouvés sur les mesures globales de la f_0 rapportées précédemment, au moins entre le positif et le négatif et le positif et le neutre. Cela indique que les différences trouvées par nos mesures globales (empan vocal, f_0 moyenne, minimum et maximum) pourraient avoir capturé des différences qui sont localisées le long de la courbe de la f_0 plutôt qu'une tendance générale sur l'ensemble de cette courbe.

négatif
 neutre
 positif

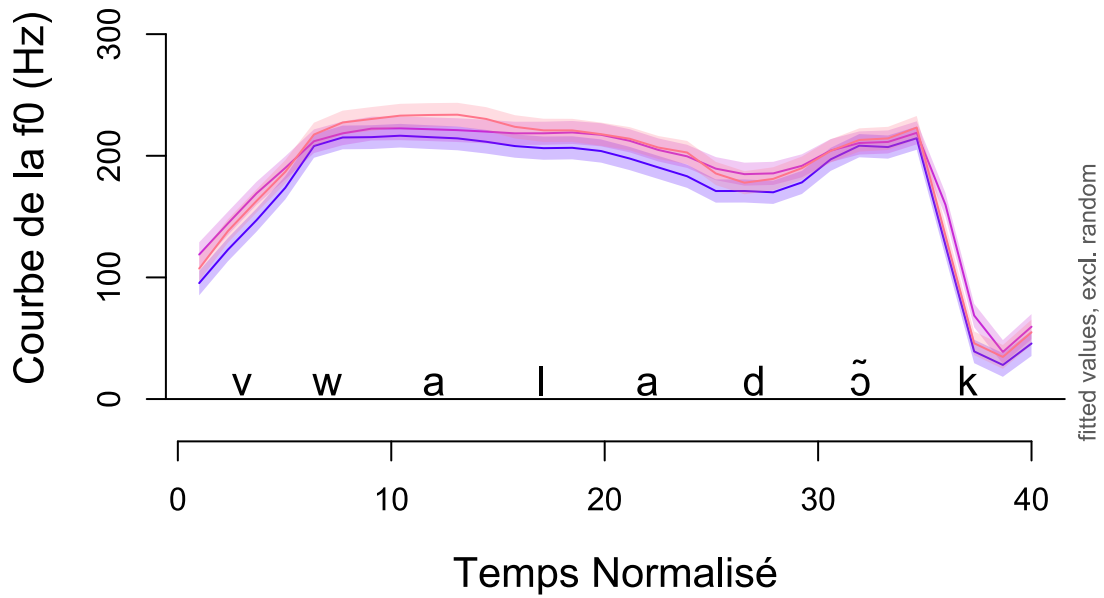


Figure 28 – Courbes lissées de la f0 au sein des DMs par valence

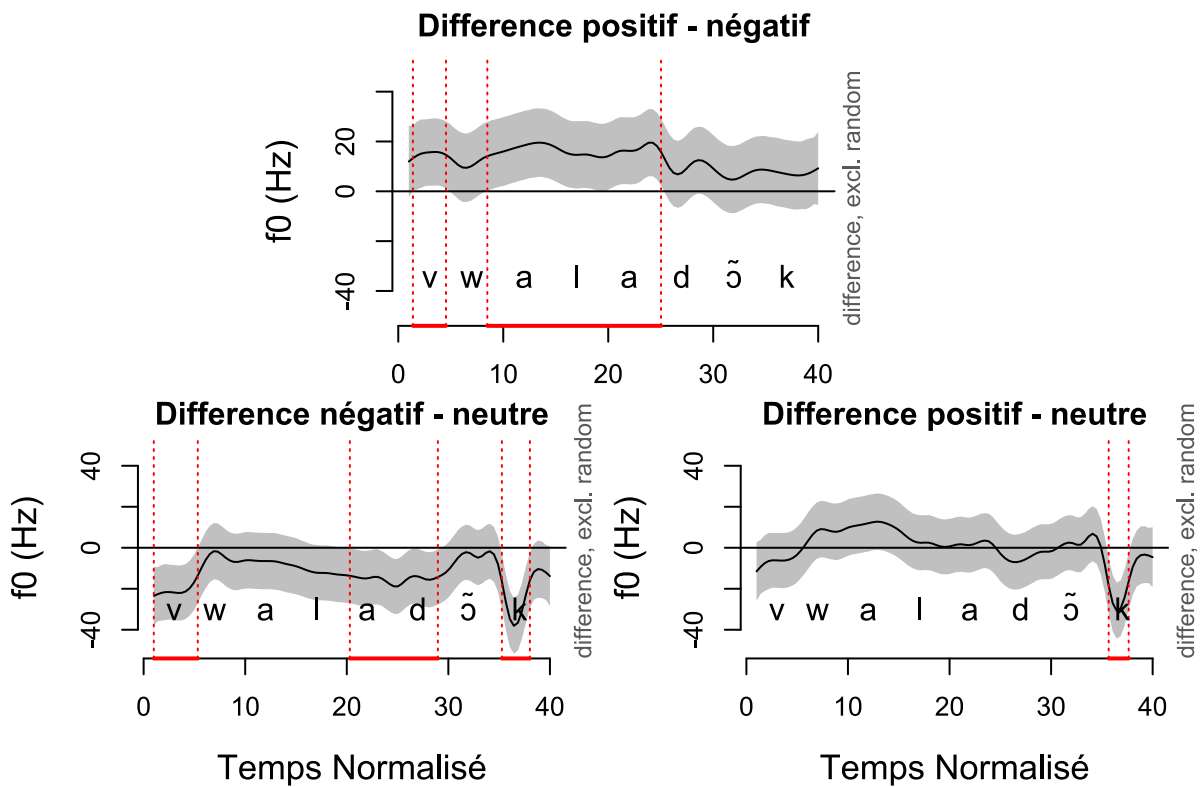


Figure 29 – Différences significatives de la f0 au sein des DMs

3.2. Mesures des données temporelles

3.2.1. Au sein de la préface

Nous trouvons des différences au niveau de la durée des préfaces (pauses incluses) et de la durée des pauses produites dans les préfaces, uniquement entre les valences positive et négative. Contrairement à nos attentes cependant, les mesures prises pour la vitesse d'articulation ne sont pas concluantes. Le Tableau 7 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

<i>Valence</i>	<i>Vitesse d'articulation (syll/sec)</i>		<i>Durée énoncés (sec)</i>		<i>Durée pauses (sec)</i>	
	<i>Moy.</i>	<i>SD</i>	<i>Moy.</i>	<i>SD</i>	<i>Moy.</i>	<i>SD</i>
<i>Négative</i>	6,027	0,804	2,996	0,784	0,434	0,304
<i>Positive</i>	6,085	0,742	2,899	0,723	0,379	0,277
<i>Neutre</i>	5,965	0,783	2,963	0,776	0,363	0,268

Tableau 7 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les préfaces

Les graphiques réalisés pour la durée de l'énoncé (Figure 30) et des pauses (Figure 31) montrent un allongement lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive : pour la durée d'énoncé [$\beta = 0.03$, $SE = 0.011$, $t = 2.67$, $p < 0.01$], et la durée des pauses [$\beta = 0.08$, $SE = 0.037$, $t = 2.34$, $p < 0.025$]. En revanche, aucune différence n'apparaît entre les valences négative et neutre ([$p = 0.4$] pour la durée d'énoncé, [$p=0.07$] pour la durée des pauses) ou positive et neutre ([$p = 0.3$] pour la durée d'énoncé, [$p=0.09$] pour la durée des pauses).

Les mesures prises pour la vitesse d'articulation (Annexe 5) n'indiquent aucune différence de vitesse entre les trois valences possibles de nouvelle pour la production des préfaces [$p > 0.025$].

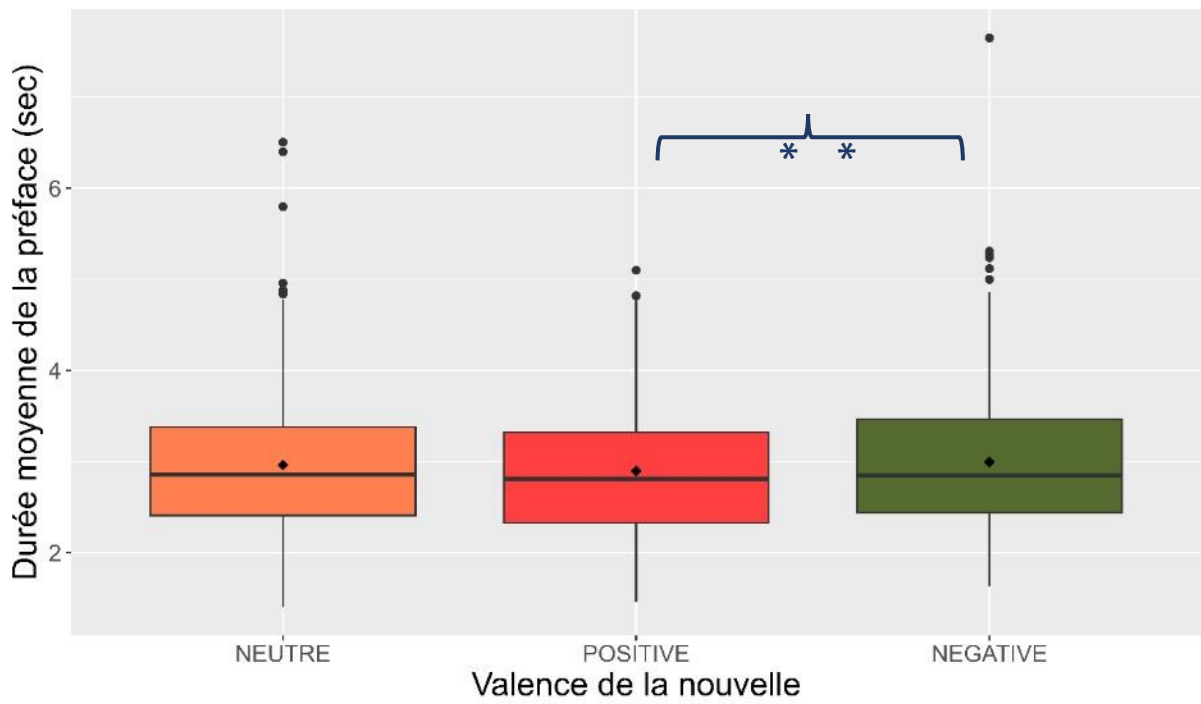


Figure 30 – Durée de la préface (sec) par valence de la nouvelle

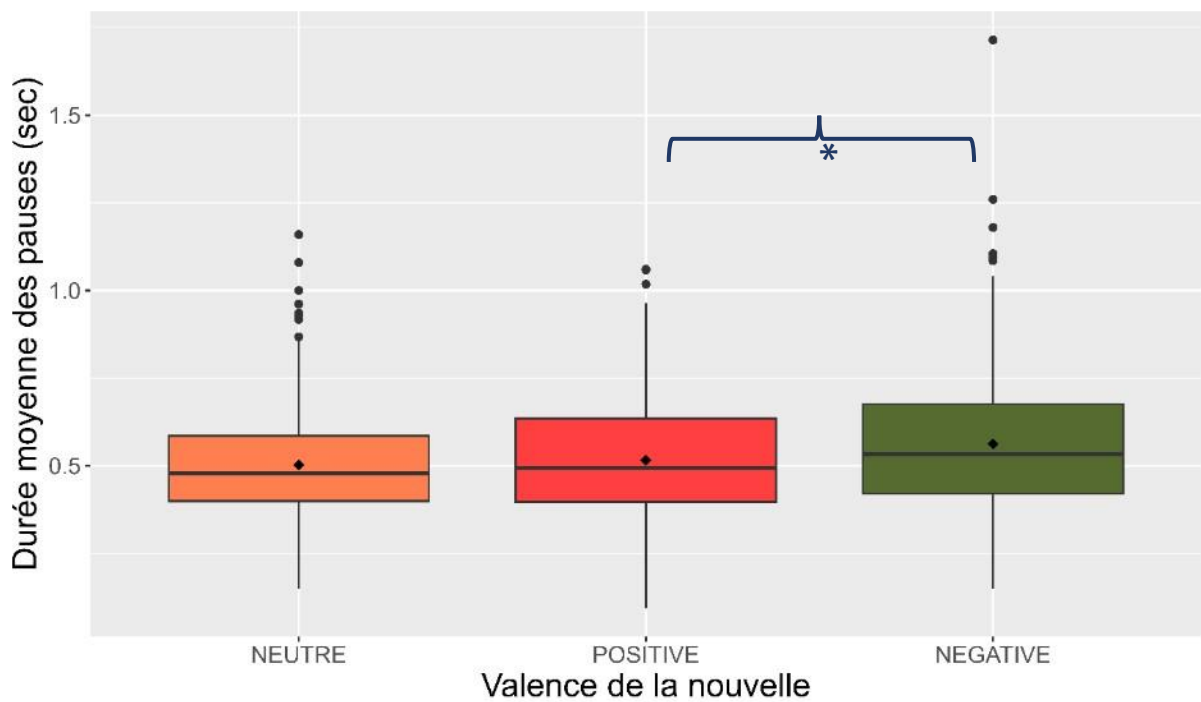


Figure 31 – Durée des pauses (sec) dans les préfaces par valence de la nouvelle

3.2.2. Au sein des marqueurs discursifs « voilà donc »

Nous trouvons des différences de longueur au niveau de la durée des marqueurs discursifs. Cependant de même que pour la préface, les mesures prises pour la vitesse d'articulation ne sont pas concluantes. Le Tableau 8 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

<i>Valence</i>	<i>Vitesse d'articulation (syll/sec)</i>		<i>Durée énoncés (sec)</i>		<i>Durée pauses (sec)</i>	
	<i>Moy.</i>	<i>SD</i>	<i>Moy.</i>	<i>SD</i>	<i>Moy.</i>	<i>SD</i>
<i>Négative</i>	6,409	1,569	0,586	0,164	0,010	0,082
<i>Positive</i>	6,354	1,548	0,579	0,140	0,009	0,061
<i>Neutre</i>	6,380	1,490	0,547	0,118	0,003	0,028

Tableau 8 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les marqueurs discursifs

Le graphique réalisé pour la durée de l'énoncé (Figure 32) montre, à l'inverse de ce que nous avons trouvé dans la préface, des différences significatives pour le négatif comparé au neutre avec un allongement de l'énoncé lorsque la valence de la nouvelle est négative [$\beta = 0.05$, $SE = 0.021$, $t = 2.76$, $p < 0.01$]. De même, une différence ressort entre le positif et le neutre, avec un allongement lorsque la valence de la nouvelle est positive [$\beta = 0.05$, $SE = 0.021$, $t = 2.44$, $p < 0.025$]. En revanche, aucune différence n'apparaît entre les valences négative et positive [$p = 0.7$].

Les mesures prises pour la vitesse d'articulation (Annexe 5) n'indiquent aucune différence de vitesse entre les trois valences possibles de nouvelle pour la production des marqueurs discursifs [$p > 0.025$].

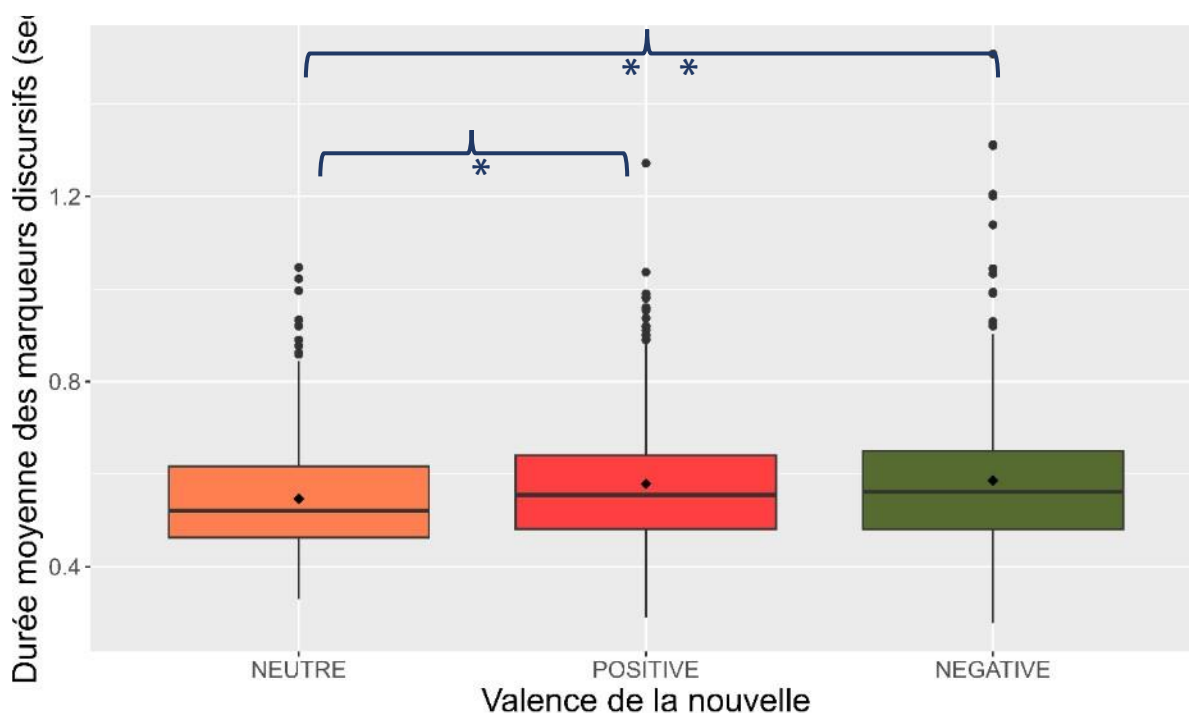


Figure 32 – Durée des marqueurs discursifs (sec) par valence de la nouvelle

3.2.3. Au sein des séquences d’annonces

Nous trouvons des différences au niveau de la durée des marqueurs discursifs et de la durée des pauses produites dans les séquences d’annonce, pour presque toutes les comparaisons de valences. De même que pour la préface et les marqueurs discursifs, les mesures prises pour la vitesse d’articulation ne sont pas concluantes. Le Tableau 9 présente les valeurs moyennes et les déviations-standard pour chacun des éléments observés.

Valence	Vitesse d’articulation (syll/sec)		Durée énoncés (sec)		Durée pauses (sec)	
	Moy.	SD	Moy.	SD	Moy.	SD
Négative	5,891	0,779	6,610	2,105	1,006	0,514
Positive	5,874	0,743	5,446	1,026	0,786	0,423
Neutre	5,758	0,754	4,935	1,550	0,805	0,448

Tableau 9 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les séquences d’annonce

Les graphiques réalisés pour la durée de l’énoncé (Figure 33) et des pauses (Figure 34) montrent un allongement lorsque la valence de la nouvelle est négative par rapport à

lorsqu'elle est positive : pour la durée d'énoncé [$\beta = 0.16$, $SE = 0.01$, $t = 11.44$, $p < 0.0001$], et la durée des pauses [$\beta = 0.22$, $SE = 0.056$, $t = 3.99$, $p < 0.0001$]. Des différences apparaissent également entre les valences négative et neutre, avec un allongement lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre : pour la durée d'énoncé [$\beta = 0.3$, $SE = 0.01$, $t = 20.84$, $p < 0.0001$], et la durée des pauses [$\beta = 0.22$, $SE = 0.056$, $t = 4.08$, $p < 0.0001$]. Les graphiques montrent que la différence entre les valences positive et neutre ne ressort que pour la durée de l'énoncé, avec un allongement de la durée de la séquence d'annonce lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre [$\beta = 0.13$, $SE = 0.01$, $t = 9.4$, $p < 0.0001$]. Cette différence ne ressort pas pour la durée des pauses [$p=0.9$].

Pour la vitesse d'articulation au sein des séquences d'annonces, nous ne trouvons aucun effet significatif de la valence de la nouvelle quelles que soient les valences comparées [$p > 0.025$].

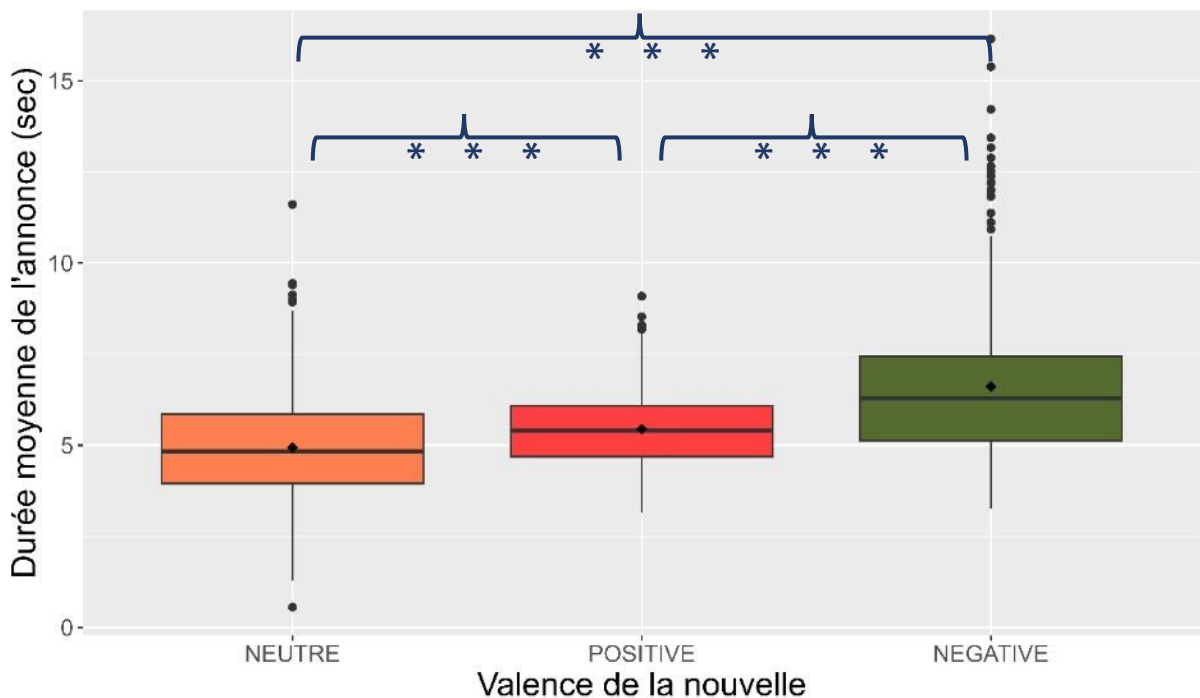


Figure 33 – Durée de la séquence d'annonce dans les annonces par valence de la nouvelle

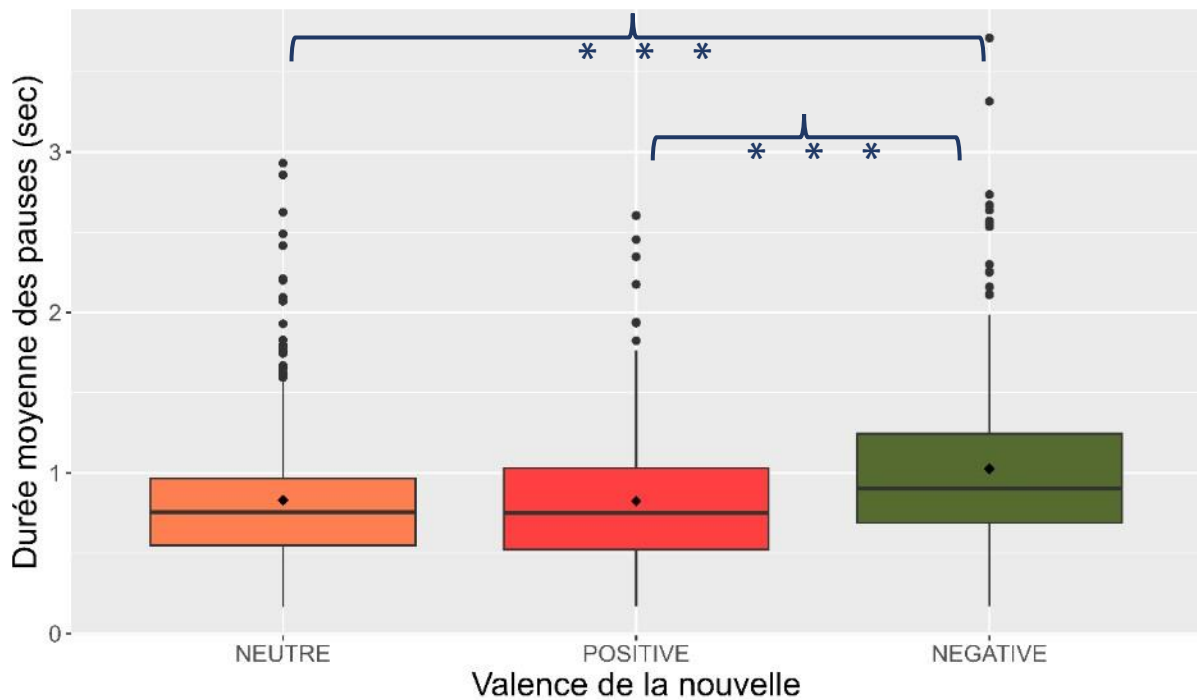


Figure 34 – Durée des pauses (sec) dans les annonces par valence de la nouvelle

3.3. Caractéristiques spectrales

3.3.1. Au sein de la préface

Pour les différentes mesures de la HNR dans les préfaces, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Nous procédons donc à une inspection visuelle des données (Figure 35 et Figure 36 pour la HNR05, Figure 37 et Figure 38 pour la HNR15, Figure 39 et Figure 40 pour la HNR25, Figure 41 et Figure 42 pour la HNR35).

Les résultats principaux montrent des différences significatives entre les valences positive et négative dans les 4 mesures de HNR. Elles se trouvent notamment en tout début de préface, sur « je ». Le ratio son bruit est plus bas (donc la voix plus soufflée ou craquée) lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive. Nous trouvons également des différences significatives avec un ratio plus bas donc une voix plus modale, lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre, au sein du mot « vous » (HNR15, HNR25, et HNR35).

Nous trouvons d'autres différences significatives, que nous interprétons différemment car de façon générale, elles indiquent que les segments seraient régulièrement dévoisés

plus que soufflés ou craqués dans les préfaces, ce facteur ne présentant pas de régularité en fonction de la valence. L'instabilité de détection de la HNR dans ces phonèmes nous a poussé à les écarter de l'analyse. Nous trouvons des différences significatives pour les quatre mesures de la HNR sein du mot « appelle » sur la voyelle /ε/. Le ratio son bruit y est plus bas lorsque la valence de la nouvelle est neutre par rapport à lorsqu'elle est positive ou négative, et lorsqu'elle est positive par rapport à lorsqu'elle est négative. Ces différences se situent sur une portion de HNR négative, et nous ne pouvons pas interpréter la voix comme soufflée, craquée ou modale dans ces conditions. Elle est simplement non voisée, car produite sur une consonne, et probablement dévoisée sur le /ε/ suivant. Il en est de même à la jonction de « vous appelle » (HNR25, HNR35) : le ratio indique un dévoisement du /z/ en /s/, qui n'interviendrait que dans les annonces positives. A l'inverse sur le mot « sujet », nous trouvons une différence significative sur la consonne voisée /ʒ/, avec une HNR négative indiquant un dévoisement du /ʒ/ en /ʃ/ dans la condition négative par rapport à la positive. Les coefficients paramétriques excluant la prise en compte du temps, présentés en Annexe 8, ne démontrent aucune différence significative quelle que soit la valence [$p > 0,025$]. Cela signifie que la prise en compte du temps est nécessaire pour observer les différences entre les valences, car plutôt que de se situer de façon globale sur l'ensemble des segments, elle serait localisée à certains points précis (un mot, une syllabe, un phonème).

négatif
 neutre
 positif

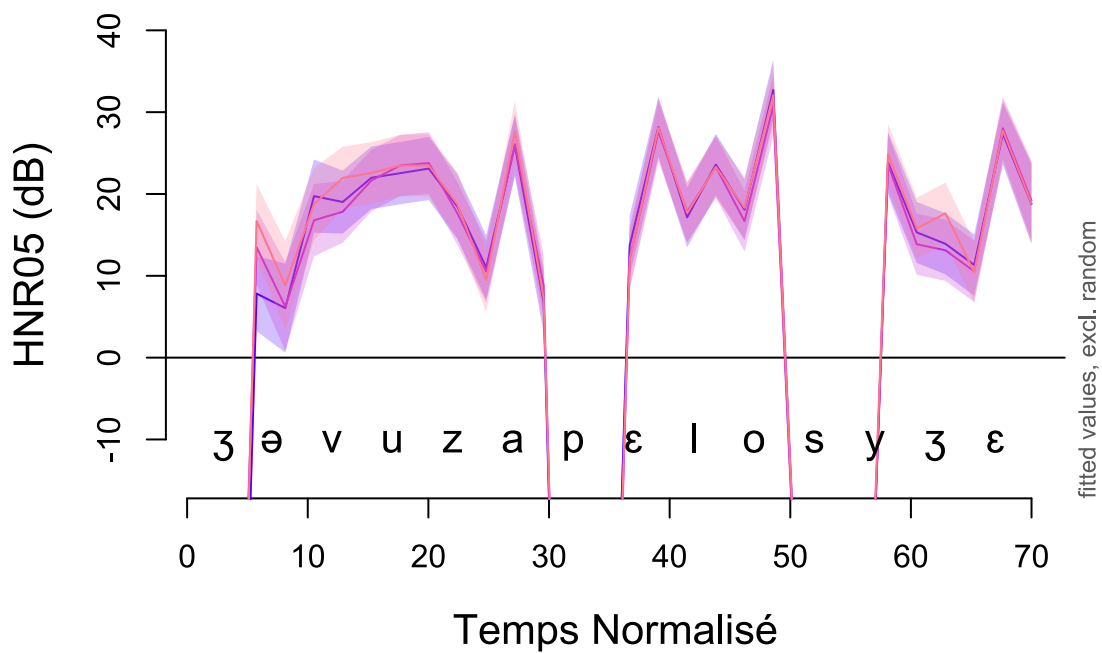


Figure 35 – HNR05 au sein de la préface par valence

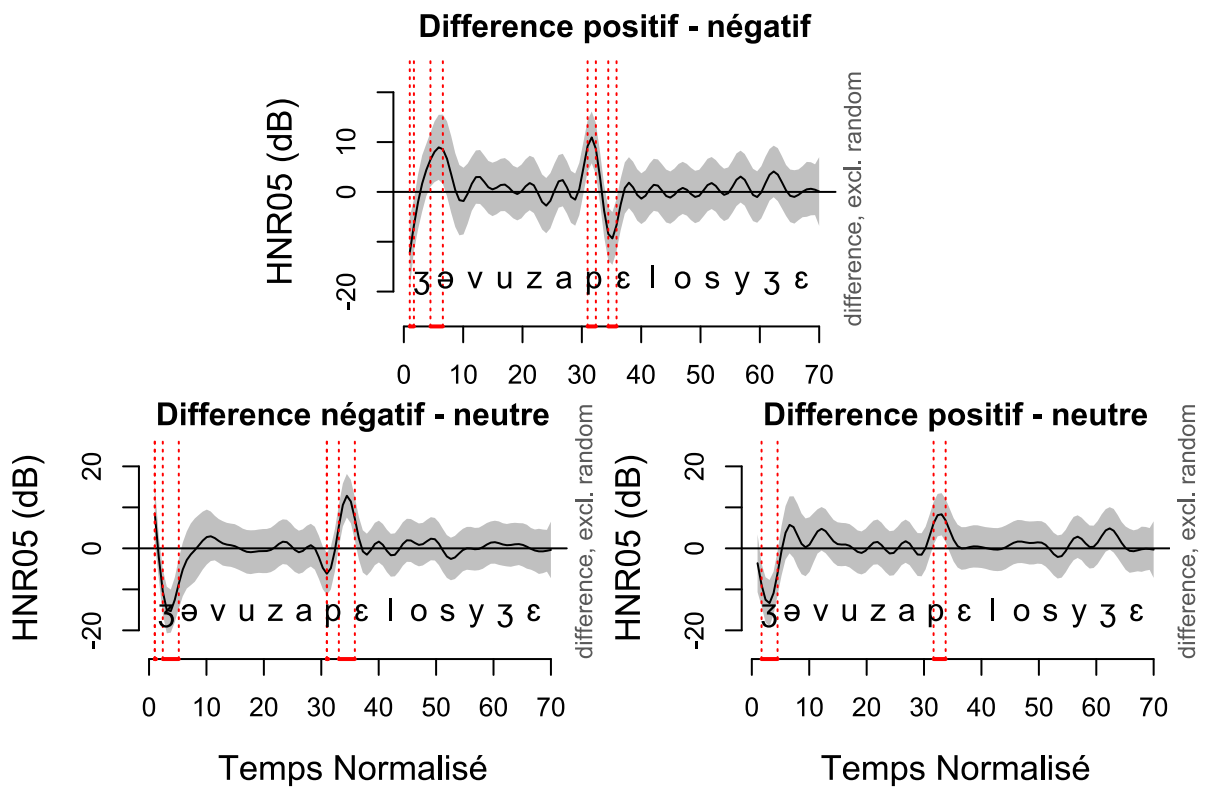


Figure 36 – Différences significatives de HNR05 au sein de la préface

négatif
 neutre
 positif

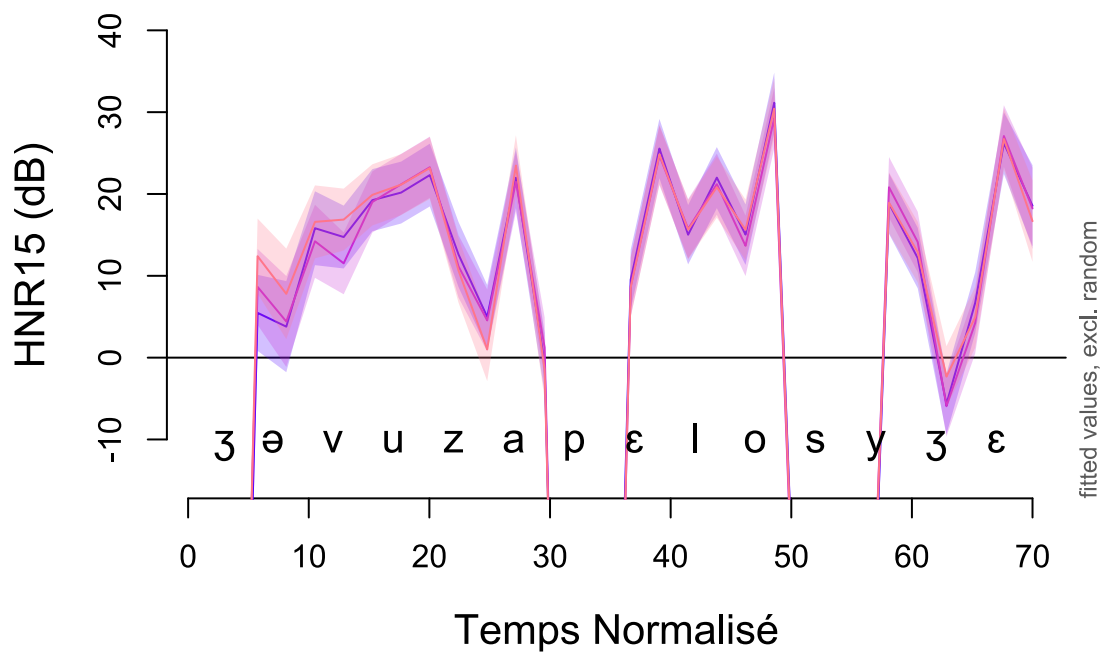


Figure 37 – HNR15 au sein de la préface par valence

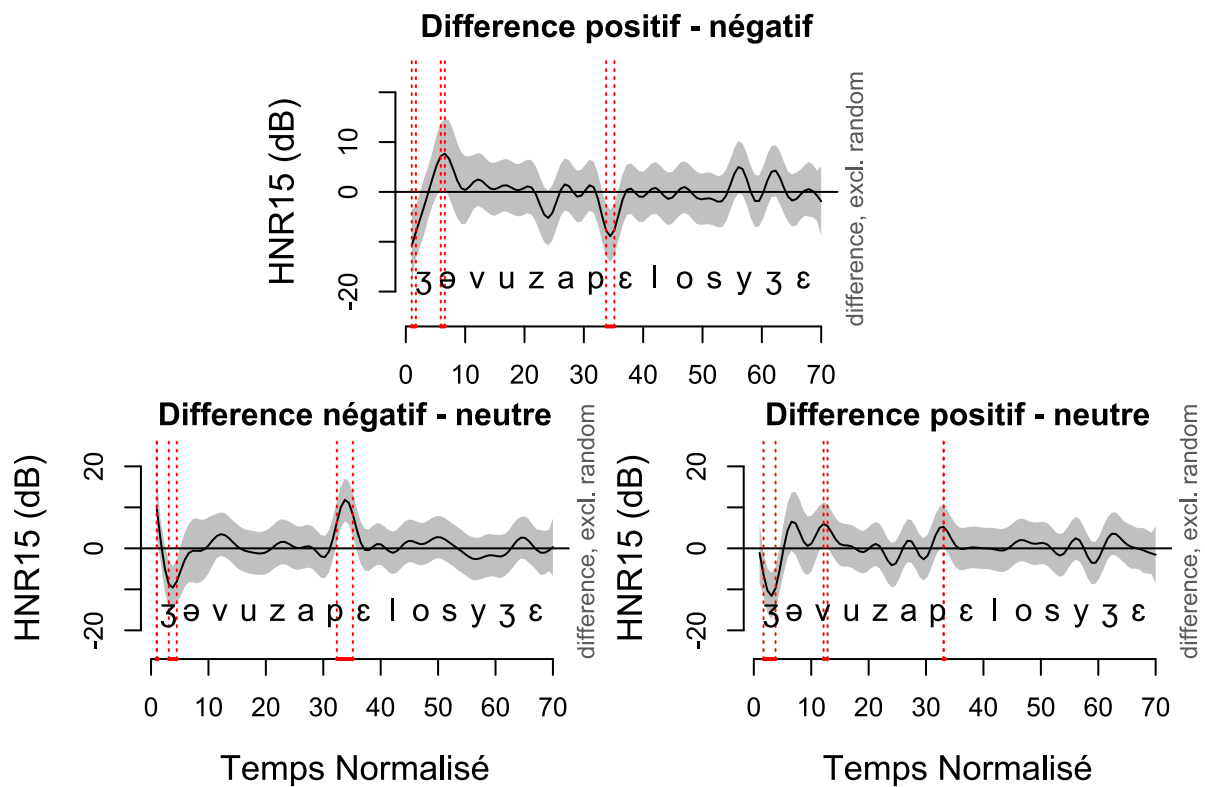


Figure 38 – Différences significatives de HNR15 au sein de la préface

négatif
neutre
positif

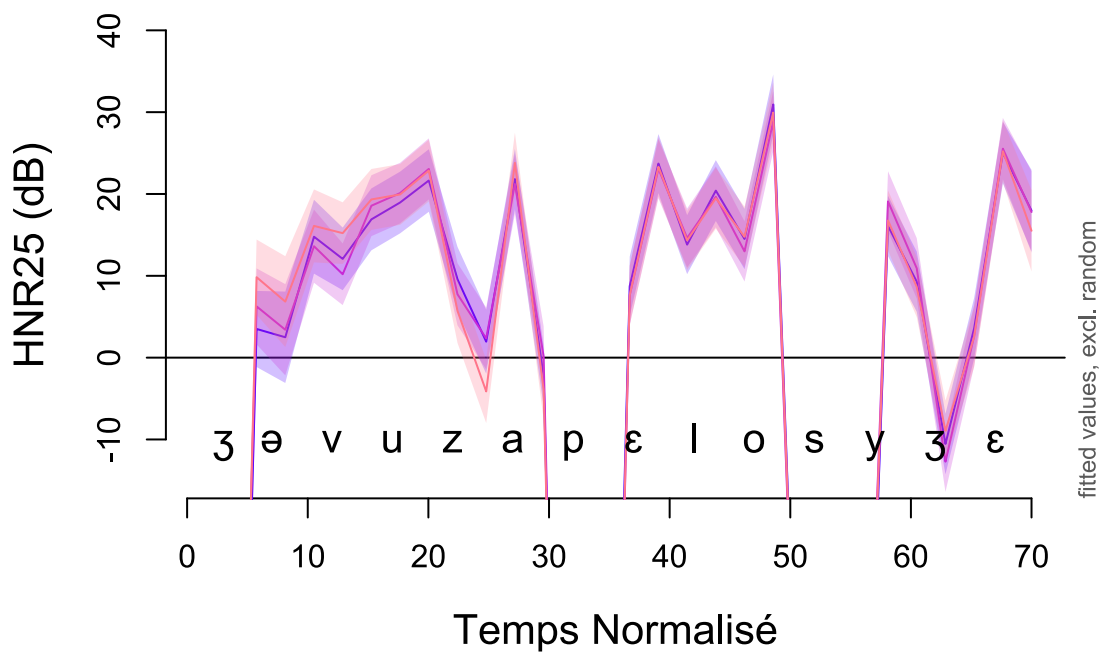


Figure 39 – HNR25 au sein de la préface par valence

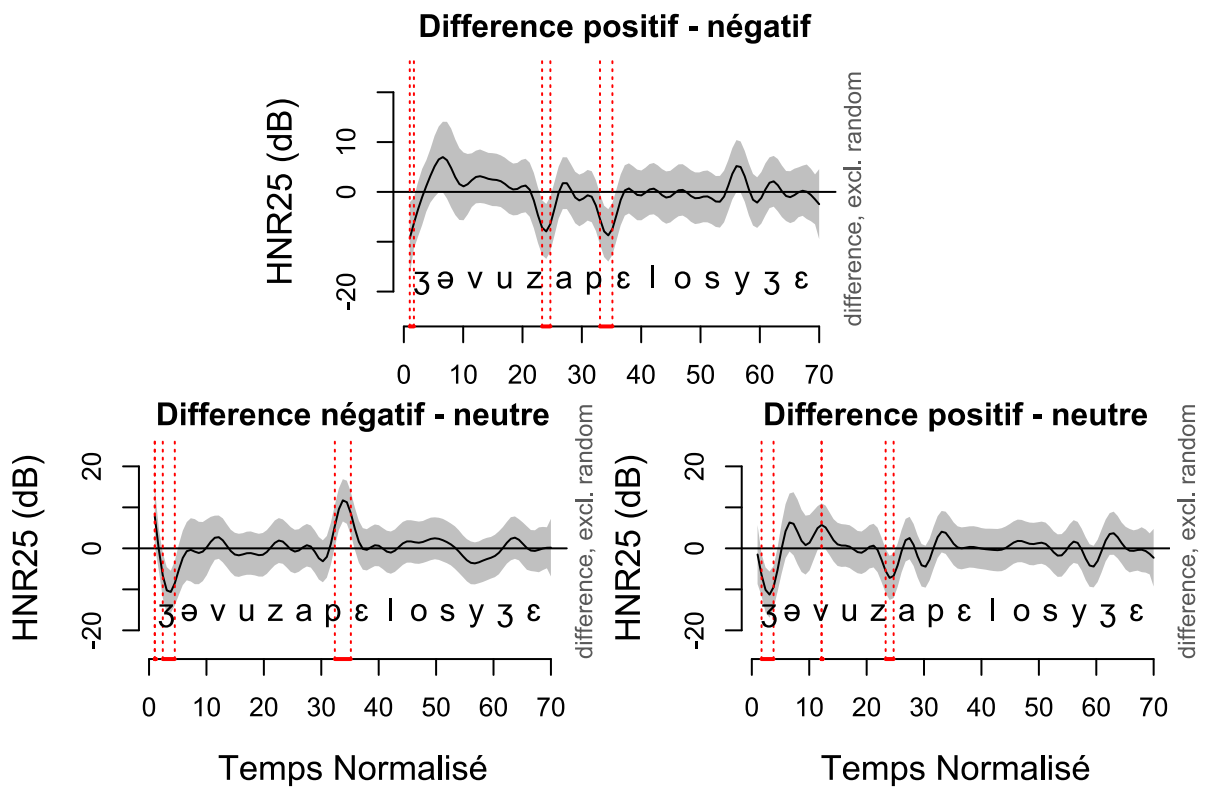


Figure 40 – Différences significatives de HNR25 au sein de la préface

négatif
 neutre
 positif

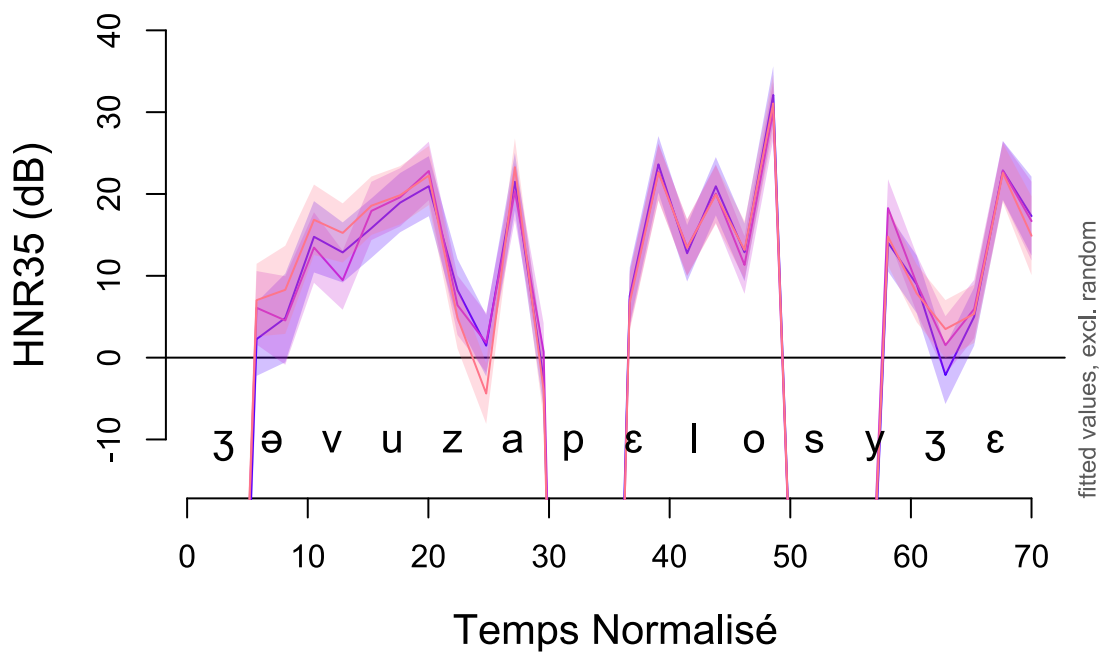


Figure 41 – HNR35 au sein de la préface par valence

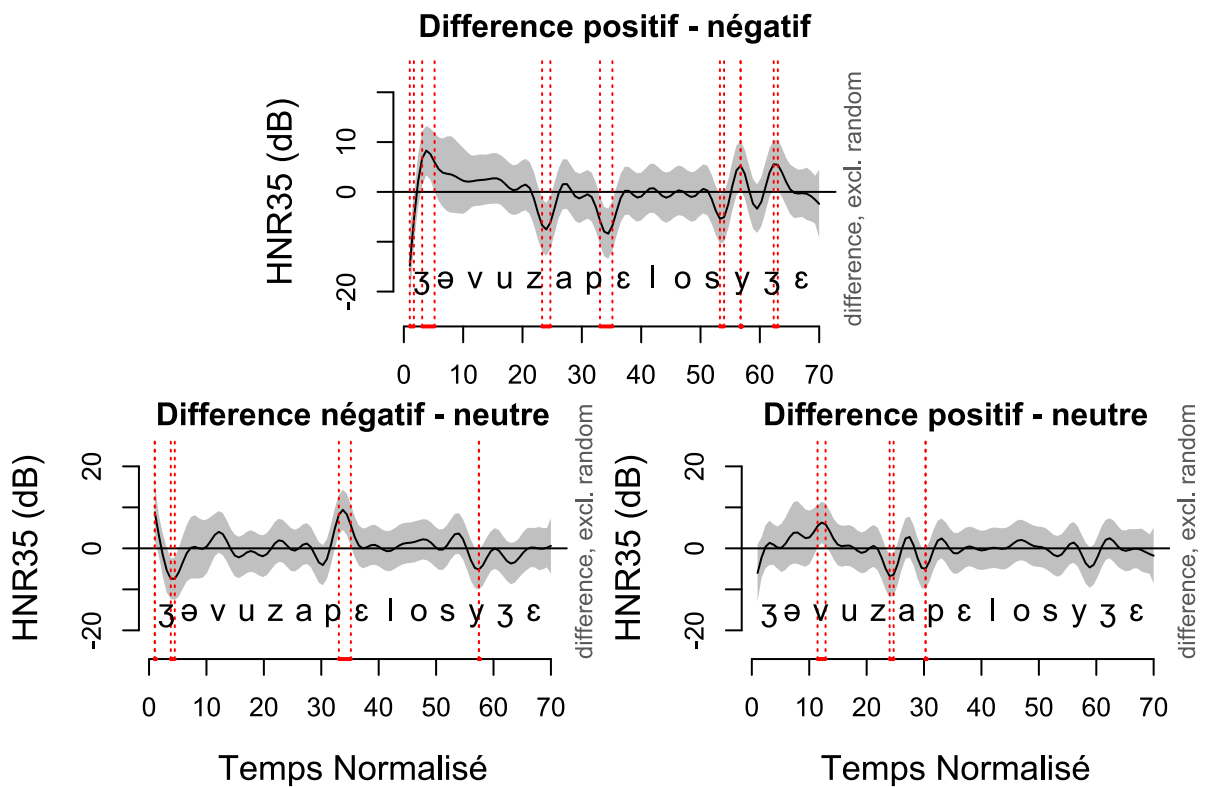


Figure 42 – Différences significatives de HNR35 au sein de la préface

Pour la différence entre les deux premières harmoniques (H1-H2) dans les préfaces, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Nous procédons donc à une inspection visuelle des données (Figure 43, Figure 44).

Nous trouvons une différence H1-H2 significativement plus haute, lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive, sur une grande portion de l'énoncé. Sur une portion similaire de l'énoncé, nous retrouvons également que H1-H2 est significativement plus haute lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre. La voix paraît donc plutôt soufflée lorsque nous annonçons une nouvelle négative. Enfin, H1-H2 est significativement plus basse lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre. Cela indique que la voix serait plus soufflée dans le neutre que lorsque la valence de la nouvelle est positive. Cette différence est peu représentée sur la courbe par rapport à ce que nous trouvons en comparant les autres valences. La différence entre H1 et H2 est rarement basse sur l'ensemble de la courbe, ce qui signifie que la voix n'est pas craquée mais plutôt soufflée ou modale. Les coefficients paramétriques présentés en Annexe 8 ne démontrent aucune différence significative quelle que soit la valence [$p > 0,025$]. La prise en compte du temps est donc nécessaire pour observer les différences entre les valences, qui pourront être localisées à certains points précis de l'énoncé.

négatif
 neutre
 positif

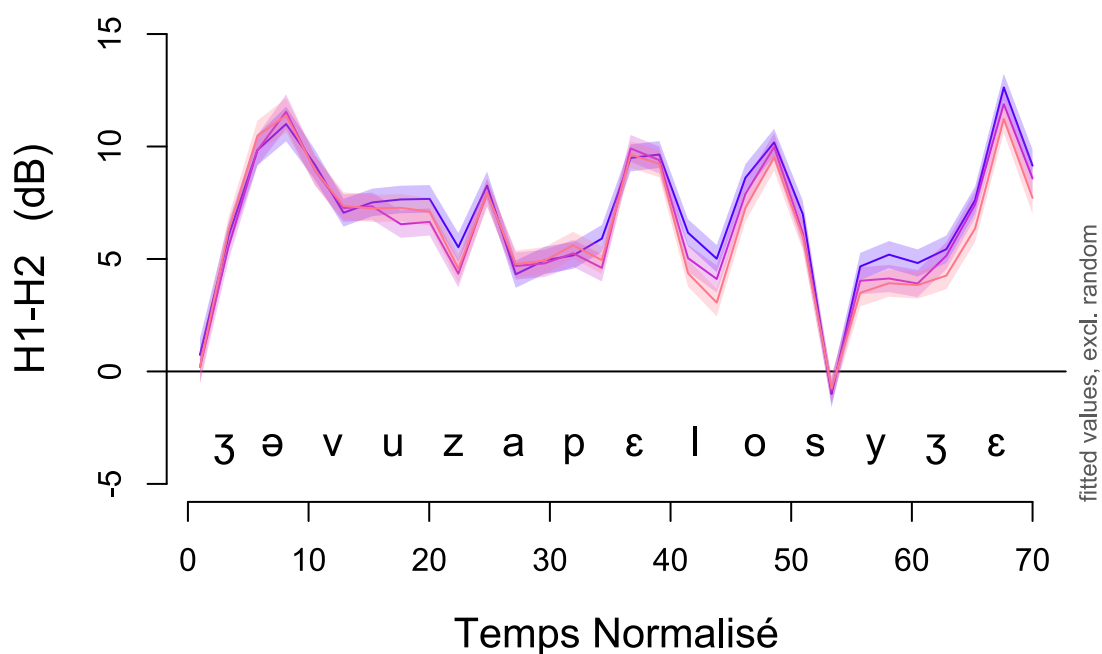


Figure 43 – H1-H2 au sein de la préface par valence

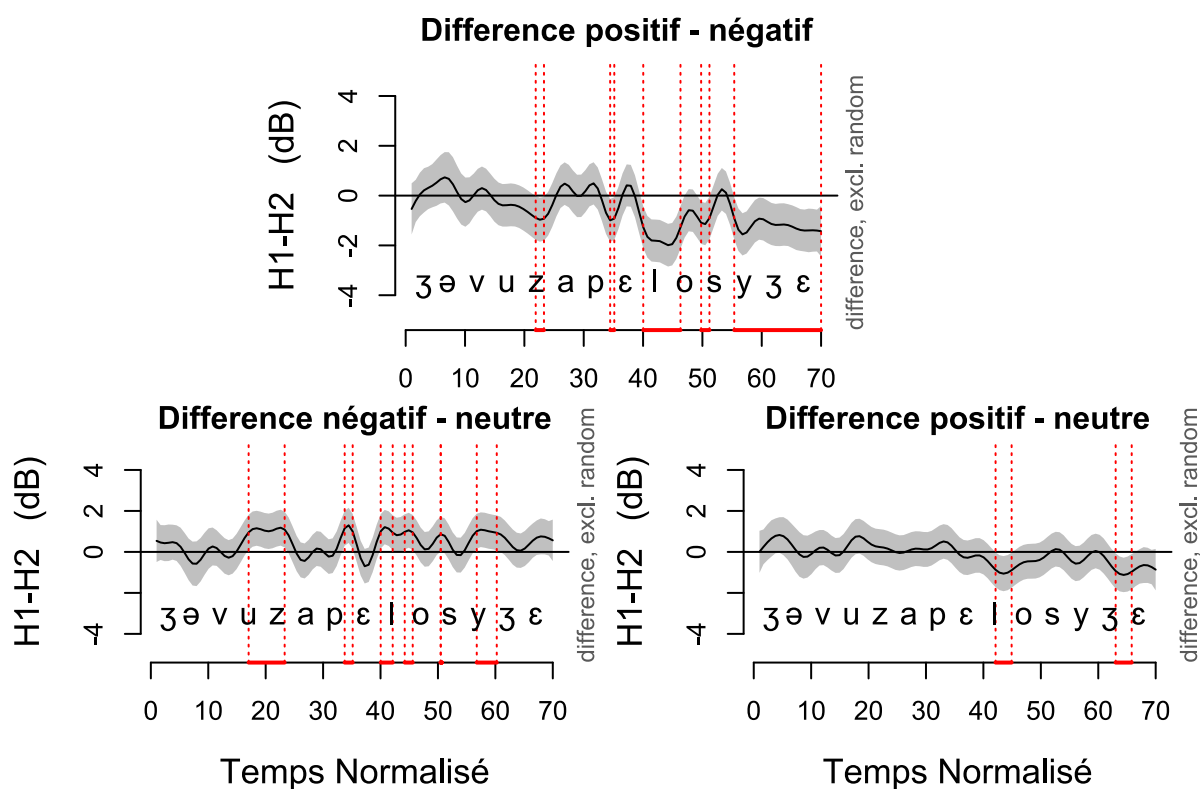


Figure 44 – Différences significatives de H1-H2 au sein de la préface

Pour la proéminence du pic cepstral lissé (CPPs) dans les préfaces, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Nous procédons donc à une inspection visuelle des données (Figure 45, Figure 46).

Les résultats des analyses du CPPs sont moins concluants que ceux de la HNR ou de H1-H2, bien que les analyses montrent des effets similaires en fonction de la valence. Nous trouvons en effet un CPPs significativement moins prononcé, donc une voix plus soufflée (Stoller et Dixon, 2016), lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive, mais sur seulement deux segments. Nous trouvons également une différence indiquant une voix plus soufflée pour le négatif que pour le neutre, cependant cette différence est presque inexistante. Enfin, aucune différence significative n'a été trouvée entre les valences positives et neutres. Les coefficients paramétriques présentés en Annexe 8 ne démontrent aucune différence significative quelle que soit la valence [$p > 0,025$]. La prise en compte du temps est donc nécessaire pour observer les différences entre les valences, qui pourront être localisées à certains points précis de l'énoncé.

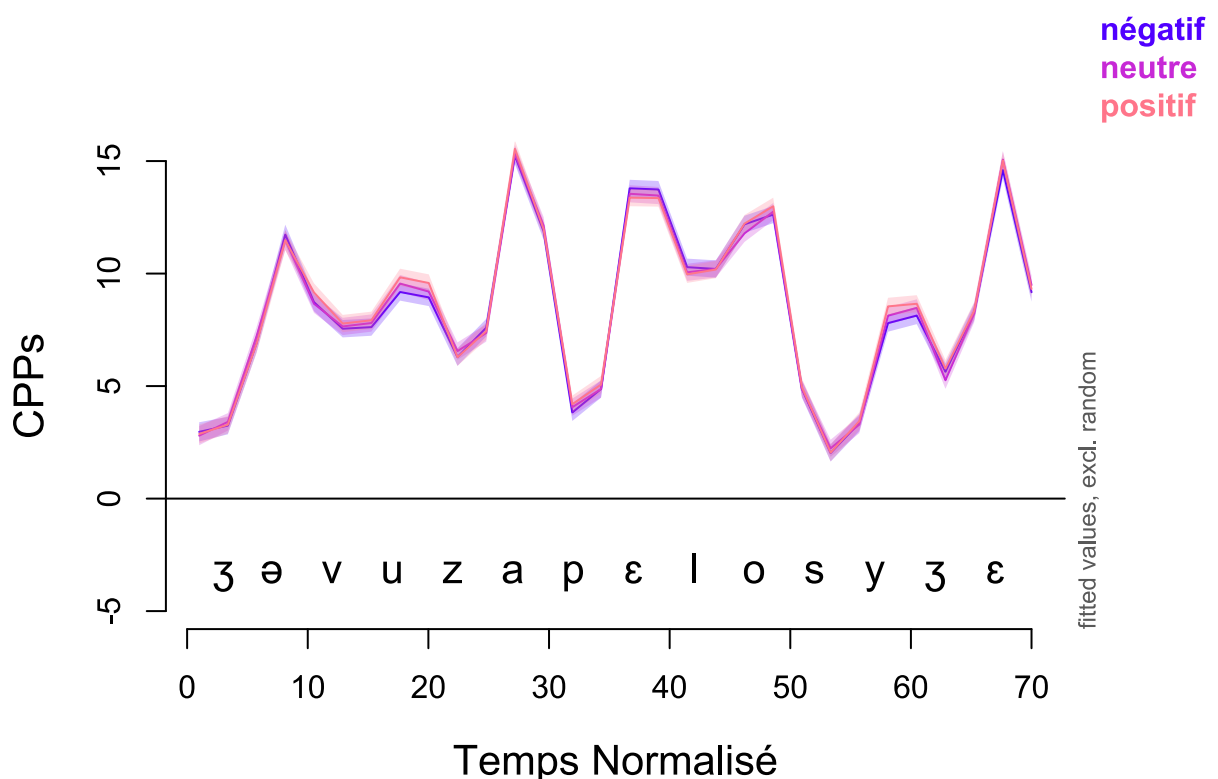


Figure 45 – CPPs au sein de la préface par valence

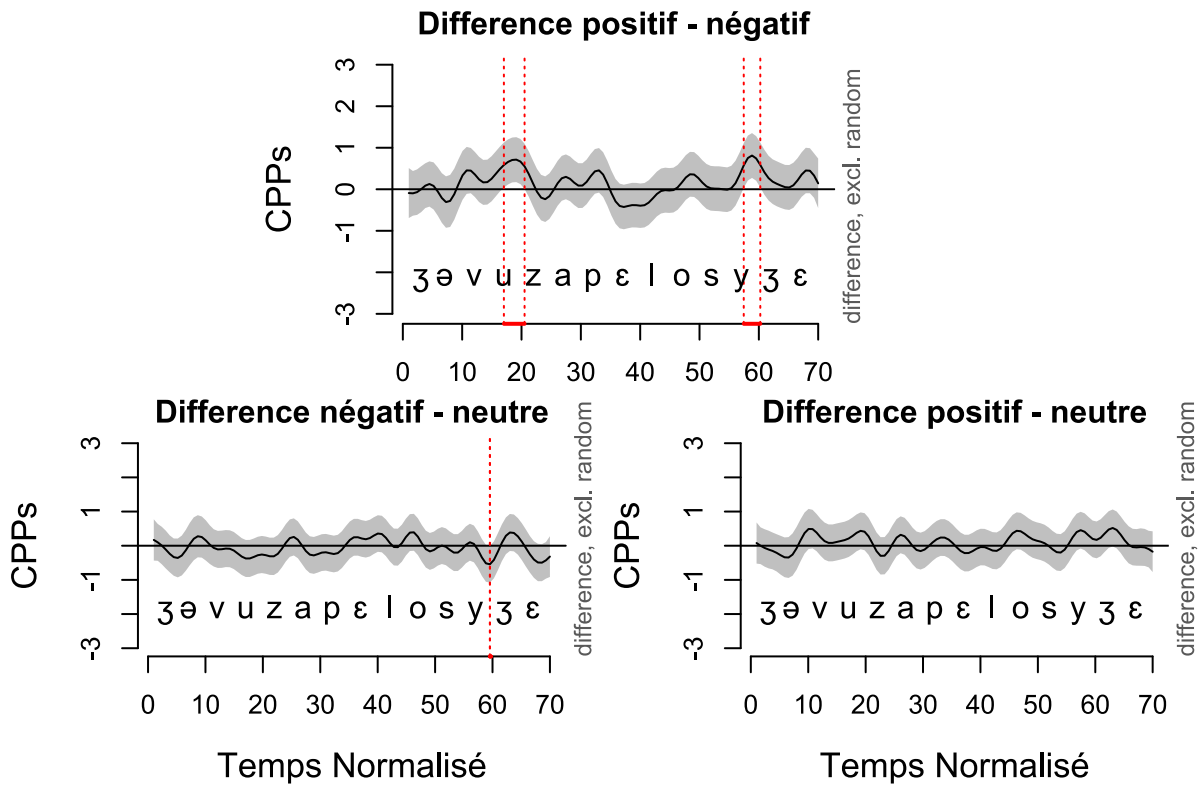


Figure 46 – Différences significatives du CPPs au sein de la préface

3.3.2. Au sein des marqueurs discursifs « voilà donc »

Pour les différentes mesures de la HNR dans les marqueurs discursifs, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Nous procédons donc à une inspection visuelle des données (Figure 47 et Figure 48 pour la HNR05, Figure 49 et Figure 50 pour la HNR15, Figure 51 et Figure 52 pour la HNR25, Figure 53 et Figure 54 pour la HNR35).

Il est à noter que le premier segment /v/ des marqueurs discursifs ne semblent pas permettre un calcul correct de la HNR sur le début de la consonne, pas plus que le segment /k/ en fin de marqueurs. Le segment /v/ semble avoir été souvent dévoisé en /f/, à l'instar de ce que nous avons pu observer dans la préface. Cela pourrait expliquer la détection d'une HNR négative extrêmement basse.

Nous trouvons des différences significatives au début de « donc » sur la consonne voisée. Le ratio son bruit est significativement plus bas donc la voix plus soufflée lorsque la valence est positive ou négative par rapport à lorsqu'elle est neutre dans les 4 mesures de la HNR (seule la HNR05 ne présente pas la différence positif – neutre). Lorsque la valence est neutre, la HNR semble rester positive, alors qu'elle passe en négatif pour les valences positives et négatives, indiquant une voix plus modale pour le neutre, et des dévoisements pour le positif et le négatif. Ainsi, la voix apparaît plus soufflée ou dévoisée lorsque la valence d'une nouvelle est négative que lorsqu'elle est positive ou neutre. Cela est cohérent avec la préface. En revanche à l'inverse de la préface, la valence positive semble induire une voix plus soufflée ou dévoisée lorsque la valence est positive plutôt que neutre. Ces différences sont cependant très localisées et peu nombreuses.

négatif
neutre
positif

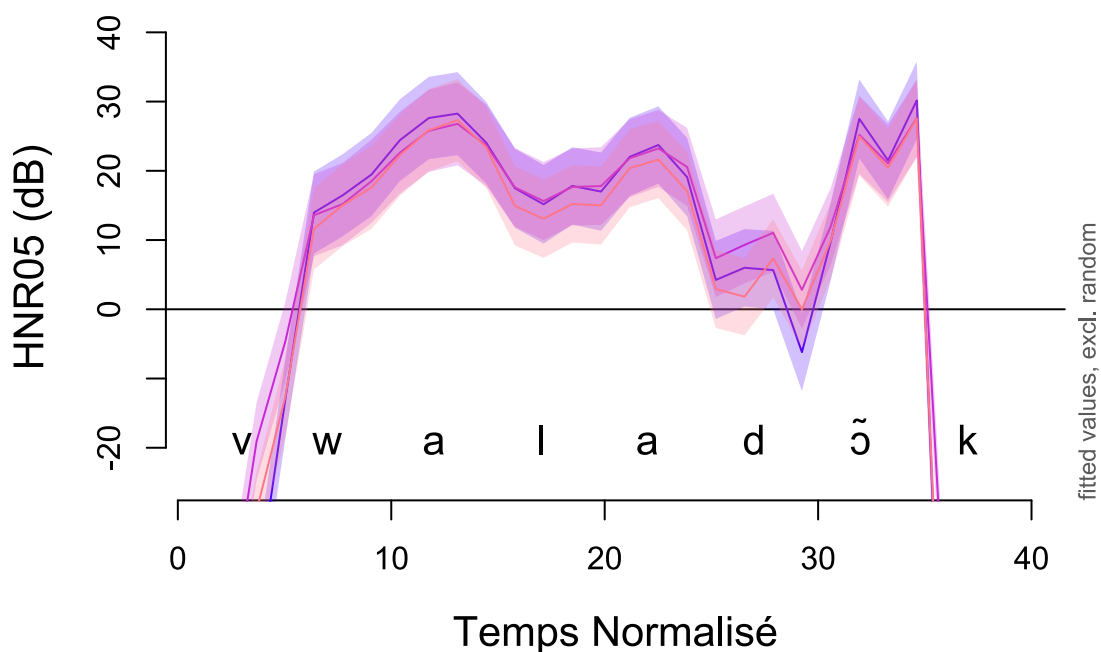


Figure 47 – HNR05 au sein des DMs par valence

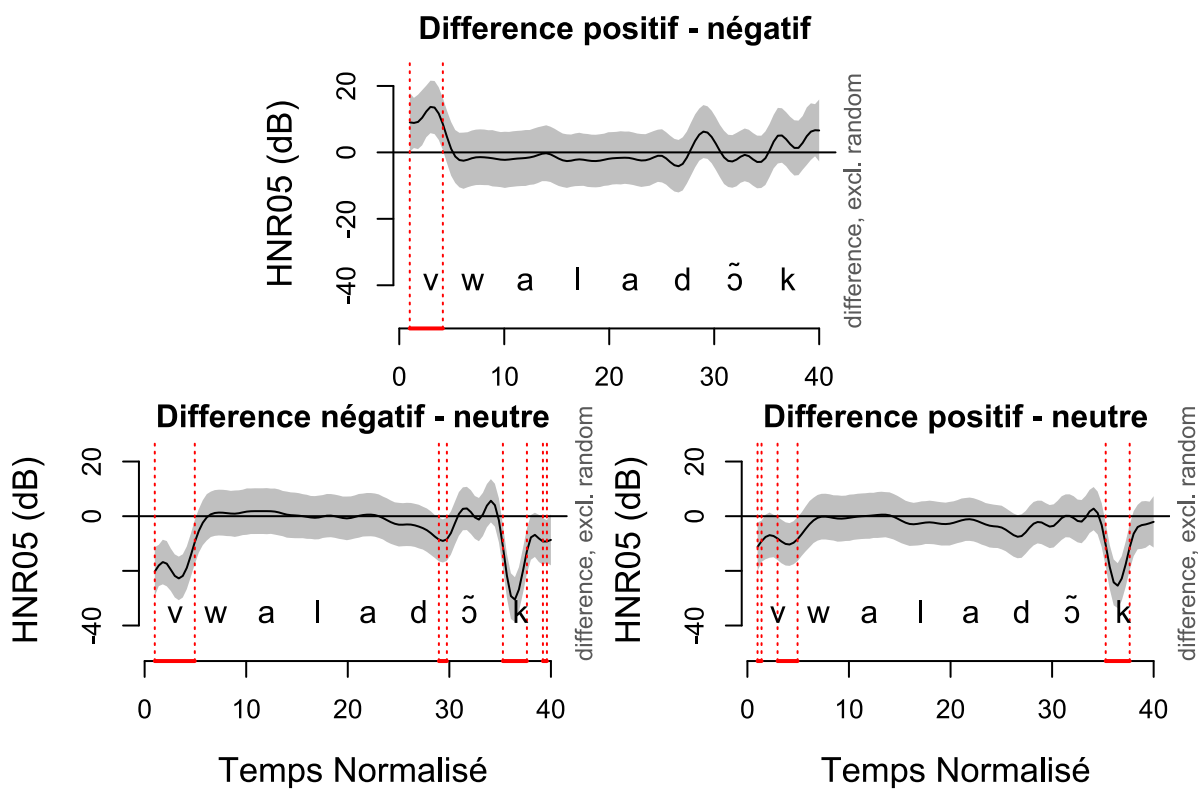


Figure 48 – Différences significatives de HNR05 au sein des DMs

négatif
 neutre
 positif

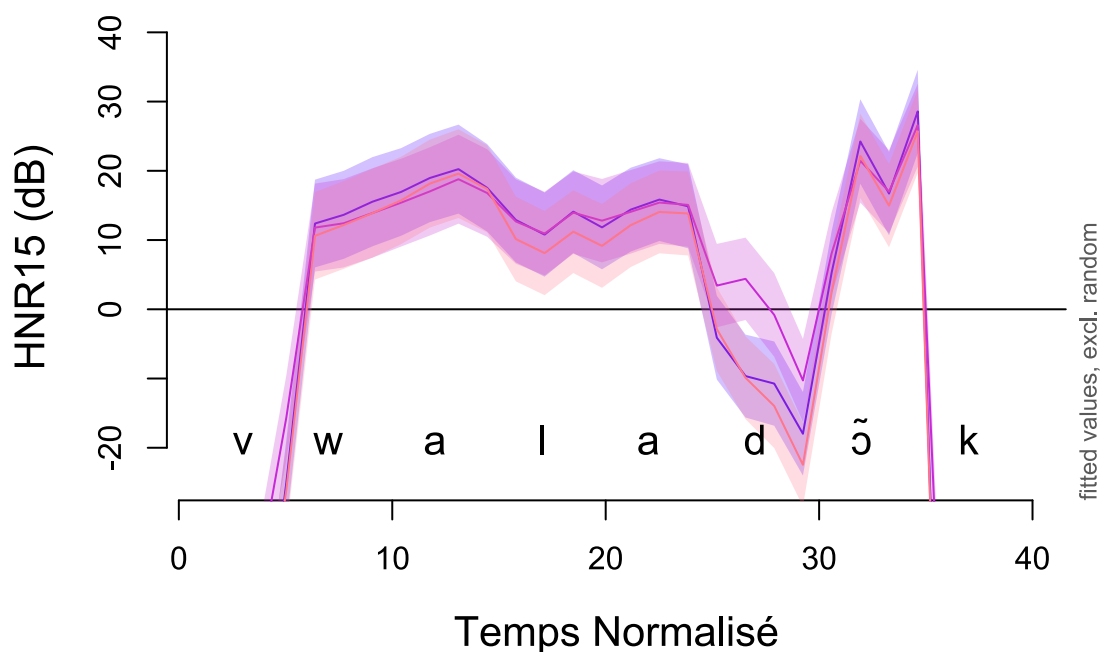


Figure 49 – HNR15 au sein des DMs par valence

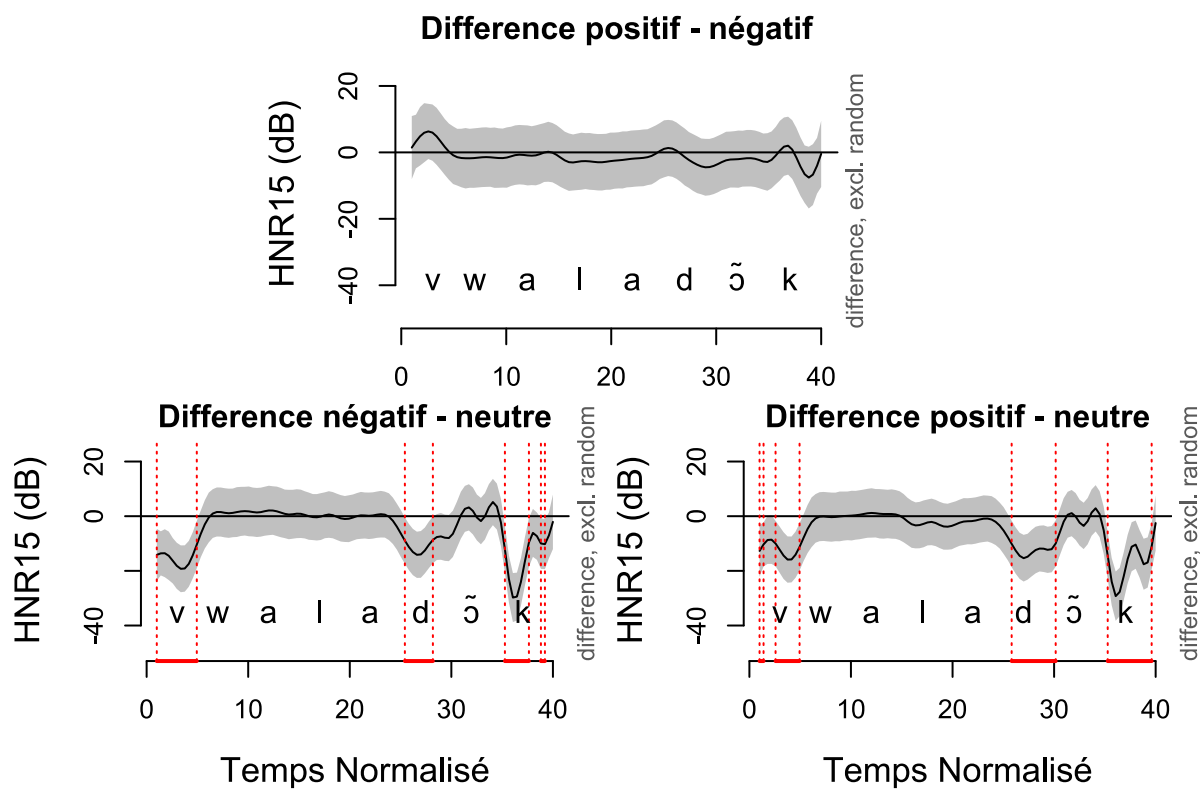


Figure 50 – Différences significatives de HNR15 au sein des DMs

négatif
 neutre
 positif

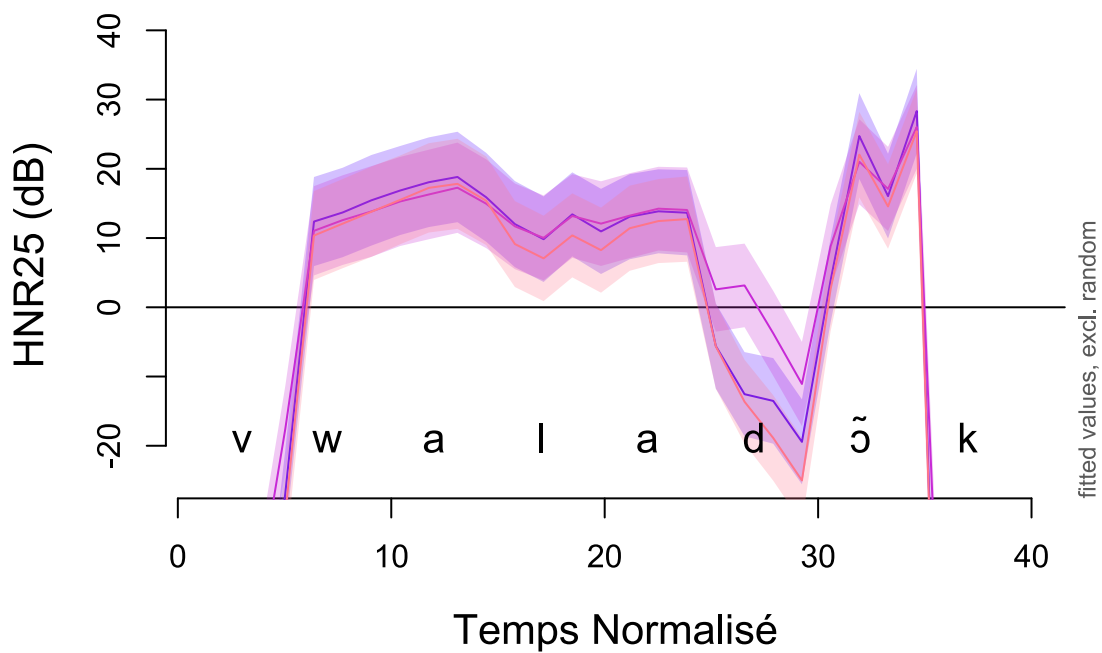


Figure 51 – HNR25 au sein des DMs par valence

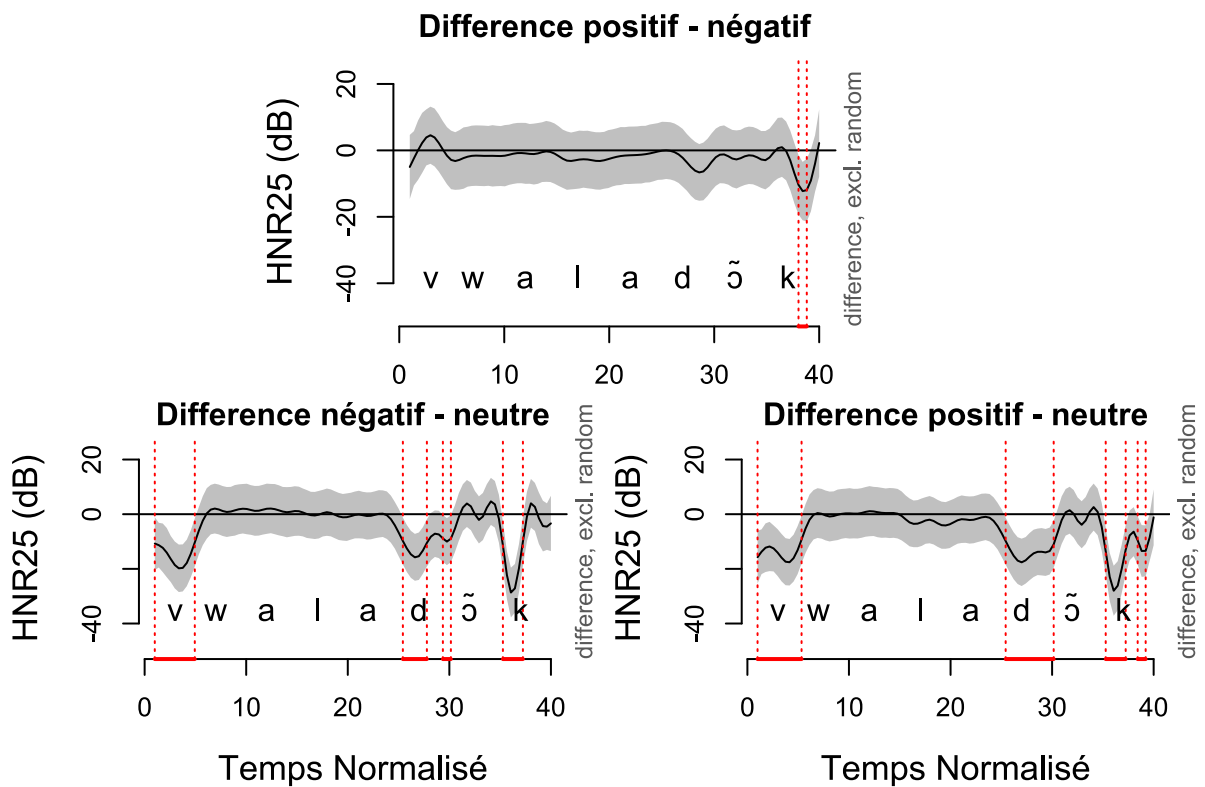


Figure 52 – Différences significatives de HNR25 au sein des DMs

négatif
 neutre
 positif

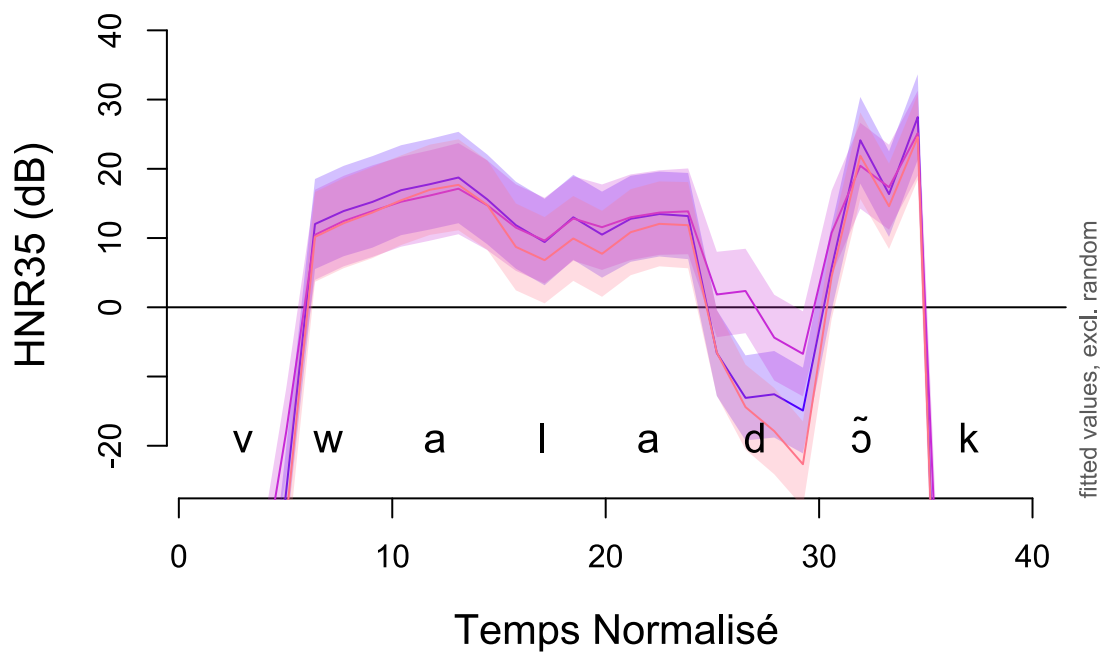


Figure 53 – HNR35 au sein des DMs par valence

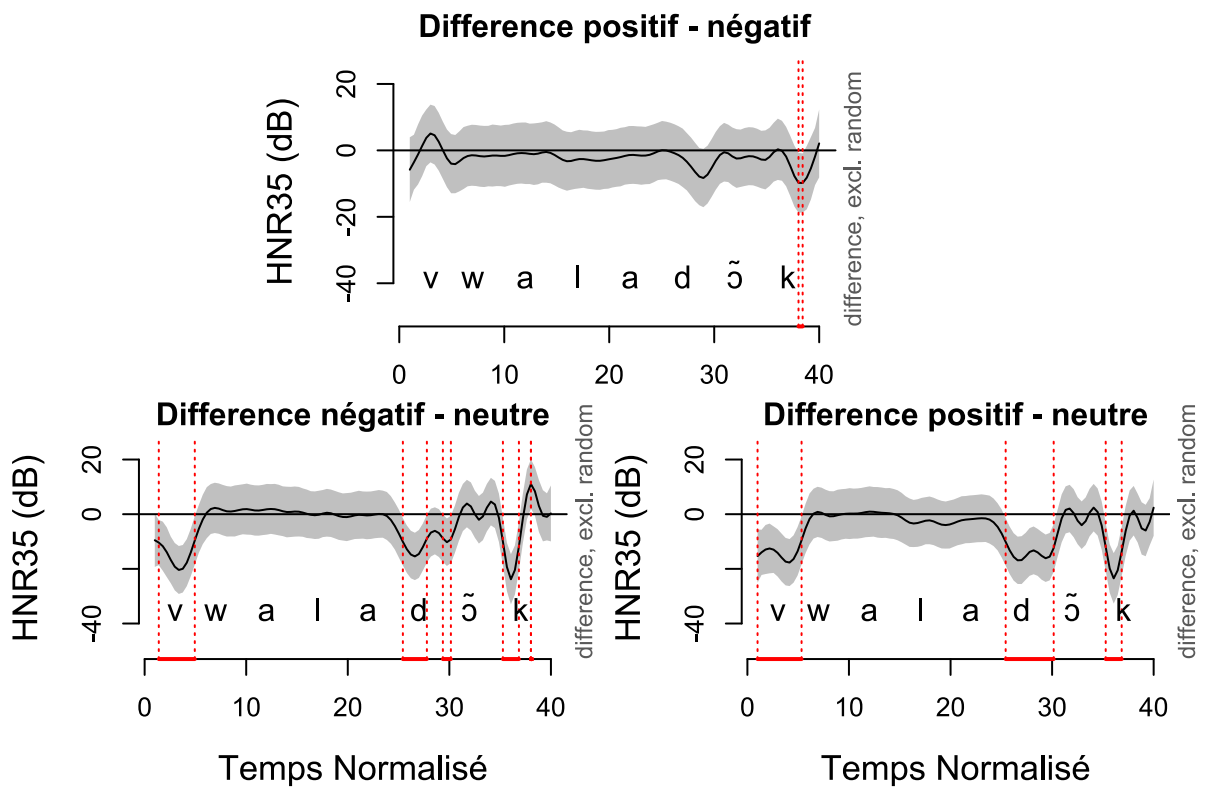


Figure 54 – Différences significatives de HNR35 au sein des DMs

Pour la différence entre les deux premières harmoniques (H1-H2) dans les marqueurs discursifs, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). Nous procédons donc à une inspection visuelle des données (Figure 55, Figure 56).

Peu de différences apparaissent dans les mesures des harmoniques pour les marqueurs discursifs. Nous trouvons une différence H1-H2 significativement plus haute, donc une voix plus soufflée, lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive ou neutre, sur « donc ». H1-H2 est aussi significativement plus bas donc la voix plus modale pour le négatif par rapport au neutre, sur « voilà » (/l/), et lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre à différents endroits de l'énoncé. Enfin, H1-H2 est significativement plus bas lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre, donc la voix est plus modale pour le positif. Ces résultats sont cohérents avec ceux trouvés dans le reste des analyses, en dehors d'une petite variation dans la comparaison négatif – neutre, puisque c'est le neutre qui est plus modal dans les préfaces, et le négatif plus soufflé. Les différences significatives se trouvent principalement en fin de marqueurs discursifs, sur les phonèmes /õk/, et au milieu de « voilà » sur le phonème /l/. Les coefficients paramétriques présentés en Annexe 8 ne démontrent aucune différence significative quelle que soit la valence [$p > 0,025$]. La prise en compte du temps est donc nécessaire pour observer les différences entre les valences, qui pourront être localisées à certains points précis de l'énoncé.

négatif
 neutre
 positif

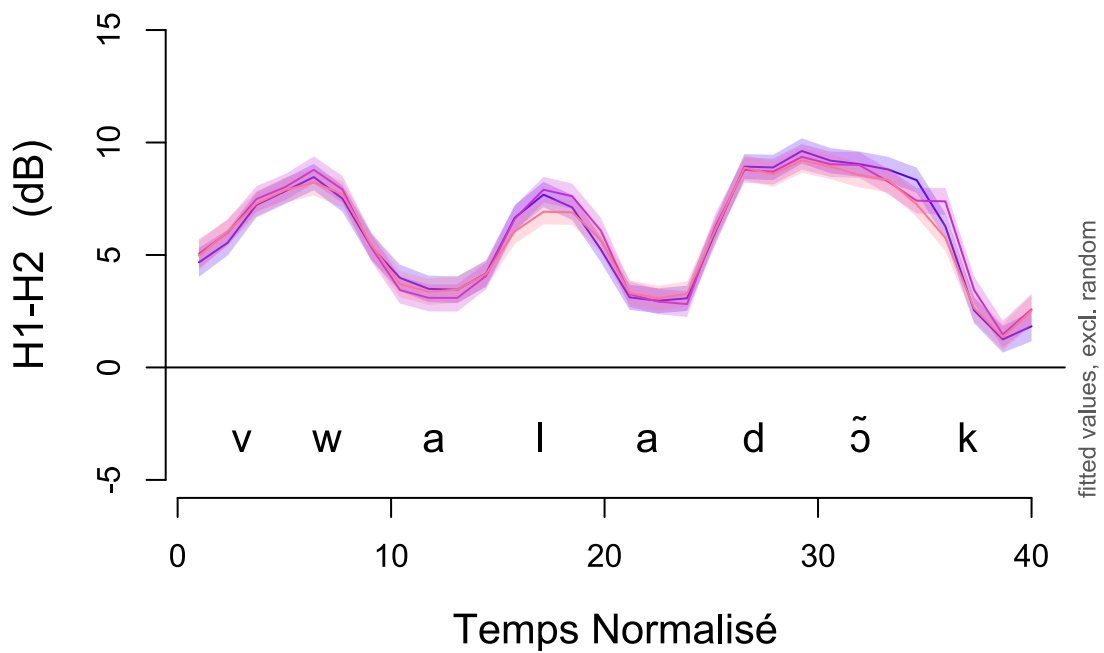


Figure 55 – H1-H2 au sein des DMs par valence

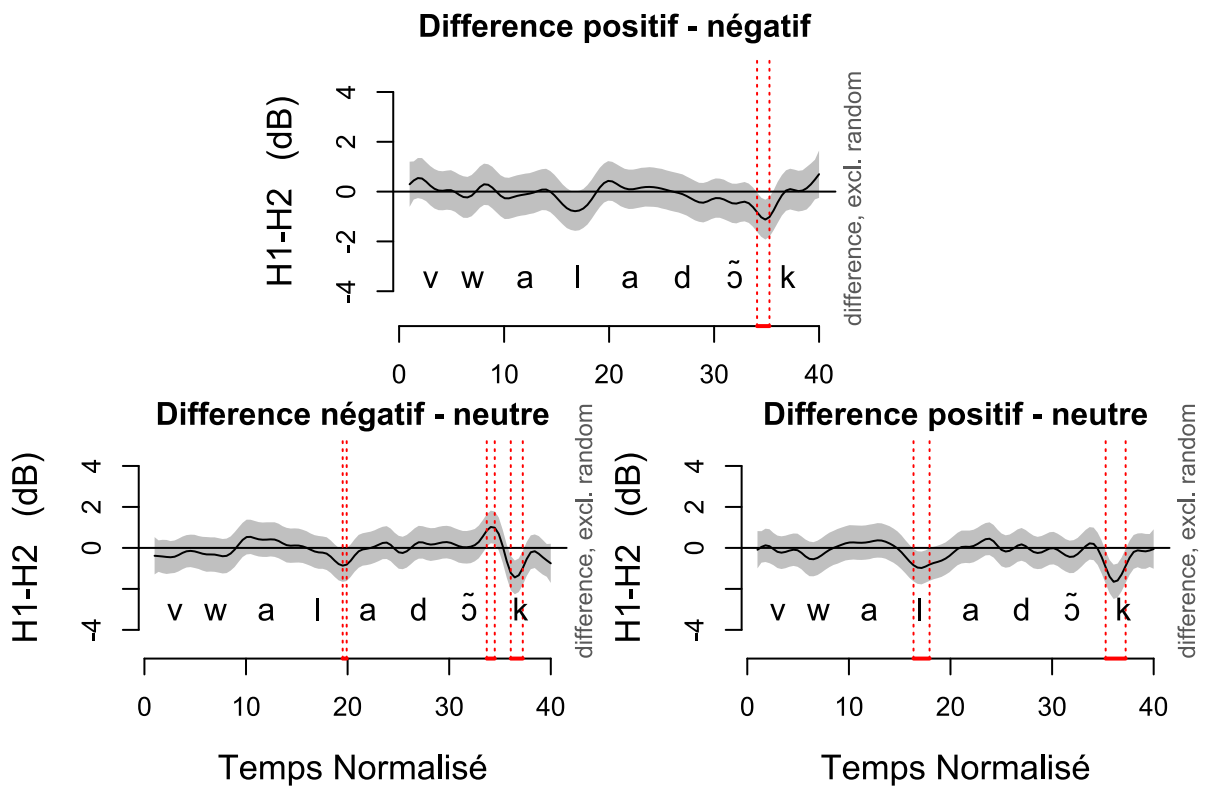


Figure 56 – Différences significatives de H1-H2 au sein des DMs

Pour la proéminence du pic cepstral (CPPs) dans les marqueurs discursifs, les valeurs p et edf des termes lissés ainsi que de la pente aléatoire indiquent que les prédicteurs sont bien non-linéaires (cf. Annexe 8). L'inspection visuelle des données (Figure 57, Figure 58) démontre que presque aucune différence significative n'est trouvée pour ces mesures hormis une voix plus soufflée pour la valence négative par rapport à la neutre au tout début des marqueurs discursifs. Ce résultat est cohérent avec les précédentes mesures.

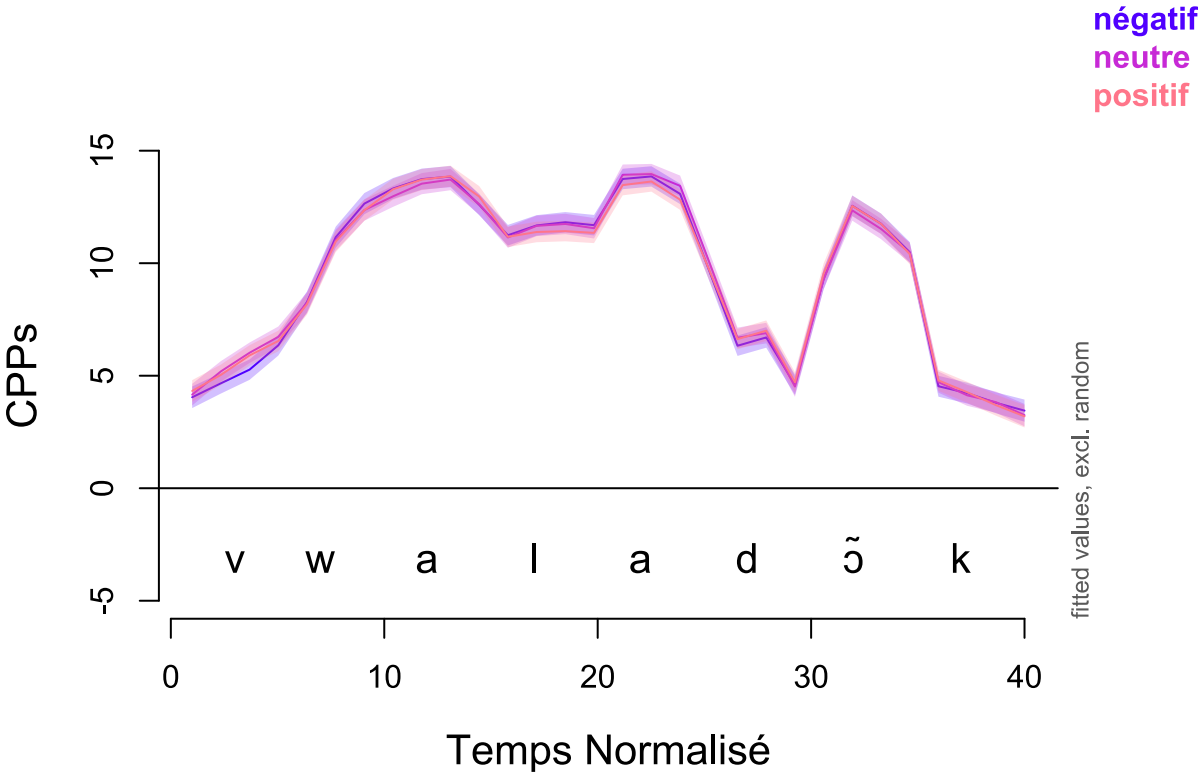


Figure 57 – CPPs au sein des DMs par valence

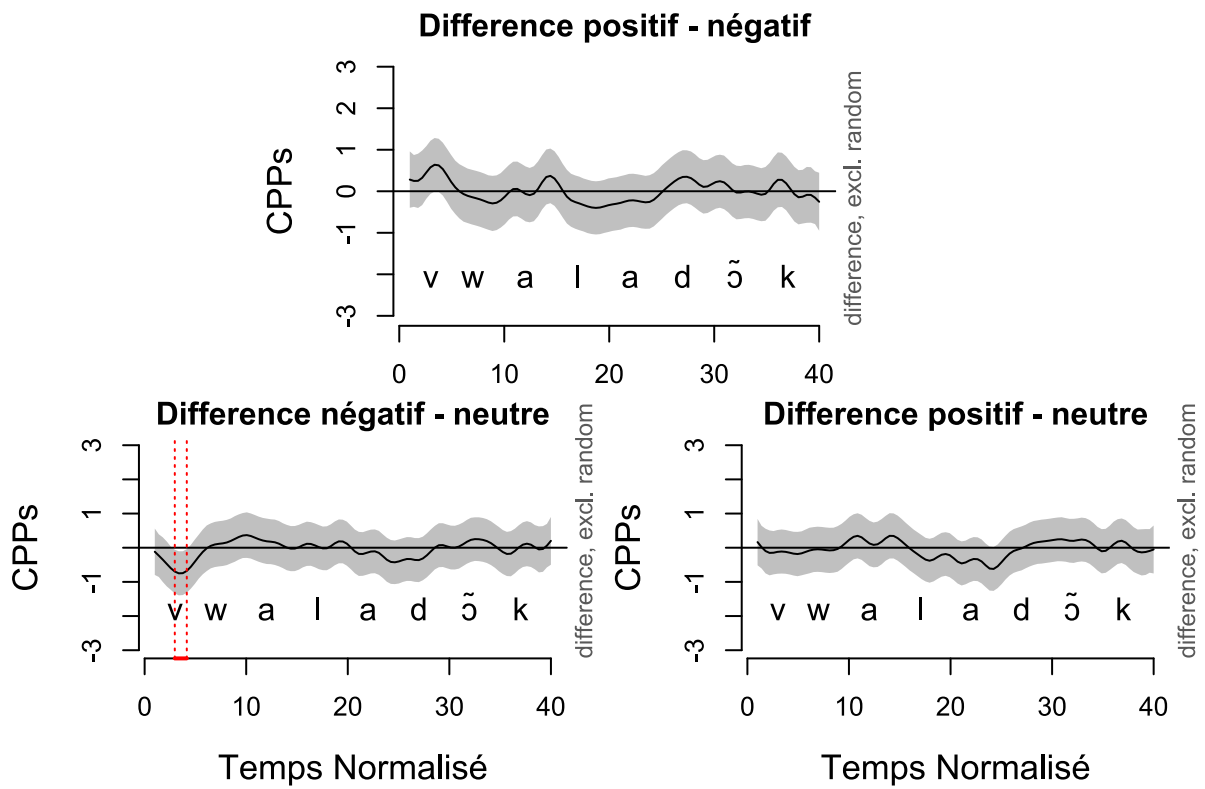


Figure 58 – Différences significatives du CPPs au sein des DMs

4. Discussion

Dans ce chapitre, suivant les travaux de Swerts et Hirschberg (2010) et ceux de Freese et Maynard (1998), nous nous sommes focalisés sur les variations phonétiques continues en tant que potentiels véhicules d'indices de projection prosodique. Nous avons présenté le corpus créé, ainsi que les analyses phonétiques réalisées sur la première étude proposée dans ce travail. La prosodie pouvant être portée à travers des variations de caractéristiques phonétiques, nous avons fait l'hypothèse que certaines variations pourraient apparaître en fonction de la valence de la nouvelle à annoncer. Nous souhaitons examiner la présence et la nature de ces corrélats phonétiques. Nous voulions aussi vérifier si ces indices de projection étaient consistants au cours des différentes parties de l'annonce.

Valence	Préface			MDs			Annonce		
	Neutre	Positive	Négative	Neutre	Positive	Négative	Neutre	Positive	Négative
Empan vocal	<>	>	<	>	>	<	<>	>	<
f0 moyenne	<>	>	<	>	>	<	<>	>	<
f0 maximum	<>	>	<	<>	>	<	<>	>	<
f0 minimum	-	>	<	-	-	-	>	>	<
Courbe de f0	<>	>	<	>	>	<	X	X	X
Durée énoncé	-	>	<	<	>	>	<	<>	>
Durée pauses	-	>	<	X	X	X	<	<	>
Vitesse d'art.	-	-	-	-	-	-	-	-	-
HNR	<	>	<	>	<	<	X	X	X
H1-H2	<>	<	>	<>	<	>>	X	X	X
CPPS	>	>	<	>	-	<	X	X	X

Tableau 10 – Tableau récapitulatif des différences phonétiques entre les trois valences, au sein des différents énoncés (< valeur basse, > valeur haute, <> intermédiaire, _ non significatif, >< inconsistant, X NA)

Les résultats principaux, récapitulés de façon impressionniste dans le Tableau 10, vont dans le sens de notre première hypothèse pour cette étude : les locuteurs qui s'apprêtent à annoncer une nouvelle à un interlocuteur semblent bel et bien produire des indices de projection prosodiques au moyen de caractéristiques phonétiques continues. Les indices semblent être utilisés dans chacune des trois parties étudiées : préface, marqueurs discursifs, annonce. Le contraste le plus fort est trouvé entre l'annonce d'une nouvelle positive et celle d'une nouvelle négative. De façon générale, nous avons trouvé que lorsque l'on s'apprête à annoncer une nouvelle négative, nous produisons les énoncés avec des valeurs de la f_0 abaissées, résultant de la baisse des valeurs minimum, maximum et moyennes de la f_0 , ainsi que de l'étrécissement de l'empan vocal. Nous les produisons également avec des durées plus longues pour l'énoncé et pour les pauses. Enfin, nous produisons les énoncés avec une voix plus soufflée, résultant de la baisse de la HNR et du CPPs, et de la hausse de H1-H2. Le second contraste le plus présent intervient entre les valences négative et neutre. Dans cette opposition, la prosodie lorsque la valence de la nouvelle est négative présente également par rapport au neutre des valeurs de la f_0 abaissées, des durées d'énoncés et de pauses allongés et une voix plus soufflée. Le contraste le moins prégnant est celui qui oppose les énoncés produits avant l'annonce d'une nouvelle positive ou neutre. Lorsque les effets ressortent, les énoncés prononcés quand la nouvelle est positive plutôt que neutre démontrent des valeurs de la f_0 élevées, des durées d'énoncés plus longues et une voix plus modale. Ces résultats sont en cohérence avec la littérature sur la prosodie affective des émotions comme la joie et la tristesse (Bänziger & Scherer, 2005; Berry & Brown, 2019; Goudbeek & Scherer, 2010; Scherer, 2003). Les résultats sont aussi en cohérence avec la littérature sur la prosodie affective utilisée dans les annonces de nouvelles, qui ont trouvé des tendances similaires à celles de la joie pour les annonces positives, et similaires à celles de la tristesse pour les annonces négatives (Freese & Maynard, 1998; Swerts & Hirschberg, 2010).

Les résultats de notre étude nous permettent de supporter l'idée généralement acceptée que la prosodie affective est caractérisée par des variations de la fréquence fondamentale, de durée et de caractéristiques spectrales. Suivant Planalp (1993) et Guerrero et al. (1996), nous voyons l'affect comme un phénomène social et interpersonnel (bien que nous ne rejetions pas le caractère personnel de l'affect). Il peut donc jouer un rôle dans l'interaction et être utilisé à des fins communicatives. Nous croyons que la présence et la cohérence de variations phonétiques dépendantes de la valence affective

d'une nouvelle, et que le fait que cette présence puisse être en amont même de l'annonce d'une nouvelle est une démonstration du rôle interactionnel projectif de la prosodie affective.

En revanche, nous avons émis l'hypothèse que la prosodie produite lorsque la nouvelle est neutre serait plus ou moins intermédiaire entre valeurs négatives et positives. Nos résultats ne valident pas cette hypothèse et ne vont pas dans le même sens que ceux des études sur la perception de la parole affective (Audibert, 2008; Audibert et al., 2005; Laukka et al., 2012). En effet dans ces études, la prosodie du neutre et de la tristesse ou l'anxiété, des affects à valence négative, sont soit confondues par les auditeurs si tous les indices phonétiques ne sont pas présents (pour la tristesse), soit proche acoustiquement (avec un contour de la f_0 similaire à celui de l'anxiété). Ainsi, si la prosodie des nouvelles neutres n'est pas équidistante du positif et du négatif, nous aurions pu penser qu'elle serait plus proche du négatif. Cependant, elle est plus proche du positif, comme on peut le voir à l'observation de la magnitude des effets sur l'empan vocal, la f_0 maximum et la f_0 moyenne. De plus certaines caractéristiques acoustiques (notamment pour la f_0 minimum), ne sont pas significatives entre le positif et le neutre. Cette proximité du neutre et du positif pourrait être liée à des enjeux conversationnels liés à la nature de l'interaction : la séquence d'annonce. En effet, annoncer des nouvelles peut être une tâche complexe si l'on souhaite la délivrer avec succès et donc réussir l'interaction, tout en protégeant sa propre face et celle de l'interactant (Heritage, 1989). Annoncer des nouvelles bonnes ou neutres pourrait être une action conversationnelle relativement sûre car plaisante ou au moins neutre. Nos nouvelles « neutres » sont en effet des appels où l'on déplace une réunion d'une demi-heure, confirme une information déjà connue ou donne de nouvelles informations sans contenu émotionnel. La nouvelle négative en revanche serait plus périlleuse. Des études cognitives et psychophysiologiques sur les annonces de nouvelles par les médias ont en effet démontré que les annonces de nouvelles négatives élicitent des réactions plus fortes et plus longues chez les personnes qui les entendent par rapport aux annonces de bonnes nouvelles. Ces études évoquent comme raison le biais de négativité généralement observé chez l'être humain, qui pousse à porter plus d'attention et mieux retenir ce qui est négatif (M. E. Grabe & Kamhawi, 2006; Soroka et al., 2019; Soroka & McAdams, 2015). Au vu du caractère important et interactionnellement sensible de l'annonce de nouvelles négatives, nous pouvons donc penser que les locuteurs vont essayer de montrer l'importance de ce qu'ils vont dire en

démontrant leur attention vis-à-vis de ce qu'ils sont en train de dire par une plus forte présence d'indices de projection, ce qui pourrait les avoir poussés à plus démarquer la valence négative des deux autres. Nous allons maintenant nous attarder quelque peu sur chacune des caractéristiques acoustiques étudiées.

Les variations de la f_0 semblent être l'un des corrélats les plus récurrents et les plus évidents de la valence affective dans notre corpus. Nous avons trouvé que la f_0 est une caractéristique acoustique qui présente des différences significatives dans presque tous les énoncés. Les mesures globales tendent toujours dans la même direction, quelle que soit la partie de discours observée : dans la préface, les marqueurs du discours « voilà donc », et la séquence d'annonce elle-même. Au sein de ces trois parties, les données présentent des différences quelles que soient les valences comparées, sur l'empan vocal utilisé par les locutrices dans l'enregistrement du corpus, la f_0 moyenne, et les valeurs maximales. Les valeurs de la f_0 minimum apparaissent comme les plus inconsistantes car elles sont plus souvent non significatives que les autres données. Ce résultat n'est pas en accord avec l'étude de Liscombe et al. (2003), qui avaient trouvé que la f_0 minimum était notamment un bon prédicteur pour la joie. Nous pensons que le peu de résultats trouvés sur la f_0 minimum, dans les marqueurs discursifs particulièrement, pourrait être une conséquence d'un effet « de sol ». Entre les différents énoncés, c'est justement sur les marqueurs discursifs que la prosodie semble porter le moins d'indices de la valence affective, avec notamment une différence entre positif et neutre qui ne ressort que pour la f_0 maximum. Les marqueurs étant, en général, produits dans un empan vocal plus bas que le reste de l'annonce (préface, annonce), il est possible que cela ait laissé moins de marge aux participants pour varier ce point, étant déjà dans la partie basse de leur empan vocal personnel. Le fait que les effets soient moins forts dans les marqueurs discursifs pourrait également être expliqué par deux autres facteurs. D'une part, ces énoncés sont bien plus courts que les deux autres et donnent ainsi moins d'espace à de la variation, comme l'atteste l'empan vocal beaucoup plus étroit, situé autour des 50 Hertz dans les marqueurs discursifs contre un empan situé autour des 75 Hertz dans les préfaces et les annonces. D'autre part, la préface étant un espace conversationnel dédié à la projection, les locuteurs pourraient s'être plus appuyés sur ces énoncés que sur le reste pour véhiculer des indices de projection.

Un autre résultat majeur de cette étude réside dans ce que nous a révélé l'observation de la courbe de la f_0 en fonction du temps, en début de préface (« Je vous appelle au sujet »). Nous avons trouvé des effets allant dans le même sens que les caractéristiques phonétiques trouvées dans les mesures générales de la f_0 . Ces mesures sont significatives sur la quasi-totalité de l'énoncé pour la comparaison entre la valence positive et la valence négative. Ainsi, en comparant les mesures globales moyennées sur un énoncé avec des données représentées au fil du temps, nos résultats permettent de supporter la validité d'utilisation de mesures phonétiques globales pour l'observation de la f_0 dans l'analyse des affects (Berry & Brown, 2019; Scherer, 2003). Ce résultat est plutôt en contradiction avec l'une des hypothèses de cette thèse : le fait que nous devrions peut-être éviter de prendre uniquement des mesures globales de la f_0 au profit de mesures plus précises présentes à un niveau potentiellement phonologique, sur des points-clés de l'énoncé. Cette découverte est nuancée par les résultats trouvés au sein des marqueurs discursifs. Pour ceux-ci, l'ajout du facteur temps dans l'analyse de la f_0 montre que les différences significatives entre les valences sont généralement concentrées sur le premier marqueur « voilà » et pas du tout sur le second, « donc ». On peut alors penser que ce n'est pas l'entièreté de l'énoncé, mais seulement une partie, qui est utilisée comme support de la prosodie affective. Il est ainsi possible que les différences de f_0 soient situées en début d'énoncé (« je vous appelle au sujet » et « voilà ») plutôt qu'à la fin de l'énoncé (« de vos examens médicaux », « donc »). En effet, nous n'avons pas pu faire l'observation de la f_0 en fonction du temps sur la seconde partie de la préface et il est possible qu'elle se comporte de la même manière que « donc ». Cette interprétation de nos résultats est soutenue par une seconde observation. Si les différences entre valences positive et négative sont ressorties non seulement sur les graphiques des modèles GAMM mais aussi dans les termes paramétriques qui ne prennent pas en compte la mesure du temps, ce n'est pas le cas des différences entre le négatif et le neutre ou le positif et le neutre. Cela est surprenant puisque nos mesures globales démontraient de telles différences comme significatives. Il se pourrait donc, à nouveau, que les différences significatives soient situées après la portion « je vous appelle » sur la partie changeante que nous n'avons pu analyser.

Ces résultats des mesures globales et temporelles de la f_0 ont donc démontré la présence de différences significatives entre les valences. Cependant, nous nous demandons si ces indices de projection liés à la f_0 sont suffisamment forts pour être

perceptibles. Nous avons donc regardé la magnitude des différences de f_0 . Les variations de l'empan vocal présentent des différences assez marquées dans la préface (13 Hz d'écart dans la préface, 15 Hz dans les marqueurs discursifs, et 26 Hz dans l'annonce) entre le positif et le négatif. Ces différences sont bien moins marquées (5 Hz dans la préface, 6 Hz dans les DMs, et 15 Hz dans l'annonce) entre les valences positive et neutre par exemple. Il a été montré que les différences de f_0 dans la parole (calculées en semi-tons) en affectent la perception, et notamment son intelligibilité (Assmann, 1999; Drullman & Bronkhorst, 2004). Les mots seraient mieux reconnus lorsque les différences de f_0 sont plus grandes, et moins bien reconnues lorsqu'elles sont plus petites. En conséquence, les différences de hauteur de la f_0 observées entre la valence positive et la valence négative, puisqu'elles sont plus grandes, ont plus de chances perçues par les auditeurs. Les tendances décrites pour l'empan vocal sont les mêmes pour les valeurs moyennes, et les valeurs maximums de la f_0 . Les valeurs minimums, qui présentaient rarement des différences significatives, présentent des différences entre les valences comprises entre 0 et 5 Hertz, tout au plus, soit bien plus petites que celles de l'empan vocal. Il est aussi à noter qu'en moyenne, les marqueurs discursifs sont produits plus bas, avec une moyenne de la f_0 plus proche des 200 Hertz que des 225 Hertz que l'on peut retrouver dans la préface ou les séquences d'annonce, et des valeurs maximales autour des 225 Hertz plus que des 270 Hertz de la préface et de la séquence d'annonce.

Nous allons maintenant nous attarder sur la seconde grande catégorie de corrélats phonétiques que nous avons observé : les données temporelles. Nous avons trouvé pour celles-ci des différences significatives de durée dans la préface, les marqueurs discursifs et l'annonce. Malgré la présence d'effets de la valence sur les données temporelles, ce ne semble pas être un véhicule principal de la valence affective pour la projection d'indices. En effet bien que les différences soient régulières pour les séquences d'annonces, elles ressortent moins souvent pour les préfaces et les marqueurs discursifs. Cette tendance se retrouve aussi pour la durée des pauses, plus marquée dans la séquence d'annonce, moins marquée dans la préface. Nous n'avons pu prendre cette mesure sur les marqueurs discursifs, trop peu de pauses y étant réalisées. Les tendances retrouvées dans la préface et l'annonce pour la durée des énoncés et des pauses (allongées pour les nouvelles négatives, écourtées pour les positives et les neutres) sont différentes de celles trouvées pour les marqueurs discursifs (la durée des énoncés est écourtée pour la valence neutre

par rapport aux deux autres). Ces résultats sont cohérents avec la littérature sur la tristesse, l'étude de Freese et Maynard (1998), et nos hypothèses. Ils sont en revanche en accord avec les résultats de Swerts et Hirschberg (2010), qui avaient trouvé que la parole était plus rapide lorsque l'annonce de nouvelle était négative. En conclusion, il semble que la temporalité des préfaces et des marqueurs de discours joue un rôle dans la production d'indices de projection de la valence, bien que moins présent que celui des variations de la f_0 . La projection d'indices pour différencier entre nouvelles négatives et neutres, et positives et neutres semble plutôt se faire au sein des marqueurs du discours, avec une durée plus longue de ceux-ci lorsque l'on a une valence affective marquée (positif, négatif), et plus courte lorsque l'on a une valence neutre. Contrairement à la littérature sur la tristesse et la joie (Berry & Brown, 2019; Scherer, 2003) et sur les annonces de nouvelles (Freese & Maynard, 1998), nous n'avons pas trouvé de différences de vitesse d'articulation, ce quelle que soit la valence de la nouvelle. Les différences de durées qui sont ressorties restent relativement faibles en termes d'amplitude, et peu nombreuses, surtout dans les portions de discours qui viennent avant la séquence d'annonce (préface, marqueurs discursifs). Cette différence pourrait être due au fait que notre corpus est un corpus d'annonces lues et non un corpus naturel, les tâches de lecture pouvant influencer la vitesse d'articulation comparé à d'autres types de tâches de production (Avanzi, 2013; Crystal & House, 1990).

Nous avons observé la magnitude des différences de durée dans les préfaces, marqueurs discursifs, et séquences d'annonce. Si l'on observe les variations de durée de l'énoncé en fonction de la valence, nous voyons que la différence significative entre le positif et le négatif trouvée dans la préface est de 100 millisecondes. En revanche, celle trouvée entre le neutre et les valences marquées (positif, négatif) est d'environ 35 millisecondes, donc beaucoup moins marquée. Dans les séquences d'annonce en revanche, les différences de durée semblent plus marquées. Elle est d'environ 1 seconde entre le négatif et le positif, et de presque 2 secondes entre le négatif et le neutre. La différence entre le positif et le neutre est de presque 1 seconde. Ces trois différences ont ainsi plus de chances d'être perçues que celles trouvées dans les autres énoncés. De même, la durée des pauses entre le positif et le négatif dans la préface diffère de 55 millisecondes, alors que dans l'annonce, elle diffère de 224 millisecondes.

Les dernières mesures sur lesquelles nous nous attardons sont celles des caractéristiques spectrales de la voix, caractérisées par les analyses de la HNR, de H1-H2 ; et du CPPs. Ces données mettent elles aussi en évidence des différences significatives au sein des préfaces et des marqueurs discursifs, selon la valence de la nouvelle à annoncer. Cela montre que nous utilisons aussi ce paramètre vocal comme véhicule pour projeter des indices phonétiques sur la valence de la nouvelle. Les résultats principaux de ces analyses indiquent une tendance à utiliser une voix soufflée lorsque l'on s'apprête à annoncer une nouvelle négative particulièrement. A l'inverse, la voix produite lorsque la nouvelle à annoncer est positive est plutôt modale. Le neutre est généralement situé entre les deux. Ces résultats sont cohérents avec la littérature et notamment, l'idée que la voix soufflée soit utilisée pour établir une certaine intimité (Laver, 1980) et démontrer l'attention que l'on porte à l'interlocuteur (Audibert, 2008; Campbell & Mokhtari, 2004), ce qui nous apparaît comme crucial pour annoncer des nouvelles négatives interactionnellement et affectivement sensibles. De plus, Freese et Maynard (1998) ont trouvé que les annonces de nouvelles positives ont tendance à être produites avec une voix « normale » qui correspond à ce que nous appelons « voix modale », et que les annonces de nouvelles négatives ont tendance à être produites avec une voix soufflée, parfois craquée.

L'observation des mesures de la HNR présente beaucoup de variabilité au sein de la préface, avec notamment des segments qui apparaissent comme dévoisés d'après les valeurs négatives de la HNR. Nous trouvons malgré tout que la valence positive est différenciée par une voix plus modale vis-à-vis des deux autres qui sont plutôt soufflées. Dans les marqueurs discursifs, les mesures de la HNR sont plus régulières mais ne montrent pas tout à fait les mêmes tendances que pour la préface. Le ratio harmoniques – bruit est en effet plus haut donc la voix plus modale lorsque la valence de la nouvelle à annoncer est neutre, par rapport à lorsque la valence de la nouvelle est positive ou négative, où la HNR devient parfois négative, indiquant un dévoisement qui ne serait pas présent pour le neutre. Les mesures de la différence entre les deux premières harmoniques ont montré, elles, beaucoup plus de différences significatives, et leur emplacement au sein des préfaces et des marqueurs discursifs reste très consistant à travers les différentes valences. Les mesures indiquent une voix plus modale lorsque la valence de la nouvelle à annoncer est positive, par rapport à lorsqu'elle est négative ou neutre, et plus modale également pour le neutre par rapport au négatif. Des différences

dans les mesures de la proéminence du pic cepstral lissé (CPPs) ont été trouvées lorsque l'on compare les annonces positives ou négatives, et indiquent elles aussi une voix plus soufflée lorsque l'on s'apprête à annoncer une nouvelle négative. Les différences sont cependant vraiment peu nombreuses, et le CPPs ne semble donc pas être l'une des mesures acoustiques les plus pertinentes pour observer la projection prosodique d'indices de la valence affective.

Notre corpus a permis de faire ressortir divers indices de projection phonétique de la valence affective, ce malgré la diversité des contextes sociaux d'annonce, et en dépit du fait que les nouvelles annoncées n'étaient jamais supposées susciter d'affects forts. La relation entre l'annonceur et celui qui reçoit la nouvelle dans notre corpus était en effet toujours soit une relation de service où la personne servie dépend de l'autre (ex, une hôtesse de l'air appelant une cliente), soit une relation d'autorité (ex, professeur – élève, RH – employée). Nous avons donc observé un contexte particulier, qui peut amener une utilisation des indices de projection de valence affective différente d'un contexte familial ou amical, par exemple. Dans ce cadre professionnel, nous pouvons penser que les indices phonétiques si nous en trouvons, auront une amplitude faible. En effet il a été démontré que maintenir le professionnalisme dans son travail est lié à la gestion des affects. Ceux-ci devraient être démontrés de façon dite « appropriée ». Pour Kramer et Hess (2002), ce caractère approprié de la démonstration d'affects correspond en général, pour les affects négatifs, à les masquer. Malgré tout, des différences significatives ont pu être rapportées.

Nous avons aussi observé des nouvelles dites « neutres », qui sont factuelles et concernent des détails anodins comme des changements d'heure ou de lieu de rendez-vous. Principalement choisies comme point zéro de comparaison, nous avons néanmoins pu mettre en évidence le fait qu'elles sont effectivement distinctes à la fois des nouvelles positives et des nouvelles négatives, bien que les indices de projection soient moins réguliers que dans les autres cas. La prosodie de la nouvelle neutre semble effectivement se situer entre le positif et le négatif, mais pas de façon équidistante entre les deux. De façon générale, la magnitude des différences, même lorsqu'elles sont significatives, montre une différence moins marquée entre les nouvelles positives et les nouvelles neutres, alors que les nouvelles négatives semblent toujours plus clairement distinguées par rapport aux deux autres valences grâce aux indices de projection. Ainsi, bien qu'acoustiquement, nous trouvons des différences entre les valences positive et neutre,

il est possible que celles-ci, si elles sont trop petites, puissent ne pas être perçues par un auditeur.

L'utilisation de modèles GAMM, et donc la prise en compte du temps sur l'analyse des données, nous a permis d'inclure une analyse des caractéristiques spectrales et de la qualité de voix, suivant la proposition de Swerts et Hirschberg (2010) pour agrandir nos connaissances sur les « prédicteurs émotionnels » ou affectifs, de façon plus générale. De plus, cela nous permet aussi d'avoir un regard plus précis sur les caractéristiques observées. Nous pouvons voir le développement des caractéristiques spectrales et de la f_0 dans le temps, et ainsi repérer les espaces où l'on retrouve plus souvent des différences significatives. Dans la préface, ces espaces sont assez changeants. Dans les marqueurs discursifs, nous avons pu remarquer grâce à la prise en compte du temps que c'est sur « voilà » que se concentrent les différences significatives de la f_0 , alors que les caractéristiques acoustiques ressortent plutôt sur les premiers et les derniers phonèmes de l'énoncé. Notre étude a permis de démontrer le rôle de la prosodie affective dans la projection d'indices de valence. Ainsi, celle-ci pourrait être utilisée, de la même façon que les préfaces, pour projeter des informations affectives. La projection de telles informations devrait, si les indices sont suffisamment forts, permettre aux interlocuteurs récepteurs de la nouvelle de percevoir les indices, anticiper la valence affective de ce qui va être dit, et se préparer affectivement en conséquence. Cette fonction de projection affective de la prosodie peut être utilisée sur des énoncés, en parallèle d'autres fonctions inhérentes auxdits énoncés (comme la structuration de discours des DMs, et la projection verbale des préfaces, contenu verbal de l'annonce).

Malgré ce qu'il a permis de démontrer, ce travail présente de nombreuses limites qui viennent nuancer nos résultats. D'une part, et comme nous l'avons déjà mentionné, le corpus que nous avons enregistré est un corpus lu, et non naturel, qui a pu influencer la vitesse d'articulation, et qui ne permet peut-être pas d'observer tous les indices de projection prosodique que nous pourrions trouver en contexte naturel. Les conditions d'enregistrement ne nous ont par ailleurs pas permis de faire l'analyse de l'intensité. En effet, nos enregistrements ayant été faits pendant la crise sanitaire du COVID-19, nous avons finalement dû les réaliser avec un micro sur pied, dont on ne peut pas contrôler la distance au visage du locuteur. La variabilité verbale en fin de préface (« Je vous appelle au sujet de vos examens médicaux » vs « Je vous appelle au sujet du voyage scolaire ») avait été conservée afin de créer des annonces les plus naturelles possible et afin de

pouvoir varier les thématiques. Ce choix ne nous a pas permis de faire une comparaison complète des caractéristiques spectrales, que nous avons dû limiter au début de la préface « Je vous appelle au sujet », et aux marqueurs discursifs « voilà donc ». Cela crée un espace où nous ne pouvons pas observer les indices de projection au niveau des caractéristiques spectrales. C'est pourtant un emplacement crucial de la préface, car situé sur la fin, juste avant le pivot des marqueurs discursifs et le passage à la séquence d'annonce. Cette zone pourrait donc contenir elle aussi des variations spectrales importantes dues à la valence affective de la nouvelle. Dans de futurs travaux, il serait intéressant d'observer ce qui se passe après la séquence d'annonce faite dans l'appel, et de voir si l'on retrouve encore des caractéristiques acoustiques liées à la valence positive ou négative de la nouvelle, dans la séquence de clôture. Il serait aussi intéressant de comparer la séquence d'ouverture de l'appel téléphonique à la préface, afin de voir si les indices de projection sont présents dès le début de l'appel, et de voir à quel endroit, de l'ouverture ou de la préface, ils sont le plus forts.

Chapitre 3 – Etude phonologique de la parole en annonce de nouvelle

1. Problématique

Dans ce chapitre, nous présentons l'analyse intonationnelle pour le corpus acoustique décrit dans le Chapitre 2. Cela nous permet de répondre à l'une de nos questions de recherche : regarder si la valence affective est portée dans la préface par des indices de projection phonologiques, au travers de patrons intonationnels. Ainsi, nous souhaitons voir si le locuteur qui annonce une nouvelle projette des indices phonologiques de la valence affective. Si tel est le cas, nous souhaitons découvrir quels indices phonologiques entrent en jeu dans cette projection. Comme pour les analyses phonétiques, nous regarderons si les indices de projection affective sont présents dans un énoncé, en sus des autres sens portés par l'intonation. Si la prosodie affective portée par des indices phonétiques est un concept globalement admis par la communauté (Scherer, 2003), ce n'est pas le cas pour l'intonation, comme nous l'avons développé dans la section 1.2.5 de l'Etat de l'Art sur la dualité de structure. En effet, malgré les multiples sens que peut porter l'intonation, il a longtemps été considéré qu'elle ne porte que du sens linguistique, le sens paralinguistique étant laissé aux variations phonétiques de la prosodie. Pourtant depuis quelques années, des études sur le sens paralinguistique, et notamment affectif de l'intonation, voient le jour : nous décrivons trois études clés sur le sujet dans la section 1.3.3 de l'Etat de l'Art (Cao et al., 2014; Liscombe et al., 2003; Wang et al., 2020). Comme dans ces trois travaux, nous croyons en l'existence d'une « intonation affective » et par conséquent, à la possibilité de découvrir et de décrire des unités phonologiques intonationnelles de l'affect. Nous analyserons donc l'intonation, au travers de la préface et des marqueurs discursifs « voilà » et « donc ». L'analyse de l'annonce n'a pas été faite, car nous devons pouvoir comparer l'intonation au sein de productions similaires au niveau segmental, ce qui n'est pas le cas des annonces. Nous revenons sur ce point dans la section méthodologique de ce chapitre. Pour l'heure, nous présentons le raisonnement qui entoure ces analyses, les questionnements rattachés aux notions utilisées, et les hypothèses spécifiques pour cette étude. Pour cela, nous nous attarderons sur deux éléments essentiels à nos analyses et particulièrement, à nos hypothèses : l'identité et la place des configurations de fin d'IP (accent mélodique + ton de frontière) en français, ainsi

que l'intonation utilisée sur les marqueurs discursifs « voilà » et « donc » selon leur fonction de structuration de l'énoncé.

1.1. Les configurations de fin d'IP en français

Nous avons fait la description de la phonologie du français dans l'Etat de l'Art général de ce travail. Nous y avons établi que trois types de frontières prosodiques existent dans cette langue : le syntagme accentuel « AP », le syntagme intermédiaire « ip » (D'Imperio & Michelas, 2014; Jun & Fougeron, 2000) et le syntagme intonational « IP ». A la fin d'une IP, se trouve ce que l'on appelle la configuration de fin d'IP, suivant les termes « *intonation phrase final* » de Post (2000). Elle est constituée de l'accent mélodique (T*) de la dernière AP et du ton de frontière d'IP (T%).

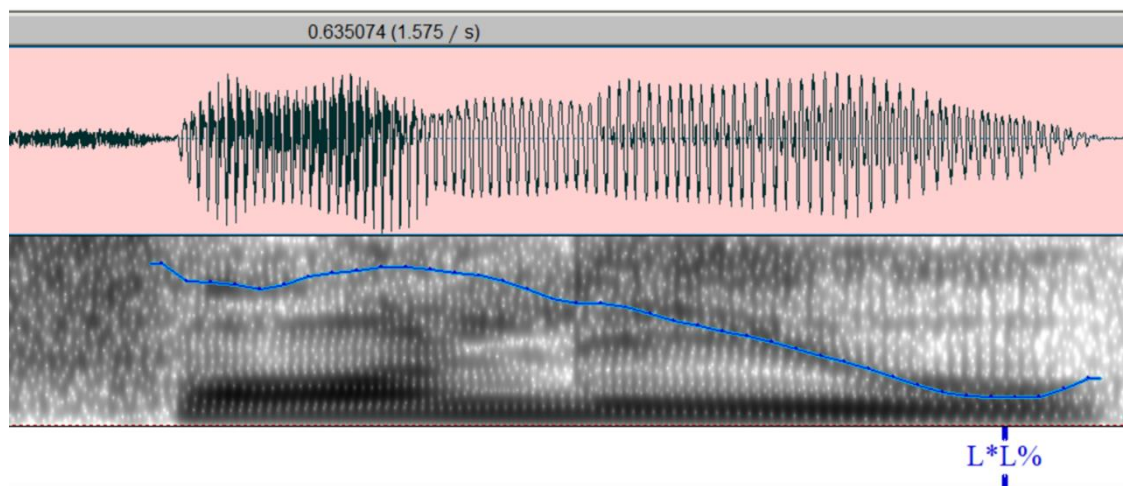


Figure 59 – Exemple de configuration de fin d'IP

C'est cette partie du patron prosodique que nous avons choisi d'étudier dans la préface, à la fin de la préface. Les configurations de fin d'IP en français correspondent en quelque sorte aux « configurations nucléaires » que l'on retrouve dans d'autres langues. Le terme « configuration nucléaire » ne fait cependant pas l'unanimité dans le cadre de l'étude de la phonologie intonative du français, où la notion « d'accent nucléaire » n'est pas toujours considérée comme pertinente. En effet, l'accent dit « nucléaire », qui correspondrait à celui de fin d'IP, est traditionnellement considéré comme similaire aux accents mélodiques produits dans les AP qui ne sont pas finales (Post, 2000). Cependant, certaines études utilisent le terme (Di Cristo, 1999), et d'autres argumentent en faveur de

l'inclusion de cette différenciation sur la base de différences significatives d'alignement et entre accents mélodiques en position de fin d'IP, et ceux qui viennent dans les AP précédentes (D'Imperio et al., 2006). Notre travail ne permettant pas une différenciation entre ces deux manières de nommer la configuration, nous utiliserons le terme « configuration de fin d'IP », qui est plus général.

Les configurations de fin d'IP sont donc situées à la fin de la dernière AP d'une IP, et sont marquées par un ton de frontière d'IP (T%). Si nous trouvons une frontière d'ip à la fin de nos préfaces, la configuration n'est alors pas tout à fait la même. Cependant de tels cas devraient être rares, la syntaxe des préfaces nous poussant à attendre une frontière d'IP entre préface et marqueurs discursifs. L'intérêt de l'étude de la configuration de fin d'IP réside dans l'importance qu'elle a au sein de la phrase prosodique, car elle contient l'accent mélodique et le ton de frontière d'IP. De plus l'accent mélodique, lorsqu'il est dans cette position, joue un rôle important dans le port du sens intonatif. Cependant les études se sont attardées sur le sens linguistique, (Delais-Roussarie et al., 2015), et nous voulons voir si c'est le cas aussi pour le sens paralinguistique. Les accents de fin d'AP qui ne sont pas en fin d'IP, en revanche, sont vus comme moins saillants dans d'autres langues. On considère souvent que les variations perçues sur ces accents « pré-nucléaires » sont simplement des conséquences de variations de l'accent nucléaire, ce qui justifierait de s'intéresser plutôt aux configurations pré-nucléaires même lorsque ces variations apparaissent : voir Calhoun (2010), Ladd (2008), Grabe et al. (2005), et Welby (2003), cités dans Petrone & Niebuhr (2014).

Nous suivons pour l'analyse des configurations de fin d'IP la description proposée par Delais-Roussarie et al. (2015) pour l'intonation du français. La technique utilisée pour obtenir leur corpus dans cette étude est une « *Discourse Completion Task* » (Prieto & Frota, 2015). Les DCT permettent de proposer aux participants un questionnaire visant à éliciter plusieurs types de phrases spécifiques. Nous prévoyons que nous pourrions trouver six possibilités en fin d'IP, correspondant aux six configurations décrites pour le français : L*L%, L*H%, H*H%, H*L%, H+H*L% et H+! H*!H%.

Cela signifie qu'il est possible d'annoter trois tons de frontière différents en français : L%, H%, et !H%. De même, il existe quatre accents mélodiques dont deux sont bitonaux : L*, H*, H+H*, H+!H*. Nous pourrions cependant trouver des préfaces qui ne se terminent pas par une frontière d'IP, mais plutôt par une frontière d'ip. Nous faisons l'hypothèse que

nous obtiendrions alors des modèles de fin de syntagme similaires à ceux des IP, avec T- et non T%.

Les trois configurations les plus régulièrement utilisées en français de façon générale sont : L*L%, H*H% et H*L% (Delais-Roussarie et al., 2015). Ce sont donc celles que nous avons le plus de chances de retrouver au sein de notre corpus. Le patron le plus fréquent pour les affirmations est L*L%.

Bien que la description du français proposée dans le chapitre de Delais-Roussarie et al. (2015) soit concentrée sur la description des fonctions linguistiques des patrons et configurations de fin d'IP, le sens paralinguistique est mentionné. Dans cette description en effet, les accents mélodiques bitonaux H+!H* et H+H* sont considérés comme pouvant être utilisés pour porter un sens paralinguistique attitudinal (la contradiction pour l'un, et l'évidence pour l'autre). De même, il est mentionné que les questions ouvertes (*wh-questions*) peuvent porter un sens attitudinal par rapport à ce que l'interlocuteur dit, en fonction des tons produits. Notamment, si la question est portée par la configuration de fin d'IP LH*L%, elle pourrait véhiculer une attitude de déception vis-à-vis de l'interlocuteur. Si les tons et configurations de fin d'IP peuvent véhiculer un sens attitudinal, et donc paralinguistique, nous pensons qu'il devrait alors être possible que certaines configurations puissent aussi porter un sens affectif plus général que l'attitude spécifique (valence positive, valence négative). De plus, un corpus tel que celui de Delais-Roussarie et al. (2015), aura moins de chance d'éliciter des productions affectives (puisque ce n'est pas son but), par rapport à notre corpus. Il est donc possible que nous puissions découvrir des configurations différentes, ou bien des variantes phonétiques de ces configurations, qui soient dépendantes du type d'affect porté par les énoncés de notre corpus.

Par ailleurs, nous avons décrit dans l'état de l'art de ce document des études sur l'anglais américain, qui toutes se sont intéressées au sens affectif qui peut être porté par l'intonation. Cao et al. (2014) se sont, comme nous, intéressés aux accents mélodiques et aux tons de frontières d'IP. Ils ont trouvé que certains tons de frontière notamment, étaient plus présents selon l'affect produit : H-L% par exemple, était plus souvent produit lorsque l'affect était la peur ou la tristesse. Bien qu'aucune hypothèse ne soit formulée sur les raisons de cette découverte dans l'article, nous pensons que cette configuration pourrait être fréquente dans la tristesse car elle se finit sur un ton de frontière L%, congruente avec le fait que la tristesse soit généralement réalisée avec un empan vocal

plus bas et des valeurs minimum, maximum et moyennes de la f0 abaissées (Berry & Brown, 2019; Liscombe et al., 2003; Scherer, 2003). Wang et al. (2020) quant à eux se sont concentrés sur les accents mélodiques et les tons de phrase T- des syntagmes intermédiaires. Ils ont, par exemple, trouvé que les accents mélodiques !H* et H* étaient des paramètres phonologiques corrélés de façon générale avec la valence affective, plutôt qu'avec l'excitation. C'est-à-dire que ce paramètre varie en fonction de la valence, et que l'on peut s'attendre le retrouver comme plus présent à certains degrés de la valence (par exemple, positive) et moins présent ailleurs (par exemple, négative). Ainsi, la configuration de fin d'IP semble pouvoir porter des variations de sens affectif en anglais américain. Nous souhaitons observer si cela est aussi possible en français.

1.2. Description phonologique de « voilà » et « donc »

Dans les marqueurs discursifs « voilà » et « donc », nous avons choisi d'annoter les patrons prosodiques dans l'ensemble de l'énoncé. Nous avons pu faire ce choix car les énoncés, courts et identiques, nous permettent une bonne comparaison de l'intonation entre les unités. Cela nous permet par ailleurs de comparer les patrons intonatifs aux contours prosodiques décrits dans la littérature présentée ci-dessous. Les énoncés pourraient, de plus, contenir des indices importants en raison de la proximité des marqueurs avec la préface et l'annonce, et leur position stratégique entre les deux.

Nous avons établi dans l'état de l'art de ce document la notion de marqueurs discursifs et leur utilisation dans l'interaction. Ils sont avant tout, comme leur nom l'indique, faits pour marquer et donc structurer le discours. Cependant, nous avons également noté des cas dans lesquels ils peuvent être utilisés en tant que préface. Nous avons donc pris le parti de les étudier, là encore, séparément de la préface en tant que telle. Notre hypothèse générale pour cette partie de l'annonce est que, malgré leurs fonctions, les marqueurs discursifs « voilà » et « donc » pourraient eux aussi permettre la projection d'indices de la valence affective à travers des variations phonologiques des patrons intonatifs qui les composent.

Pour pouvoir étudier cette potentielle projection, nous devons commencer par nous intéresser aux études existantes sur l'intonation dans les marqueurs discursifs, et aux liens qui ont été proposés entre les fonctions discursives de « voilà » / « donc » et certains

patrons intonatifs. Dans une étude sur corpus, Petit (2020) décrit les caractéristiques prosodiques de « voilà », dans le contexte de discours médiatiques en français standard. Le corpus est constitué d'une centaine d'enregistrements de plusieurs émissions télévisées. Petit (2020) retient pour son étude plusieurs sens possibles à « voilà » : validation, conclusion, validation-conclusion, « combler le discours », présentatif, introductif, ou encore, il peut être associé à une complétive (« voilà que »). L'emploi qui retiendra le plus notre attention ici est le « voilà » conclusif, car c'est un emploi très probable dans notre corpus (pour conclure la préface avant de passer à la séquence d'annonce). Il peut être produit en simple conclusion : « Attendons de voir quelle va être leur progression c'est ça qui nous intéresse *voilà* » (Petit, 2020, p. 158), ou il peut être produit pour conclure et poursuivre le discours :

« Il n'y a qu'une seule ligne au Front National c'est celle qui nous porte à 30% d'intentions de vote aujourd'hui à la présidentielle celle qui nous a mis en tête aux Européennes celle qui fait que nous n'avons jamais eu autant d'adhérents *voilà* donc ceux qui s'exprimeraient comme ça de manière euh anonyme pour remettre en cause cette ligne sont juste stupides. »

Exemple de « voilà » utilisé pour sa fonction conclusive, extrait de Petit (2020, p. 158)

Dans cet exemple, nous remarquons que « voilà » est utilisé conjointement avec le marqueur « donc ». Bien que celui-ci ne soit pas discuté dans le texte, cela nous permet de vérifier hors de notre corpus la possibilité d'utiliser « voilà » et « donc » à la suite l'un de l'autre. Lorsque « voilà » est employé de façon à conclure un discours, Petit (2020) distingue trois tendances à l'analyse des patrons intonatifs avec lesquels il est produit, que nous représentons dans la Figure 60 avec une schématisation des contours, et une proposition de correspondance ToBI. « Voilà » peut avoir une mélodie descendante, pour clore un sujet et ne pas revenir dessus. Il peut également avoir une mélodie montante lorsque l'on cherche à clore un sujet et convaincre autrui en même temps. Enfin, il peut avoir une mélodie en « forme de cloche » pour les cas où l'on veut clore un sujet en insistant ou en soulignant un élément de ce qui a été dit comme une évidence. Il est noté que l'implication affective du locuteur peut parfois éliciter la mélodie en forme de cloche, de même qu'un allongement de la durée de l'énoncé lorsque le locuteur est impliqué dans

ce qu'il dit (Petit, 2020). Cette dernière utilisation de « voilà » nous intéresse tout particulièrement au vu de son utilisation en contexte d'implication émotionnelle. En effet, on pourrait s'attendre à retrouver la forme de cloche dans une annonce de nouvelle, si le locuteur utilise « voilà » comme marqueur discursif et cherche à montrer son implication émotionnelle, ou est émotionnellement impliqué par la nouvelle qu'il cherche à annoncer. En revanche il est probable qu'il soit moins utilisé s'il n'y a pas d'implication émotionnelle, comme ce pourrait être le cas dans une nouvelle neutre. Suivant notre transcription des descriptions des patrons prosodiques, un « voilà » réalisé avec un contour de la f0 montant sera probablement transcrit par un patron LH* seul, ou bien LH*L-, ou LH*L% selon son placement dans l'IP. De même, un patron descendant sera transcrit HL*, HL*L- ou HL*L%. Un patron « en cloche » sera transcrit LHiL*, LHiL*L- ou LHiL*L%.

Après avoir décrit les fonctions que peut prendre « voilà », il est temps de nous intéresser à « donc ». Ferré (2011) décrit les corrélats acoustiques et prosodiques de trois marqueurs du discours (« donc », « alors », « en fait ») selon leurs fonctions. Similaires à celles de « voilà », elles peuvent être : conclusion, inférence, résomption, auto-correction, et ponctuation. A travers l'analyse d'extraits du corpus du CID (Bertrand et al., 2006), Ferré (2011) observe les marqueurs du discours utilisés par les interlocuteurs dans des diades (conversations à deux). L'analyse prosodique montre que lorsque le marqueur « donc » est utilisé pour conclure, il aura tendance à être produit avec un contour plat de la f0, dans la partie basse de l'empan vocal du locuteur (contour présenté dans la Figure 60). Il peut aussi se trouver à l'initiale de l'IP, être précédé d'une pause silencieuse, et être produit avec une voix craquée ou silencieuse. Lorsque le marqueur a fonction de reprise du sujet (« *topic resumption* »), il se trouvera plutôt en position médiane dans l'IP, et précédé d'une pause silencieuse s'il n'est pas précédé déjà d'un second marqueur discursif. Suivant notre transcription des descriptions des patrons prosodiques, un « donc » réalisé avec un contour de la f0 plat dans la partie basse de l'empan vocal sera probablement transcrit par un patron L* seul, ou bien L*L-, ou L*L% selon son placement dans l'IP.





	Fonction	Contour	Schématisation	ToBI
<i>Voilà</i> ¹	Conclusion de sujet	« tombant »		HiL*(L%)
	Conclusion forte	« en cloche »		LHiL*(L%)
	Persuasion	« montant »		LH*(H%)
<i>Donc</i> ²	Conclusion de sujet	« plat »		L*(L%)

Figure 60 – Schématisation et correspondance ToBI des contours phonologiques associés aux fonctions des marqueurs discursifs « voilà »¹ (Petit, 2020) et « donc »² (Ferre, 2011)

Une dernière étude (Haileselassie, 2015) a observé le comportement de « voilà » et « donc » lorsqu'ils sont utilisés côte à côte. Dans cette étude, il est démontré que « voilà » employé seul peut être utilisé pour faire la gestion de l'organisation d'une séquence d'actions : il permet de marquer la transition entre la fin de la première action, et le début de la seconde. Il pourrait ainsi être adopté pour marquer la transition entre la préface, et l'annonce de nouvelle. Lorsqu'il est utilisé à côté de « donc », « voilà » a toujours pour fonction de mettre en lumière la séquence qui vient, pendant que « donc » accomplit l'action de transition. C'est une fonction que nous pourrions retrouver dans notre corpus.

08 d'une vie personnelle pas facile, euh
09 → de moments pénibles, voilà, donc c'est comme
08 toutes choses de la vie, vous savez il y a

Exemple d'utilisation de « voilà donc » ensemble, extrait de Haileselassie (2015, p. 185)

Il semble ainsi que les deux marqueurs du discours « voilà » et « donc » puissent être utilisés de multiples manières grâce à leurs différentes fonctions. Il ressort des études présentées que des patrons prosodiques peuvent être attachés à certaines de leurs fonctions. On peut alors se demander si la projection prosodique de la valence d'une nouvelle à venir est possible avec l'intonation sur ces marqueurs discursifs, et dans quelle mesure elle serait présente. Il est possible que les fonctions de structuration du discours soient les seules à pouvoir faire varier les patrons intonatifs, et qu'aucune projection intonationnelle affective ne soit faite en plus de la projection structurelle.

1.3. Hypothèses

Au niveau des patrons phonologiques, notre étude est exploratoire pour les préfaces vis-à-vis du lien de la phonologie avec le paralinguistique. Nous souhaitons voir si nous observons des patrons phonologiques récurrents en fonction de la valence de la nouvelle annoncée, de même qu'il a pu être remarqué que certaines configurations revenaient selon le sens linguistique porté (type de phrase) et même l'attitude du locuteur (Delais-Roussarie et al., 2015). Comme dans les études présentées dans l'état de l'art général de ce document (Cao et al., 2014; Liscombe et al., 2003; Wang et al., 2020), les configurations de fin d'IP pourraient porter du sens paralinguistique affectif. Il est probable que les annonces de nouvelles téléphoniques fassent principalement produire aux locuteurs des affirmations à focalisation large, caractérisées par le patron L*L%. Nous faisons également l'hypothèse le Code de Fréquence (Gussenhoven, 2002; Ohala, 1983; Rodero, 2011), pourrait exercer une influence sur les tons réalisés en fin d'IP dans les préfaces. La hauteur de f0 élevée étant généralement perçue comme non-menaçante, nous pourrions retrouver plus de configurations H*H% dans les préfaces venant avant l'annonce d'une nouvelle positive ou neutre. A l'inverse la f0 abaissée étant généralement perçue comme menaçante, nous pourrions retrouver plus de configurations L*L% dans les préfaces venant avant l'annonce d'une nouvelle négative. Cela serait aussi cohérent avec les résultats de notre étude phonétique, car nous avons trouvé que la f0 était abaissée lorsque la nouvelle était négative, et élevée lorsqu'elle était positive.

Pour les marqueurs discursifs « voilà » et « donc », nous faisons l'hypothèse que les locuteurs devant annoncer une nouvelle vont souvent utiliser un patron « en cloche » (LHiL*) ou un patron « descendant » (HiL*) pour « voilà ». Le patron en cloche, généralement employé pour conclure un sujet en insistant sur la conclusion (Petit, 2020), pourrait être utilisé plus souvent pour les mauvaises nouvelles, afin de marquer plus fortement la transition et préparer l'interlocuteur à l'arrivée de la nouvelle en insistant sur le passage entre préface et nouvelle. Le patron montant (LH*) généralement utilisé lorsque le locuteur a une volonté de convaincre son auditeur n'est pas attendu dans notre corpus, car ce n'est pas une fonction que l'on pourrait attendre dans une préface à une annonce de nouvelle. Pour « donc », nous faisons l'hypothèse que le patron « plat » (L*) pourrait être régulièrement utilisé pour marquer la fonction conclusive. Comme pour la préface, il est également possible que les patrons les plus fréquents lorsque l'on annonce une nouvelle négative se terminent sur des configurations de fin d'IP L*L%, et que lorsque

l'on annonce une nouvelle positive ou neutre, les patrons les plus fréquents se finissent avec H*H%.

Cependant, si la phonologie dans les préfaces et les marqueurs discursifs (« voilà », « donc ») n'est pas un véhicule de valence affective, nous faisons l'hypothèse que nous trouverons tout de même des différences dans la réalisation phonétique des patrons phonologiques, au niveau de la valeur mélodique et de l'alignement des cibles tonales. Pour la valeur mélodique, nous pensons que les valeurs suivront la tendance générale retrouvée dans les analyses phonétiques larges décrites en Chapitre 2. Ainsi lorsque la nouvelle est négative, nous faisons l'hypothèse que les tons hauts et bas seront réalisés en cibles tonales plus basses que lorsque la nouvelle est positive ou neutre, et qu'elles seront réalisées en cibles tonales plus hautes lorsque la nouvelle est positive que lorsqu'elle est négative. Ces analyses nous permettent de mieux voir la variation inhérente aux patrons intonatifs à des pics de valeur intonative situés précisément dans le temps. Cela nous permet de voir si les valeurs mélodiques sont congruentes avec les analyses plus globales de la f_0 , ou si elles se comportent différemment. Par ailleurs, de telles analyses nous permettraient de comparer nos résultats avec ceux de Freese et Maynard (1998) sur les annonces de nouvelles. Ils avaient en effet trouvé qu'elles étaient produites avec un ton descendant en cas de mauvaise nouvelle, ce qui équivaldrait dans notre cadre théorique à la réalisation de cibles basses avec une valeur mélodique abaissée pour les nouvelles négatives, ou bien à une occurrence plus importante de tons bas.

Nous avons également choisi de regarder l'alignement des cibles tonales. Il a été démontré que celles-ci pouvaient marquer des contrastes de sens linguistiques dans différentes langues et variétés : en italien (D'Imperio, 2000; D'Imperio & House, 1997; Gili Fivela, 2002; Petrone, 2008), en allemand (Niebuhr & Kohler, 2004), ou encore en anglais (Dilley & Heffner, 2013). Cette analyse du sens paralinguistique que l'alignement pourrait porter est exploratoire, car ce type d'analyses n'a à notre connaissance pas encore été réalisé pour la valence affective dans la littérature. Nous n'attendons donc pas de variation pour l'alignement en fonction de la valence.

2. Méthodologie

2.1. Annotation phonologique

Dans notre transcription des configurations de fin d'IP dans la préface, deux tons sont annotés : l'accent mélodique et le ton de frontière. Le ton qui réalise l'accent mélodique de fin d'AP peut être soit H*, soit L*, et est associé au bord droit du syntagme qui le contient (Jun & Fougeron, 2002). Pour les tons de frontière, si la frontière prosodique était forte, marquée par un allongement de la dernière voyelle, et parfois une pause suivant la frontière, alors nous annotions un ton T% de frontière d'IP. Si la frontière était marquée par un allongement moins fort, nous annotions T- pour la frontière d'ip (Jun & Fougeron, 2002). Dans notre transcription des marqueurs discursifs « voilà » et « donc », le marquage des tons bas différencie L1, le ton initial de la phrase, et L2 le ton qui précède H*. La cible haute Hi de la montée intonative LHi est généralement annotée dans les deux premières syllabes de l'AP, car le ton est associé à la frontière gauche de l'AP (Jun & Fougeron, 2000; Welby, 2003; Welby & Loevenbruck, 2006). La cible basse L1 de la montée LHi a été marquée à la frontière gauche et au bord gauche du premier mot de contenu de l'AP. Hi est annoté à la suite de L1, mais ne possède pas d'ancrage segmental (Welby, 2006) et son lien aux premières syllabes de l'AP n'est que faible (Jun & Fougeron, 2002). La cible basse L2 de l'accent de hauteur LH*, conformément à la phonologie du français, a été annotée soit sur la même syllabe que H*, soit sur la précédente. En effet, celle-ci ne possède pas non plus d'ancrage spécifique au niveau segmental (Welby & Loevenbruck, 2005, 2006). La cible H* a été annotée dans la dernière syllabe pleine de l'AP, associée au bord droit du syntagme (Jun & Fougeron, 2002). Enfin, les tons de frontière d'IP (par exemple H%) ou d'ip (par exemple, H-) sont annotés dans la dernière syllabe du syntagme. Leur présence est marquée par un allongement de ladite syllabe (plus fort pour l'IP), et par la présence optionnelle d'une pause après les tons de frontière d'IP (Jun & Fougeron, 2002).

L'annotation phonologique des fichiers a été réalisée manuellement sous Praat (Boersma & Weenink, 2023). Nous avons réutilisé, pour cette annotation, les fichiers TextGrids transcrits pour les analyses phonétiques décrites dans le Chapitre 2. Nous avons conservé deux tires de ces transcriptions : SyllAlign, qui contient la transcription

alignée des syllabes, et TokenAlign, qui contient la transcription alignée des mots. Ces deux tires nous étaient utiles pour repérer visuellement les sections à annoter. Nous avons ajouté trois tires en plus de ces deux-là, pour l'annotation de l'intonation. Deux tires étaient dédiées à l'intonation et les phrases phonologiques : la tire « tones » pour l'annotation des tons, et la tire « boundaries » pour l'annotation des frontières prosodiques. Les frontières prosodiques ont été annotées aux trois niveaux de phrase intonactionnelle décrits pour l'intonation du français : IP, ip, AP. A cela, nous avons ajouté une tire « miscellaneous » (abrégé « misc »), afin de pouvoir noter nos doutes et réflexions (par exemple, si nous n'étions pas sûrs de l'annotation d'un ton comme L ou H, ou si la présence d'une frontière prosodique était difficile à déterminer).

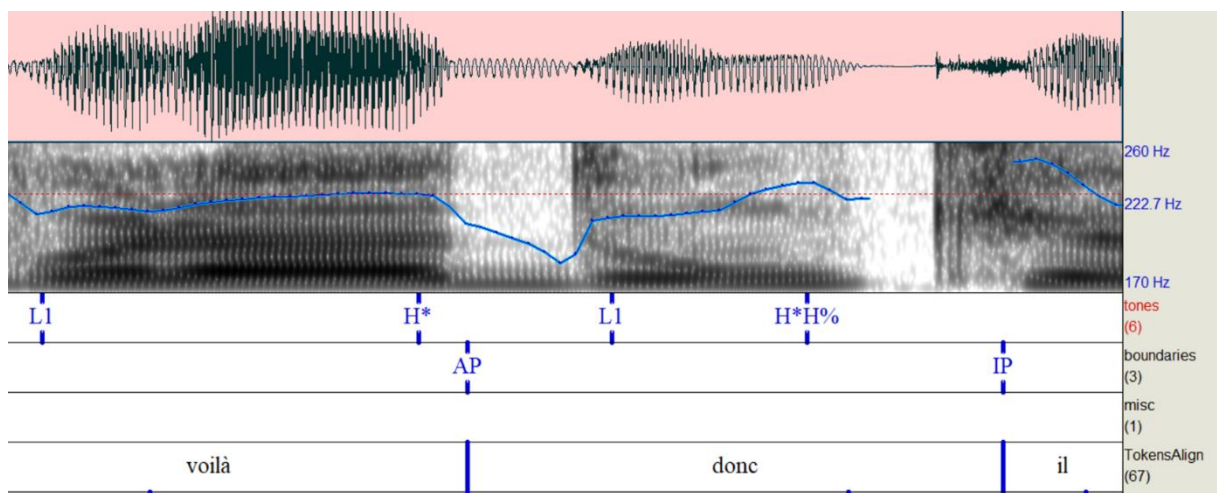


Figure 61 – Fichiers audio et TextGrid alignés, avec chaque tire utilisée pour l'annotation : *tones* (avec par exemple le ton L marqué L1), *boundaries* (où / marque le début de l'unité, AP la fin du syntagme accentuel, et IP la fin du syntagme intonactionnel, *misc*, *SyllAlign*, *TokenAlign*)

2.2. Accord inter-annotateurs

Pour vérifier la fiabilité de nos annotations, nous avons fait un accord inter-annotateurs dans les préfaces, ainsi que dans les marqueurs discursifs « voilà » et « donc ». L'autrice de ce document (annotateur 1), a annoté toutes les préfaces et tous les marqueurs discursifs. Ensuite, une seconde personne (annotateur 2) a annoté 20% des données de la préface, et 20% des données des marqueurs discursifs de façon indépendante. La sélection des items annotés par l'annotateur 2 a été faite de façon aléatoire. L'annotateur 1 et l'annotateur 2 ont eu un rendez-vous avant l'annotation des

20% du corpus, afin de s'assurer de suivre la même méthode d'annotation et de l'utilisation de labels identiques. L'annotateur 2 est linguiste, avec une expertise en prosodie et en annotation prosodique utilisant la convention d'annotation F-ToBI.

Nous avons calculé l'accord inter-annotateur à l'aide des tests des coefficients d'accord de Gwet (2001). Nous avons choisi ce test au lieu de tests plus communément utilisés dans la littérature (comme le kappa de Cohen, le kappa de Fleiss ou l'alpha de Krippendorph) car il est plus robuste à la prévalence des traits, et a déjà été utilisé en linguistique (Rohrer et al., 2022). La « prévalence des traits » se produit lorsqu'il y a une distribution asymétrique des catégories analysées (Wongpakaran et al., 2013), c'est-à-dire lorsqu'une catégorie est représentée plus fréquemment qu'une autre dans un corpus. Par exemple, si les évaluateurs annotant un corpus sur le sourire utilisent la catégorie « souriant » 70 % du temps et la catégorie « froncer les sourcils » 30 % du temps, il y a prévalence du trait « souriant ». La formule du kappa de Cohen (corrigée ou non) permet un calcul prenant en compte le risque qu'il y ait des accords de hasard entre les deux annotateurs (« *chance agreement* »). C'est-à-dire que ce sont des cas où les annotateurs sont d'accord, mais que cet accord est dû uniquement à la chance. Cependant dans le cas d'asymétrie des données que nous avons décrites, le kappa de Cohen, le kappa de Fleiss ou l'alpha de Krippendorph peuvent être artificiellement inférieurs ou supérieurs à ce qu'ils devraient en surestimant ou sous-estimant les accords de hasards. La cause de ce problème tient à deux paradoxes nommés « paradoxes de Kappa » par Feinstein et Cicchetti (1990). Le premier paradoxe intervient lorsque nous avons une prévalence d'un trait, que l'on peut vérifier par la présence d'un déséquilibre fort dans les totaux marginaux (compte total d'annotations pour un item dans les données) du tableau d'annotation horizontalement ou verticalement, selon si l'un des annotateurs a très peu utilisé une catégorie d'annotation par rapport aux autres catégories possibles (Wongpakaran et al., 2013). Dans un tel cas, le kappa de Cohen risque d'abaisser artificiellement le coefficient d'accord. Le second paradoxe intervient lorsque le déséquilibre est asymétrique dans ces totaux marginaux, c'est-à-dire lorsqu'une catégorie A est moins utilisée par un annotateur, et qu'une catégorie B, différente, est moins utilisée par l'autre annotateur. Dans un tel cas, le kappa de Cohen risque d'augmenter artificiellement le coefficient d'accord.

Comme nos données présentent un tel déséquilibre, nous avons écarté le kappa et l'alpha, sujet aux mêmes paradoxes. En effet dans nos données, certains traits sont plus prédominants que d'autres. Elles entrent dans la première catégorie de paradoxes : plus

de cas de L* que de H*, plus de cas de présence d'une frontière entre « *voilà* » et « *donc* » que de cas sans frontière. Cela rend les coefficients de Gwet plus appropriés pour vérifier l'accord inter-annotateur dans nos données, car les deux s'ajustent pour prendre en compte l'accord de hasard en prenant en compte des poids différents pour les désaccords, basés sur un calcul de la gravité du désaccord (Blood & Spratt, 2007). Dans la formule ci-dessous, qui est la formule générale des coefficients de Gwet, p_a représente la probabilité d'accord globale (par chance, ou non), et p_e est la probabilité que l'accord soit par chance. La manière dont ces deux probabilités sont calculées peut différer, selon le calcul choisi : coefficient d'accord de premier ordre AC1, et coefficient d'accord de second ordre AC2.

$$AC1 = \frac{p_a - p_e}{1 - p_e}$$

Pour Wongpakaran et al. (2013) le premier coefficient de Gwet, appelé AC1, possède un coefficient de fiabilité inter-évaluateurs stable car il est moins affecté par la prévalence et les probabilités marginales que le Kappa de Cohen. AC1 est adapté pour les données catégorielles non ordonnées. Il est défini comme la probabilité conditionnelle que deux évaluateurs choisis au hasard soient d'accord, étant donné qu'il n'y a pas d'accord de hasard (Gwet, 2001). A l'inverse, le second coefficient AC2 est adapté pour les données catégorielles ordonnées, car il prend en compte la possibilité de mauvaises classifications : c'est-à-dire, lorsque les annotateurs voient un item de la même façon, mais le classifient différemment.

Pour évaluer l'accord inter-annotateur, nous avons suivi les étapes proposées par Breen et al. (2012), afin de sélectionner les différentes catégories d'annotation à comparer. Nous avons comparé dans les préfaces les annotations pour le type d'accent mélodique (H* ou L*), et avons ajouté une comparaison pour le type de frontière prosodique (H%, L%, H-, L-). La première comparaison est réalisée avec AC1, la seconde avec AC2. Au sein des marqueurs discursifs, nous avons comparé, suivant le même article, le type d'accent mélodique (H* ou L*) avec AC1, la présence d'une frontière prosodique (présence ou absence) avec AC1, et la force de la frontière prosodique (IP, ip, AP ou absence) avec AC2. Nous avons pour cela utilisé le package R *irrCAC* (Gwet, 2019). Pour analyser les résultats, nous avons suivi la catégorisation d'Altman (1990) rapportée dans Bouvier et al. (2021) : $AC1 \leq 0,20$ = mauvais accord ; $0,21 \leq AC1 \leq 0,40$ = faible accord ; $0,41 \leq AC1 \leq 0,60$ = accord modéré ; $0,61 \leq AC1 \leq 0,80$ = bon accord ; et $AC1 \geq 0,81$ = très bon accord.

2.3. Critères pour l’alignement des cibles tonales

2.3.1. Critères pour la préface

Les tons étaient annotés sur la tire de points, ce qui nous permettait de marquer l’emplacement exact de la cible, et donc de prendre en compte l’alignement. Nous avons suivi plusieurs critères pour l’alignement tonal. Il a été réalisé entièrement à la main sous Praat. Après une première écoute de l’audio, qui nous permettait de déterminer quelles cibles tonales étaient produites, nous écoutions le mot et la syllabe sur lesquels le ton était produit plusieurs fois afin de mieux repérer son ancrage temporel au sein de la syllabe, à l’oreille mais aussi visuellement. Sur le plan visuel, nous affichions la courbe de la f_0 sur le spectrogramme correspondant à l’audio. Suivant les méthodes décrites dans la littérature sur l’italien ((D’Imperio, 2002; voir Petrone, 2008 et références dedans; Prieto et al., 1995) et sur le français (Kaminskaia, 2015; Welby & Loevenbruck, 2006) nous avons pris en compte les règles d’association des tons du français que nous avons décrit en section 2.1 pour faire cette première recherche visuelle, et cherchions donc en priorité la cible tonale à cet emplacement. Si le ton était aligné dans la syllabe, ou à la frontière du syntagme auquel il était associé, nous placions le point à la hauteur maximale (H^* , H^- , $H\%$) ou minimale (L^* , L^- , $L\%$) de la courbe. Nous avons pris en compte les possibles perturbations de la f_0 , lorsque le locuteur avait produit une voix craquée, ou lorsque le point maximal ou minimal était placé au sein d’une consonne voisée. Dans ces cas-là, la cible tonale a été annotée au point le plus proche du minimum ou du maximum. De façon arbitraire mais systématique, nous avons dans ces cas-là annoté la cible à gauche de la zone de perturbation. Les tons placés sur un plateau de f_0 étaient annotés sur le point initial du plateau (Welby, 2006; Welby & Loevenbruck, 2006). Pour prendre en compte le fait que les accents mélodiques et les tons de frontière peuvent être réalisés en même temps dans certains cas où le ton de frontière « préempte » l’accent mélodique, c’est-à-dire qu’il prend sa place (Dohen et al., 2004), nous avons annoté ces occurrences sur un seul point dans la tire. La Figure 62 présente les cibles tonales L^* et $L\%$ lorsqu’elles sont réalisées au même point en fin de préface, à la fin de la descente intonative. La Figure 63 présente les cibles $L^*H\%$ lorsqu’ils sont réalisés sur la même syllabe, la dernière de l’IP. L^* est placé juste avant la montée intonative, et $H\%$ est placé au point maximum. La Figure

64 présente les cibles H* et L%, où H* est réalisé sur la dernière syllabe pleine de l'IP, au point maximum de la f0. L% en revanche, est réalisé à la frontière droite de l'IP, soit dans la syllabe qui suit, terminée par un schwa. Il est placé en fin de descente intonative, juste avant une perturbation de la f0.

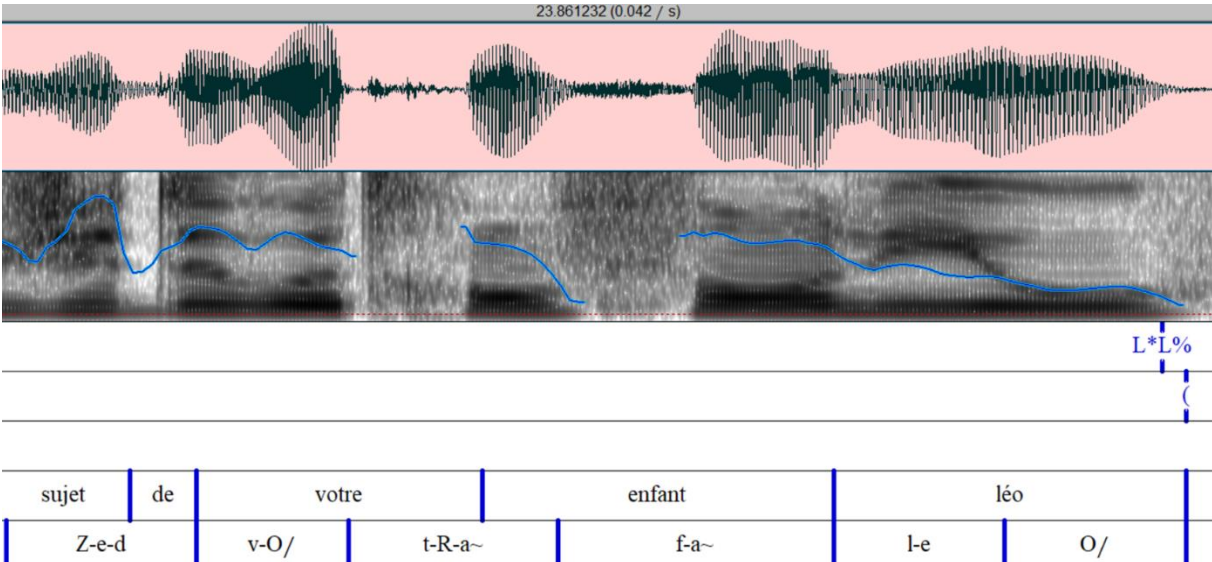


Figure 62 – Spectrogramme, courbe de f0 et TextGrid pour la préface *Au sujet de votre enfant Léo*, prononcée par la locutrice P04. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales L* et L% réalisés au même point

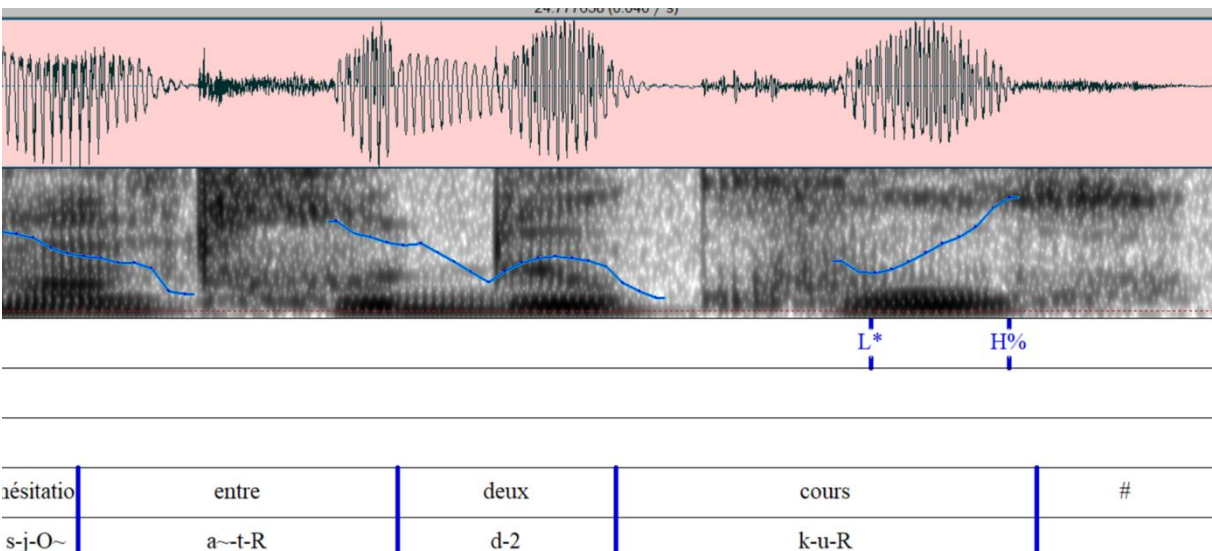


Figure 63 – Spectrogramme, courbe de f0 et TextGrid pour la préface *Au sujet de votre hésitation entre deux cours*, prononcée par la locutrice P03. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales L* et H%

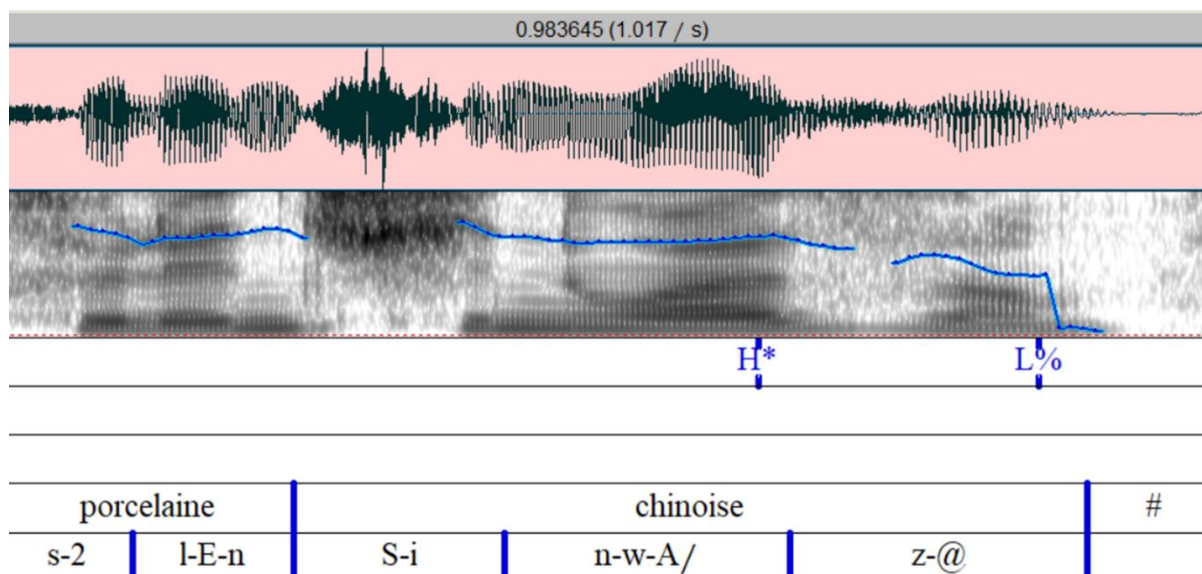


Figure 64 – Spectrogramme, courbe de f0 et TextGrid pour la préface *Au sujet de votre vase en porcelaine chinoise*, prononcée par la locutrice P02. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales H* et L%, avec une détection de la f0 perturbée en fin de préface

Pour calculer l’alignement et la valeur mélodique, les valeurs de la fréquence fondamentale aux points annotés ont ensuite été extraites à l’aide d’un script Praat personnel, réalisant les extractions avec la fonction de base du logiciel pour prendre la valeur de la f0 à un point donné. Le script extrayait aussi le marquage temporel de la cible tonale, et ceux du début et de la fin de la syllabe qui le contenait, et du mot qui le contenait. Nous avons ensuite effectué un calcul de la longueur des mots et des syllabes en secondes, et un calcul de la distance entre la fin du mot (ou de la syllabe) et la cible tonale en secondes. Enfin, nous avons calculé la différence en pourcentage entre l’emplacement de la cible et la fin du mot ou de la syllabe afin d’obtenir une normalisation de la distance (D’Imperio et al., 2014; Fivela & Savino, 2003; Kalaldehy et al., 2009).

2.3.2. Critères pour « voilà donc »

Nous avons suivi des critères similaires à ceux décrits pour la configuration de fin d’IP afin d’annoter les marqueurs discursifs. Après une première écoute de l’énoncé pour déterminer les cibles tonales produites, nous faisons des écoutes aidées du visuel, pour repérer l’ancrage temporel des cibles tonales. Nous avons pris en compte les règles d’association des tons du français que nous venons de décrire pour la première recherche visuelle. Les tons cibles Hi et L2 (le second ton L de l’AP)n’ayant pas d’association fixe (Jun & Fougeron, 2002), ils étaient cherchés dans les syllabes où ils sont communément

trouvés. Hi est généralement trouvé soit à la place de L1 (le premier ton L de l'AP) si celui-ci n'est pas réalisé, en frontière gauche de l'AP, soit dans l'une des deux premières syllabes de l'AP, pour nous /vwa/ et /la/. L2 est réalisé sur la dernière syllabe pleine de l'AP avec H*, ou sur la syllabe précédente, pour nous /dōk/ ou /la/. Comme pour la préface, si le ton était aligné dans la syllabe, ou au bord du syntagme auquel il était associé, nous placions le point à la hauteur maximale (H*, H-, H%, Hi) ou minimale (L*, L-, L%, L1, L2) de la courbe. Nous avons pris en compte les possibles perturbations de la f0, lorsque le locuteur produit une voix craquée, ou lorsque le point maximal ou minimal était placé au sein d'une consonne voisée. Par exemple, L1 étant réalisé dans la première syllabe de l'AP à la frontière gauche, nous devons prendre en compte les perturbations de la f0 présentes dans la première consonne voisée : /v/. Dans ces cas-là, la cible tonale a été annotée au point le plus proche du minimum ou du maximum, hors de la zone de perturbation (pour L1, généralement juste au début de la voyelle /w/. Comme dans la préface, les tons placés sur un plateau de f0 étaient annotés sur le point final du plateau, et lorsque les accents mélodiques et les tons de frontière étaient réalisés en même temps, nous avons annoté ces occurrences sur un seul point dans la tire. Nous présentons ci-dessous quatre exemples d'annotation phonologique des marqueurs discursifs, réalisés « voilà », « donc » ou « voilà donc ». La Figure 65 présente un exemple du patron sous-jacent du français lorsque les tons sont réalisés en surface, sur « voilà donc ». Nous voyons que L1 et Hi sont réalisés dans la même première syllabe de l'AP, L1 rattaché à la frontière gauche. L2 est réalisé dans la syllabe qui précède l'accent mélodique H*, annotée juste avant le début de la montée intonative. La Figure 66 présente un exemple du patron LH* lorsque « voilà » est réalisé seul. On voit que contrairement à la figure précédente, le /v/ est dévoisé et la cible est placée dès le début de la courbe de f0 sur la voyelle /w/. La Figure 67 montre un exemple du patron LHiL*, où L* est réalisé au même endroit que le ton de frontière L-. La Figure 68 montre le patron L* réalisé sur « donc », où L* est réalisé au même endroit que le ton de frontière L%.

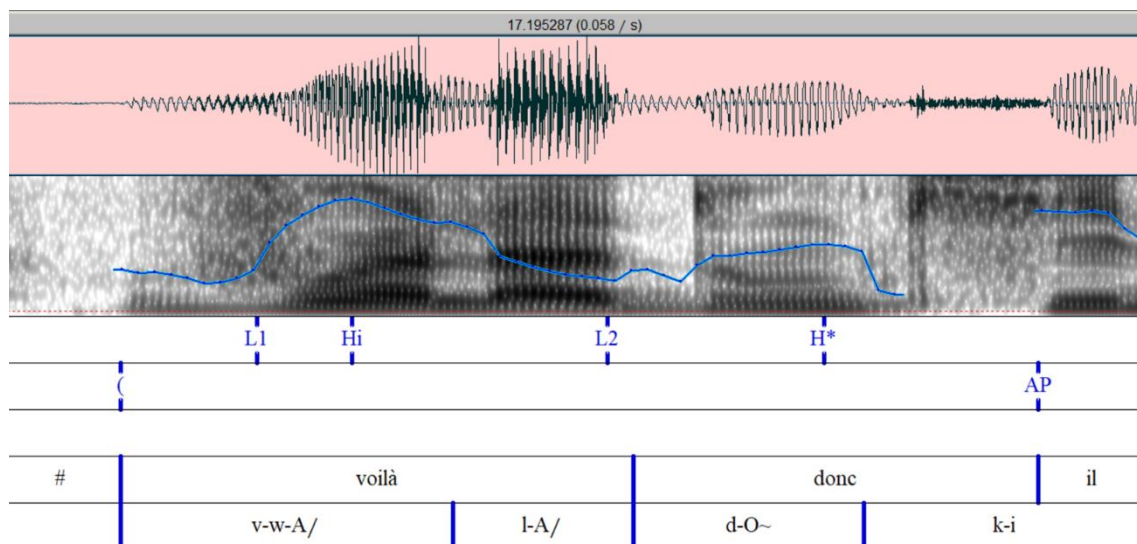


Figure 65 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « voilà donc », prononcé par la locutrice P02 (item 6). Dans le TextGrid, nous voyons un exemple d'annotation du patron sous-jacent du français, où toutes les cibles font surface : LHilH*. Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial, L2 marque le ton L de l'accent final

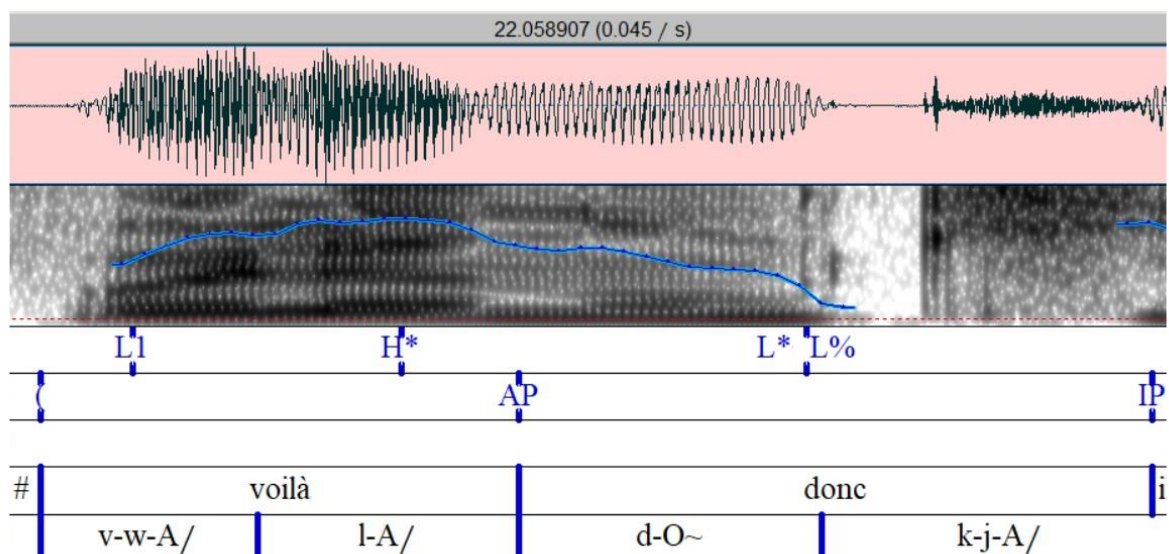


Figure 66 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « voilà donc », prononcée par la locutrice P06 (item 10). Dans le TextGrid, nous voyons un exemple d'annotation du patron LH*. Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial

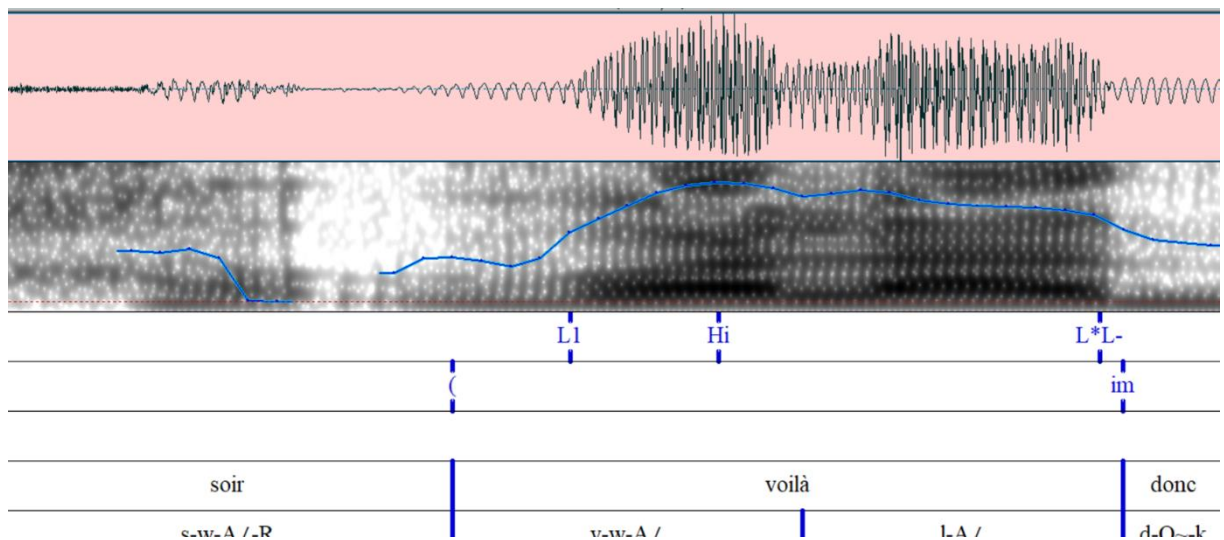


Figure 67 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « *voilà* », prononcée par la locutrice P07 (item 2). Dans le TextGrid, nous voyons un exemple d'annotation du patron LHL*, où L* et L- sont réalisés ensemble. Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial

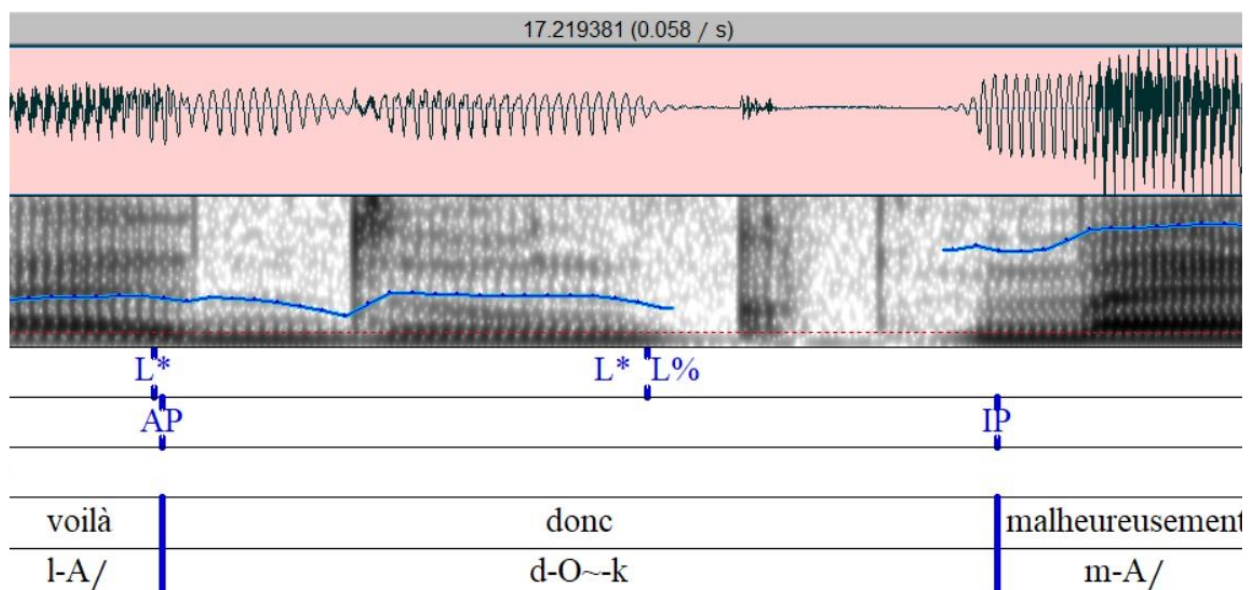


Figure 68 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « *donc* », prononcée par la locutrice P30 (item 9). Dans le TextGrid, nous voyons un exemple d'annotation du patron L*, où L* et L% sont réalisés ensemble

Pour calculer l’alignement et la valeur mélodique, les valeurs de la fréquence fondamentale aux points annotés ont ensuite été extraites à l’aide du même script Praat que celui utilisé dans la préface. L’alignement du ton L1 a été calculé par rapport à la première syllabe de l’énoncé lorsqu’il était réalisé. Le placement des tons Hi et L2 était observé par rapport à la première syllabe du début de l’unité prosodique pour Hi (la même que L1), et à la dernière syllabe de l’unité prosodique pour L2 (la même que L*).

2.4. Analyses statistiques

Pour les variables phonologiques, nous avons utilisé des modèles mixtes logit séparément sur chacun des patrons les plus fréquents : pour les préfaces, puis pour les marqueurs discursifs « voilà » et « donc ». L’utilisation d’un patron ou d’un ton dans les 1078 productions enregistrées dans le corpus est représentée par la variable « présence ou absence » dans chaque occurrence d’un énoncé. Les valeurs assignées sont « 1 » pour la présence et « 0 » pour l’absence d’un patron tonal spécifique. Les modèles logit sont basés sur les distributions binomiales (scores z, Generalized Linear Model, GLM). Il est possible de l’utiliser pour représenter des variables binomiales telles que celles que nous utilisons (1 et 0), notamment pour un événement intonational (R. H. Baayen et al., 2008). Nous avons testé l’effet de la VALENCE (positive, négative, neutre) sur chacune des variables (les patrons principaux produits, et la présence ou non d’une frontière prosodique entre « voilà » et « donc »). Notre modèle maximal comportait quatre éléments spécifiés aléatoires. Tout d’abord, nous avons inclus un intercepte aléatoire ITEM à douze niveaux (les 12 scénarios du corpus). Nous avons également inclus un intercepte aléatoire LOCUTEUR à trente niveaux (car 30 locutrices dans notre corpus). Enfin, nous avons ajouté deux pentes aléatoires : VALENCE par ITEM et VALENCE par LOCUTEUR. Les interceptes et pentes aléatoires ont ensuite été enlevées ou conservées dans le modèle, selon les résultats d’un test de vraisemblance réalisé avec la fonction *anova()* de R, qui permet de comparer les modèles emboîtés pour déterminer quel modèle est le plus ajusté aux données. Suivant Pinheiro & Bates (2000), nous avons choisi le modèle avec les plus petites valeurs BIC et AIC, qui représente le mieux nos données, en prenant également en compte la valeur p. Le modèle final inclut uniquement la pente aléatoire VALENCE par LOCUTEUR. Nous avons, pour chaque modèle, opéré un changement de l’intercepte sur les données avec la fonction *relevel*, avec dans un premier temps la condition « négatif »

comme intercepte, puis « positif » après *relevel*. Cela nous permet de comparer directement les nouvelles positives aux nouvelles neutres. Pour cette raison, nous avons corrigé le niveau alpha de la *p-value* de 0.05 à 0.025 (0.05/2 modèles exécutés par chaque variable dépendante).

Modèle GLMER utilisé pour les analyses des évènements intonationnels :

glmer(variable~Valence+(1+Valence|Locuteur))

Pour les variables phonétiques (la valeur mélodique des cibles tonales et l'alignement dans les patrons phonologiques), nous avons utilisé des modèles linéaires mixtes (Linear Mixed Effect Model, LMER, Pinheiro & Bates, 2000) dans les préfaces et dans les marqueurs discursifs car ces variables sont continues. Nous avons testé l'effet de la VALENCE (positive, négative, neutre) sur chacune des variables. Notre modèle maximal comportait, comme pour les modèles précédents, quatre éléments spécifiés aléatoires : un intercepte aléatoire ITEM, un intercepte aléatoire LOCUTEUR, et deux pentes aléatoires, VALENCE par ITEM et VALENCE par LOCUTEUR. Nous avons réalisé des ANOVA pour faire la comparaison des modèles. Le modèle final inclut LOCUTEUR et ITEM comme interceptes aléatoires, mais aucune pente. Là encore, nous avons réalisé un *relevel* afin de pouvoir comparer le positif et le neutre, mettant le niveau alpha, de nouveau, à 0.025.

Modèle LMER utilisé pour les analyses d'alignement et de valeur mélodique :

lmer(variable~Valence+ (1|Locuteur) + (1|Item))

3. Résultats

3.1. Accord inter-annotateur

3.1.1. Préface

Pour l'observation de l'accord inter-annotateurs au sein des préfaces, nous avons obtenu un coefficient d'accord inter-annotateur faible de 0.37 pour le type d'accent mélodique (sur lequel nous revenons dans la section discussion de ce chapitre), et un très bon coefficient d'accord de 0.90 pour le type de frontière prosodique. Nous présentons ci-dessous les matrices de confusion pour les deux types de tons observés dans ces énoncés. Ces matrices représentent le nombre de fois (en nombre d'occurrences) où un item a été rangé dans une catégorie ou une autre par les deux annotateurs. En tout, 180 items ont été annotés. Les annotateurs ont noté 180 accents mélodiques, et 180 tons de frontière. Les matrices permettent de visualiser les catégories qui présentent ou non des accords, et d'avoir une meilleure idée de la nature des désaccords. Par exemple pour l'accent mélodique qui a obtenu un mauvais score inter-annotateur, nous voyons que le désaccord est principalement situé sur le fait que l'annotateur 1 a souvent catégorisé des tons comme L*, que l'annotateur 2 a catégorisé comme H*. Pour les deux annotateurs, nous notons une forte représentation de L* par rapport à H* pour les accents mélodiques. Pour les tons de frontière, malgré la présence de quelques frontières d'ip (4), le type de frontière le plus représenté est celui des IP. Le ton L% est le plus souvent produit par rapport au ton H%.

		Annotateur 1	
		H*	L*
Annotateur 2	H*	31	58
	L*	4	87

Tableau 11 – Matrice de confusion : nombre d'occurrence du type d'accent mélodique dans la préface

		Annotateur 1			
		H%	L%	H-	L-
Annotateur 2	H%	47	2	0	0
	L%	9	97	0	0
	H-	7	1	2	0
	L-	0	11	1	2

Tableau 12 – Matrice de confusion : nombre d’occurrence du type de ton de frontière dans la préface

3.1.2. « voilà donc »

Pour l’observation de l’accord inter-annotateurs au sein des marqueurs discursifs, nous avons obtenu un coefficient d’accord inter-annotateur modéré de 0.50 pour le type d’accent mélodique, un bon coefficient d’accord de 0.63 pour la présence ou absence d’une frontière prosodique, et un bon coefficient d’accord de 0.75 pour la force de la frontière prosodique. Les matrices de confusion ci-dessous présentent à nouveau, par nombre d’occurrences, les catégories d’accents mélodiques, de présence ou absence d’une frontière, et de force de la frontière prosodique. Elles permettent de comparer le nombre de fois où un annotateur a rangé un item dans une catégorie ou une autre, et le nombre de fois où il n’y a pas eu d’accord sur la catégorie. En tout, 180 items audios ont été annotés par les deux annotateurs pour la mesure de l’accord. Les annotateurs ont noté 261 accents mélodiques, et 180 lieux où une frontière prosodique pouvait être présente ou non, et 182 types de tons de frontière (absence incluse).

		Annotateur 1	
		H*	L*
Annotateur 2	H*	52	29
	L*	42	136

Tableau 13 – Matrice de confusion : nombre d’occurrence du type de d’accent mélodique dans les marqueurs discursifs

		Annotateur 1	
		Pas de frontière	Frontière
Annotateur 2	Pas de frontière	16	24
	Frontière	20	120

Tableau 14 – Matrice de confusion : nombre d’occurrence en fonction force de la frontière prosodique dans les marqueurs discursifs

		Annotateur 1			
		IP	ip	AP	Absence
Annotateur 2	IP	19	4	0	1
	ip	3	22	7	0
	AP	5	19	43	22
	Absence	2	4	13	18

Tableau 15 – Matrice de confusion : nombre d’occurrence de la présence ou l’absence d’une frontière prosodique entre « voilà » et « donc »

3.2. Analyses des patrons phonologiques

3.2.1. Préface

Nous avons observé, pour la préface, les quatre configurations de fin d'IP les plus fréquentes sur la base de l'annotateur 1 : L*L%, L*H%, H*H%, H*L% (l'ensemble des configurations est proposé en Annexe 11). Elles représentent 96% du corpus. Les autres patrons étaient L*L-, L*H-, H*H-, H*L-. Les modèles glmer n'ont pas, de manière générale, montré d'effets de la valence sur l'intonation dans la préface, presque toutes les valeurs p étant au-dessous du niveau alpha de 0.025. La seule exception est la configuration L*L% qui présente une différence significative entre le négatif et le neutre, étant utilisée 16% du temps lorsque la valence de la nouvelle est négative, contre 13% (Tableau 15) lorsqu'elle est neutre [$\beta = 0.11$, SE = 0.04, $z = 2.4$, $p = 0.015$]. Les configurations de fin d'IP semblent donc être principalement produites indépendamment de la valence de la préface.

Patron	Négatif		Positif		Neutre		Total	
	%	nb	%	nb	%	nb	%	nb
L*L%	16%	177	14%	153	13%	137	43%	467
L*H%	10%	105	11%	116	13%	140	33%	361
H*L%	4%	40	4%	45	4%	42	12%	127
H*H%	2%	26	3%	32	2%	26	8%	84
autre	1%	12	1%	14	1%	15	4%	41
Total	33%	360	33%	360	33%	360	100%	1 080

Tableau 16 – Distribution des configurations de fin d'IP dans la préface, selon la valence de la nouvelle

3.2.2. « voilà donc »

Pour l'analyse des patrons phonologiques produits dans les marqueurs discursifs, nous avons procédé à une première observation. Nous avons regardé si les marqueurs étaient réalisés en une seule unité prosodique, ou en deux. C'est-à-dire [voilà donc]_{AP}, ou [voilà]_{AP}[donc]_{AP}. Nous avons aussi regardé s'il y a une influence de la valence de la nouvelle à annoncer sur cette différence de production. Nous voyons dans le Tableau 16 que la valence affective de la nouvelle ne semble pas influencer la production des

marqueurs discursifs en une, ou deux unités prosodiques, lorsque l'on compare le négatif au positif [$p = 0.6$], le négatif au neutre [$p = 0.7$], ou le positif au neutre [$p = 0.3$]. Il semble que la production des marqueurs en deux unités prosodiques soit le cas le plus fréquent, car il représente 69% du corpus dans son ensemble.

Nb. unité	Négatif	Positif	Neutre	Total
2 unités	70%	69%	68%	69%
1 unité	30%	31%	32%	31%

Tableau 17 – Choix d'une ou deux unités phonologiques dans les marqueurs discursifs

Au vu de cette distribution des patrons phonologiques sur les marqueurs discursifs, nous avons mené une analyse séparément sur chacune des possibilités : « voilà » réalisé seul sur une unité et « donc » réalisé seul sur une unité ([voilà]AP[donc]AP), ou « voilà donc » réalisés ensemble sur une unité ([voilà donc]AP). Dans le Tableau 18, nous présentons le pourcentage et le nombre d'apparitions des patrons phonologiques les plus récurrents selon la valence des nouvelles à venir (l'ensemble des patrons par marqueur est proposé en Annexe 13). Nous nous concentrons sur les 4 patrons les plus récurrents pour chacune des trois possibilités (présentées désormais comme : « voilà », « donc », « voilà donc »). Comme nous pouvons le voir dans le tableau ci-dessous, ils sont répartis relativement uniformément sur les trois valences. Les patrons sur « voilà donc » semblent contenir une très forte variabilité.

Marqueur discursif	Patron	Négatif		Positif		Neutre		Total	
		%	nb	%	nb	%	nb	%	nb
Voilà	<i>HiL*L%</i>	31%	78	26%	65	18%	44	25%	187
	<i>HiL*</i>	27%	67	15%	37	22%	55	21%	159
	<i>HiL*L-</i>	10%	26	8%	20	9%	22	9%	68
	<i>LH*</i>	8%	19	8%	21	10%	24	9%	64
	<i>Autre</i>	24%	61	43%	107	41%	101	36%	269
	<i>Total</i>	100%	251	100%	250	100%	246	100%	747
Donc	<i>L*</i>	20%	50	24%	61	28%	69	24%	180
	<i>L*L%</i>	18%	45	16%	39	19%	46	17%	130
	<i>L*L-</i>	16%	41	13%	33	17%	43	16%	117
	<i>H*H-</i>	11%	28	12%	30	12%	29	12%	87
	<i>Autre</i>	35%	87	35%	87	24%	59	31%	233
	<i>Total</i>	100%	251	100%	250	100%	246	100%	747
Voilà donc	<i>LLH*H%</i>	28%	31	21%	23	18%	21	23%	75
	<i>HiL*L%</i>	8%	9	7%	8	17%	19	11%	36
	<i>LHiLH*H%</i>	6%	6	15%	17	10%	11	10%	34
	<i>LHiL*L%</i>	2%	2	13%	14	11%	12	8%	28
	<i>Autre</i>	56%	61	44%	48	45%	51	48%	160
	<i>Total</i>	100%	109	100%	110	100%	114	100%	333

Tableau 18 – Distribution des patrons phonologiques dans les marqueurs discursifs, selon la valence de la nouvelle

L'analyse du marqueur « voilà » lorsqu'il est produit seul est la seule à présenter une différence significative. Cette exception est le patron *HiL*L%* produit sur « voilà », qui présente une différence significative entre le négatif et le neutre. Il est plus souvent produit lorsque la valence de la nouvelle est négative (31%) par rapport à lorsqu'elle est neutre (18%) [$\beta = 0.11$, $SE = 0.04$, $z = 2.51$, $p = 0.017$], mais pas par rapport à lorsqu'elle est positive [$p = 0.2$]. Nous n'avons pas non plus trouvé de différence significative entre la valence positive et la neutre [$p = 0.07$]. Pour les trois autres patrons principaux trouvés dans « voilà », ainsi que pour tous les patrons de « donc » et « voilà donc », la valeur p est toujours située au-dessus du niveau alpha de 0.025. Cela indique qu'il n'y a pas eu de patron significativement plus produit selon la valence de la nouvelle à annoncer dans ces cas-là.

3.2.3. Valeur mélodique et alignement

Nous allons maintenant présenter les résultats des analyses sur l'alignement et la valeur mélodique des cibles tonales dans les préfaces et les marqueurs discursifs. Pour la préface, nous commençons par présenter les résultats des modèles sur la valeur mélodique d'une cible, selon l'ordre suivant : L%, H%, L*, H*. Puis nous présentons les résultats d'alignement pour L* et H*. Nous n'avons pas réalisé d'analyses de hauteur sur les tons L- et H- produits au sein des préfaces, leur quantité étant insuffisante, avec seulement 15 occurrences pour L- et 25 occurrences de H-. Pour chaque ton étudié pour son alignement, nous observons l'alignement de la cible tonale en fonction de l'association du ton, au niveau de la fin de la dernière syllabe, ainsi que l'alignement au niveau de la fin du dernier mot.

Pour les marqueurs discursifs, nous avons réalisé une analyse des tons en trois parties séparées de la même manière que pour les analyses des patrons phonologiques. Ainsi, selon la manière dont ils ont été produits, les tons sont analysés au sein de « voilà donc » en un bloc ou séparément, soit : « voilà donc », « voilà », et « donc ». Nous n'avons pas réalisé d'analyses sur les tons L- et H- produits dans les marqueurs discursifs « voilà donc », leur quantité par valence étant insuffisante. Par exemple, nous observons seulement 17 occurrences de L-, et 10 occurrences de H- pour le neutre. Nous n'avons pas réalisé d'analyses sur le ton H- produit dans le marqueur discursif « voilà », sa quantité par valence étant aussi insuffisante pour les analyses statistiques (par exemple, seulement 14 occurrences pour le positif). Nous n'avons pas réalisé d'analyses sur les tons L1, Hi et L2 produits dans le marqueur discursif « donc », leur quantité étant insuffisante par valence (par exemple, seulement 4 occurrences de L1, 13 occurrences de Hi pour le neutre, aucune occurrence de L2).

Nous présenterons, pour chaque ton évalué, ses réalisations en cibles tonales dans les différents groupements de marqueurs discursifs en unités prosodiques. Dans les graphiques pour l'alignement, 100 correspond au début de la syllabe, 0 correspond à la fin, les valeurs négatives correspondent à ce qui est réalisé après la syllabe. Les graphiques des données non significatives ainsi que les résultats complets des modèles sont proposés en annexe pour la préface (Annexe 11, Annexe 12) et pour les marqueurs discursifs (Annexe 13, Annexe 14).

3.2.4. Préface

Valeur mélodique des tons

Dans un premier temps, nous présentons la valeur mélodique des tons de frontière L% et H%, ainsi que des accents mélodiques L* et H*. Conformément à nos attentes, nous trouvons des différences significatives de hauteur en fonction de la valence de la nouvelle. Le Tableau 19 ci-dessous présente les valeurs moyennes et déviations-standard, en Hertz, pour la hauteur des cibles tonales.

	Négatif		Positif		Neutre	
	Moy	SD	Moy	SD	Moy	SD
H%	240,2	47,4	276,9	72,8	276,6	74,5
L%	176,2	34,3	180,6	34,9	181,4	35,3
H*	255,6	42,8	250,8	41,7	263,4	48,5
L*	181,8	34,7	192,8	39,1	186,8	37,2

Tableau 19 – Valeurs moyennes et SD (Hertz) pour les cibles tonales évaluées au sein de la préface

Les graphiques (Figure 69, Figure 70, Figure 71) indiquent que les cibles tonales sont souvent réalisées plus bas lorsque la valence est négative, par rapport à lorsqu'elle est positive. Cette différence se retrouve sur L% [$\beta = -5.77$, SE = 2.13, $t = -2.7$, $p < 0.01$], sur H% [$\beta = -34.79$, SE = 6.39, $t = -5.43$, $p < 0.0001$] et sur L* [$\beta = -8.38$, SE = 1.9, $t = -4.33$, $p < 0.0001$], mais pas sur H* [$p = 0.5$].

Les Figure 70, Figure 71, et Figure 72 indiquent également des cibles tonales réalisées significativement plus bas lorsque la valence est négative, par rapport à lorsqu'elle est neutre. Cette différence se retrouve sur H% [$\beta = -38.14$, SE = 6.25, $t = -6.09$, $p < 0.0001$], sur L* [$\beta = -4.67$, SE = 1.9, $t = -2.42$, $p < 0.025$], et sur H* [$\beta = -17.13$, SE = 5.87, $t = -2.91$, $p < 0.01$], mais pas sur L%* [$p = 0.1$].

Aucune différence significative n'a été trouvée lorsque l'on compare la valence positive à la valence neutre, quelle que soit la cible tonale : L% [$p = 0.2$], H% [$p = 0.5$], L* [$p = 0.5$], ou H* [$p = 0.2$].

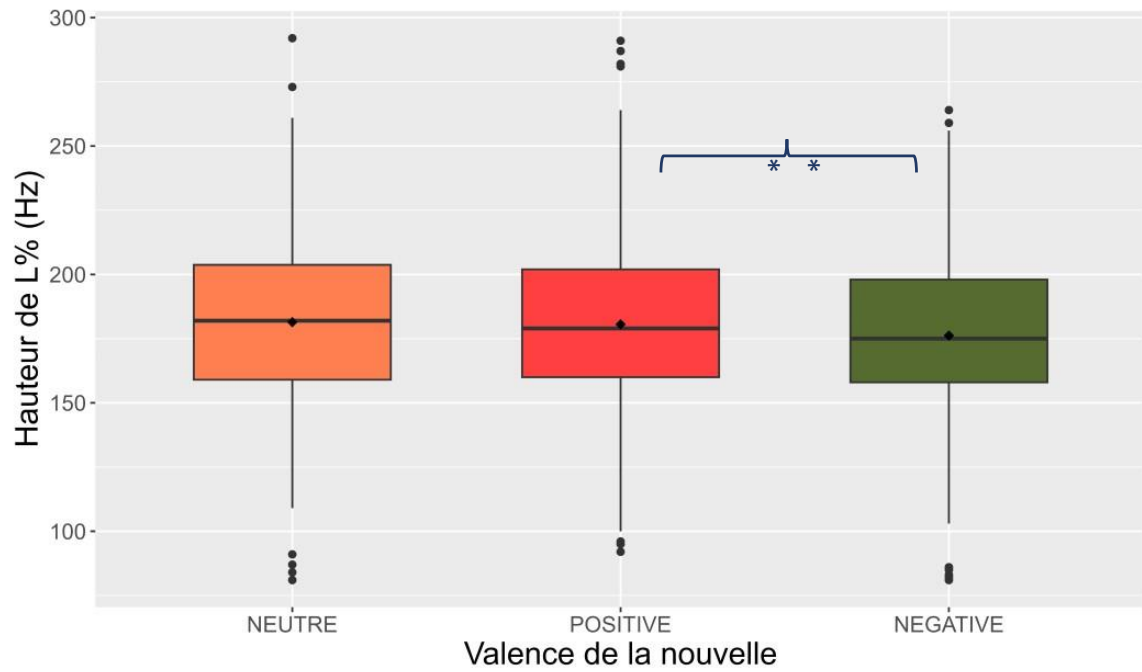


Figure 69 – Valeur mélodique des tons L% en fin de préface

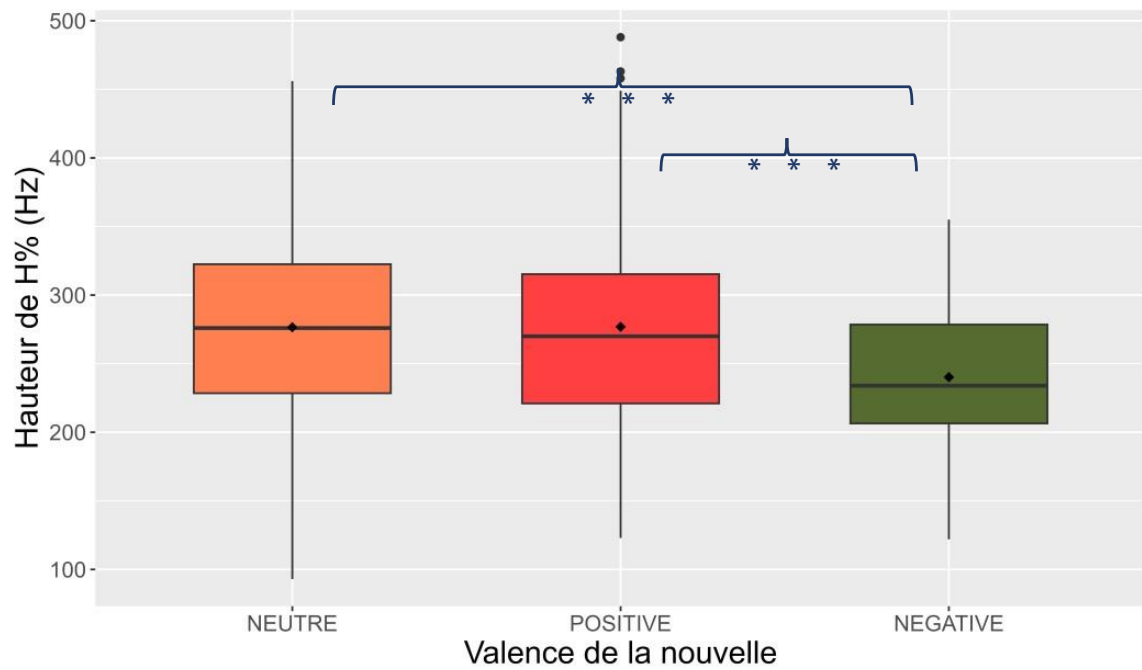


Figure 70 – Valeur mélodique des tons H% en fin de préface

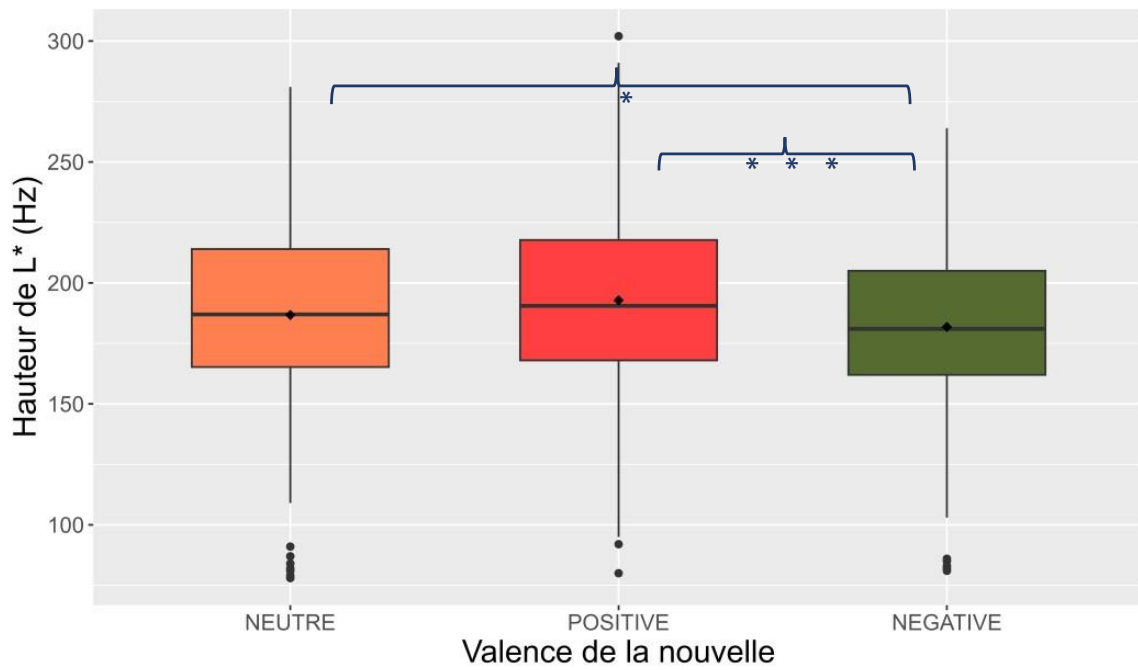


Figure 71 – Valeur mélodique des tons L* en fin de préface

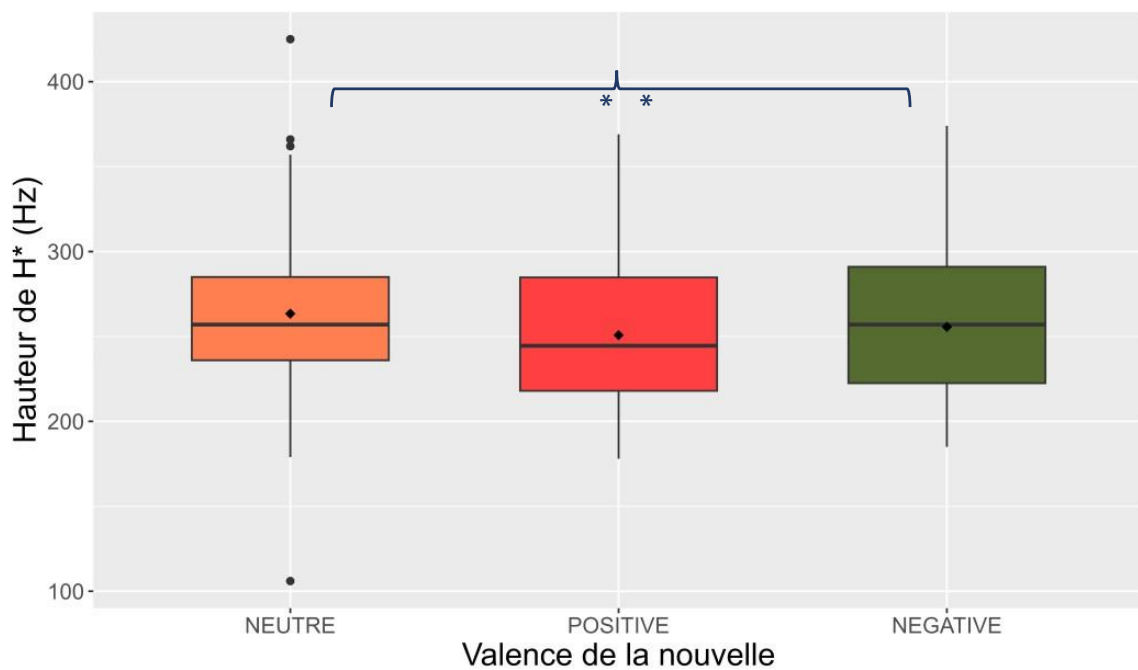


Figure 72 – Valeur mélodique des tons H* en fin de préface

Alignement des tons

Dans un second temps, nous présentons les résultats des analyses d'alignement des accents mélodiques L* et H*. Les analyses ont démontré qu'il n'existe aucune différence significative d'alignement des cibles dans la préface, qu'elle que soit la valence de la nouvelle à annoncer [$p > 0.025$].

3.2.5. « voilà donc »

Valeur mélodique des tons

Nous présentons maintenant les résultats des analyses de valeur mélodique réalisés au sein des marqueurs discursifs « voilà donc ». Nous présentons les résultats détaillés selon l'énoncé sur lequel ils étaient produits (« voilà », « donc », ou « voilà donc »). Nous trouvons de manière générale que les tons sont réalisés plus bas lorsque la valence de la nouvelle est négative, et plus hauts lorsqu'elle est positive. La valeur mélodique des tons produits avant l'annonce d'une nouvelle positive est généralement assez proche de celle produite lorsque la nouvelle est neutre. Cependant, lorsque nous trouvons des différences significatives entre ces deux valences, les tons sont produits avec une valeur mélodique plus haute lorsque la valence a une nouvelle positive plutôt que neutre.

Nous commençons par les résultats trouvés dans l'analyse des tons qui sont ancrés à une frontière prosodique, soit les tons de frontière d'IP et ceux d'ip (frontière droite), et L1 (frontière gauche). Le Tableau 20 ci-dessous présente les valeurs moyennes et déviations-standard, en Hertz, pour la hauteur des cibles tonales.

MD	Ton	<i>Négatif</i>		<i>Positif</i>		<i>Neutre</i>	
		Moy	SD	Moy	SD	Moy	SD
Voilà	L%	170,0	25,8	184,7	32,8	177,7	30,3
	H%	231,0	26,1	262,6	34,4	258,6	34,8
	L1	205,8	28,9	216,6	30,3	209,3	22,0
Donc	L%	183,1	26,6	193,1	22,0	186,8	32,9
	H%	215,8	23,2	244,2	38,6	236,4	37,8
	L1	201,4	27,7	221,5	30,1	213,8	26,7
Voilà donc	L%	175,5	28,6	172,6	31,1	183,1	39,4
	H%	281,7	43,7	309,9	47,1	299,1	38,1
	L1	203,5	20,9	215,6	36,8	204,1	27,8

Tableau 20 – Valeurs moyennes et SD (Hertz) pour les cibles tonales L%, H% et L1 évaluées au sein des marqueurs discursifs

D'après les trois figures (Figure 73, Figure 74, Figure 75) nous pouvons voir des différences de hauteur pour ton L%, H% et L1 entre les valences négative et positive, et parfois entre les valences négative et neutre.

Tout d'abord lorsque les tons sont produits sur l'énoncé « voilà », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la

nouvelle est négative, par rapport à lorsqu'elle est positive pour L% [$\beta = -10.87$, SE = 3.58, $t = -3.03$, $p < 0.01$], et pour H% [$\beta = -34.89$, SE = 10.64, $t = -3.27$, $p < 0.01$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour H% [$\beta = -26.84$, SE = 10.04, $t = -2.67$, $p < 0.01$], mais pas pour L% [$p = 0.1$]. Enfin, nous ne trouvons pas de différence significative entre le positif et le neutre pour L% [$p = 0.2$] ou H% [$p = 0.3$].

Lorsque les tons sont produits sur l'énoncé « donc », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour L% [$\beta = -10.19$, SE = 4.11, $t = -2.47$, $p < 0.025$], et H% [$\beta = -20.12$, SE = 6.11, $t = -3.29$, $p < 0.01$]. Nous ne trouvons pas de différence significative entre le négatif et le neutre pour L% [$p = 0.8$] ou H% [$p = 0.3$]. Enfin, nous trouvons une différence significative avec un ton réalisé plus haut lorsque la valence de la nouvelle est positive, par rapport à lorsqu'elle est neutre pour L% [$\beta = 9.339$, SE = 4.07, $t = 2.29$, $p < 0.025$], mais pas pour H% [$p = 0.06$].

Lorsque les tons sont produits sur l'énoncé « voilà donc », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour H% [$\beta = -32.29$, SE = 5.27, $t = -6.11$, $p < 0.0001$], mais pas pour L% [$p = 0.6$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour H% [$\beta = -24.12$, SE = 5.48, $t = -4.39$, $p < 0.0001$], mais pas L% [$p = 0.4$]. Enfin, nous ne trouvons pas de différence significative entre le positif et le neutre pour L% [$p = 0.7$] ou H% [$p = 0.1$].

Nous présentons maintenant les analyses réalisées sur L1 (Figure 75). Nous avons trouvé peu de différences significatives sur ce ton, car uniquement lorsqu'il est réalisé sur l'énoncé « voilà ». Dans ce cadre-là L1 est réalisé significativement plus bas lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive [$\beta = -10.79$, SE = 2.35, $t = -4.58$, $p < 0.0001$]. Il est réalisé plus haut lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre [$\beta = 8.48$, SE = 2.08, $t = 4.06$, $p < 0.0001$]. Nous n'avons trouvé aucune différence significative sur « voilà » entre la valence positive et la valence neutre, et aucune sur « voilà donc » [$p > 0.025$].

Aucune différence de valeur mélodique n'a été trouvée sur les tons de frontière d'ip, sur L- lorsqu'il était réalisé sur « voilà donc », entre les valences négatives et neutres pour H-, et entre les valences positives et neutres pour L- sur « voilà » [$p > 0.025$].

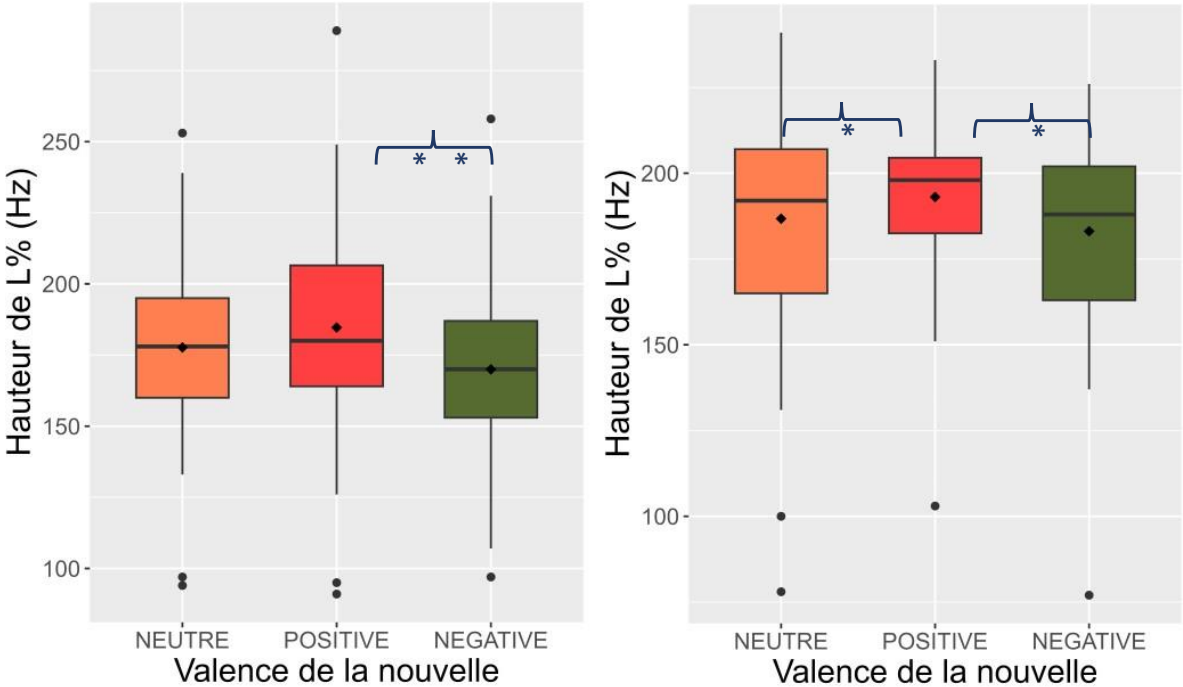


Figure 73 – Valeur mélodique des tons L% dans « voilà » (à gauche), et « donc » (à droite)

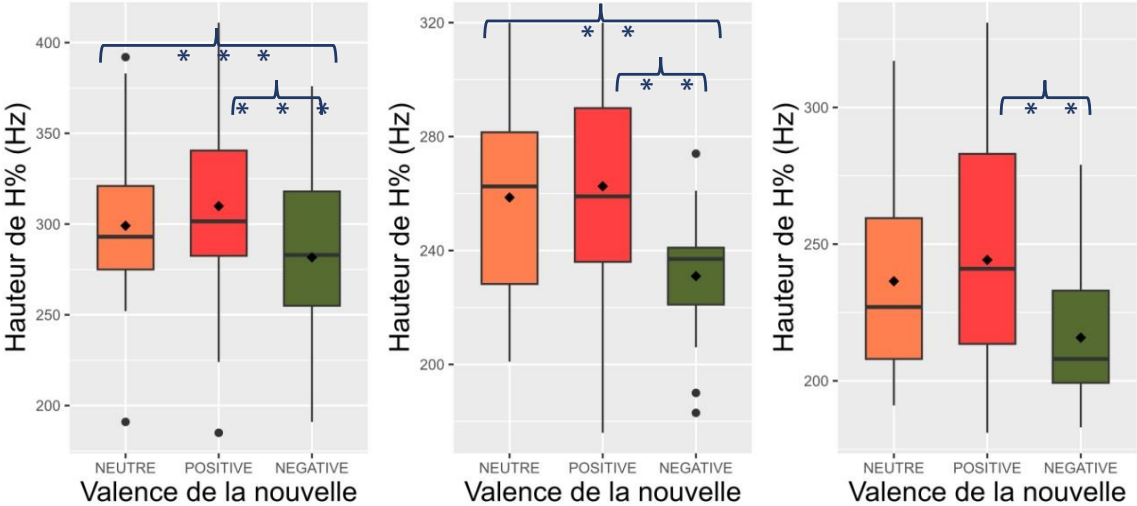


Figure 74 – Valeur mélodique des tons H% dans « voilà donc » (à gauche), « voilà » (au milieu) et « donc » (à droite)

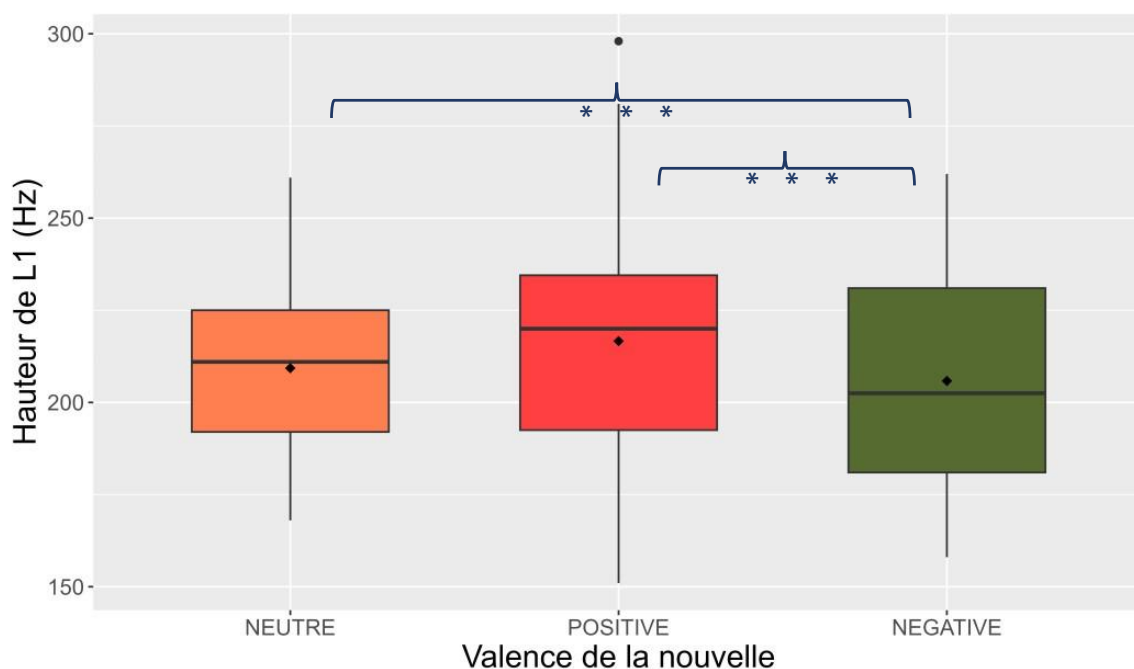


Figure 75 – Valeur mélodique de L1 dans « voilà »

Nous présentons maintenant les résultats trouvés dans l'analyse des accents mélodiques, qui suivent la même tendance que les tons de frontière. Le Tableau 21 ci-dessous présente les valeurs moyennes et déviations-standard, en Hertz, pour la hauteur des cibles tonales.

MD	Ton	<i>Négatif</i>		<i>Positif</i>		<i>Neutre</i>	
		Moy	SD	Moy	SD	Moy	SD
Voilà	<i>L*</i>	179,6	26,9	190,0	28,5	189,2	29,4
	<i>H*</i>	237,1	23,3	251,1	32,7	244,7	30,8
Donc	<i>L*</i>	189,4	26,3	195,3	23,9	192,4	29,5
	<i>H*</i>	210,6	24,2	234,0	34,8	225,3	32,3
Voilà donc	<i>L*</i>	184,0	21,8	180,6	30,7	184,1	30,0
	<i>H*</i>	268,8	45,7	295,8	56,6	285,4	44,2

Tableau 21 – Valeurs moyennes et SD (Hertz) pour les cibles tonales *L** et *H** évaluées au sein des marqueurs discursifs

Nous voyons dans les figures (Figure 76, Figure 77) que la différence entre valence positive et valence négative semble être la plus marquée, que ce soit sur le ton *H** ou *L**. La différence entre les valences négative et neutre semble bien marquée également. En

revanche, il ne semble pas toujours y avoir de différence significative de hauteur des cibles tonales lorsque nous comparons les valences positive et neutre.

Tout d'abord lorsque les tons sont produits sur l'énoncé « voilà », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour L* [$\beta = -7.19$, SE = 2.18, $t = -3.29$, $p < 0.01$], et pour H* [$\beta = -25.2$, SE = 4.5, $t = -5.59$, $p < 0.0001$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour L* [$\beta = -6.03$, SE = 2.21, $t = -2.72$, $p < 0.01$], et pour H* [$\beta = -16.47$, SE = 4.35, $t = -3.77$, $p < 0.0001$]. Enfin, nous trouvons une différence significative avec un ton réalisé plus haut lorsque la valence de la nouvelle est positive, par rapport à lorsqu'elle est neutre pour H* [$\beta = 8.73$, SE = 3.6, $t = 2.42$, $p < 0.025$] mais pas pour L* [$p = 0.6$].

Lorsque les tons sont produits sur l'énoncé « donc », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour L* [$\beta = -6.15$, SE = 1.9, $t = -3.23$, $p < 0.01$], et H* [$\beta = -15.59$, SE = 2.87, $t = -5.42$, $p < 0.0001$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour H* [$\beta = -7.46$, SE = 3.04, $t = -2.45$, $p < 0.025$] mais pas pour L* [$p = 0.4$]. Enfin, nous trouvons une différence significative avec un ton réalisé plus haut lorsque la valence de la nouvelle est positive, par rapport à lorsqu'elle est neutre pour L* [$\beta = 4.75$, SE = 1.83, $t = 2.58$, $p < 0.01$], mais pas pour H* [$\beta = 8.13$, SE = 3.07, $t = 2.64$, $p < 0.01$].

Lorsque les tons sont produits sur l'énoncé « voilà donc », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour H* [$\beta = -29.87$, SE = 4.58, $t = -6.51$, $p < 0.0001$], mais pas pour L* [$p = 0.8$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour H* [$\beta = -24.96$, SE = 4.58, $t = -5.44$, $p < 0.001$], mais pas L* [$p = 0.6$]. Enfin, nous ne trouvons pas de différence significative entre le positif et le neutre pour L* [$p = 0.8$] ou H* [$p = 0.2$].

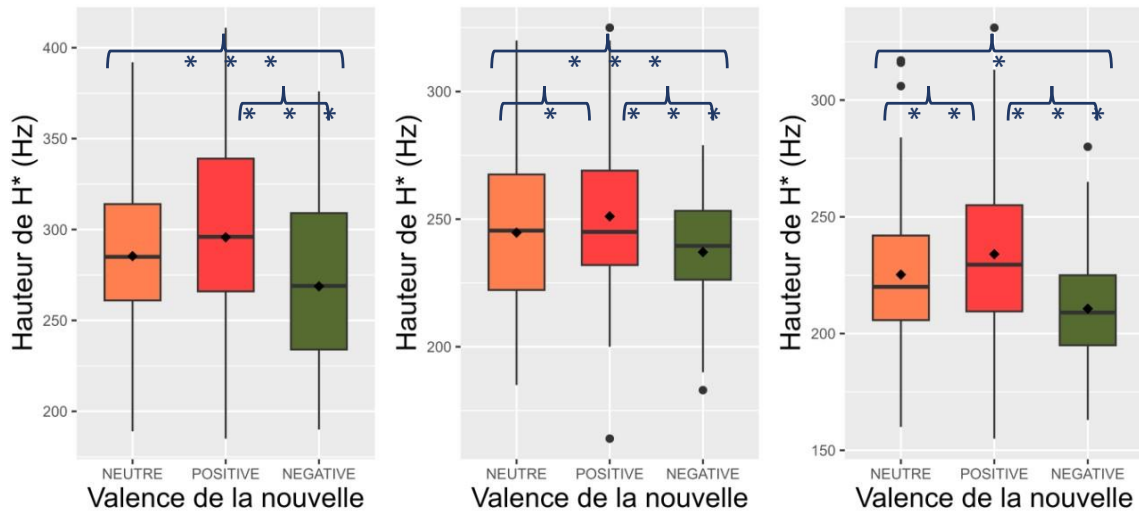


Figure 76 – Valeur mélodique des tons H* dans « voilà donc » (à gauche), « voilà » (au milieu) et « donc » (à droite)

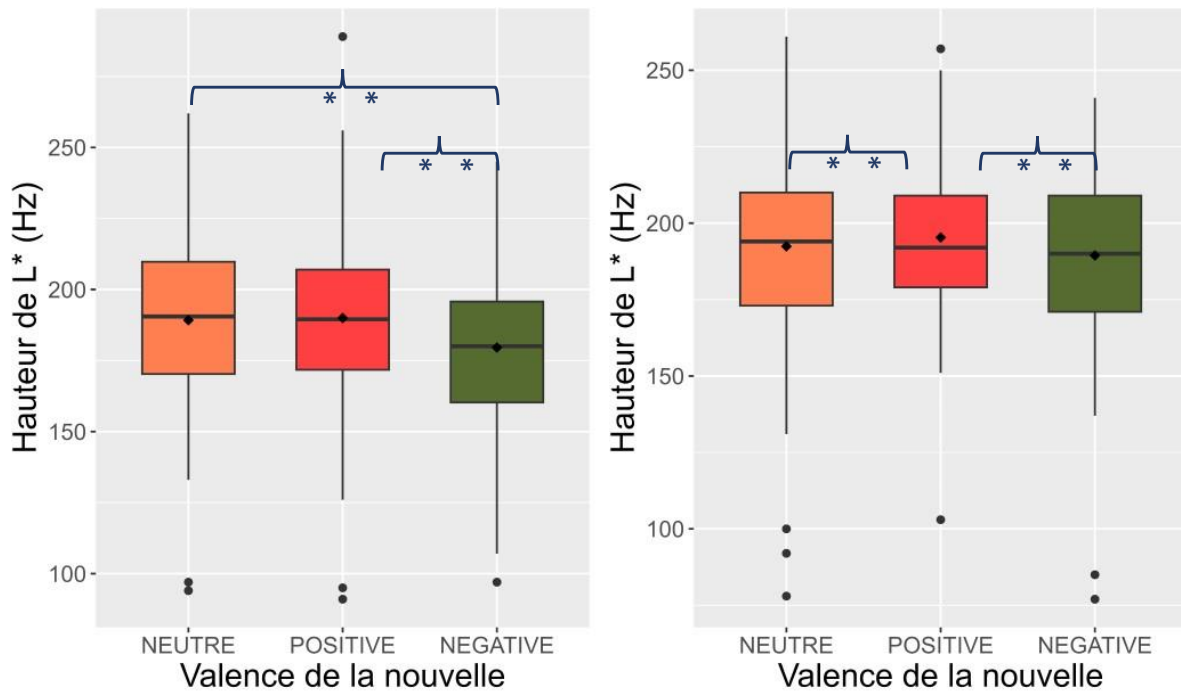


Figure 77 – Valeur mélodique des tons L* dans « voilà » (à gauche) et « donc » (à droite)

Nous allons maintenant nous attarder sur les cibles tonales réalisées pour les tons L2 et Hi, qui n'ont pas d'ancrage segmental spécifique, mais sont contraints d'être produits en séquence linéaire (Hi, puis L2), et tous deux apparaissent entre L1 et H*. Le Tableau 22

ci-dessous présente les valeurs moyennes et déviations-standard, en Hertz, pour la hauteur des cibles tonales.

MD	Ton	<i>Négatif</i>		<i>Positif</i>		<i>Neutre</i>	
		Moy	SD	Moy	SD	Moy	SD
Voilà	<i>Hi</i>	218,6	23,1	238,1	28,3	226,4	23,1
	<i>L2</i>	208,0	25,9	189,0	30,2	191,4	29,7
Voilà donc	<i>Hi</i>	221,2	22,7	254,3	41,5	239,9	34,3
	<i>L2</i>	204,7	25,7	209,4	32,9	198,2	31,3

Tableau 22 – Valeurs moyennes et SD (Hertz) pour les cibles tonales Hi et L2 évaluées au sein des marqueurs discursifs

D'après les graphiques (Figure 78 pour Hi et Figure 79 pour L2), Hi semble présenter des différences significatives entre chacune des valences. Les différences sont moins évidentes pour le ton L2, avec des valeurs médianes et moyennes beaucoup plus proches.

Tout d'abord lorsque les tons sont produits sur l'énoncé « voilà », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour Hi [$\beta = -17.36$, SE = 1.5, t = -11.52, $p < 0.0001$], mais pas pour L2 [$p = 0.2$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour Hi [$\beta = -7.42$, SE = 1.53, t = -4.84, $p < 0.0001$], mais pas pour L2 [$p = 0.2$]. Enfin, nous trouvons une différence significative avec un ton réalisé plus haut lorsque la valence de la nouvelle est positive, par rapport à lorsqu'elle est neutre pour Hi [$\beta = 9.93$, SE = 1.52, t = 6.51, $p < 0.0001$] mais pas pour L2 [$p = 0.9$].

Lorsque les tons sont produits sur l'énoncé « voilà donc », nous trouvons des différences significatives de hauteur avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est positive pour Hi [$\beta = -29.3$, SE = 4.35, t = -6.73, $p < 0.0001$], et pour L2 [$\beta = -8.09$, SE = 3.39, t = -2.38, $p < 0.025$]. Nous trouvons également une différence significative avec un ton réalisé plus bas lorsque la valence de la nouvelle est négative, par rapport à lorsqu'elle est neutre pour Hi [$\beta = -11.8$, SE = 4.3, t = -2.74, $p < 0.01$], mais pas L2 [$p = 0.6$]. Enfin, nous trouvons une différence significative avec un ton réalisé plus haut lorsque la valence de la nouvelle est positive, par rapport à lorsqu'elle est neutre pour Hi [$\beta = 17.5$, SE = 3.6, t = 4.85, $p < 0.0001$] mais pas pour L2 [$p = 0.5$].

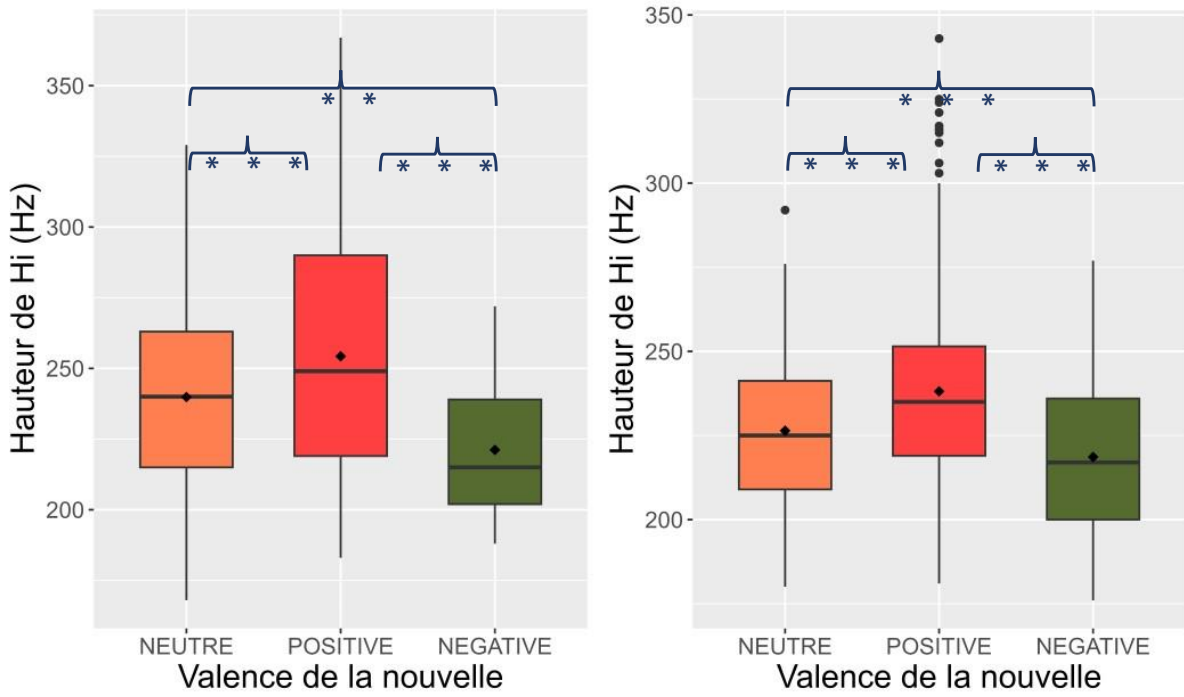


Figure 78 – Valeur mélodique des tons Hi dans « voilà donc » (à gauche), « voilà » (à droite)

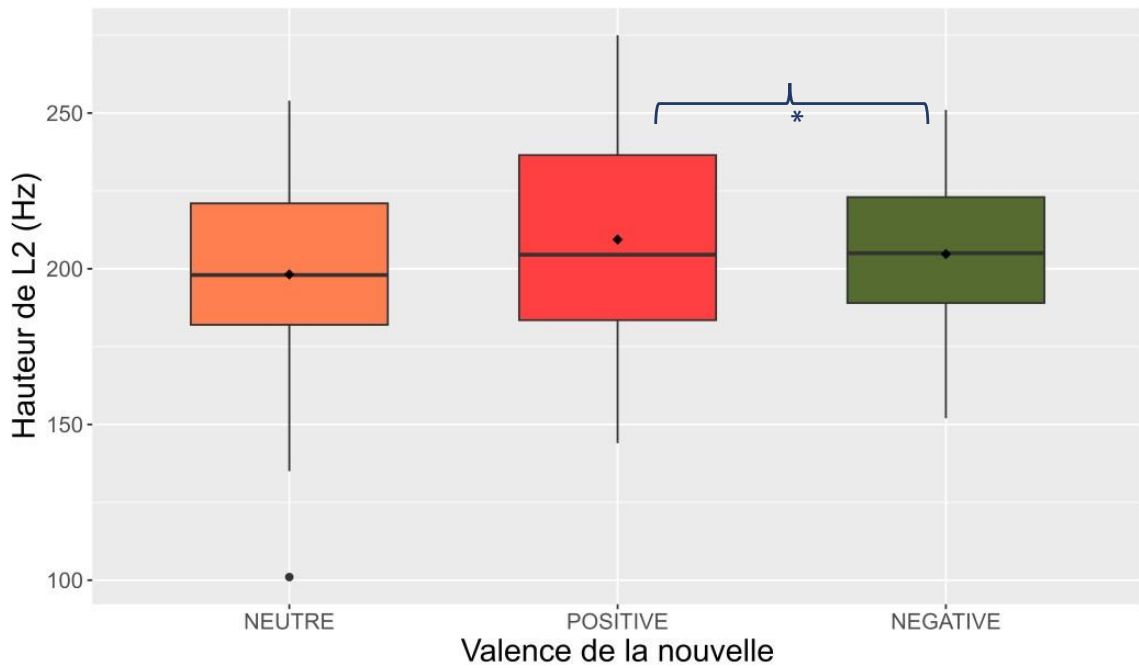


Figure 79 – Valeur mélodique de L2 dans « voilà donc »

Alignement tonal

Nous présentons maintenant les résultats des analyses de l'alignement tonal pour les différents tons extraits dans notre corpus. Ces mesures sont prises par rapport à la syllabe. De la même façon de dans la section précédente sur la hauteur mélodique, nous précisons lorsque les tons ont été réalisés sur l'énoncé « voilà », « donc », ou « voilà donc ». Le Tableau 23 ci-dessous présente les valeurs moyennes et déviations-standard, en pourcentages, pour l'alignement des cibles tonales à la syllabe.

MD	Ton	<i>Négatif</i>		<i>Positif</i>		<i>Neutre</i>	
		Alignement (%)		Alignement (%)		Alignement (%)	
		Moy	SD	Moy	SD	Moy	SD
Voilà	<i>L*</i>	20,9	20,5	19,2	17,5	17,8	18,4
	<i>H*</i>	16,2	8	17,7	23,2	19	12,4
Donc	<i>L*</i>	24,8	28,2	33,2	34,1	29,8	41,3
	<i>H*</i>	21,3	23	21,9	20,9	23	19,1
Voilà donc	<i>L*</i>	29,1	31,7	33,7	29,9	26,1	28,4
	<i>H*</i>	18,3	20,9	27,9	32,2	18,7	23,6

Tableau 23 – Valeurs moyennes et SD d'alignement (%) pour les cibles tonales *L** et *H** évaluées au sein des marqueurs discursifs

D'après la Figure 80, le ton *H** semble présenter des différences entre la nouvelle positive et les nouvelles neutre et négative. D'après la Figure 81, le ton *L** porte plutôt une distinction entre la valence négative et la valence neutre.

Tout d'abord lorsque les tons sont produits sur l'énoncé « voilà », nous ne trouvons pas de différence significative d'alignement entre le négatif et le positif pour *L** [$p = 0.06$] ou *H** [$p = 0.5$]. Nous trouvons en revanche que le ton est réalisé significativement plus tôt lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre pour *L** [$\beta = 4.18$, $SE = 1.75$, $t = 2.37$, $p < 0.025$], mais pas pour *H** [$p = 0.1$]. Enfin, nous ne trouvons pas de différence significative entre le positif et le neutre pour *L** [$p = 0.6$] ou *H** [$p = 0.3$].

Lorsque les tons sont produits sur l'énoncé « donc », nous ne trouvons aucune différence significative pour les tons *H** ou *L** [$p > 0.025$].

Lorsque les tons sont produits sur l'énoncé « voilà donc », nous trouvons que le ton est réalisé significativement plus tard lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est positive pour *H** [$\beta = -10.96$, $SE = 3.99$, $t = -2.74$, $p < 0.01$], mais pas pour *L** [$p = 0.4$]. Nous ne trouvons pas de différence significative entre le négatif et

le neutre pour L* [p = 0.7] ou H* [p = 0.9]. En revanche, nous nous trouvons que le ton est réalisé significativement plus tôt lorsque la valence de la nouvelle est positive par rapport à lorsqu'elle est neutre pour H* [$\beta = 11.06$, SE = 4.08, t = 2.7, p < 0.01] mais pas pour L* [p = 0.2].

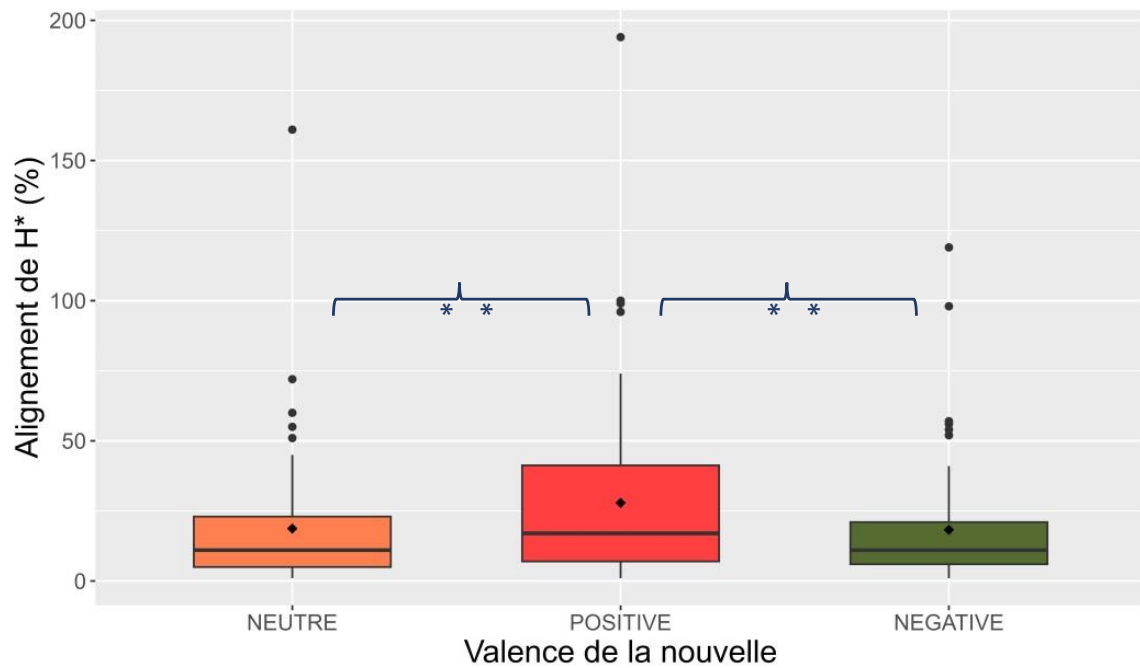


Figure 80 – Alignement des tons H* à la syllabe au sein de « voilà donc »

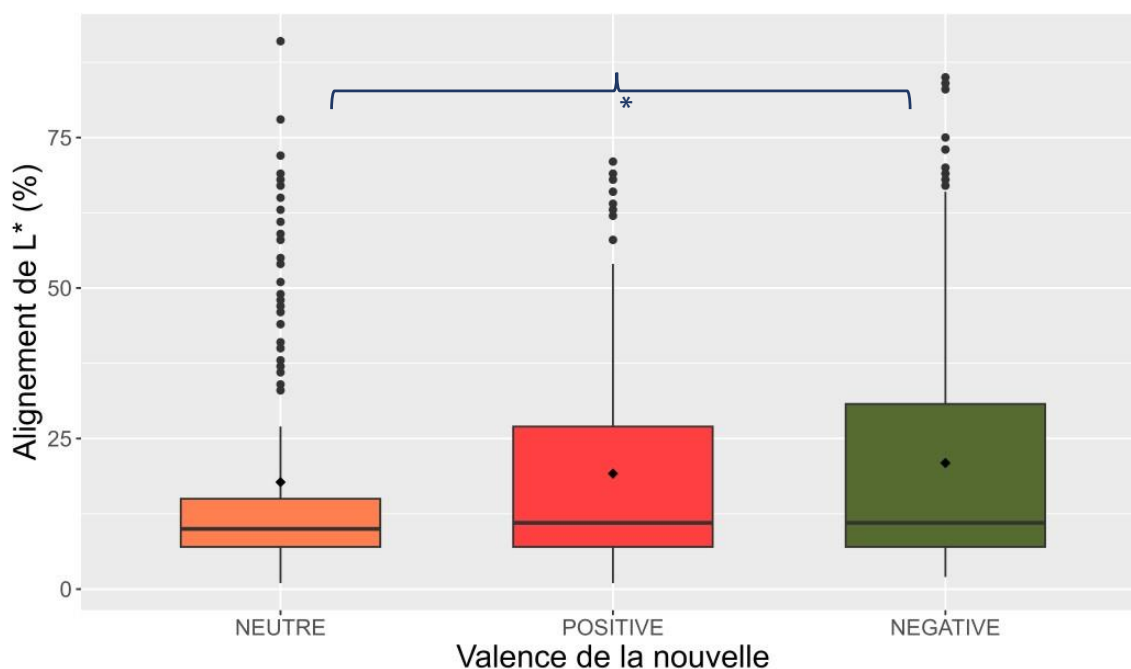


Figure 81 – Alignement des tons L* à la syllabe au sein de « voilà »

Nous présentons maintenant les résultats des analyses d'alignement réalisées sur L2 et Hi. Le Tableau 24 ci-dessous présente les valeurs moyennes et déviations standard, en pourcentages, pour l'alignement des cibles tonales à la syllabe.

DMs	Ton	<i>Négatif</i>		<i>Positif</i>		<i>Neutre</i>	
		Alignement (%)		Alignement (%)		Alignement (%)	
		Moy	SD	Moy	SD	Moy	SD
Voilà	<i>Hi</i>	41,2	17	37,3	22,5	39,1	21,9
	<i>L2</i>	61,8	36,5	71	12,6	69,9	22,3
Voilà donc	<i>Hi</i>	28,3	40,2	19,2	34,5	32,1	36,3
	<i>L2</i>	7,7	5,8	8,5	7,3	11,6	17,4

Tableau 24 – Valeurs moyennes et SD d'alignement (%) pour les cibles tonales Hi et L2 évaluées au sein des marqueurs discursifs

L'observation de la Figure 82 pour Hi nous permet de remarquer que les cibles tonales semblent présenter des différences entre les valences positive et neutre seulement. Rien ne semble ressortir pour L2, dont les graphiques sont en annexe.

Une seule différence significative est trouvée, sur l'énoncé « voilà donc ». Nous trouvons un ton réalisé significativement plus tard lorsque la valence de la nouvelle est

positive par rapport à lorsqu'elle est neutre pour Hi [$\beta = -14.89$, $SE = 5.05$, $t = -2.94$, $p < 0.01$], mais pas pour L2 [$p = 0.4$]. Toutes les autres différences sont au dessus du niveau alpha.

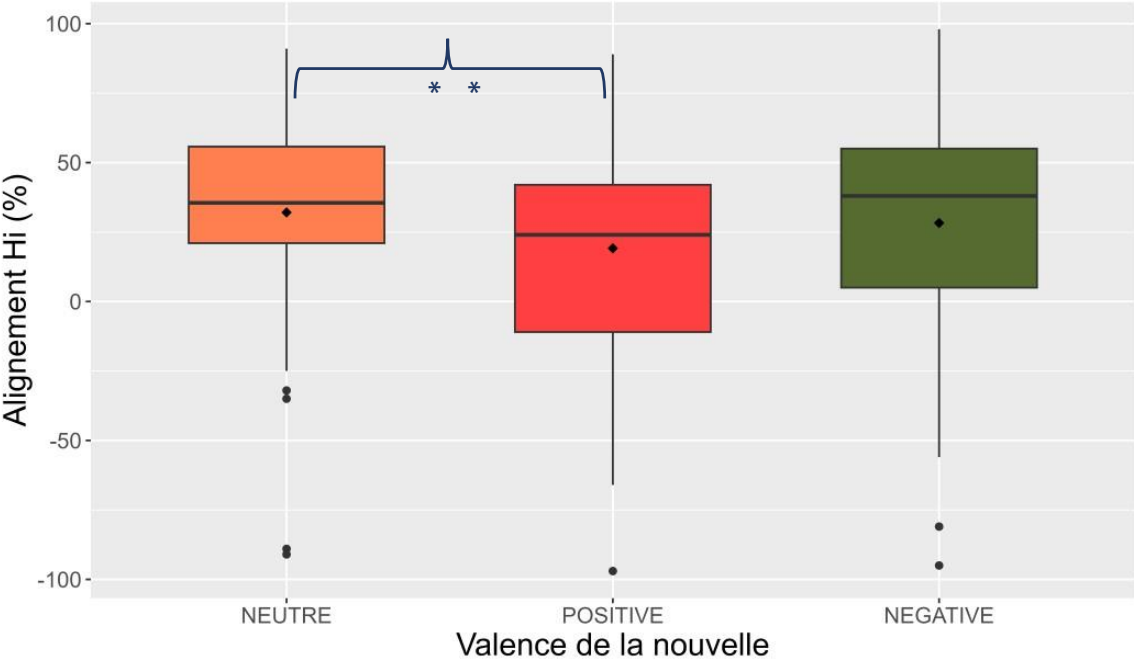


Figure 82 – Alignement de Hi au sein de « voilà donc »

4. Discussion

Dans ce chapitre, nous avons cherché à voir si la valence affective est portée dans la préface par des indices de projection phonologiques, au travers de patrons intonationnels. Afin d'observer cela, nous avons annoté phonologiquement le corpus d'annonces de nouvelles présenté dans le Chapitre 2, en nous focalisant sur la configuration de fin d'IP à la fin de la préface, c'est-à-dire le dernier accent mélodique, et le dernier ton de frontière de la dernière IP, et en nous focalisant sur l'ensemble du ou des patrons phonologiques associés à « voilà » et « donc » dans les marqueurs discursifs. Nous avons évalué la présence ou l'absence des patrons, ainsi que la réalisation phonétique des cibles tonales produites en termes de valeur mélodique et d'alignement.

Mesure		Préface			MDs		
		Neutre	Positive	Négative	Neutre	Positive	Négative
Présence d'un patron	L*L%	<	_	>	X	X	X
	HiL*L%	X	X	X	<	_	>
Hauteur mélodique	L%	_	>	<	<	>	<
	H%	>	>	<	>	>	<
	L*	>	>	<	<>	>	<
	H*	>	_	<	<>	>	<
	L1	X	X	X	>	>	<
	Hi	X	X	X	<>	>	<
	L2	X	X	X	_	>	<
Alignement	L*	_	_	_	<	_	>
	H*	_	_	_	<	>	<
	L1	X	X	X	_	_	_
	Hi	X	X	X	>	<	_
	L2	X	X	X	_	_	_

Tableau 25 – Tableau récapitulatif des différences phonologiques et phonétiques entre les trois valences, au sein des différents énoncés (< moins présent / plus tard, > valeur haute, <> intermédiaire, _ non significatif, >< inconsistant, X NA)

Il y a trois résultats majeurs dans cette étude : ceux concernant les patrons phonologiques, ceux concernant la valeur mélodique des cibles tonales, et ceux concernant l'alignement de ces cibles. Nous avons trouvé que L*L% dans la préface, et

HiL*L% dans le marqueur discursif « voilà » étaient plus utilisés dans les préfaces négatives que dans les préfaces neutres. Au niveau de la valeur mélodique des cibles tonales, nous avons trouvé des différences significatives avec des tons réalisés plus bas lorsque la valence de la nouvelle est négative, et plus haut lorsque la valence de la nouvelle est positive. La hauteur des cibles pour le neutre est régulièrement assez proche de celle que l'on trouve pour le positif, bien que la différence entre les deux soit parfois suffisamment marquée pour être significative. Enfin, nous avons aussi trouvé des différences d'alignement des cibles tonales. Cependant, les résultats sont très irréguliers et varient en fonction de la cible tonale observée, particulièrement pour les valences positive et négative. Pour la valence neutre, les accents mélodiques sont réalisés plus tard que pour les autres valences.

Les découvertes concernant l'annotation phonologique sont donc en accord avec la littérature (Cao et al., 2014; Liscombe et al., 2003; Wang et al., 2020), car comme eux, nous avons trouvé que la valence affective d'un énoncé peut être portée plus régulièrement par certains patrons que par d'autres. Les patrons trouvés dans la préface et le marqueur discursif « voilà » présentent des tons bas pour l'accent mélodique et le ton de frontière. Ce résultat est cohérent les résultats phonétiques trouvés au Chapitre 2 et avec l'hypothèse que nous avons formulée quant à l'influence que le Code de Fréquence (Gussenhoven, 2002; Ohala, 1983; Rodero, 2011) pourrait avoir sur les patrons des configurations de fin d'IP de la préface et des marqueurs discursifs, en lien avec la nature menaçante de l'annonce de nouvelle négative qui serait formulée plutôt avec des tons bas. Le fait que nous n'ayons trouvé que deux patrons phonologiques qui semblent porter du sens affectif suggère néanmoins que ceux-ci ne jouent pas un rôle prépondérant dans le port d'indices sur la valence affective de la nouvelle qui doit être annoncée.

Les résultats concernant la valeur mélodique des cibles tonales sont en ligne avec les résultats d'analyses acoustiques globales et localisées proposées pour la f0 dans le Chapitre 2, et donc avec la littérature sur les affects (Berry & Brown, 2019; Scherer, 2003). Ils sont également cohérents avec les hypothèses que nous avons formulées sur la base des travaux de Freese et Maynard (1998) sur les annonces de nouvelles. En effet, l'étude citée avait décrit des contours prosodiques hauts et soutenus avant de tomber sur la syllabe finale lorsque la valence de la nouvelle annoncée était positive, ce qui nous poussait à attendre des cibles H plus hautes pour la valence positive. Ces découvertes

confirment notre hypothèse selon laquelle les tendances globales trouvées dans les analyses de la f_0 se retrouvent également à certains points clé de l'énoncé.

Nous n'avons pas formulé d'hypothèses sur notre exploration de l'alignement des cibles tonales au sein de la syllabe, ce paramètre n'ayant pas été étudié à ce jour, à notre connaissance, en lien avec des informations paralinguistiques telles que la valence affective d'une nouvelle. Les résultats présentés indiquent quelques différences significatives qui restent peu nombreuses et instables, au sein des marqueurs discursifs. Lorsqu'elles apparaissent, elles présentent beaucoup de variabilité. L'alignement n'apparaît donc pas comme un indice de projection de valence dans la préface, et son utilisation dans les marqueurs discursifs reste cantonnée à une différenciation entre le neutre et les autres valences (positive, négative).

Avant de présenter dans le détail notre discussion des résultats, nous devons discuter de ceux de l'accord inter-annotateurs. S'ils ont globalement permis de nous assurer de la validité de nos annotations, ils montrent aussi un désaccord entre les annotateurs 1 et 2 quant à la catégorisation des accents mélodiques dans les préfaces, avec un coefficient de 0.37, soit un « accord faible ». Le tableau des contingences (Section 3.1.1) pour les 2 catégories de ce test montre que l'annotateur 1 a très souvent annoté comme L^* des tons qui ont été annotés comme H^* par l'annotateur 2. Il serait intéressant de regarder si le désaccord entre les annotateurs est plus représenté en fonction de la valence affective de la nouvelle. Par exemple, nous pourrions regarder si c'est lorsque la préface de la nouvelle est négative et les valeurs de la f_0 abaissées que le désaccord est le plus présent.

Cette annotation devrait, par la suite, être révisée tout d'abord par une discussion plus approfondie des critères d'annotation, puis par une ré-annotation. Herman et McGory (2002) ont montré dans une étude sur la similarité conceptuelle des tons de l'intonation que si certains accents mélodiques et les tons périphériques sont considérés comme similaires pour un annotateur, cela augmente le risque de désaccord entre annotateurs. Il serait donc possible dans notre corpus, les deux catégories H^* et L^* ont été considérées comme similaires par l'un des annotateurs. Nous écartons cependant cette hypothèse, car Herman et McGory (2002) ont trouvé que les catégories H^* et L^* en anglais américain sont généralement considérées comme dissimilaires en anglais américain, il est donc probable qu'elles le soient aussi en français.

Le résultat principal de nos analyses phonologiques est la présence accrue de deux patrons lorsque la valence de la nouvelle est négative par rapport à lorsqu'elle est neutre : la configuration de fin d'IP L*L% dans la préface (16 % négatif – 13 % neutre), et le patron HiL*L% dans le marqueur « voilà » (31 % négatif – 18 % neutre). Ils correspondent tous deux au patron le plus représenté dans leur unité, toutes valences confondues. Cette différence est la seule qui a été trouvée de toutes nos analyses phonologiques, mais elle est en accord avec plusieurs éléments de la littérature. Tout d'abord, l'omniprésence de ces deux patrons n'a rien d'étonnant : tous deux comportent L*L%, la configuration de fin d'IP généralement produite sur les affirmations en français, lorsque l'on marque un focus large ou étroit (Delais-Roussarie et al., 2015). Nous ajoutons que Cao et al. (2014) ont eux aussi trouvé, en anglais américain, que la plupart de leurs données montraient un patron final déclaratif L-L%, lié aux affirmations. En revanche, nos résultats diffèrent en ce que le patron trouvé pour l'anglais américain était moins présent lorsque l'énoncé avait pour affect la peur, la joie ou la tristesse. Or, nous aurions pu penser que s'il était moins présent pour la tristesse, il aurait également été moins présent lorsque nous produisons des énoncés à valence négative. Nos résultats sont en revanche en ligne avec ceux de Stibbard (2000) sur l'anglais britannique. Il avait en effet trouvé que l'accent mélodique L* était plus présent lorsque l'émotion observée était la tristesse, ce qui indique pour lui une tendance à produire un ton bas qu'il considère en accord avec la littérature sur les affects. C'est-à-dire, la tendance à produire des énoncés dans un registre plus bas lorsque l'affect produit est la tristesse (Berry & Brown, 2019; Scherer, 2003).

Le patron HiL*L% produit sur « voilà », est par ailleurs le patron que nous avons décrit comme la correspondance du contour « descendant » décrit par Petit (2020), et il est un corrélat de la fonction de conclusion de sujet du marqueur discursif en question. C'est l'un des patrons que nous nous attendions donc à retrouver dans notre corpus, pour conclure la partie de la préface et passer à la séquence d'annonce de nouvelle. Les trois patrons principaux trouvés pour « voilà » sont des variantes de ce contour descendant, avec différents marquages de frontières prosodiques : HiL*L%, HiL* et HiL*L-. Par ailleurs, le ton Hi présent dans ce patron est un ton décrit comme pouvant marquer la frontière gauche du focus de l'énoncé (D'Imperio et al., 2012; Frota & Prieto, 2015; German & D'Imperio, 2015). Il n'est donc pas surprenant de le retrouver utilisé ici, adjoint à une configuration de fin d'IP portant une fonction d'affirmation avec marquage de focus large ou étroit. Nous avons également fait l'hypothèse que le patron que nous pourrions

trouver sur « voilà » lorsque nous devons marquer une valence négative serait celui de conclusion forte en « cloche » (LHiL*), mais celui-ci n'est pas apparu comme l'un des patrons principaux dans notre corpus. Bien que nous ne trouvions pas de différence en fonction de la valence affective dans les marqueurs discursifs « donc », les patrons les plus fréquents sont ceux que nous attendions vis-à-vis des fonctions discursives de « donc ». C'est le patron « plat » conclusif proposé par Haileselassie (2015) qui est le plus présent : L*, L*L% et L*L-. Il est possible que l'absence de résultats significatifs sur les autres patrons, tant au sein des marqueurs discursifs que de la préface, soit due au fait qu'ils soient moins représentés dans notre corpus. Nos résultats nous poussent à penser que pour marquer la valence négative, les configurations de fin d'IP basses sont préférées, en lien avec l'abaissement général de la f0. Cela expliquerait la présence des patrons que nous avons trouvés, par un effet principal de la configuration de fin d'IP. Hi pourrait alors mettre en exergue, par contraste, la configuration.

Le second résultat majeur de cette étude est trouvé vis-à-vis de la valeur mélodique des cibles tonales. De façon générale, nous avons trouvé des différences significatives cohérentes avec les résultats du Chapitre 2 précédent : les tons sont réalisés systématiquement plus bas dans le négatif que dans le positif ou le neutre, et systématiquement plus bas dans le neutre que dans le positif, lorsqu'une différence est trouvée entre ces deux conditions. Ce résultat est cohérent avec la littérature sur les affects (Berry & Brown, 2019; Scherer, 2003), comme nous l'avons évoqué au début de cette discussion. Ils sont aussi consistants avec ceux de Cao et al. (2014), qui avaient par exemple trouvé que dans le patron H*!H-L%, la cible H* était généralement très haute, et ce, particulièrement pour la joie, un affect que nous rapprochons de notre valence affective positive. Cet effet trouvé pour la joie et la valence positive pourrait aussi être le résultat d'un niveau plus haut d'excitation que pour d'autres affects (ex. tristesse). Ils avaient par ailleurs noté que ToBI ne propose pas de façon de rendre compte de ce type de différences de hauteur systématique. Pourtant, ils soulignent que cette différence en particulier pourrait jouer un rôle dans la désambiguïsation du patron selon s'il est réalisé avec joie ou peur. Nous nous attardons maintenant quelque peu sur les résultats de la valeur mélodique trouvés en fonction des différents tons observés. Dans la préface comme dans les marqueurs discursifs, les tons sont généralement réalisés plus haut, lorsque la valence de la nouvelle est positive, plus bas lorsque la valence de la nouvelle est négative,

avec une hauteur intermédiaire lorsqu'elle est neutre. Les seules exceptions sont situées dans la préface. Il se pourrait que l'absence d'effets de la valence positive sur H* et de la valence négative sur L% vienne d'un effet de plafond et de sol lié à l'empan vocal des locuteurs (Y. Chen & Gussenhoven, 2008). Les résultats généraux corroborent le parti-pris dans l'étude des affects que les mesures globales de la hauteur de la f0 sont un bon moyen d'évaluer les différences entre les affects et émotions, puisque les mesures globales, les mesures dans le temps et les mesures des cibles tonales présentent tous des tendances similaires (voir les résultats du Chapitre 2 et les mesures prises par Audibert et al. (2005); Banse et Scherer (1996); Berry et Brown (2019); Johnstone et Scherer (2000); Scherer (2003)). Les différences entre les valences neutre et positive ou neutre et négative sont les moins consistantes entre les deux énoncés. Il arrive que la différence ne se joue qu'à quelques Hertz de différence (par exemple, 4Hz) ; cette différence n'est alors pas audible.

Le troisième résultat majeur de notre étude correspond à la présence de différences d'alignement à la syllabe des cibles tonales en fonction de la valence de la nouvelle. Ces différences sont très peu nombreuses, et présentes uniquement au sein des marqueurs discursifs. Lorsqu'elles apparaissent, elles présentent beaucoup de variabilité. La différence significative la plus stable correspond au fait que les cibles tonales des accents mélodiques H* et L* sont réalisés plus tard lorsque la valence de la nouvelle à annoncer est neutre, comme nous pouvons le voir dans le Tableau 25. Nous avons trouvé dans notre Chapitre 2 que lorsque la valence de la nouvelle est marquée (positive, négative), il semble y avoir une certaine tendance au dévoisement de consonnes normalement voisées. De telles tendances auraient ainsi un effet sur la fréquence fondamentale, qui n'est plus présente sur les consonnes non-voisées et pourrait donc s'arrêter plus tôt sur certaines parties de la courbe. Cela pourrait nous avoir poussés lors de l'annotation à annoter les tons en bordure de dévoisements plus tôt, donc plus tôt lorsque la valence est négative ou positive par rapport à lorsqu'elle est neutre, ce qui expliquerait la différence trouvée.

Nous nous attardons maintenant quelque peu sur les résultats d'alignement trouvés en fonction de l'énoncé observé. Dans la préface, nous ne trouvons aucune différence significative d'alignement. Au sein des marqueurs discursifs, les effets sont trouvés lorsque nous observons les tons réalisés sur « voilà donc » ou « voilà », jamais sur « donc ». Nous trouvons que le ton H* est réalisé plus tôt lorsque la valence de la nouvelle

est positive, et plus tard lorsqu'elle est négative. Les cibles L* sont à l'inverse produites plus tôt pour le négatif par rapport au positif. Hi est produit plus tard lorsque la valence est positive par rapport à lorsqu'elle est neutre. Hi, qui n'a pas d'ancrage segmental précis (Welby, 2003, 2006), présente une forte déviation-standard d'alignement par rapport à leurs valeurs moyennes dans chaque valence. Cela est cohérent avec le fait qu'ils n'ont pas d'ancrage, et peuvent ainsi être réalisés à des points assez variés dans l'énoncé, et avec la littérature sur l'ancrage segmental et ce type de tons (Welby & Loevenbruck, 2005, 2006).

En conclusion, bien que nous n'ayons pas trouvé une grande quantité d'effets significatifs quant à la présence d'un patron intonatif particulier en fonction de la valence, notre étude démontre néanmoins que leur rôle dans la projection d'information affective mérite notre attention, puisqu'il semble pouvoir être porté à travers des indices phonologiques et leur implantation phonétique dans la parole. Il serait intéressant d'observer les patrons produits en fonction de l'affect dans des situations qui poussent les participants à produire une plus grande variété de patrons, s'éloignant ainsi des patrons standards pour les affirmations. Cela nous permettrait de voir si l'effet que nous avons trouvé s'étend à d'autres patrons lorsqu'ils sont produits en plus grande quantité. Nous aurions ainsi une meilleure appréhension du lien entre sens paralinguistique et patrons intonatifs, car cela nous permettrait de voir si les patrons préférés en fonction d'un affect sont toujours des patrons déjà décrits pour des fonctions linguistiques, où s'il existe des patrons qui ne seraient produits qu'en fonction d'un certain affect.

Une proposition additionnelle pour de futures recherches pourrait être d'observer les configurations qui viennent avant la fin de l'IP (généralement appelées « pré-nucléaires »), écartées de notre analyse des préfaces, plutôt que les configurations de fin d'IP. Ces configurations ont fait l'objet de moins d'attention et leur étude plus approfondie a été suggérée par certains auteurs (Delais-Roussarie & Herment, 2018; Petrone, 2008; Petrone & Niebuhr, 2014) pour la phonologie. Cette partie du patron intonatif est, en effet, considérée comme potentiellement moins chargée d'informations linguistiques comme nous l'avons évoqué au début de ce chapitre, car ne se situant pas à une frontière prosodique. Cependant, ces configurations précédant la fin d'IP pourraient se révéler peu informatives, si l'on se fie à nos résultats concernant « voilà donc » (car le résultat significatif trouvé est en configuration de fin d'IP) : il faudrait alors s'assurer d'avoir une

grande quantité de productions dans chaque configuration afin d'avoir une plus grande chance de trouver des effets.

La valeur mélodique des cibles tonales est apparue comme un indice constant de la différence entre valence positive et négative, avec des cibles réalisées plus haut pour le positif et plus bas pour le négatif. La direction de cet effet n'est jamais inversée. La différence d'alignement pour les cibles tonales des accents mélodiques semble n'être un indice constant de la valence affective que pour marquer la valence neutre par rapport aux deux autres. L'alignement présente en dehors de ce résultat des différences inconsistantes et peu nombreuses, donc peu fiables. Ces deux paramètres semblent néanmoins apporter des informations qui se complètent pour projeter la valence de la nouvelle qui va être annoncée. Comme Cao et al. (2014), nous suggérons que ToBI, pour rendre compte de la parole affective, bénéficierait de l'ajout d'un moyen de rendre compte de ces différences d'alignement, qui jouent un rôle central dans l'encodage des contrastes intonatifs (Prieto, 2012). Notre étude, si elle n'a pas démontré un rôle prépondérant de variations intonationnelles d'ordre phonologique dans la projection de la valence affective, n'exclut pas cette possibilité. Par ailleurs, l'étude démontre qu'à minima, des différences systématiques au niveau des cibles tonales apparaissent, influencées par le caractère négatif, positif, ou neutre de la nouvelle à annoncer. Ces différences ne peuvent donc être écartées et il serait utile de continuer à explorer le caractère informatif que pourraient avoir des différences localisées de hauteur par rapport à des différences globales, pour véhiculer les affects. L'idée d'une théorie métrique des différences de hauteur mélodique, bien que moins répandue que celle de la phonologie AM, a déjà été proposée par Ladd (1993), et pourrait être un moyen d'inclure les informations de valeur mélodique de façon phonologique pour décrire le sens. Les variations d'alignement semblent elles aussi devoir être prises en compte pour notre compréhension de l'intonation affective. Cependant, il serait nécessaire d'étendre ce type d'analyses à d'autres études, corpus, et langues, afin de mieux comprendre quelles variations sont systématiquement réalisées selon la valence affective, ou selon d'autres informations paralinguistiques.

Malgré ce que cette étude nous permet de démontrer, nous manquons d'informations sur l'intonation dans les préfaces, que nous n'avons pu annoter en entier. Par conséquent, il est possible que nous soyons passés à côté d'indices cruciaux de la valence dans les configurations pré-nucléaires de la préface, par exemple. De plus, la nature même de notre

corpus, avec des nouvelles dont la valence affective n'est jamais extrême, peut avoir influencé nos données. Il est possible que nous ayons pu trouver de plus fortes variations si nous avons utilisé un corpus démontrant des affects plus forts. Enfin, les résultats de notre analyse de l'accord inter-annotateurs dans la préface n'ont pas été concluants pour les accents mélodiques. Nous ne pouvons donc pas être certains de la fiabilité des données extraites en termes de valeur mélodique et d'alignement dans ces cas-là. Afin d'obtenir une plus grande assurance quant à nos données sur les accents mélodiques, il faudrait revoir les critères d'annotation de ceux-ci et refaire les analyses.

Chapitre 4 – Perception des préfaces aux annonces de nouvelles

1. Problématique

Nous allons maintenant présenter une étude perceptive qui a été faite sur les préfaces extraites du corpus d'annonces de nouvelles téléphoniques. L'objectif était de voir si la valence affective projetée dans la préface par le locuteur est utilisée par l'interlocuteur pour se préparer à l'annonce. Nous avons demandé à des auditeurs d'écouter les préfaces (présentées en isolement) et de choisir parmi trois choix possibles en fonction de leur réaction affective : une réponse qui dénote une réaction affective positive, une qui dénote une réaction affective négative, et une qui dénote une réaction affective neutre. L'objectif de cette consigne est de voir quelle est la réaction affective instinctive des auditeurs, en fonction de la valence de la nouvelle véhiculée par la prosodie dans la préface (que nous appellerons maintenant « valence de la nouvelle » par soucis de simplicité). Nous croyons en effet que si les auditeurs choisissent une réponse en fonction de leur réaction affective, ils choisiront la réponse qui correspond à la valence perçue pour la nouvelle à venir : c'est-à-dire par exemple, une réaction positive à l'écoute d'une préface perçue comme amenant une nouvelle positive. Nous avons mesuré la proportion de réponse en fonction de la valence de la nouvelle rattachée au stimulus écouté, et en fonction du genre du participant. Nous avons également observé le lien entre la congruence de la réponse et le temps de réaction des participants. La « congruence » entre un stimulus et une réponse arrive lorsque, par exemple, la préface utilisée en stimulus a été produite avant l'annonce d'une nouvelle positive (et porte donc des indices prosodiques de valence positive), et l'auditeur choisit une réponse qui dénote une réaction affective positive. Les temps de réaction ont été enregistrés afin de nous permettre un regard sur les processus cognitifs sous-jacent à la tâche (la compréhension de la prosodie comme positive, neutre ou négative, notamment). A la fin de l'étude de perception, nous avons proposé aux auditeurs de remplir un questionnaire d'empathie, afin de pouvoir explorer l'influence du niveau d'empathie d'un interlocuteur sur la temporalité de perception des indices de valence affective.

Nous avons souhaité observer la perception que l'on a d'une préface en fonction de sa prosodie uniquement, par rapport à la réaction affective élicitée par la préface. Par exemple, est-ce que la préface « *Je vous appelle au sujet de votre chat Minouche qui était*

malade », si elle a été produite avant d'annoncer que le chat est guéri, et est écoutée seule, élicitera une réaction affective plutôt positive de la part d'un auditeur naïf ? Suivant Swerts et Hirschberg (2010), nous souhaitons savoir si les indices de projection produits dans les préfaces à l'annonce d'une nouvelle sont récupérés par l'interlocuteur (ici, nos auditeurs), pour anticiper la valence affective de la nouvelle et se préparer à la recevoir sur le plan affectif. Dans leur étude sur le néerlandais, Swerts et Hirschberg (2010) avaient tronqué des portions de signal de leur corpus correspondant à l'ouverture de l'appel, et la préface (un exemple de leur corpus est présenté dans l'état de l'art de ce document). Les auditeurs devaient écouter et juger ces portions dans une tâche à choix forcé comme amenant la nouvelle de l'obtention d'un travail (nouvelle positive) ou non (nouvelle négative). Cette étude avait trouvé que les auditeurs sont capables d'inférer la nature positive ou négative d'une nouvelle à partir de variations prosodiques présentes dans la préface à l'annonce. Certains des stimuli avaient été tronqués de messages vocaux dans lesquels l'appelant faisait ensuite l'annonce de la nouvelle. D'autres avaient été tronqués de messages vocaux dans lesquels l'appelant n'annonçait pas la nouvelle, et demandait que l'interlocuteur rappelle pour connaître la réponse. Cette différence a permis de démontrer que la valence perçue en amont de l'annonce ne relève pas simplement de l'humeur actuelle du locuteur. Elle relève de la valence de la nouvelle qu'il doit annoncer : en effet, seuls les stimuli tronqués de messages où la nouvelle était ensuite réellement annoncée ont permis aux auditeurs de correctement anticiper la valence de la nouvelle, au-delà du niveau de la chance. Swerts et Hirschberg (2010) mettent en avant un questionnement quant au fait que des indices sur le contenu affectif de ce qui vient pourraient aider les auditeurs à se préparer mentalement au contenu des nouvelles. Cette « préparation mentale » renvoie pour nous tant à une préparation interactionnelle qu'à une préparation affective à la valence de la nouvelle. En anticipant la valence affective de la nouvelle qui va leur être annoncée, un interlocuteur pourra se préparer à répondre de manière alignée voire affiliée au locuteur. S'il anticipe une nouvelle positive par exemple, il pourra mieux adapter sa réponse en faisant une réponse positive. En faisant attention à s'aligner et s'affilier, l'interlocuteur préserve la face des interactants (la sienne, et celle du locuteur) mais aussi la progressivité de l'interaction et son potentiel succès.

Nous souhaitons regarder les aspects de cette préparation, et en particulier la préparation affective de l'interlocuteur. Pour analyser cela, nous avons aussi tronqué une

portion de l'annonce dans notre étude : la « préface ». Nous souhaitons en effet observer le rôle de ces énoncés en particulier.

Dans les études de perception de la prosodie (linguistique ou paralinguistique), les tâches d'identification telles que la nôtre sont très souvent combinées à l'enregistrement des temps de réaction (RT, de l'anglais « *reaction time* »). En effet, il a été montré (Pisoni & Tash, 1974) que le temps de réaction s'allonge lorsque l'identification est incertaine ou inconsistante, alors qu'il se raccourcit lorsque l'identification est certaine et consistante. Le temps de réaction peut donc fournir des informations précieuses sur la rapidité avec laquelle les affects sont identifiés et traités, car il est supposé refléter la difficulté de la tâche d'identification pour les participants (Massaro, 1987).

Nous allons maintenant décrire le raisonnement derrière les choix de réponse proposés aux participants dans la tâche à choix forcé : les feedbacks. Par ailleurs, puisque nous nous intéressons à la réaction affective des individus, nous explorerons l'influence que le genre et le niveau d'empathie pourraient avoir sur les réactions affectives. Nous nous attardons donc sur ces notions dans les sections suivante.

1.1. Notion de feedbacks

Nous avons établi, dans l'Etat de l'Art de ce document, le rôle des préfaces dans les annonces de nouvelles, et de façon générale le déroulé dans le temps de ces annonces. Nous avons établi notamment la différence entre les préfaces et les pré-séquences. Les préfaces permettent aux locuteurs de préparer l'annonce juste avant de la faire, sans laisser à l'interlocuteur le temps de faire une réponse. Les pré-séquences laissent à l'interlocuteur l'opportunité de faire une réponse avant l'annonce, en la devinant ou la refusant, par exemple. Nous nous concentrons sur la notion de préface, dans laquelle l'interlocuteur n'a donc pas l'opportunité de prendre un tour de parole à moins de la couper. Cependant, l'Analyse Conversationnelle et la Linguistique Interactionnelle ont démontré que lorsque l'on s'adresse à nous, il est courant que, sans pour autant couper la parole de l'autre, nous ayons des réactions orales ou gestuelles pendant que le locuteur est en train de parler. Par exemple, si le locuteur dit « Et tu vois là, je lui ai dit d'arrêter ! Il commençait à être agaçant ! », on peut imaginer que l'interlocuteur produise quelque chose comme « hm-hm », « ouais », « non » pendant que l'énoncé est fait (après « arrêter », par exemple). L'interlocuteur est alors en train de fournir un feedback au

locuteur qui raconte une histoire pour indiquer qu'il entend, comprend et exprime son accord ou son désaccord avec ce qui est dit. Il est donc possible que, sans pour autant couper la parole après la préface du locuteur, l'interlocuteur fasse un feedback, pendant ou après cette préface. Ces feedbacks sont régulièrement étudiés dans le cadre de l'analyse de l'interaction (Boudin et al., 2022; Prévot et al., 2013; Prévot & Bertrand, 2012), et particulièrement dans les interactions d'annonces de nouvelles médicales (Mastriani et al., 2018; Porhet et al., 2017).

Allwood et al. (1992) décrivent les feedbacks comme des mécanismes interactionnels qui permettent aux participants d'échanger des informations à propos de fonctions de communication de base sur ce qui est dit. Ils permettent de transmettre des informations par rapport au consentement et à la capacité que l'on a pour réaliser ces fonctions. Premièrement, faire un feedback permet de gérer le contact entre les interactants, c'est-à-dire qu'il permet de montrer que l'on veut et que l'on est capable de continuer une interaction. Deuxièmement, il permet de faire part du fait que l'on veut et que l'on est capable de percevoir le message véhiculé. Troisièmement, le feedback permet aussi de faire savoir que l'on est d'accord pour écouter, et capable de comprendre le message perçu. Quatrièmement, il permet de montrer que l'on veut et que l'on peut transmettre des informations attitudinales vis-à-vis du message, du locuteur, ou de ce qu'il dit. Ces quatre éléments sont cruciaux dans la communication, et donc dans l'annonce d'une nouvelle. Ils servent à montrer la capacité et la volonté à continuer l'interaction d'annonce, percevoir, écouter et comprendre la nouvelle, et enfin à transmettre l'attitude affective que l'on a vis-à-vis de la préface, l'annonce, et la nouvelle en elle-même.

Les feedbacks peuvent être « génériques » ou « spécifiques » (Bavelas et al., 2000). Les feedbacks génériques servent principalement à montrer que l'on est attentif et que l'on comprend ce qui est dit. Bertrand (2021) liste par exemple les feedbacks génériques verbaux en français tels que « mh », « oui », « d'accord », ou les feedbacks mimo-gestuels tels que les hochements de tête. A ceux-là s'opposent les feedbacks spécifiques, qui ajoutent du sens vis-à-vis de ce qui est dit par le locuteur. Ils servent à commenter et peuvent être caractérisés par des énoncés brefs, des gestes, des grimaces, ou une prosodie spécifique. Ils peuvent véhiculer des informations sur les affects de l'interlocuteur tels que leur *stance* ou leur état émotionnel (Bertrand, 2021), et peuvent donc être utilisés en réaction aux préfaces d'annonces de nouvelles en fonction de la valence des nouvelles.

Les feedbacks sont donc réalisés lorsque nous sommes dans la position de l'interlocuteur. La dimension à la fois interactionnelle et affective qu'ils peuvent porter les rend pertinents pour observer la réaction affective des interlocuteurs aux indices prosodiques de projection de l'affect véhiculés dans la préface. Porhet et al. (2017) ont par exemple étudié les feedbacks dans le cadre de l'entraînement des soignants à l'annonce de nouvelles et de la création d'agents virtuels de communication pour ces entraînements, ce qui démontre leur importance cruciale dans les interactions d'annonce de nouvelle pour aider au déroulé de l'annonce.

1.2. Perception des affects : facteurs de variabilité

Comme nous souhaitons observer la réaction affective des participants dans notre étude, il est apparu pertinent de nous intéresser aux facteurs individuels qui peuvent influencer la perception des affects. Un grand nombre d'études sur la perception des émotions et autres affects ont été réalisées. Elles ont permis de démontrer la capacité des êtres humains à décoder les émotions portées par la prosodie (Banse & Scherer, 1996; Johnstone & Scherer, 2000; Scherer, 1989). Elles ont aussi permis de démontrer que les affects ne sont pas reconnus aussi bien les uns que les autres selon l'individu.

Le premier facteur de variabilité auquel nous nous intéresserons dans l'étude de perception est celui du genre. Il a, en effet, été démontré dans plusieurs études qu'ils ne perçoivent pas la parole affective de la même façon. Schirmer et al. (2002, 2004) ont notamment réalisé deux études sur l'allemand pour observer les différences entre ces deux genres sur leur manière de traiter la parole affective. Particulièrement, elles ont regardé quels indices sont utilisés par les individus entre indices verbaux, et indices prosodiques. La première étude (Schirmer et al., 2002) s'est intéressée à l'influence de la prosodie (par ex., sur des phrases telles que « *Yesterday, she had her final exam* ») sur le traitement neuronal de mots positifs ou négatifs présentés de façon écrite (ex. « *success* » vs « *failure* »). Pour cela, les participants écoutaient par exemple une phrase prononcée avec une prosodie joyeuse, et devaient ensuite déterminer entre deux stimuli écrits lequel était un mot et lequel était un pseudo-mot. Il a été découvert que la prosodie module la réaction neuronale aux mots et le temps de réaction dans le choix des participants. La prosodie de l'énoncé facilitait la reconnaissance du mot lorsque la valence affective était

congruente avec celle du mot (ex, prosodie joyeuse + « *success* »). L'étude a démontré que l'influence de la prosodie affective n'était pas la même pour les hommes et pour les femmes. En effet, les hommes ont porté plus d'attention au contenu affectif contenu directement dans les mots. Les femmes en revanche, ont démontré par leur réaction neuronale et leurs temps de réaction qu'elles reconnaissaient les mots par rapport aux pseudo-mots plus rapidement que les hommes, en s'aidant plus de la prosodie affective. Schirmer et al. (2002) soulignent que cela indique que plutôt qu'une différence hommes-femmes traditionnelle où l'on considère les femmes comme plus « émotionnelles » et les hommes comme moins, la différence réside plutôt dans une manière différente de traiter les informations affectives. Les femmes seraient plus sensibles et réactives à l'affect porté par la prosodie. L'étude suivante (Schirmer et al., 2004) porte elle aussi sur les réactions neuronales (fMRI) à la parole affective en fonction du genre. Les stimuli présentés aux participants étaient des mots positifs et négatifs prononcés avec une prosodie joyeuse ou de colère. L'effet de la congruence ou de l'incongruence entre valeur affective du mot et valeur affective de la prosodie était observé sur la réaction neuronale des hommes et des femmes. L'étude a permis de démontrer que les femmes réagissaient plus vite à l'incongruence entre valence du mot et valence de la prosodie. Les hommes réagissaient aussi, mais plus lentement.

Les résultats de ces études ont par la suite été corroborés par Lausen et Schacht (2018) qui ont démontré qu'à l'écoute de pseudo-énoncés et pseudo-phrases portant une prosodie affective, les femmes ont obtenu de meilleures performances de reconnaissance que les hommes. De même Beaucousin et al. (2011) ont corroboré les résultats de (Schirmer et al., 2002, 2004) dans une étude sur l'activité cérébrale déclenchée par l'écoute de parole affective. Ils ont trouvé que bien que les hommes et les femmes présentent des fonctionnements similaires dans l'activité cérébrale, des différences apparaissent au sein de ces fonctionnements. Les hommes ont plus utilisé les indices lexicaux pour comprendre la parole affective que les femmes. Nous observerons donc aussi la différence de perception des préfaces de notre étude en fonction du genre des participants.

Barrett (1998) s'est intéressé aux différences individuelles de perception et à l'origine de ces différences dans une étude sur l'intéroception vis-à-vis de l'affect. L'intéroception correspond à la capacité d'un individu à percevoir et comprendre les signaux internes de

son propre corps. Barret (1998) a trouvé que selon la manière dont l'individu perçoit ses propres émotions, il s'attardera sur différents aspects de l'affect. Certaines personnes sont plus focalisées sur le caractère plaisant ou non des affects, et feront donc plus attention à l'axe de la valence. D'autres personnes sont plus focalisées sur les sensations internes ressenties et feront donc plus attention à l'axe de l'excitation. Dans une étude plus récente, Papousek et al. (2014) ont observé l'influence des différences individuelles de perception des affects sur la réponse d'un individu face à l'humour. Ils ont pour cela présenté des séquences de dessin animé à des participants, et les ont fait remplir plusieurs échelles ou sous-échelles permettant de comprendre la perception individuelle des affects. La première échelle est la « *Self-report emotional ability scale* » qui évalue la capacité à percevoir les affects des autres, et la capacité à réguler nos propres affects. La seconde est le « *Mayer-Salovey-Caruso emotional intelligence test* », qui évalue la perception et la gestion des affects. La dernière échelle est la « *Center for Epidemiologic Studies depression scale* », qui évalue l'humeur et donne une indication sur la dépression. L'étude a démontré que les différences de score sur ces échelles étaient corrélées avec des réponses affectives différentes à l'humour. La perception des affects semble donc sujette à de nombreux facteurs de variation liés à la compréhension, la gestion et la perception individuelle que l'on en a. Les individus ne semblent notamment pas être tous aussi sensibles à la valence les uns que les autres, ce qui pourrait participer à de la variation inter-individuelle chez les auditeurs dans la perception des préfaces de notre étude.

Certaines études s'intéressent à un autre facteur de variabilité dans la perception des affects : l'empathie. Bien que régulièrement confondue avec les notions présentées plus haut (intéroception, habilité émotionnelle, intelligence émotionnelle) dans la littérature (Olderbak & Wilhelm, 2017), elle ne concerne pas tout à fait le même phénomène. L'empathie joue un rôle fondamental dans notre compréhension des émotions et des interactions humaines. C'est une notion complexe et multidimensionnelle, sujette à de nombreuses controverses. Lawrence et al. (2004) décrivent deux pans de l'empathie. D'une part, l'empathie cognitive correspondrait à la capacité intellectuelle à appréhender les affects d'autrui. D'autre part, l'empathie affective correspondrait à la réaction affective qu'un individu aura face aux ressentis d'une autre personne. Des études plus récentes ont proposé d'autres catégorisations des types d'empathie. Notamment, une catégorisation en trois types d'empathie existe et c'est à celle-ci que nous nous intéressons, car elle est utilisée dans le test d'empathie que nous avons choisi. Carre et al. (2013) listent, tout

d'abord, la contagion émotionnelle. Cela correspondrait au fait de reproduire automatiquement l'émotion d'autrui pour soi-même. Puis, il y a l'empathie cognitive, qui comme dans la définition précédente, correspond à la capacité à comprendre l'émotion d'autrui. La dernière catégorie est la déconnexion émotionnelle, c'est-à-dire le fait de savoir se protéger du stress, de la douleur, et de ce qui peut fortement impacter l'affect.

La variété dans la compréhension et les catégorisations de l'empathie a donné naissance à de nombreux tests d'empathie, qui ont été proposés pour tenter de mesurer le plus précisément possible l'empathie et ses sous-catégories. Parmi les tests qui ont été développés pour évaluer cette capacité chez les individus, on retrouve, par exemple : l'évaluation du Quotient d'Empathie (EQ) (Baron-Cohen et Wheelwright, 2004), l'Index de Réactivité Interpersonnelle (Keaton et al., 2017), l'Echelle Emotionnelle Multidimensionnelle d'Empathie (Caruso & Mayer, 1998), ou encore le Basic Empathy Scale (BES) (Jolliffe & Farrington, 2006). C'est à ce dernier que nous nous intéressons ici, car il permet une évaluation du niveau d'empathie à la fois de façon précise (car il évalue la contagion, la déconnexion, et l'empathie cognitive) et de façon globale en prenant en compte tous ces facteurs comme faisant partie de l'empathie. L'utilisation de ce test nous permettra d'explorer les effets que pourraient avoir les différences individuelles de niveau d'empathie sur la perception des affects portés par la prosodie dans notre étude. Il a été démontré dans différentes études utilisant le test d'empathie BES ou l'une de ses variantes que les hommes et les femmes n'obtiennent pas le même type de score sur le test. Les femmes et les hommes présentent des différences dans la manière dont ils sont empathiques (Carre et al., 2013; D'Ambrosio et al., 2009; Jolliffe & Farrington, 2006). L'empathie joue donc un rôle important dans la compréhension qu'un individu peut avoir des affects d'autrui, et il apparaît que des différences de genre peuvent jouer un rôle sur le niveau d'empathie de chacun. Ces différences dans les scores d'empathie sont soutenues par des différences de genre dans les processus neuronaux engagés pour évaluer la réaction affective personnelle à l'affect des autres personnes (Schulte-Rüther et al., 2008).

Aziz-Zadeh et al. (2010) ce sont intéressés à l'activité des aires du cerveau (fMRI) en fonction de la perception de la prosodie, de sa production, et en lien avec l'empathie et la « capacité prosodique ». Dans cette étude, il a été demandé aux participants de produire ou d'écouter des phrases ne portant aucun sens avec différentes intonations pendant que leur activité neuronale était enregistrée. L'activité enregistrée a ensuite été corrélée avec

des scores d'empathie des participants. Cette corrélation a permis de mettre en lumière le fait que les individus avec une empathie plus haute démontraient un plus fort taux d'activité dans certaines régions du cerveau que ce qui est normalement trouvé en moyenne pour la perception de la prosodie. Elle met aussi en avant le fait que les individus avec une empathie haute avaient une activité neuronale moins forte dans ces régions lorsque la prosodie perçue était non-affective comparé à lorsque la prosodie était affective. Aziz-Zadeh et al. (2010) interprètent ce résultat comme en ligne avec l'idée que certains aspects de l'empathie liés aux stimuli affectifs se basent sur un procédé de simulation des affects lié au système des neurones miroirs.

Dans une étude multimodale concentrée sur la perception de la prosodie neutre, Regenbogen et al. (2012) ont évalué, entre autres, l'influence de l'utilisation d'une prosodie neutre sur la perception d'indices affectifs afin d'avoir une idée de l'influence de la prosodie sur l'empathie. Pour cela ils ont montré à des participants des vidéos d'histoires courtes racontées par des acteurs tout en enregistrant leur réaction électrodermale. L'une des conditions présentait les histoires racontées avec une prosodie neutre, alors que les indices faciaux véhiculaient toujours des informations affectives. Ils ont démontré que la prosodie neutre élicitait moins de réaction affective empathique de la part des participants.

1.3. Hypothèses

Suivant la littérature (Swerts & Hirschberg, 2010), nous pensons que les participants à l'étude seront capables de déceler les indices de projection prosodique. Ainsi, ils devraient reconnaître quelles préfaces ont été produites avant une annonce de nouvelle positive, négative, ou neutre. Nous faisons l'hypothèse qu'en conséquence de cette reconnaissance, ils choisiront, parmi les feedbacks proposés, celui qui sera congruent avec la valence affective de la nouvelle. Ainsi, si la préface écoutée a été produite avant l'annonce d'une nouvelle positive, nous faisons l'hypothèse que l'auditeur, reconnaissant les indices de projection prosodiques de la valence, aura une réaction affective positive. En conséquence, il choisira comme réponse dans la tâche à choix forcé le feedback qui porte une valence positive. De même si la valence de la nouvelle était négative, nous pensons que l'écoute de la préface à une nouvelle négative créera une réaction affective négative chez l'auditeur qui choisira le feedback négatif. Enfin, si la valence de la nouvelle était

neutre, nous faisons l'hypothèse que l'auditeur aura une réaction affective neutre et choisira le feedback neutre.

L'étude de perception a été divisée en trois tâches qui opposent les valences deux à deux. Dans chaque tâche, seuls deux feedbacks sont proposés comme réponses : ceux qui correspondent aux valences évaluées. Nous avons des hypothèses spécifiques pour chaque tâche. La tâche 1 comporte des préfaces prononcées avant l'annonce d'une nouvelle positive ou négative. Nous croyons que la tâche 1 sera celle où les valences auront le plus de chances d'être reconnues, et où en conséquence les feedbacks affectifs élicités seront le plus souvent congruents avec la valence entendue, démontrant une préparation affective régulière. En effet, les valences positive et négative sont celles qui présentent le plus de différences significatives et des différences plus marquées au niveau acoustique, d'après notre étude de production. La tâche 2 comporte des préfaces prononcées avant l'annonce d'une nouvelle négative ou neutre. Nous nous attendons à ce que les valences puissent être plus confondues que dans la tâche 1, notamment car Audibert et al. (2005) ont démontré que la prosodie de la tristesse (que nous rapprochons du négatif) est facilement confondue avec celle du neutre. Il est possible que les préfaces neutres élicitent plus de réactions affectives congruentes, car si les participants doutent de leur perception nous pouvons nous attendre à ce que leur réaction affective soit neutre plus que négative. Par ailleurs, notre étude phonétique a démontré la présence de plusieurs différences acoustiques dans les préfaces entre ces deux valences de nouvelles. Nous devrions donc voir apparaître des différences perceptives. La tâche 3 comporte des préfaces prononcées avant l'annonce d'une nouvelle négative ou neutre. Nous nous attendons à ce que cette tâche démontre le moins de préparation affective, les auditeurs ayant plus de risques de mélanger les deux valences. En effet acoustiquement, les valences positives et neutres avaient élicité des productions similaires sur la préface, avec des effets plus faibles que pour les autres valences. Ainsi, les valences positives et neutres pourraient éliciter moins de réponses affectives congruentes.

Nous pensons que lorsque les participants choisiront un feedback congruent avec la valence de la nouvelle, le temps de réaction sera plus court que si les valences sont incongruentes. En effet, nous pensons par exemple que les indices de projection d'une valence, lorsqu'ils sont perçus, font un effet de « *priming* » pour la réponse qui porte la même valence. En effet, Pell et al. (2022) ont par exemple démontré que dans une tâche de reconnaissance d'affects sur les expressions faciales, l'écoute d'un audio en amont, s'il

porte une valence affective congruente avec la prosodie, facilite la reconnaissance de l'expression faciale. Ainsi, nous pensons que la valence affective portée dans la préface facilitera la vitesse de réaction lorsque le feedback sélectionné par l'auditeur portera une valence en accord avec celle de la prosodie entendue. En revanche, si l'auditeur choisit la réponse incongruente, le temps de réaction ne sera pas facilité et sera donc plus lent (Hommel, 1997). Nous notons cependant que les découvertes de Pell et al. (2022) relèvent du domaine visuel et de la prédiction multimodale. Il est donc possible que les stimuli auditifs que nous utilisons ne présentent pas exactement le même aspect « facilitatoire » et de priming pour le choix des feedbacks de notre étude.

Nous pouvons nous attendre à ce que vis-à-vis du genre uniquement, les femmes aient des temps de réaction plus rapides et les hommes plus lents, car les préfaces que nous ferons écouter comporteront uniquement des indices prosodiques de la valence, et non lexicaux (Beaucousin et al., 2011; Besson et al., 2002; Lausen & Schacht, 2018; Schirmer et al., 2002, 2004). En revanche, nous ne pensons pas que le genre impactera la proportion des réponses selon la valence.

Nous faisons par ailleurs l'hypothèse que les personnes qui auront obtenu un haut score d'empathie auront une meilleure reconnaissance de la prosodie émotionnelle et donc de la valence dans les préfaces. De même, un temps de réaction plus court pourrait être observé (Aziz-Zadeh et al., 2010). Il est possible que les scores d'empathie puissent être corrélés avec ceux des hommes et des femmes, puisque les deux genres présentent des différences dans leurs scores d'empathie (Carre et al., 2013; D'Ambrosio et al., 2009; Jolliffe & Farrington, 2006). En regardant à la fois le genre et l'empathie, nous pourrions explorer le lien que ces deux facteurs entretiennent.

2. Tâche de perception des indices de projection affective

2.1. Matériel

Nous présentons dans cette section la construction du matériel utilisé pour notre étude de perception : les stimuli audios utilisés et les feedbacks proposés comme choix de réponse écrite à ces stimuli. Nous avons dans un premier temps sélectionné le type de réponses qui sont proposées en choix forcé aux participants, dans une tâche de validation présentée comme la « phase de construction 1 ». Nous devions en effet nous assurer que pour chaque valence de nouvelle incluse dans notre étude (positive, négative, neutre), nous proposerions un feedback portant une valence affective identique (positive, négative, neutre). Puis, nous présenterons le matériel utilisé dans l'étude principale (phase de construction 2).

2.1.1. Phase de construction 1

Le but de cette première étape de construction était d'obtenir trois feedbacks qui seraient proposés comme choix de réponse à l'écoute des préfaces, et à choisir en fonction de la réaction affective de l'auditeur : un feedback à valence positive, un feedback à valence négative, un feedback à valence neutre. Les trois feedbacks sélectionnés sont le résultat d'une étude préliminaire constituée de 3 tâches de décision à choix forcé, faites dans une première phase de construction de l'étude. La tâche 1 nous a permis de vérifier que le feedback positif et le feedback négatif sont bien perçus comme distincts. La tâche 2 portait sur le feedback négatif et le neutre, et la 3 sur le feedback positif et le neutre. Nous allons commencer par décrire cette première phase, qui nous a permis de valider « ah oui » comme positif, « oh non » comme négatif, et « j'écoute » comme neutre.

Cent trente-six participants de langue maternelle française ont pris part à cette étude préliminaire (77 femmes, 42 hommes, 3 non binaires, âge moyen de 25,8 ans, SD = 4,2 ans). Le niveau moyen d'études supérieures (niveau 4 de la Cife, ou Classification internationale type de l'éducation) était de 3,7 ans (SD = 1,7). Les participants, recrutés via réseaux sociaux, n'avaient pas de troubles de la vue ou de la lecture. Ils n'ont pas été

rémunérés, et ont validé un formulaire de consentement en ligne. 43 participants ont fait la première tâche (qui oppose des préfaces écrites négatives et positives), 44 la deuxième (qui oppose des préfaces écrites négatives et neutres) et 49 la troisième (qui oppose des préfaces écrites positives et neutres). Les tâches étaient toutes trois coupées en 2 listes. Ainsi, chaque liste a été évaluée par au moins 21 participants. Les participants ont lu 12 préfaces écrites qui contiennent un indice verbal de la valence de la nouvelle : ex. « J'ai une bonne nouvelle au sujet de vos examens médicaux ». Le but d'avoir les préfaces écrites était de nous assurer que la valence des nouvelles serait explicite et d'écartier ainsi toute ambiguïté sur celle-ci. Une schématisation de la tâche est présentée en Figure 83.

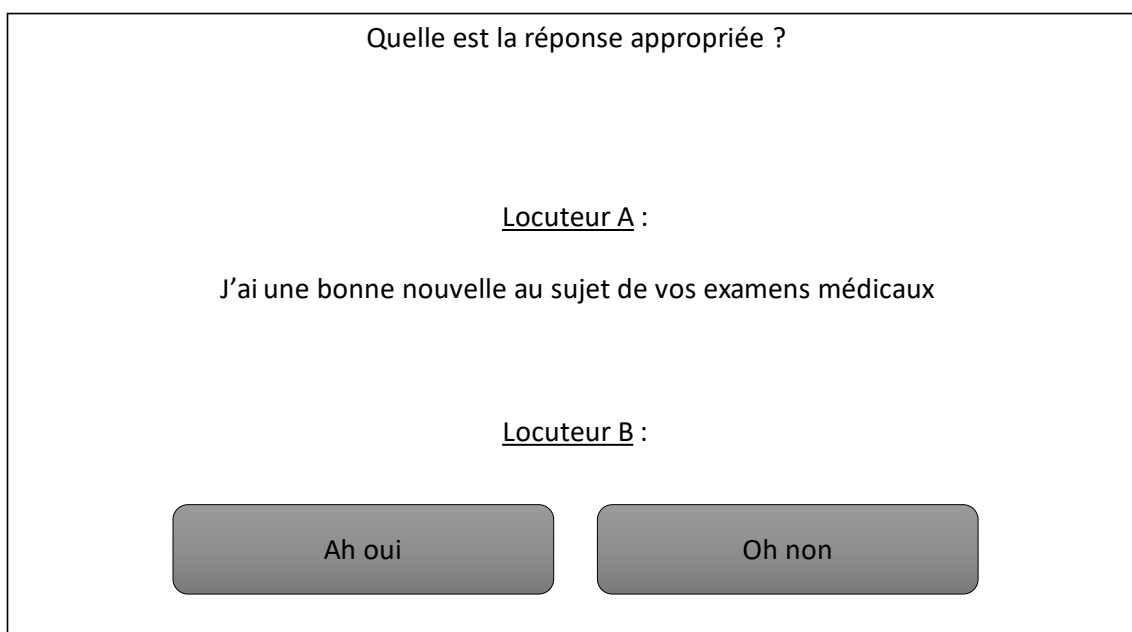


Figure 83 – Illustration du design expérimental de la tâche de validation

Chaque stimulus est inspiré du corpus d'annonces, l'ensemble des stimuli étant proposé en Annexe 15. L'instruction était de lire chaque préface et de choisir la réponse appropriée (parmi deux feedbacks écrits proposés). Pour la tâche 1, qui présentait des valences de nouvelles positives et négatives, les feedbacks proposés étaient « ah oui », « oh non ». Pour la tâche 2, qui présentait des valences de nouvelles négatives et neutres, les feedbacks proposés étaient « oh non », « j'écoute ». Pour la tâche 3, qui présentait des valences de nouvelles positives et neutres, les feedbacks proposés étaient « ah oui », « j'écoute ». Chaque tâche suit la structure : 2 listes contenant 12 scénarios en 2 valences (24 items), divisées en 2 blocs (alternant l'ordre des valences et réponses). Chaque

participant ne faisait qu'une tâche et lisait 72 items seulement (7 minutes) afin de limiter les abandons (Hoerger, 2010).

Nous avons réalisé un modèle logit (GLMER) pour chaque tâche, testant l'effet de la VALENCE (positive, négative, neutre) sur le CHOIX du feedback. La présence d'un choix était marquée 1, son absence (donc présence de l'autre choix) marquée 0. Les modèles initiaux comportaient deux interceptes aléatoires (PARTICIPANT et ITEM), et deux pentes aléatoires (VALENCE par ITEM et VALENCE par AUDITEUR). Ils ont ensuite été enlevés ou conservés dans le modèle ; pour cela nous avons effectué des comparaisons de modèles avec le test du rapport de vraisemblance par élimination progressive en utilisant la fonction *Anova()* de R. Les modèles finaux comprennent deux interceptes aléatoires : ITEM et PARTICIPANT. Les modèles montrent une augmentation significative du choix de « ah oui » (94%) par rapport à « oh non » (6%) lorsque la valence de la nouvelle est positive [$\beta = 4.6$, SE = 0.002, $z = 1946$, $p < 0.0001$] par rapport à lorsqu'elle est négative (« ah oui » 19%, « oh non » 81%) dans la tâche 1. Nous trouvons une augmentation significative du choix de « j'écoute » (88%) par rapport à « oh non » (18%) lorsque la valence de la nouvelle est neutre [$\beta = 2.5$, SE = 0.1, $z = 6.32$, $p < 0.0001$] par rapport à lorsqu'elle est négative (« oh non » 48%, « j'écoute » 52%) dans la tâche 2. Enfin, nous trouvons une augmentation significative du choix de « j'écoute » (69%) par rapport à « ah oui » (31%) lorsque la valence de la nouvelle est neutre [$\beta = 0.8$, SE = 0.1, $z = 13.24$, $p < 0.0001$] par rapport à lorsqu'elle est positive (« ah oui » 48%, « j'écoute » 52%) dans la tâche 3. Ces résultats confirment nos hypothèses de départ et valident l'utilisation des feedbacks dans l'étude principale : « ah oui » porte une valence positive, « j'écoute » une valence neutre et « oh non » une valence négative.

2.1.2. Phase de construction 2

Après avoir validé les feedbacks comme étant rattachés aux valences positive, négative et neutre, nous avons construit l'étude principale de perception. Dans cette étude, les préfaces étaient les préfaces originales extraites dans notre corpus acoustique (donc sans contenu verbal pour projeter la valence affective de la nouvelle). Les préfaces étaient présentées à l'écoute de façon isolée pour évaluer le rôle de la prosodie affective sur la réaction affective des auditeurs. L'étude principale était articulée en trois tâches d'identification à deux choix. La tâche 1 présentait des préfaces produites avant l'annonce de nouvelles positives ou négatives. En conséquence, les deux choix proposés étaient :

« ah oui », « oh non ». La tâche 2 présentait des préfaces produites avant l'annonce de nouvelles négatives ou neutres. Les choix proposés étaient : « oh non », « j'écoute ». La tâche 3 présentait des préfaces produites avant l'annonce de nouvelles positives ou neutres. Les choix proposés étaient : « ah oui », « j'écoute ». Cette division nous permet d'opposer deux à deux les valences, et de modéliser ainsi de façon plus simple et claire les différences de réaction affective et de temps de réaction élicités. Les participants devaient faire un choix entre les deux options de réponse proposées après chaque écoute. Le temps de réaction était pris en compte en plus de la décision.

Les extraits utilisés en stimuli correspondent à la partie de la préface uniquement : « Je vous appelle au sujet de vos examens médicaux ». Nous utilisons comme stimuli chacun des 12 scénarios de nouvelles. Les marqueurs discursifs ont été exclus des audios afin de ne conserver que la partie de la préface. Ainsi, nous observons le rôle anticipatoire de la prosodie dans les préfaces d'annonce spécifiquement. C'est par ailleurs dans la préface que les indices prosodiques de la valence étaient les plus consistants, d'après les analyses rapportées dans les précédents chapitres. Nous rappelons que ces préfaces ne contiennent aucun indice verbal sur la valence affective de la nouvelle qui suit. Toutes ont été produites avant des nouvelles positives, négatives, et neutres. D'après nos analyses des chapitres 2 et 3, ces préfaces contiennent de la variation prosodique en fonction de la valence de la nouvelle. Cette variation correspond, selon nous, à la projection d'indices de valence affective, faite par le locuteur pour préparer l'interlocuteur à recevoir l'annonce.

Pour réaliser l'étude, le logiciel E-Prime 3.0 (Psychology Software Tools, 2020) a été choisi. La tâche était réalisée dans une salle informatique du Laboratoire Parole et Langage. Le logiciel permet de mesurer les temps de réaction en même temps que les décisions des participants. Le décompte du temps de réaction démarre à l'*onset* de l'audio. Pour pouvoir faire cette mesure, nous avons choisi le logiciel e-prime plutôt qu'une étude en ligne, car nous n'avions pas d'idée de la taille des effets que nous allions trouver et avons donc choisi un mode de passation plus fiable face au bruit et aux effets de petite taille. E-prime est, en effet, reconnu comme l'un des logiciels les plus précis et performants pour la mesure des temps de réaction.

A cette partie centrale de l'étude, nous avons ajouté la Basic Empathy Scale (BES, voir Annexe 16), (Jolliffe & Farrington, 2006). Ce test évalue en 20 questions l'empathie. Les participants répondent aux questions posées sur une échelle de Lickert à 7 points allant de « pas du tout d'accord » à « tout à fait d'accord » avec la phrase qui est proposée. Le

test a, par ailleurs, été validé pour une utilisation en français (D'Ambrosio et al., 2009). Il a été adapté plus tard sous le nom "BES-A" (Carre et al., 2013), afin de l'utiliser avec une population adulte, le BES ayant initialement été mis en place pour les populations adolescentes. Dans cette version « BES-A », que nous utiliserons dans notre étude, nous prendrons la note « générale » d'empathie englobant les trois sous-tests (empathie cognitive, contagion émotionnelle, déconnexion émotionnelle), sur un score allant de 0 à 100 (0 = le minimum ; 100 = le maximum). Grâce à ce score général, nous ferons la différence entre les personnes à l'empathie haute, et les personnes à l'empathie basse. Pour déterminer les seuils d'empathie haute et basse, nous suivons Esteve-Gibert et al. (2020), et avons donc pris la médiane des scores sur l'ensemble des 110 participants, à travers les 3 études. La médiane est à 62. De 47 (minimum) à 61, le niveau d'empathie est bas. De 62 à 75 (maximum), il est haut.

2.1. Participants

Cent-dix participants de langue maternelle française ont pris part aux tâches de perception (dont 66 femmes, 42 hommes, et 2 non-binaires, âge moyen de 20,8 ans, SD = 2,5). 38 personnes ont participé à la tâche 1 (dont 22 femmes, 14 hommes, et 2 non-binaires). 36 personnes ont participé à la tâche 2 (dont 22 femmes et 14 hommes). 36 personnes ont participé à la tâche 3 (dont 22 femmes et 14 hommes). Aucun participant n'avait de trouble de la vue non corrigé, ou de troubles de l'audition. Nous avons accepté les personnes gauchères dans l'étude, en faisant attention de contrebalancer leur présence entre les listes : nous avons 2 gauchers par liste, 1 homme et une femme à chaque fois. Les participants étaient rémunérés 10 euros en tickets cadeaux, pour 20 minutes. Nous avons calculé les années d'éducation à partir des années effectuées après le niveau 3 d'étude, d'après la Classification internationale type de l'éducation (Cite). Le niveau 3 correspond en France au niveau du baccalauréat. Le niveau 4 correspond de façon internationale à l'enseignement supérieur. Le nombre d'années moyen d'études supérieures était de 2,3 ans (SD = 1,6). Le recrutement a été réalisé via les réseaux sociaux, le bouche à oreille, et la distribution de flyers. Les participants ont signé un formulaire de consentement pour l'étude. L'étude de perception présentée ici a été validée avec celle de production par le comité éthique d'Aix-Marseille Université (référence : 2021-01-07-02).

2.2. Procédure

Chacune des trois tâches était divisée en deux listes pour contrebalancer l'ordre d'apparition des feedbacks à l'écran. Dans la première liste, si une option de réponse (ex., « ah oui ») est présentée à gauche et l'autre (ex., « oh non ») à droite, alors elles étaient inversées dans la seconde liste. Au sein de chacune des listes, il y avait deux occurrences de chacune des douze préfaces. Une occurrence dans une valence affective (ex, préface dite avant la nouvelle négative), et une seconde occurrence dans l'autre valence affective présente dans l'étude (ex, préface dite avant la nouvelle positive). Les listes sont divisées en 2 blocs dans lesquels la valence des préfaces s'inverse : une préface proposée dans sa valence positive en bloc 1 sera présentée dans sa valence négative en bloc 2 (et *vice-versa*). Chaque liste a été présentée à au moins 18 participants.

Au sein des trois tâches, nous avons dû faire un choix quant à l'item qui serait entendu pour chaque préface. Le corpus ayant été enregistré par 30 locutrices, nous ne pouvions faire écouter chaque audio à tous les participants : l'écoute de 1080 audios aurait constitué une tâche trop longue pour les auditeurs, dont l'attention se serait amenuisée au fil de la tâche. Nous avons donc sélectionné certaines des voix des 30 locutrices, et avons fait varier la voix qui était entendue selon la tâche et la liste. Dans chaque liste, il y avait 4 voix différentes : 2 par valence émotionnelle. Ces voix étaient randomisées sur les items. Les 4 voix étaient différentes pour chaque liste : nous avons donc utilisé 24 voix de locutrices différentes (4 voix x 3 tâches x 2 listes). Puisque 6 voix allaient être écartées, nous avons sélectionné les voix où les locutrices semblaient avoir, subjectivement, mieux réussi l'enregistrement (peu d'hésitations liées à la lecture, pas d'impression que le texte était lu sans que la locutrice ne se soit mis à la place de quelqu'un qui fait une réelle annonce de nouvelle). La durée de passation de la tâche d'écoute et de choix forcé pour un participant était de 10 minutes. Une schématisation de la tâche est présentée en Figure 84.

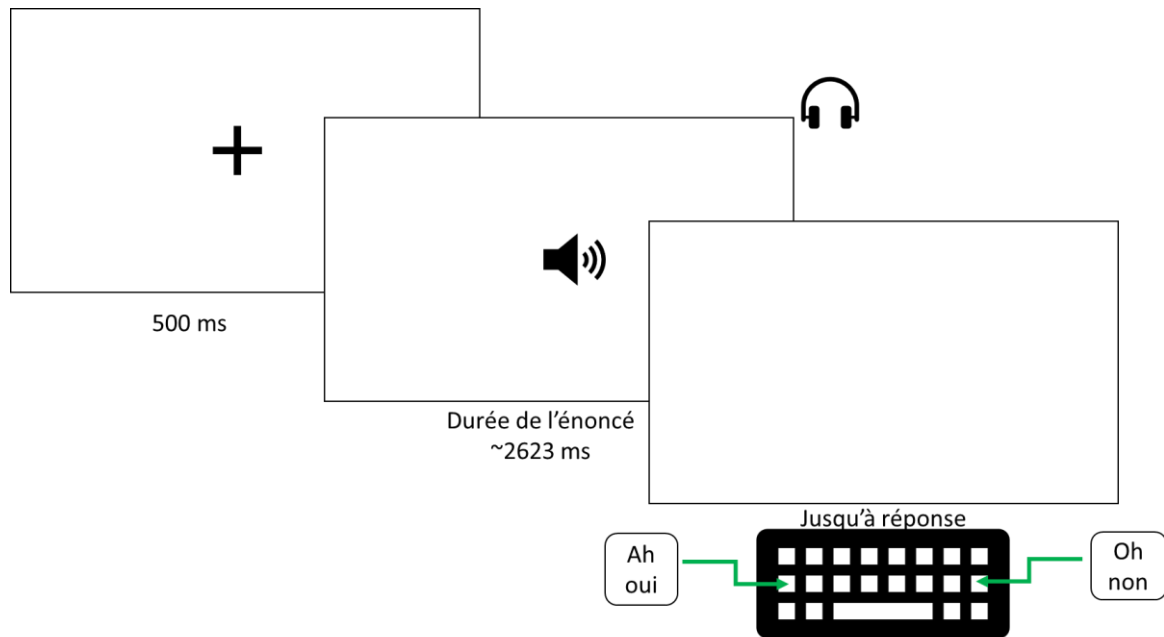


Figure 84 – Illustration du design expérimental de la tâche principale

Pour chaque tâche, il était demandé aux participants de sélectionner l'option de réponse qui était la plus proche de leur réaction émotionnelle parmi les deux feedbacks proposés. Il était demandé aux participants de ne pas se préoccuper de s'ils auraient réellement prononcé la réponse à l'oral, simplement de choisir celle qui était proche de leur réaction « émotionnelle » (affective). Le but de cette consigne et des choix de feedbacks proposés était de fournir aux auditeurs qui participaient à l'étude un cadre interactionnel aux stimuli audios qu'ils entendaient, afin d'observer leur réaction affective dans ce contexte. A l'oral, nous précisions notre attente d'un choix fait non pas en fonction de ce que l'auditeur aurait répondu à l'oral, mais de ce qui se rapproche le plus de leur ressenti.

En proposant des feedbacks tels que « ah oui » et « oh non » (valence marquée) et « j'écoute » (neutre), nous avons cherché à reproduire un contexte interactionnel afin qu'ils puissent mieux s'identifier à la personne qui reçoit l'annonce de nouvelle et afin qu'ils puissent récupérer les indices de projection et anticiper la nouvelle. Le choix de cette consigne a été fait après deux pilotes. Le premier pilote, réalisé informellement sur trois personnes du laboratoire, a testé la compréhension de la consigne qui a été revue, car floue et trop courte. Puis nous avons commencé la tâche avec 14 participants avec pour consigne de choisir la « réponse appropriée ». Cela a poussé les participants à choisir la réponse en prenant en compte le contexte social et la politesse, ce qui n'est pas notre

but. Nous avons alors modifié les consignes pour qu'elles mentionnent précisément le ressenti affectif.

2.3. Analyses statistiques

Pour chacune des trois tâches de l'étude principale, nous avons réalisé deux modèles statistiques pour évaluer l'effet de la valence de la nouvelle sur les choix des participants (modèles logit) et sur les temps de réaction associés (modèles linéaires).

En premier lieu, nous avons réalisé une série de modèles mixtes logit (GLMER) basé sur les distributions binomiales, afin de regarder les CHOIX de réponse avec le système de « présence ou absence » dans chaque occurrence d'un énoncé. Les valeurs assignées sont « 1 » pour le positif et « 0 » pour le négatif. Le neutre change en conséquence. Nous avons regardé que ces réponses faites en fonction de la valence de la nouvelle avaient été choisies au-delà du niveau de la chance. Dans ces modèles, nous observons donc uniquement l'effet de la VALENCE (2 niveaux : négatif vs positif pour la tâche 1 par exemple) sur le CHOIX de la réponse. Notre modèle maximal comportait cinq éléments spécifiés aléatoires. Tout d'abord, pour rendre compte de la variation inter-items, nous avons inclus un intercepte aléatoire ITEM à douze niveaux (les 12 scénarios du corpus). Pour prendre en compte la variation par qui peut être introduite par la variabilité individuelle, nous avons inclus un intercepte aléatoire AUDITEUR (à 38 ou 36 niveaux selon la tâche). Enfin, pour prendre en compte la variance qui pourrait être causée par les différences de voix utilisées dans les stimuli audio, nous avons ajouté un intercepte aléatoire VOIX (à 8 niveaux). Nous avons aussi ajouté deux pentes aléatoires : VALENCE par ITEM et VALENCE par AUDITEUR, pour prendre mieux en compte les variations spécifiquement causées sur la perception par valence par ces facteurs. Nous avons ensuite effectué des comparaisons de modèles avec le test du rapport de vraisemblance par élimination progressive en utilisant la fonction *Anova()* de R (Pinheiro & Bates, 2000). Le modèle final ne contient que la pente VALENCE par AUDITEUR. L'observation de la significativité de l'intercepte du modèle nous indiquera à chaque fois si pour la valence utilisée en intercepte, un choix de réponse a été fait au niveau de la chance, ou au-delà du niveau de la chance (sous, ou sur le niveau). Cette information combinée à l'observation des graphiques nous permettra de déterminer dans quels cas la préface a élicité chez l'interlocuteur une réaction affective significativement au-dessus du niveau de la chance selon la valence de la nouvelle. Nous

avons, pour chaque modèle, opéré un changement de l'intercepte sur les données avec la fonction *relevel*, avec dans un premier temps une valence comme intercepte, puis la seconde valence après *relevel*. Cela nous permet de comparer la significativité de l'intercepte pour les deux valences présentes dans le modèle. Pour cette raison, nous avons corrigé le niveau alpha de la *p-value* de 0.05 à 0.025 (0.05/2 modèles exécutés par chaque variable dépendante).

Modèle GLMER utilisé pour les analyses du niveau de la chance :

glmer(data = tâche1, Choix~Valence +(1+Valence|Auditeur), family="binomial")

Pour les choix des participants, nous avons également fait une seconde série de modèles mixtes logit. Nous avons testé l'effet de la VALENCE (2 niveaux : négatif vs positif pour la tâche 1 par exemple) et du GENRE (2 niveaux : homme vs femme) sur le CHOIX de la réponse. L'ajout de la variable dépendante GENRE a été validé par un test de vraisemblance avec la fonction *anova()*, qui nous a permis de comparer un modèle sans cette variable, ou avec. Le test a révélé que cette variable faisait une différence significative pour la tâche 1, mais pas pour les deux autres. Par soucis de cohérence dans l'analyse des trois tâches, nous avons décidé de garder cette variable dans les tâches 2 et 3 également. Les données des participants non-binaires n'ont pu être incluses dans la comparaison statistique par genre du fait de leur petit nombre (2). Nous sommes partis du même modèle maximal avec cinq éléments spécifiés aléatoires identiques à ceux du modèle plus « simple ». Dans le modèle final, la pente aléatoire AUDITEUR par VALENCE a été incluse.

Modèle GLMER utilisé pour les analyses du choix :

*glmer(data = tâche1, Choix~Valence*Genre +(1+Valence|Auditeur), family="binomial")*

La deuxième variable que nous avons modélisée est les temps de réaction (RT) associés aux choix des participants. Pour cela, nous avons effectué une série de modèles mixtes linéaires (LMER), pour tester l'effet de la VALENCE (négatif vs positif, négatif vs neutre, ou positif vs neutre), de la CONGRUENCE (congruent vs incongruent, c'est-à-dire si le participant a choisi une réponse dont la valence correspond à celle de la nouvelle) et du GENRE (homme vs femme) sur le temps de réaction (RT) en log du participant. Le modèle maximal incluait en intercepte aléatoires AUDITEUR, ITEM et VOIX, avec le même nombre de

niveaux respectivement que dans le modèle GLMER, pour prendre en compte les variations dues à ces trois facteurs. Le modèle incluait aussi deux pentes aléatoires : VALENCE par ITEM et VALENCE par AUDITEUR. Ce modèle maximal ne convergeant pas, il est ensuite comparé à un second modèle encaissé, plus simple, dans lequel une seule pente aléatoire est incluse. Nous avons effectué des comparaisons de modèles avec le test du rapport de vraisemblance par élimination progressive en utilisant la fonction *anova()* de R. Le modèle final obtenu n'inclut aucune pente aléatoire, mais il inclut les trois interceptes aléatoires.

Modèle LMER utilisé pour les analyses du temps de réaction :

$$\text{lmer}(\text{data} = \text{t\^a}che1, \log(RT) \sim \text{Valence} * \text{Congruence} * \text{Genre} + (1|\text{Auditeur}) + (1|\text{Item}) \\ + (1|\text{Voix}))$$

Nous avons, enfin, réalisé un modèle LMER sur un tableau de données combinant celles des trois tâches. Cela nous permet de tester l'effet de la CONGRUENCE, du GENRE, et de l'EMPATHIE sur les RT des participants. Ce modèle avait donc pour but de voir les effets de l'empathie et du genre sur les temps de réaction des participants, mais il nous permet également de voir si les informations empathie et genre étaient liées voire redondantes. Le modèle maximal incluait, comme précédemment, les interceptes aléatoires AUDITEUR, ITEM et VOIX, ainsi que les deux pentes aléatoires VALENCE par ITEM et VALENCE par AUDITEUR. Le modèle final obtenu après avoir comparé avec les tests ANOVA n'inclut aucune pente aléatoire, mais inclut AUDITEUR, ITEM et VOIX comme interceptes aléatoires.

Modèle LMER utilisé pour les analyses du temps de réaction (empathie) :

$$\text{lmer}(\text{data} = 3\text{t\^a}ches, \log(RT) \sim \text{Congruence} * \text{Genre} * \text{Empathie} + (1|\text{Auditeur}) + (1|\text{Item}) \\ + (1|\text{Voix}))$$

Pour chaque modèle réalisé lorsque nous avons plus d'une variable dépendante, nous avons effectué une comparaison par paire au sein du modèle. Cela a été réalisé avec la fonction *emmeans* (Searle et al., 1980) de R. Elle permet d'afficher chacune des interactions possibles calculées par le modèle, deux à deux. Au vu des grandes quantités de contrastes à observer dans nos données, la fonction *emmeans* nous a paru la solution la plus optimale pour rendre compte de nos données, plutôt qu'un simple *relevel*. Par

ailleurs, la fonction *emmeans* opère une correction automatique des valeurs p (par la méthode Turkey).

3. Résultats

Avant l'analyse des résultats, nous avons éliminé les *outliers* en nous basant sur les temps de réaction, éliminant les temps de réaction en même temps que les choix de réponse associés. Nous avons pour cela suivi les recommandations de Baayen et Milin (2010). Pour voir où se trouvaient les *outliers*, nous avons observé les graphiques de densité du temps de réaction (voir Annexe 17). Après observation visuelle des données, nous avons pu remarquer quelques *outliers*, avec notamment un temps de réaction qui montait jusqu'à 1 minute. Les réponses avant la fin de l'audio (durée moyenne, 2623 ms) étaient autorisées dans la tâche, mais ces occurrences sont rares. Nous avons éliminé, dans chacune des tâches, les temps de réaction de plus de 8 secondes, ainsi qu'un temps de réaction extrêmement court qui indique que l'auditeur a répondu seulement 1,5 seconde après l'onset de l'audio (une seule occurrence). Sur la tâche 1, 6% des observations, soit 54 observations sur 912 ont été exclues. Sur la tâche 2, 1% des observations, soit 9 observations sur 864 ont été exclues. Sur la tâche 3, 1% des observations, soit 5 observations sur 864 ont été exclues. Baayen et Millin (2010) proposent d'éliminer les temps de réaction très courts (<5 ms après le début du stimulus) ou absurdelement longs. Pour des tâches de décision visuelle lexicale par exemple, ils conseillent de ne pas conserver de temps de réaction de plus de 5 secondes. Nous avons conservé un seuil plus haut, car la tâche utilisée n'est pas visuelle mais auditive, et nous devons prendre en compte le temps d'écoute inclus dans le temps de réaction. Les résumés complets des résultats des modèles avec *emmeans* pour les temps de réaction sont proposés en Annexe 18.

3.1. Tâche 1 : Positif – Négatif

Dans la tâche 1, les auditeurs (22 femmes, 14 hommes) ont jugé des préfaces qui avaient été produites avant des annonces de nouvelles positives ou négatives. Ils avaient donc le choix entre une réponse positive « ah oui » et une réponse négative « oh non ». Le Tableau 26 présente les pourcentages de réponses dans chaque condition. Nous pouvons voir que les préfaces de nouvelles positives semblent obtenir un plus fort taux de réponses congruentes que celles des nouvelles négatives. La réponse « ah oui » est en effet plus souvent choisie lorsque la nouvelle est positive, mais elle semble aussi souvent choisie lorsque la nouvelle est négative.

		Choix du feedback				
		Ah oui		Oh non		Total
		Compte	%	Compte	%	Compte
Valence de la nouvelle	Positif	259	62%	156	38%	415
	Négatif	211	50%	208	50%	419
	Total	470	112%	364	88%	834

Tableau 26 – Tâche 1. Proportion du choix de feedback en fonction de la valence de la nouvelle dans la préface

Dans un premier temps, nous avons déterminé si le choix d'une réponse est utilisé au-delà du niveau de la chance en fonction de la valence de la nouvelle. Lorsque la valence négative est choisie comme intercepte, le choix de réponse « ah oui » n'est pas significatif [$\beta = 0.02$, $SE = 0.17$, $z = 0.11$, $p = 0.9$]. Cela signifie que la réponse « ah oui » a été choisie au niveau de la chance (50% du temps) lorsque la valence de la nouvelle était négative (choix aléatoire entre « ah oui » et « oh non »). Lorsque la valence positive est choisie comme intercepte en revanche, la valeur p est significative [$\beta = 0.54$, $SE = 0.16$, $z = 3.85$, $p < 0.0001$]. Cela signifie que « ah oui » a été choisi au-delà du niveau de la chance (62% du temps) lorsque la valence de la nouvelle est positive.

Dans un second temps, l'observation des graphiques de la Figure 85 couplée au résultat du modèle glmer complexe (Valence x Genre) nous permet d'évaluer la proportion de réponses pour un feedback donné en fonction de la valence de la nouvelle véhiculée par la préface et du genre du participant. Le graphique (spéculaires) présente les données pour « ah oui » (à gauche) et « oh non » (à droite). Nous voyons que les femmes choisissent

« ah oui » significativement plus souvent lorsque la valence de la nouvelle est positive (59%) par rapport à lorsqu'elle est négative (42%) [$\beta = 0.71$, $SE = 0.2$, $z = 3.55$, $p > 0.01$]. En revanche, les hommes ne choisissent pas « ah oui » significativement plus souvent que la valence soit positive (67%) ou négative (62%) [$\beta = 0.21$, $SE = 0.26$, $z = 3.81$, $p = 0.8$]. En d'autres termes, les hommes semblent donc moins bien distinguer les valences affectives et se préparer plus souvent à quelque chose de positif. Lorsque la valence de la nouvelle était négative, les femmes (59%) ont moins souvent répondu « ah oui » que les hommes (67%) [$\beta = -0.9$, $SE = 0.33$, $z = -2.74$, $p < 0.05$]. En revanche lorsqu'elle était positive, nous ne trouvons pas de différence entre les deux genres [$\beta = -0.39$, $SE = 0.28$, $z = -1.42$, $p = 0.4$].

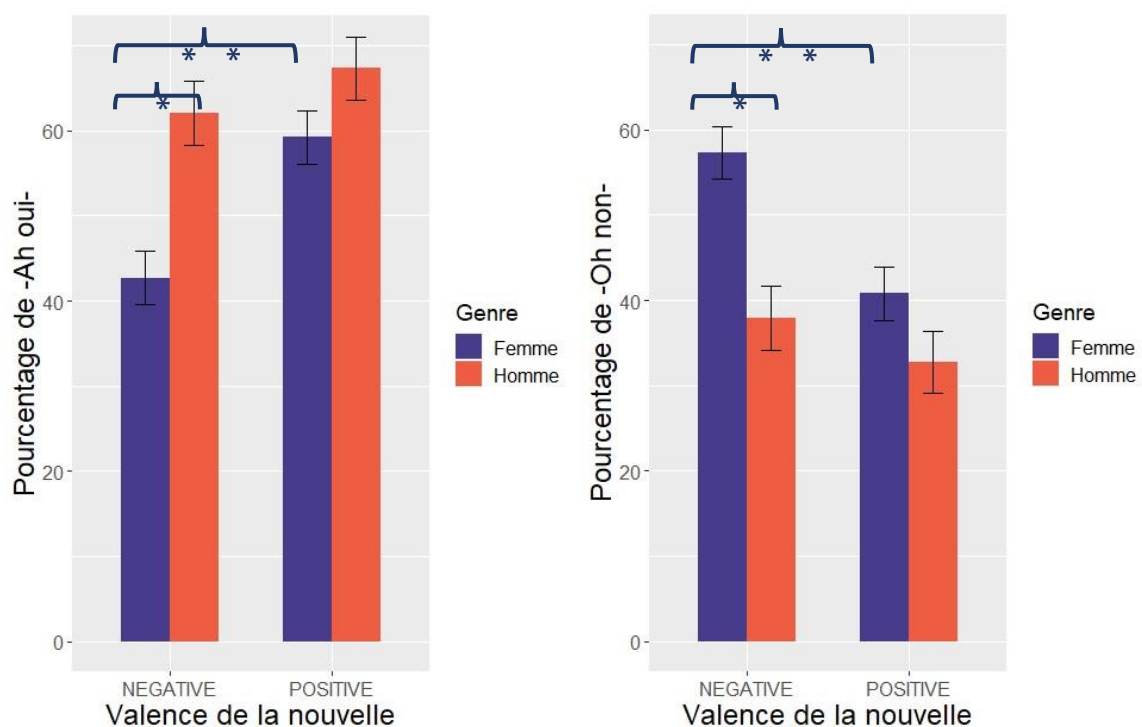


Figure 85 – Proportion de « ah oui » ou « oh non » et erreur-standard par valence et par genre dans la tâche 1

Nous présentons maintenant les différences dans le temps de réaction des participants en fonction de plusieurs facteurs : la valence de la nouvelle, la congruence de la réponse avec cette valence, et le genre du participant. Le Tableau 27 présente les valeurs moyennes de temps de réaction en fonction des trois facteurs.

<i>Valence</i>	<i>Genre</i>	<i>Congruence</i>	<i>RT moyen (ms)</i>	<i>SD (ms)</i>
<i>Négative</i>	Femme	Congruent	3658	1122
		Incongruent	3696	952
		Total	3677	1037
	Homme	Congruent	3714	1138
		Incongruent	3987	1115
		Total	3850	1126
<i>Positive</i>	Femme	Congruent	3436	885
		Incongruent	3534	1029
		Total	3485	957
	Homme	Congruent	3592	1070
		Incongruent	3631	962
		Total	3612	1016

Tableau 27 – Tâche 1. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse

Le graphique présenté Figure 86 illustre les différences dans le temps de réaction des participants en fonction des mêmes facteurs que le tableau. Les temps de réaction semblent être assez constants au travers de ces variables. Les analyses n'ont effectivement révélé aucune différence significative entre les conditions. Il n'y a pas de différence significative de temps de réaction entre les deux valences négative et positive, selon le genre, selon la congruence des réponses, ou dans les interactions entre les facteurs, la valeur p étant toujours au-dessus du seuil de significativité [$p > 0.05$]. Cela veut dire que le temps de réaction n'est impacté par aucun des facteurs lorsque la valence négative est comparée à la valence positive.

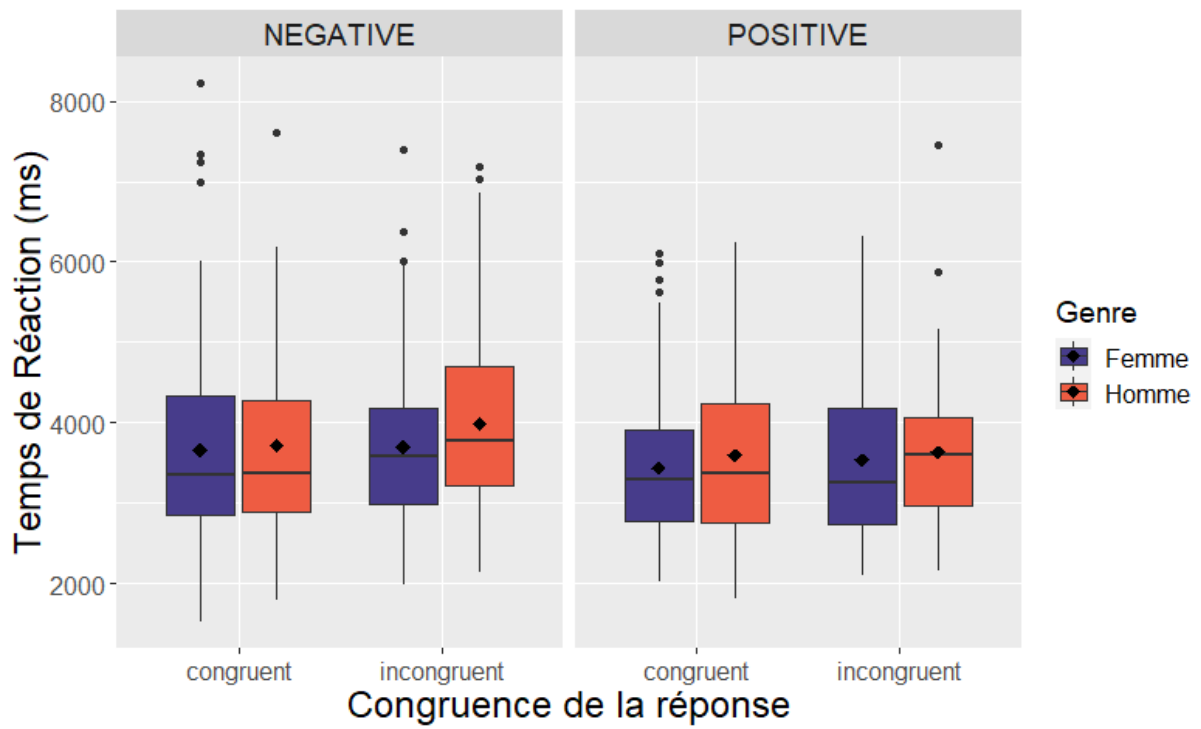


Figure 86 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 1

3.2. Tâche 2 : Négatif – Neutre

Dans la tâche 2, les auditeurs (22 femmes, 14 hommes) ont jugé des préfaces qui avaient été produites avant des annonces de nouvelles négatives ou neutres. Ils avaient donc le choix entre une réponse négative « oh non » et une réponse neutre « j’écoute ». Le Tableau 28 présente les pourcentages de réponses dans chaque condition. Nous pouvons voir que la nouvelle neutre paraît obtenir un plus fort taux de réponses congruentes que la nouvelle négative. La réponse « j’écoute » semble plus souvent choisie lorsque la préface est neutre, mais elle est presque tout autant choisie lorsque la préface est négative. En comparaison, le feedback « oh non » est bien moins souvent choisi.

		Choix du feedback				
		Oh non		J’écoute		Total
		Compte	%	Compte	%	Compte
Valence de la nouvelle	Négatif	158	38%	257	62%	415
	Neutre	138	33%	278	67%	416
	Total	296	71%	535	129%	831

Tableau 28 – Tâche 2. Proportion moyenne du choix de feedback en fonction de la valence de la nouvelle dans la préface

Dans un premier temps, nous avons déterminé si le choix d’une réponse est utilisé au-delà du niveau de la chance en fonction de la valence de la nouvelle. Lorsque la valence négative est choisie comme intercepte, le choix de réponse « j’écoute » est significatif [$\beta = 0.53$, $SE = 0.14$, $z = 3.64$, $p < 0.001$]. Cela indique que la réponse « j’écoute » a été choisie au niveau de la chance (62% du temps) lorsque la valence de la nouvelle était négative, donc que les préfaces n’ont pas élicité une réponse affective congruente avec la valence de la nouvelle. Lorsque la valence neutre est choisie comme intercepte, la valeur p est également significative [$\beta = 0.73$, $SE = 0.13$, $z = 5.55$, $p < 0.0001$]. Cela signifie que la réponse « j’écoute » (congruente) a été choisi au-delà du niveau de la chance (67% du temps) lorsque la valence de la nouvelle est neutre. La réponse « oh non » est donc toujours choisie en dessous du niveau de la chance.

Dans un second temps, l’observation des graphiques de la Figure 87 couplée au résultat du modèle glmer complexe nous permet d’évaluer la proportion de réponses pour un feedback donné en fonction de la valence de la nouvelle véhiculée par la préface et du

genre du participant. La figure présente les données pour « j'écoute » (à gauche) et « oh non » (à droite). Nous voyons que la proportion de choix « j'écoute » ne varie pas significativement entre la nouvelle négative et la neutre, que ce soit pour les femmes [$\beta = 0.27$, $SE = 0.19$, $z = 1.4$, $p = 0.4$] ou les hommes [$\beta = 0.08$, $SE = 0.25$, $z = 0.33$, $p = 0.9$]. Nous ne trouvons pas non plus de différences significatives au sein d'une même valence entre femmes et hommes, pour le négatif [$\beta = -0.36$, $SE = 0.29$, $z = -1.22$, $p = 0.6$], et pour le neutre [$\beta = -0.17$, $SE = 0.26$, $z = -0.66$, $p = 0.9$].

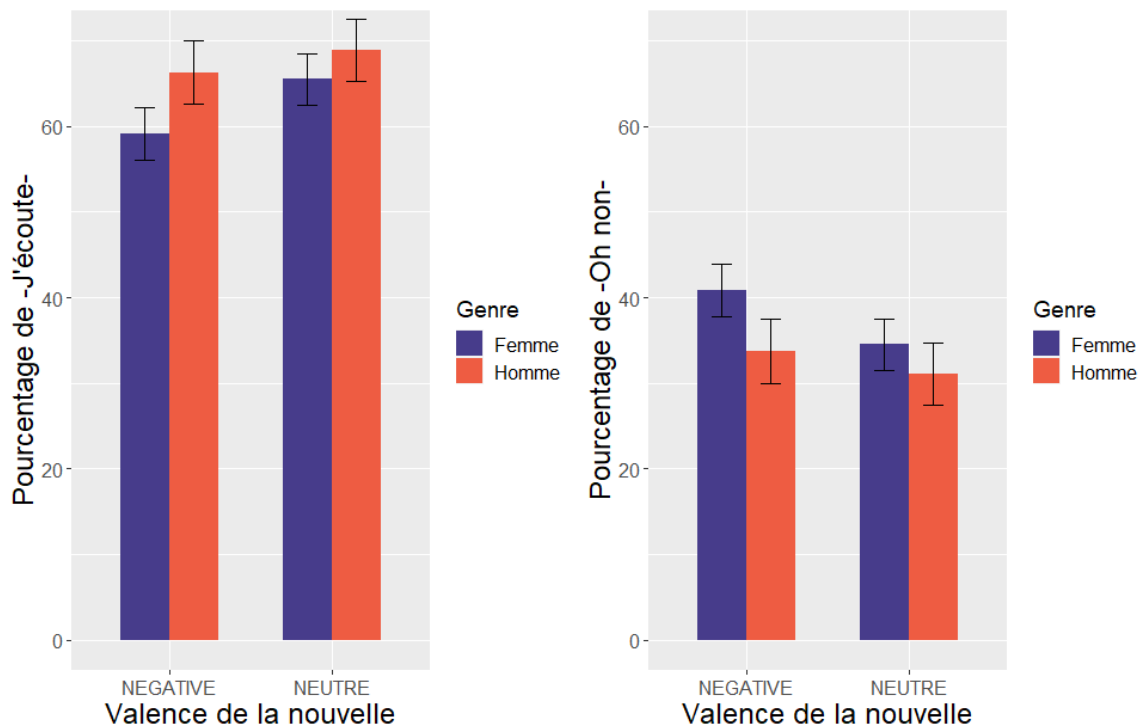


Figure 87 – Proportion de « j'écoute » ou « oh non » et erreur-standard par valence et par genre dans la tâche 2

Nous présentons maintenant les différences dans le temps de réaction des participants en fonction de plusieurs facteurs : la valence de la nouvelle, la congruence de la réponse avec cette valence, et le genre du participant. Le Tableau 29 présente les valeurs moyennes de temps de réaction en fonction des trois facteurs.

Valence	Genre	Congruence	RT moyen (ms)	SD (ms)
<i>Négative</i>	Femme	Congruent	3777	792
		Incongruent	3694	845
		Total	3735	819
	Homme	Congruent	4184	863
		Incongruent	3937	1071
		Total	4061	967
<i>Positive</i>	Femme	Congruent	3359	809
		Incongruent	3432	894
		Total	3396	852
	Homme	Congruent	3610	928
		Incongruent	3698	872
		Total	3654	900

Tableau 29 – Tâche 2. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse

Le graphique présenté Figure 88 illustre les différences dans le temps de réaction des participants en fonction des mêmes facteurs que le tableau. Les analyses n'ont révélé aucune différence significative entre les conditions. Il n'y a pas de différence significative de temps de réaction entre les deux valences négative et neutre, selon le genre, selon la congruence des réponses, ou dans les interactions entre les facteurs, la valeur p étant toujours au-dessus du seuil de significativité [$p > 0.05$]. Cela veut dire que le temps de réaction n'est impacté par aucun des facteurs lorsque la valence négative est comparée à la valence neutre.

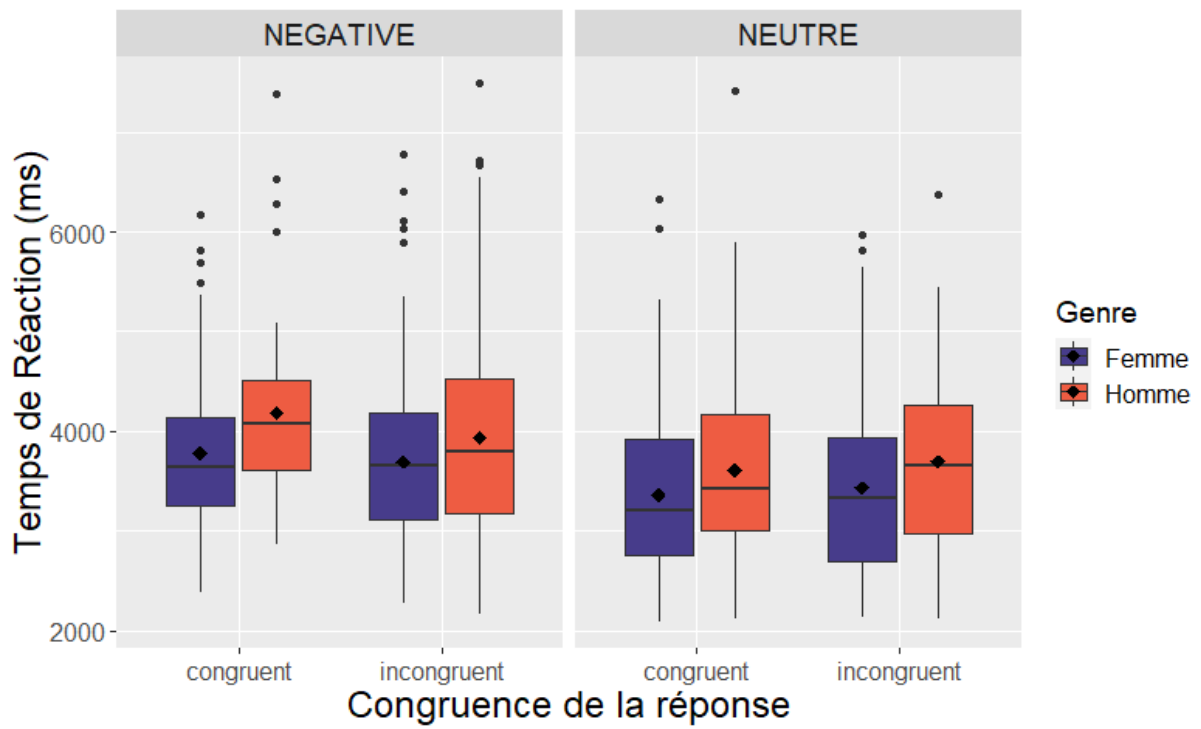


Figure 88 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 2

3.3. Tâche 3 : Positif – Neutre

Dans la tâche 3, les auditeurs (22 femmes, 14 hommes) ont jugé des préfaces qui avaient été produites avant des annonces de nouvelles positives ou neutres. Ils avaient donc le choix entre une réponse positive « ah oui » et une réponse neutre « j’écoute ». Le Tableau 30 présente les pourcentages de réponses dans chaque condition. Nous pouvons voir que la préface de la nouvelle positive semble obtenir un plus fort taux de réponses congruentes que celle de la nouvelle neutre, cette différence étant très peu marquée. Les réponses semblent avoir été choisies aléatoirement, indépendamment de la valence de la nouvelle. La réponse « ah oui » est ainsi souvent choisie lorsque la nouvelle est neutre (60% du temps), mais elle est presque tout autant choisie lorsque la nouvelle est positive (56% du temps). En comparaison, le feedback neutre « j’écoute » est choisi entre 40 et 44 % du temps.

		Choix du feedback				
		Ah oui		J’écoute		Total
		Compte	%	Compte	%	Compte
Valence de la nouvelle	Positif	242	56%	188	44%	430
	Neutre	257	60%	172	40%	429
	Total	499	116%	360	84%	859

Tableau 30 – Tâche 3. Proportion du choix de feedback en fonction de la valence de la nouvelle dans la préface

Dans un premier temps, nous avons déterminé si le choix d’une réponse est utilisé au-delà du niveau de la chance en fonction de la valence de la nouvelle. Lorsque la valence positive est choisie comme intercepte, le choix de réponse « ah oui » n’est pas significatif [$\beta = 0.25$, $SE = 0.11$, $z = 2.24$, $p = 0.025$]. Cela signifie que la réponse « ah oui » a été choisie au niveau de la chance (56% du temps) lorsque la valence de la nouvelle est positive (choix aléatoire entre « ah oui » et « j’écoute »). Donc les préfaces aux nouvelles positives n’ont pas élicité une réponse affective congruente avec la valence de la nouvelle. Lorsque la valence neutre est choisie comme intercepte, la valeur p est en revanche significative [$\beta = 0.41$, $SE = 0.12$, $z = 3.43$, $p < 0.001$]. Cela signifie que « ah oui » a été choisi au-delà du niveau de la chance (60% du temps) lorsque la valence de la nouvelle est neutre. Aucune des deux valences véhiculées par la préface n’a donc élicité une réaction affective congruente avec la nouvelle.

Dans un second temps, l'observation des graphiques de la Figure 89 couplée au résultat du modèle glmer complexe nous permet d'évaluer la proportion de réponses pour un feedback donné en fonction de la valence de la nouvelle véhiculée par la préface et du genre du participant. La figure présente les données pour « ah oui » (à gauche) et « j'écoute » (à droite). Nous voyons que la proportion de choix « ah oui » ne varie pas significativement entre la nouvelle positive et la neutre, que ce soit pour les femmes [$\beta = -0.23$, SE = 0.18, $z = -1.32$, $p = 0.5$] ou les hommes [$\beta = -0.02$, SE = 0.22, $z = -0.11$, $p = 0.9$]. Nous ne trouvons pas non plus de différences significatives au sein d'une même valence entre femmes et hommes, pour le positif [$\beta = 0.01$, SE = 0.23, $z = 0.08$, $p = 0.9$], et pour le neutre [$\beta = 0.23$, SE = 0.24, $z = 0.94$, $p = 0.7$]. Ces différents résultats indiquent que la valence de la nouvelle portée par la préface et le genre n'ont pas influencé le choix des participants.

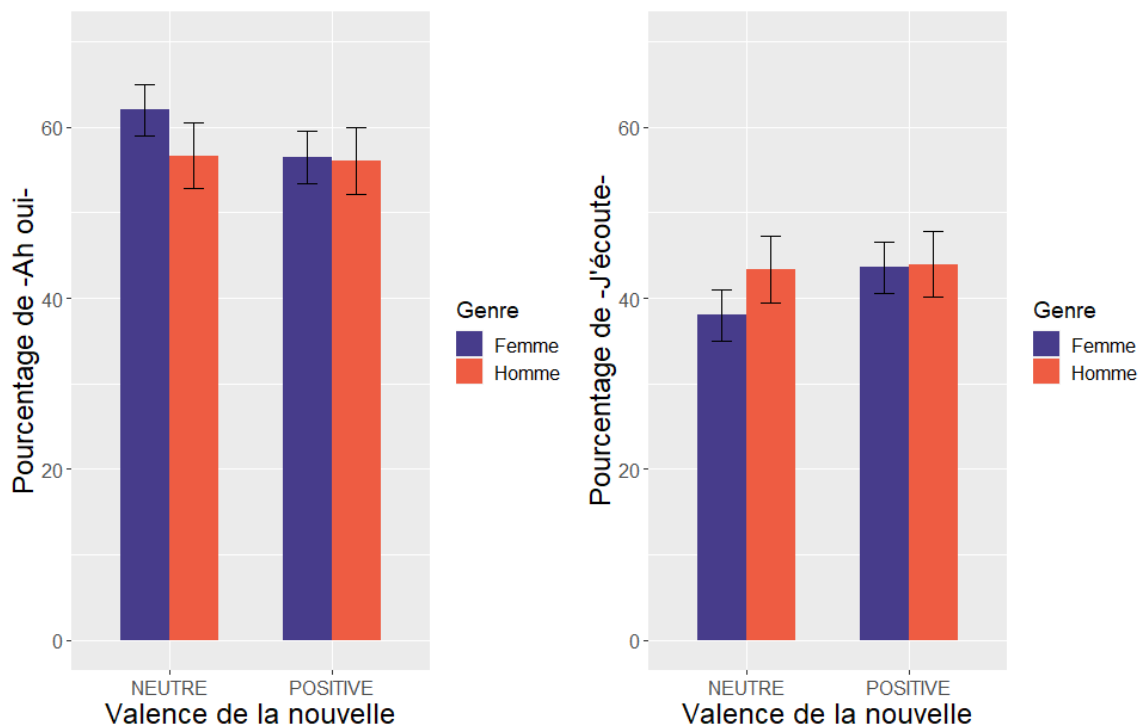


Figure 89 – Proportion de « ah oui » ou « j'écoute » et erreur-standard par valence et par genre dans la tâche 3

Nous présentons maintenant les différences dans le temps de réaction des participants en fonction de plusieurs facteurs : la valence de la nouvelle, la congruence de la réponse avec cette valence, et le genre du participant. Le Tableau 31 présente les valeurs moyennes de temps de réaction en fonction des trois facteurs.

Valence	Genre	Congruence	RT moyen (ms)	SD (ms)
<i>Positive</i>	Femme	Congruent	3535	851
		Incongruent	3694	864
		Total	3614	858
	Homme	Congruent	3762	883
		Incongruent	3957	975
		Total	3859	929
<i>Neutre</i>	Femme	Congruent	3782	1147
		Incongruent	3711	1124
		Total	3746	1135
	Homme	Congruent	3973	1166
		Incongruent	3968	1108
		Total	3970	1137

Tableau 31 – Tâche 3. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse

Le graphique présenté Figure 90 illustre les différences dans le temps de réaction des participants en fonction des mêmes facteurs que le tableau. Les temps de réaction semblent être assez constants au travers de ces variables. Les analyses n'ont effectivement révélé aucune différence significative entre les conditions. Il n'y a pas de différence significative de temps de réaction entre les deux valences positive et neutre, selon le genre, selon la congruence des réponses, ou dans les interactions entre les facteurs, la valeur p étant toujours au-dessus du seuil de significativité [$p > 0.05$].

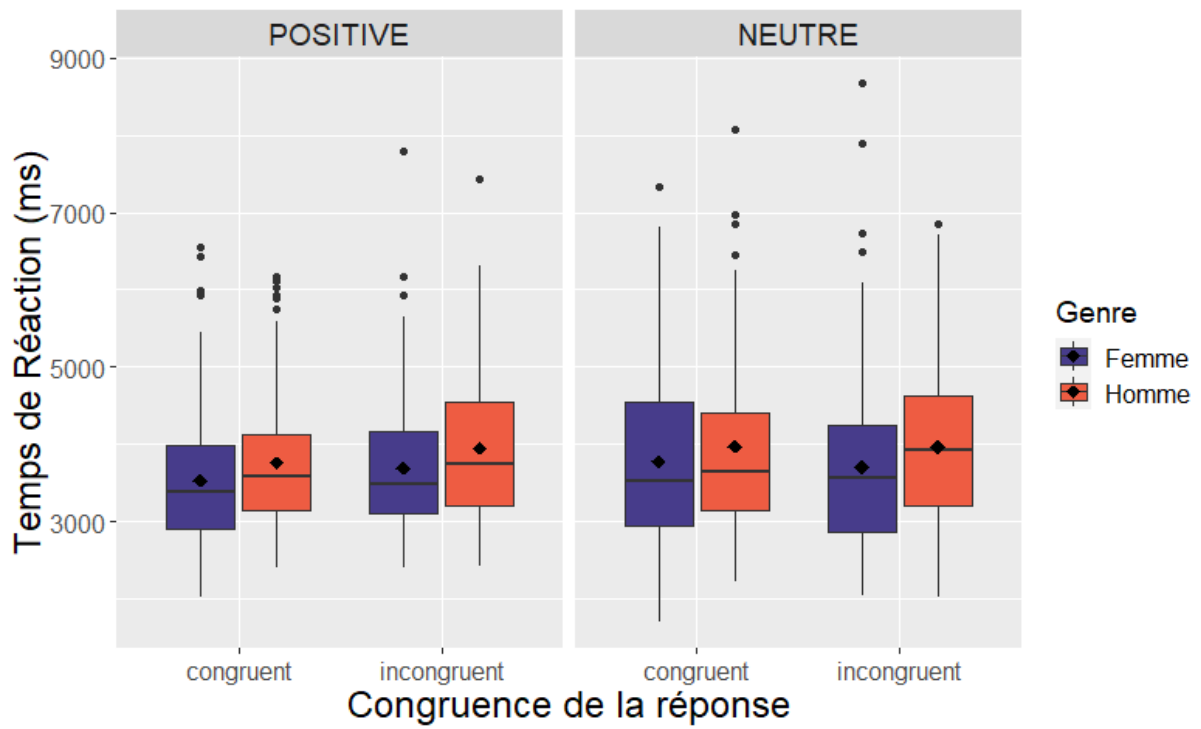


Figure 90 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 3

3.4. Observation de l'empathie à travers les trois tâches

Nous avons fait des comparaisons au travers des trois tâches, afin de voir si le niveau d'empathie et le genre peuvent moduler les temps de réaction des auditeurs. Nous avons en effet fait l'hypothèse que les hommes pourraient avoir des temps de réaction plus longs, car ils sont plus attentifs aux indices verbaux qu'aux indices prosodiques, à l'inverse des femmes. En effet, bien qu'ils soient également attentifs aux indices prosodiques, ils mettent plus de temps à les utiliser que les femmes (Schirmer et al., 2002, 2004). Ici, nous nous concentrons donc spécifiquement les temps de réaction, la proportion de réponses en fonction de la valence différant au sein des trois tâches. Nous avons également fait l'hypothèse que les personnes ayant un haut score d'empathie auront un temps de réaction plus court que celui des personnes avec un score bas, du fait du lien neuronal entre la perception de la prosodie et l'empathie (Aziz-Zadeh et al., 2010). Enfin, nous avons fait l'hypothèse qu'il pourrait y avoir une interaction pour les temps de réaction en fonction du genre et du niveau d'empathie des participants, car les scores d'empathie démontrent des différences entre les genres dans la littérature (Carre et al., 2013; D'Ambrosio et al., 2009; Jolliffe & Farrington, 2006).

Nous avons choisi de n'observer les variations en fonction de l'empathie qu'en regardant les données des trois tâches assemblées. Ce choix a été fait, car le nombre d'auditeurs par tâche et selon le genre était trop petit pour évaluer l'apport de ce facteur, comme nous pouvons le voir dans le Tableau 32 ci-dessous. Le tableau présente le nombre d'auditeurs dans chaque tâche, par genre et par niveau d'empathie. Le compte correspond au nombre exact de personnes, et le pourcentage représente la proportion que cela représente par rapport au nombre total de participants à la tâche. En moyenne sur les trois tâches, et dans chacune d'entre elles, les femmes ont obtenu plus souvent un score d'empathie haut que les hommes.

Tâche 1 (Négatif vs Positif)					
<i>Empathie</i>		<i>Basse</i>		<i>Haute</i>	
		<i>Compte</i>	<i>%</i>	<i>Compte</i>	<i>%</i>
Genre	Hommes	5	14%	9	25%
	Femmes	6	17%	16	44%
	Total	11	31%	25	69%
Tâche 2 (Négatif vs Neutre)					
<i>Empathie</i>		<i>Basse</i>		<i>Haute</i>	
		<i>Compte</i>	<i>%</i>	<i>Compte</i>	<i>%</i>
Genre	Hommes	5	14%	9	26%
	Femmes	6	17%	15	43%
	Total	11	31%	24	69%
Tâche 3 (Positif vs Neutre)					
<i>Empathie</i>		<i>Basse</i>		<i>Haute</i>	
		<i>Compte</i>	<i>%</i>	<i>Compte</i>	<i>%</i>
Genre	Hommes	4	11%	10	28%
	Femmes	5	14%	17	47%
	Total	9	25%	27	75%
Tâche 1 + Tâche 2 + Tâche 3					
<i>Empathie</i>		<i>Basse</i>		<i>Haute</i>	
		<i>Compte</i>	<i>%</i>	<i>Compte</i>	<i>%</i>
Genre	Hommes	23	21%	19	18%
	Femmes	24	22%	41	38%
	Total	47	44%	60	56%

Tableau 32 – Proportion du niveau d'empathie par genre dans chaque tâche

Nous présentons maintenant les différences dans le temps de réaction des participants en fonction de plusieurs facteurs : la congruence de la valence du feedback avec la valence de la nouvelle portée par la préface, le genre du participant, et le niveau d'empathie. Le Tableau 33 présente les valeurs moyennes de temps de réaction en fonction des trois facteurs.

<i>Congruence</i>	<i>Genre</i>	<i>Empathie</i>	<i>RT moyen (ms)</i>	<i>SD (ms)</i>
<i>Congruent</i>	Femme	Basse	3486	999
		Haute	3612	912
		Total	3549	956
	Homme	Basse	3618	956
		Haute	3933	1076
		Total	3776	1016
<i>Incongruent</i>	Femme	Basse	3425	981
		Haute	3774	934
		Total	3600	958
	Homme	Basse	3799	1025
		Haute	4014	1060
		Total	3907	1043

Tableau 33 – Comparaison inter-tâches. Temps de reaction moyen et SD en fonction du genre, du niveau d'empathie et de la congruence de la réponse

Le graphique présenté Figure 91 illustre les différences dans le temps de réaction des participants en fonction des mêmes facteurs que le tableau. Lorsque le feedback de réponse choisi par l'auditeur est congruent, nous voyons une tendance à augmenter le temps de réaction progressivement entre les femmes à l'empathie basse (3612 ms) et les hommes à empathie haute (3933 ms), le seul contraste significatif étant celui entre, justement, les femmes à l'empathie basse et les hommes à empathie haute [$\beta = -0.14$, $SE = 0.04$, $t = -3.5$, $p < 0.05$]. Il n'y a pas différences significatives entre les femmes à l'empathie basse et les hommes à l'empathie haute [$p = 0.6$], les femmes à l'empathie basse et les femmes à l'empathie haute [$p = 0.9$], les femmes à l'empathie haute et les hautes à l'empathie haute [$p = 0.1$], les hommes à l'empathie basse et les femmes à l'empathie haute [$p = 0.9$], et les hommes à l'empathie basse et les hommes à l'empathie haute [$p = 0.6$]. Ce schéma est le même lorsque nous regardons le temps de réaction si la réponse choisie est incongruente. Nous voyons la même tendance à augmenter le temps de réaction

progressivement entre les femmes à l'empathie basse (3425 ms) et les hommes à empathie haute (4014 ms), le seul contraste significatif étant celui entre, à nouveau, les femmes à l'empathie basse et les hommes à empathie haute [$\beta = -0.12$, $SE = 0.04$, $t = -3.12$, $p < 0.05$]. Il n'y a pas de différences significatives entre les femmes à l'empathie basse et les hommes à l'empathie haute [$p = 0.6$], les femmes à l'empathie basse et les femmes à l'empathie haute [$p = 0.5$], les femmes à l'empathie haute et les hommes à l'empathie haute [$p = 0.7$], les hommes à l'empathie basse et les femmes à l'empathie haute [$p = 1$], et les hommes à l'empathie basse et les hommes à l'empathie haute [$p = 0.8$].

Ainsi, nous voyons que le temps de réaction des individus n'a pas été influencé par la congruence de la réponse. Malgré de légères tendances visibles sur le graphique, aucune différence significative n'apparaît au sein d'un seul genre, et aucune n'apparaît au sein d'un seul niveau d'empathie. En revanche, nous notons une influence croisée du genre et de l'empathie.

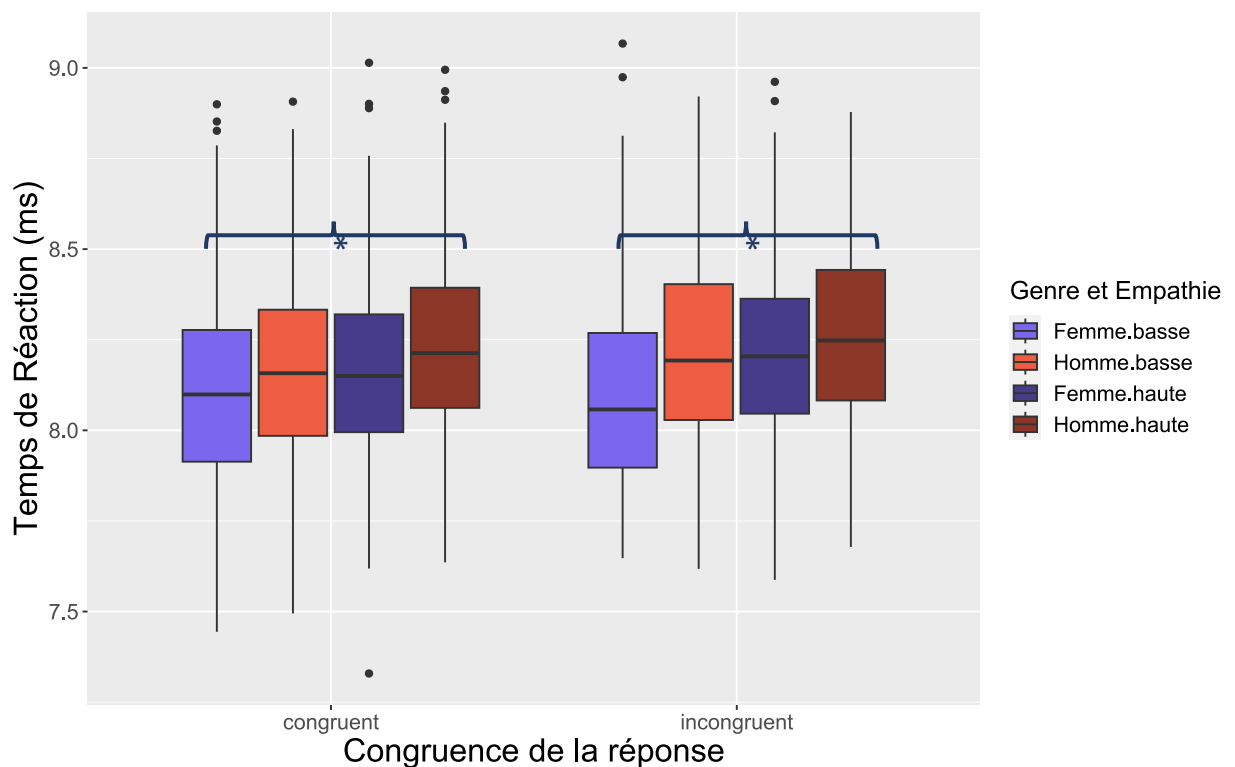


Figure 91 – Temps de réaction (ms) en fonction de la congruence de la réponse et du genre et du niveau d'empathie, à travers les trois tâches

4. Discussion

Le but de notre étude était d'observer la perception de l'interlocuteur qui reçoit une annonce de nouvelles de la part d'autrui. Nous avons voulu regarder s'il est capable de récupérer des indices de projection de la préface sur la valence affective de la nouvelle, et si cette anticipation influence sa réaction affective. Pour cela, nous avons créé trois tâches où, à l'écoute de préfaces d'annonces, des auditeurs devaient choisir une réponse accordée à leur « réaction émotionnelle ». Notre hypothèse générale est qu'un interlocuteur va utiliser les indices de projection affective pour se préparer affectivement à la nouvelle. Précisément, les auditeurs dans notre étude devraient faire la différence entre des préfaces réalisées avant une nouvelle positive, négative, ou neutre, et choisir une réponse affective en accord avec la valence de la nouvelle.

		Choix		
		Valence neutre	Valence positive	Valence négative
Ah oui	Total	_	>	<
	Hommes	_	>	>
	Femmes	_	>	<
Oh non	Total	<	<	<
	Hommes	<	<	<
	Femmes	<	<	<
J'écoute	Total	<	_	>
	Hommes	>	_	>
	Femmes	>	_	>
Temps de réaction				
		Congruence	Incongruence	
Congruence x Genre		_	_	X
Genre x Empathie				X
Femmes	Basse	<	<	X
	Haute	_	_	X
Hommes	Basse	_	_	X
	Haute	>	>	X

Tableau 34 – Tableau récapitulatif des choix selon la valence et le genre et des RT selon la congruence, le genre et l'empathie (< moins choisi / RT plus court , > plus choisi / RT plus long, _ non significatif, X NA)

Dans l'ensemble, les analyses (résumées dans le Tableau 25) suggèrent que les préfaces peuvent éliciter une perception différente chez l'interlocuteur en fonction de la valence de la nouvelle. Cette découverte est cohérente avec les résultats de l'étude de perception de Swerts et Hirschberg (2010), où les nouvelles positives et négatives avaient été reconnues comme telles. Le premier résultat important est révélé par les modèles glmer « simples » (ne regardant que l'effet de la valence sur le choix de feedback) : « ah oui » est significativement plus souvent sélectionné lorsque la valence de la nouvelle est positive. De même « j'écoute » est significativement plus souvent sélectionné lorsque la valence de la nouvelle est neutre. En revanche, lorsqu'elles sont opposées entre elles, les valences positive et neutre n'élicitent pas de réponses congruentes, ce qui semble indiquer qu'elles pourraient être confondues, en ligne avec nos résultats des analyses phonétiques et phonologiques de production. La réponse « oh non » est généralement moins utilisée, « j'écoute » lui étant même significativement préféré. Ce résultat n'est pas congruent avec la littérature sur la perception de la prosodie (Audibert et al., 2005) où la tristesse, que nous rattachons à la valence négative, est normalement bien reconnue. De même, ce n'est pas non plus en ligne avec les résultats de notre étude de production, la valence négative étant distinguée des deux autres par de nombreux indices de projection phonétiques et quelques indices phonologiques. Nous attribuons ces résultats à des facteurs interactionnels de protection de la face (P. Brown & Levinson, 1987; Goffman, 1955, 1967) et de préférence (Sacks, 1973; Schegloff, 1972) que les auditeurs pourraient avoir pris en compte pour sélectionner le feedback. Le second résultat important correspond à des différences significatives dans le choix des feedbacks en fonction du genre de l'auditeur. Les différences significatives ressortent principalement dans la tâche 1, lorsque les préfaces projettent des nouvelles positives et négatives. Nous avons trouvé que les femmes répondent plus souvent « ah oui » que « oh non » lorsque la nouvelle est positive par rapport à lorsqu'elle est négative. Nous ne retrouvons pas cette distinction chez les hommes : ils répondent « ah oui » plus souvent, que la nouvelle soit positive ou négative. Les femmes paraissent donc plus se préparer affectivement en fonction de la valence de la nouvelle que les hommes. Cette découverte n'est pas en ligne avec les études sur la perception de la prosodie affective par les hommes et les femmes (Beaucousin et al., 2011; Besson et al., 2002; Schirmer et al., 2002, 2004). En effet, si les différences entre les genres étaient présentes au niveau des processus cognitifs engagés et du décours temporel de la

reconnaissance de la prosodie, elles n'étaient pas présentes au niveau de la capacité à reconnaître la prosodie affective. Le troisième résultat important de notre étude correspond aux effets croisés du genre et du niveau d'empathie des auditeurs sur le temps de réaction. Nous avons trouvé que les hommes à l'empathie haute répondent plus lentement que les femmes à l'empathie basse. Cela indique des tendances sur une influence de genre et de l'empathie qui ne ressortent que lorsque ces facteurs sont combinés. Ce résultat est partiellement cohérent avec la littérature et nos hypothèses. Le fait que les hommes puissent mettre plus de temps à réagir à la prosodie affective par rapport aux femmes est cohérent avec les études de Schirmer et al. (2002, 2004). En revanche, nous nous attendions à ce que les liens entre le niveau d'empathie et la reconnaissance de la prosodie (Aziz-Zadeh et al., 2010) poussent les personnes avec un haut niveau d'empathie à répondre plus vite, car ils reconnaîtraient mieux la prosodie affective. Or, la tendance est plutôt inverse. Aziz-Zadeh et al. (2010) évoquent dans leur étude le rôle potentiel des neurones miroirs et d'une simulation des affects qui serait engagée pour la perception et la reconnaissance de la prosodie affective en particulier. Si de tels procédés sont engagés, il est possible que les auditeurs avec un haut niveau d'empathie les aient plus utilisés et que cela ait délayé leur choix de réponse, contrairement aux personnes à empathie basse qui pourraient avoir moins engagé ce processus de simulation, et répondraient donc plus vite.

Le premier résultat important de cette étude correspond au fait que les indices de projection prosodiques ont été récupérés et utilisés par les auditeurs pour se préparer affectivement à l'annonce lorsque les préfaces écoutées avaient été produites avant l'annonce d'une nouvelle positive ou avant l'annonce d'une nouvelle neutre, dans les tâches 1 et 2. Dans les deux cas, les auditeurs ont plus souvent sélectionné la réponse affective congruente avec la valence entendue : le feedback positif « ah oui » pour les nouvelles positives et le feedback neutre « j'écoute » pour les nouvelles neutres. En revanche, les valences de nouvelles portées par la prosodie n'ont pas toujours élicité une réaction affective congruente. En effet dans ces deux tâches, la réponse congruente avec la nouvelle négative aurait été le feedback négatif « oh non ». Pourtant, nous voyons que « ah oui » et « oh non » ont été choisis au niveau de la chance quand la nouvelle était négative, ce qui indique qu'ils ont été choisis aléatoirement. Dans la tâche 2, c'est « j'écoute » qui a été sélectionné au-dessus du niveau de la chance, que la nouvelle soit

négative ou neutre. La valence négative a donc élicité peu de réponses avec le feedback congruent « oh non », ce qui indique les auditeurs se sont moins préparés à la valence affective de ces nouvelles. Elles pourraient donc avoir été mal reconnues. Dans la tâche 3 le choix « ah oui » est utilisé aléatoirement entre les nouvelles positives et neutres. Cela indique que les deux valences pourraient avoir été confondues. Le fait que la valence positive ait élicité des feedbacks positifs « ah oui » dans la tâche 1 est en ligne avec nos hypothèses, et avec le fait que les nouvelles positives avaient été bien reconnues dans l'étude de Swerts et Hirschberg (2010). L'ensemble des résultats est également cohérent avec l'étude de perception de Audibert et al. (2005). Ils avaient en effet trouvé que la joie (que nous rapprochons du positif) est identifiée 80% du temps, alors que la tristesse (que nous rapprochons du négatif) ne l'est qu'environ 60% du temps. Le neutre est, dans leur étude, confondu 48% du temps avec la tristesse. Nos données sont partiellement en désaccord avec ce dernier élément, la prosodie des nouvelles neutres semblant avoir élicité une préparation affective similaire à celle des nouvelles positives (sur la tâche 3). De plus, cela est cohérent avec notre étude de production, qui a démontré que la prosodie produite dans les préfaces de nouvelles neutres et positives était parfois semblable. Les résultats indiquent donc que les indices de projection de la prosodie affective sont utilisés par les interlocuteurs lorsqu'ils sont en interaction, mais pas qu'elle que soit la valence de la nouvelle. A l'écoute des préfaces d'annonces, les interlocuteurs sont capables de récupérer et d'utiliser les indices de projection non seulement pour anticiper la valence affective de la nouvelle (Swerts & Hirschberg, 2010), mais aussi pour se préparer sur le plan affectif. Cette découverte confirme notre hypothèse principale sur le rôle de préparation affective que peut jouer la prosodie auprès des interlocuteurs. Les indices de projection portés par la prosodie dans les préfaces semblent plus préparer l'interlocuteur lorsqu'ils opposent l'annonce d'une nouvelle qui pourrait mettre en danger la face des interactant (négative), à une nouvelle qui est interactionnellement moins dangereuse et pour laquelle les auditeurs se préparent plus (positive, neutre). Une telle différence est cohérente avec la littérature interactionnelle sur les notions de face (P. Brown & Levinson, 1987; Goffman, 1955, 1967) et de préférence (Sacks, 1973; Schegloff, 1972).

En revanche, il est surprenant que la valence négative n'ait pas élicité de réaction affective congruente, étant donné les résultats de nos analyses phonétiques et phonologiques de production : la valence négative était toujours celle qui présentait des indices de projection très distincts des deux autres valences de nouvelle. Les nouvelles

positives et neutres ont, elles, élicité une réponse affective congruente (visible par le choix d'un feedback dont la valence est congruente avec celle de la nouvelle) lorsqu'elles étaient opposées au négatif. Nous nous questionnons donc sur le fait que l'absence de préparation affective aux nouvelles négatives puisse être due à une autre raison qu'une seule mauvaise reconnaissance de la valence négative de la nouvelle. Nous nous demandons si le choix du feedback a réellement été guidé uniquement par la valence de la nouvelle portée par la prosodie de la préface et par le ressenti affectif de l'auditeur, ou si un autre facteur est entré en jeu dans le choix de réponse.

Nous pensons qu'en dépit de la consigne donnée aux participants de choisir la réponse correspondant à leur « réaction émotionnelle » et non celle qu'ils auraient dit à l'oral, ils ont pu prendre en compte le contexte interactionnel dans leur choix de réponse et sélectionner malgré tout une réaction en pensant à ce qu'ils auraient répondu oralement. En effet, il est possible que nous n'ayons pas assez insisté sur le fait de ne pas se préoccuper de ce qui aurait été répondu à l'oral. Mais surtout, il est possible que le choix de feedbacks verbaux n'ait pas été le plus adapté, puisqu'ils sont supposés être un élément interactionnel avant tout ce qui a pu induire en erreur les auditeurs au cours de la tâche. Si l'attention des auditeurs s'est bien focalisée sur les aspects interactionnels, nos résultats pourraient être en ligne avec les règles interactionnelles décrites dans l'état de l'art de cette thèse, et notamment en fonction des règles de préférence et de protection de la face (Brown & Levinson, 1987; Goffman, 1955, 1967). Répondre « oh non » à une préface de nouvelle indépendamment de sa valence pourrait en effet correspondre à une réponse non-préférée, car elle laisse entendre que l'interlocuteur réagit, d'avance, négativement à la nouvelle. De plus, une telle réponse, même en simple feedback, ressemble à un refus de la part de l'interlocuteur de recevoir l'annonce, ce qui peut la rendre difficile à faire pour le locuteur, et mettre en danger la face des interactants. Il serait alors beaucoup plus simple pour l'interlocuteur de répondre un feedback positif « ah oui » ou neutre « j'écoute » malgré la valence négative de la nouvelle, car ceux-ci permettent de répondre sans mettre en danger les faces de chacun, en choisissant des réponses préférées qui laissent le locuteur continuer son annonce. Mais alors, pourquoi choisir « ah oui » autant que « j'écoute » lorsque la nouvelle est neutre ? Dans ce cadre-là, la notion de mise en danger de la face des interactants ne devrait pas être prégnante, et la « réponse préférée » devrait être celle qui est la plus affiliée (donc un feedback neutre). Pourtant, « ah oui » n'est pas la réponse la plus affiliée pour le neutre puisqu'elle a été

validée comme étant à valence positive dans notre étude préliminaire de construction de l'étude, et elle est malgré tout régulièrement choisie.

Selon nous, ce fait peut être interprété de multiples façons. Une première explication pourrait être que les préfaces orales de nouvelles neutres et positives ne présentant que des différences phonétiques subtiles entre elles, elles pourraient simplement avoir été confondues. Une telle possibilité serait cohérente avec les résultats de l'étude de production. Il est ainsi possible que les participants aient interprété la prosodie dans les préfaces neutres comme étant positive ou *vice-versa*. Il est également possible que cette égalité de représentation des réponses relève d'une différence dans la fréquence d'utilisation des feedbacks : le feedback « ah oui » pourrait être simplement plus régulièrement utilisé que le feedback « j'écoute » dans la vie de tous les jours, et aurait pu être parfois préféré à celui-ci pour cette raison en dépit de la valence. Ce phénomène n'a cependant pas été observé dans la tâche de validation. Nous penchons donc plus vers la possibilité que les préfaces projetant des nouvelles positives ou neutres aient été confondues.

Nous présentons une autre interprétation. Pour bien comprendre celle-ci, nous allons nous attarder quelque peu sur les résultats de l'étude de validation. Dans cette étude préliminaire, nous avons validé les trois feedbacks comme étant bien positif (« ah oui »), négatif (« oh non ») et neutre (« j'écoute »). Si l'on s'attarde à l'observation des résultats de façon minutieuse, nous remarquons certains éléments intéressants. Lorsque l'on oppose valence positive et valence négative, « ah oui » et « oh non » sont clairement catégorisés comme positif pour l'un et négatif pour l'autre. En revanche, lorsque ces valences de nouvelle sont opposées à la valence neutre, les feedbacks sont moins souvent choisis par rapport à l'option de feedback neutre « j'écoute » (voir Annexe 19). Il est significativement préféré lorsque la nouvelle est neutre : 88% du temps lorsqu'elle est opposée à la valence négative, et 69% lorsqu'elle est opposée à la valence positive. Cependant, « j'écoute » est aussi très utilisé lorsque la valence de la nouvelle est négative ou positive, car il a été choisi dans les deux cas environ 52% du temps. Il est possible que l'absence d'autres indices que la seule prosodie de la préface, sortie de tout contexte, pousse les personnes à choisir le feedback « j'écoute », plus neutre, pour prendre moins de risque interactionnel et s'assurer de ne pas se désaffilier du point de vue émotionnel et attitudinal (« *stance* ») proposé par le locuteur (Stivers, 2008). Le fait d'utiliser un feedback neutre serait alors plus adapté, permettant de signifier de l'attention sans

risquer la désaffiliation. Le choix de « j'écoute » face à une préface qui annonce une nouvelle positive reste en revanche étonnant. Il se pourrait que d'autres facteurs de fréquence d'utilisation des mots ou de contexte social entrent en compte, et induisent un certain type de réponse plutôt qu'un autre. Au vu de ces résultats de la tâche de validation, nous nous serions attendus à ce que dans la tâche principale le feedback neutre « j'écoute » soit également préféré lorsque l'on oppose valence neutre et valence positive. Or, ce n'est pas le cas. Serait-il alors possible que la réponse « ah oui » ait été considérée comme un feedback neutre ? Notre feedback « ah oui » est très proche du feedback générique « oui » (Bertrand, 2021). La fonction de ce feedback est celle de continuation : elle laisse savoir au locuteur qu'il peut continuer à parler, et dans notre cas à annoncer. Cette fonction de continuation est celle que nous attribuons à « j'écoute » dans notre étude. Si « ah oui » et « j'écoute » ont tous deux été perçus comme génériques dans notre tâche principale, alors cela pourrait expliquer le fait que tous deux soient indifféremment utilisés à l'écoute des préfaces neutres. Alternativement, les participants à l'étude ont parfois mentionné deux choses : le feedback « j'écoute » serait plus professionnel et plus poli, et plus cohérent si l'on ne s'attend pas à recevoir la nouvelle. Le feedback « ah oui » serait plus familier, et plus cohérent si l'on s'attend à recevoir la nouvelle. Ces facteurs de politesse, de familiarité et de surprise évoqués sont pour nous un autre indicateur de la mécompréhension de notre consigne pour les tâches de perception. Ils pourraient avoir influencé le choix des participants à l'étude, qui seraient alors potentiellement influencés dans leur réaction affective par la thématique de la nouvelle.

Un deuxième résultat important de notre étude correspond aux différences significatives trouvées pour le choix des réponses en fonction du genre du participant. Nous avons trouvé dans la tâche 1 que les hommes choisissaient plus souvent la réponse « ah oui » que les femmes lorsque la valence de la nouvelle est positive. Cependant, ils sélectionnent également ce feedback plus souvent lorsque la nouvelle est négative : ils ont donc préféré « ah oui » indépendamment de la valence de la nouvelle. Dans la tâche 2, hommes comme femmes réagissaient plus souvent de façon congruente à la valence neutre (avec le feedback « j'écoute ») qu'à la valence négative. Ils ont aussi plus souvent utilisé « j'écoute » lorsque la valence de la nouvelle était négative. Dans la tâche 3, nous n'avons trouvé aucune différence significative. Les différences trouvées entre hommes et femmes dans la tâche 1 ne sont pas en accord avec nos hypothèses ou avec la littérature

sur la perception de la prosodie affective en fonction du genre (Beaucousin et al., 2011; Besson et al., 2002; Schirmer et al., 2002). Nous pensions en effet que bien que le genre puisse impacter le temps de réaction, il n'impacterait pas le choix de réponse. Le choix de « ah oui » indifférencié entre les valences pour les hommes, et différencié entre les valences pour les femmes, pourrait indiquer que les femmes font plus attention aux indices de projection prosodiques et se préparent donc plus sur le plan affectif en fonction de la valence perçue de la nouvelle.

Additionnellement, il est possible que les femmes soient plus enclines à choisir un feedback à valence marquée (« ah oui », « oh non ») lorsque la valence de la nouvelle est elle-même marquée (positive/négative). Une hypothèse pour expliquer cette différence pourrait être une tendance des femmes (dépendamment du contexte) à utiliser un langage plus « affiliatif » que les hommes, qui eux utiliseraient un langage plus assertif (Leaper & Ayres, 2007). Ainsi, il est possible que les femmes aient eu plus tendance que les hommes à choisir un feedback en affiliation avec la valence perçue dans la préface. Il a par ailleurs été démontré que lorsqu'un interlocuteur interrompt le locuteur en utilisant des « réponses minimales » (que l'on a par la suite appelé feedbacks), les femmes étaient là encore plus prônes à interrompre pour montrer leur accord et s'affilier au locuteur, là où les interruptions des hommes seraient plus neutres avec une tendance au désaccord et à la désaffiliation, et à garder une distance avec le locuteur (Makri-Tsilipakou, 1994).

Vis-à-vis des temps de réaction, nous avons trouvé que les temps de réaction ne semblent pas affectés par la valence de la nouvelle et le genre dans les tâches 1, 2 et 3. L'étude combinée des temps moyens de réaction sur les trois tâches nous a également montré que le genre seul ou le niveau d'empathie seul n'influencent pas significativement le temps de réaction des auditeurs. En revanche, le temps de réaction était affecté par une interaction entre les facteurs de genre et de niveau d'empathie. Les hommes avec un niveau d'empathie haut ont mis en moyenne plus de temps à répondre que les femmes avec un niveau d'empathie bas. A l'inverse, les hommes avec un niveau d'empathie bas (3618 ms) et les femmes avec un niveau d'empathie haut (3612 ms) ont eu des temps de réaction très similaires. Ceci tend à indiquer que bien que la différence ne ressorte pas statistiquement, les femmes pourraient généralement avoir un temps de réaction un peu plus court que celui des hommes. Cette tendance serait en accord avec nos hypothèses initiales sur le genre et le temps de réaction, ainsi qu'en accord avec la littérature

(Beaucousin et al., 2011; Besson et al., 2002; Schirmer et al., 2002, 2004). Les femmes s'appuieraient donc plus fortement sur la prosodie affective que les hommes pour comprendre le sens affectif d'un énoncé, et que les hommes mettraient plus de temps à utiliser la prosodie affective qui n'est pas l'indice qu'ils utilisent en priorité. La différence croisée entre empathie et genre montre également une tendance de différence de temps de réaction en fonction de l'empathie, qui n'est pas non plus ressortie au niveau statistique. Cette tendance indique que les personnes à l'empathie basse répondraient plus vite que celles à l'empathie haute. Ce n'est pas en ligne avec notre hypothèse initiale selon laquelle les auditeurs avec un niveau d'empathie plus haut auraient des temps de réaction plus rapides (Aziz-Zadeh et al., 2010), et nous pensons que les temps de réaction pourraient s'allonger chez les personnes avec un plus haut niveau d'empathie en conséquence de certains processus cognitifs liés aux neurones miroirs. Enfin, nous trouvés que les temps de réaction ne semblent pas affectés par le facteur de congruence entre la valence de la nouvelle projetée par la préface, et la valence rattachée au feedback choisi comme réaction affective. Cela va à l'encontre de notre hypothèse selon laquelle ils se raccourciraient lorsque la réponse choisie serait congruente, grâce à un effet de *priming* de la valence dans la préface (Hommel, 1997; Pell et al., 2022).

Cette étude nous a donc permis d'observer des différences de réaction de la part des participants en fonction de la valence de la nouvelle projetée par la préface qu'ils écoutent. Elle nous a permis de démontrer qu'une différence est perçue entre les préfaces selon la valence de la nouvelle qu'elles précèdent. Cela montre que les interlocuteurs sont capables de récupérer les indices de projection prosodique de la valence affective produits dans la préface par les locuteurs, en accord avec les résultats de Swerts et Hirschberg (2010). De plus, notre étude confirme qu'il existe une forme de préparation interactionnelle de l'interlocuteur, lorsqu'il anticipe la valence de la nouvelle. L'étude laisse néanmoins la place à quelques doutes quant à l'interprétation exacte des choix de feedbacks faits par les participants. Nous avons émis l'hypothèse que la mise en place de l'étude selon un cadre interactionnel (écoute de la préface et proposition entre deux choix de feedbacks écrits) a pu pousser les participants à choisir le feedback qu'ils auraient répondu à l'oral malgré nos consignes, et donc à ne pas choisir leur réponse uniquement en fonction de leur réaction affective. Cela les aurait ainsi fait sélectionner leur réaction selon les fonctionnements interactionnels de la préférence et de la protection de la face.

La structure même de l'étude est donc problématique, car elle cause un flou entre l'aspect interactionnel et l'aspect purement perceptif et de réaction affective que nous souhaitons vérifier.

Nous pensons que cela aurait pu être évité. Nous pourrions à l'avenir proposer une étude comportementale à choix forcé où les participants écoutent les mêmes audios, et notent s'ils s'attendent à recevoir une nouvelle positive ou négative en sélectionnant « bonne nouvelle », « mauvaise nouvelle », ou « nouvelle neutre ». Cela permettrait une meilleure appréhension de la reconnaissance des indices prosodiques de la valence affective par les auditeurs, car aucune influence de la préférence ou de la protection de la face n'interviendrait dans ces réponses. De même, il serait intéressant de s'attarder sur une étude où, contrairement à celle que nous avons présentée dans le chapitre, les participants devraient sélectionner la réponse « appropriée » à la préface entendue. Nous vérifierions alors comment la perception de la valence de la nouvelle dans la préface influence le choix de feedback verbal produit par un interlocuteur. Enfin, nous pourrions aussi envisager une étude dans laquelle des mesures électrodermales seraient prises. Ces mesures seraient un moyen d'observer la réaction physiologique des individus à l'écoute des préfaces, et donc un moyen d'observer la réaction affective aux indices projetés de la valence d'une autre façon. Cette observation permettrait de vérifier que les interlocuteurs à qui l'on annonce une nouvelle, positive, négative ou neutre, sont bien préparés sur le plan affectif par la perception d'indices prosodiques de la valence affective de la nouvelle.

Discussion générale

Cette thèse s'est donnée pour objectif d'évaluer la fonction projective de la prosodie affective. Le thème des indices de projection prosodique, leur identité et leur rôle, a été abordé en production et en perception. En particulier, nous avons exploré comment la valence affective des annonces de nouvelles peut être portée par des indices phonétiques et phonologiques en amont de l'annonce, dans la préface produite juste avant ainsi que dans des marqueurs discursifs, et si ces indices sont perçus et utilisés par l'interlocuteur pour se préparer affectivement. Le second but de cette thèse était donc de vérifier comment la prosodie affective est produite dans les annonces de nouvelles en fonction de la valence affective (positive, négative, ou neutre), par quels canaux elle est véhiculée, et quelle réaction elle peut éliciter chez l'interlocuteur.

Nous allons commencer ce chapitre par le résumé des découvertes les plus importantes de notre thèse. Puis, nous allons faire le point sur les apports méthodologiques que ce travail peut offrir. Nous tenterons aussi d'analyser ce que nos résultats peuvent apporter aux problématiques plus générales liées à l'étude phonétique et phonologique de la voix affective et à la multidisciplinarité.

Résumé des attentes et des découvertes

Notre travail a démarré par la création d'un corpus d'annonces de nouvelles et l'exploration de la production phonétique (Chapitre 2) et phonologique (Chapitre 3) de la prosodie dans les préfaces et les marqueurs discursifs produits juste avant la nouvelle. Ce choix est motivé en particulier par un travail de Swerts et Hirschberg (2010) sur les indices anticipatoires véhiculés dans la variation acoustique de la prosodie dans les préfaces aux annonces de nouvelles. En ligne avec leurs résultats, notre travail démontre que certains indices phonétiques sont bel et bien produits par les locuteurs, et perçus par les interlocuteurs dans la préface à l'annonce d'une nouvelle. Additionnellement, nous avons aussi trouvé de la variation au niveau phonologique de l'intonation et son implémentation phonétique, ainsi qu'une influence sur la perception et la réaction affective des interlocuteurs à qui l'annonce est destinée.

En accord avec Swerts et Hirschberg (2010), nous avons trouvé dans le chapitre 2 que des variations phonétiques sont produites systématiquement par les locuteurs en fonction de la valence affective de la nouvelle qu'ils s'apprêtent à annoncer, au niveau de la préface ainsi que des marqueurs discursifs « voilà » et « donc » produits avant la séquence d'annonce. Spécifiquement, nous avons vu que des variations de la hauteur de la fréquence fondamentale des locuteurs et des variations de durée des pauses et des énoncés sont produites en fonction de la valence affective de la nouvelle, qui pouvait être positive, négative, ou neutre. Suivant la suggestion faite en conclusion de l'article de Swerts et Hirschberg (2010), nous avons aussi cherché de potentiels indices de variation de la valence au niveau des caractéristiques spectrales de la voix, et en avons trouvé. En cohérence avec la littérature (Berry & Brown, 2019; Freese & Maynard, 1998; Scherer, 2003; Swerts & Hirschberg, 2010), nous attendions, pour la fréquence fondamentale, des valeurs maximum, minimum et moyennes élevées, et un empan vocal élargi lorsque la valence de la nouvelle à annoncer était positive. A l'inverse, nous attendions des valeurs abaissées lorsque la valence de la nouvelle était négative. Nous avons aussi fait plusieurs mesures relatives à la temporalité des énoncés : vitesse d'articulation, durée des énoncés, durée des pauses. Nous pensions trouver une vitesse d'articulation plus lente, et des durées plus longues d'énoncés et de pauses lorsque la valence de la nouvelle était négative, et au contraire une vitesse accélérée et des durées plus courtes pour les nouvelles positives. Pour les caractéristiques spectrales de la voix, nous avons pris trois mesures régulièrement utilisées pour ce type d'analyses : la différence entre les deux premières harmoniques (H1-H2) et le ratio bruit harmoniques (HNR) (Eskenazi et al., 1990; Jiang & Pell, 2017; Krom, 1995; Kuo, 2013; Yang et al., 2020; Yumoto et al., 1982), ainsi que la proéminence du pic cepstral lissée (CPPs) (Stoller & Dixon, 2016; Yoon et al., 2020). Chacune de ces mesures permet de rendre compte de la voix lorsqu'elle est soufflée par rapport à lorsqu'elle est modale. Nous nous attendions à trouver une voix qui serait plus soufflée pour la valence négative, et une voix plus modale pour la valence positive, selon les différences trouvées pour la tristesse et la joie (Johnstone & Scherer, 2000; Laukkanen et al., 1996; Scherer, 2003; Scherer, 1986; Yuan et al., 2002). Pour chacune des mesures mentionnées jusqu'à maintenant, les productions en fonction de la valence positive ou négative étaient aussi comparées à celles où la valence de la nouvelle était neutre. Nous pensions que les valeurs des préfaces produites pour les nouvelles neutres seraient situées entre les deux autres : valeurs de f_0 plus basses, durées plus longues et

vitesse d'articulation plus lente, voix plus soufflée que pour le positif, mais valeurs de f_0 plus hautes, durées plus courtes et vitesse d'articulation plus rapide, voix plus modale que pour le négatif. Dans l'ensemble, nos résultats ont confirmé les hypothèses pour les valeurs de la fréquence fondamentale, et cela se voyait principalement au niveau de la moyenne de la f_0 et de l'empan vocal, ainsi que pour le maximum. La f_0 minimum semble être moins informative. Nos hypothèses sont aussi confirmées pour les caractéristiques spectrales, les mesures les plus informatives étant celles de la HNR et de H1-H2, le CPPs n'ayant lui montré que peu de différences significatives. Les mesures de durées ont démontré des différences en fonction de la valence, mais les elles sont assez faibles, et ne se retrouvent pas dans toutes les zones observées. La vitesse de parole enfin, n'a démontré aucune différence significative en fonction de la valence de la nouvelle, contrairement à nos attentes vis-à-vis de la littérature (Scherer, 2003; Swerts & Hirschberg, 2010; Vidrascu & Devillers, 2007). Lorsque la valence de la nouvelle était neutre, les mesures prises (ex. f_0 , durées) avaient tendance à se rapprocher de celles de la nouvelle positive, bien que la différence reste régulièrement suffisamment constante pour être considérée comme significative.

Les effets des mesures spectrales sont complexes à interpréter, car contrairement aux autres mesures qui peuvent être moyennées sur l'ensemble de l'énoncé, il est plus juste pour les mesures spectrales de faire une comparaison au sein de phonèmes identiques ou de même type (Bitouk et al., 2010). Nous avons donc utilisé un modèle statistique sur lequel nous revenons plus tard dans cette conclusion. Il nous permet de comparer des énoncés phonologiquement identiques, phonème par phonème. Ainsi, les caractéristiques spectrales étaient évaluées au sein de chaque segment de l'énoncé observé. En conséquence, certains segments (ex. consonnes non-voisées) ne démontraient aucune différence spectrale. Ainsi, si nous trouvons plusieurs phonèmes sur lesquels la voix est plus soufflée lorsque la valence est négative, ces effets de « voix soufflée » ne sont pas constants tout au long de l'énoncé contrairement aux valeurs de f_0 , qui présentent des résultats beaucoup plus linéaires.

En plus de causer des variations phonétiques dans la prosodie, la valence d'une nouvelle semble causer de la variation au niveau de patrons phonologiques, et particulièrement sur l'implantation phonétique de ceux-ci, par la valeur mélodique et l'alignement des cibles tonales (Chapitre 3). Le choix de l'ajout de telles analyses est

motivé par le fait que depuis quelques années, certaines études remettent en question la notion de dualité de structure et s'attaquent à l'exploration des corrélats phonologiques intonationnels de l'affect (Cao et al., 2014; Liscombe et al., 2003; Stibbard, 2000; Wang et al., 2020). Nous nous attendions à trouver un grand nombre de patrons phonologiques rattachés à des fonctions linguistiques, mais aussi à l'apparition de patrons liés au sens paralinguistique. Cette étude était exploratoire. Nous avons néanmoins émis l'hypothèse qu'en vertu du Code de Fréquence (Gussenhoven, 2002; Ohala, 1983; Rodero, 2011) nous pourrions trouver des patrons démontrant un plus grand nombre de tons hauts dans la préface lorsque la valence de la nouvelle est positive ou neutre (en lien avec l'aspect non-menaçant d'une f_0 élevée) et un plus grand nombre de tons bas lorsque la valence de la nouvelle est négative (en lien avec l'aspect menaçant d'une f_0 abaissée). Nous avons aussi proposé une influence des valeurs trouvées pour la f_0 dans notre analyse acoustique, qui sont congruentes avec cette hypothèse. Nous avons trouvé des différences significatives, avec des patrons comportant des tons bas plus représentés lorsque la nouvelle est négative par rapport à lorsqu'elle est neutre, dans les préfaces et dans les marqueurs discursifs. Cela corrobore ce qui a été trouvé dans les études concernant l'existence d'un lien entre affect et phonologie intonationnelle (Cao et al., 2014; Liscombe et al., 2003; Stibbard, 2000; Wang et al., 2020). Les résultats sont donc en ligne avec nos hypothèses. Nous nous attendions à trouver des patrons intonationnels rattachés à la valence affective de la nouvelle et, pour les patrons qui seraient rattachés au sens linguistique, nous attendions différents patrons selon ce qui était observé. Dans la préface, nous attendions des configurations de fin d'IP qui marquent l'affirmation, soit $L^*L\%$ (Delais-Roussarie et al., 2015), nos préfaces étant justement des affirmations. Dans les marqueurs discursifs « voilà » et « donc », nous attendions des patrons cohérents avec les fonctions structurelles des marqueurs, notamment : HiL^* et $LHiL^*$ pour la conclusion sur « voilà » (Petit, 2020), et L^* pour la conclusion sur « donc » (Ferré, 2011). L'implantation phonétique de ces patrons phonologiques présente aussi de la variation phonétique, dans la façon dont les patrons sont réalisés, horizontalement au niveau de l'alignement, et verticalement au niveau de la valeur mélodique (Ladd, 2008). Nous avons fait l'hypothèse que la valeur mélodique des cibles tonales réalisées dans les patrons serait cohérente avec la littérature (Berry & Brown, 2019; Freese & Maynard, 1998; Scherer, 2003) et nos résultats des analyses acoustiques pour la fréquence fondamentale : des cibles réalisées plus bas lorsque la valence est négative, plus haut lorsqu'elle est positive, entre les deux

lorsqu'elle est neutre. Les mesures de hauteur mélodique ont confirmé ces hypothèses : les effets tendent dans la même direction que nos mesures globales de la f_0 , et dans la même direction que ce qui est décrit dans la littérature sur les affects et sur les annonces de nouvelles. Les mesures d'alignement enfin, étaient exploratoires. Nous avons trouvé des différences d'alignement des cibles tonales en fonction de la valence de la nouvelle, mais ces résultats comportent beaucoup de variabilité et ne se retrouvent qu'au sein des marqueurs discursifs. Le résultat le plus stable correspond au fait que les cibles tonales des accents mélodiques H^* et L^* sont produites plus tôt la plupart du temps lorsque la valence de la nouvelle est neutre, par rapport à lorsqu'elle est positive ou négative.

Enfin, nous avons souhaité regarder si les indices de projection produits dans les préfaces aux annonces de nouvelles aidaient l'interlocuteur à se préparer à la valence de la nouvelle en regardant si les indices agissaient sur leur ressenti affectif (Chapitre 4), suivant Swerts et Hirschberg (2010). Nous avons trouvé, en ligne avec cette étude, que les auditeurs font une différence entre les préfaces selon la valence affective de la nouvelle. Nous avons observé cette perception en demandant aux participants de choisir la réponse se rapprochant le plus de leur ressenti affectif entre trois feedbacks : « ah oui » (positif), « oh non » (négatif), « j'écoute » (neutre). Nous nous attendions à ce que les auditeurs choisissent le feedback dont la valence serait accordée avec leur perception de la valence affective de la préface. Les résultats sont cohérents avec nos hypothèses : nous avons effectivement trouvé des différences dans le choix des réponses affectives par les participants, qui ont souvent choisi la réponse « ah oui » lorsque la valence de la nouvelle projetée par la préface était positive, et « j'écoute » lorsque la valence de la nouvelle était neutre. Le fait que la valence positive soit bien différenciée du négatif n'est pas étonnant au vu des fortes différences phonétiques observées entre les deux dans notre étude de production. Le résultat trouvé pour le neutre est aussi cohérent avec les résultats des études de production. En revanche lorsque les valences neutres et positives étaient opposées entre elles pour la perception, elles ont mal été distinguées l'une de l'autre et ont élicité des choix de réaction plutôt positives. Cela est aussi cohérent avec les résultats observés pour leurs mesures phonétiques, qui n'étaient pas toujours significatives et dont les effets étaient moins importants que pour les autres comparaisons de valences. Cependant, les effets de l'étude recèlent une certaine ambiguïté : il semble y avoir un effet des règles interactionnelles de protection de la face (P. Brown & Levinson, 1987; Goffman,

1955, 1967) et de préférence (Heritage, 1989; Sacks, 1973; Schegloff, 1972) qui entrent en jeu dans le choix des auditeurs, malgré les consignes données au début des tâches. La réponse « oh non » est, en effet, systématiquement moins choisie que les deux autres même lorsque la valence positive et la valence négative sont opposées, alors qu'elles sont phonétiquement les plus distinctes. Au-delà des différences causées par la valence de la nouvelle, nous avons exploré les effets du genre sur les réponses données par les participants, et avons trouvé que les femmes semblent plus faire la différence en termes de préparation affective pour les nouvelles positives et neutres que les hommes. De plus, nous avons trouvé que les hommes avec un niveau d'empathie bas ont eu des temps de réaction plus longs en moyenne que ceux des femmes avec un niveau d'empathie haut. Ce résultat est cohérent avec nos hypothèses et la littérature (Schirmer et al., 2002, 2004), les femmes utilisant les indices prosodiques plus et plus vite que les hommes, qui eux utilisent plus les indices verbaux de la valence affective. L'empathie haute semble avoir ralenti le temps de réaction par rapport à l'empathie basse, bien que ce facteur doive être associé à celui du genre pour ressortir statistiquement. Cela allait à l'encontre de nos hypothèses mais est cohérent avec l'idée que l'empathie pousse les individus à simuler les affects pour reconnaître la prosodie affective (Aziz-Zadeh et al., 2010), et avec le fait que cette simulation pourrait ralentir le temps de réaction.

Apports méthodologiques

Au cours de cette thèse, nous avons utilisé plusieurs méthodologies de traitement statistique pour mieux évaluer les corrélats phonétiques et phonologiques des indices de projection de l'affect. Les deux modèles statistiques les plus utilisés étaient les modèles mixtes : le modèle LMER (*Linear Mixed-Effects Regression*) et le modèle GLMER (*Generalized Linear Mixed-Effects Regression*). Ce sont des approches traditionnelles couramment utilisées dans les études en linguistique et en prosodie. Cependant, l'un de leurs inconvénients majeurs réside dans leur incapacité à modéliser efficacement les relations non linéaires, puisqu'ils assument une relation linéaire entre les variables dépendantes et indépendantes, ce qui limite leur applicabilité dans de nombreuses situations du monde réel. Dans nos données, nous avons justement voulu vérifier si la fréquence fondamentale est produite avec des différences de hauteur en fonction de la valence affective de manière constante, linéaire, au cours de l'unité observée, ou si cette variation était située à certains points clés de l'énoncé. De même pour faire l'analyse des

caractéristiques spectrales de la voix, nous avons dû trouver un autre type de modèle pour observer le déroulement au cours de l'énoncé, segment par segment. Les mesures spectrales sont en effet dépendantes du phonème observé, et nous ne pouvons donc pas comparer des phonèmes différents, ou faire des mesures globales pour montrer les tendances qui se dégagent en fonction de la valence. Nous nous sommes donc tournés vers les modèles GAMM (*Generalized Additive Mixed-Effects Models*), suivant l'exemple de travaux en linguistique (Porretta et al., 2016; Wieling, 2018; Xiao et al., 2023). Ils offrent la flexibilité nécessaire pour modéliser des relations non linéaires (ex. séries temporelles) en incorporant des termes lisses en plus des variables indépendantes et des effets aléatoires, ce qui permet de capturer des tendances complexes (Sóskuthy, 2017; Wood, 2017). Cette méthode nous a permis de modéliser les courbes de H1-H2, de la HNR et du CPPs, et y repérer les points temporels sur lesquels les différences sont trouvées. Les GAMM produisent aussi des valeurs p pour les données analysées de façon linéaire (faisant ainsi un calcul sur l'ensemble de l'énoncé sans prise en compte du temps), et nous avons pu noter qu'elles n'étaient pas significatives. En revanche, nous trouvions des régions temporelles précises sur lesquelles nous voyions apparaître des différences significatives. Le fait que certaines régions ressortaient comme significatives et d'autres non confirme le caractère non-linéaire des caractéristiques spectrales. Pour les données de la courbe de f0 en revanche, les valeurs p ont montré des différences significatives lorsque les données étaient considérées comme ayant une relation linéaire. On voit par ailleurs sur la courbe de f0 modélisée par valences que les différences (surtout entre les valences négative et positive) sont étendues sur l'ensemble de la courbe, que ce soit sur la préface ou sur les marqueurs discursifs. Les zones non significatives correspondent principalement aux consonnes occlusives ou fricatives, soit des segments où la détection de la f0 est perturbée. Les résultats de cette analyse suggèrent que les mesures globales de f0 sont adaptées à l'étude des affects, et qu'une analyse temporelle n'est pas forcément nécessaire. L'utilisation de modèles GAM dans l'étude de la parole affective serait donc intéressante pour rendre compte des données non-linéaires, lorsque des phrases identiques sont comparées. Il serait intéressant d'étendre ce type d'analyse à d'autres mesures traditionnellement prises de façon globale pour vérifier, à nouveau, si celles-ci sont pertinentes, ou si nous devons nous intéresser plus étroitement à l'emplacement du signal où les différences peuvent avoir lieu. Par exemple, les GAMM pourraient être un

outil d'observation de la durée des syllabes au cours de l'énoncé, ou de la vitesse d'articulation.

Un autre élément intéressant de nos analyses statistiques est le coefficient d'accord utilisé pour rendre compte de l'accord inter-annotateurs réalisé sur les annotations phonologiques. Nous avons utilisé une méthodologie de mesure de l'accord inter-annotateur différente de mesures plus standard : le coefficient d'accord de Gwet (2001). Nous avons fait ce choix en raison de la robustesse des coefficients d'accord de Gwet à l'égard de la prévalence des traits, un problème qui se pose lorsque les catégories analysées sont distribuées de manière asymétrique (Wongpakaran et al., 2013). Par exemple, si une catégorie est utilisée beaucoup plus fréquemment qu'une autre dans un corpus, cela crée une prévalence d'un trait sur les autres. Les autres tests courants, comme le kappa de Cohen, peuvent sous-estimer ou surestimer les accords de hasard dans de telles situations. Nos données présentaient un tel déséquilibre, d'où notre choix d'utiliser les coefficients de Gwet. Nous avons utilisé l'AC1 pour nos données catégorielles non ordonnées et l'AC2 pour nos données catégorielles ordonnées (Blood & Spratt, 2007). Dans notre étude, les coefficients de Gwet se sont avérés être une méthode robuste pour évaluer l'accord inter-annotateur compte tenu de la nature de nos données. Les coefficients AC1 et AC2 de Gwet nous apparaissent alors comme pertinents pour notre domaine d'études, et nous appuyons sur le fait que ces mesures devraient être envisagées comme une alternative intéressante aux coefficients traditionnellement utilisés, dans la recherche en sciences du langage.

Apports théoriques

Le Chapitre 2 de cette thèse présente des analyses phonologiques de l'affect, un champ d'études longtemps écarté par les chercheurs en intonation. En effet, les chercheurs utilisant les différentes théories de la phonologie de l'intonation (y compris la théorie AM) ont longtemps considéré que le sens porté par la prosodie fait face à une structure bicéphale. D'un côté, le sens linguistique peut être porté par des unités intonationnelles phonologiques ainsi que par des variations phonétiques continues que l'on trouve au niveau acoustique. D'un autre côté, le sens paralinguistique ne peut être porté que par le second canal, c'est-à-dire les variations phonétiques continues. Cette dualité de structure (Gussenhoven, 2004; Ladd, 2008; Prieto, 2015), ancrée dans la tradition linguistique, a poussé les chercheurs sur la prosodie affective à laisser de côté l'idée que des variations

intonationnelles phonologiques puissent véhiculer des informations affectives. Bien que le sujet ne soit jamais complètement oublié et même régulièrement mentionné, notamment à cause de la difficulté première à distinguer réellement le sens linguistique du sens paralinguistique (Ladd, 2008; Prieto, 2015), il n'est jamais mentionné que pour être écarté en raison notamment de sa complexité. Cette mise à l'écart a permis à la recherche de se concentrer vers un même but pendant de nombreuses années : la compréhension du rôle de patrons intonationnels dans le port d'informations linguistiques telles que le type de phrase, ou le focus. Bien que nous soyons loin d'avoir résolu les mystères du sens linguistique, de nombreux outils ont été proposés pour décrire et comprendre l'intonation, de la manière la plus fiable possible, et de façon à pouvoir comparer les découvertes de façon cross-linguistique. Notamment, la convention d'annotation ToBI (Silverman et al., 1992) et ses variantes permettent la comparaison du sens intonational pour de multiples langues. Le sujet complexe du sens paralinguistique devient depuis quelques années un champ d'étude attractif auquel les chercheurs recommencent à s'attaquer. Le sens affectif de même, fait depuis quelques temps l'objet d'études réalisées dans le cadre de la théorie AM notamment (Cao et al., 2014; Liscombe et al., 2003; Stibbard, 2001; Wang et al., 2020). Les études sur le sujet, encore rares, se concentrent sur la langue anglaise, dans sa variété américaine ou britannique. Cette remise en cause de la dualité de structure et ce nouvel élan d'intérêt étant extrêmement récents, il n'est pas étonnant que la littérature sur d'autres langues et familles de langues manque. Nous avons proposé ici une étude sur le français, en nous intéressant à l'affect en tant que « valence » qui peut être positive, négative ou neutre, en fonction de la teneur affective de nouvelles à annoncer. En ligne avec les découvertes des études mentionnées, nous avons trouvé un lien entre valence affective et patrons phonologiques. Cependant, nous n'avons trouvé que peu de résultats significatifs, et ne pouvons donc tirer beaucoup de conclusions. Premièrement, il semble que des patrons liés au sens linguistique peuvent porter une différence de sens affectif par leur présence plus ou moins forte dans certains contextes. Nous avons en effet trouvé que lorsque la valence affective était négative, les patrons principalement attendus pour leurs fonctions linguistiques étaient plus représentés que lorsque la valence de la nouvelle était neutre. Deuxièmement, puisque le sens affectif ne semble pas entièrement étranger à l'intonation, nous argumentons qu'il mérite que nous nous penchions sur le sujet dans de futures études, au travers de langues, et de catégorisations de l'affect, des émotions, ou des attitudes diverses et variées. Au-delà

des patrons purement phonologiques, nous avons découvert une variation dans l'implantation phonétique des tons en cibles tonales, sur le plan de la valeur mélodique ainsi que celui de l'alignement. Ce type d'analyse est aussi rare que celui des patrons phonologiques, si ce n'est plus. Nos résultats indiquent que ces différences pourraient être des indicateurs intéressants de la valence affective. Il serait donc pertinent dans les futures études de prendre en compte ces facteurs également. Par ailleurs, Cao et al. (2014) ont suggéré que ToBI pourrait être revu afin d'inclure ces informations de façon systématique dans l'annotation phonologique pour rendre compte des variations de sens. Au vu de nos découvertes, nous adhérons à cette proposition.

En ligne avec la volonté générale d'interdisciplinarité actuellement encouragée dans la recherche, nous avons utilisé tout au long de nos travaux des définitions, des concepts et des raisonnements qui nous viennent de l'Analyse Conversationnelle (Sacks et al., 1978) et de la Linguistique Interactionnelle (Couper-Kuhlen & Selting, 1996; Ochs et al., 1996), pour comprendre le rôle projectif de la prosodie sur le sens affectif. Nous nous sommes servis des analyses faites sur le fonctionnement interactionnel des annonces de nouvelles pour définir les énoncés que nous souhaitions observer, et créer notre corpus en conséquence : notamment, la préface, les marqueurs discursifs, et la séquence d'annonce. Cela nous a permis de formuler des hypothèses ancrées dans l'analyse de l'interaction autour d'une notion centrale à notre travail : la projection (Auer, 2005; Ford & Thompson, 1996; Goodwin, 2002; Schegloff, 1996), et son pendant perceptif, l'anticipation. Nous avons cherché à comprendre comment la prosodie peut apporter des indices essentiels dans les préfaces, liés à la valence affective de séquences interactionnelles qui sont préparées par le locuteur. Nous avons proposé des raisons pour l'utilisation de ces indices dans la préface aux annonces de nouvelles, telles que la « préparation ». Le locuteur projette des indices, proposant ainsi une trajectoire (Streeck, 1995) à la fois interactionnelle (de quoi va-t-on parler) et affective (quelle affectivité est liée à la séquence). Cette proposition de trajectoire permet au locuteur rendre pertinente la séquence suivante et de s'assurer de donner des clés d'anticipation à l'interlocuteur. Ainsi, l'interlocuteur a l'opportunité par les plans verbal et prosodique de comprendre que l'on démarre une séquence d'annonce, et de comprendre la teneur affective proposée pour cette séquence d'annonce. Cela augmente ses chances de pouvoir refuser l'annonce, s'il n'en veut pas, et s'il l'accepte de s'accorder au locuteur en termes d'alignement et

d'affiliation (Stivers, 2008). S'il accepte l'annonce, l'alignement et l'affiliation, s'ils sont atteints, permettent d'assurer au maximum la progressivité de l'interaction et son succès. Ce succès de l'interaction est en Linguistique Interactionnelle considéré comme un but commun tacite, inhérent à toute interaction. Le refus de l'annonce en amont permet d'éviter des situations complexes mettant en danger la face des interactants, que chacun tend à protéger. Ainsi, s'assurer en amont que l'annonce ne sera pas refusée, et limiter le risque de désaffiliation et désalignement est pour le locuteur essentiel, s'il veut réussir l'interaction (entre-autres ici, aller au bout de l'annonce), et s'il veut le faire en protégeant au maximum sa face et celle de celui qui reçoit la nouvelle.

Nos études permettent d'apporter de nouvelles informations sur des phénomènes attestés de façon empirique dans la littérature sur les interactions. Nous avons trouvé des différences systématiques en fonction de la valence affective de la nouvelle, dans des portions interactionnelles en partie dédiées à la projection d'indices sur la séquence suivante, la préface. Nous avons trouvé des indices produits aux niveaux phonétique et phonologique, et avons pu observer que ces indices sont perçus et utilisés par les interlocuteurs pour se préparer sur le plan affectif. Bien que les résultats de notre étude de perception aient pu être influencés par les fonctionnements interactionnels de préférence et de protection de la face, ils démontrent un certain degré de préparation sur le plan affectif de la part des auditeurs. Ceux-ci ont, en effet, fait des choix de « réaction émotionnelle » différents en fonction de la prosodie de la préface et de la valence de nouvelle qu'elle porte. Tout cela démontre pour nous le rôle projectif que peut prendre la prosodie, particulièrement pour porter le sens affectif, dans les préfaces d'annonces mais aussi ailleurs comme dans les marqueurs discursifs (pour nous, « voilà » et « donc »), et dans la séquence d'annonce elle-même. L'étude des annonces de nouvelles semble avoir été un type d'interaction pertinent et adapté pour comprendre et observer le rôle de la projection dans la préface aux annonces de nouvelles, le rôle que joue la prosodie dans la projection d'indices de valence spécifiquement, et la présence d'indices phonétiques et phonologiques pour porter cette projection prosodique.

Bien entendu, nos travaux se sont cantonnés à un type spécifique d'interaction : les annonces de nouvelles. Cela limite la portée de notre étude, et le rôle projectif de la prosodie affective mériterait d'être observé dans d'autres types d'interactions, telles que les pré-blaques, les pré-histoires, ou encore les pré-demandes (Fox, 2015; Levinson, 1983;

Lindström et al., 2005; Sacks, 1989; Schegloff, 2007). De plus, nous avons évoqué au cours des précédents chapitres les multiples défauts présents dans nos études. Tout d'abord, la manière dont a été enregistré notre corpus ne nous permet pas de faire des analyses de l'intensité, une mesure pourtant régulièrement prise par les chercheurs en prosodie affective, et qui varie en fonction de l'affect. Nous sommes donc potentiellement passés à côté d'indices additionnels de la valence affective, qui peuvent être déterminants dans la perception des auditeurs et dont nous ne pouvons attester. De plus, les annotations phonologiques que nous avons réalisées se cantonnent à des portions spécifiques du syntagme intonational dans la préface. Nous n'avons pas fait l'observation de configurations de fin d'AP autres de celles de fin d'IP dans les préfaces, ni l'observation de tons tels que ceux de la montée initiale. Par ailleurs, l'annotation des accents mélodiques dans les préfaces présente un problème non négligeable car nous avons obtenu pour celle-ci un accord inter-annotateurs faible, ce qui indique que l'annotation de ces tons n'est pas fiable. Cette portion des configurations mériterait d'être réannotée, et les analyses subséquentes d'être refaites en conséquence. Enfin, comme nous l'avons déjà évoqué, notre étude de perception pourrait avoir mesuré un phénomène lié aux règles interactionnelles qui entrent en jeu dans la préparation de l'interlocuteur, en plus d'avoir mesuré la préparation affective de celui-ci. Notre étude ne nous permet pas de déterminer avec complète assurance quelle part des effets trouvés sont liés à la préparation interactionnelle, et quelle part des effets sont liés à la préparation affective. Il est possible par exemple, que les auditeurs aient régulièrement hésité à faire le choix d'une réponse affective négative même si une préface leur apparaissait négative, ayant considéré que qu'une réponse à valence négative ne serait soit pas appropriée, soit trop risquée interactionnellement (avec notamment un risque de désaffiliation au cas où ils se tromperait sur la nature de la nouvelle). Enfin, des notions de politesse en fonction du contexte social de la nouvelle ont pu rentrer en jeu. Cette étude pourrait bénéficier de l'ajout d'une seconde tâche comportementale, simplement pour vérifier la perception des indices projectifs sans rendre compte de leur cadre interactionnel.

Ouverture sociale

En dehors des apports pour la recherche fondamentale et la compréhension des interactions humaines par la parole, notre travail a une importance sociétale. En effet, l'annonce de nouvelles est une interaction cruciale dans de nombreuses situations,

notamment dans le cadre professionnel ou institutionnel. Le cadre professionnel ayant élicité le plus de recherche sur la thématique est le domaine médical. Dans ce domaine, souvent en oncologie, le nombre d'études sur le sujet de l'annonce de nouvelles (phénomène souvent appelé « *breaking bad news* »), témoigne du désir et de la nécessité pour les soignants de se former à l'annonce de nouvelles, pour faciliter la gestion émotionnelle de l'annonce, pour faciliter le déroulement d'un traitement médical après annonce, mais aussi pour aider au mieux le patient à prendre conscience de la situation tout en essayant de protéger son bien-être psychologique (Boyd, 2005; Buckman, 2005; Freese & Maynard, 1998; Gueneuc et al., 2021; Maguire, 1998; Maynard, 2017; Vandekieft, 2001). La question de la formation des soignant a récemment fait l'objet de recherches pour créer des outils d'interaction homme-machine afin d'entraîner à l'annonce avec une interaction simulée virtuelle (Ochs et al., 2018).

Si la recherche sur les annonces de nouvelles en contexte médicale semble être la plus prolifique, il existe d'autres domaines dans lesquels on peut s'intéresser aux mécanismes de l'annonce de nouvelle. Tracy et Craig (2019) se sont par exemple attardés sur les annonces de décisions de justice sur des litiges aux États-Unis afin de montrer qu'il serait pertinent de normer la façon d'annoncer ces nouvelles, pour permettre aux juges de mieux les communiquer. McClenahan et Lofland (1976) se sont intéressés aux tactiques mises en place par les maréchaux adjoints (« *deputy US marshal* ») aux États-Unis lorsqu'ils doivent requérir la participation d'un individu à un procès, ou encore annoncer à une personne qu'elle sera gardée en détention provisoire en attendant une décision de justice. Mais nous pouvons aussi imaginer l'importance d'une annonce de nouvelle dans les meilleures conditions dans d'autres situations, comme lorsqu'un chef d'équipe annonce une décision à un employé, lorsqu'un patron licencie ou embauche, ou encore lorsqu'un professeur appelle un parent pour parler d'un événement positif ou négatif à l'école. D'un point de vue social, l'importance de la compréhension des interactions d'annonces de nouvelles pour améliorer au maximum l'acceptation par la personne concernée et le ressenti de celui qui annonce ne fait aucun doute. La compréhension du rôle projectif de la prosodie est donc cruciale pour permettre le développement de formations holistiques pour les soignants, et tout autre individu ayant endossé un rôle social amenant à de telles interactions.

Bibliographie

- Allwood, J., Nivre, J., & Ahlsén, E. (1992). On the semantics and pragmatics of linguistic feedback. *Journal of semantics*, 9(1), 1-26.
- Alpert, M., Kurtzberg, R. L., & Friedhoff, A. J. (1963). Transient voice changes associated with emotional stimuli. *Archives of General Psychiatry*, 8(4), 362-365.
<https://doi.org/10.1001/archpsyc.1963.01720100052006>
- Altarriba, J. (2012). Emotion and Mood : Over 120 Years of Contemplation and Exploration in The American Journal of Psychology. *The American Journal of Psychology*, 125(4), 409-422. <https://doi.org/10.5406/amerjpsyc.125.4.0409>
- Altman, D. G. (1990). *Practical Statistics for Medical Research*. Chapman and Hall, CRC.
<https://doi.org/10.1201/9780429258589>
- Arnhold, A., & Kyröläinen, A.-J. (2017). Modelling the Interplay of Multiple Cues in Prosodic Focus Marking. *Laboratory Phonology*, 8(1), 4.
<https://doi.org/10.5334/labphon.78>
- Arnold, D., Wagner, P., & Baayen, R. H. (2013). Using generalized additive models and random forests to model prosodic prominence in German. *Proceedings of Interspeech 2013*, 272-276. <https://doi.org/10.21437/Interspeech.2013-82>
- Arvaniti, A. (2020). The Phonetics of Prosody. In M. Aronoff (Éd.), *Oxford Research Encyclopedia of Linguistics*.
<https://doi.org/10.1093/acrefore/9780199384655.013.411>
- Arvaniti, A. (2022). The Autosegmental-Metrical Model of Intonational Phonology. In J. Barnes & S. Shattuck-Hufnagel (Éds.), *Prosodic Theory and Practice* (p. 24-64). The MIT Press. <https://direct.mit.edu/books/oa-edited->

volume/5259/chapter/3572254/The-Autosegmental-Metrical-Model-of-Intonational

- Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment : The case of Greek prenuclear accents. *Journal of Phonetics*, 26(1), 3-25.
<https://doi.org/10.1006/jpho.1997.0063>
- Assmann, P. (1999). Fundamental frequency and the intelligibility of competing voices. *Proceedings of the 14th International Congress of Phonetic Sciences*, 179-182.
- Audibert, N. (2008). *Prosodie de la parole expressive : Dimensionnalité d'énoncés méthodologiquement contrôlés authentiques et actés* [Linguistique, Institut National Polytechnique de Grenoble - INPG]. <https://tel.archives-ouvertes.fr/tel-00489924>
- Audibert, N., Aubergé, V., & Rilliard, A. (2005). *The prosodic dimensions of emotion in speech : The relative weights of parameters*. 525-528.
<https://doi.org/10.21437/Interspeech.2005-332>
- Audibert, N., Vincent, D., Aubergé, V., & Rosec, O. (2006). Evaluation of expressive speech resynthesis. *Corpora for research on Emotion and Affect*, 37-40. <http://www.lrec-conf.org/proceedings/lrec2006/workshops/W09/Emotion-proceeding.pdf#page=46>
- Auer, P. (2005). Projection in Interaction and Projection in Grammar. *Text - Interdisciplinary Journal for the Study of Discourse*, 25(1), 7-36.
<https://doi.org/10.1515/text.2005.25.1.7>
- Avanzi, M. (2013). Notes de recherche sur l'accentuation et le phrasé prosodique à la lumière des corpus de français. *Travaux neuchâtelois de linguistique*, 59, Article 59.
<https://doi.org/10.26034/tranel.2013.2942>

- Aziz-Zadeh, L., Sheng, T., & Gheytañchi, A. (2010). Common Premotor Regions for the Perception and Production of Prosody and Correlations with Empathy and Prosodic Ability. *PLoS One*, 5(1), e8759. <https://doi.org/10.1371/journal.pone.0008759>
- Baayen, H., & Millin, P. (2010). Analyzing Reaction Times. *International Journal of Psychological Research*, 3(2), 12-28. <https://doi.org/10.21500/20112084.807>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>
- Bachorowski, J.-A., & Owren, M. J. (1995). Vocal Expression of Emotion: Acoustic Properties of Speech Are Associated With Emotional Intensity and Context. *Psychological Science*, 6(4), 219-224. <https://doi.org/10.1111/j.1467-9280.1995.tb00596.x>
- Bakker, I., van der Voordt, T., Vink, P., & de Boon, J. (2014). Pleasure, Arousal, Dominance: Mehrabian and Russell revisited. *Current Psychology*, 33(3), 405-421. <https://doi.org/10.1007/s12144-014-9219-4>
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614-636. <https://doi.org/10.1037/0022-3514.70.3.614>
- Bänziger, T., Grandjean, D., Bernard, P. J., Klasmeyer, G., & Scherer, K. R. (2001). *Prosodie de l'émotion: Étude de l'encodage et du décodage*. <https://www.semanticscholar.org/paper/Prosodie-de-l%27%C3%A9motion-%3A-%C3%A9tude-de-l%27encodage-et-du-B%3%A4nziger-Grandjean/53b040e8f121c13e4fe20e3bc142b7bed9609f34>

- Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, 46(3), 252-267. <https://doi.org/10.1016/j.specom.2005.02.016>
- Barkhuysen, P., Krahmer, E., & Swerts, M. (2005). Problem detection in human-machine interactions based on facial expressions of users. *Speech Communication*, 45(3), 343-359. <https://doi.org/10.1016/j.specom.2004.10.004>
- Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient : An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders*, 34(2), 163-175. <https://doi.org/10.1023/B:JADD.0000022607.19833.00>
- Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion*, 12(4), 579-599. <https://doi.org/10.1080/026999398379574>
- Batliner, A., Fischer, K., Huber, R., Spilker, J., & Noth, E. (2000). Desperately Seeking Emotions or : Actors, Wizards, and Human BeingsSpeech & Emotion. *ISCA tutorial and research workshop. ITRW on Speech and Emotion*, Newcastle, Northern Ireland, UK. https://www.isca-speech.org/archive_open/speech_emotion/spem_195.html
- Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79(6), 941-952. <https://doi.org/10.1037/0022-3514.79.6.941>
- Beaucousin, V., Zago, L., Hervé, P.-Y., Strelnikov, K., Crivello, F., Mazoyer, B., & Tzourio-Mazoyer, N. (2011). Sex-dependent modulation of activity in the neural networks engaged during emotional speech comprehension. *Brain Research*, 1390, 108-117. <https://doi.org/10.1016/j.brainres.2011.03.043>

- Beckman, M., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The Original ToBi System and the Evolution of the ToBi Framework. In S.-A. Jun (Éd.), *Prosodic Typology: The Phonology of Intonation and Phrasing* (p. 9-54). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199249633.003.0002>
- Beckman, M., & Pierrehumbert, J. (1986). Intonational Structure in Japanese and English. *Phonology Yearbook*, 3, 255-309. <https://doi.org/10.1017/S095267570000066X>
- Beedie, C., Terry, P., & Lane, A. (2005). Distinctions between emotion and mood. *Cognition and Emotion*, 19(6), 847-878. <https://doi.org/10.1080/02699930541000057>
- Benedetti, J. (2012). *The Art of the Actor: The Essential History of Acting from Classical Times to the Present Day*. Routledge.
- Beňuš, S., Gravano, A., & Hirschberg, J. (2007). Prosody, emotions, and... « whatever ». *8th Annual Conference of the International Speech Communication Association*, 2629-2632. <https://doi.org/10.21437/Interspeech.2007-691>
- Beňuš, Š., & Rusko, M. (2009). Prosodic Characteristics and Emotional Meanings of Slovak Hot-Spot Words. In A. Esposito & R. Vích (Éds.), *Cross-Modal Analysis of Speech, Gestures, Gaze and Facial Expressions* (p. 18-27). Springer. https://doi.org/10.1007/978-3-642-03320-9_3
- Berry, M., & Brown, S. (2019). Acting in action : Prosodic analysis of character portrayal during acting. *Journal of Experimental Psychology: General*, 148(8), 1407-1425. <https://doi.org/10.1037/xge0000624>
- Bertrand, R. (2021). *Linguistique Interactionnelle : Du Corpus à l'Expérimentation* [Thesis, Aix Marseille Université]. <https://hal.science/tel-03500277>
- Bertrand, R., Blache, P., Espesser, R., Ferré, G., Meunier, C., Priego-Valverde, B., & Rauzy, S. (2006). Le CID-Corpus of Interactional Data- : Protocoles, conventions,

- annotations. *Travaux interdisciplinaires du Laboratoire parole et langage d'Aix-en-Provence (TIPA)*, 25, 25-55.
- Bertrand, R., German, J. S., Herment, S., Hirst, D. J., Michelas, A., Petrone, C., Portes, C., Tortel, A., & Welby, P. (2022). La prosodie au Laboratoire Parole et Langage : Histoire, recherches actuelles et perspectives. *Travaux Interdisciplinaires sur la Parole et le Langage*, 38. <https://doi.org/10.4000/tipa.5210>
- Besson, M., Magne, C., & Schön, D. (2002). Emotional prosody: Sex differences in sensitivity to speech melody. *Trends in Cognitive Sciences*, 6(10), 405-407. [https://doi.org/10.1016/S1364-6613\(02\)01975-7](https://doi.org/10.1016/S1364-6613(02)01975-7)
- Bigi, B. (2015). SPPAS - Multi-lingual approaches to the automatic annotation of speech. *The Phonetician. Journal of the International Society of Phonetic Sciences*, 111-112(ISSN:0741-6164), 54-69.
- Bigi, B., & Hirst, D. (2012). *SPeech Phonetization Alignment and Syllabification (SPPAS) : A tool for the automatic analysis of speech prosody*. 19. <https://hal.archives-ouvertes.fr/hal-00983699>
- Bigi, B., & Meunier, C. (2018). Automatic Segmentation of Spontaneous Speech. *Revista de Estudos da Linguagem*, 26(4). <https://doi.org/10.17851/2237-2083.26.4.1489-1530>
- Bigi, B., Meunier, C., Nesterenko, I., & Bertrand, R. (2010). Automatic detection of syllable boundaries in spontaneous speech. *7th International conference on Language Resources and Evaluation (LREC 2010)*, 3285-3292. <https://hal.archives-ouvertes.fr/hal-01393609>
- Bigi, B., & Priego-Valverde, B. (2019). Search for Inter-Pausal Units : Application to Cheese! corpus. *9th Language & Technology Conference: Human Language*

- Technologies as a Challenge for Computer Science and Linguistics*, 289-293.
<https://hal.science/hal-02428485>
- Bitouk, D., Verma, R., & Nenkova, A. (2010). Class-level spectral features for emotion recognition. *Speech Communication*, 52(7), 613-625.
<https://doi.org/10.1016/j.specom.2010.02.010>
- Blakemore, D. (1987). *Semantic Constraints on Relevance*. Blackwell.
[https://www.scirp.org/\(S\(351jmbntvnsjt1aadkozje\)\)/reference/referencespapers.aspx?referenceid=1502267](https://www.scirp.org/(S(351jmbntvnsjt1aadkozje))/reference/referencespapers.aspx?referenceid=1502267)
- Blood, E., & Spratt, K. F. (2007). *Disagreement on Agreement : Two Alternative Agreement Coefficients*. 186, 1-12.
- Boersma, P., & Weenink, D. (2023). *Praat : Doing phonetics by computer* (6.3.18) [Logiciel].
<http://www.praat.org/>
- Bolden, G. B. (2006). Little Words That Matter : Discourse Markers “So” and “Oh” and the Doing of Other-Attentiveness in Social Interaction. *Journal of Communication*, 56(4), 661-688. <https://doi.org/10.1111/j.1460-2466.2006.00314.x>
- Bolden, G. B. (2008). “So What’s Up?” : Using the Discourse Marker So to Launch Conversational Business. *Research on Language and Social Interaction*, 41(3), 302-337. <https://doi.org/10.1080/08351810802237909>
- Bolden, G. B. (2009). Implementing incipient actions : The discourse marker ‘so’ in English conversation. *Journal of Pragmatics*, 41(5), 974-998.
<https://doi.org/10.1016/j.pragma.2008.10.004>
- Bolinger, D. (1964). Around the Edge of Language : Intonation. *Harvard Educational Review*, 34(2), 282-296. <https://doi.org/10.17763/haer.34.2.4474051q78442216>

- Bonner, M. R. (1943). Changes in the speech pattern under emotional tension. *The American Journal of Psychology*, 56(2), 262-273.
<https://doi.org/10.2307/1417508>
- Bono, J. E., & Ilies, R. (2006). Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, 17(4), 317-334.
<https://doi.org/10.1016/j.leaqua.2006.04.008>
- Borràs-Comes, J., Sichel-Bazin, R., & Prieto, P. (2015). Vocative intonation preferences are sensitive to politeness factors. *Language and Speech*, 58(Pt 1), 68-83.
<https://doi.org/10.1177/0023830914565441>
- Boudin, A., Bertrand, R., Rauzy, S., Ochs, M., & Blache, P. (2022). A Multimodal Model for Predicting Feedback Position and Type During Conversation. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4143334>
- Bouvier, L., Monetta, L., Laforce, R. J., Vitali, P., Bocti, C., & Martel-Sauvageau, V. (2021). Progressive apraxia of speech in Quebec French speakers: A case series. *International Journal of Language & Communication Disorders*, 56(3), 528-548.
<https://doi.org/10.1111/1460-6984.12606>
- Boyd, E. A. (2005). Book Review : Bad News, Good News: Conversational Order in Everyday Talk and Clinical Settings. *Journal of Language and Social Psychology*, 24(3), 318-320. <https://doi.org/10.1177/0261927X05278184>
- Breen, M., Dilley, L. C., Kraemer, J., & Gibson, E. (2012). Inter-transcriber reliability for two systems of prosodic annotation : ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch). *Corpus Linguistics and Linguistic Theory*, 8(2), 277-312.
<https://doi.org/10.1515/cllt-2012-0011>
- Brestoff, R. (1995). *The Great Acting Teachers and Their Methods*. Smith and Kraus, Inc.

- Brown, L., & Prieto, P. (2017). (Im)politeness : Prosody and Gesture. In J. Culpeper, M. Haugh, & D. Z. Kádár (Éds.), *The Palgrave Handbook of Linguistic (Im)politeness* (p. 357-379). Palgrave Macmillan UK eBooks. https://link.springer.com/chapter/10.1057/978-1-137-37508-7_14
- Brown, L., & Prieto, P. (2021). Gesture and Prosody in Multimodal Communication. In D. Z. Kádár, M. Terkourafi, & M. Haugh (Éds.), *The Cambridge Handbook of Sociopragmatics* (p. 430-453). Cambridge University Press. <https://doi.org/10.1017/9781108954105.023>
- Brown, P., & Levinson, S. C. (1987). *Politeness : Some Universals in Language Usage*. Cambridge University Press.
- Bruce, G. (1977). *Swedish word accents in sentence perspective* (Vol. 12, p. 12). LiberLäromedel/Gleerup.
- Brusco, P., Vidal, J., Beňuš, Š., & Gravano, A. (2020). A cross-linguistic analysis of the temporal dynamics of turn-taking cues using machine learning as a descriptive tool. *Speech Communication*, 125, 24-40. <https://doi.org/10.1016/j.specom.2020.09.004>
- Bryant, G. A. (2010). Prosodic Contrasts in Ironic Speech. *Discourse Processes*, 47(7), 545-566. <https://doi.org/10.1080/01638530903531972>
- Buckman, R. A. (2005). Breaking bad news : The S-P-I-K-E-S strategy. *Community Oncology*, 2(2), 138-142. [https://doi.org/10.1016/S1548-5315\(11\)70867-1](https://doi.org/10.1016/S1548-5315(11)70867-1)
- Caballero, J. A., Vergis, N., Jiang, X., & Pell, M. D. (2018). The sound of im/politeness. *Speech Communication*, 102, 39-53. <https://doi.org/10.1016/j.specom.2018.06.004>
- Calhoun, S. (2010). How does informativeness affect prosodic prominence? *Language and Cognitive Processes*, 25(7-9), 1099-1140. <https://doi.org/10.1080/01690965.2010.491682>

- Campbell, N., & Mokhtari, P. (2004). Voice Quality : The 4 th Prosodic Dimension. *Proc. of the 15th Int. Congress of Phonetic Sciences*, 3, 2417-2420.
<https://www.semanticscholar.org/paper/Voice-Quality-%3A-the-4-th-Prosodic-Dimension-Campbell-Mokhtari/f6fa2c18456ba8a15c85fc4c3490b89e84a2ce2b>
- Cao, H., Beňuš, Š., Gur, R. C., Verma, R., & Nenkova, A. (2014). Prosodic cues for emotion : Analysis with discrete characterization of intonation. *Speech prosody (Urbana, Ill.)*, 2014, 130-134. <https://doi.org/10.21437/SpeechProsody.2014-14>
- Carre, A., Stefaniak, N., d'Ambrosio, F., Bensalah, L., Besche-Richard, C., & d'Ambrosio, F. (2013). The Basic Empathy Scale in Adults (BES-A) : Factor structure of a revised form. *Psychological Assessment*, 25(3), 679-691.
<https://doi.org/10.1037/a0032297>
- Caruso, D. R., & Mayer, J. D. (1998). The Multidimensional Emotional Empathy Scale (MDEES). *UNH Personality Lab*, 21.
- Cashion, J. L., Cody, M. J., & Erickson, K. V. (1986). 'You'll Love This One...' An Exploration Into Joke-Prefacing Devices. *Journal of Language and Social Psychology*, 5(4), 303-312. <https://doi.org/10.1177/0261927X8600500405>
- Cesare, A.-M. D. (2011). L'italien ecco et les français voici, voilà. Regards croisés sur leurs emplois dans les textes écrits. *Langages*, n° 184(4), 51-67.
- Chai, Y., & Garellek, M. (2022). On H1-H2 as an acoustic measure of linguistic phonation type. *The Journal of the Acoustical Society of America*, 152(3), 1856-1870.
<https://doi.org/10.1121/10.0014175>
- Chen, A., Gussenhoven, C., & Rietveld, T. (2004). Language-Specificity in the Perception of Paralinguistic Intonational Meaning. *Language and Speech*, 47(4), 311-349.
<https://doi.org/10.1177/00238309040470040101>

- Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36(4), 724-746.
<https://doi.org/10.1016/j.wocn.2008.06.003>
- Clark, H. H. (1996). *Using language*. Cambridge university press.
- Col, G., Danino, C., & Bikialo, S. (2020). *Polysémie, usages et fonctions de « voilà »*. Walter de Gruyter GmbH & Co KG.
- Col, G., Danino, C., & Rault, J. (2015). Éléments de cartographie des emplois de voilà en vue d'une analyse instructionnelle. *Revue de Sémantique et Pragmatique*, 37, 37-59.
<https://doi.org/10.4000/rsp.1015>
- Couper-Kuhlen, E., & Selting, M. (1996). *Prosody in conversation : Interactional studies* (Cambridge University Press). Cambridge University Press.
- Couper-Kuhlen, E., & Selting, M. (2017). *Interactional Linguistics : Studying Language in Social Interaction*. Cambridge University Press.
- Cruttenden, A. (1997). *Intonation*. Cambridge University Press.
- Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *The Journal of the Acoustical Society of America*, 88(1), 101-112. <https://doi.org/10.1121/1.399955>
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25(4), 385-400. [https://doi.org/10.1016/0749-596X\(86\)90033-1](https://doi.org/10.1016/0749-596X(86)90033-1)
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121. <https://doi.org/10.1037/0096-1523.14.1.113>
- D'Ambrosio, F., Olivier, M., Didon, D., & Besche, C. (2009). The basic empathy scale : A French validation of a measure of empathy in youth. *Personality and Individual*

- Diefferences, 46, 160-165. *Personality and Individual Differences*, 46, 160-165.
<https://doi.org/10.1016/j.paid.2008.09.020>
- Dargnat, M. (2020). Subjectivité et projection : Le cas des particules discursives. *SHS Web of Conferences*, 78, 12007. <https://doi.org/10.1051/shsconf/20207812007>
- Delahaie, J. (2013). Vers une analyse sémantique (presque) unitaire des multiples emplois de voilà à l'écrit et à l'oral. *Revue de Sémantique et Pragmatique*, 33-34, 99-120.
- Delais-Roussarie, E., & Herment, S. (2018). Intonation et interprétation des questions : Un puzzle pluridimensionnel. In M.-J. Béguelin, A. Coveney, & A. Guryev (Éds.), *L'interrogative en français* (p. 51-78). Peter Lang. <https://halshs.archives-ouvertes.fr/halshs-01425412>
- Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S.-A., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H.-Y. (2015). Intonational phonology of French : Developing a ToBI system for French. In S. Frota & P. Prieto (Éds.), *Intonation in Romance* (p. 0). Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199685332.003.0003>
- Delattre, P. (1966). Les Dix Intonations de base du français. *The French Review*, 40, 141-146.
- Deliens, G., Antoniou, K., Clin, E., Ostashchenko, E., & Kissine, M. (2018). Context, facial expression and prosody in irony processing. *Journal of Memory and Language*, 99, 35-48. <https://doi.org/10.1016/j.jml.2017.10.001>
- Di Cristo, A. (1999). Vers une modélisation de l'accentuation du français : Première partie. *Journal of french language studies*, 9(2), 143-179.
- Dilley, L. C., & Heffner, C. C. (2013). The role of f0 alignment in distinguishing intonation categories : Evidence from American english. *Journal of Speech Sciences*, 3(1), Article 1. <https://doi.org/10.20396/joss.v3i1.15039>

- D'Imperio, M. (2000). *The role of perception in defining tonal targets and their alignment* [The Ohio State University].
<https://search.proquest.com/openview/001e2ef22ff1f6105571576c2d0745c4/1?pq-origsite=gscholar&cbl=18750&diss=y>
- D'Imperio, M. (2002). Language-specific and universal constraints on tonal alignment : The nature of targets and « anchors ». *Proc. Speech Prosody 2002*, 101-106.
- D'Imperio, M., Bertrand, R., Di Cristo, A., & Portes, C. (2006). The phonology and phonetics of prenuclear and nuclear accents in French. *Proceedings of ISCA Tutorial and Research Workshop on Experimental Linguistics*, 121-124.
<https://doi.org/10.36505/ExLing-2006/01/0024/000024>
- D'Imperio, M., Cavone, R., & Petrone, C. (2014). Phonetic and phonological imitation of intonation in two varieties of Italian. *Frontiers in Psychology*, 5.
<https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01226>
- D'Imperio, M., German, J. S., & Michelas, A. (2012). A multi-level approach to focus, phrasing and intonation in French. In *Prosody and Meaning* (p. 11-34). De Gruyter eBooks. <https://doi.org/10.1515/9783110261790.11>
- D'Imperio, M., & House, D. (1997). Perception of questions and statements in Neapolitan Italian. *Fifth European Conference on Speech Communication and Technology*.
https://www.academia.edu/download/67572877/Perception_Of_Questions_And_Statements_I20210609-6715-sfjkjv.pdf
- D'Imperio, M., & Michelas, A. (2014). Pitch scaling and the internal structuring of the Intonation Phrase in French. *Phonology*, 31(1), 95-122.
<https://doi.org/10.1017/S0952675714000049>

- Dohen, M., Løevenbruck, H., Cathiard, M.-A., & Schwartz, J.-L. (2004). Visual perception of contrastive focus in reiterant French speech. *Speech Communication*, 44(1), 155-172. <https://doi.org/10.1016/j.specom.2004.10.009>
- Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech : Towards a new generation of databases. *Speech Communication*, 40(1-2), 33-60. [https://doi.org/10.1016/S0167-6393\(02\)00070-5](https://doi.org/10.1016/S0167-6393(02)00070-5)
- Drew, P. (1984). Speakers' reporting in invitation sequences. In M. Atkinson & J. Heritage (Éds.), *Structures of Social Action : Studies in Conversation Analysis* (p. 129-151). Cambridge University Press. <https://emcawiki.net/bibtex/browser.php?key=Drew1984&bib=emca.bib>
- Drullman, R., & Bronkhorst, A. W. (2004). Speech perception and talker segregation : Effects of level, pitch, and tactile support with multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, 116(5), 3090-3098. <https://doi.org/10.1121/1.1802535>
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99(3), 550-553. <https://doi.org/10.1037/0033-295X.99.3.550>
- Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384-392. <https://doi.org/10.1037/0003-066X.48.4.384>
- Elordieta, G., & Prieto, P. (Éds.). (2012). *Prosody and Meaning*. De Gruyter Mouton. <https://doi.org/10.1515/9783110261790>
- Eskenazi, L., Childers, D. G., & Hicks, D. M. (1990). Acoustic correlates of vocal quality. *Journal of Speech and Hearing Research*, 33(2), 298-306. <https://doi.org/10.1044/jshr.3302.298>
- Esteve-Gibert, N., Schafer, A. J., Hemforth, B., Portes, C., Pozniak, C., & D'Imperio, M. (2020). Empathy influences how listeners interpret intonation and meaning when words

- are ambiguous. *Memory & Cognition*, 48(4), 566-580.
<https://doi.org/10.3758/s13421-019-00990-w>
- Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa : I. The problems of two paradoxes. *Journal of Clinical Epidemiology*, 43(6), 543-549.
[https://doi.org/10.1016/0895-4356\(90\)90158-1](https://doi.org/10.1016/0895-4356(90)90158-1)
- Fernandez, R., & Picard, R. (2011). Recognizing affect from speech prosody using hierarchical graphical models. *Speech Communication*, 53(9), 1088-1103.
<https://doi.org/10.1016/j.specom.2011.05.003>
- Ferrari, A., & Rossari, C. (1994). De donc à dunque e quindi : Les connexions par raisonnement inférentiel. *Cahiers de linguistique française*, 15, 7-49.
- Ferré, G. (2011). Multimodal Analysis of Discourse Markers « donc », « alors » and « en fait » in Conversational French. *ICPhS VXII*, 671-674. <https://hal.archives-ouvertes.fr/hal-00617521>
- Fivela, B. G., & Savino, M. (2003). *Segments, Syllables and Tonal Alignment : A Study on Two Varieties of Italian*.
- Ford, C. E., & Thompson, S. A. (1996). Interactional units in conversation : Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Éds.), *Interaction and Grammar* (p. 134-184). Cambridge University Press. <https://doi.org/10.1017/CB09780511620874.003>
- Fox, B. (2015). On the notion of pre-request. *Discourse Studies*, 17(1), 41-63.
<https://doi.org/10.1177/1461445614557762>
- Fraser, B. (1988). Types of English Discourse Markers. *Acta Linguistica Hungarica*, 38(1/4), 19-33.
- Freese, J., & Maynard, D. W. (1998). Prosodic Features of Bad News and Good News in Conversation. *Language in Society*, 27(2), 195-219.

- Frota, S., & Prieto, P. (2015). Intonation in Romance : Systemic similarities and differences. In S. Frota & P. Prieto (Éds.), *Intonation in Romance* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0011>
- Fry, D. B. (1955). Duration and Intensity as Physical Correlates of Linguistic Stress. *The Journal of the Acoustical Society of America*, 27(4), 765-768. <https://doi.org/10.1121/1.1908022>
- Gao, J., & Hallé, P. (2015). The role of voice quality in Shanghai tone perception. *Proceedings of the 18th International Congress of Phonetic Sciences*, 1-5. <https://www.semanticscholar.org/paper/The-role-of-voice-quality-in-Shanghai-tone-Gao-Hall%C3%A9/4d5e44858ad95b04636698a34422040316342a9c>
- Garellek, M. (2019). The phonetics of voice 1. In W. Katz F. & P. Assmann F. (Éds.), *The Routledge Handbook of Phonetics* (1st éd., p. 32). Routledge. https://www.researchgate.net/publication/311558717_The_phonetics_of_voice
- Geluykens, R., & Swerts, M. (1994). Prosodic cues to discourse boundaries in experimental dialouges. *Speech Communication*, 15(1), 69-77. [https://doi.org/10.1016/0167-6393\(94\)90042-6](https://doi.org/10.1016/0167-6393(94)90042-6)
- German, J. S., & D'Imperio, M. (2015). The Status of the Initial Rise as a Marker of Focus in French. *Language and Speech*, 59(2), 165-195. <https://doi.org/10.1177/0023830915583082>
- Gili Fivela, B. (2002). Tonal alignment in two Pisa Italian peak accents. In B. Bel & I. Marlien (Éds.), *Proceedings of the Speech Prosody 2002 Conference* (p. 339-342).
- Gobl, C., & Chasaide, A. N. (2000). Testing affective correlates of voice quality through analysis and resynthesis. *Proc. ITRW on Speech and Emotion*, 178-183. https://www.isca-speech.org/archive/speechemotion_2000/gobl00_speechemotion.html

- Goffman, E. (1955). On Face-Work. *Psychiatry*, 18(3), 213-231.
<https://doi.org/10.1080/00332747.1955.11023008>
- Goffman, E. (1967). *Interaction Ritual: Essays in Face-To-Face Behavior*. By Erving Goffman. Aldine Publishing Company. <https://doi.org/10.2307/2574751>
- Goffman, E. (1969). *Strategic interaction*. New York : Ballantine Books, Philadelphia : University of Pennsylvania Press.
<http://archive.org/details/strategicinteracgoff00goff>
- Goldsmith, J. (1976). *Autosegmental phonology* [Massachusetts Institute of Technology. Dept. of Foreign Literatures and Linguistics].
<http://www.ai.mit.edu/projects/dm/theses/goldsmith76.pdf>
- Goldstein, T. R., & Bloom, P. (2011). The mind on stage : Why cognitive scientists should study acting. *Trends in Cognitive Sciences*, 15(4), 141-142.
<https://doi.org/10.1016/j.tics.2011.02.003>
- Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. In G. Psathas (Éd.), *Everyday Language : Studies in Ethnomethodology* (New York, p. 97-121).
- Goodwin, C. (1995). The negotiation of coherence within conversation. In M. A. Gernsbacher & T. Givón (Éds.), *Coherence in spontaneous text* (p. 117-137). John Benjamins Publishing.
<https://books.google.fr/books?id=RfOCjJ312ycC&dq=the+negotiation+of+coherence&hl=fr&lr=>
- Goodwin, C. (2002). Time in Action. *Current Anthropology*, 43(4), 19-35.
<https://doi.org/10.1086/339566>
- Goodwin, C., & Heritage, J. (1990). Conversation Analysis. *Annual Review of Anthropology*, 19, 283-307. JSTOR.

- Goodwin, M., Cekaite, A., & Goodwin, C. (2012). Emotion as Stance. In *Emotion in Interaction* (p. 16-40).
<https://doi.org/10.1093/acprof:oso/9780199730735.003.0002>
- Gorrall, B. K., Curtis, J., Little, T. D., & Panko, P. (2016). Innovations in Measurement : Visual Analog Scales and Retrospective Pretest Self-Report Designs. *Actualidades En Psicología*, 30(120), Article 120. <https://doi.org/10.15517/ap.v30i120.22932>
- Goudbeek, M., & Scherer, K. R. (2010). Beyond arousal : Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America*, 128(3), 1322-1336. <https://doi.org/10.1121/1.3466853>
- Grabe, E., Kochanski, G., & Coleman, J. (2005). The intonation of native accent varieties in the British Isles : Potential for miscommunication. *English Pronunciation Models: A Changing Scene*, 21, 311-337.
- Grabe, E., Peter, G., & Johan, P. (2004). Intonational Variation in urban dialects of English spoken in the British Isles. In *Regional variation in intonation* (p. 9-31).
 Linguistische Arbeiten.
- Grabe, M. E., & Kamhawi, R. (2006). Hard Wired for Negative News? Gender Differences in Processing Broadcast News. *Communication Research*, 33(5), 346-369.
<https://doi.org/10.1177/0093650206291479>
- Grice, H. P. (1975). Logic and conversation. *Speech Acts*, 3, 41-58.
- Grice, M. (1995). *The intonation of interrogation in Palermo Italian; implication for intonation theory* (Niemeyer).
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*, 12(3), 175-204.
- Gueneuc, A., Dagher, C., Rameh, G., Haddad, G., Hivernaud, D., Mousty, E., Favre, R., Ville, Y., & Chalouhi, G. E. (2021). Announcing fetal pathology : Challenges encountered by

- physicians and potential role of simulation in training for breaking bad news. *Journal of Gynecology Obstetrics and Human Reproduction*, 50(4), 102044. <https://doi.org/10.1016/j.jogoh.2020.102044>
- Guerrero, L. K., Andersen, P. A., & Trost, M. R. (1996). Communication and emotion : Basic concepts and approaches. In P. A. Andersen & L. K. Guerrero (Éds.), *Handbook of Communication and Emotion* (p. 3-27). Academic Press. <https://doi.org/10.1016/B978-012057770-5/50003-5>
- Gussenhoven, C. (1984). *On the Grammar and Semantics of Sentence Accents* (De Gruyter Mouton). <https://doi.org/10.1515/9783110859263>
- Gussenhoven, C. (2002). *Intonation and interpretation : Phonetics and phonology*. 47-57. <https://repository.uibn.ru.nl/handle/2066/250706>
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press.
- Gwet, K. (2001). *Handbook of inter-rater reliability : How to estimate the level of agreement between two or multiple raters*. STATAXIS Publishing Company.
- Gwet, K. L. (2019). *irrCAC : Computing Chance-Corrected Agreement Coefficients (CAC) (1.0)* [Logiciel]. <https://cran.r-project.org/web/packages/irrCAC/index.html>
- Haileselassie, A. (2015). *Voilà, an orientation shift marker in modern French discourse : A conversation analytic perspective* [University of Illinois at Urbana-Champaign]. <https://www.proquest.com/openview/8cbbae7fab4849072d849c1d535e719c/1?pq-origsite=gscholar&cbl=18750>
- Hakulinen, A., & Couper-Kuhlen, E. (2015). Insisting on 'my side' : Siis-prefaced utterances in Finnish. *Journal of Pragmatics*, 75, 111-130. <https://doi.org/10.1016/j.pragma.2014.10.009>

- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10(1), 41-53.
- Hansen, M.-B. M. (1997). Alors and donc in spoken French : A reanalysis. *Journal of Pragmatics*, 28(2), 153-187. [https://doi.org/10.1016/S0378-2166\(96\)00086-0](https://doi.org/10.1016/S0378-2166(96)00086-0)
- Harré, R. (1984). *Social elements as mind*. 57(2), 127-135. <https://doi.org/10.1111/j.2044-8341.1984.tb01591.x>
- Harré, R. (1986). Mind as a Social Formation. In J. Margolis, M. Krausz, & R. M. Burian (Éds.), *Rationality, Relativism and the Human Sciences* (p. 91-106). Springer Netherlands. https://doi.org/10.1007/978-94-009-4362-9_6
- Havrdová, Z., & Morávek, M. (1979). Changes of the voice expression during suggestively influenced states of experiencing. *Activitas Nervosa Superior*, 21(1), 33-35.
- Heldner, M., Włodarczak, M., Beňuš, Š., & Gravano, A. (2019). Voice quality as a turn-taking cue. *Interspeech 2019, Graz, Austria, 15-19 September 2019*, 4165-4169. <https://www.diva-portal.org/smash/record.jsf?pid=diva2:1354344>
- Heritage, J. (1989). Current developments in conversation analysis. In *Conversation : An interdisciplinary perspective* (p. 21-47). D. Roger & P. Bull.
- Heritage, J. (2015). Well-prefaced turns in English conversation : A conversation analytic perspective. *Journal of Pragmatics*, 88, 88-104. <https://doi.org/10.1016/j.pragma.2015.08.008>
- Heritage, J., & Sorjonen, M.-L. (1994). Constituting and Maintaining Activities across Sequences : And-Prefacing as a Feature of Question Design. *Language in Society*, 23(1), 1-29. <https://doi.org/10.1017/S0047404500017656>
- Herman, R., & McGory, J. T. (2002). The Conceptual Similarity of Intonational Tones and its Effects on Intertranscriber Reliability. *Language and Speech*, 45(1), 1-36. <https://doi.org/10.1177/00238309020450010101>

- Hicks, J. W. (1979). An acoustical/temporal analysis of the effect of situational stress on speech. In H. Hollien & P. Hollien (Éds.), *Current Issues in the Phonetic Sciences : Proceedings of the IPS-77 Congress* (p. 279-284). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.9.32hic>
- Hoerger, M. (2010). Participant dropout as a function of survey length in internet-mediated university studies: Implications for study design and voluntary participation in psychological research. *Cyberpsychology, Behavior and Social Networking*, 13(6), 697-700. <https://doi.org/10.1089/cyber.2009.0445>
- Hommel, B. (1997). Interactions between stimulus-stimulus congruence and stimulus-response compatibility. *Psychological Research*, 59(4), 248-260. <https://doi.org/10.1007/BF00439302>
- Hussain, Q. (2021). Phonetic correlates of laryngeal and place contrasts of Burushaski. *Speech Communication*, 126, 71-89. <https://doi.org/10.1016/j.specom.2020.11.001>
- Icht, M., Wiznitsker Rassis-tal, H., & Lotan, M. (2021). Can the Vocal Expression of Intellectually Disabled Individuals Be Used as a Pain Indicator? Initial Findings Supporting a Possible Novice Assessment Method. *Frontiers in Psychology*, 12. <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.655202>
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. MIT Press.
- Jacoby, S., & Ochs, E. (1995). Co-Construction : An Introduction. *Research on Language and Social Interaction*, 28(3), 171-183. https://doi.org/10.1207/s15327973rlsi2803_1
- Jeong, S., & Potts, C. (2016). Intonational sentence-type conventions for perlocutionary effects: An experimental investigation. *Semantics and Linguistic Theory*, 1-22. <https://doi.org/10.3765/salt.v26i0.3787>

- Jiang, X., & Pell, M. D. (2017). The sound of confidence and doubt. *Speech Communication*, 88, 106-126. <https://doi.org/10.1016/j.specom.2017.01.011>
- Johnstone, T., & Scherer, K. R. (2000). Vocal communication of Emotion. *Handbook of Emotions*, 2, 220-235.
- Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence*, 29(4), 589-611. <https://doi.org/10.1016/j.adolescence.2005.08.010>
- Jucker, A. H. (1993). The discourse marker well : A relevance-theoretical account. *Journal of Pragmatics*, 19(5), 435-452. [https://doi.org/10.1016/0378-2166\(93\)90004-9](https://doi.org/10.1016/0378-2166(93)90004-9)
- Jun, S.-A., & Fougeron, C. (2000). A Phonological Model of French Intonation. In A. Botinis (Éd.), *Intonation : Analysis, Modelling and Technology* (p. 209-242). Springer Netherlands. https://doi.org/10.1007/978-94-011-4317-2_10
- Jun, S.-A., & Fougeron, C. (2002). The Realizations of the Accentual Phrase in French Intonation. *Probus*, 14(1), 147-172. <https://doi.org/10.1515/prbs.2002.002>
- Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion (Washington, D.C.)*, 1(4), 381-412. <https://doi.org/10.1037/1528-3542.1.4.381>
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance : Different channels, same code? *Psychological Bulletin*, 129(5), 770-814. <https://doi.org/10.1037/0033-2909.129.5.770>
- Juslin, P. N., Laukka, P., & Bänziger, T. (2018). The Mirror to Our Soul? Comparisons of Spontaneous and Posed Vocal Expression of Emotion. *Journal of Nonverbal Behavior*, 42(1), 1-40. <https://doi.org/10.1007/s10919-017-0268-x>
- Juslin, P. N., & Scherer, K. R. (2005). Vocal expression of affect. In *The new handbook of methods in nonverbal behavior research* (p. 65-135). Oxford University Press.

- Kaiser, J., & Božil, T. (2018). *Impact of the GSM AMR Codec on Automatic Vowel Formant Measurement in Praat and VoiceSauce*. 1-4.
<https://doi.org/10.1109/TSP.2018.8441185>
- Kalaldehy, R., Dorn, A., & Ní Chasaide, A. (2009). Tonal alignment in three varieties of hiberno-English. *Interspeech 2009*, 2443-2446.
<https://doi.org/10.21437/Interspeech.2009-310>
- Kaland, C., Swerts, M., & Himmelmann, N. P. (2023). Red and blue bananas : Time-series f0 analysis of contrastively focused noun phrases in Papuan Malay and Dutch. *Journal of Phonetics*, 96, 101200. <https://doi.org/10.1016/j.wocn.2022.101200>
- Kaminskaia, S. (2015). Tonal patterns, associations, and alignment of peaks in regional French. *WORD*, 61(2), 101-140.
<https://doi.org/10.1080/00437956.2015.1033174>
- Karlsson, I., Bänziger, T., Dankovicová, J., Johnstone, T., Lindberg, J., Melin, H., Nolan, F., & Scherer, K. R. (2000). Speaker verification with elicited speaking styles in the VeriVox project. *Speech Communication*, 31, 121-129.
[https://doi.org/10.1016/S0167-6393\(99\)00073-4](https://doi.org/10.1016/S0167-6393(99)00073-4)
- Karunaratne, I., Atukorale, A., & Perera, H. (2011). Surveillance of human-computer interactions : A way forward to detection of users' Psychological Distress. *2011 IEEE Colloquium on Humanities, Science and Engineering, CHUSER 2011*.
<https://doi.org/10.1109/CHUSER.2011.6163779>
- Keaton, S. A., Worthington, D. L., & Bodie, D. (2017). Interpersonal Reactivity Index (IRI). In *The Sourcebook of Listening Research : Methodology and Measures* (p. 340-347). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781119102991.ch34>
- Keevallik, L. (2012). Compromising progressivity : 'No'-prefacing in estonian. *Pragmatics*, 22(1), 119-146. <https://doi.org/10.1075/prag.22.1.05kee>

- Keevallik, L. (2016). Estonian no(o)(h) in turns and sequences : Families of function: A Family of Discourse Markers Across the Languages of Europe and Beyond. In P. Auer & Y. Maschler, *NU / NÁ: A Family of Discourse Markers Across the Languages of Europe and Beyond*. De Gruyter. <https://doi.org/10.1515/9783110348989-007>
- Kemp, R. (2012). *Embodied Acting : What Neuroscience Tells Us About Performance*. Routledge.
- Kirby, J. (2018). *PraatSauce : Praat-based tools for spectral analysis (0.2.4)* [Logiciel]. <https://github.com/kirbyj/praatsauce>
- Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, 3(3), 129-140. [https://doi.org/10.1016/S0095-4470\(19\)31360-9](https://doi.org/10.1016/S0095-4470(19)31360-9)
- Kohler, K. J. (1987). Categorical pitch perception. In U. Viks (Éd.), *Proceedings of the 11th International Congress of Phonetic Sciences* (Vol. 5, p. 331-333).
- Konijn, E. (2000). *Acting Emotions : Shaping Emotions on Stage*. Amsterdam University Press. <https://doi.org/10.5117/9789053564448>
- Krahmer, E., & Swerts, M. (2009). Audiovisual Prosody—Introduction to the Special Issue. *Language and Speech*, 52(2-3), 129-133. <https://doi.org/10.1177/0023830909103164>
- Kramer, M. W., & Hess, J. A. (2002). Communication Rules for the Display of Emotions in Organizational Settings. *Management Communication Quarterly*, 16(1), 66-80. <https://doi.org/10.1177/0893318902161003>
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality : Review, tutorial, and a framework for future research. *Journal of Speech and Hearing Research*, 36(1), 21-40. <https://doi.org/10.1044/jshr.3601.21>

- Krom, G. (1995). Some spectral correlates of pathological breathy and rough voice quality for different types of vowel fragments. *Journal of Speech and Hearing Research*, 38(4), 794-811. <https://doi.org/10.1044/jshr.3804.794>
- Kuo, G. (2013). Perceived prosodic boundaries in Taiwanese and Swedish. *The Journal of the Acoustical Society of America*, 133, 3573. <https://doi.org/10.1121/1.4806551>
- Kyaw, W. T., & Sagisaka, Y. (2017). *Cross-modal Sentiment Information Expression of Voice Source Characteristics using Image Texture Features*. <https://onlineresource.ucsy.edu.mm/handle/123456789/896>
- Ladd, D. R. (2008). *Intonational Phonology*. Cambridge University Press.
- Ladd, D. R., & Arvaniti, A. (2023). Prosodic Prominence Across Languages. *Annual Review of Linguistics*, 9(1), 171-193. <https://doi.org/10.1146/annurev-linguistics-031120-101954>
- Ladd, D. R., Silverman, K., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *Journal of The Acoustical Society of America - JACOUST SOC AMER*, 78(2), 435-444. <https://doi.org/10.1121/1.392466>
- Ladd, R. (1993). In Defense of a Metrical Theory of Intonational Downstep. In K. Snider & H. van der Hulst (Éds.), *The Phonology of Tone: The Representation of Tonal Register* (p. 109-132). De Gruyter Mouton. <https://doi.org/10.1515/9783110869378.109>
- Laerhoven, H. van, Zaag-Loonen, H. van der, & Derkx, B. H. F. (2004). A comparison of Likert scale and visual analogue scales as response options in children's questionnaires. *Acta Paediatrica*, 93(6), 830-835. <https://doi.org/10.1111/j.1651-2227.2004.tb03026.x>

- Lancia, L., Krasovitsky, G., & Stuntebeck, F. (2019). Coordinative patterns underlying cross-linguistic rhythmic differences. *Journal of Phonetics*, 72, 66-80. <https://doi.org/10.1016/j.wocn.2018.08.004>
- Lange, C. G. (1922). The emotions. In I. A. Haupt (Éd.), *The emotions, Vol. 1.* (p. 33-90). Williams & Wilkins Co. <https://doi.org/10.1037/10735-002>
- Laukka, P. (2005). Categorical perception of vocal emotion expressions. *Emotion*, 5(3), 277-295. <https://doi.org/10.1037/1528-3542.5.3.277>
- Laukka, P., Audibert, N., & Aubergé, V. (2012). Exploring the determinants of the graded structure of vocal emotion expressions. *Cognition and Emotion*, 26(4), 710-719. <https://doi.org/10.1080/02699931.2011.602047>
- Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition and Emotion*, 19(5), 633-653. <https://doi.org/10.1080/02699930441000445>
- Laukka, P., Linnman, C., Åhs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgård, Å., Appel, L., Fredrikson, M., & Furmark, T. (2008). In a nervous voice : Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior*, 32(4), 195-214. <https://doi.org/10.1007/s10919-008-0055-9>
- Laukka, P., Neiberg, D., Forsell, M., Karlsson, I., & Elenius, K. (2011). Expression of affect in spontaneous speech : Acoustic correlates and automatic detection of irritation and resignation. *Computer Speech & Language*, 25(1), 84-104. <https://doi.org/10.1016/j.csl.2010.03.004>
- Laukkanen, A.-M., Vilkmán, E., Alku, P., & Oksanen, H. (1996). On the perception of emotions in speech : The role of voice quality. *Logopedics Phoniatrics Vocology*, 22(4), 157-168. <https://doi.org/10.3109/14015439709075330>

- Lausen, A., & Schacht, A. (2018). Gender Differences in the Recognition of Vocal Emotions. *Frontiers in Psychology*, 9(882), 22. <https://doi.org/10.3389/fpsyg.2018.00882>
- Laver, J. (1980). The phonetic description of voice quality | Semantic Scholar. *Cambridge Studies in Linguistics London*, 31, 186-201.
- Laver, J. (1999). Towards an Integrated Theory of Nonverbal Communication. *Proceedings of the 14th International Congress of Phonetic Sciences-ICPhS99*, 2433-2436. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_2433.html
- Laver, J. (1994). Principles of phonetics. *Journal of Linguistics | Cambridge Core*, xxviii+502.
- Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy : Reliability and validity of the Empathy Quotient. *Psychological Medicine*, 34(5), 911-919. <https://doi.org/10.1017/S0033291703001624>
- Leeper, C., & Ayres, M. M. (2007). A Meta-Analytic Review of Gender Variations in Adults' Language Use : Talkativeness, Affiliative Speech, and Assertive Speech. *Personality and Social Psychology Review*, 11(4), 328-363. <https://doi.org/10.1177/1088868307302221>
- Lehiste, I. (1973). *Phonetic disambiguation of syntactic ambiguity*. 53(1_Supplement), 380-380.
- Lennes, M. (2017). *SpeCT - The Speech Corpus Toolkit for Praat (v1.0.0) : First release on GitHub* [Logiciel]. <https://doi.org/10.5281/zenodo.375923>
- Lerner, G. H. (1993). Collectivities in action : Establishing the relevance of conjoined participation in conversation. *Text - Interdisciplinary Journal for the Study of Discourse*, 13(2), 213-246. <https://doi.org/10.1515/text.1.1993.13.2.213>

- Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511813313>
- Levitan, R., Gravano, A., Willson, L., Beňuš, Š., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies*, 11-19. <https://aclanthology.org/N12-1002.pdf>
- Lieberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic inquiry*, 8(2), 249-336.
- Lieberman, M. Y. (1975). *The intonational system of English*. [Thesis, Massachusetts Institute of Technology]. <https://dspace.mit.edu/handle/1721.1/27376>
- Lieberman, M. Y., & Pierrehumbert, J. B. (2005). A metric for the height of certain pitch peaks in English. *The Journal of the Acoustical Society of America*, 66(S1), S64. <https://doi.org/10.1121/1.2017882>
- Lindström, A., Hakulinen, A., & Selting, M. (2005). Language as social action. In *Syntax and lexis in conversation : Studies on the use of linguistic resources in talk-in interaction* (p. 209-230). John Benjamins Publishing Company. <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-280308>
- Liscombe, J., Venditti, J., & Hirschberg, J. (2003). Classifying Subject Ratings of Emotional Speech Using Acoustic Features. *EUROSPEECH 2003 - INTERSPEECH 2003*, 4.
- Maguire, P. (1998). Breaking bad news. *European Journal of Surgical Oncology*, 24(3), 188-191. [https://doi.org/10.1016/S0748-7983\(98\)92929-8](https://doi.org/10.1016/S0748-7983(98)92929-8)
- Makri-Tsilipakou, M. (1994). Interruption revisited: Affiliative vs. disaffiliative intervention. *Journal of Pragmatics*, 21(4), 401-426. [https://doi.org/10.1016/0378-2166\(94\)90012-4](https://doi.org/10.1016/0378-2166(94)90012-4)

- Markel, N. N., Bein, M. F., & Phillis, J. A. (1973). The relationship between words and tone-of-voice. *Language and Speech*, 16(1), 15-21.
<https://doi.org/10.1177/002383097301600102>
- Martínez Garcia, M. N. (2020). *Prosodically Conditioned Realization of Voiced Stops and Vowels in Yucatecan Spanish* [Text.thesis.doctoral, Universität zu Köln].
<http://www.uni-koeln.de/>
- Maschler, Y., & Schiffrin, D. (2015). Discourse Markers Language, Meaning, and Context. In *The Handbook of Discourse Analysis* (p. 189-221). John Wiley & Sons, Ltd.
<https://doi.org/10.1002/9781118584194.ch9>
- Massaro, D. W. (1987). *Speech Perception By Ear and Eye : A Paradigm for Psychological Inquiry*. Psychology Press. <https://doi.org/10.4324/9781315808253>
- Mastriani, C., Petrone, C., Bertrand, R., & Ochs, M. (2018). *Rôle de la prosodie dans la perception du feedback 'oui' dans un contexte médical*. 7.
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and Prosodic Effects on Word Segmentation in Infants. *Cognitive Psychology*, 38(4), 465-494.
<https://doi.org/10.1006/cogp.1999.0721>
- Mauchand, M., Vergis, N., & Pell, M. D. (2020). Irony, Prosody, and Social Impressions of Affective Stance. *Discourse Processes*, 57(2), 141-157.
<https://doi.org/10.1080/0163853X.2019.1581588>
- Maynard, D. (2011). Defensive mechanisms : « I mean » prefaced utterances in complaint and other conversational sequences. *Conversational Repair and Human Understanding*, 30, 315-374. <https://doi.org/10.1017/cbo9780511757464.007>
- Maynard, D. W. (1996). On « Realization » in Everyday Life : The Forecasting of Bad News as a Social Relation. *American Sociological Review*, 61(1), 109-131.
<https://doi.org/10.2307/2096409>

- Maynard, D. W. (2017). Delivering bad news in emergency care medicine. *Acute Medicine & Surgery*, 4(1), 3-11. <https://doi.org/10.1002/ams2.210>
- Mazzocconi, C., Petrone, C., & Champagne-Lavau, M. (2022). *Does laughter play a role in irony comprehension?* <https://hal.science/hal-03878579>
- McClenahan, L., & Lofland, J. (1976). Bearing Bad News: Tactics of the Deputy U.S. Marshal. *Sociology of Work and Occupations*, 3(3), 251-272. <https://doi.org/10.1177/003803857600300301>
- Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology* (p. xii, 266). The MIT Press.
- Michelas, A., & D'Imperio, M. (2010). *Durational Cues and Prosodic Phrasing in French : Evidence for the Intermediate Phrase*. 4. <https://hal.science/hal-00463205>
- Michelas, A., & D'Imperio, M. (2009, juin 17). *Is there an Intermediate Phrase in French?* Phonetics and Phonologie in Iberia, Las Palmas de Gran Canaria, France. <https://hal.science/hal-00433565>
- Mol, H., & Uhlenbeck, E. M. (1955). The linguistic relevance of intensity in stress. *Lingua*, 5, 205-213. [https://doi.org/10.1016/0024-3841\(55\)90010-3](https://doi.org/10.1016/0024-3841(55)90010-3)
- Mondada, L. (1999). L'organisation séquentielle des ressources linguistiques dans l'élaboration collective des descriptions. *Langage & société*, 89(1), 9-36. <https://doi.org/10.3406/lsoc.1999.2882>
- Mozziconacci, S. (2002). *Prosody and emotions*.
- Mühlhoff, R. (2019). Affective disposition. In J. Slaby & C. Von Scheve (Éds.), *Affective Societies : Key concepts* (Routledge Studies in Affective Societies, p. 119-130). <https://doi.org/10.4324/9781351039260-10>

- Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech :
A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, 93(2), 1097-1108. <https://doi.org/10.1121/1.405558>
- Nadeu, M., & Prieto, P. (2011). Pitch range, gestural information, and perceived politeness in Catalan. *Journal of Pragmatics*, 43(3), 841-854. <https://doi.org/10.1016/j.pragma.2010.09.015>
- Nakatani, C., Hirschberg, J., & Grosz, B. (1995). Discourse structure in spoken language :
Studies on speech corpora. *AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation (1995)*. <https://dash.harvard.edu/handle/1/2580299>
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Foris.
- Nespor, M., & Vogel, I. (2007). *Prosodic Phonology: With a New Foreword*. Walter de Gruyter.
- Niebuhr, O. (2007). The Signalling of German Rising-Falling Intonation Categories – The Interplay of Synchronization, Shape, and Height. *Phonetica*, 64(2-3), 174-193. <https://doi.org/10.1159/0000107915>
- Niebuhr, O., & Kohler, K. (2004, juin 3). *Perception and Cognitive Processing of Tonal Alignment in German*. International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages.
- Ochs, E., Schegloff, E. A., & Thompson, S. A. (1996). *Interaction and Grammar*. Cambridge University Press.
- Ochs, E., & Schieffelin, B. (1989). Language has a heart. *Text - Interdisciplinary Journal for the Study of Discourse*, 9(1), 7-26. <https://doi.org/10.1515/text.1.1989.9.1.7>
- Ochs, M., Blache, P., de Montcheuil, G., Pergandi, J.-M., Saubesty, J., Francon, D., & Mestre, D. (2018, mai). A Semi-autonomous System for Creating a Human-Machine

- Interaction Corpus in Virtual Reality : Application to the ACORFORMed System for Training Doctors to Break Bad News. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. LREC 2018, Miyazaki, Japan. <https://aclanthology.org/L18-1467>
- Ohala, J. J. (1983). Cross-Language Use of Pitch : An Ethological View. *Phonetica*, 40(1), 1-18. <https://doi.org/10.1159/000261678>
- Olderbak, S., & Wilhelm, O. (2017). Emotion perception and empathy : An individual differences test of relations. *Emotion*, 17(7), 1092-1106. <https://doi.org/10.1037/emo0000308>
- Owren, M. J., & Bachorowski, J.-A. (2007). Measuring emotion-related vocal acoustics. In *Handbook of emotion elicitation and assessment* (p. 239-266). Oxford University Press.
- Paeschke, A., Kienast, M., & Sendlmeier, W. F. (1999). F0-Contours in Emotional Speech. *Proc. 14th Int. Congress of Phonetic Sciences*, 2, 929-932. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_0929.html
- Papousek, I., Schuler, G., Lackner, H. K., Samson, A., & Freudenthaler, H. H. (2014). Experimentally observed responses to humor are related to individual differences in emotion perception and regulation in everyday life. *HUMOR*, 27(2), 271-286. <https://doi.org/10.1515/humor-2014-0018>
- Pell, M. D., Sethi, S., Rigoulot, S., Rothermich, K., Liu, P., & Jiang, X. (2022). Emotional voices modulate perception and predictions about an upcoming face. *Cortex*, 149, 148-164. <https://doi.org/10.1016/j.cortex.2021.12.017>

- Pellet, S. H. (2005). *The Development of Competence in French Interlanguage Pragmatics : The Case of the Discourse Marker 'donc'* [Philosophy]. Faculty of the Graduate School of The University of Texas.
- Pernet, C., & Belin, P. (2012). The Role of Pitch and Timbre in Voice Gender Categorization. *Frontiers in Psychology*, 3. <https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00023>
- Petit, M. (2020). La prosodie de voilà en français dans le discours médiatique. In *Polysémie, usages et fonctions de « voilà »* (p. 153-178). De Gruyter. <https://www.degruyter.com/document/doi/10.1515/9783110622454-005/html>
- Petrone, C. (2008). Le rôle de la variabilité phonétique dans la représentation des contours intonatifs et de leur sens. *Travaux interdisciplinaires sur la parole et le langage*, 27, Article 27. <https://doi.org/10.4000/tipa.410>
- Petrone, C., D'Alessandro, D., & Falk, S. (2021). Working memory differences in prosodic imitation. *Journal of Phonetics*, 89, 101100. <https://doi.org/10.1016/j.wocn.2021.101100>
- Petrone, C., German, E. S., German, J. S., & Ito, K. (2019, août 5). *Effects of Tune and Interspeaker Differences on the Interpretation of Requests and Offers*. International Congress of Phonetic Sciences ICPhS 2019, Melbourne, Australia. <https://hal.science/hal-02110414>
- Petrone, C., & Niebuhr, O. (2014). On the Intonation of German Intonation Questions : The Role of the Prenuclear Region. *Language and Speech*, 57(1), 108-146. <https://doi.org/10.1177/0023830913495651>
- Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of Phonetics*, 18(3), 375-394. [https://doi.org/10.1016/S0095-4470\(19\)30380-8](https://doi.org/10.1016/S0095-4470(19)30380-8)

- Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* [Thesis, Massachusetts Institute of Technology].
<https://dspace.mit.edu/handle/1721.1/16065>
- Pierrehumbert, J. B., & Steele, S. A. (2009). Categories of Tonal Alignment in English. *Phonetica*, 46(4), 181-196. <https://doi.org/10.1159/000261842>
- Pierrehumbert, J., & Beckman, M. (1988). Japanese Tone Structure. In *The Journal of the Association of Teachers of Japanese*. MIT Press. <https://doi.org/10.2307/489057>
- Pinheiro, J. C., & Bates, D. M. (Éds.). (2000). Linear Mixed-Effects Models : Basic Concepts and Examples. In *Mixed-Effects Models in S and S-PLUS* (p. 3-56). Springer.
https://doi.org/10.1007/0-387-22747-4_1
- Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, 15(2), 285-290.
<https://doi.org/10.3758/BF03213946>
- Planalp, S. (1993). Communication, cognition, and emotion. *Communication Monographs*, 60(1), 3-9. <https://doi.org/10.1080/03637759309376288>
- Porhet, C., Ochs, M., Saubesty, J., de Montcheuil, G., & Bertrand, R. (2017). Mining a multimodal corpus of doctor's training for virtual patient's feedbacks. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 473-478.
<https://doi.org/10.1145/3136755.3136816>
- Porretta, V., Tucker, B. V., & Järvikivi, J. (2016). The influence of gradient foreign accentedness and listener experience on word recognition. *Journal of Phonetics*, 58, 1-21. <https://doi.org/10.1016/j.wocn.2016.05.006>
- Portes, C. (2004). *Prosodie et économie du discours: Spécificité phonétique, écologie discursive et portée pragmatique de l'intonation d'implication* [Phdthesis, Université de Provence - Aix-Marseille I]. <https://theses.hal.science/tel-00296742>

- Post, B. (2000). *Tonal and Phrasal structures in French Intonation*. Université de Nijmegen.
- Pourquery, D. (2014). *Les mots de l'époque : 100 tics, trouvailles et autres extravagances du langage quotidien*. Autrement.
- Prévot, L., & Bertrand, R. (2012, janvier 1). CoFee-Toward a multidimensional analysis of conversational feedback, the case of French language. *Proceedings of the Workshop on Feedback Behaviors*. Workshop on Feedback Behaviors Authors:, Portland, USA. <https://hal.archives-ouvertes.fr/hal-01231913>
- Prévot, L., Bigi, B., & Bertrand, R. (2013). A quantitative view of feedback lexical markers in conversational French. *Proceedings of the SIGDIAL 2013 Conference*, 87-91. <https://aclanthology.org/W13-4011>
- Price, D. D., Staud, R., & Robinson, M. E. (2012). How should we use the visual analogue scale (VAS) in rehabilitation outcomes? II: Visual analogue scales as ratio scales: An alternative to the view of Kersten et al. *Journal of Rehabilitation Medicine*, 44(9), Article 9. <https://doi.org/10.2340/16501977-1031>
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, 90(6), 2956-2970. <https://doi.org/10.1121/1.401770>
- Priego-Valverde, B., Nguyen, N., & Bertrand, R. (2022). Alignement, affiliation et trajectoire interactionnelle dans la conversation. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, 38, Article 38. <https://doi.org/10.4000/tipa.4819>
- Prieto, P. (1998). The scaling of the L values in Spanish downstepping contours. *Journal of Phonetics*, 26(3), 261-282. <https://doi.org/10.1006/jpho.1998.0074>
- Prieto, P. (2012). Experimental methods and paradigms for prosodic analysis. In C. Fougeron & M. Huffman (Éds.), *The Oxford Handbook of Laboratory Phonology* (Oxford University Press, p. 528-537).

- Prieto, P. (2015). Intonational meaning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(4), 371-381. <https://doi.org/10.1002/wcs.1352>
- Prieto, P. (2003). *The Search for Phonological Targets in the Tonal Space : H1 scaling and alignment in five sentence-types in Peninsular Spanish*. <https://www.semanticscholar.org/paper/The-Search-for-Phonological-Targets-in-the-Tonal-H1-Prieto/aa0df9b5d0cb21946267cf7763986a744867c42c>
- Prieto, P., D'Imperio, M., & Fivela, B. G. (2005). Pitch Accent Alignment in Romance : Primary and Secondary Associations with Metrical Structure. *Language and Speech*, 48(4), 359-396. <https://doi.org/10.1177/00238309050480040301>
- Prieto, P., & Frota, S. (2015). Introduction. In S. Frota & P. Prieto (Éds.), *Intonation in Romance* (p. 1-8). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0001>
- Prieto, P., Van Santen, J., & Hirschberg, J. (1995). Tonal alignment patterns in Spanish. *Journal of Phonetics*, 23(4), 429-451. <https://doi.org/10.1006/jpho.1995.0032>
- Psychology Software Tools. (2020). *E-Prime* (3.0) [Logiciel]. <https://support.pstnet.com/>
- Reips, U.-D., & Funke, F. (2008). Interval-level measurement with visual analogue scales in Internet-based research : VAS Generator. *Behavior Research Methods*, 40(3), 699-704. <https://doi.org/10.3758/BRM.40.3.699>
- Rhee, N., Chen, A., & Kuang, J. (2019). Integration of spectral cues in the development of Mandarin tone production. *Proceedings of the 19th International Congress of Phonetic Sciences*, 3135-3138. <https://dspace.library.uu.nl/handle/1874/391812>
- Rilliard, A., Takaaki, S., Matin, J.-C., Erickson, D., & Aubergé, V. (2009). Multimodal Indices to Japanese and French Prosodically Expressed Social Affects. *Language and Speech*, 52(2-3), 223-243. <https://doi.org/10.1177/0023830909103171>

- Rivière, E., Klein, M., & Champagne-Lavau, M. (2018). Using context and prosody in irony understanding: Variability amongst individuals. *Journal of Pragmatics*, *138*, 165-172. <https://doi.org/10.1016/j.pragma.2018.10.006>
- Rodero, E. (2011). Intonation and Emotion : Influence of Pitch Levels and Contour Type on Creating Emotions. *Journal of Voice*, *25*(1), 25-34. <https://doi.org/10.1016/j.jvoice.2010.02.002>
- Roessler, R., & Lester, J. W. (1979). Vocal pattern in anxiety. In W. E. Fann, A. D. Pokorny, I. Koracau, & R. L. Williams (Éds.), *Phenomenology and Treatment of Anxiety* (Spectrum).
- Rohrer, P. L., Florit-Pons, J., Vilà-Giménez, I., & Prieto, P. (2022). Children Use Non-referential Gestures in Narrative Speech to Mark Discourse Elements Which Update Common Ground. *Frontiers in Psychology*, *12*. <https://doi.org/10.3389/fpsyg.2021.661339>
- Rooth, M. (1985). *Association with focus (montague grammar, semantics, only, even)* [University of Massachusetts, Amherst]. <https://hdl.handle.net/1813/28568>
- Rooth, M. (1992). A Theory of Focus Interpretation. *Natural Language Semantics*, *1*(1), 75-116. <https://doi.org/10.1007/BF02342617>
- Rossi, M. (2001). 6. L'intonation. *Modèles linguistiques*, *XXII-1*(43), Article 43. <https://doi.org/10.4000/ml.1463>
- Roy, J., Cole, J., & Mahrt, T. (2017). Individual differences and patterns of convergence in prosody perception. *Laboratory Phonology*, *8*(1), 22. <https://doi.org/10.5334/labphon.108>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161-1178. <https://doi.org/10.1037/h0077714>

- Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145-172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Sacks, H. (1973). The preference for agreement in natural conversation. *Linguistic Institute, Ann Arbor, Michigan*.
- Sacks, H. (1989). An Analysis of the Course of a Joke's Telling in Conversation. In R. Bauman & J. Sherzer (Éds.), *Explorations in the Ethnography of Speaking* (2^e éd., p. 337-353). Cambridge University Press. <https://doi.org/10.1017/CBO9780511611810.022>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50(4), 696-735. JSTOR. <https://doi.org/10.2307/412243>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A Simplest Systematics for the Organization of Turn Taking for Conversation. In J. Schenkein (Éd.), *Studies in the Organization of Conversational Interaction* (p. 7-55). Academic Press. <https://doi.org/10.1016/B978-0-12-623550-0.50008-2>
- Schechner, R., Bala, S., Kochhar-Lindgren, K., Chibici-Revneanu, C., Harle, R., Tsoumas, J., Lawrence, D. W., Ghosh, S., Acharya, I., Purkayastha, S. D., Grossberg, R., Ghatak, P., Shetty, Y., McAllister, B., Basu, S. P., Sesan, A. A., Bhattacharya, B., Mukherjee, A., Konar, R., ... Ghosh, S. (2013). *Performance Studies: Rupkatha Journal on Interdisciplinary Studies in Humanities, Volume V, Number 2, 2013*. Aesthetics Media Services.
- Schegloff, E. (1996). Turn organization : One intersection of grammar and interaction. In E. Ochs & E. A. Schegloff (Éds.), *Interaction and grammar* (p. 52-133). Cambridge University Press. <https://doi.org/10.1017/CBO9780511620874.002>

- Schegloff, E. (2006). On possible. *Discourse Studies*, 8(1), 141-157.
<https://doi.org/10.1177/1461445606059563>
- Schegloff, E. A. (1972). Notes on a conversational practice : Formulating place. In D. Sudnow (Éd.), *Studies in Social Interaction* (p. 75-119). Free Press.
- Schegloff, E. A. (1979). The Relevance of Repair to Syntax-for-Conversation. In *Discourse and Syntax* (p. 261-286). Brill. https://doi.org/10.1163/9789004368897_012
- Schegloff, E. A. (1980). Preliminaries to Preliminaries : "Can I Ask You a Question?" *Sociological Inquiry*, 50(3-4), 104-152. <https://doi.org/10.1111/j.1475-682X.1980.tb00018.x>
- Schegloff, E. A. (2007). *Sequence Organization in Interaction : A Primer in Conversation Analysis I*. Cambridge University Press.
- Schegloff, E. A., & Sacks, H. (1973). *Opening up Closings*. 8(4), 289-327.
<https://doi.org/10.1515/semi.1973.8.4.289>
- Scherer, K. R. (1979). *Nonlinguistic Vocal Indicators of Emotion and Psychopathology* (C. E. Izard, Éd.; p. 493-529). Springer US. https://doi.org/10.1007/978-1-4613-2892-6_18
- Scherer, K. R. (1985). Vocal Affect Signaling : A Comparative Approach. In J. S. Rosenblatt, C. Beer, M.-C. Busnel, & P. J. B. Slater (Éds.), *Advances in the Study of Behavior* (Vol. 15, p. 189-244). Academic Press. [https://doi.org/10.1016/S0065-3454\(08\)60490-8](https://doi.org/10.1016/S0065-3454(08)60490-8)
- Scherer, K. R. (1986). Vocal affect expression : A review and a model for future research. *Psychological Bulletin*, 99(2), 143-165. <https://doi.org/10.1037/0033-2909.99.2.143>
- Scherer, K. R. (1989). Vocal correlates of emotional arousal and affective disturbance. In *Handbook of social psychophysiology* (p. 165-197). John Wiley & Sons.

- Scherer, K. R. (2003). Vocal communication of emotion : A review of research paradigms. *Speech Communication, 40*(1-2), 227-256. [https://doi.org/10.1016/S0167-6393\(02\)00084-5](https://doi.org/10.1016/S0167-6393(02)00084-5)
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511611841>
- Schirmer, A., Kotz, S. A., & Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*(2), 228-233. [https://doi.org/10.1016/S0926-6410\(02\)00108-8](https://doi.org/10.1016/S0926-6410(02)00108-8)
- Schirmer, A., Zysset, S., Kotz, S. A., & Yves von Cramon, D. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *NeuroImage, 21*(3), 1114-1123. <https://doi.org/10.1016/j.neuroimage.2003.10.048>
- Schourup, L. (1999). Discourse markers. *Lingua, 107*(3), 227-265. [https://doi.org/10.1016/S0024-3841\(96\)90026-1](https://doi.org/10.1016/S0024-3841(96)90026-1)
- Schröder, M. (2004). Dimensional Emotion Representation as a Basis for Speech Synthesis with Non-extreme Emotions. In E. André, L. Dybkjær, W. Minker, & P. Heisterkamp (Éds.), *Affective Dialogue Systems* (Vol. 3068, p. 209-220). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-24842-2_21
- Schröder, M., Cowie, R., Douglas-Cowie, E., Westerdijk, M., & Gielen, S. (2001). *Acoustic Correlates of Emotion Dimensions in View of Speech Synthesis. 1*, 87-90. <https://doi.org/10.21437/Eurospeech.2001-34>
- Schulte-Rüther, M., Markowitsch, H. J., Shah, N. J., Fink, G. R., & Piefke, M. (2008). Gender differences in brain networks supporting empathy. *NeuroImage, 42*(1), 393-403. <https://doi.org/10.1016/j.neuroimage.2008.04.180>

- Searle, S. R., Speed, F. M., & Milliken, G. A. (1980). Population Marginal Means in the Linear Model : An Alternative to Least Squares Means. *The American Statistician*, 34(4), 216-221. <https://doi.org/10.2307/2684063>
- Selkirk, E. (1984). *Phonology and Syntax : The Relation between Sound and Structure*. The MIT Press. <https://mitpress.mit.edu/9780262690980/phonology-and-syntax/>
- Selkirk, E. O. (1980). The Role of Prosodic Categories in English Word Stress. *Linguistic Inquiry*, 11(3), 563-605.
- Selting, M. (2010). Affectivity in conversational strorytelling : An analysis of displays of anger or indignation in complaint stories. *Pragmatics*, 20(2), 229-277.
- Seo, Y.-S., & Huh, J.-H. (2019). Automatic Emotion-Based Music Classification for Supporting Intelligent IoT Applications. *Electronics*, 8(2), 164. <https://doi.org/10.3390/electronics8020164>
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press.
- Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. (2009). VoiceSauce : A program for voice analysis. *The Journal of the Acoustical Society of America*, 126(4_Supplément). <https://doi.org/10.1121/1.3248865>
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI : A standard for labeling English prosody. *Proceedings of ICSLP*, 867-870. <https://doi.org/10.21437/ICSLP.1992-260>
- Skinner, E. R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness; and a determination of the pitch and force of the subjective concepts of ordinary, soft and loud tones.

- Communications Monographs*, 2(1), 81-137.
<https://doi.org/10.1080/03637753509374833>
- Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, 116(38), 18888-18892. <https://doi.org/10.1073/pnas.1908369116>
- Soroka, S., & McAdams, S. (2015). News, Politics, and Negativity. *Political Communication*, 32(1), 1-22. <https://doi.org/10.1080/10584609.2014.881942>
- Sóskuthy, M. (2017a). Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction. *arXiv (Cornell University)*.
<http://export.arxiv.org/pdf/1703.05339>
- Sóskuthy, M. (2017b, mars 15). *Generalised additive mixed models for dynamic analysis in linguistics: A practical introduction*. arXiv.Org.
<https://arxiv.org/abs/1703.05339v1>
- Sperber, D., & Wilson, D. (2004). *Relevance Theory*. (L. R. Horn & G. Ward, Éds.). The Handbook of Pragmatics. <https://www.dan.sperber.fr/?p=93>
- Stibbard, R. (2000). Automated Extraction Of Tobi Annotation Data From The Reading/Leeds Emotional Speech Corpus. *Proceedings of the ISCA ITRW on Speech and Emotion*, 6.
https://www.researchgate.net/publication/2319384_Automated_Extraction_Of_Tobi_Annotation_Data_From_The_Readingleeds_Emotional_Speech_Corpus
- Stibbard, R. (2001). *Vocal expressions of emotions in non-laboratory speech: An investigation of the Reading/Leeds Emotion in Speech Project annotation data*. [Ph.D., University of Reading].
<https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.343222>

- Stivers, T. (2008). Stance, Alignment, and Affiliation During Storytelling : When Nodding Is a Token of Affiliation. *Research on Language & Social Interaction*, 41(1), 31-57. <https://doi.org/10.1080/08351810701691123>
- Stivers, T., & Robinson, J. D. (2006). A preference for progressivity in interaction. *Language in Society*, 35(3), 367-392. <https://doi.org/10.1017/S0047404506060179>
- Stoller, D., & Dixon, S. (2016, mai 13). Analysis and classification of phonation modes in singing. *17th International Society for Music Information Retrieval. ISMIR 2016*, NY city, USA. <https://qmro.qmul.ac.uk/xmlui/handle/123456789/13500>
- Stranovská, E., Munková, D., & Munk, M. (2016). A 'Broad Categorizer'and Language Modeling in Relation to Selected Expressive Factors. *Procedia - Social and Behavioral Sciences*, 217, 923-928. <https://doi.org/10.1016/j.sbspro.2016.02.050>
- Streeck, J. (1995). On projection. In E. N. Goody (Éd.), *Social Intelligence and Interaction : Expressions and implications of the social bias in human intelligence* (p. 87-110). Cambridge University Press. <https://doi.org/10.1017/CBO9780511621710.007>
- Swerts, M., & Hirschberg, J. (2010). Prosodic predictors of upcoming positive or negative content in spoken messages. *The Journal of the Acoustical Society of America*, 128(3), 1337-1345. <https://doi.org/10.1121/1.3466875>
- Terasaki, A. (1976). Pre-announcement sequences in conversation. *Social Science Working Paper 99*.
- Terasaki, A. K. (2004). Pre-announcement sequences in conversation. In G. H. Lerner (Éd.), *Conversation Analysis: Studies from the first generation* (p. 171-223). John Benjamins Publishing Company. <https://doi.org/10.1075/pbns.125.11ter>

- Thompson, W., Peter, V., Olsen, K. N., & Stevens, C. J. (2012). The effect of intensity on relative pitch. *Quarterly Journal of Experimental Psychology*, 65(10), 2054-2072. <https://doi.org/10.1080/17470218.2012.678369>
- Tieu, L., Pasternak, R., Schlenker, P., & Chemla, E. (2018). Co-speech gesture projection : Evidence from inferential judgments. *Glossa: A Journal of General Linguistics*, 3(1), Article 1. <https://doi.org/10.5334/gjgl.580>
- Tilsen, S. (2019). Space and time in models of speech rhythm. *Annals of the New York Academy of Sciences*, 1453(1), 47-66. <https://doi.org/10.1111/nyas.14102>
- Tolkmitt, F. J., & Scherer, K. R. (1986). Effect of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology. Human Perception and Performance*, 12(3), 302-313. <https://doi.org/10.1037//0096-1523.12.3.302>
- Tracy, K., & Craig, R. T. (2019). Decision announcements in small claims court : A recurring judge dilemma. *Language and Dialogue*, 9(1), 65-83. <https://doi.org/10.1075/ld.00032.tra>
- Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, 4(1), 93-118. <https://doi.org/10.1515/lp-2013-0005>
- Vandekieft, G. K. (2001). Breaking Bad News. *American Family Physician*, 64(12), 1975-1979.
- Vanrell, M. del M. (2007). A tonal scaling contrast in Majorcan Catalan interrogatives. *Journal of Portuguese Linguistics*, 6(1), Article 1. <https://doi.org/10.5334/jpl.148>
- Vidrascu, L., & Devillers, L. (2007, août). Five emotion classes detection in real-world call center data : The use of various types of paralinguistic features. *Proc. Inter. workshop on Paralinguistic Speech between models and data, ParaLing. ParaLing'07*, Saarbrücken, Allemagne.

- Vlemings, J. (2003). The discourse use of French *donc* in imperative sentences. *Journal of Pragmatics*, 35(7), 1095-1112. [https://doi.org/10.1016/S0378-2166\(03\)00024-9](https://doi.org/10.1016/S0378-2166(03)00024-9)
- Wang, W., Watters, P. A., Cao, X., Shen, L., & Li, B. (2020). Significance of Phonological Features in Speech Emotion Recognition. *International Journal of Speech Technology*, 23(3), 633-642. <https://doi.org/10.1007/s10772-020-09734-7>
- Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech*, 46(1), 53-81. <https://doi.org/10.1177/00238309030460010401>
- Welby, P. (2006). French intonational structure : Evidence from tonal alignment. *Journal of Phonetics*, 34(3), 343-371. <https://doi.org/10.1016/j.wocn.2005.09.001>
- Welby, P., & Loevenbruck, H. (2005). Segmental “Anchorage” and the French late rise. *Interspeech 2005: the 9th Annual Conference on Speech Communication and Technology*, 2369-2372. <https://hal.science/hal-00371967>
- Welby, P., & Loevenbruck, H. (2006). Anchored down in Anchorage : Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics / Rivista di linguistica*, 18, 74-124.
- Wells, B., & Macfarlane, S. (1998). Prosody as an Interactional Resource : Turn-projection and Overlap. *Language and Speech*, 41(3-4), 265-294. <https://doi.org/10.1177/002383099804100403>
- Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health*, 13(4), 227-236. <https://doi.org/10.1002/nur.4770130405>
- Wichmann, A. (2000). The attitudinal effects of prosody, and how they relate to emotion. *ITRW on Speech and Emotion*, 143-148. https://www.isca-speech.org/archive_open/speech_emotion/spem_143.html

- Wichmann, A. (2011). Prosody and pragmatic effects. In *Pragmatics of Society* (p. 181-214). De Gruyter Mouton. <https://doi.org/10.1515/9783110214420.181>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* [R]. <https://ggplot2.tidyverse.org>
- Wieling, M. (2018). Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English. *Journal of Phonetics*, 70, 86-116. <https://doi.org/10.1016/j.wocn.2018.03.002>
- Wongpakaran, N., Wongpakaran, T., Wedding, D., & Gwet, K. L. (2013). A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: A study conducted with personality disorder samples. *BMC Medical Research Methodology*, 13(1), 61. <https://doi.org/10.1186/1471-2288-13-61>
- Wood, S. (2017). Generalized Additive Models: An Introduction With R. In *Generalized Additive Models: An Introduction with R, Second Edition*. CRC press. <https://doi.org/10.1201/9781315370279>
- Xiao, X., Kuhnert, B., Audibert, N., Locqueville, G., Pillot-Loiseau, C., Zhang, H., & d'Alessandro, C. (2023). *Performative Vocal Synthesis for Foreign Language Intonation Practice*. 1. <https://doi.org/10.1145/3544548.3581210>
- Yang, Z., Huynh, J., Tabata, R., Cestero, N., Aharoni, T., & Hirschberg, J. (2020). What Makes a Speaker Charismatic? Producing and Perceiving Charismatic Speech. *Proc. 10th International Conference on Speech Prosody*, 685-689. <https://doi.org/10.21437/SpeechProsody.2020-140>
- Yoon, D., Audibert, N., & Fougeron, C. (2020). Effets du sexe et de la langue parlée sur la production de la parole chez les locuteurs coréens et français. In C. Benoit, C. Braud, L. Huber, D. Langlois, S. Ouni, S. Pogodalla, & S. Schneider (Éds.), *6e*

- conférence conjointe Journées d'Études sur la Parole (JEP, 33e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition). Volume 1 : Journées d'Études sur la Parole (p. 645-653). ATALA. <https://hal.science/hal-02798598>
- Yuan, J., Shen, L., & Chen, F. (2002, septembre 16). *The acoustic realization of anger, fear, joy and sadness in Chinese*. <https://doi.org/10.21437/ICSLP.2002-556>
- Yumoto, E. (1983). The quantitative evaluation of hoarseness. *Archives of Otolaryngology (Chicago, Ill.: 1960)*, 109(1), 48-52.
- Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, 71(6), 1544-1549. <https://doi.org/10.1121/1.387808>
- Zarrilli, P. B., & Hulton, P. (2009). *Psychophysical Acting : An Intercultural Approach After Stanislavski*. Routledge.
- Zénone, A. (1981). Marqueurs de consécution : Le cas de donc. *Cahiers de linguistique française*, 2, 113-139.

Table des figures

Figure 1 – Exemple de hiérarchie de préférence entre les réponses	33
Figure 2 – Représentation schématique du cours de l'interaction.....	35
Figure 3 – Représentation schématique de la différence entre pré-annonce et préface à l'annonce, proposée d'après la distinction de Maynard (2017)	44
Figure 4 – Représentation du poids des syllabes, extrait de Liberman et Prince (1977)	60
Figure 5 – Exemple de représentation des tons dans la théorie AM (à droite), d'après Ladd (2008)	62
Figure 6 – Patron tonal sous-jacent du français	65
Figure 7 – Les cinq réalisations possibles du patron tonal sous-jacent, d'après Jun et Fougeron (2000, 2002)	66
Figure 8 – Représentation schématique de la hiérarchie entre les niveaux de la phrase prosodique.....	68
Figure 9 – Résumé des types d'affects, exemples, et caractéristiques, extrait de Karunaratne et al. (2011).....	74
Figure 10 – Modèle circomplexe (Russel, 2003, p.143).....	79
Figure 11 – Tableau des différences acoustiques entre des émotions discrètes, extrait de Scherer (2003)	87
Figure 12 – Exemples de messages vocaux d'annonce avec décision, extrait de Swerts et Hirschberg (2010)	96
Figure 13 – Présentation des stimuli pour l'enregistrement des annonces de nouvelles	116
Figure 14 – Fichiers Praat et audio avec les tires d'annotation automatique : Transcription, PhonAlign, TokensAlign et SyllAlign	118
Figure 15 – Empan de la f0 (Hz) dans les préfaces par valence de la nouvelle	127
Figure 16 – Moyenne de f0 (Hz) dans les préfaces par valence de la nouvelle	127
Figure 17 – f0 maximum (Hz) dans les préfaces par valence de la nouvelle.....	128
Figure 18 – f0 minimum (Hz) dans les préfaces par valence de la nouvelle	128

Figure 19 – Empan de la f0 (Hz) dans les marqueurs discursifs par valence de la nouvelle.....	130
Figure 20 – Moyenne de f0 (Hz) dans les marqueurs discursifs par valence de la nouvelle.....	131
Figure 21 – f0 maximum (Hz) dans les marqueurs discursifs par valence de la nouvelle.....	131
Figure 22 – Empan de la f0 (Hz) dans les annonces par valence de la nouvelle.....	133
Figure 23 – Moyenne de f0 (Hz) dans les annonces par valence de la nouvelle.....	133
Figure 24 – f0 maximum (Hz) dans les annonces par valence de la nouvelle.....	134
Figure 25 – f0 minimum (Hz) dans les annonces par valence de la nouvelle.....	134
Figure 26 – Courbes lissées de la f0 au sein de la préface par valence.....	137
Figure 27 – Différences significatives de la f0 au sein de la préface.....	137
Figure 28 – Courbes lissées de la f0 au sein des DMs par valence.....	139
Figure 29 – Différences significatives de la f0 au sein des DMs.....	139
Figure 30 – Durée de la préface (sec) par valence de la nouvelle.....	141
Figure 31 – Durée des pauses (sec) dans les préfaces par valence de la nouvelle.....	141
Figure 32 – Durée des marqueurs discursifs (sec) par valence de la nouvelle.....	143
Figure 33 – Durée de la séquence d’annonce dans les annonces par valence de la nouvelle.....	144
Figure 34 – Durée des pauses (sec) dans les annonces par valence de la nouvelle.....	145
Figure 35 – HNR05 au sein de la préface par valence.....	147
Figure 36 – Différences significatives de HNR05 au sein de la préface.....	147
Figure 37 – HNR15 au sein de la préface par valence.....	148
Figure 38 – Différences significatives de HNR15 au sein de la préface.....	148
Figure 39 – HNR25 au sein de la préface par valence.....	149
Figure 40 – Différences significatives de HNR25 au sein de la préface.....	149
Figure 41 – HNR35 au sein de la préface par valence.....	150
Figure 42 – Différences significatives de HNR35 au sein de la préface.....	150
Figure 43 – H1-H2 au sein de la préface par valence.....	152
Figure 44 – Différences significatives de H1-H2 au sein de la préface.....	152
Figure 45 – CPPs au sein de la préface par valence.....	153
Figure 46 – Différences significatives du CPPs au sein de la préface.....	154
Figure 47 – HNR05 au sein des DMs par valence.....	156

Figure 48 – Différences significatives de HNR05 au sein des DMs.....	156
Figure 49 – HNR15 au sein des DMs par valence.....	157
Figure 50 – Différences significatives de HNR15 au sein des DMs.....	157
Figure 51 – HNR25 au sein des DMs par valence.....	158
Figure 52 – Différences significatives de HNR25 au sein des DMs.....	158
Figure 53 – HNR35 au sein des DMs par valence.....	159
Figure 54 – Différences significatives de HNR35 au sein des DMs.....	159
Figure 55 – H1-H2 au sein des DMs par valence.....	161
Figure 56 – Différences significatives de H1-H2 au sein des DMs.....	161
Figure 57 – CPPs au sein des DMs par valence.....	162
Figure 58 – Différences significatives du CPPs au sein des DMs.....	163
Figure 59 – Exemple de configuration de fin d’IP.....	178
Figure 60 – Schématisation et correspondance ToBI des contours phonologiques associés aux fonctions des marqueurs discursifs « voilà » ¹ (Petit, 2020) et « donc » ² (Ferre, 2011).....	184
Figure 61 – Fichiers audio et TextGrid alignés, avec chaque tire utilisée pour l’annotation : <i>tones</i> (avec par exemple le ton L marqué L1), <i>boundaries</i> (où (marque le début de l’unité, AP la fin du syntagme accentuel, et IP la fin du syntagme intonational, <i>misc</i> , <i>SyllAlign</i> , <i>TokenAlign</i>).....	188
Figure 62 – Spectrogramme, courbe de f0 et TextGrid pour la préface <i>Au sujet de votre enfant Léo</i> , prononcée par la locutrice P04. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales L* et L% réalisés au même point.....	192
Figure 63 – Spectrogramme, courbe de f0 et TextGrid pour la préface <i>Au sujet de votre hésitation entre deux cours</i> , prononcée par la locutrice P03. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales L* et H%.....	192
Figure 64 – Spectrogramme, courbe de f0 et TextGrid pour la préface <i>Au sujet de votre vase en porcelaine chinoise</i> , prononcée par la locutrice P02. Dans le TextGrid, nous voyons un exemple d’annotation des cibles tonales H* et L%, avec une détection de la f0 perturbée en fin de préface.....	193
Figure 65 – Spectrogramme, courbe de f0 et TextGrid pour l’énoncé « <i>voilà donc</i> », prononcé par la locutrice P02 (item 6). Dans le TextGrid, nous voyons un exemple d’annotation du patron sous-jacent du français, où toutes les cibles font surface : LHiLH*.	

Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial, L2 marque le ton L de l'accent final	195
Figure 66 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « <i>voilà donc</i> », prononcée par la locutrice P06 (item 10). Dans le TextGrid, nous voyons un exemple d'annotation du patron LH*. Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial	195
Figure 67 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « <i>voilà</i> », prononcée par la locutrice P07 (item 2). Dans le TextGrid, nous voyons un exemple d'annotation du patron LHL*, où L* et L- sont réalisés ensemble. Le symbole (marque le début de l'unité, L1 marque le ton L de l'accent initial	196
Figure 68 – Spectrogramme, courbe de f0 et TextGrid pour l'énoncé « <i>donc</i> », prononcé par la locutrice P30 (item 9). Dans le TextGrid, nous voyons un exemple d'annotation du patron L*, où L* et L% sont réalisés ensemble	196
Figure 69 – Valeur mélodique des tons L% en fin de préface	207
Figure 70 – Valeur mélodique des tons H% en fin de préface	207
Figure 71 – Valeur mélodique des tons L* en fin de préface	208
Figure 72 – Valeur mélodique des tons H* en fin de préface	208
Figure 73 – Valeur mélodique des tons L% dans « <i>voilà</i> » (à gauche), et « <i>donc</i> » (à droite)	212
Figure 74 – Valeur mélodique des tons H% dans « <i>voilà donc</i> » (à gauche), « <i>voilà</i> » (au milieu) et « <i>donc</i> » (à droite)	212
Figure 75 – Valeur mélodique de L1 dans « <i>voilà</i> »	213
Figure 76 – Valeur mélodique des tons H* dans « <i>voilà donc</i> » (à gauche), « <i>voilà</i> » (au milieu) et « <i>donc</i> » (à droite)	215
Figure 77 – Valeur mélodique des tons L* dans « <i>voilà</i> » (à gauche) et « <i>donc</i> » (à droite)	215
Figure 78 – Valeur mélodique des tons Hi dans « <i>voilà donc</i> » (à gauche), « <i>voilà</i> » (à droite)	217
Figure 79 – Valeur mélodique de L2 dans « <i>voilà donc</i> »	217
Figure 80 – Alignement des tons H* à la syllabe au sein de « <i>voilà donc</i> »	219
Figure 81 – Alignement des tons L* à la syllabe au sein de « <i>voilà</i> »	220
Figure 82 – Alignement de Hi au sein de « <i>voilà donc</i> »	221
Figure 83 – Illustration du design experimental de la tâche de validation	245

Figure 84 – Illustration du design experimental de la tâche principale	250
Figure 85 – Proportion de « ah oui » ou « oh non » et erreur-standard par valence et par genre dans la tâche 1.....	257
Figure 86 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 1.....	259
Figure 87 – Proportion de « j’écoute » ou « oh non » et erreur-standard par valence et par genre dans la tâche 2.....	261
Figure 88 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 2.....	263
Figure 89 – Proportion de « ah oui » ou « j’écoute » et erreur-standard par valence et par genre dans la tâche 3.....	265
Figure 90 – Temps de réaction (ms) pour chaque valence en fonction de la congruence de la réponse et du genre, dans la tâche 3.....	267
Figure 91 – Temps de réaction (ms) en fonction de la congruence de la réponse et du genre et du niveau d’empathie, à travers les trois tâches.....	271

Liste des tableaux

Tableau 1 – Exemple de scénario d’annonce de nouvelle.....	112
Tableau 2 – Présentation de la structure des appels d’annonce de nouvelle.....	114
Tableau 3 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans la préface	126
Tableau 4 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans les marqueurs discursifs	129
Tableau 5 – Tableau des valeurs moyennes et déviations-standard (Hz) pour les mesures globales de la f0 dans les séquences d’annonce	132
Tableau 6 – Coefficients paramétriques et termes lisses du modèle GAMM réalisé sur la courbe de f0 au sein des préfaces.....	135
Tableau 7 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les préfaces	140
Tableau 8 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les marqueurs discursifs	142
Tableau 9 – Tableau des valeurs moyennes et déviations-standard (syll/sec et sec) pour les mesures globales de données temporelles dans les séquences d’annonce	143
Tableau 10 – Tableau récapitulatif des différences phonétiques entre les trois valences, au sein des différents énoncés (< valeur basse, > valeur haute, <> intermédiaire, _ non significatif, >< inconsistant, X NA)	164
Tableau 11 – Matrice de confusion : nombre d’occurrence du type d’accent mélodique dans la préface.....	199
Tableau 12 – Matrice de confusion : nombre d’occurrence du type de ton de frontière dans la préface.....	200
Tableau 12 – Matrice de confusion : nombre d’occurrence du type de d’accent mélodique dans les marqueurs discursifs.....	200
Tableau 14 – Matrice de confusion : nombre d’occurrence en fonction force de la frontière prosodique dans les marqueurs discursifs.....	201
Tableau 15 – Matrice de confusion : nombre d’occurrence de la présence ou l’absence d’une frontière prosodique entre « voilà » et « donc ».....	201

Tableau 16 – Distribution des configurations de fin d’IP dans la préface, selon la valence de la nouvelle.....	202
Tableau 17 – Choix d’une ou deux unités phonologiques dans les marqueurs discursifs	203
Tableau 18 – Distribution des patrons phonologiques dans les marqueurs discursifs, selon la valence de la nouvelle	204
Tableau 19 – Valeurs moyennes et SD (Hertz) pour les cibles tonales évaluées au sein de la préface	206
Tableau 20 – Valeurs moyennes et SD (Hertz) pour les cibles tonales L%, H% et L1 évaluées au sein des marqueurs discursifs.....	210
Tableau 21 – Valeurs moyennes et SD (Hertz) pour les cibles tonales L* et H* évaluées au sein des marqueurs discursifs	213
Tableau 22 – Valeurs moyennes et SD (Hertz) pour les cibles tonales Hi et L2 évaluées au sein des marqueurs discursifs	216
Tableau 23 – Valeurs moyennes et SD d’alignement (%) pour les cibles tonales L* et H* évaluées au sein des marqueurs discursifs.....	218
Tableau 24 – Valeurs moyennes et SD d’alignement (%) pour les cibles tonales Hi et L2 évaluées au sein des marqueurs discursifs.....	220
Tableau 25 – Tableau récapitulatif des différences phonologiques et phonétiques entre les trois valences, au sein des différents énoncés (< moins présent / plus tard , > valeur haute, <> intermédiaire, _ non significatif, >< inconsistant, X NA).....	222
Tableau 26 – Tâche 1. Proportion du choix de feedback en fonction de la valence de la nouvelle dans la préface.....	256
Tableau 27 – Tâche 1. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse.....	258
Tableau 28 – Tâche 2. Proportion moyenne du choix de feedback en fonction de la valence de la nouvelle dans la préface	260
Tableau 29 – Tâche 2. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse.....	262
Tableau 30 – Tâche 3. Proportion du choix de feedback en fonction de la valence de la nouvelle dans la préface.....	264
Tableau 31 – Tâche 3. Temps de reaction moyen et SD en fonction de la valence de la nouvelle, du genre et de la congruence de la réponse.....	266

Tableau 32 – Proportion du niveau d'empathie par genre dans chaque tâche	269
Tableau 33 – Comparaison inter-tâches. Temps de reaction moyen et SD en fonction du genre, du niveau d'empathie et de la congruence de la réponse.....	270
Tableau 34 – Tableau récapitulatif des choix selon la valence et le genre et des RT selon la congruence, le genre et l'empathie (< moins choisi / RT plus court , > plus choisi / RT plus longue, _ non significatif, X NA)	272

Annexes

A. Etude 1 – Analyses phonétiques.

Annexe 1. Scénarios initiaux (21) – étude préliminaire 1

Recrutement 1

NEUTRE : Vous êtes directrice/directeur des ressources humaines et vous devez tenir informée l'une des candidates au poste. Vous avez bien reçu sa candidature. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes directrice/directeur des ressources humaines et vous devez tenir informée l'une des candidates au poste. Vous avez fait votre choix : elle a obtenu le poste. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes directrice/directeur des ressources humaines et vous devez tenir informée l'une des candidates au poste. Vous avez fait votre choix : elle n'a pas été retenue pour le poste. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Médicale 2

NEUTRE : Vous êtes médecin et vous recevez une patiente régulière. Vous allez lui faire un examen. Aujourd'hui, vous vous êtes installé dans le cabinet en face du vôtre car ce dernier est en rénovation. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes médecin et vous avez reçu une patiente. Vous lui avez fait faire un examen récemment : ses analyses sont meilleures que la dernière fois. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes médecin et vous recevez une patiente. Vous lui avez fait faire un examen récemment : ses analyses sont mauvaises et vous devez en discuter. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Aéroport 3

NEUTRE : Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol a été déplacé du terminal 1 au terminal 4. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol est un long courrier et elle a un siège « classe économique ». Vous pouvez la relocaliser en première classe pour le même prix. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes chef d'escale et vous vous occupez d'une cliente. Son vol a été annulé et vous n'avez pas trouvé de solution : elle ne pourra prendre l'avion que le lendemain. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Juridique 4

NEUTRE : Vous êtes avocate/avocat et l'une de vos clientes approche de la date de son procès. Vous avez pris connaissance d'un déplacement du lieu du procès : il aura lieu en salle 3 et non en salle 2. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes avocate/avocat et l'une de vos clientes approche de la date de son procès. Vous avez pris connaissance de nouvelles informations dans le dossier : elles sont en faveur de son innocence. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes avocate/avocat et l'une de vos clientes approche de la date de son procès. Vous avez pris connaissance de nouvelles informations dans le dossier : elles sont incriminantes. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Hôtel vacances 5

NEUTRE : Vous êtes hôtelière/hôtelier et une cliente veut réserver une chambre. Elle veut la chambre 33 : laquelle se situe au quatrième étage. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes hôtelière/hôtelier et une cliente veut réserver une chambre. Elle veut la chambre 33 : vous allez lui offrir une chambre plus spacieuse pour le même prix. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes hôtelière/hôtelier et une cliente veut réserver une chambre. Elle veut la chambre 33 : elle est déjà prise et vous ne pouvez donc pas accepter la réservation. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Entreprise 6

NEUTRE : Vous êtes directrice/directeur des ressources humaines et une employée a récemment fait une demande d'augmentation. Vous allez évaluer la demande : vous la tiendrez au courant dans les plus brefs délais. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes directrice/directeur des ressources humaines et une employée a récemment fait une demande d'augmentation. Vous avez réfléchi : vous allez l'augmenter. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes directrice/directeur des ressources humaines et une employée a récemment fait une demande d'augmentation. Vous avez réfléchi : vous ne pouvez pas l'augmenter. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Voyage scolaire 7

NEUTRE : Vous êtes professeur et organisez un voyage scolaire. Vous êtes en contact avec un parent d'élève. Le point de rassemblement pour le voyage a changé et il est maintenant devant l'école. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes professeur et organisez un voyage scolaire. Vous êtes en contact avec un parent d'élève. L'école a pu débloquer des fonds et va financer une plus grande partie du voyage que prévu. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes professeur et organisez un voyage scolaire. Vous êtes en contact avec un parent d'élève. L'école n'a pas pu débloquer les fonds annoncés et le voyage va être plus cher que prévu. Vous l'appellez et laissez un message sur son répondeur pour lui annoncer.

Camping 8

NEUTRE : Vous êtes gérante/gérant d'un camping et l'une de vos employées s'occupe du nettoyage des emplacements. Vous avez passé une commande pour des outils de

nettoyage : ils arrivent dans une semaine. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes gérante/gérant d'un camping et l'une de vos employées s'occupe du nettoyage des emplacements. Vous avez passé une commande pour des outils de nettoyage : ils sont performants et feront gagner du temps. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes gérante/gérant d'un camping et l'une de vos employées s'occupe du nettoyage des emplacements. Vous avez passé une commande pour des outils de nettoyage : elle est arrivée et il en manque la moitié. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Entraîneur sportif 9

NEUTRE : Vous êtes coach et la sportive que vous entraînez vous a récemment parlé d'une compétition à laquelle elle veut participer. Vous vous êtes renseigné.e : la ville d'accueil sera Metz. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes coach et la sportive que vous entraînez vous a récemment parlé d'une compétition à laquelle elle veut participer. Vous vous êtes renseigné.e : vous acceptez de la faire participer. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes coach et la sportive que vous entraînez vous a récemment parlé d'une compétition à laquelle elle veut participer. Vous vous êtes renseigné.e : vous n'allez pas la faire participer. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Vétérinaire 10

NEUTRE : Vous êtes vétérinaire et une propriétaire vous a récemment amené un chat malade. Un changement d'emploi du temps est survenu donc elle devra venir le chercher 30 minutes plus tard. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes vétérinaire et une propriétaire vous a récemment amené un chat malade. Il est bien soigné et est guéri. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes vétérinaire et une propriétaire vous a récemment amené un chat malade. Le traitement pour la maladie qu'il a est très cher. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Paysan/producteur laitier 11

NEUTRE : Vous êtes productrice/producteur de lait et vous travaillez avec une associée. Vous avez changé les vaches de champ cet après-midi : elles sont dans celui en face de la ferme. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes productrice/producteur de lait et vous travaillez avec une associée. Vous avez changé les vaches de champ cette après-midi : vous avez aidé l'une d'elles à vêler et le veau est en bonne santé. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes productrice/producteur de lait et vous travaillez avec une associée. Vous avez changé les vaches de champ cette après-midi : vous avez trouvé l'une d'entre-elles morte. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Librairie 12

NEUTRE : Vous êtes libraire et avez une employée. Vous avez opéré quelques changements dernièrement qui ont modifié les rayons : les livres fantasy ne seront plus mis dans le rayon jeunes adultes mais à côté des livres de science-fiction. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes libraire et avez une employée. Vous avez opéré quelques changements dernièrement qui ont modifié les rayons : cela a augmenté votre chiffre d'affaires et vous allez pouvoir donner une prime à votre employée. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes libraire et avez une employée. Vous avez opéré quelques changements dernièrement qui ont modifié les rayons : cela a fait baisser votre chiffre d'affaires et vous n'allez pouvoir l'employer qu'à temps partiel à compter de septembre. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

En cuisine 13

NEUTRE : Vous êtes cheffe/chef cuisinier/cuisinière et avez une cliente. Vous avez commandé des ingrédients pour une fête récemment : vous lui amènerez tout samedi. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes cheffe/chef cuisinier et avez une cliente. Vous avez commandé des ingrédients pour une fête récemment : le fournisseur vous a offert des ingrédients en plus pour votre fidélité et vous proposez de faire avec un plat supplémentaire avec sans frais additionnels. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes cheffe/chef cuisinier et avez une cliente. Vous avez commandé des ingrédients pour une fête récemment : à cause de grèves de transports, ils arriveront trop tard. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Agence de communication 14

NEUTRE : Vous êtes directrice/directeur dans une agence de communication et dirigez une chef de projet. Vous avez pensé à un destinataire très en vogue : vous avez fouillé dans vos papiers pour retrouver son adresse mail et la lui donner. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes directrice/directeur dans une agence de communication et dirigez une chef de projet. Vous avez pensé à un destinataire très en vogue : le connaissant personnellement, vous l'avez contacté vous-même et il a accepté le travail. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes directrice/directeur dans une agence de communication et dirigez une chef de projet. Vous avez pensé à un destinataire très en vogue : après avoir d'abord accepté, il vient de vous appeler pour annuler au dernier moment. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Commissaire-priseur 15

NEUTRE : Vous êtes commissaire-priseur et vous vous occupez d'expertises et de ventes aux enchères pour une cliente. On vous a demandé d'expertiser un vase en porcelaine

chinoise : vous venez de finir. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes commissaire-priseur et vous vous occupez d'expertises et de ventes aux enchères pour une cliente. On vous a demandé d'expertiser un vase en porcelaine chinoise : il a une plus grande valeur que ce qu'elle pensait. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes commissaire-priseur et vous vous occupez d'expertises et de ventes aux enchères pour une cliente. On vous a demandé d'expertiser un vase en porcelaine chinoise : elle a été victime de fraude car ce n'est qu'une copie récente ayant beaucoup moins de valeur. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Musée 16

NEUTRE : Vous êtes conservateur de musée et une maitresse d'école vient chaque année avec sa classe. Depuis quelques années vous changez la collection et l'organisation un peu avant sa visite : vous venez de vous décider sur ce que vous allez montrer aux enfants. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes conservateur de musée et une maitresse d'école vient chaque année avec sa classe. Depuis quelques années vous changez la collection et l'organisation un peu avant sa visite : cette année la nouvelle exposition correspond tout à fait au programme de la classe. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes conservateur de musée et une maitresse d'école vient chaque année avec sa classe. Depuis quelques années vous changez la collection et l'organisation un peu avant sa visite : cette année vous ne pouvez pas la changer, donc les enfants qui l'ont déjà vue vont s'ennuyer. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Fleuriste 17

NEUTRE : Vous êtes fleuriste et avez une employée qui vous aide pour la vente et les arrangements floraux. La Saint-Valentin arrive bientôt : la cargaison de fleurs vient d'être livrée. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes fleuriste et avez une employée qui vous aide pour la vente et les arrangements floraux. La Saint-Valentin arrive bientôt : après cette date vous la formerez pour reprendre la boutique. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes fleuriste et avez une employée qui vous aide pour la vente et les arrangements floraux. La Saint-Valentin arrive bientôt : dû à une récente grêle vous n'aurez pas de fleurs à vendre et elle ne recevra donc pas la prime habituelle. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Mode 18

NEUTRE : Vous êtes créatrice/créateur de mode et employez régulièrement une couturière. Vous êtes en train de créer une nouvelle collection de vêtements : les rubans que vous avez commandés lui seront livrés demain. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes créatrice/créateur de mode et employez régulièrement une couturière. Vous êtes en train de créer une nouvelle collection de vêtements. Vous voulez lui offrir un CDI à la fin de ce contrat. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes créatrice/créateur de mode et employez régulièrement une couturière. Vous êtes en train de créer une nouvelle collection de vêtements : la couturière a commencé à coudre des robes mais la couleur ne vous convient finalement pas donc vous voulez qu'elle refasse tout. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Administration 19

NEUTRE : Vous êtes directrice/directeur d'équipe et gérez une cadre. Elle vous a récemment demandé à passer en télétravail le vendredi : vous avez fait remonter la demande à votre supérieur et attendez la réponse. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes directrice/directeur d'équipe et gérez une cadre. Elle vous a récemment demandé à passer en télétravail le vendredi : cela ne posera pas de problème donc vous allez accepter. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes directrice/directeur d'équipe et gérez une cadre. Elle vous a récemment demandé à passer en télétravail le vendredi : une autre collègue est déjà en télétravail ce jour-là donc vous allez refuser, elle doit venir sur place même si elle habite loin. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Garderie 20

NEUTRE : Vous êtes gestionnaire d'une garderie et un enfant est sur liste d'attente alors que sa mère doit reprendre le travail rapidement. La nouvelle année va commencer : la décision sera prise au plus tôt. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes gestionnaire d'une garderie et un enfant est sur liste d'attente alors que sa mère doit reprendre le travail rapidement. La nouvelle année va commencer : l'enfant est accepté. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes gestionnaire d'une garderie et un enfant est sur liste d'attente alors que sa mère doit reprendre le travail rapidement. La nouvelle année va commencer : l'enfant n'est pas accepté. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Scolarité 21

NEUTRE : Vous êtes directrice d'un département de la faculté et avez reçu un mail d'une étudiante. Elle veut prendre rendez-vous car elle hésite entre deux cours proposés au même horaire : vous décidez de la date du rendez-vous pour en discuter de vive-voix. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

BONNE : Vous êtes directrice d'un département de la faculté et avez reçu un mail d'une étudiante. Elle veut prendre rendez-vous car elle hésite entre deux cours proposés au

même horaire : vous avez trouvé une solution pour lui permettre d'assister aux deux. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

MAUVAISE : Vous êtes directrice d'un département de la faculté et avez reçu un mail d'une étudiante. Elle veut prendre rendez-vous car elle hésite entre deux cours proposés au même horaire : vous vous êtes renseignée et elle ne peut en fait assister à aucun des deux car ils ne sont pas à son niveau. Vous l'appellez et laissez un message sur son répondeur pour l'en informer.

Annexe 2. Résultats – tâche préliminaire 1

Score obtenu	Code	Valence	Item	Explication
-65,9	1M	Négatif	1	Neutre plus haut que positif
87,375	1N	Neutre	1	
70,91666667	1B	Positif	1	
-83,5	2M	Négatif	2	Neutre et positif trop proches
51,3	2N	Neutre	2	
67,8	2B	Positif	2	
-	3M	Négatif	3	Ok
78,91666667	3N	Neutre	3	
-10,75	3B	Positif	3	
85,5	3B	Positif	3	Neutre et positif trop proches
-51,875	4M	Négatif	4	
56,2	4N	Neutre	4	
56,33333333	4B	Positif	4	Ok
-68	5M	Négatif	5	
-13,75	5N	Neutre	5	
96,3	5B	Positif	5	Ok
-	6M	Négatif	6	
60,91666667	6N	Neutre	6	
8,125	6N	Neutre	6	Neutre et positif trop proches
89,72727273	6B	Positif	6	
-52,8	7M	Négatif	7	
74,77777778	7N	Neutre	7	Ok, mais le tout tend vers le positif
72,08333333	7B	Positif	7	
-59,6	8M	Négatif	8	
17,66666667	8N	Neutre	8	Ok, mais le tout tend vers le positif
88,7	8B	Positif	8	
-	9M	Négatif	9	
49,41666667	9N	Neutre	9	Ok, mais le tout tend vers le positif
39,2	9N	Neutre	9	
83,09090909	9B	Positif	9	
60,85714286	10N	Négatif	10	Neutre et positif trop proches
-84	10M	Négatif	10	
78,58333333	10B	Positif	10	
-47,1	11M	Négatif	11	
-17	11N	Neutre	11	

100	11B	Positif	11	Ok, malgré négatif un peu haut
-56,5	12M	Négatif	12	
0,5	12N	Neutre	12	Ok
83,90909091	12B	Positif	12	
-58,4	13M	Négatif	13	
80	13N	Neutre	13	Neutre et positif trop proches
79,66666667	13B	Positif	13	
-	14M	Négatif	14	
59,36363636	14N	Neutre	14	Ok
48	14B	Positif	14	
82,4	15M	Négatif	15	
31,41666667	15N	Neutre	15	Négatif pas assez bas
56,85714286	15B	Positif	15	
75,9	16M	Négatif	16	
-49,7	16N	Neutre	16	Neutre et positif trop proches
60,5	16B	Positif	16	
78,45454545	17M	Négatif	17	
-	17N	Neutre	17	Ok
66,81818182	17B	Positif	17	
36	18M	Négatif	18	
75,44444444	18N	Neutre	18	Ok
-	18B	Positif	18	
69,18181818	19M	Négatif	19	
-22	19N	Neutre	19	Ok, mais tout tend vers positif
86,90909091	19B	Positif	19	
-67,9	20M	Négatif	20	
37,14285714	20N	Neutre	20	Ok
82,08333333	20B	Positif	20	
-64,8	21M	Négatif	21	
43,77777778	21N	Neutre	21	Neutre trop haut
88,9	21B	Positif	21	
-	21M	Négatif	21	
71,83333333	21N	Neutre	21	
62	21B	Positif	21	
86,72727273	21B	Positif	21	

Annexe 3. Productions libres – tâche préliminaire 2

Locuteur 1

NOUVELLE NEUTRE

1 : Oui bonjour Madame Monat c'est Camille Dumont euh le docteur Camille Dumont euh j'vous appelais euh pour votre rdv de de 10h de tout à l'heure euh alors ne soyez pas étonnée en fait j'me suis installée dans le cabinet juste en face de d'habitude hum parce que mon cabinet est en rénovation euh voilà si vous avez un doute vous pouvez m'appeler bien sûr à tout à l'heure au revoir

2 : Oui bonjour Madame Amélie Monat ? euh je suis le chef d'escalier Camille Dumont et je vous appelais pour vous é vous dire que votre vol a été déplacé euh il se retrouvera au terminal 4 et non plus au terminal 1 voilà euh rendez-vous bien au terminal 4 merci au revoir

3 : Heu oui bonjour madame euh Monat euh c'est Camille Dumont euh euh la gestionnaire de la Niblanc euh j'ai bien reçu votre message concernant le petit Gabriel euh c c'était pour vous dire que pour l'instant il est toujours sur liste d'attente euh on sait que la nouvelle année commence doc euh vraiment on va vrai revenir vers vous euh très prochainement euh fin de semaine au plus tard euh pour vous dire ce qu'il en est merci au revoir

4 : Allo Amélie c'était euh Camille je vous appelais euh pour vous dire que euh on a bien reçu votre demande d'augmentation et qu'on est en train de l'étudier euh je reviendrais vers vous dans les plus brefs délais donc euh d'ici euh d'ici demain ou après-demain maximum en vous souhaitant une bonne journée au revoir

5 : Euh allo Amélie euh c'est Camille Dumont euh j'vous appelais pour vous dire que on a bien passé commande pour les outils de nettoyage donc euh vous devriez les avoir dans une semaine au revoir

6 : Bonjour Madame Monat euh c'était Camille Dumont euh le vétérinaire euh j'essayais de vous joindre pour vous dire qu'il faudrait que vous passiez chercher votre chat euh un peu plus tard que prévu euh on avait dit h il faudrait que vous passiez vers 10h30 voilà si si ça ne vous dérange pas au revoir

7 : Hum allo madame Monat donc euh j'ai essayé de vous joindre c'est euh Camille Dumont le professeur de Jeremy euh je vous appelais pour vous dire que le point de rassemblement pour le voyage a changé donc euh on se retrouve devant l'école demain à 10h donc euh voilà rendez-vous devant l'école à demain au revoir

8 : Euh allo madame Monat donc je vous appelais pour vous dire que c'est Camille Dumont pardon euh j'vous appelais pour vous dire que on a fini d'expertiser votre vase en porcelaine donc vous pouvez venir euh si vous voulez dans la semaine euh pour qu'on vous donne les résultats euh mh voilà alors ah on est disponibles tous les jours entre 10h et 18h donc n'hésitez pas à nous dire ce qui vous arrange le plus à bientôt au revoir

9 : Oui allô Amélie c'était Camille c'était juste pour te dire qu'on a bien fait passer ta demande de télétravail euh à Jean Charles euh voilà il m'a dit qu'il revenait très vite vers moi donc euh on te tient au courant dans la semaine bonne journée au revoir

10 : Bonjour madame Monat euh c'est Camille Dumont euh je vous appelais par rapport au procès de la semaine prochaine euh le procès aura lieu finalement et non pas d'en salle 2 comme annoncé au départ donc venez bien euh salle 3 à la semaine prochaine au revoir

11 : Oui : bonjour Amélie c'est Camille euh j'avais juste un petit message à te faire passer avant qu't'arrives euh cette après-midi euh on a changé les rayons donc euh ne sois pas perturbée maintenant euh les livres fantasy sont à côté des livres de science-fiction donc on avait parlé la semaine dernière et voilà ça paraît mieux que à côté de jeune adulte donc euh voilà c'est tout c'était tout donc euh au revoir

12 : Oui bonjour Amélie euh c'est Camille Dumont euh j'vous appelais parce que j'ai bien reçu votre euh mail euh pour euh pour discuter des de votre choix de de cours euh si vous voulez je suis disponible euh demain demain matin je serai disponible à 8h en salle des professeurs jusqu'à jusqu'à midi en tout cas je suis à l'université jusqu'à midi donc euh revenez vers moi pour l'horaire qui vous convient euh je trouverai un moment merci à bientôt

BONNE NOUVELLE

13 : Bonjour madame Monat euh c'est le docteur Dumont euh je vous rappelais par rapport à au à vos résultats d'examens euh c'était juste pour vous dire que les analyses sont vraiment meilleures que la dernière donc c'est c'est su c'est vraiment bon signe donc bah euh je vous expliquerai tout ça plus en détails mais euh donc euh on se voit dans 2 semaines j crois que votre prochain rdv est dans 2 semaines donc pas besoin de l'avancer donc voilà à très vite merci au revoir

14 : Bonjour madame Monat c'est Camille Dumont le chef d'escalier euh je vous appelais juste pour vous dire que on a la possibilité de vous surclasser euh si vous le souhaitez il y a un siège qui s'est libéré en 1ère classe et jpeux vous le faire au même prix donc voilà c'est si ça vous intéresse vous pouvez me rappeler euh à mon numéro euh voilà j crois quvous l'avez déjà euh à à bientôt merci

15 : Allo madame Monat c'est Camille Dumont le la la gestionnaire de la garderie de la Niblanc euh je vous appelle parce que bonne nouvelle euh Gabriel euh n'est plus sur liste d'attente il est accepté donc voilà vous pouvez l'amener lundi prochain on sera ravis de l'accueillir à très vite au revoir

16 : Oui euh allo Amélie c'était Camille euh bah je t'appelais pour de dire que et bha super nouvelle euh ta demande d'augmentation est acceptée donc euh voilà si tu veux on se retrouve euh euh demain ou après-demain quand quand tu es dispo pour formaliser ça merci

17 : Oui bonjour Amélie c'était donc Camille j't'appelais pour te dire que ça y est les outils de nettoyage sont arrivés et tu verras y sont fin comme on en avait discuté ils sont vraiment performants et ça va t'faire gagner du temps j pense donc euh voilà tu me diras ce que t'en penses à très vite

18 : Bonjour madame Monat c'est euh la vétérinaire Camille Dumont alors je vous appelais pour vous dire que votre chat et ça y est il est il est soigné il est en pleine forme donc euh vous pouvez venir le chercher euh euh dès demain euh voilà les horaires de la clinique c'est euh entre 9h et 17h donc euh passez quand vous le souhaitez

19 : Hum allo madame euh Monat euh c'est Camille Dumont euh le professeur de Jeremie euh je vous appelle par rapport au voyage scolaire euh voilà on a une bonne nouvelle euh l'école a pu débloquent des fonds et donc on va pouvoir financer euh plus de 60% euh du prix total donc on va vous faire parvenir euh le nouveau tarif mais euh voilà je voulais vous annoncer la bonne nouvelle par téléphone euh on se tient au courant au revoir

20 : Euh bonjour madame euh Monat donc c'est Camille Dumont euh le commissaire-priseur euh je vous appelais parce qu'on a terminé d'expertiser euh votre vase euh en porcelaine chinoise et il s'avère que ce n'est pas pas une reproduction comme on pensait mais euh un vase euh Ming du coup donc euh c'est c'est super donc on va vous faire un estimatif plus précis mais il peut vraiment prendre beaucoup de :: de valeur euh aux enchères donc euh rappelez-moi quand quand vous le souhaitez et euh qu'on en parle euh voilà plus précisément à très vite au revoir

21 : Allo Amélie c'est Camille euh j't'appelais par rapport à ta demande de télétravail donc c'est bon ça pose aucun souci tu pourras prendre ton télétravail euh ton jour de télétravail euh tous les vendredis voilà à très vite

22 : Bonjour madame Monat euh c'est Camille Dumont euh j'vous appelais pour vous dire que j'ai pris connaissance de nouvelle information dans votre dossier et vraiment ça va vraiment nous aider à prouver votre innocence euh lors du procès donc c'était vraiment pour vous dire qu'il fallait arriver avec optimisme euh lors de procès ça va bien se passer à très vite

23 : Allo Amélie c'était Camille de La Niblanc euh j't'appelai pour te dire que les derniers changements qu'on a faits sur les rayons fantasy ont on vraiment bien marché là on commence à avoir les résultats après 3 mois et ça ça a vraiment fait augmenter énormément le chiffre d'affaires on est vraiment euh un petit peu bah surpris que ça ait autant marché donc euh bah la bonne nouvelle c'est que :: tu vas pouvoir avoir une prime euh pour Noël voilà donc on va on en discutera euh plus plus précisément ensemble euh demain mais voilà je voulais déjà te dire pas message à demain

24 : Bonjour Amélie euh c'est Camille Dumont j'vous appelais pour vous dire que j'ai trouvé une solution euh à votre problème euh en fait vous allez pouvoir assister aux deux cours parce que les horaires ont été décalés donc euh si ça vous intéresse toujours et que vraiment vous arrivez pas à vous décider voilà y a y a cette possibilité donc tenez moi tenez moi au courant de votre choix et puis voilà à bientôt au revoir

MAUVAISE NOUVELLE

25 : Euh bonjour madame Monat c'est euh le docteur Camille Dumont euh voilà je vous appelais par rapport aux résultats de de votre dernier examen bon euh bon les analyses sont pas très très bonnes donc euh rien de trop inquiétant non plus mais euh quand même j'aimerais bien vous voir euh assez rapidement euh est-ce que est-ce qu'un rdv euh la semaine prochaine euh mettons euh lundi euh dans la en fin de journée ou euh quand vous voulez vous conviendrait ? euh voilà c rappelez-moi bon surtout vous inquiétez pas trop ce c'est pas non plus la catastrophe mais euh faudrait quand même qu'on en discute voilà rappelez-moi au revoir

26 : Bonjour madame Monat c'est Camille Dumont le chez d'escalier j'vous appelais par rapport à v à votre vol qui a été annulé euh voilà je suis vraiment désolée on a fait tout ce

qu'on pouvait mais on a pas trouvé de solution y a aucun y a aucun vol euh qui part aujourd'hui donc euh je vous ai trouvé un vol en revanche pour euh demain matin euh eh tôt à partir de 8h donc j'espère que ça vous irait donc vous pouvez me rappeler du coup euh et on v on vous le réservera si ça vous convient à euh merci à très vite au revoir

27 : Bonjour madame Monat euh c'est Camille Dumont la gestionnaire de la Niblanc j'vous appelais par rapport au petit Gabriel et à sa place en crèche euh je suis vraiment désolée mais euh il n'est pas euh accepté on été vraiment euh complets euh donc voilà je s je sais que c'est compliqué euh euh c'que j'peux euh v si vous l'souhaitez euh je peux vous tenir au courant si jamais y a des des désistements euh des fois ça nous arrive au 2ème semestre voilà j'espère que vous allez pouvoir trouver une autre solution euh voilà merci au revoir

28 : Oui bonjour Amélie euh c'est Camille euh j't'appelais par rapport à ta demande d'augmentation voilà j'aurais bien aimé en parler de vive voix mais j'arrive pas à te joindre euh du coup euh ben on a réfléchi mais on va pas pouvoir l'accepter euh j'espère que euh voilà fin t'as déjà été augmentée l'année dernière et euh on peut pas augmenter la euh tous les ans là comme ça donc euh beh si tu veux on peut se faire un rdv pour en parler la semaine prochaine euh voilà pour que tu comprennes un petit peu les raisons en tout cas c'est pas par rapport au travail effectué bien sûr euh voilà tiens moi au courant au revoir

29 : Euh allo Amélie c'est Camille euh je t'appelais par rapport à la commande des outils de nettoyage donc ils sont bien arrivés mais euh manque de pot il en manque la moitié voilà c'est c'est un petit peu embêtant surtout que je sais que t'avais prévu pas mal de choses cette semaine euh ils doivent nous dire euh quand ça va arriver le reste mais euh pour l'instant on n'en sait pas plus mais euh voilà je suis désolée c'est pas de chance voilà à très vite merci

30 : Oui allo madame Monat euh c'est euh le vétérinaire voilà je vous appelais euh par rapport à votre chat donc on a on a trouvé ce qu'il avait euh par contre euh les soins sont assez chers donc si vous voulez vous pouvez repasser euh demain euh on regardera ce qu'il est possible de faire euh voilà à très vite

31 : Euh allo madame Monat donc c'est Camille Dumont la professeur de de Jeremy euh j'vous appelais pour vous dire que finalement l'école n'a pas pu débloquer les fonds que qui étaient annoncés pour le voyage ça va nous couter un petit peu plus cher euh j'vous ai j'ai renvoyé à tous les parents d'élèves euh pour vous dire ça et comme je je sais que c'était un tout petit peu plus compliqué pour vous euh tenez moi tenez moi au courant et euh dites-nous si vous voulez que Jérémie participe ou pas du coup voilà jsuis désolée au revoir

32 : Oui euh bonjour euh bonjour Amélie c'est euh Camille Dumont le commissaire-priseur euh j'vous appelais euh pour vous parler du vase euh de en porcelaine chinoise que vous avez déposé la semaine dernière je suis au regret de vous annoncer que c'est une c'est une copie euh une copie récente euh donc euh très bien faite hein ce qui explique euh la la confusion mais je pense que vous avez été victime de fraude malheureusement donc bon il a quand même une certaine valeur euh pour les ventes aux enchères ça ces vases-là partent parfois quand même très bien pour la dans des sur des marchés comme la chine mais mais heu c'est sûr que ce sera moins haut que ce que vous espériez donc euh on se tient au courant au revoir

33 : Oui allo Amélie c'est Camille euh je t'appelle par rapport à ta demande de télétravail du vendredi euh je suis désolée mais en fait il y a déjà trop de membres de l'équipe euh qui sont pas à ce jour-là il faut vraiment qu'on ait quelqu'un pour assurer le relais euh donc pour le moment on va te demander de de revenir euh on pourra reparler de tout ça euh peut être l'an prochain si besoin mais pour le moment on va rester sur sur euh l'organisation qui est déjà mise en place voilà donc il faudrait que tu prennes un autre jour si tu veux quand même prendre une journée de télétravail à très vite au revoir

34 : Euh bonjour madame Monat c'est Camille Dumont euh je vous appelle euh parce que la date du procès approche et donc j'ai pris connaissance de nouvelles informations dans vot dossier euh qui sont assez incriminantes donc euh ça va peut-être changer un peu notre stratégie donc euh si vous voulez on peut prendre rendez-vous euh d'ici la fin de semaine euh pour voir un petit peu comment on va aborder ça donc euh voilà à très vite merci

35 : Euh bonjour Amélie c'était Camille euh j't'appelais parce que euh on a une mauvaise euh nouvelle euh notre chi le chiffre d'affaire a assez baissé euh pour ce semestre euh voilà c'est pt'être dû euh au changement euh qu'on a fait sur les rayons c'pas c'pas sûr à 100% mais on se dit que peut-être c'était pas le meilleur choix stratégique là pour le rayon fantasy en tout cas je voilà j'ai le regret de t'annoncer que :: j'vais pas pouvoir euh t'employer à temps plein euh à compter de septembre euh je sais qu'ça doit être assez :: difficile comme nouvelle mais là euh j'ai vraiment pas le choix euh fin bon ça va pas être possible fin bon si tu veux on peut en parler demain moi je serai sur place euh voilà pour voir comment euh tu veux t'organiser savoir si tu veux rester quand même ou pas voilà hh bon bonne journée quand même et puis à demain euh au revoir

36 : Bonjour Amélie euh c'est Camille Dumont euh je vous appelais euh pour vous dire que malheureusement euh les deux cours entre lesquels vous hésitez sont pas de votre niveau euh c'est réservé euh en master euh donc vous allez pas pouvoir y assister donc euh voilà l'administration m'a dit que c'est pas possible ce que je vous propose de faire c'est que si vraiment vous avez envie euh d'en d'avoir euh d fin de de d'acquérir euh des connaissances sur ces cours là on peut demander à des élèves de master de vous passer leur notes mais euh voilà c'est le maximum de ce que je peux faire au revoir

Locuteur 2

NOUVELLE NEUTRE

1 : Oui bonjour Madame Monat le docteur Dumont à l'appareil je vous appelle pour vous informer qu'aujourd'hui nous aurons ren euh RENDEZ-VOUS en face de mon cabinet normal puisqu'il est en rénovation donc voilà je vous remercie de bien vouloir euh écouter ce message avant de venir me voir à tout à l'heure bonne journée

2 : Oui bonjour Madame Monat ici le chef d'escalle Camille Dumont euh je vous appelle pour vous informer que votre vol a été déplacé du terminal 1 au terminal 4 voilà je vous remercie à tout à l'heure

3 : bonjour madame Monat ici Camille Dumont de la garderie euh la Niblanc euh voilà je suis au courant que votre enfant est sur liste d'attente nous allons faire au plus vite mais malheureusement pour l'instant nous n'avons pas de possibilité de libération de de place

pour votre enfant donc euh la la décision sera prise euh au plus tôt euh j'essaierai de vous inf euh de vous en informer très rapidement je vous remercie de votre compréhension bonne journée

4 : Bonjour madame Monat ici Camille Dumont des ressources humaines je viens vous faire part de donc nous avons bien reçu votre dossier euh et votre demande d'augmentation nous allons l'évaluer au plus vite et je vous tiendrai au courant le plus rapidement possible je vous remercie pour votre demande bonne journée

5 : Oui bonjour Amélie c'est Camille de la Niblanc je voulais simplement vous informer que nous avons donc passé commande euh des outils pour le nettoyage euh que vous allez pouvoir utiliser ils arriveront d'ici une semaine voilà bonne journée

6 : Bonjour Madame Monat ici le vétérinaire Dumont je vous appelle en euh au sujet de votre chat donc que vous avez déposé ce matin euh malheureusement donc nous avons un petit retard de 30 minutes si ça ne vous embête pas de venir le chercher plutôt euh 30 minutes plus tard voilà je vous remercie à tout à l'heure

7 : Bonjour madame Monat ici le professeur Dumont euh je vous appelle au sujet le v du voyage scolaire euh simplement pour vous dire que le point de rassemblement n'est plus euh au point préalablement établi mais il est plutôt devant l'école maintenant donc voilà JE VOULAIS simplement QUE VOUS Ayez le message avant d'amener votre enfant tout à l'heure je vous remercie à tout à l'heure

8 : Bonjour madame Monat ici la commissaire-priseur euh Dumont je vous appelle au sujet de l'expertise de votre vase en porcelaine euh voilà on vient de la finir donc c'était simplement pour vous informer que vous pouvez passer le récupérer voilà bonne journée

9 : Bonjour Amélie euh ici Camille Camille Dumont euh je v vous tiens simplement informée du fait que j'ai euh fait la demande auprès de notre supérieur pour votre demande en télétravail pour le vendredi voilà j'attends encore ça réponse mais je vous tiens informée le plus vite possible merci beaucoup bonne journée

10 : Bonjour madame Monat ici maitre Dumont je vous appelle pour vous innformer que l'audience aura lieu en salle 3 et non en salle 2 mercredi prochain voilà donc j'espère que vous aurez le message à temps merci beaucoup bonne journée

11 : Bonjour Amélie c c'est Monat de la Niblanc je vous appelle simplement pour vous dire qu'on a fait qua quelques changements de dernière minute euh nous avons déplacé les livres fantasy euh du rayon jeune adulte euh à côté du rayon de science-fiction voilà donc je simplement pour vous dire ne vous inquiétez pas la prochaine fois que vous arrivez à la librairie les ray les livres fantasy sont à COTE des science-fictions voilà bonne journée au revoir

12 : bonjour mademoiselle Monat veu euh je suis la directrice du département de la faculté donc j'ai reçu votre message euh JE VOUS PROpose un rdv le mardi 29 à 11h si vous êtes disponible pour que nous puissions discuter euh des deux cours qui vous intéressent qui sont au même horaire voilà je vous serai euh je vous remercierai je re je vous remercie de me renvoyer un message rapidement pour me confirmer si l'heure et la date vous conviennent voilà je vous remercie bonne journée au revoir

BONNE NOUVELLE

13 : Bonjour madame Monat c'est le médecin Dumont à l'appareil voilà je vous appelle pour vous informer que j'ai reçu vos analyses qui sont bien meilleures que la dernière fois donc je j'en suis vraiment ravie pour vous euh voilà j'espère que nous nous reverrons très vite pour un prochain rdv d'ici là portez-vous bien bonne journée au revoir

14 : Bonjour madame Monat je suis la chef d'escale Camille Dumont je vous appelle pour vous informer que nous avons la possibilité de vous relocaliser en première place à la place de votre siège en classe économique euh ceci pour le même prix donc voilà c'estu c'est une très très bonne offre donc si jamais vous êtes intéressée merci de me rappeler très rapidement je vous remercie bonne journée au revoir

15 : Bonjour madame Monat c'est la gestionnaire Camille Dumont à l'appareil j'ai le plaisir de vous appeler pour vous annoncer que votre enfant a été accepté à la Niblanc à la garderie de la Niblanc donc voilà pour cette nouvelle année euh voilà j'en suis absolument ravie j'espère qu'il va bien s'entendre avec les autres enfants voilà je vous remercierai de bien vouloir me rappeler pour me confirmer que vous vous le vous voulez toujours le mettre à notre garderie d'ici là portez-vous bien au revoir

16 : Bonjour madame Monat c'est la directrice Camille Dumont des ressources humaines je vous appelle pour vous informer qu que nous avons bien réfléchi et nous sommes prêts à vous augmenter voilà donc je suis ravie euh je veux simplement vous demander de me rappeler pour que nous puissions en discuter discuter d des modalités mais voilà d'ici là je vous souhaite une bonne journée à bientôt

17 : Bonjour Amélie c'est Camille de la du camping de la Niblanc voilà euh je vous appelle pour vous informer que les outils que nous avons commandés pour le nettoyage sont performants très performants ils nous permettront de gagner du temps euh sur notre temps de travail de nettoyage du camping voilà donc j'en suis absolument ravie euh voilà je voulais simplement t'partager la bonne nouvelle avec vous et je vous souhaite une bonne journée et à très bientôt au revoir

18 : Bonjour madame Monat c'est le vétérinaire Dumont à l'appareil je vous appelle au sujet de votre chat que vous nous avez laissé la semaine dernière j'ai t'une bonne nouvelle il est bien soigné et bien bien guéri donc il n'a plus il n'aura plus aucu aucune séquelle il n'aura aucune séquelle donc voilà je je vous appelle pour vous proposer de venir le chercher dès ce soir si vous êtes disponible et euh voilà merci de me rappeler assez rapidement bonne journée

19 : Bonjour madame Monat c'est la professeure d'histoire geo Camille Dumont je vous appelle au sujet de voyage scolaire et j'ai une bonne nouvelle puisque l'école a pu débloquer des fonds et va pouvoir financer une grande partie du voyage voilà donc euh j'espère que euh j'espère que vous aurez la nouvelle assez rapidement euh d'ici là bon weekend et euh à lundi j'espère au revoir

20 : Bonjour madame Monat ici la commissaire-priseur Camille Dumont de la Niblanc jvous appelle au sujet de l'expertise de votre vase en porcelaine chinoise alors nous avons une bonne nouvelle puisqu'il s'agit ici d'uneuh d'un d'un vase qui a une grande valeur et qui a même une plus grande valeur que ce que vous pensiez donc voilà je vous propose que vous me rappeliez assez rapidement pour que nous puissions en discuter de vive voix voilà bonne journée au revoir

21 : Bonjour euh bonjour Amélie c'est la directrice Amélie Dumont à l'appareil je vous appelle au sujet de votre demande pour passer en télétravail le vendredi alors j'ai une bonne nouvelle puisque ce ce ns'ra aucun ça n'pose aucun problème à la boîte et donc vous allez pouvoir v euh télétravailler le vendredi voilà euh j'en suis ravie euh je vous souhaite un bon weekend et à très vite au revoir

22 : Bonjour madame Monat ici maître Dumont à l'appareil j'ai une bonne nouvelle puisque donc j'ai pris connaissance de nouvelles informations sur le dossier concernant votre procès et elles sont en faveur de votre innocence donc euh voilà nous allons essayer de nous voir au plus vite pour pouvoir euh monter un dossier en béton je vous remercie de bien vouloir me rappeler très très rapidement et voilà pour qu'on puisse euh fixer un rdv bonne journée au revoir

23 : bonjour Amélie c'est Camille de La Niblanc voilà j'ai une bonne nouvelle à vous annoncer puisque que bah grâce aux changements que nous avons opérés dans les rayons récemment nous avons pu euh augmenter notre chiffre d'affaires et donc euh bonne nouvelle pour vous et pour moi puisque nous allons pouvoir vous augmenter vous allez pouvoir toucher une prime ce mois-ci euh voilà je voulais simplement vous en informer hm par téléphone mais nous en reparlerons lundi quand nous nous verrons voilà bon weekend au revoir

24 : Bonjour mademoiselle Monat Monat c'est la directrice du département Camille Dumont à l'appareil voilà je vous appelle au sujet de votre mail par rapport aux deux cours qui vous intéressaient mais qui étaient à la même au même horaire voilà j'ai trouvé une solution pour vous permettre d'assister aux deux cours donc voilà euh je vous remercie de bien vouloir me rappeler très rapidement pour que je puisse vous expliquer en quoi cela consiste voilà bonne journée bonne fin de journée au revoir

MAUVAISE NOUVELLE

25 : Bonjour madame Monat ici le docteur Dumont à l'appareil voilà je me permet de vous joindre puisque nous avons reçu vos analyses et malheureusement elles sont plus mauvaises que la dernière fois donc voilà j'aimerais bien en discuter directement avec vous puisque c'est un petit peu difficile à faire euh sur un message vocal donc euh voilà je vous remercie de bien vouloir me rappeler le plus rapidement possible pour que nous puissions fixer un RENDEZ-VOUS très rapidement merci bonne journée au revoir

26 : Bonjour madame Monat ici le chef d'escale Camille Dumont euh voilà je suis désolée je suis au regret de vous annoncer que votre avion n'a été annulé votre vol a été annulé pour ce soir et malheureusement nous n'avons pas trouvé d'autre solution euh voilà donc nous allons vous réserver une chambre d'hôtel mais donc vous ne pourrez pas prendre l'avion avant demain matin voilà j'en suis vraiment navrée et j'espère que voilà vous comprendrez notre impossibilité de vous faire voyager ce soir merci beaucoup et bonne fin de journée au revoir

27 : Bonjour madame Monat ici Camille Dumont de la garderie de la Niblanc euh voilà je vous appelle pour vous informer que votre enfant malheureusement ne pourra pas être accepté cette année puisqu'en effet il y avait énormément d'enfants sur la liste d'attente et malheureusement tous ont été euh affectés à notre garderie donc euh voilà je suis au regret de vous annoncer que nous ne pourrions pas prendre votre enfant j'espère que vous

aurez rapidement une solution pour vous retourner je vous souhaite une bonne journée une bonne semaine au revoir

28 : Bonjour madame Monat bonjour Amélie c'est Camille Camille Dumont des ressources humaines euh voilà suite à votre demande d'augmentation nous avons bien réfléchi et malheureusement de fait que nous n'avons pas beaucoup de fonds aux ressources humaines je suis désolée de vous annoncer que nous ne pourrons pas vous augmenter euh en tout cas pas ce mois-ci voilà je je suis je suis vraiment navrée euh je j'espère que nous pourrons rapidement trouver une solution le mois prochain voilà bonne journée au revoir

29 : Bonjour Amélie ici la gé la gérante du camping Camille Camille Dumont euh voilà je vous appelle au sujet de LA COMMANDE que nous avons fait pour les out les outils de nettoyage des emplacements euh voilà elle est arrivée mais il manque la moitié des outils voilà malheureusement donc vous ne pourrez pas commencer euh la la vous ne pourrez pas commencer le nettoyage cette semaine vous allez devoir attendre la semaine prochaine voilà j'en suis navrée euh b bonne journée euh à bientôt

30 : Bonjour madame Monat ici le vétérinaire Camille Dumont euh je je vous appelle au sujet de votre chat que vous nous avez déposé la semaine dernière euh alors il existe un traitement pour la maladie qu'il a malheureusement il est très très cher donc je voulais en discuter directement avec vous voir ce que vous en pensiez voilà je préférerais en discuter directement de vive voix avec vous donc si nous pouvions fixer in rdv ce serait plus facile à à voilà à dire je vous remercie vous souhaite une bonne fin de journée au revoir

31 : Bonjour madame Dumont ici le professeur d'histoire geo Camille Dumont euh je vous appelle au sujet du voyage scolaire voilà nous avons malh nous n'avons malheureu malheureusement pas pu débloquer les fonds nécessaires pour le voyage euh qui en effet coute plus cher que prévu vu les circonstances actuelles voilà donc je voulais voir avec vous ré directement si vous comptez toujours peut-être essayer de trouver une solution pour que Morgan continue et vienne avec nous au voyage scolaire ou si vous pensez plutôt le retirer de la liste euh voilà je vous remercie bonne fin de journée au revoir

32 : Bonjour madame Monat ici le commissaire-priseur Camille Dumont je vous appelle au sujet de l'expertise de votre vase en porcelaine chinoise euh voilà je suis au regret de vous dire que vous avez été victime d'une fraude car euh ce vase euh n'est qu'une copie récente euh qui a évidemment beaucoup moins de valeur que l'original donc euh voilà malheureusement nous ne pourrons pas expertiser ce cette euh cet objet qui n'a aucune valeur pour nous voilà donc j'en suis absolument navrée pour vous parce que ça reste une belle pièce mais voilà je j'espère que vous aurez le message assez rapidement et que vous pourrez me rappeler à bientôt au revoir

33 : Bonjour Amélie ici Camille Dumont euh à l'appareil je vous appelle au sujet de votre demande pour passer en télétravail le vendredi euh voilà donc euh malheureusement euh vous ne pourrez pas passer en télétravail puisque monique est déjà en télétravail ce jour-là donc nous avons besoin de vous et je suis vraiment navrée de devoir vous faire venir le vendredi même si vous habitez loin et je sais ô combien vous préféreriez bosser depuis chez vous mais malheureusement nous n'avons pas d'autre solution pour le moment voilà

j'en suis navrée je j'espère que vous ne m'en tiendrez pas trop rigueur bonne journée au revoir

34 : Bonjour madame Monat ici l'maître ici le maitre Dumont à l'appareil euh jvous appelle pour euh vous informer que j'ai pris connaissance de nouvelles informations concernant votre dossier et malheureusement pour vous elles sont plutôt incriminantes euh donc je voudrais en parler directement avec vous euh voilà donc je vous remercie de bien vouloir me rappeler pour que nous puissions fixer un rdv très rapidement pour essayer de trouver une solution je vous remercie bonne journée au revoir

35 : Bonjour Amélie ici Camille de la garderie la Niblanc je vous appelle euh voilà au sujet de :: NOTRE CHIFFRE d'affaire puisque nous avons comme vous le savez euh modifié les rayons quelque peu récemment et euh voilà nous n'avons malheureusement pas pu augmenter notre chiffre d'affaire et nous n'allons pas pouvoir vous garder en temps plein à partir de septembre donvoilà j'en suis navrée nous v allons devoir vous passer à temps partiel euh si nous voulons pouvoir garder la librairie ouverte voilà je vous remercie je vous je suis vraiment désolée euh voilà je je je vous souhaite une bonne soirée au revoir

36 : Bonjour madame Monat ici la directrice Camille Dumont du département de la faculté j'ai bien reçu votre mail euh voilà euh par rapport au rdv par rapport au deux cours qui vous intéressaient qui sont au même horaire je me suis renseignée eu vous ne pourrez en fait malheureusement assister à aucun des deux car ils ne sont pas à votre niveau de compétence euh voilà vous êtes en L2 et malheureusement ce ces ces cours-là sont pour des niveau L3 donc voilà j'en suis navrée j'espère que vous pourrez y assister l'année prochaine en tout cas n'hésitez pas à revenir vers moi à ce moment-là voilà bonne journée au revoir

Locuteur 3

NOUVELLE NEUTRE

1 : Oui bonjour Madame Monat ici Camille Dumont je vous appelle car mon cabinet étant en rénovation j'ai déménagé juste en face. Merci d'avance, au revoir

2 : Oui bonjour Madame Monat ici Monsieur Domont je vous appelle pour vous avertir car l'embarquement de votre vol a été déplacé au terminal 4 au revoir

3 : Bonjour madame Monat ici la garderie de la Niblanc je vous appelle pour vous informer que nous avons bien reçu votre dossier votre enfant est actuellement sur liste d'attente une dcision sera prise dans les prochains jours nous vous informerons dès que possible bonne journée et à bientôt

4 : Bonjour Amélie c'est Camille je t'appelle pour te dire que j'ai bien reçu ta demande d'augmentation je vais voir ce que je peux faire je te tiens au courant bonne journée et à bientôt

5 : Bonjour Amélie c'est Camille c'était pour t'avertir que le matériel de nettoyage pour le campung doit arriver dans la semaine merci d'avance bonne journée

6 : Bonjour Madame Monat c'est Camille Dumont du service vétérinaire je vous appelle pour vous informer que vous pourrez récupérer votre chat à 11h plutôt que 10h30 merci d'avance, bonne journée

7 : Bonjour madame Monat ici Monsieur Dumont je vous appelle car le rassemblement pour le voyage scolaire a changé il est maintenant devant l'école merci d'avance à bientôt

8 : Bonjour madame Monat ici monsieur Dumont de la Niblanc je vous appelle pour vous informer que nous avons terminé l'expertise de votre vase rappelez-moi dès que possible pour que nous puissions en discuter bonne journée

9 : Bonjour Amélie c'est Camille je vous appelle pour vous informer que j'ai fait remonter votre demande pour le télétravail je vous informe dès que je reçois une réponse bonne journée

10 : Bonjour madame Monat ici monsieur Dumont je vous appelle pour vous informer concernant le lieu du procès il se situera en salle trois et non plus en salle 2 merci à vous bonne journée et à bientôt

11 : Bonjour Amélie c'est Camille je t'appelle pour te prévenir que nous avons effectué un changement aujourd'hui nous avons mis les livres de fantasy à côté du rayon jeunes adultes dorénavant je préférerais te prévenir pour que tu ne sois pas surprise bonne journée

12 : Bonjour Amélie c'est monsieur Dumont je t'appelle suite à ton hésitation entre les deux cours si tu veux nous pouvons en discuter demain midi après mon cours rappelle moi pour me faire part de ta réponse bonne journée

BONNE NOUVELLE

13 : Bonjour madame Monat ici monsieur Dumont je vous appelle suite à vos dernières analyses qui sont meilleures que les précédentes nous pourrions en discuter lors de notre prochaine entrevue rappelez-moi dès que possible bonne journée

14 : Bonjour madame Monat ici monsieur Dumont je vous appelle concernant votre vol vous aviez un siège en classe économique mais euh si vous le souhaitez-vous pouvez avoir un siège en première classe pour le même prix rappelez-moi dès que possible si vous souhaitez ce surclassement bonne journée

15 : Bonjour madame Monat ici la garderie de la Niblanc je vous appelle pour vous informer que votre enfant a été accepté dans notre établissement bonne journée

16 : Bonjour Amélie c'est Camille je t'appelle suite à notre dernière conversation après réflexion je pense effectivement que tu mérites cette augmentation nous en parlerons plus en détail dès que nous nous revoyons bonne journée et à bientôt

17 : Bonjour Amélie c'est Camille je t'appelle pour t'informer que nous recevrons bientôt de nouveaux outils de nettoyage qui seront plus pratiques que les anciens bonne journée à bientôt

18 : Bonjour madame Monat ici monsieur Dumont je vous appelle pour vous informer que votre chat a bien guéri vous pouvez désormais passer le récupérer bonne journée à bientôt

19 : Bonjour madame Monat ici monsieur Dumont je vous appelle concernant le voyage scolaire nous avons réussi à débloquer plus de fonds pour le financement du voyage nous allons donc pouvoir revoir à la baisse les participations des parents d'élèves bonne journée et à bientôt

20 : Bonjour madame Monat ici monsieur Dumont de la Niblanc je vous appelle concernant votre vase je voulais vous informer qu'il a plus de valeur que ce que vous

pensiez nous pourrions (en) discuter de vive voix rappelez-moi dès que possible bonne journée

21 : Bonjour Amélie c'est Camille je voulais te dire qu'il n'y a aucun souci à faire du télétravail le vendredi nous pourrions en discuter dès que nous nous verrons bonne journée

22 : Bonjour madame Monat ici monsieur Dumont je vous appelle pur vous informer que suite à de nouveaux éléments dans votre dossier le procès dans les prochains jours devrait bien se dérouler, tout porte à croire que vous êtes innocent bonne journée à bientôt

23 : Bonjour Amélie c'est Camille je t'appelle pour t'informer que suite à nos bénéfices ces derniers mois avec ton idée de modifications de rayons nous allons pouvoir t'augmenter rappelle moi dès que possible pour que nous en discussions bonne journée

24 : Bonjour Amélie c'est monsieur Dumont je voulais t'informer que suite à notre dernière discussion j'ai pu trouver une solution pour les cours en effet l'un des deux cours contient un second groupe auquel tu pourrais te greffer rappelle moi dès que possible pour que nous puissions en discuter à bientôt

MAUVAISE NOUVELLE

25 : Bonjour madame Monat ici le docteur Dumont je vous appelle suite à vos analyses qui ne sont pas très bonnes nous pourrions en discuter de vive voix rappelez-moi dès que possible pour prendre rendez-vous sur ce, bonne journée

26 : Bonjour madame Monat ici monsieur Dumont j'ai le regret de vous annoncer que votre vol a été annulé la seule solution que j'ai pu trouver pour le moment est de décaler le vol au lendemain appelez-moi le plus rapidement possible pour que nous puissions en discuter à bientôt

27 : Bonjour madame Monat ici monsieur Dumont dla garderie de la Niblanc j'ai lregret de vous annoncer que notre garderie est complète cette année nous n'allons malheureusement pas pouvoir garder votre enfant bien à vous bonne journée

28 : Bonjour Amélie c'est Camille nous n'allons malheureusement pas pouvoir vous augmenter cette année l'entreprise n'ayant pas fait les bénéfices escomptés mais je prends en compte votre demande pour une prochaine fois bonne journée

29 : Bonjour Amélie ici Camille je voulais t'avertir concernant les outils qu'on nous avons reçus qu'il en manque la moitié on ne pourra donc pas les utiliser pour le moment il va falloir faire avec en attendant de recevoir le reste malheureusement bonne journée

30 : Bonjour madame Monat ici docteur Dumont du cabinet vétérinaire j'vous appelle concernant votre chat le traitement pour sa maladie est plus cher que prévu nous pourrions en discuter de vive voix rappelez-moi dès que possible bonne journée

31 : Bonjour madame Monat ici monsieur Dumont je vous appelle concernant le voyage scolaire nous n'avons pas pu débloquer autant de fonds que nous le pensions il faudra donc revoir à la hausse la participation des parents d'élèves pour ce voyage rappelez-moi dès que possible pour que nous puissions en discuter à bientôt

32 : Bonjour madame Monat ici monsieur Dumont de la Niblanc je vous appelle suite à votre vase il s'agit en effet malheureusement d'un faux d'une copie récente et non pas d'un vase de porcelaine chinoise bonne journée et à bientôt

33 : Bonjour Amélie c'est Camille je t'appelle pour te dire que Suzane est déjà en télétravail le vendredi nous ne pourrions pas nous permettre d'avoir deux personnes absentes ce jour-là malheureusement rappelle moi dès que possible pour qu'on puisse en discuter à bientôt

34 : Bonjour madame Monat ici monsieur Dumont de nouveaux éléments dans le dossier de l'affaire viennent d'arriver ils ne sont pas très bons nous pourrions en discuter de vive voix rappelez-moi dès que possible pour que nous fixions un RENDEZ-VOUS bonne journée à bientôt

35 : Bonjour Amélie c'est Camille de la librairie c'était pour te dire que suite à notre baisse de chiffres d'affaires à partir de septembre nous risquons de devoir t'employer qu'à temps partiel nous pourrions en discuter rappelle moi dès que possible à bientôt

36 : Bonjour Amélie c'est monsieur Dumont je me suis renseigné concernant les deux cours es-tu sûre d'avoir le niveau pour les deux il s'agit de cours d'un niveau supérieur à celui que tu es censée avoir rappelle moi dès que possible pour que nous en discutions à bientôt

Locuteur 4

NOUVELLE NEUTRE

1 : Bonjour c'est votre médecin Camille Dumont je vous appelle pour vous dire que nous avons installé notre cabinet en face de là où on était d'habitude car l'autre est en rénovation donc je vous attends en face euh bonne journée au revoir

2 : C'est Camille Dumont votre chef d'escale c'était pour vous informer que votre vol a été déplacé du terminal 1 au terminal 4 donc euh veuillez prendre en considération cette information euh j'ai bonne euh journée au revoir

3 : Bonjour c'est Camille Dumont la gestionnaire de la garderie donc c'était pour vous informer que la décision sera prise au plus tôt pour vous dire si on peut prendre votre enfant ou pas à la rentrée donc euh voilà je vous laisse et j'vous dis bonne journée

4 : Camille Dumont je vous appelle car j'ai bien pris en compte votre demande d'augmentation est on vous donnera une réponse dans les plus brefs délais bonne journée au revoir

5 : Bonjour c'est Camille Dumont je vous appelle pour vous dire que la commande que j'ai effectuée pour vous pour les outils de nettoyage va arriver dans la semaine dans une semaine pardon euh je vous dis bonne journée et à bientôt

6 : Bonjour c'est Camille Dumont euh la vétérinaire j'vous appelle parce qu'en fait euh le y a eu un changement d'emploi et j'aurais bien aimé que vous veniez chercher votre chat euh d 30 minutes plus tard s'il vous plait donc je vous dis à tout à l'heure merci beaucoup au revoir

7 : Bonjour c'est Camille Dumont euh la professeure de euh Kevin euh j'vous appelle car le point de rassemblement pour le voyage scolaire a changé il est maintenant devant l'école donc je vous dis à tout à l'heure et bonne journée

8 : Bonjour c'est Camille Dumont la commissaire-priseur j'vous appelle par rapport euh à l'expertise de votre vase en porcelaine chinoise donc euh et on l'a fini donc je vous ::S

APPELLE justement pour vous demander de venir nous voir dès que vous avez la possibilité voilà merci de me rappeler bonne journée merci au revoir

9 : Bonjour c'est Camille Dumont euh je vous appelle car j'ai bien pris en compte votre demande de télétravail pour le vendredi j'ai fait part de votre demande à au direct a euh mon supérieur à votre supérieur il vous informera de la réponse au plus tôt bonne journée au revoir

10 : Bonjour c'est Camille Dumont euh je vous appelle parce que j'ai pris connaissance d'un déplacement du lieu du procès donc il y aura lieu à la place 2 à la place de la salle 3 voilà donc euh je vous souhaite une bonne journée au revoir

11 : C'est Camille Dumont euh je vous appelle pour te dire que j'ai changé modifié les rayons donc j'ai mis des livres fantasy euh à côté des livres de sciences fiction à la place de les avoir dans le rayon euh jeune adulte donc voilà je TE DIS A Bientôt au revoir

12 : Bonjour c'est Camille Dumont je vous appelle parce qu'en fait j'aurais souhaité qu'on prenne qu'on fixe un RENDEZ-VOUS pour parler des deux cours que vous hésitez donc veuillez me rappeler euh s'il vous plait euh à tout à l'heure au revoir

BONNE NOUVELLE

13 : C'est Camille Dumont votre médecin donc je vous appelle pour vous donner une bonne nouvelle euh vos analyses de la d sont meilleures que la dernière fois donc voilà on en reparlera plus mieux de vive voix bonne journée au revoir

14 : Bonjour c'est Camille Dumont euh votre chef d'escale je vous appelle parce qu'en fait euh vous pouvez bénéficier d'une classe première classe pour le même prix euh à la place de votre classe économique donc euh je vous tenez moi informée si cela vous intéresse bonne journée au revoir

15 : Bonjour c'est Camille Dumont euh la gestionnaire de la garderie euh je vous appelle parce qu'en fait euh votre enfant est accepté pour la rentrée bon voilà on en discutera plus une prochaine fois je vous dis bonne journée au revoir

16 : Bonjour c'est madame euh c'est Camille Dumont euh je vous appelle parce que suite à votre demande euh j'ai décidé de vous augmenter le salaire donc euh nous en discuterons bonne journée au revoir

17 : Bonjour c'est Camille Dumont euh je vous appelle parce qu'en fait nous avons passé une commande d'outils de nettoyage donc euh normalement ils sont beaucoup plus performants et de ce fait ils vont faire gagner beaucoup de temps donc voilà je vais vous laisser et je vous dis une bonne journée au revoir

18 : Bonjour c'est Camille Dumont euh la vétérinaire euh je vous appelle parce que votre chat va t va bien on l'a bien soigné et il est complètement guéri vous pouvez venir le rechercher dès que vous pouvez bonne journée au revoir

19 : Bonjour c'est Camille Dumont euh la professeure de Kevin euh je vous appelle car euh car l'école a pu débloquer des fonds donc euh on va pouvoir participer sur une plus grande partie euh du financement que prévu euh voilà je vous dis bonne journée et à bientôt

20 : Bonjour c'est Camille Dumont la commissaire-priseur je vous appelle car euhm la valeur de votre vase est plus importante que vous le pensiez donc euh nous en discuterons ultérieurement bonne journée au revoir

21 : Bonjour c'est Camille Dumont je vous appelleuh suite à votre demandeuh de télétravail pour le vendredi donc elle a été acceptée donc on se voit euh au travail bonne journée au revoir

22 : Bonjour c'est Camille Dumont votre avocate je vous appelle car j'ai pris connaissance de nouvelles informations qui sont euh en votre faveur je voudrais en rediscuter avec vous euh plus tard bonne journée au revoir

23 : Bonjour c'est Camille Dumont je vous appelle car euh suite à la modification des rayons euh nous avons eu une augmentation des chiffres d'affaire donc par consé euh par conséquent euh j'vais pouvoir vous DONNER UNEU prime euh nous en discuterons plus tard bonne journée au revoir

24 : Bonjour c'est Camille Dumont euh en fait j'ai regardé plus précisément votre cas et finalement j'ai trouvé une solution euh vous pouvez assister aux deux cours car ils ne sont pas en même temps bonne journée au revoir

MAUVAISE NOUVELLE

25 : Bonjour c'est Camille Dumont votre médecin euh je vous appelle car en fait euh les analyses que nous avons faites la dernière fois sont pas bonnes il faudrait en reparler le plus tôt possible bonne journée au revoir

26 : Bonjour c'est Camille Dumont votre chef d'escale euh je vous appelle car j'ai une mauvaise nouvelle à vous annoncer euh votre vol a été annulé et nous n'avons pas pu trouver de solution il y aurait un vol juste pour le lendemain veuillez me rappeler s'il vous plaît bonne journée au revoir

27 : Bonjour c'est Camille Dumont la gestionnaire de la garderie euh je vous appelle car euh votre demande pour la garderie de votre enfant a é n'a pas été acceptée je suis désolée nous n'avons pas pu trouver de solution euh VOUS V POUVEZ euh nous recontacter pour qu'on en reparle et voilà bonne journée au revoir

28 : Bonjour c'est Camille Dumont je vous appelle parce qu'en fait euh j'ai bien réfléchi à votre demande d'augmentation mais malheureusement je n'ai pas la possibilité de vous augmenter euh donc voilà je vous dis à euh bonne journée et à bientôt

29 : Bonjour c'est Camille Dumont euh je vous appelle car euh la commande que j'avais effectué la dernière fois est arrivée mais malheureusement il en manque la moitié donc euh je vous tiens au courant dès que l'autre moitié sera arrivée bonne journée au revoir

30 : Bonjour c'est Camille Dumont la vétérinaire je vous appelle car euh votre chat a une maladie dont le traitement coute euh très cher donc euh je vous APPELLE POUR voir qu'estce que vous vous souhaitez faire donc euh veuillez me rappeler le plus rapidement possible merci au revoir

31 : Bonjour c'est Camille Dumont [avaproferum (la professeure hum ?)] de Kevin je vous appelle car a l'école n'a pas pu malheureusement débloquer les fonds qu'ils avaient annoncé donc par conséquent le voyage va être plus cher donc euh veuillez me rappeler pour euh qu'on puisse parler de la situation et vois ce que vous voulez faire bonne journée au revoir

32 : Bonjour c'est Camille Dumont la commissaire-priseur je vous appelle pour euh vous annoncer une mauvaise nouvelle euh le vase que vous vouliez nous faire expertiser est en fait une copie récente donc vous avez été fraudé donc elle a beaucoup moins de valeur

donc euh si vous voulez plus d'informations vous pouvez me recontacter bonne journée au revoir

33 : Bonjour c'est Camille Dumont euh je vous appelle car euh je ne peux pas vous mettre en télétravail le vendredi car déjà une autre collègue y est déjà donc je suis désolée on pourra en discuter plus tard si vous voulez bonne journée au revoir

34 : Bonjour c'est Camille Dumont votre avocate je vous appelle car euh j'ai pris connaissance euh de nouvelles informations dans le dossier donc elles ne sont pas du tout en vot faveur il faudrait qu'on se voit pour en parler bonne journée au revoir

35 : Bonjour c'est Camille Dumont je vous appelle car suite à la baisse euh de chiffre d'affaires dont on est victimes euh je ne vais pas pouvoir vous garder à temps complet donc euh vous part euh vous devrez passer à temps partiel à compter de septembre j'suis vraiment désolée euh vous pouvez bien sûr me rappeler pour euh plus de renseignements bonne journée au revoir

36 : euh je vous appelle car en fait euh les deux cours que vous VOULIEZ PREndre ne sont en fait pas adaptés pour vous donc euh je peux vous conseiller des cours qui le seront bonne journée au revoir

Annexe 4. Scripts pour enregistrement du corpus – étude principale

ITEM	NEUTRE	POSITIF	NEGATIF
1	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre médecin. Je vous appelle au sujet de vos examens médicaux, voilà donc pour tout à l'heure mon cabinet étant en rénovation vous devrez venir au cabinet juste en face pour notre rendez- vous. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre médecin. Je vous appelle au sujet de vos examens médicaux, voilà donc vos analyses de la dernière fois sont bien meilleures qu'aux examens précédents et on pourra en discuter au prochain rendez- vous. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre médecin. Je vous appelle au sujet de vos examens médicaux, voilà donc vos analyses de la dernière fois sont moins bonnes qu'aux examens précédents et il faudrait que l'on en discute rapidement. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
2	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre chef d'escale. Je vous appelle au sujet de votre vol de ce soir, voilà donc l'embarquement a été déplacé du terminal 1 au terminal 4. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre chef d'escale. Je vous appelle au sujet de votre vol de ce soir, voilà donc nous vous proposons de changer votre place en classe économique pour une place en première classe, au même prix. Donc je me tiens à	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre chef d'escale. Je vous appelle au sujet de votre vol de ce soir, voilà donc il a été annulé et il m'est impossible de vous faire partir aujourd'hui sur un autre vol, nous allons vous réserver un hôtel. Donc je me tiens à

		disposition pour toute question, bonne journée à bientôt	disposition pour toute question, bonne journée à bientôt
3	Bonjour Madame Monat c'est Camille Dumont à l'appareil, gestionnaire de la garderie de la Niblane. Je vous appelle au sujet de votre enfant Léo, voilà donc nous pourrons bientôt vous dire s'il est accepté ou non pour cette année. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, gestionnaire de la garderie de la Niblane. Je vous appelle au sujet de votre enfant Léo, voilà donc il est accepté et nous l'accueillerons avec plaisir à partir de la rentrée prochaine. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, gestionnaire de la garderie de la Niblane. Je vous appelle au sujet de votre enfant Léo, voilà donc il est malheureusement refusé pour cette année, nous n'avons pas assez de place, il vous faudra trouver une autre solution. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
4	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction des ressources humaines. Je vous appelle au sujet de votre demande d'augmentation, voilà donc je l'ai bien reçue et je reviendrai vers vous après l'avoir étudiée. Donc je me tiens à disposition pour toute	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction des ressources humaines. Je vous appelle au sujet de votre demande d'augmentation, voilà donc je vais l'accepter, il faudrait que l'on se voie pour discuter du nouveau montant de votre salaire. Donc je	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction des ressources humaines. Je vous appelle au sujet de votre demande d'augmentation, voilà donc je vais devoir la refuser pour diverses raisons dont nous discuterons demain. Donc je me tiens à

question, bonne journée à bientôt	me tiens à disposition pour toute question, bonne journée à bientôt	disposition pour toute question, bonne journée à bientôt
--------------------------------------	---	--

5	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la gestion du camping. Je vous appelle au sujet de notre commande pour de nouveaux outils de nettoyage, voilà donc ils devraient arriver dans une semaine. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la gestion du camping. Je vous appelle au sujet de notre commande pour de nouveaux outils de nettoyage, voilà donc ils viennent d'arriver et sont visiblement très performants, vous aller gagner beaucoup de temps et d'énergie. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la gestion du camping. Je vous appelle au sujet de notre commande pour de nouveaux outils de nettoyage, voilà donc il en manque la moitié, ils sont arrivés tout à l'heure, il va falloir appeler le service après- vente. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
----------	---	--	--

6	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre vétérinaire. Je vous appelle au sujet de votre chat Minouche qui était malade, voilà donc il faudrait que vous veniez	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre vétérinaire. Je vous appelle au sujet de votre chat Minouche qui était malade, voilà donc il est désormais bien guéri et	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre vétérinaire. Je vous appelle au sujet de votre chat Minouche qui était malade, voilà donc il l'est toujours
----------	--	---	--

	le chercher plutôt vers 11h au lieu de 10h30 si possible. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	en pleine forme, vous pourrez venir le chercher dès demain. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	malheureusement et le traitement pour sa maladie coute très cher. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
7	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le professeur d'histoire-géo de Léo. Je vous appelle au sujet du voyage scolaire, voilà donc le point de rassemblement pour le voyage scolaire a changé, il est maintenant devant l'école. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le professeur d'histoire-géo de Léo. Je vous appelle au sujet du voyage scolaire, voilà donc l'école a pu débloquer des fonds et va pouvoir financer une grande partie du voyage. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le professeur d'histoire-géo de Léo. Je vous appelle au sujet du voyage scolaire, voilà donc l'école n'a pas pu débloquer les fonds que nous pensions pour financer le voyage et il faudra revoir à la hausse la participation des parents d'élèves. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
8	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le commissaire-priseur. Je vous appelle au sujet de votre vase en porcelaine chinoise, voilà donc on a fini de l'expertiser et vous pourrez passer dans la semaine pour le	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le commissaire-priseur. Je vous appelle au sujet de votre vase en porcelaine chinoise, voilà donc il a beaucoup plus de valeur que ce que l'on pensait et ce serait bien que l'on	Bonjour Madame Monat c'est Camille Dumont à l'appareil, le commissaire-priseur. Je vous appelle au sujet de votre vase en porcelaine chinoise, voilà donc vous avez été fraudé c'est en fait une copie du vase original qui a donc

récupérer et que l'on vous donne les résultats. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	en discute rapidement. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	beaucoup moins de valeur. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
--	---	---

9	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction. Je vous appelle au sujet de votre demande de télétravail, voilà donc j'ai bien fait remonter que vous voudriez le vendredi, je vous informe dès que je reçois une réponse. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction. Je vous appelle au sujet de votre demande de télétravail, voilà donc elle a été acceptée, vous pourrez télétravailler tous les vendredis à partir de la semaine prochaine. Donc je me tiens à disposition pour toute question, bonne journée à bientôt	Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction. Je vous appelle au sujet de votre demande de télétravail, voilà donc malheureusement le vendredi est déjà pris par une autre collègue donc il faudrait que vous choisissiez un autre jour. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
----------	--	--	--

10	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre avocat. Je vous appelle au sujet de votre procès à venir, voilà donc il est déplacé, il aura lieu en salle 3 et non en salle 2 comme annoncé au départ. Donc je me tiens à disposition pour toute	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre avocat. Je vous appelle au sujet de votre procès à venir, voilà donc il y a des informations en faveur de votre innocence qui ont été ajoutées au dossier, il faudrait que l'on en discute rapidement. Donc je me	Bonjour Madame Monat c'est Camille Dumont à l'appareil, votre avocat. Je vous appelle au sujet de votre procès à venir, voilà donc il y a des informations malheureusement incriminantes pour vous qui ont été ajoutées au dossier, il faudrait que l'on en discute
-----------	--	---	---

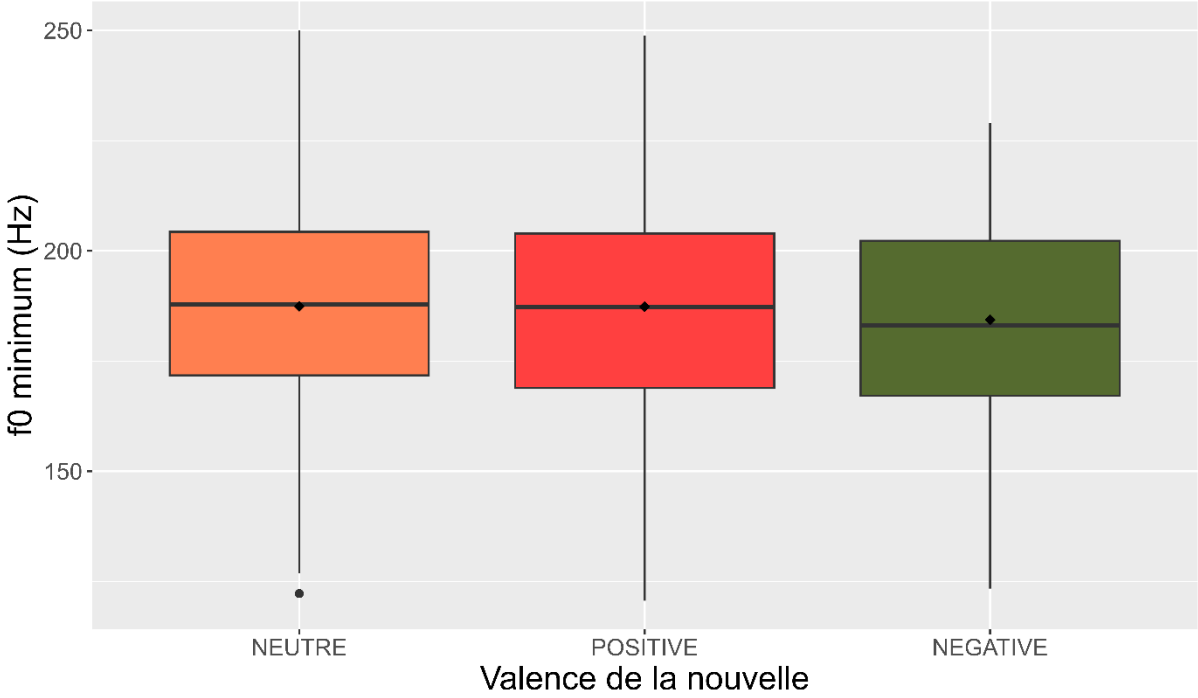
question, bonne journée à bientôt	tiens à disposition pour toute question, bonne journée à bientôt	rapidement. Donc je me tiens à disposition pour toute question, bonne journée à bientôt
--------------------------------------	--	--

11	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la librairie. Je vous appelle au sujet des changements opérés au niveau des rayons, voilà donc nous avons mis les livres de fantasy à côté du rayon jeunes adultes, je préférais vous prévenir pour que vous ne soyez pas surprise. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la librairie. Je vous appelle au sujet des changements opérés au niveau des rayons, voilà donc grâce à ça notre chiffre d'affaires a augmenté et ça veut dire que je vais pouvoir vous donner une belle prime pour Noël. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la librairie. Je vous appelle au sujet des changements opérés au niveau des rayons, voilà donc à cause de ça je ne vais pas pouvoir vous garder à temps complet à partir de septembre, parce que depuis qu'ils ont été faits notre chiffre d'affaires a beaucoup baissé et je ne peux plus me permettre de payer le salaire d'un temps complet. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>
-----------	---	--	---

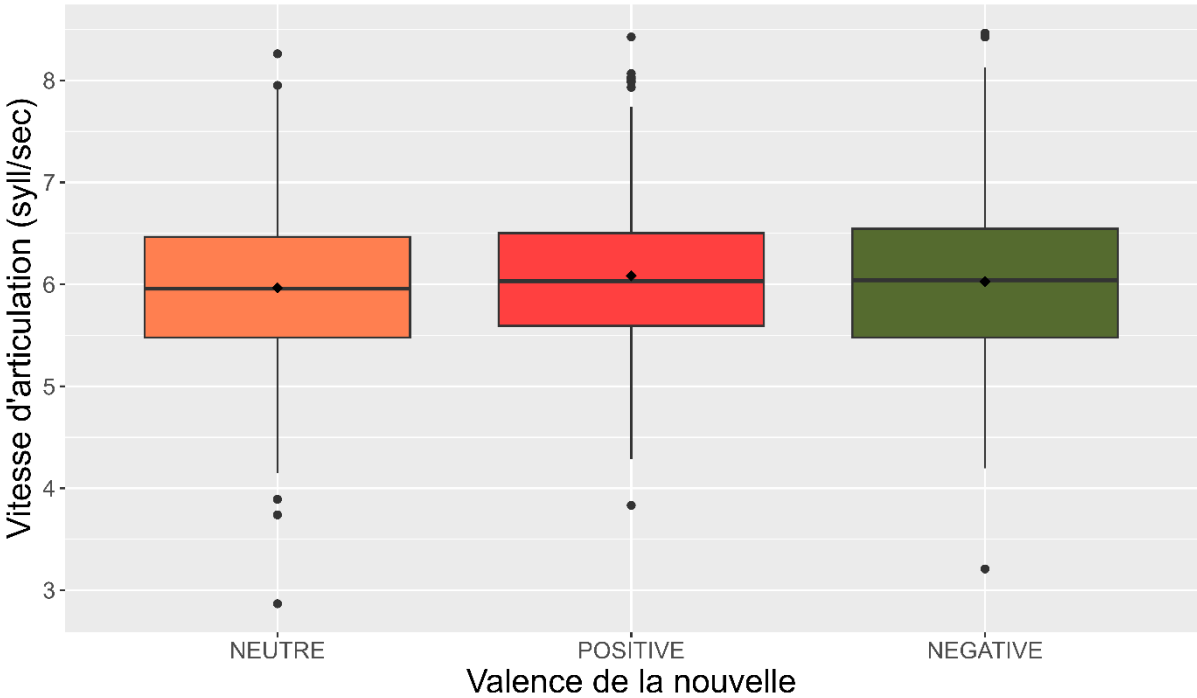
12	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction du département d'histoire. Je vous appelle au sujet de votre mail concernant votre hésitation entre deux cours, voilà donc nous pouvons en discuter demain à la fin de mon cours de 13h si ça vous convient. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction du département d'histoire. Je vous appelle au sujet de votre mail concernant votre hésitation entre deux cours, voilà donc vous allez en fait pouvoir assister aux deux cours si vous le souhaitez toujours car ils ne sont plus au même horaire. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>	<p>Bonjour Madame Monat c'est Camille Dumont à l'appareil, de la direction du département d'histoire. Je vous appelle au sujet de votre mail concernant votre hésitation entre deux cours, voilà donc les deux cours ne sont pas adaptés pour vous en fait, ils sont de niveau master, il faudrait que vous en choisissiez d'autres assez rapidement. Donc je me tiens à disposition pour toute question, bonne journée à bientôt</p>
-----------	---	--	---

Annexe 5. Graphiques des résultats non significatifs – étude principale

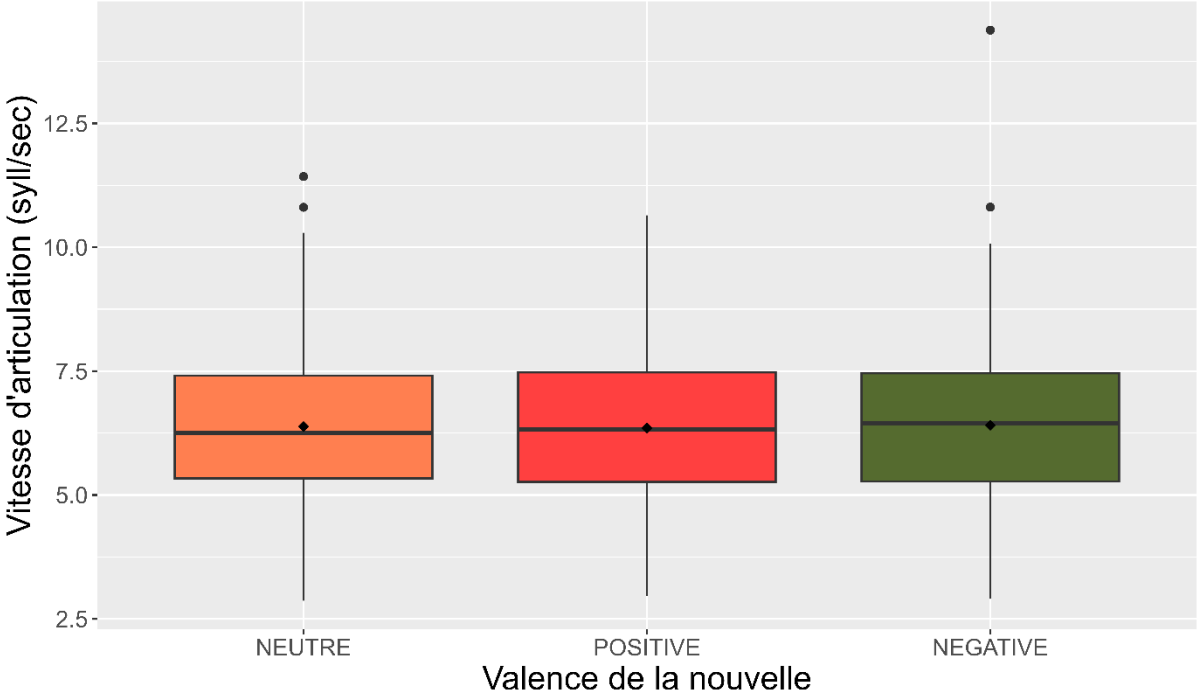
Graphique pour la f0 minimum (DMs).



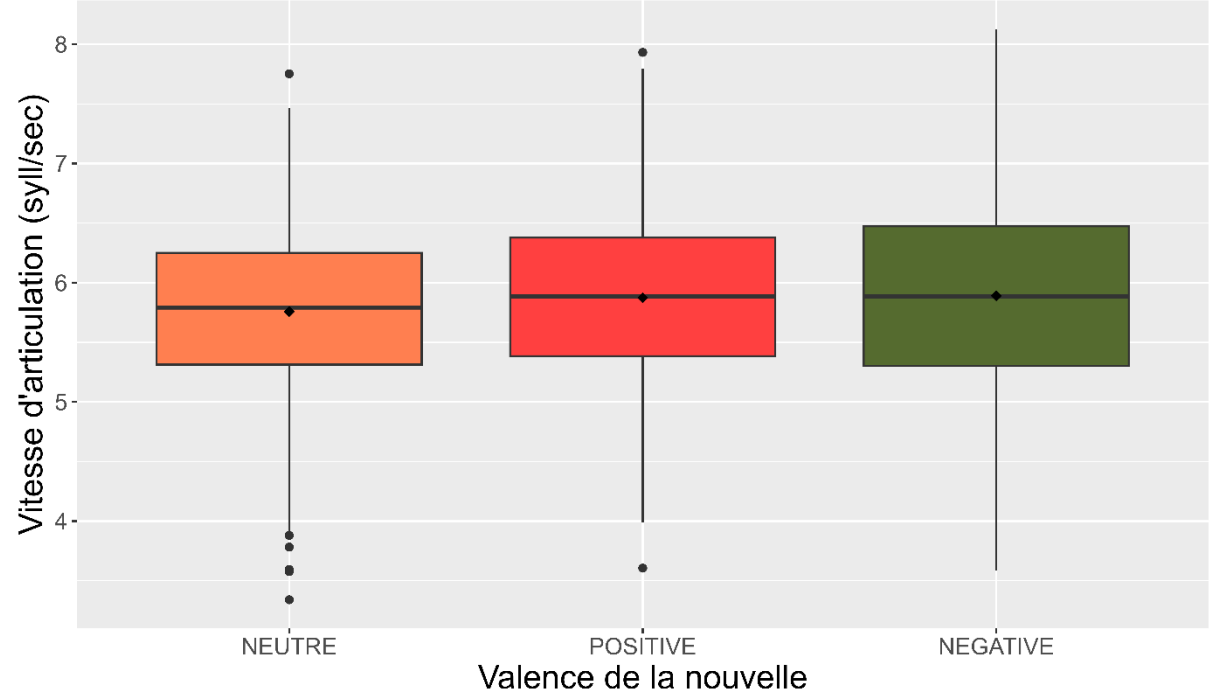
Graphique pour la vitesse d'articulation (préfaces).



Graphique pour la vitesse d'articulation (DMs).



Graphique pour la vitesse d'articulation (annonces).



Annexe 6. Résumés des modèles LMER réalisés sur les données de la f0

Préface - Empan vocal						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	4,25	0,04	96,28	40,10	< 0,0001
fixed	Valence NEUTRE	0,11	0,02	5,21	57,91	< 0,0001
fixed	Valence POSITIVE	0,16	0,02	7,53	57,58	< 0,0001
Préface (relevel) - Empan vocal						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	4,42	0,04	100,00	40,10	< 0,0001
fixed	Valence NEGATIVE	-0,16	0,02	-7,53	57,58	< 0,0001
fixed	Valence NEUTRE	-0,05	0,02	-2,31	57,91	0,0244
Préface - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	224,64	3,98	56,44	34,67	< 0,0001
fixed	Valence NEUTRE	6,02	1,59	3,79	57,99	0,0004
fixed	Valence POSITIVE	11,45	1,59	7,22	57,84	< 0,0001
Préface (relevel) - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,46	0,02	310,50	34,44	< 0,0001
fixed	Valence NEGATIVE	-0,05	0,01	-7,34	57,81	< 0,0001
fixed	Valence NEUTRE	-0,02	0,01	-3,40	57,97	0,0012
Préface - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	264,17	4,77	55,36	36,35	< 0,0001
fixed	Valence NEUTRE	9,92	2,07	4,80	57,94	< 0,0001
fixed	Valence POSITIVE	17,30	2,07	8,38	57,75	< 0,0001
Préface (relevel) - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,63	0,02	312,71	36,00	< 0,0001
fixed	Valence NEGATIVE	-0,06	0,01	-8,34	57,71	< 0,0001
fixed	Valence NEUTRE	-0,03	0,01	-3,37	57,90	0,0014
Préface - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,25	0,02	256,79	35,80	< 0,0001
fixed	Valence NEUTRE	0,01	0,01	1,72	58,08	0,0915
fixed	Valence POSITIVE	0,02	0,01	2,88	57,55	0,0056
Préface (relevel) - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,27	0,02	257,82	35,80	< 0,0001
fixed	Valence NEGATIVE	-0,02	0,01	-2,88	57,55	0,0056
fixed	Valence NEUTRE	-0,01	0,01	-1,16	58,02	0,2521
Marqueurs discursifs - Empan vocal						
effect	term	β	SE	t	df	p.value

fixed	(Intercept)	3,78	0,08	48,18	35,92	< 0,0001
fixed	Valence NEUTRE	0,19	0,04	4,21	58,11	0,0001
fixed	Valence POSITIVE	0,27	0,04	6,22	57,98	< 0,0001
Marqueurs discursifs (relevel) - Empan vocal						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	4,06	0,08	51,67	35,92	< 0,0001
fixed	Valence NEGATIVE	-0,27	0,04	-6,22	57,98	< 0,0001
fixed	Valence NEUTRE	-0,09	0,04	-2,01	58,11	0,0494
Marqueurs discursifs - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,31	0,02	277,13	36,48	< 0,0001
fixed	Valence NEUTRE	0,03	0,01	3,45	58,05	0,001
fixed	Valence POSITIVE	0,06	0,01	5,73	57,97	< 0,0001
Marqueurs discursifs (relevel) - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,37	0,02	280,13	36,48	< 0,0001
fixed	Valence NEGATIVE	-0,06	0,01	-5,73	57,97	< 0,0001
fixed	Valence NEUTRE	-0,02	0,01	-2,28	58,05	0,0262
Marqueurs discursifs - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,43	0,02	252,15	36,53	< 0,0001
fixed	Valence NEUTRE	0,05	0,01	4,43	58,03	< 0,0001
fixed	Valence POSITIVE	0,08	0,01	7,01	57,98	< 0,0001
Marqueurs discursifs (relevel) - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,51	0,02	256,10	36,53	< 0,0001
fixed	Valence NEGATIVE	-0,08	0,01	-7,01	57,98	< 0,0001
fixed	Valence NEUTRE	-0,03	0,01	-2,57	58,03	0,0127
Marqueurs discursifs - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,20	0,02	262,14	35,08	< 0,0001
fixed	Valence NEUTRE	0,02	0,01	1,81	58,82	0,0758
fixed	Valence POSITIVE	0,02	0,01	2,18	58,25	0,0334
Marqueurs discursifs (relevel) - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,22	0,02	263,41	34,97	< 0,0001
fixed	Valence NEGATIVE	-0,02	0,01	-2,18	58,25	0,0334
fixed	Valence NEUTRE	0,00	0,01	-0,37	58,28	0,7151
Séquence d'annonce - Empan vocal						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	4,18	0,03	128,68	40,50	< 0,0001
fixed	Valence NEUTRE	0,16	0,01	12,44	1032,05	< 0,0001
fixed	Valence POSITIVE	0,33	0,01	25,30	1032,01	< 0,0001
Séquence d'annonce (relevel) - Empan vocal						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	4,51	0,03	138,87	40,50	< 0,0001

fixed	Valence NEGATIVE	-0,33	0,01	-25,30	1032,01	< 0,0001
fixed	Valence NEUTRE	-0,17	0,01	-12,77	1032,05	< 0,0001
Séquence d'annonce - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,36	0,02	298,87	34,46	< 0,0001
fixed	Valence NEUTRE	0,04	0,01	6,16	58,00	< 0,0001
fixed	Valence POSITIVE	0,07	0,01	10,23	57,84	< 0,0001
Séquence d'annonce (relevel) - f0 moyenne						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,44	0,02	302,99	34,46	< 0,0001
fixed	Valence NEGATIVE	-0,07	0,01	-10,23	57,84	< 0,0001
fixed	Valence NEUTRE	-0,03	0,01	-4,06	58,00	0,0001
Séquence d'annonce - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,52	0,02	328,44	37,95	< 0,0001
fixed	Valence NEUTRE	0,06	0,01	6,76	58,00	< 0,0001
fixed	Valence POSITIVE	0,11	0,01	13,02	57,86	< 0,0001
Séquence d'annonce (relevel) - f0 maximum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,63	0,02	335,24	37,95	< 0,0001
fixed	Valence NEGATIVE	-0,11	0,01	-13,02	57,86	< 0,0001
fixed	Valence NEUTRE	-0,05	0,01	-6,25	58,00	< 0,0001
Séquence d'annonce - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,21	0,02	256,29	32,13	< 0,0001
fixed	Valence NEUTRE	0,02	0,01	2,69	58,75	0,0092
fixed	Valence POSITIVE	0,02	0,01	3,38	58,40	0,0013
Séquence d'annonce (relevel) - f0 minimum						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	5,23	0,02	257,30	32,13	< 0,0001
fixed	Valence NEGATIVE	-0,02	0,01	-3,38	58,40	0,0013
fixed	Valence NEUTRE	0,00	0,01	-0,68	58,71	0,4973

Annexe 7. Résumés des modèles LMER réalisés sur les données temporelles

Préface - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,07	0,06	16,94	14,15	< 0,0001
fixed	Valence NEUTRE	-0,01	0,01	-0,84	57,95	0,4024
fixed	Valence POSITIVE	-0,03	0,01	-2,67	57,73	0,0098
Préface (relevel) - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,03	0,06	16,44	14,15	< 0,0001
fixed	Valence NEGATIVE	0,03	0,01	2,67	57,73	0,0098
fixed	Valence NEUTRE	0,02	0,01	1,83	57,95	0,073
Préface - durée des pauses						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,71	0,06	-12,21	47,61	< 0,0001
fixed	Valence NEUTRE	-0,09	0,04	-2,35	51,72	0,0227
fixed	Valence POSITIVE	-0,08	0,04	-2,23	50,25	0,0301
Préface (relevel) - durée des pauses						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,79	0,06	-13,62	48,05	< 0,0001
fixed	Valence cNEGATIVE	0,08	0,04	2,23	50,25	0,0301
fixed	Valence cNEUTRE	0,00	0,04	-0,12	52,70	0,9022
Préface - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,79	0,02	72,80	26,64	< 0,0001
fixed	Valence NEUTRE	-0,01	0,01	-0,96	57,80	0,3417
fixed	Valence POSITIVE	0,01	0,01	1,10	57,61	0,2748
Préface (relevel) - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,80	0,02	73,26	26,64	< 0,0001
fixed	Valence NEGATIVE	-0,01	0,01	-1,10	57,61	0,2748
fixed	Valence NEUTRE	-0,02	0,01	-2,06	57,80	0,0439
Marqueurs discursifs - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,57	0,03	-17,90	49,18	< 0,0001
fixed	Valence NEUTRE	-0,06	0,02	-2,76	58,08	0,0077
fixed	Valence POSITIVE	-0,01	0,02	-0,32	57,97	0,7486
Marqueurs discursifs (relevel) - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,57	0,03	-18,12	49,15	< 0,0001
fixed	Valence NEGATIVE	0,01	0,02	0,32	57,97	0,7486
fixed	Valence NEUTRE	-0,05	0,02	-2,44	58,02	0,0178
Marqueurs discursifs - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value

fixed	(Intercept)	1,83	0,03	53,77	36,32	< 0,0001
fixed	Valence NEUTRE	0,00	0,02	-0,06	58,13	0,9511
fixed	Valence POSITIVE	-0,01	0,02	-0,47	57,91	0,6435
Marqueurs discursifs (relevel) - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,82	0,03	53,52	36,30	< 0,0001
fixed	Valence NEGATIVE	0,01	0,02	0,47	57,91	0,6435
fixed	Valence NEUTRE	0,01	0,02	0,40	58,02	0,6882
Séquence d'annonce - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,84	0,06	31,52	14,51	< 0,0001
fixed	Valence NEUTRE	-0,30	0,01	-20,84	1037,00	< 0,0001
fixed	Valence POSITIVE	-0,17	0,01	-11,44	1037,00	< 0,0001
Séquence d'annonce (relevel) - durée de l'énoncé						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,68	0,06	28,67	14,51	< 0,0001
fixed	Valence NEGATIVE	0,17	0,01	11,44	1037,00	< 0,0001
fixed	Valence NEUTRE	-0,14	0,01	-9,40	1037,00	< 0,0001
Séquence d'annonce - durée des pauses						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,08	0,07	-1,24	59,16	0,2214
fixed	Valence NEUTRE	-0,23	0,06	-4,08	55,27	0,0001
fixed	Valence POSITIVE	-0,22	0,06	-4,00	55,44	0,0002
Séquence d'annonce (relevel) - durée des pauses						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	-0,31	0,07	-4,57	59,42	< 0,0001
fixed	Valence cNEGATIVE	0,22	0,06	4,00	55,44	0,0002
fixed	Valence cNEUTRE	0,00	0,06	-0,09	55,63	0,932
Séquence d'annonce - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,76	0,02	87,85	46,99	< 0,0001
fixed	Valence NEUTRE	-0,02	0,01	-2,07	58,00	0,043
fixed	Valence POSITIVE	0,00	0,01	-0,20	58,00	0,8436
Séquence d'annonce (relevel) - Vitesse d'articulation						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	1,76	0,02	87,74	46,99	< 0,0001
fixed	Valence NEGATIVE	0,00	0,01	0,20	58,00	0,8436
fixed	Valence NEUTRE	-0,02	0,01	-1,87	58,00	0,0664

Annexe 8. Segments significatifs et résumés des GAMMs

Courbe de la f0 dans la préface

Positif vs Négatif : /zə/ de 3.7 à 7.9, /vuz/ de 11.4 à 21.9), /za/ (de 24.6 à 27.4), /ap/ (de 28.8 à 32.3), /ɛlos/ (de 36.5 à 50.4), et /yʒɛ/ (de 56.7 à 70)

Négatif vs Neutre : /z/ (au point 3.7), /ap/ (de 29.5 à 30.9), /ɛ/ de 37.2 à 37.9

Positif vs Neutre : /p/ au point 31.6, et /o/ de 46.3 à 49.0

Coefficients paramétriques	β	SE	t	p
Intercept	205.517	3.921	52.413	<0.0001
Valence neutre	6.263	5.545	1.129	0.258
Valence positif	13.316	5.545	2.401	0.0163
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.86	29	1130.6	<0.0001
s(Temps):Valence neutre	28.88	29	1227.9	<0.0001
s(Temps):Valence positif	28.89	29	1309.5	<0.0001
s(Participant, Valence)	86.14	87	100.5	<0.0001

Courbe de la f0 dans les marqueurs discursifs

Positif vs Négatif : 1.3 à 4.5 et 8.4 à 25 /v/ et /wala/

Négatif vs Neutre : /v/ 1 à 5.3, /ad/ de 20.3 à 28.9 et /k/ de 35.2 à 38

Positif vs Neutre : /k/ de 35.6 à 37.6

Coefficients paramétriques	β	SE	t	p
Intercept	174.060	4.223	41.212	<0.0001
Valence neutre	12.904	5.973	2.160	0.0308
Valence positif	12.577	5.973	2.106	0.0352
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	26.59	28.51	400.52	<0.0001
s(Temps):Valence neutre	25.52	28.02	330.28	<0.0001
s(Temps):Valence positif	26.60	28.52	414.86	<0.0001
s(Participant, Valence)	85.81	87.00	72.73	<0.0001

CPPs dans la préface

Positif vs Négatif : /uz/ de 17 à 20.5, et /yʒ/ de 57.4 à 60.2

Négatif vs Neutre : /y/ au point 59.5

Positif vs Neutre : aucune différence significative

Coefficients paramétriques	β	SE	t	p
Intercept	8.46970	0.17454	48.526	<0.0001

Valence neutre	0.06015	0.24682	0.244	0.807
Valence positif	0.15314	0.24683	0.620	0.535
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.91	29	1661.6	<0.0001
s(Temps):Valence neutre	28.92	29	1653.5	<0.0001
s(Temps):Valence positif	28.92	29	1659.5	<0.0001
s(Participant, Valence)	86.19	87	106.6	<0.0001

H1-H2 dans la préface

Positif vs Négatif : /z/ de 21.9 à 23.3, /pɛ/ de 34.4 à 35.1, /lo/ de 40 à 46.3, /os/ de 49.7 à 51.1, et /yzɛ/ de 55.3 à 70

Négatif vs Neutre : /a/ de 17 à 23.3, /p/ de 33.7 à 35.1, /uɜ/ de 40 à 42.1, /pɛ/ de 44.2 à 45.6, /s/ au point 50.4, et /yz/ de 56.7 à 60.2

Positif vs Neutre : /l/ de 42.1 à 44.9, et /zɛ/ de 63 à 65.8

Coefficients paramétriques	β	SE	t	p
Intercept	6.7641	0.2455	27.549	<0.0001
Valence neutre	-0.4686	0.3472	-1.350	0.177
Valence positif	-0.6120	0.3472	-1.763	0.078
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.56	28.98	193.25	<0.0001
s(Temps):Valence neutre	28.55	28.98	200.76	<0.0001
s(Temps):Valence positif	28.50	28.98	195.13	<0.0001
s(Participant, Valence)	85.10	87.00	45.18	<0.0001

HNR05 dans la préface

Positif vs Négatif : /zə/ de 4.4 à 6.5, /pɛ/ de 34.4 à 35.8

Négatif vs Neutre : /pɛ/ de 33 à 35.8

Positif vs Neutre : aucune différence significative

Coefficients paramétriques	β	SE	t	p
Intercept	-11.40101	1.118233	-10.196	<0.0001
Valence neutre	0.001154	1.580958	0.001	0.999
Valence positif	0.594900	1.581059	0.376	0.707
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.93	29	1510.24	<0.0001
s(Temps):Valence neutre	28.92	29	1488.93	<0.0001
s(Temps):Valence positif	28.93	29	1521.21	<0.0001
s(Participant, Valence)	80.85	87	13.13	<0.0001

HNR15 dans la préface

Positif vs Négatif : /ə/ 5.8 à 6.5, /pɛ/ de 33.7 à 35.1

Négatif vs Neutre : /pɛ/ de 32.3 à 35.1

Positif vs Neutre : /v/ : de 12.1 à 12.8

Coefficients paramétriques	β	SE	t	p
Intercept	- 17.6760	0.9680	-18.259	<0.0001
Valence neutre	-0.5009	1.3684	-0.366	0.714
Valence positif	-0.3439	1.3686	-0.251	0.802
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.92	29	1495.889	<0.0001
s(Temps):Valence neutre	28.92	29	1528.607	<0.0001
s(Temps):Valence positif	28.92	29	1555.906	<0.0001
s(Participant, Valence)	77.92	87	8.572	<0.0001

HNR25 dans la préface

Positif vs Négatif : /pɛ/ de 33 à 35.1, /z/ de 23 à 24.6

Négatif vs Neutre : /pɛ/ de 32.3 à 35.1

Positif vs Neutre : /v/ au point 12.1, /z/ de 23.3 à 24.6

Coefficients paramétriques	β	SE	t	p
Intercept	- 19.3664	0.9252	-20.931	<0.0001
Valence neutre	-0.1814	1.3078	-0.139	0.89
Valence positif	-0.4348	1.3080	-0.332	0.74
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.91	29	1457.392	<0.0001
s(Temps):Valence neutre	28.91	29	1484.574	<0.0001
s(Temps):Valence positif	28.92	29	1514.681	<0.0001
s(Participant, Valence)	76.87	87	7.586	<0.0001

HNR35 dans la préface

Positif vs Négatif : /zə/ de 3 à 5.1, /pɛ/ de 33 à 35.1, /z/ de 23.3 à 24.6, /ʒ/ de 62.3 à 63

Négatif vs Neutre : /pɛ/ de 33 à 35.1

Positif vs Neutre : /v/ de 11.4 à 12.8, /z/ de 24 à 24.6

Coefficients paramétriques	β	SE	t	p
Intercept	- 17.4521	0.8170	-21.362	<0.0001
Valence neutre	-0.1314	1.1547	-0.114	0.909
Valence positif	-0.2274	1.1548	-0.197	0.844
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.90	29	1355.752	<0.0001

s(Temps):Valence neutre	28.91	29	1408.462	<0.0001
s(Temps):Valence positif	28.91	29	1422.808	<0.0001
s(Participant, Valence)	74.49	87	5.944	<0.0001

CPPs dans les marqueurs discursifs

Positif vs Négatif : aucune différence significative

Négatif vs Neutre : de 2.9 à 4.1 /v/

Positif vs Neutre : aucune différence significative

Coefficients paramétriques	ß	SE	t	p
Intercept	9.37355	0.19545	47.960	<0.0001
Valence neutre	0.05357	0.27642	0.194	0.846
Valence positif	0.00686	0.27641	0.025	0.980
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	28.43	28.97	748.33	<0.0001
s(Temps):Valence neutre	28.43	28.97	691.62	<0.0001
s(Temps):Valence positif	28.41	28.97	687.64	<0.0001
s(Participant, Valence)	85.60	87.00	59.98	<0.0001

H1-H2 dans les marqueurs discursifs

Positif vs Négatif : /õ/ de 34 à 35.2

Négatif vs Neutre : /õ/ de 33.6 à 34.4, /k/ de 35.2 à 37.2 et /l/ de 19.5 à 19.9

Positif vs Neutre : /k/ de 36 à 37.2, et /l/ de 16.3 à 17.9

Coefficients paramétriques	ß	SE	t	p
Intercept	6.04657	0.19121	31.623	<0.0001
Valence neutre	0.09525	0.27047	0.352	0.725
Valence positif	-	0.27041	-0.358	0.720
0.09682				
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	24.48	27.42	108.23	<0.0001
s(Temps):Valence neutre	25.57	28.05	97.59	<0.0001
s(Temps):Valence positif	24.19	27.24	91.33	<0.0001
s(Participant, Valence)	82.09	87.00	16.89	<0.0001

HNR05 dans les marqueurs discursifs

Positif vs Négatif : /v/ de 1 à 4.1

Négatif vs Neutre : /v/ de 1 à 4.9, /d/, /õk/ de 35.2 à 37.6 et de 39.2 à 39.6

Positif vs Neutre : /v/ 1 à 1.3 et 2.9 à 4.9, /õk/ de 35.2 à 37.6

Coefficients paramétriques	ß	SE	t	p
Intercept	-11.5049	1.9051	-6.039	1.56e-09
Valence neutre	4.7742	2.6948	1.772	0.0765
Valence positif	0.6921	2.6943	0.257	0.7973
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	27.57	28.82	598.93	<0.0001
s(Temps):Valence neutre	26.78	28.58	471.76	<0.0001
s(Temps):Valence positif	27.37	28.77	526.51	<0.0001
s(Participant, Valence)	82.49	87.00	17.87	<0.0001

HNR15 dans les marqueurs discursifs

Positif vs Négatif : aucune différence significative

Négatif vs Neutre : /v/ de 1 à 4.9, /d/, /õk/ de 35.2 à 37.6 et de 38.8 à 39.2

Positif vs Neutre : /v/ de 1 à 1.3 et 2.5 à 4.9, /õk/ de 35.2 à 39.6, /d/

Coefficients paramétriques	ß	SE	t	p
Intercept	-19.585	2.037	-9.616	<0.0001
Valence neutre	4.782	2.881	1.660	0.097
Valence positif	-1.224	2.880	-0.425	0.671
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	27.94	28.90	552.39	<0.0001
s(Temps):Valence neutre	27.25	28.73	448.54	<0.0001
s(Temps):Valence positif	27.88	28.89	528.78	<0.0001
s(Participant, Valence)	82.52	87.00	17.96	<0.0001

HNR25 dans les marqueurs discursifs

Positif vs Négatif : /õk/ de 38 à 38.8

Négatif vs Neutre : /v/ de 1 à 4.9, /d/, /õk/ de 35.2 à 37.2

Positif vs Neutre : /v/ de 1 à 5.3, /d/, /õk/ de 35.2 à 37.2

Coefficients paramétriques	ß	SE	t	p
Intercept	-20.273	2.062	-9.834	<0.0001
Valence neutre	4.367	2.916	1.497	0.134
Valence positif	-2.009	2.916	-0.689	0.491
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	27.99	28.91	511.60	<0.0001
s(Temps):Valence neutre	27.20	28.72	436.35	<0.0001
s(Temps):Valence positif	27.90	28.89	507.20	<0.0001
s(Participant, Valence)	82.47	87.00	17.82	<0.0001

HNR35 dans les marqueurs discursifs

Positif vs Négatif : /zə/ de 3 à 5.1, /õk/ de 38 à 38.4

Négatif vs Neutre : /v/ de 1.3 à 4.9, /d/, /õk/ de 35.2 à 36.8 et au point 38

Positif vs Neutre : /v/ de 1 - 5.3, /d/, /õk/ de 35.2 à 36.8

Coefficients paramétriques	ß	SE	t	p
Intercept	-18.990	2.088	-9.094	<0.0001
Valence neutre	3.613	2.954	1.223	0.221
Valence positif	-2.135	2.953	-0.723	0.470
TERMES LISSES	Edf	Ref.df	F	p-value
s(Temps):Valence négatif	27.91	28.89	450.25	<0.0001
s(Temps):Valence neutre	27.03	28.67	403.76	<0.0001
s(Temps):Valence positif	27.82	28.88	445.03	<0.0001
s(Participant, Valence)	82.50	87.00	17.98	<0.0001

B. Etude 1 – Analyses phonologiques.

Annexe 9. Ensemble des patrons produits dans la préface

PATRON	NOMBRE	%
L*L%	467	43%
L*H%	361	33%
H*L%	127	12%
H*H%	84	8%
L*L-	12	1%
L*H-	12	1%
H*H-	12	1%
H*L-	2	0%

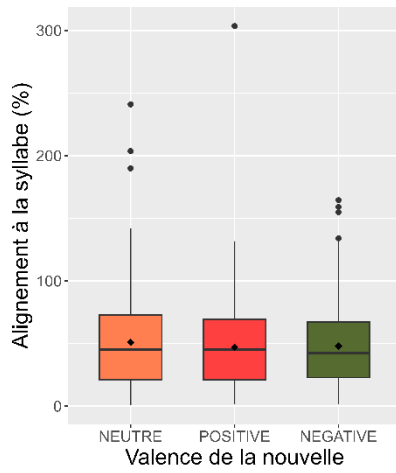
Annexe 10. Ensemble des patrons produits dans les marqueurs discursifs

VOILA			DONC			VOILA DONC		
PATRON	Nombre	%	Patron	Nombre	%	Patron	Nombre	%
HIL*L%	187	25%	L*	180	24%	L1L2H*H%	75	23%
HIL*	159	21%	L*L%	131	18%	HiL*L%	36	11%
HIL*L-	68	9%	L*L-	117	16%	L1HiL2H*H%	34	10%
L1H*	64	9%	H*H-	87	12%	L1HiL*L%	28	8%
L1HIL*L%	56	7%	H*	66	9%	HiL2H*H%	27	8%
L1HIL*	42	6%	H*H%	63	8%	HiL*L-	27	8%
L1H*H%	41	5%	L1H*H%	29	4%	HiL*	19	6%
HIL2H*	24	3%	HiL*L%	20	3%	L1L2H*H-	14	4%
L1HIL*L-	23	3%	HiL*L-	15	2%	L1HiL*L-	12	4%
L1H*H-	21	3%	L1H*	8	1%	HiL2H*H-	9	3%
HIL2H*H%	13	2%	HiL*	7	1%	L1H*H-	6	2%
HIL2H*H-	11	1%	L1H*H-	7	1%	L1HiL*	6	2%
HIH*	6	1%	H*L%	6	1%	HiL2H*	5	2%
L1L2H*H%	4	1%	L*H%	3	0%	L1HiL2H*H-	5	2%
L1L2H*H-	4	1%	L*H-	3	0%	L1HiH*H-	4	1%
L1L*	3	0%	H*L-	2	0%	HiL2H*L%	3	1%
L1HIH*H%	3	0%	HiL2H*	1	0%	L1H*	3	1%
L1HIL2H*H%	3	0%	HiL*H-	1	0%	L1H*H%	2	1%
HIL*H-	2	0%	L1L2H*	1	0%	L1H*L%	2	1%
L1HIH*	2	0%				L1H*L-	2	1%

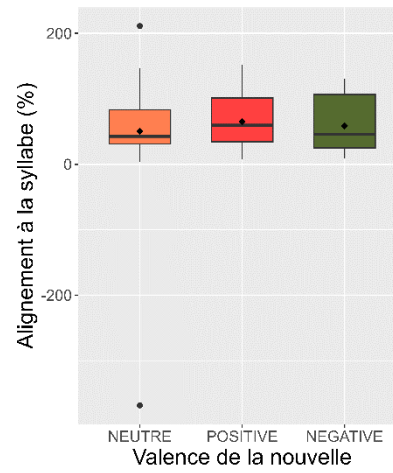
L1HIH*H-	2	0%		L1HiH*H%	2	1%
HIL2H*L-	1	0%		NA	2	1%
L1H*L%	1	0%		H*H-	1	0%
L1H*L-	1	0%		HiL2H*HiL*L%	1	0%
L1L*L-	1	0%		HiL*H-	1	0%
L1HIL2H*	1	0%		HiL*LH*	1	0%
L1HIL2H*H-	1	0%		L1HiH*L-	1	0%
L1HIL*H%	1	0%		L1HiL2H*	1	0%
L1HIL*H-	1	0%		L1HiL2L*L%	1	0%
L1L2H*	1	0%		L1HiL*L	1	0%
				L1L2H*	1	0%
				L1L2H*L%	1	0%

Annexe 11. Graphiques des résultats non significatifs dans la préface

*Alignement : Pitch accent L**



*Alignement : Pitch accent H**



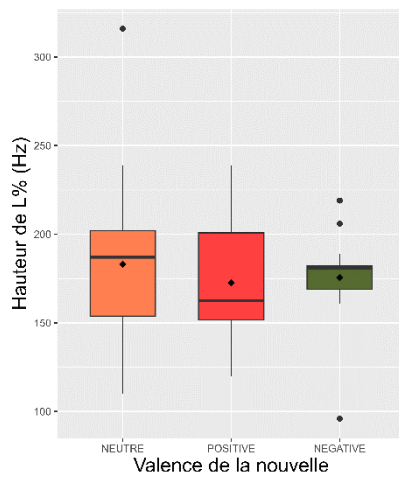
Annexe 12. Résumés des modèles réalisés sur la valeur mélodique et l'alignement dans la préface

Préface - hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	182,62	5,46	33,45	42,65	< 0,0001
fixed	Valence POSITIVE	2,68	2,27	1,18	557,17	0,2371
fixed	Valence NEGATIVE	-3,09	2,20	-1,40	556,21	0,1611
Préface (relevel) - hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	185,30	5,44	34,06	42,05	< 0,0001
fixed	Valence NEGATIVE	-5,77	2,14	-2,70	555,76	0,0071
fixed	Valence NEUTRE	-2,68	2,27	-1,18	557,17	0,2371
Préface - hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	274,28	9,36	29,30	42,57	< 0,0001
fixed	Valence POSITIVE	-3,35	6,10	-0,55	413,53	0,5838
fixed	Valence NEGATIVE	-38,14	6,26	-6,09	412,20	< 0,0001
Préface (relevel) - hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	270,94	9,54	28,40	45,49	< 0,0001
fixed	Valence NEGATIVE	-34,80	6,40	-5,44	408,78	< 0,0001
fixed	Valence NEUTRE	3,35	6,10	0,55	413,53	0,5838
Préface - hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	189,53	5,78	32,81	39,67	< 0,0001
fixed	Valence POSITIVE	3,71	1,93	1,93	811,43	0,0544
fixed	Valence NEGATIVE	-4,67	1,93	-2,42	811,64	0,0155
Préface (relevel) - hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	193,24	5,78	33,43	39,77	< 0,0001
fixed	Valence NEGATIVE	-8,38	1,93	-4,34	811,54	< 0,0001
fixed	Valence NEUTRE	-3,71	1,93	-1,93	811,43	0,0544
Préface - hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	261,37	7,55	34,64	38,65	< 0,0001
fixed	Valence POSITIVE	-6,15	5,80	-1,06	211,60	0,2898
fixed	Valence NEGATIVE	-17,14	5,87	-2,92	207,63	0,0039
Préface (relevel) - hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	255,21	7,43	34,33	36,62	< 0,0001
fixed	Valence NEGATIVE	-10,98	5,77	-1,90	208,77	0,0584
fixed	Valence NEUTRE	6,15	5,80	1,06	211,60	0,2898
Préface - alignement : L*						

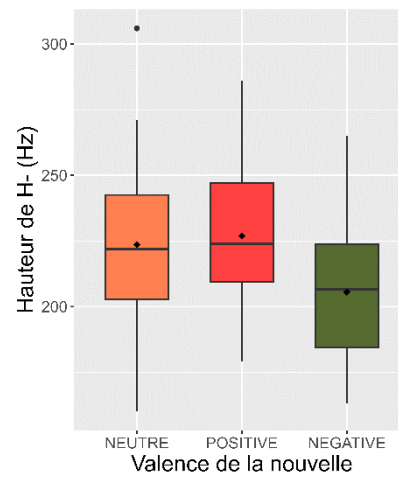
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	52,61	5,39	9,76	19,54	< 0,0001
fixed	Valence POSITIVE	-2,75	2,37	-1,16	814,34	0,2466
fixed	Valence NEGATIVE	-2,64	2,37	-1,12	816,40	0,2649
Préface (relevel) - alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	49,86	5,39	9,24	19,63	< 0,0001
fixed	Valence NEGATIVE	0,10	2,38	0,04	815,96	0,9656
fixed	Valence NEUTRE	2,75	2,37	1,16	814,34	0,2466
Préface - alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	43,92	10,50	4,18	33,84	0,0002
fixed	Valence POSITIVE	5,55	6,26	0,89	196,95	0,3762
fixed	Valence NEGATIVE	7,74	6,32	1,22	190,88	0,2222
Préface (relevel) - alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	49,47	10,45	4,74	33,16	< 0,0001
fixed	Valence NEGATIVE	2,19	6,15	0,36	191,99	0,722
fixed	Valence NEUTRE	-5,55	6,26	-0,89	196,95	0,3762

Annexe 13. Graphiques des résultats non significatifs dans les marqueurs discursifs

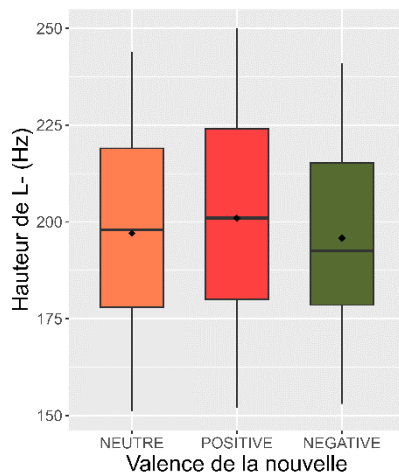
Valeur mélodique : Ton de frontière L% - « voilà donc »



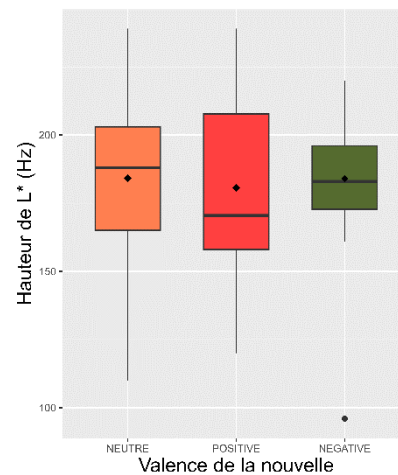
Valeur mélodique : Ton de frontière H- - « donc »



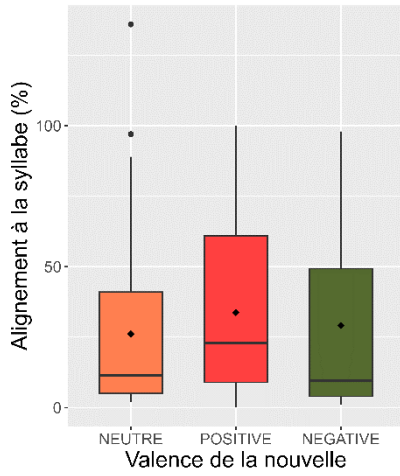
Valeur mélodique : Ton de frontière L- - « donc »



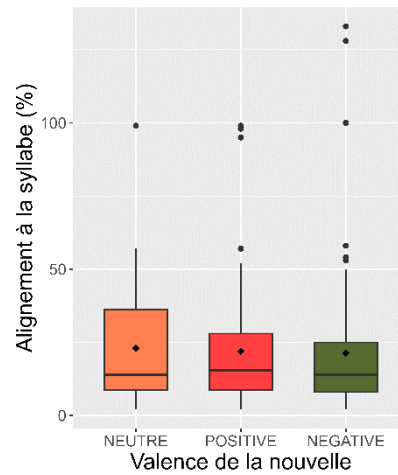
Valeur mélodique : Pitch accent L* - « voilà donc »



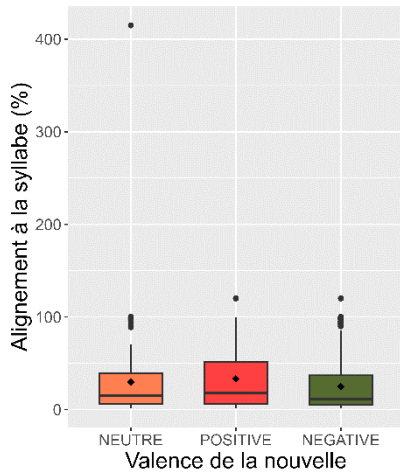
Alignement : Pitch accent L* - « voilà donc »



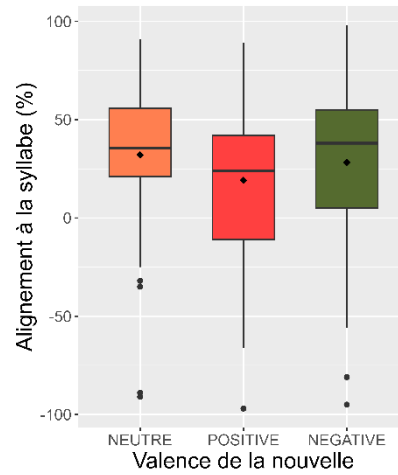
Alignement : Pitch accent H* - « donc »



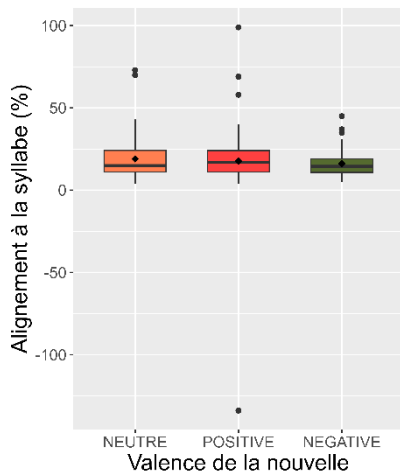
Alignement : Pitch accent L* - « donc »



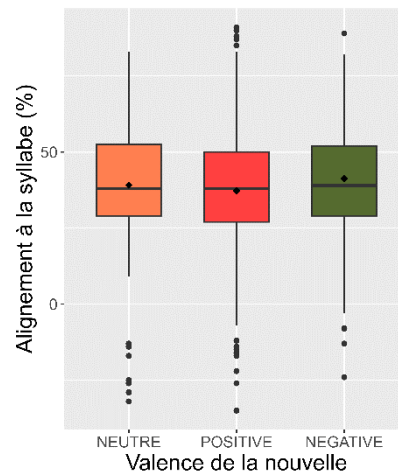
Alignement : Ton Hi - « voilà donc »



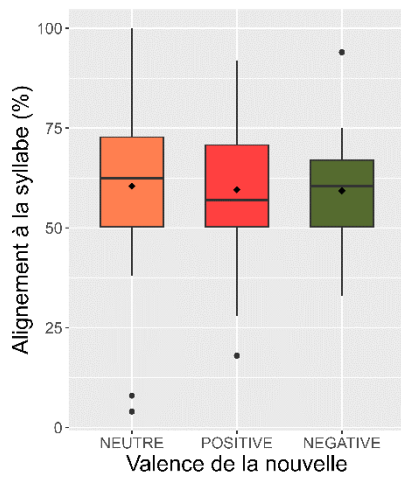
Alignement : Pitch accent H* - « voilà »



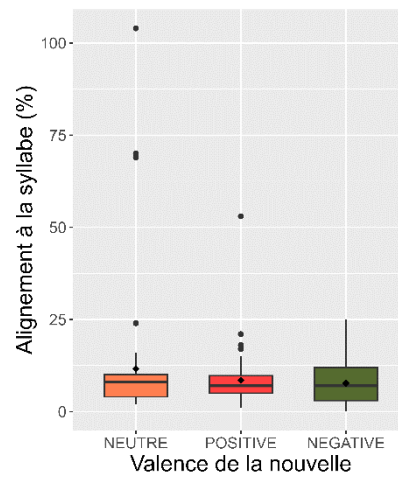
Alignement : Ton Hi - « voilà »



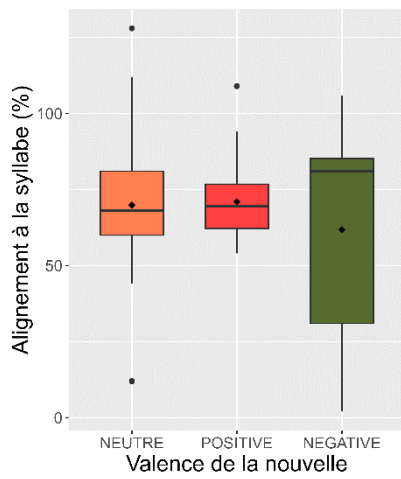
Alignement : Ton L1 - « voilà donc »



Alignement : Ton L2 - « voilà donc »



Alignement : Ton L2 - « voilà »



Annexe 14. Résumés des modèles réalisés sur la valeur mélodique et l’alignement dans les marqueurs discursifs

Marqueurs discursifs - Voilà, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	182,28	5,39	33,82	33,07	< 0,0001
fixed	Valence POSITIVE	4,99	4,10	1,22	228,99	0,2251
fixed	Valence NEGATIVE	-5,88	4,08	-1,44	229,09	0,1504
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	187,27	5,23	35,80	26,90	< 0,0001
fixed	Valence NEUTRE	-4,99	4,10	-1,22	228,99	0,2251
fixed	Valence NEGATIVE	-10,88	3,58	-3,04	223,66	0,0027
Marqueurs discursifs - Donc, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	185,62	5,16	36,00	39,17	< 0,0001
fixed	Valence POSITIVE	9,34	4,07	2,29	140,44	0,0234
fixed	Valence NEGATIVE	-0,85	4,13	-0,21	142,50	0,8371
Marqueurs discursifs (relevel) - Donc, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	194,96	5,19	37,60	41,20	< 0,0001
fixed	Valence NEUTRE	-9,34	4,07	-2,29	140,44	0,0234
fixed	Valence NEGATIVE	-10,19	4,11	-2,48	137,57	0,0144
Marqueurs discursifs - Voilà donc, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	193,77	7,44	26,04	22,13	< 0,0001
fixed	Valence POSITIVE	-2,43	6,77	-0,36	63,69	0,7209
fixed	Valence NEGATIVE	-7,92	10,61	-0,75	73,46	0,4579
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : L%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	191,34	8,01	23,88	28,76	< 0,0001
fixed	Valence NEUTRE	2,43	6,77	0,36	63,69	0,7209
fixed	Valence NEGATIVE	-5,49	10,84	-0,51	72,68	0,6142
Marqueurs discursifs - Voilà, hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	248,24	8,71	28,50	24,95	< 0,0001
fixed	Valence POSITIVE	8,05	8,47	0,95	57,91	0,346
fixed	Valence NEGATIVE	-26,84	10,04	-2,67	55,93	0,0098
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	256,29	8,24	31,11	19,65	< 0,0001
fixed	Valence NEUTRE	-8,05	8,47	-0,95	57,91	0,346
fixed	Valence NEGATIVE	-34,89	10,64	-3,28	59,89	0,0017
Marqueurs discursifs - Donc, hauteur mélodique : H%						

effect	term	β	SE	t	df	p.value
fixed	(Intercept)	228,28	8,72	26,18	24,81	< 0,0001
fixed	Valence POSITIVE	13,04	6,88	1,89	72,69	0,0621
fixed	Valence NEGATIVE	-7,09	6,84	-1,04	65,72	0,3042
Marqueurs discursifs (relevel) - Donc, hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	241,32	8,14	29,64	18,56	< 0,0001
fixed	Valence NEUTRE	-13,04	6,88	-1,89	72,69	0,0621
fixed	Valence NEGATIVE	-20,12	6,11	-3,29	69,89	0,0016
Marqueurs discursifs - Voilà donc, hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	267,05	12,52	21,33	15,46	< 0,0001
fixed	Valence POSITIVE	8,17	5,40	1,51	128,62	0,1327
fixed	Valence NEGATIVE	-24,12	5,49	-4,40	128,67	< 0,0001
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : H%						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	275,22	12,59	21,86	15,77	< 0,0001
fixed	Valence NEUTRE	-8,17	5,40	-1,51	128,62	0,1327
fixed	Valence NEGATIVE	-32,29	5,28	-6,12	127,97	< 0,0001
Marqueurs discursifs - Voilà, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	192,97	4,85	39,76	30,83	< 0,0001
fixed	Valence POSITIVE	1,16	2,23	0,52	504,39	0,602
fixed	Valence NEGATIVE	-6,03	2,21	-2,73	507,81	0,0067
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	194,13	4,86	39,96	30,72	< 0,0001
fixed	Valence NEUTRE	-1,16	2,23	-0,52	504,39	0,602
fixed	Valence NEGATIVE	-7,19	2,18	-3,30	506,30	0,001
Marqueurs discursifs - Donc, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	191,73	4,52	42,38	29,56	< 0,0001
fixed	Valence POSITIVE	4,76	1,84	2,59	445,13	0,0099
fixed	Valence NEGATIVE	-1,40	1,87	-0,75	444,09	0,4539
Marqueurs discursifs (relevel) - Donc, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	196,49	4,55	43,23	30,06	< 0,0001
fixed	Valence NEUTRE	-4,76	1,84	-2,59	445,13	0,0099
fixed	Valence NEGATIVE	-6,16	1,90	-3,23	442,62	0,0013
Marqueurs discursifs - Voilà donc, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	189,86	5,15	36,83	33,35	< 0,0001
fixed	Valence POSITIVE	-0,96	3,83	-0,25	120,25	0,8022
fixed	Valence NEGATIVE	-1,97	4,88	-0,40	131,84	0,6867
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : L*						
effect	term	β	SE	t	df	p.value

fixed	(Intercept)	188,90	5,19	36,40	34,39	< 0,0001
fixed	Valence NEUTRE	0,96	3,83	0,25	120,25	0,8022
fixed	Valence NEGATIVE	-1,01	4,88	-0,21	130,98	0,8365
Marqueurs discursifs - Voilà, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	241,20	5,68	42,48	26,60	< 0,0001
fixed	Valence POSITIVE	8,73	3,61	2,42	183,59	0,0165
fixed	Valence NEGATIVE	-16,47	4,37	-3,77	186,43	0,0002
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	249,93	5,73	43,62	28,09	< 0,0001
fixed	Valence NEUTRE	-8,73	3,61	-2,42	183,59	0,0165
fixed	Valence NEGATIVE	-25,20	4,50	-5,60	190,25	< 0,0001
Marqueurs discursifs - Donc, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	223,41	5,80	38,51	27,98	< 0,0001
fixed	Valence POSITIVE	8,13	3,08	2,64	240,15	0,0088
fixed	Valence NEGATIVE	-7,46	3,05	-2,45	239,97	0,015
Marqueurs discursifs (relevel) - Donc, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	231,54	5,74	40,36	26,78	< 0,0001
fixed	Valence NEUTRE	-8,13	3,08	-2,64	240,15	0,0088
fixed	Valence NEGATIVE	-15,59	2,88	-5,42	239,98	< 0,0001
Marqueurs discursifs - Voilà donc, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	251,93	9,82	25,66	25,57	< 0,0001
fixed	Valence POSITIVE	4,92	4,71	1,04	184,76	0,2975
fixed	Valence NEGATIVE	-24,96	4,59	-5,44	185,40	< 0,0001
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	256,85	9,96	25,80	26,85	< 0,0001
fixed	Valence NEUTRE	-4,92	4,71	-1,04	184,76	0,2975
fixed	Valence NEGATIVE	-29,88	4,59	-6,52	184,41	< 0,0001
Marqueurs discursifs - Voilà, hauteur mélodique : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	204,79	4,61	44,40	30,84	< 0,0001
fixed	Valence POSITIVE	8,48	2,08	4,07	249,60	0,0001
fixed	Valence NEGATIVE	-2,32	2,34	-0,99	250,65	0,3239
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	213,27	4,60	46,32	30,62	< 0,0001
fixed	Valence NEUTRE	-8,48	2,08	-4,07	249,60	0,0001
fixed	Valence NEGATIVE	-10,80	2,36	-4,58	250,28	< 0,0001
Marqueurs discursifs - Voilà donc, hauteur mélodique : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	203,50	6,63	30,68	25,61	< 0,0001

fixed	Valence POSITIVE	3,17	2,95	1,08	184,29	0,2836
fixed	Valence NEGATIVE	0,68	2,98	0,23	184,07	0,8183
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	206,67	6,63	31,17	25,53	< 0,0001
fixed	Valence NEUTRE	-3,17	2,95	-1,08	184,29	0,2836
fixed	Valence NEGATIVE	-2,49	2,99	-0,83	184,42	0,4067
Marqueurs discursifs - Voilà, hauteur mélodique : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	199,08	9,88	20,15	10,10	< 0,0001
fixed	Valence POSITIVE	-0,77	6,12	-0,13	50,74	0,9006
fixed	Valence NEGATIVE	8,20	7,53	1,09	53,32	0,2811
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	198,31	10,15	19,54	11,00	< 0,0001
fixed	Valence NEUTRE	0,77	6,12	0,13	50,74	0,9006
fixed	Valence NEGATIVE	8,96	7,51	1,19	51,66	0,2378
Marqueurs discursifs - Voilà donc, hauteur mélodique : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	196,89	7,12	27,65	19,66	< 0,0001
fixed	Valence POSITIVE	6,69	3,46	1,94	158,51	0,0547
fixed	Valence NEGATIVE	-1,40	3,42	-0,41	159,72	0,682
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	203,58	7,19	28,30	20,40	< 0,0001
fixed	Valence NEUTRE	-6,69	3,46	-1,94	158,51	0,0547
fixed	Valence NEGATIVE	-8,09	3,40	-2,38	159,11	0,0184
Marqueurs discursifs - Voilà, hauteur mélodique : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	228,74	4,27	53,55	29,34	< 0,0001
fixed	Valence POSITIVE	9,94	1,52	6,52	568,64	< 0,0001
fixed	Valence NEGATIVE	-7,43	1,53	-4,84	569,01	< 0,0001
Marqueurs discursifs (relevel) - Voilà, hauteur mélodique : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	238,68	4,26	56,00	29,01	< 0,0001
fixed	Valence NEUTRE	-9,94	1,52	-6,52	568,64	< 0,0001
fixed	Valence NEGATIVE	-17,37	1,51	-11,53	568,29	< 0,0001
Marqueurs discursifs - Voilà donc, hauteur mélodique : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	229,82	5,57	41,27	38,96	< 0,0001
fixed	Valence POSITIVE	17,51	3,60	4,86	206,12	< 0,0001
fixed	Valence NEGATIVE	-11,80	4,30	-2,74	215,17	0,0066
Marqueurs discursifs (relevel) - Voilà donc, hauteur mélodique : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	247,32	5,54	44,61	38,35	< 0,0001
fixed	Valence NEUTRE	-17,51	3,60	-4,86	206,12	< 0,0001

fixed	Valence NEGATIVE	-29,31	4,35	-6,73	217,23	< 0,0001
Marqueurs discursifs - Voilà, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	15,64	2,28	6,85	48,81	< 0,0001
fixed	Valence POSITIVE	0,93	1,78	0,52	514,87	0,6028
fixed	Valence NEGATIVE	4,18	1,76	2,38	521,49	0,0178
Marqueurs discursifs (relevel) - Voilà, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	16,57	2,27	7,29	46,66	< 0,0001
fixed	Valence NEUTRE	-0,93	1,78	-0,52	514,87	0,6028
fixed	Valence NEGATIVE	3,26	1,73	1,88	521,74	0,0608
Marqueurs discursifs - Donc, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	30,00	3,34	8,97	45,57	< 0,0001
fixed	Valence POSITIVE	3,59	3,83	0,94	458,82	0,3487
fixed	Valence NEGATIVE	-4,47	3,85	-1,16	462,80	0,2462
Marqueurs discursifs (relevel) - Donc, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	33,59	3,44	9,76	50,53	< 0,0001
fixed	Valence NEUTRE	-3,59	3,83	-0,94	458,82	0,3487
fixed	Valence NEGATIVE	-8,07	3,94	-2,05	462,82	0,0411
Marqueurs discursifs - Voilà donc, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	25,49	4,68	5,45	55,86	< 0,0001
fixed	Valence POSITIVE	7,32	5,70	1,28	132,85	0,2012
fixed	Valence NEGATIVE	1,75	6,62	0,26	124,05	0,7922
Marqueurs discursifs (relevel) - Voilà donc, alignement : L*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	32,81	4,78	6,86	58,29	< 0,0001
fixed	Valence NEUTRE	-7,32	5,70	-1,28	132,85	0,2012
fixed	Valence NEGATIVE	-5,57	6,68	-0,83	125,58	0,406
Marqueurs discursifs - Voilà, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	21,44	4,98	4,31	16,30	0,0005
fixed	Valence POSITIVE	-2,05	2,19	-0,93	177,46	0,3513
fixed	Valence NEGATIVE	-3,59	2,68	-1,34	181,93	0,1821
Marqueurs discursifs (relevel) - Voilà, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	19,39	4,99	3,88	16,59	0,0012
fixed	Valence NEUTRE	2,05	2,19	0,93	177,46	0,3513
fixed	Valence NEGATIVE	-1,54	2,76	-0,56	183,31	0,5782
Marqueurs discursifs - Donc, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	23,10	2,68	8,63	54,25	< 0,0001
fixed	Valence POSITIVE	-1,27	3,25	-0,39	255,70	0,6955
fixed	Valence NEGATIVE	-1,81	3,21	-0,56	254,24	0,5728

Marqueurs discursifs (relevel) - Donc, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	21,82	2,49	8,77	40,40	< 0,0001
fixed	Valence NEUTRE	1,27	3,25	0,39	255,70	0,6955
fixed	Valence NEGATIVE	-0,54	3,05	-0,18	255,32	0,8602
Marqueurs discursifs - Voilà donc, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	22,59	4,71	4,80	33,67	< 0,0001
fixed	Valence POSITIVE	11,07	4,09	2,71	183,40	0,0074
fixed	Valence NEGATIVE	0,10	3,97	0,03	184,43	0,98
Marqueurs discursifs (relevel) - Voilà donc, alignement : H*						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	33,66	4,86	6,93	35,07	< 0,0001
fixed	Valence NEUTRE	-11,07	4,09	-2,71	183,40	0,0074
fixed	Valence NEGATIVE	-10,97	4,00	-2,75	184,59	0,0066
Marqueurs discursifs - Voilà, alignement : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	62,84	2,08	30,26	42,87	< 0,0001
fixed	Valence POSITIVE	0,29	2,10	0,14	259,82	0,8911
fixed	Valence NEGATIVE	-3,27	2,35	-1,39	258,42	0,1648
Marqueurs discursifs (relevel) - Voilà, alignement : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	63,13	2,06	30,70	41,08	< 0,0001
fixed	Valence NEUTRE	-0,29	2,10	-0,14	259,82	0,8911
fixed	Valence NEGATIVE	-3,56	2,36	-1,51	259,62	0,1335
Marqueurs discursifs - Voilà donc, alignement : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	60,79	2,18	27,92	45,01	< 0,0001
fixed	Valence POSITIVE	-1,17	2,17	-0,54	200,86	0,5897
fixed	Valence NEGATIVE	-1,57	2,20	-0,71	198,77	0,4778
Marqueurs discursifs (relevel) - Voilà donc, alignement : L1						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	59,61	2,15	27,74	40,69	< 0,0001
fixed	Valence NEUTRE	1,17	2,17	0,54	200,86	0,5897
fixed	Valence NEGATIVE	-0,39	2,20	-0,18	201,53	0,8582
Marqueurs discursifs - Voilà, alignement : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	69,88	4,74	14,75	58,00	< 0,0001
fixed	Valence POSITIVE	1,12	6,92	0,16	58,00	0,872
fixed	Valence NEGATIVE	-8,09	7,91	-1,02	58,00	0,3101
Marqueurs discursifs (relevel) - Voilà, alignement : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	71,00	5,05	14,06	58,00	< 0,0001
fixed	Valence NEUTRE	-1,12	6,92	-0,16	58,00	0,872
fixed	Valence NEGATIVE	-9,21	8,10	-1,14	58,00	0,2598
Marqueurs discursifs - Voilà donc, alignement : L2						

effect	term	β	SE	t	df	p.value
fixed	(Intercept)	17,41	3,48	5,00	21,95	0,0001
fixed	Valence POSITIVE	-2,60	1,28	-2,03	158,77	0,044
fixed	Valence NEGATIVE	-2,58	1,27	-2,03	159,41	0,0438
Marqueurs discursifs (relevel) - Voilà donc, alignement : L2						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	14,81	3,51	4,22	22,46	0,0003
fixed	Valence NEUTRE	2,60	1,28	2,03	158,77	0,044
fixed	Valence NEGATIVE	0,02	1,26	0,02	159,03	0,9851
Marqueurs discursifs - Voilà, alignement : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	37,74	2,80	13,49	29,08	< 0,0001
fixed	Valence POSITIVE	-1,24	1,84	-0,67	573,34	0,5014
fixed	Valence NEGATIVE	2,09	1,85	1,13	577,59	0,2587
Marqueurs discursifs (relevel) - Voilà, alignement : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	36,50	2,77	13,18	27,61	< 0,0001
fixed	Valence NEUTRE	1,24	1,84	0,67	573,34	0,5014
fixed	Valence NEGATIVE	3,32	1,81	1,83	583,01	0,0671
Marqueurs discursifs - Voilà donc, alignement : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	35,74	5,83	6,13	33,79	< 0,0001
fixed	Valence POSITIVE	-14,90	5,05	-2,95	206,51	0,0036
fixed	Valence NEGATIVE	-7,22	5,93	-1,22	223,29	0,2245
Marqueurs discursifs (relevel) - Voilà donc, alignement : Hi						
effect	term	β	SE	t	df	p.value
fixed	(Intercept)	20,85	5,79	3,60	32,74	0,001
fixed	Valence NEUTRE	14,90	5,05	2,95	206,51	0,0036
fixed	Valence NEGATIVE	7,67	5,98	1,28	223,86	0,201

C. Etude de perception

Annexe 15. Liste des stimuli de la tâche préliminaire

Item	Positif	Négatif	Neutre
1	J'ai une bonne nouvelle au sujet de vos examens médicaux	J'ai une mauvaise nouvelle au sujet de vos examens médicaux	J'ai une nouvelle au sujet de vos examens médicaux
2	J'ai une bonne nouvelle au sujet de votre vol de ce soir	J'ai une mauvaise nouvelle au sujet de votre vol de ce soir	J'ai une nouvelle au sujet de votre vol de ce soir
3	J'ai une bonne nouvelle au sujet de votre enfant Léo	J'ai une mauvaise nouvelle au sujet de votre enfant Léo	J'ai une nouvelle au sujet de votre enfant Léo
4	J'ai une bonne nouvelle au sujet de votre demande d'augmentation	J'ai une mauvaise nouvelle au sujet de votre demande d'augmentation	J'ai une nouvelle au sujet de votre demande d'augmentation
5	J'ai une bonne nouvelle au sujet de notre commande pour de nouveaux outils de nettoyage	J'ai une mauvaise nouvelle au sujet de notre commande pour de nouveaux outils de nettoyage	J'ai une nouvelle au sujet de notre commande pour de nouveaux outils de nettoyage
6	J'ai une bonne nouvelle au sujet de votre chat Minouche qui était malade	J'ai une mauvaise nouvelle au sujet de votre chat Minouche qui était malade	J'ai une nouvelle au sujet de votre chat Minouche qui était malade

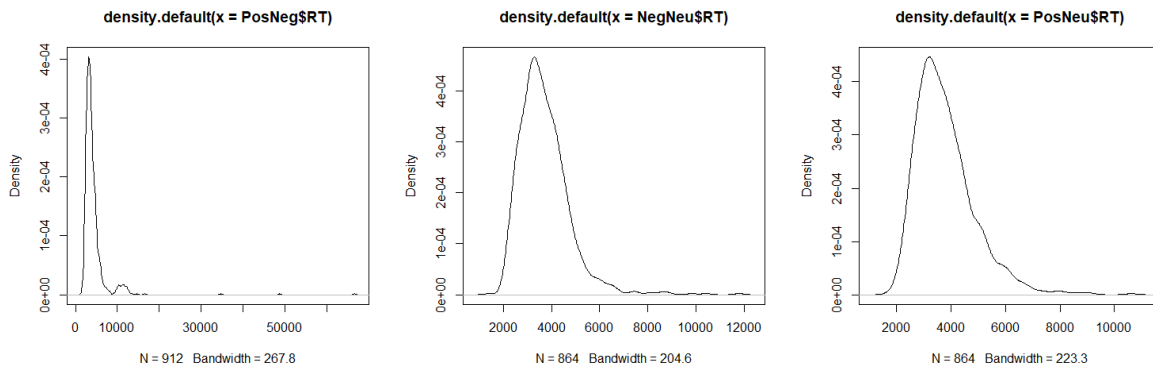
7	J'ai une bonne nouvelle au sujet du voyage scolaire	J'ai une mauvaise nouvelle au sujet du voyage scolaire	J'ai une nouvelle au sujet du voyage scolaire
8	J'ai une bonne nouvelle au sujet de votre vase en porcelaine chinoise	J'ai une mauvaise nouvelle au sujet de votre vase en porcelaine chinoise	J'ai une nouvelle au sujet de votre vase en porcelaine chinoise
9	J'ai une bonne nouvelle au sujet de votre demande de télétravail	J'ai une mauvaise nouvelle au sujet de votre demande de télétravail	J'ai une nouvelle au sujet de votre demande de télétravail
10	J'ai une bonne nouvelle au sujet de votre procès à venir	J'ai une mauvaise nouvelle au sujet de votre procès à venir	J'ai une nouvelle au sujet de votre procès à venir
11	J'ai une bonne nouvelle au sujet de des changements dans la librairie	J'ai une mauvaise nouvelle au sujet de des changements dans la librairie	J'ai une nouvelle au sujet de des changements dans la librairie
12	J'ai une bonne nouvelle au sujet de vos deux cours d'histoire	J'ai une mauvaise nouvelle au sujet de vos deux cours d'histoire	J'ai une nouvelle au sujet de vos deux cours d'histoire

Annexe 16. Questionnaire d'empathie BES-A

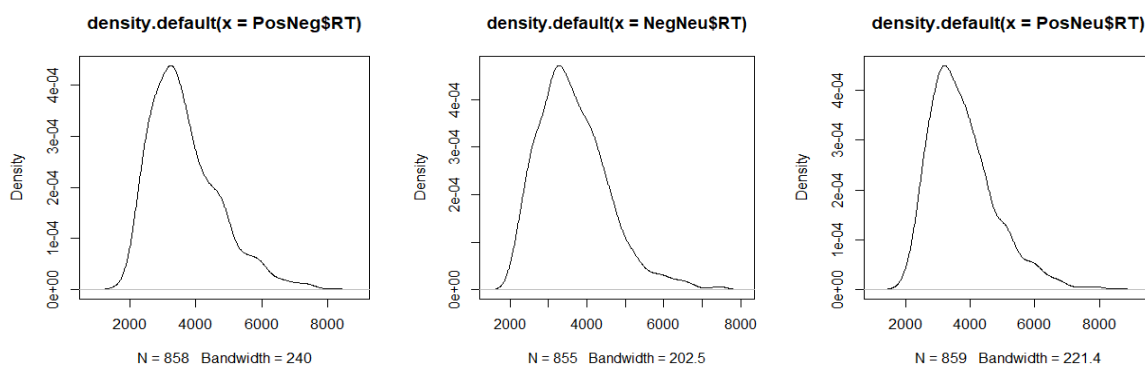
Voici des caractéristiques qui peuvent ou non vous être appliquées. Veuillez entourer une réponse pour chaque affirmation afin d'indiquer dans quelle mesure vous êtes d'accord ou pas d'accord avec l'affirmation. Répondez aussi honnêtement que vous le pouvez.	1	2	3	4	5
	pas du tout d'accord	pas d'accord	ni d'accord ni pas d'accord	d'accord	tout à fait d'accord
1. Les émotions de mes amis(ies) ne m'affectent pas beaucoup	1	2	3	4	5
2. Après avoir été avec un(e) ami(e) qui est triste, je me sens généralement triste	1	2	3	4	5
3. Je peux comprendre le bonheur d'un(e) ami(e) lorsque quelque chose marche bien pour lui/elle	1	2	3	4	5
4. Je peux être effrayé(e) par un bon film d'horreur	1	2	3	4	5
5. Je m'imprègne facilement des émotions des autres	1	2	3	4	5
6. J'ai du mal à savoir quand mes amis(ies) ont peur	1	2	3	4	5
7. Je ne suis pas triste quand je vois une personne pleurer	1	2	3	4	5
8. Les sentiments des autres ne me tracassent pas du tout	1	2	3	4	5
9. Quand quelqu'un a la sensation d'être « au plus bas », je peux comprendre ce qu'il/elle ressent	1	2	3	4	5
10. Généralement, je peux me rendre compte que mes amis(ies) sont effrayés(ées)	1	2	3	4	5
11. Souvent, je me sens triste quand je regarde des choses ou des films tristes	1	2	3	4	5
12. Souvent, je peux comprendre comment les gens se sentent avant même qu'ils ne me l'aient dit	1	2	3	4	5
13. Voir une personne se mettre en colère n'a aucun effet sur moi	1	2	3	4	5
14. Habituellement, je sais quand les gens sont joyeux	1	2	3	4	5
15. J'ai tendance à me sentir apeuré(ée) quand je suis avec des amis(ies) qui sont effrayés(ées)	1	2	3	4	5
16. Généralement, je me rends vite compte quand un(e) ami(e) est furieux(se)	1	2	3	4	5
17. Je suis souvent happé(e) par les sentiments de mes amis(ies)	1	2	3	4	5
18. Je ne ressens rien face à la tristesse de mes amis(ies)	1	2	3	4	5
19. Je ne suis généralement pas attentif(ve) aux sentiments de mes amis(ies)	1	2	3	4	5
20. J'ai du mal à comprendre quand mes amis(ies) sont heureux(ses)	1	2	3	4	5

Annexe 17. Plots de densité pour les temps de réaction dans l'étude principale

Densité dans les études Négatif vs Positif, Négatif vs Neutre, Positif vs Neutre (de gauche à droite), avant nettoyage des données :



Densité dans les études Négatif vs Positif, Négatif vs Neutre, Positif vs Neutre (de gauche à droite), après nettoyage des données :



Annexe 18. Résumés des emmeans réalisés sur le RT selon les différences tâches

Tâche 1 - RT : Valence*Congruence*Genre					
Contraste	β	SE	df	t	p-value
Négatif congr. F - Positif congr. F	0,05	0,08	5,99	0,64	0,9962
Négatif congr. F - Négatif incongr. F	-0,03	0,02	758,72	-1,18	0,9387
Négatif congr. F - Positif incongr. F	0,04	0,08	6,17	0,50	0,9992
Négatif congr. F - Négatif congr. H	-0,07	0,06	44,81	-1,27	0,9055
Négatif congr. F - Positif congr. H	-0,01	0,09	13,01	-0,10	1
Négatif congr. F - Négatif incongr. H	-0,08	0,06	40,49	-1,46	0,8244
Négatif congr. F - Positif incongr. H	0,00	0,09	13,92	0,00	1
Positif congr. F - Négatif incongr. F	-0,07	0,08	6,11	-0,98	0,9628
Positif congr. F - Positif incongr. F	-0,01	0,02	765,32	-0,45	0,9998
Positif congr. F - Négatif congr. H	-0,12	0,09	13,70	-1,31	0,8803
Positif congr. F - Positif congr. H	-0,06	0,06	39,56	-1,00	0,9711
Positif congr. F - Négatif incongr. H	-0,13	0,09	13,10	-1,42	0,8325
Positif congr. F - Positif incongr. H	-0,05	0,06	48,12	-0,81	0,9919
Négatif incongr. F - Positif incongr. F	0,06	0,08	6,35	0,83	0,9838
Négatif incongr. F - Négatif congr. H	-0,05	0,06	47,26	-0,82	0,9913
Négatif incongr. F - Positif congr. H	0,02	0,09	13,23	0,18	1
Négatif incongr. F - Négatif incongr. H	-0,06	0,06	41,91	-1,00	0,9723
Négatif incongr. F - Positif incongr. H	0,03	0,09	14,16	0,27	1
Positif incongr. F - Négatif congr. H	-0,11	0,09	13,98	-1,19	0,9209
Positif incongr. F - Positif congr. H	-0,05	0,06	42,90	-0,81	0,9919
Positif incongr. F - Négatif incongr. H	-0,12	0,09	13,41	-1,30	0,883
Positif incongr. F - Positif incongr. H	-0,04	0,06	48,77	-0,63	0,9982
Négatif congr. H - Positif congr. H	0,07	0,08	6,72	0,84	0,9829
Négatif congr. H - Négatif incongr. H	-0,01	0,03	760,31	-0,32	1
Négatif congr. H - Positif incongr. H	0,07	0,08	7,27	0,94	0,9711
Positif congr. H - Négatif incongr. H	-0,07	0,08	6,24	-0,98	0,9629
Positif congr. H - Positif incongr. H	0,01	0,03	763,16	0,30	1
Négatif incongr. H - Positif incongr. H	0,08	0,08	6,97	1,06	0,9464
Tâche 2 - RT : Valence*Congruence*Genre					
Contraste	β	SE	df	t	p-value
Négatif congr. F - Neutre congr. F	0,11	0,06	6,57	1,83	0,6222
Négatif congr. F - Négatif incongr. F	0,01	0,02	763,80	0,34	1
Négatif congr. F - Neutre incongr. F	0,10	0,06	7,07	1,75	0,666
Négatif congr. F - Négatif congr. H	-0,09	0,04	63,71	-2,05	0,4593
Négatif congr. F - Neutre congr. H	0,02	0,07	12,46	0,29	1
Négatif congr. F - Négatif incongr. H	-0,06	0,04	50,41	-1,39	0,8574
Négatif congr. F - Neutre incongr. H	0,06	0,07	13,96	0,84	0,9868
Neutre congr. F - Négatif incongr. F	-0,10	0,06	6,23	-1,73	0,6742
Neutre congr. F - Neutre incongr. F	0,00	0,02	760,22	-0,15	1
Neutre congr. F - Négatif congr. H	-0,20	0,07	13,28	-2,83	0,1662

Neutre congr. F - Neutre congr. H	-0,09	0,04	43,68	-2,15	0,3998
Neutre congr. F - Négatif incongr. H	-0,17	0,07	11,99	-2,42	0,311
Neutre congr. F - Neutre incongr. H	-0,05	0,04	60,72	-1,07	0,9609
Négatif incongr. F - Neutre incongr. F	0,10	0,06	6,81	1,64	0,7206
Négatif incongr. F - Négatif congr. H	-0,10	0,04	59,78	-2,25	0,3398
Négatif incongr. F - Neutre congr. H	0,01	0,07	12,04	0,18	1
Négatif incongr. F - Négatif incongr. H	-0,07	0,04	45,43	-1,61	0,7445
Négatif incongr. F - Neutre incongr. H	0,05	0,07	13,52	0,75	0,9934
Neutre incongr. F - Négatif congr. H	-0,20	0,07	14,05	-2,75	0,1858
Neutre incongr. F - Neutre congr. H	-0,08	0,04	52,48	-1,98	0,5066
Neutre incongr. F - Négatif incongr. H	-0,16	0,07	12,74	-2,34	0,3434
Neutre incongr. F - Neutre incongr. H	-0,04	0,05	66,85	-0,97	0,977
Négatif congr. H - Neutre congr. H	0,11	0,06	7,77	1,83	0,6217
Négatif congr. H - Négatif incongr. H	0,03	0,03	764,67	1,21	0,9277
Négatif congr. H - Neutre incongr. H	0,15	0,06	8,71	2,42	0,336
Neutre congr. H - Négatif incongr. H	-0,08	0,06	6,67	-1,34	0,8589
Neutre congr. H - Neutre incongr. H	0,04	0,03	760,52	1,47	0,8236
Négatif incongr. H - Neutre incongr. H	0,12	0,06	7,93	1,94	0,5657
Tâche 3 - RT : Valence*Congruence*Genre					
Contraste	β	SE	df	t	p-value
Neutre congr. F - Positif congr. F	0,01	0,10	6,13	0,12	1
Neutre congr. F - Neutre incongr. F	-0,02	0,02	789,62	-0,87	0,9885
Neutre congr. F - Positif incongr. F	0,01	0,10	6,19	0,08	1
Neutre congr. F - Neutre congr. H	-0,07	0,05	52,86	-1,51	0,7983
Neutre congr. F - Positif congr. H	-0,06	0,11	8,56	-0,56	0,9987
Neutre congr. F - Neutre incongr. H	-0,08	0,05	50,26	-1,67	0,7034
Neutre congr. F - Positif incongr. H	-0,05	0,11	8,68	-0,46	0,9996
Positif congr. F - Neutre incongr. F	-0,03	0,10	5,99	-0,31	1
Positif congr. F - Positif incongr. F	0,00	0,02	786,99	-0,22	1
Positif congr. F - Neutre congr. H	-0,08	0,11	8,55	-0,79	0,9893
Positif congr. F - Positif congr. H	-0,07	0,05	45,70	-1,55	0,7737
Positif congr. F - Neutre incongr. H	-0,09	0,11	8,42	-0,86	0,9832
Positif congr. F - Positif incongr. H	-0,06	0,05	49,45	-1,32	0,8879
Neutre incongr. F - Positif incongr. F	0,03	0,10	6,06	0,26	1
Neutre incongr. F - Neutre congr. H	-0,05	0,05	49,37	-1,15	0,9427
Neutre incongr. F - Positif congr. H	-0,04	0,11	8,40	-0,39	0,9999
Neutre incongr. F - Neutre incongr. H	-0,06	0,05	44,77	-1,32	0,8848
Neutre incongr. F - Positif incongr. H	-0,03	0,11	8,52	-0,29	1
Positif incongr. F - Neutre congr. H	-0,08	0,11	8,63	-0,74	0,9924
Positif incongr. F - Positif congr. H	-0,07	0,05	48,35	-1,43	0,8365
Positif incongr. F - Neutre incongr. H	-0,09	0,11	8,50	-0,81	0,9875
Positif incongr. F - Positif incongr. H	-0,06	0,05	51,55	-1,21	0,9262
Neutre congr. H - Positif congr. H	0,01	0,10	6,39	0,13	1
Neutre congr. H - Neutre incongr. H	-0,01	0,03	789,56	-0,26	1
Neutre congr. H - Positif incongr. H	0,02	0,10	6,47	0,23	1
Positif congr. H - Neutre incongr. H	-0,02	0,10	6,25	-0,20	1

Positif congr. H - Positif incongr. H	0,01	0,03	787,17	0,38	0,9999
Neutre incongr. H - Positif incongr. H	0,03	0,10	6,39	0,29	1
Tâche 1, 2 et 3 - RT : Congruence*Genre*Empathie					
Contraste	β	SE	df	t	p-value
congr. F basse - incongr. F basse	0,00	0,01	2313,87	0,31	1
congr. F basse - congr. H basse	-0,07	0,04	111,48	-1,74	0,6645
congr. F basse - incongr. H basse	-0,07	0,04	112,81	-1,71	0,6822
congr. F basse - congr. F haute	-0,04	0,03	111,58	-1,25	0,9149
congr. F basse - incongr. F haute	-0,06	0,03	113,42	-1,77	0,642
congr. F basse - congr. H haute	-0,14	0,04	110,67	-3,51	0,0147
congr. F basse - incongr. H haute	-0,12	0,04	111,80	-3,02	0,0596
incongr. F basse - congr. H basse	-0,07	0,04	113,89	-1,84	0,5967
incongr. F basse - incongr. H basse	-0,07	0,04	113,80	-1,82	0,6112
incongr. F basse - congr. F haute	-0,05	0,03	114,38	-1,37	0,8699
incongr. F basse - incongr. F haute	-0,06	0,03	114,82	-1,89	0,5612
incongr. F basse - congr. H haute	-0,15	0,04	112,74	-3,60	0,011
incongr. F basse - incongr. H haute	-0,13	0,04	112,74	-3,12	0,0454
congr. H basse - incongr. H basse	0,00	0,01	2301,70	0,06	1
congr. H basse - congr. F haute	0,02	0,03	111,29	0,71	0,9964
congr. H basse - incongr. F haute	0,01	0,03	113,63	0,19	1
congr. H basse - congr. H haute	-0,07	0,04	112,23	-1,80	0,6197
congr. H basse - incongr. H haute	-0,06	0,04	113,58	-1,33	0,8836
incongr. H basse - congr. F haute	0,02	0,03	113,20	0,69	0,9972
incongr. H basse - incongr. F haute	0,01	0,03	113,34	0,17	1
incongr. H basse - congr. H haute	-0,08	0,04	113,49	-1,82	0,6092
incongr. H basse - incongr. H haute	-0,06	0,04	113,49	-1,36	0,875
congr. F haute - incongr. F haute	-0,02	0,01	2315,26	-1,66	0,7132
congr. F haute - congr. H haute	-0,10	0,04	111,75	-2,69	0,1369
congr. F haute - incongr. H haute	-0,08	0,04	113,60	-2,16	0,3838
incongr. F haute - congr. H haute	-0,08	0,04	113,71	-2,19	0,3655
incongr. F haute - incongr. H haute	-0,06	0,04	113,62	-1,67	0,7045
congr. H haute - incongr. H haute	0,02	0,02	2305,99	1,23	0,9217

Annexe 19. Taux de réponses par feedback selon la valence, résultats de la tâche préliminaire (phase de construction 1)

Expérimentation préliminaire n°1 :

		Choix du feedback		
		<i>Ah oui</i>		<i>Oh non</i>
		<i>Compte</i>	<i>%</i>	<i>Compte</i>
Valence de la nouvelle	<i>Positive</i>	497	94%	31
	<i>Négative</i>	101	19%	427

Expérimentation préliminaire n°2 :

		Choix du feedback		
		<i>J'écoute</i>		<i>Oh non</i>
		<i>Compte</i>	<i>%</i>	<i>Compte</i>
Valence de la nouvelle	<i>Neutre</i>	454	88%	62
	<i>Négative</i>	268	52%	248

Expérimentation préliminaire n°3 :

		Choix du feedback		
		<i>J'écoute</i>		<i>Ah oui</i>
		<i>Compte</i>	<i>%</i>	<i>Compte</i>
Valence de la nouvelle	<i>Neutre</i>	404	69%	184
	<i>Positive</i>	306	52%	282