



HAL
open science

Dominance, hostilité et expressivité vocale dans le débat politique étude perceptive et acoustique du conseil municipal de Montreuil (93100)

Charlotte Kouklia

► To cite this version:

Charlotte Kouklia. Dominance, hostilité et expressivité vocale dans le débat politique étude perceptive et acoustique du conseil municipal de Montreuil (93100). Linguistique. Université Sorbonne Paris Cité, 2019. Français. NNT : 2019USPCA013 . tel-02444742

HAL Id: tel-02444742

<https://theses.hal.science/tel-02444742v1>

Submitted on 19 Jan 2020

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

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

UNIVERSITÉ SORBONNE PARIS CITÉ
UNIVERSITÉ SORBONNE NOUVELLE PARIS 3
ED 268 : « Langage et langues : description, théorisation, transmission »
Laboratoire de Phonétique et de Phonologie, UMR 7018

Thèse de doctorat en phonétique

Charlotte KOUKLIA

**Dominance, hostilité et expressivité vocale dans le débat politique :
Étude perceptive et acoustique du conseil municipal de Montreuil (93100)**

Thèse dirigée par
Mme Jacqueline VAISSIERE et M. Nicolas AUDIBERT

Soutenue le 19 janvier 2019

Jury :

Mme Anne-Catherine Simon (pré-rapporteure), Professeure, Université catholique de Louvain
Monsieur Rudolph Sock (pré-rapporteur) Professeur, Université de Strasbourg
Mme Véronique Aubergé (examinatrice), Chargée de Recherche CNRS, Université de Grenoble

Mme Jacqueline Vaissière (directrice), Professeure émérite, Université Sorbonne Nouvelle
M. Nicolas Audibert (co-directeur), Maître de Conférences, Université Sorbonne Nouvelle

Résumé court (français)

Dominance, hostilité et expressivité vocale dans le débat politique : Une étude perceptive et acoustique du conseil municipal de Montreuil (93100)

Cette thèse étudie les rapports entre l'expression d'attitudes agressives contrôlées et la perception de la dominance, à partir d'extraits de séances télévisées du conseil Municipal de Montreuil (93100) durant l'année 2013, période marquée par un climat politique vif et hostile. Un corpus a été constitué à partir d'extraits de parole spontanée de la Maire, Dominique Voynet, et de quatre de ses opposants. Les cinq locuteurs ont participé à l'enregistrement d'une relecture neutre du corpus de leurs propres extraits de parole (25 stimuli par locuteur) ainsi qu'à l'auto-évaluation perceptive de leurs stimuli (profils émotionnels), dont les résultats ont été comparés à l'évaluation perceptive des extraits par des auditeurs naïfs. Tous les locuteurs ont été perçus comme légèrement hostiles, légèrement mécontents et légèrement dominants malgré des profils émotionnels auto-évalués variables. Les extraits originaux et relus ont été comparés au niveau de leur structuration prosodico-syntaxique et de leurs caractéristiques temporelles et mélodiques. Les résultats montrent que 1) certains locuteurs semblent plus s'appuyer sur des paramètres mélodiques et d'autres sur des paramètres temporels ; 2) on peut néanmoins dégager les tendances générales concernant les corrélats dans la parole de l'hostilité et de la dominance dans notre corpus : a) des écarts entre structuration syntaxique et prosodique des extraits, b) la réduction ou l'absence d'allongements syllabiques finaux pré-pausaux, c) de fortes variations de plage de variation de F0 de part et d'autre des pauses silencieuses.

Mots-clés : parole expressive, parole politique, agressivité, perception, pauses silencieuses, structuration prosodico-syntaxique.

Abstract (english)

Dominance, hostility and vocal expressivity in political debate: A perceptual and acoustic study of the Montreuil Municipal Council (93100)

This dissertation studies the relationship between the expression of controlled aggressive attitudes and the perception of dominance, based on extracts from televised sessions of the Municipal Council of Montreuil during 2013; a period marked by a lively and hostile political climate. We constituted a corpus of spontaneous speech extracts from the Mayor, Dominique Voynet, and four of her opponents. During subsequent recording sessions, the five speakers were asked to read transcriptions of their own speech extracts in a neutral tone (25 stimuli per speaker). They also participated in a self-evaluation questionnaire that focused on the perception of emotional profiles in their own stimuli. The results of the speakers' perceptive evaluations were compared to those of a control group of naïve listeners. All speakers were perceived as slightly hostile, discontent and dominant, despite the variability observed in the self-evaluated profiles. The original and re-read extracts were compared in their prosodic-syntactic structure as well as their temporal and melodic characteristics. We show that: 1) some speakers seem to rely mostly on melodic parameters whereas others primarily use temporal parameters, 2) nevertheless, general trends emerge regarding the speech correlates of hostility and dominance in our corpus, notably: a) discrepancies between the syntactic and the prosodic structure of the extracts, b) reduction or absence of pre-pausal final syllabic lengthening, c) large variations in F0 range on both sides of silent pauses.

Keywords: expressive speech, political speech, aggressiveness, perception, silent pauses, prosodic-syntactic structures.

*À ma famille, et ma famille de cœur
Pour Nicolas Combet, sans qui...*

Remerciements

Je tiens à remercier en premier lieu mes deux directeurs de thèse, pour leur patience et leur présence dans l'accomplissement de la présente recherche doctorale.

Merci à ma directrice Jacqueline Vaissière, pour sa perspective unique sur notre domaine, sachant allier érudition et pédagogie, rigueur et curiosité. L'apprentissage que j'ai pu faire à vos côtés des caractéristiques phonétiques par leur fonctionnalité dans la chaîne de la communication parlée ont eu sur moi un impact indélébile. Merci de m'avoir transmis votre vision appliquée de la phonétique.

Merci à mon directeur Nicolas Audibert, pour toutes ces années de travail, de recherche, et de débats. N'importe qui ayant eu la chance de travailler avec toi a pu connaître ton implication rare pour ne pas dire unique dans les projets. J'aurai eu pour ma part la chance d'être ta première thésarde et d'avoir pu construire cette recherche avec toi et je t'en remercie sincèrement.

Merci à Dominique Voynet, Gaylord Le Chequer, Frédéric Molossi, Jean-Jacques Serey et Alexandre Tuillon pour leur participation essentielle dans cette thèse, et merci à tous les conseillers municipaux qui ont accepté d'échanger avec moi dans la phase préliminaire de cette étude.

Je remercie chaleureusement Anne-Catherine Simon, Rudolph Sock et Véronique Aubergé, d'avoir accepté de faire partie de mon jury.

J'envoie ici une pensée et mes respects à Isabella Poggi, pour l'inspiration apporté par ses recherches et son intérêt aux miennes.

Merci à Nicolas Combet et Alice Barbaza pour avoir été mes piliers.

Merci à Hélène, Adrian et Tesla.

Merci aux Longs Quartiers et aux douces rues quadrillées tout autour.

Merci à Maria Candea. Merci à toi Maria pour ton acuité et ta bienveillance unique.

Merci à Lucien Tisserand, mon *wingman*, je serai là quand tu termineras la tienne de thèse.

Merci à Jean-Louis Young pour les années de travail à ses côtés dans la salle informatique de l'ED 268, j'en garde un souvenir si agréable, je n'ai pas vu le temps passer.

Merci à Aron Arnold et Gregory Miras qui ont su me soutenir en me faisant rire de tout.

Merci à Mélanie Lancien et Daria D'Allessandro pour m'avoir soutenue au moment le plus important, et surtout merci d'être mes amies.

Merci à tous les copains du LPP et leur bienveillance, leur gentillesse, je pense ici en particulier à Bowei, Dji Dji, Jane qui est docteure maintenant, Emre qui toujours été là pour parler. J'embrasse aussi bien fort ici Sachie sur un air de Kate Bush et des souvenirs de Françoise.

Merci à mes relecteurs et amis chéris Hélène Jolivet, Guillaume Lamaze, Thomas Carpentier, Laurent Couret, Pauline Berger, Marine Ventura, Joseph Sainderichin, Florence Watrin, Lise Trinquand et Marine Normand.

Merci à Nora Fangel-Gustavson et Tessa Namias pour être venues souvent m'apporter de l'oxygène dans mon sous-marin de travail.

Merci à mes sœurs chéries Inès, Michèle et Louann et merci à Jérôme.

Merci à Rémi Gérard pour notre aventure parallèle.

Merci à Pascal et Eva et leur émission de radio « The Brain » qui m'a tenue compagnie pendant de longues et nombreuses nuits sans sommeil, à grands renforts de BPM enfiévrés.

Merci à tous les amis d'un jour et de toujours avec qui j'ai pu échanger au sujet de ce travail, qui a été la star de ma vie pendant ces dernières années, de Tina Thill à Tom et Théo.

Enfin, un grand merci à Lucas Trostaniecki pour m'avoir montré ce fameux débat entre Jérôme Guedj et Henri Guaino qui a été le déclencheur pour projeter mes problématiques de recherche de Master sur les rapports entre agressivité et dominance dans le domaine du politique.

Surtout, merci à mes merveilleux parents pour leur foi en moi.

Sommaire

Résumé court (français)	iii
Abstract (english).....	v
Remerciements.....	ix
Sommaire	xi
I. Introduction	1
II. Théories et pratiques pour l'étude de la parole expressive en politique.....	5
II.1 De l'émotion brute universelle à l'approche fonctionnelle socioconstructiviste	6
II.1.1. L'origine universelle des émotions fondamentales.....	6
II.1.2. Vers une considération de la perception et de l'évaluation émotionnelle.....	7
II.1.3. L'approche socioconstructiviste.....	11
II.1.4. Vers une approche fonctionnaliste de l'émotion.....	12
II.2 Comprendre les enjeux de la communication politique	14
II.2.1 La communication politique, une affaire d'image de soi (ethos).....	14
II.2.2 Signaux de dominance et agressivité dans le débat politique	16
II.2.3 L'attaque à l'image de l'autre, stratégie de dominance dans le débat.....	17
II.3 Débat politique et expressivité vocale : Caractéristiques de parole attendues.....	19
II.3.1 Introduction à l'approche phonétique	19
II.3.2 Le phonostyle de la parole politique et les enjeux du débat.....	20
II.3.3 Parole politique et dominance, de part et d'autre de la pause silencieuse	22
II.3.4 Les caractéristiques phonétiques des attitudes agressives contrôlées.....	28
III. Protocole de recueil et traitement des données.....	33
III.1. Description du conseil municipal de Montreuil en 2013	34
III.1.1. Un contexte politique tendu	34
III.1.2. Une interaction régulée	34
III.1.3. Déroulement type d'une séance	35
III.2. Des archives brutes à la constitution du corpus et sa condition de contrôle....	36
III.2.1. L'annotation large des séances étudiées (24h d'archives).....	36
III.2.2. La sélection des locuteurs	38
III.3. L'enregistrement de la condition de contrôle (corpus relecture).....	41
III.3.1. Préparation des données.....	41
III.3.2. Lieu et déroulement général des entretiens	44
III.3.3. Consignes et passation de la tâche de relecture	46
III.4. Segmentation et annotation lexico-syntaxique des données.....	46

III.4.1.	Segmentation et annotation segmentale.....	48
III.4.2.	Annotation syntaxique	48
IV.	Perception de l'hostilité et profils dimensionnels.....	53
IV.1.	Enjeux de l'analyse perceptive et choix méthodologiques	54
IV.2.	Principes pour la présentation des analyses statistiques	55
IV.3.	Étude n°1 : Hostilité perçue et estimation du destinataire du message	57
IV.3.1.	Protocole méthodologique	58
IV.3.2.	Résultats de l'étude de l'hostilité perçue	59
IV.3.3.	Résultats de l'estimation du destinataire supposé.....	63
IV.3.4.	Corrélations entre hostilité perçue et estimation du destinataire	67
IV.4.	Étude n°2 : Profils émotionnels (valence, excitation et contrôle/dominance). 68	
IV.4.1.	Protocole méthodologique de l'étude n°2.....	68
IV.4.2.	Résultats pour la valence perçue et auto-attribuée.....	70
IV.4.3.	Résultats pour l'excitation perçue et auto-évaluée	74
IV.4.4.	Résultats pour le contrôle de soi auto-évalué et la dominance perçue	77
IV.4.5.	Inter-corrélations entre dimensions perçues (auditeurs « naïfs »)	81
IV.4.6.	Corrélations entre dimensions auto-évaluées.....	83
IV.4.7.	Corrélations entre mesures de différences (<i>dVal</i> , <i>dExc</i> , <i>dDom</i>)	83
IV.5.	Analyse des rapports entre hostilité perçue et profils dimensionnels	84
IV.5.1.	Rapports entre hostilité et valence	84
IV.5.2.	Rapports entre hostilité et excitation.....	85
IV.5.3.	Corrélations entre hostilité, contrôle de soi et dominance	86
V.	Délimitation des unités d'analyse par les frontières prosodiques majeures.....	88
V.1.	Critères pour l'annotation des frontières prosodiques majeures.....	89
V.1.1.	Protocole d'annotation des frontières	89
V.1.2.	Adaptation de l'annotation dans les cas de schwas épenthétiques	91
V.1.3.	Typologie des frontières prosodiques annotées	92
V.2.	Résultats de l'annotation prosodique paramètre par paramètre.....	93
V.2.1.	Frontières prosodiques majeures annotées.....	93
V.2.2.	Frontières prosodiques marquées par des pauses.....	94
V.2.3.	Frontières marquées seulement par la présence d'une pause.....	95
V.2.4.	Frontières prosodiques marquées par F0	96
V.2.5.	Frontières prosodiques marquées par la durée.....	98
V.2.6.	Typologie multiparamétrique des frontières prosodiques majeures	99

V.3.	Annotation des contours mélodiques aux frontières prosodiques majeures	101
V.3.1.	Présentation générale des contours mélodiques annotés	103
V.3.2.	Nombre de contours montants par stimulus	105
V.3.3.	Nombre de contours en cloche.....	105
V.3.4.	Nombre de contours type focus	105
V.3.5.	Nombre de contours plats	106
V.3.6.	Nombre de contours descendants.....	107
V.3.7.	Nombre de contours terminaux.....	107
VI.	Analyses comparées entre extraits originaux et relus : pauses, durées et F0.....	108
VI.1.	Enjeux et principes méthodologiques pour l'analyse prosodique.....	109
VI.1.1.	Présentation de la méthodologie pour les statistiques.....	110
VI.2.	Analyses des pauses silencieuses.....	112
VI.2.1.	Nombre et durées des pauses silencieuses	112
VI.2.2.	Proportion de silence par stimulus.....	114
VI.2.3.	Durée relative des pauses silencieuses de type « frontière »	114
VI.3.	Analyses du débit et du rythme syllabique	116
VI.3.1.	Débit articulatoire moyen (syllabes).....	116
VI.3.2.	Variations de débit articulatoire.....	117
VI.4.	Analyses de la fréquence fondamentale (F0).....	122
VI.4.1.	Fréquence fondamentale (F0) moyenne.....	122
VI.4.2.	Variations de la F0 moyenne	125
VI.4.3.	Plage de variation de la fréquence fondamentale (F0 min – F0 max)	128
VI.5.	Inter-corrélations entre variables mélodiques et temporelles	132
VI.5.1.	Inter-corrélations entre variables temporelles.....	132
VI.6.	Inter-corrélations entre variables mélodiques.....	136
VI.6.1.	Moyenne (Hertz) et plage de variation (demi-tons) de F0.....	136
VI.7.	Inter-corrélations entre variables mélodiques et temporelles	137
VI.7.1.	Frontières marquées par la durée et plage de variation de F0.....	137
VI.7.2.	Débit articulatoire et F0 moyenne.....	137
VII.	Confrontation entre structurations prosodique et syntaxique.....	139
VII.1.	Typologie des frontières prosodico-syntaxiques analysées	140
VII.2.	Présentation des résultats	142
VII.2.1.	Répartition générale des frontières prosodico-syntaxiques	142
VII.3.	Distributions par stimulus	145

VII.3.1.	Frontières prosodiques inter-unités syntaxiques majeures	145
VII.3.2.	Frontières prosodiques inter-unités syntaxiques intermédiaires	145
VII.3.3.	Frontières prosodiques intra-unités syntaxiques	147
VII.3.4.	Proéminences prosodiques intra-mots	148
VII.4.	Localisation des pauses « frontières » dans la structure syntaxique	149
VII.4.1.	Pauses « frontières » inter-unités syntaxiques majeures	149
VII.4.2.	Pauses « frontières » inter-unités syntaxiques intermédiaires	149
VII.4.3.	Pauses « frontières » intra-syntaxiques par stimulus	150
VII.5.	Caractéristiques prosodiques selon la position syntaxique des frontières	152
VII.5.1.	Caractéristiques temporelles	152
VII.5.2.	Caractéristiques mélodiques	155
VIII.	Corrélat acoustique de l'hostilité, de la dominance, de la valence et de l'excitation.	156
VIII.1.	Présentation de la méthodologie d'analyse	157
VIII.2.	Analyses des corrélats de l'hostilité perçue	158
VIII.2.1.	Hostilité perçue et frontières prosodiques majeures	158
VIII.2.2.	Hostilité perçue et caractéristiques de F0	160
VIII.2.3.	Hostilité perçue et caractéristiques temporelles	161
VIII.2.4.	Hostilité perçue et structuration prosodico-syntaxique	164
VIII.3.	Corrélat de la dominance perçue	166
VIII.3.1.	Dominance perçue et frontières prosodiques majeures	166
VIII.3.2.	Dominance perçue et caractéristiques de F0	169
VIII.3.3.	Dominance perçue et contours mélodiques annotés	170
VIII.3.4.	Dominance perçue et caractéristiques temporelles	172
VIII.3.5.	Dominance perçue et structuration prosodico-syntaxique	173
VIII.5.	Corrélat de la valence perçue	174
VIII.5.1.	Valence perçue et frontières prosodiques majeures	174
VIII.5.2.	Valence perçue et structuration prosodico-syntaxique	175
VIII.5.3.	Valence perçue et caractéristiques de F0	176
VIII.5.4.	Valence perçue et contours mélodiques annotés	176
VIII.5.5.	Corrélat temporel de la valence perçue	177
VIII.6.	Corrélat de l'excitation perçue	179
VIII.6.1.	Excitation perçue et frontières prosodiques majeures	179
VIII.6.2.	Excitation perçue et caractéristiques de F0	181
VIII.6.3.	Excitation perçue et contours mélodiques annotés	182

VIII.6.4.	Excitation perçue et caractéristiques temporelles	183
VIII.6.5.	Excitation perçue et structuration prosodico-syntaxique.....	184
IX.	Discussion et Conclusion.....	185
IX.1.	Expressivité vocale et perception des personnes politiques en débat.....	186
IX.2.	Avantages et limites de la condition de contrôle (corpus relecture).....	191
IX.3.	Argument une analyse multifactorielle de la pause silencieuse.....	197
IX.4.	Vers une approche plurielle et pluridisciplinaire de l'expressivité vocale	201
IX.5.	Principaux apports de cette étude	204
	Références bibliographiques	205
	Table des exemples	213
	Table des tableaux.....	215
	Table des figures	219
	Annexes.....	227

I. Introduction

S'il est le théâtre de tensions et de désaccords, le débat politique est aussi l'endroit idéal pour que chacun de ses intervenants travaille à construire une image de lui-même (Charaudeau, 2015). Dans la rhétorique aristotélicienne, pour convaincre et avoir des chances d'être élu, un politicien se doit de véhiculer l'image (*ethos*) d'une personne morale, bienveillante et compétente. Pour Isabella Poggi et Francesca D'Errico (2010), la personne politique devra également apparaître comme ayant plus de pouvoir que l'autre, qui plus est en situation de débat. Cet ajout de la dominance aux dimensions de l'*ethos* s'explique selon ces chercheuses en psychologie sociale par le fait qu'on donnera difficilement son vote à une personne politique qui n'arrive pas à prendre l'ascendant sur l'autre, et ce malgré son apparente honnêteté et ses compétences.

Dans leur étude sur les signaux de dominance¹ dans le débat, Poggi & D'Errico (2010) indiquent que ceux-ci seraient entre autres constitués d'une forme visible de maîtrise de soi (contrôle des émotions) et d'autre part d'hostilité envers l'opposant ; la combinaison entre ces deux modalités comportementales contradictoires donnant fréquemment lieu à l'expression d'attitudes agressives contrôlées telles que le mépris, le sarcasme, ou l'irritation. L'objectif premier de notre recherche consiste à étudier, via la combinaison d'analyses perceptives et phonétiques, la façon dont s'exprime, sur le plan phonétique, l'expressivité vocale dans les rapports de force inhérents au débat politique. En cherchant à déterminer la composante vocale de l'expression de telles attitudes, nous nous interrogeons principalement sur l'influence de l'émergence de caractéristiques agressives sur la perception de la dominance.

Le second objectif de notre étude est méthodologique et réside dans la mise en place d'un protocole expérimental adapté à l'étude phonétique de l'expressivité dans la parole politique, champ d'étude paradoxalement marginal. En effet, malgré la place importante qu'occupent les caractéristiques vocales dans la fonction expressive du langage (Jakobson, 1963), la parole expressive est paradoxalement peu traitée en analyse politique. Une telle lacune peut, entre autres, s'expliquer par le fait que l'étude de la parole expressive produite en contexte d'interaction comporte de nombreuses limites d'ordre méthodologique.

¹ Poggi & D'Errico se réfèrent spécifiquement à la notion de « dominance » et non à celle plus large de « pouvoir » et justifient ce choix par le fait que « dominer » signifie sur le plan pragmatique « avoir plus de pouvoir » que l'opposant dans le débat politique.

Nous traiterons principalement dans la présente étude de deux problèmes méthodologiques fondamentaux pour l'analyse phonétique appliquée de corpus de parole expressive naturelle : la validation perceptive des données et leur comparabilité, garantes de la fiabilité des analyses phonétiques systématiques menées par la suite. La validation perceptive des données, indispensable à toute étude phonétique, n'est en effet pas facilitée par la nature variable et non contrôlée des énoncés d'un corpus de parole spontanée, notamment sur le plan du contenu sémantique. Lorsque l'on travaille sur de la parole expressive produite en contexte de communication, il est effectivement impératif de prendre en compte le fait que le message linguistique (contenu sémantique) est souvent lui aussi porteur d'expressivité. Outre la validation perceptive, la démarche classique dans l'analyse phonétique de l'expressivité vocale consiste à quantifier l'écart perçu et produit entre un extrait de parole expressive et une norme supposée neutre ou non-expressive, en effectuant des comparaisons systématiques sur un matériel linguistique similaire. En parole naturelle, du fait de la plus forte hétérogénéité des énoncés et complexité de leurs structures syntaxiques comparées à de la parole de laboratoire (Blanche-Benveniste, 2010), il est impossible de contrôler le contenu phonétique, syllabique, lexical, syntaxique et sémantique et ainsi d'effectuer des analyses comparatives systématiques à chacun de ces niveaux.

Nous proposons pour pallier aux difficultés énoncées une approche méthodologique adaptée reposant d'une part sur la mise en place d'évaluations perceptives plurielles visant notamment à quantifier l'impact de l'expressivité inhérente au contenu sémantique par rapport à l'expressivité vocale, et d'autre part sur l'enregistrement a posteriori par les locuteurs d'une condition de contrôle, composée de relectures calmes des transcriptions orthographiques de leurs extraits originaux.

Enfin, le troisième objectif de cette thèse, à proprement parler phonétique, est de proposer une analyse multiparamétrique des caractéristiques prosodiques relatives à l'insertion de pauses silencieuses, notamment dans la parole politique expressive. Les analyses phonétiques présentées ici s'inscrivent dans la continuité des recherches menées par Duez (1991, 1999) sur les caractéristiques temporelles de la parole politique, et notamment des rapports entre pauses silencieuses et accession au pouvoir. Pour Danielle Duez, au-delà de leur fonction structurante du discours, la fonction primordiale des pauses en parole politique est pragmatique : il s'agit là principalement d'un signal de dominance du locuteur qui montre ainsi à l'auditoire ou l'interlocuteur qu'il ne craint pas d'être interrompu du fait de la supériorité de son statut.

Nous souhaitons dans cette thèse confronter les observations de Duez (1991, 1999, 2003) à un corpus de débat politique (situation n'ayant été traitée que succinctement par Duez) ainsi qu'à la réalité perceptive du pouvoir des locuteurs, en testant notamment les rapports entre la dominance perçue par une cohorte d'auditeurs et les caractéristiques pausales, mélodiques et temporelles de part et d'autres des pauses silencieuses.

Nous détaillons à présent, à la fin de ce **chapitre I** d'introduction, le plan de la thèse et des différentes parties qui y sont développées pour répondre à la question des caractéristiques vocales pertinentes dans la perception de locuteurs comme dominants et/ou hostiles dans le débat politique.

Le **chapitre II** est consacré à l'explicitation d'un ensemble de concepts nécessaires à la bonne compréhension des enjeux et problématiques propres à l'étude phonétique de l'expressivité vocale dans le débat politique. Nous détaillons d'abord les principes théoriques relatifs à l'expression des émotions, de l'approche évolutionniste de Darwin (1872) aux théories socioconstructivistes de Averill (1980). Nous introduisons ensuite la question des enjeux de la communication politique. En partant de concepts issus de la rhétorique jusqu'à son héritage dans des théories contemporaines en socio-cognitivism, nous présentons un ensemble de stratégies de dominance attendues dans le cas du débat politique dont les deux composantes majeures et contradictoires sont la démonstration d'une maîtrise de soi et la manifestation tacite ou explicite d'hostilité. Nous fournissons alors une description des caractéristiques prosodiques d'une part relatives au signalement en politique de la dominance à travers une approche centrée sur **la pause silencieuse** (Duez, 1991, 1999), et d'autre part propres à l'expression d'attitudes agressives contrôlées.

Le **chapitre III** présente le corpus et notre protocole méthodologique d'acquisition et de traitement du corpus original et de sa contrepartie relue (corpus relecture). À partir de 24 heures d'archives, une première sélection de 200 extraits (40 extraits pour chacun des cinq locuteurs) a été soumise à validation perceptive, notamment en terme d'hostilité perçue (voir Chapitre **IV**). Un sous-ensemble de 125 stimuli (25 par locuteur) a été sélectionné et a fait l'objet d'une relecture par chaque locuteur de ses propres extraits transcrits orthographiquement, dans le cadre d'entretiens individuels auxquels ils ont bien voulu participer. Les données ont été segmentées à différents niveaux traditionnels, en phonèmes, syllabes et mots. Une annotation parallèle des frontières syntaxiques et prosodiques a également été effectuée selon les recommandations de Degand et Simon (2011) pour permettre ultimement une prise en compte des frontières prosodico-syntaxiques majeures.

Dans le **chapitre IV**, nous présentons le protocole expérimental mis en place et les résultats obtenus des **analyses perceptives** mises en place sur le corpus original. Nous présentons dans un premier temps les résultats de l'étude n°1 concernant la perception de l'hostilité et du destinataire présumé (interlocuteur particulier vs. auditoire) effectuée avec 22 participants et de l'étude n°2 portant sur une évaluation en profils émotionnels tridimensionnels (valence, excitation, et dominance) renseignés en parallèle par 24 auditeurs et auto-évalués par les locuteurs du corpus durant les entretiens individuels effectués avec eux.

Les analyses phonétiques et prosodiques sont alors détaillées sur trois chapitres :

Dans le **chapitre V**, nous présentons le protocole d'annotation semi-automatique des frontières prosodiques majeures et de l'annotation qualitative des contours mélodiques aux frontières et les résultats de ces annotations. Dans le **chapitre VI**, nous analysons les caractéristiques prosodiques fines des locuteurs à travers un ensemble de variables relatives aux pauses silencieuses, au débit et aux variations de durées segmentales ainsi qu'aux caractéristiques mélodiques du corpus original et de ses différences avec le corpus relecture. Une dernière section présente les inter-corrélations significatives entre variables temporelles et mélodiques dans le corpus original.

Dans le **chapitre VII**, nous analysons les rapports entre prosodie et syntaxe à travers une analyse croisée des frontières prosodiques et des frontières syntaxiques annotées puis nous cherchons à déterminer les corrélats prosodiques des différents types de structuration prosodico-syntaxique (**VII.2**).

Dans le **chapitre VIII**, partie centrale de nos analyses, les résultats des analyses perceptives et prosodiques sont mis en relation dans le but d'estimer l'influence de caractéristiques prosodiques spécifiques sur la perception des locuteurs par les auditeurs comme hostiles, mécontents, dominants ou agités.

Enfin, dans le **chapitre IX**, nous présentons une discussion générale et une conclusion de la thèse. En écho à nos objectifs de départ, nous revenons sur les avantages et limites de notre protocole expérimental. Enfin, au regard des résultats obtenus, nous proposons un ensemble de recommandations pour engager à une analyse plus approfondie et répandue de la composante vocale de l'expressivité en parole politique, notamment au travers de la pause silencieuse.

II. Théories et pratiques pour l'étude de la parole expressive en politique

Dans ce chapitre, nous présentons un ensemble de concepts nécessaires à la bonne compréhension des enjeux et problématiques propres à l'étude phonétique de l'expressivité vocale dans le débat politique. Du fait du caractère intrinsèquement pluridisciplinaire de cet objet d'étude, nous présentons ici un ensemble de théories et méthodologies issues de domaines aussi divers que la psychologie, la sociologie, la rhétorique, la linguistique, et la phonétique.

Dans un premier temps (**II.1 De l'émotion universelle à l'approche fonctionnelle socioconstructiviste**) la question de l'émotion est abordée en partant de sa conception évolutionniste origine biologique résultante d'instincts de survie (Darwin, 1872) jusqu'à la considération de sa construction sociale (Averill, 1980).

Nous abordons ensuite (**II.2 Comprendre les enjeux de la communication politique**) les impératifs des personnes politiques pour véhiculer une image de soi (*ethos*) favorable à l'adhésion des électeurs. En partant de concepts issus de la rhétorique jusqu'à son héritage dans des théories contemporaines en socio-cognitivism, nous présentons un ensemble de stratégies attendues dans le cas du débat politique dont les deux composantes majeures et contradictoires sont la démonstration d'une maîtrise de soi et la manifestation tacite ou explicite d'hostilité.

Dans un troisième temps (**II.3 Les caractéristiques vocales de dominance et d'agressivité présentes dans le débat politique**), pour aborder les caractéristiques vocales auxquelles s'attendre, nous fournissons une description des caractéristiques prosodiques relative au signalement en politique de la **dominance** à travers une approche centrée sur la **pause silencieuse** (Duez, 1991, 1999), et d'autre part nous présentons alors une revue de la littérature concernant les caractéristiques vocales propres aux attitudes agressives contrôlées (**II.3.2**).

Dans leur ensemble, ces informations vont nous permettre d'introduire dans cette thèse l'analyse des caractéristiques vocales d'extraits de débats au travers des rapports entre signaux phonétiques de dominance en parole politique et l'expression d'attitudes agressives pouvant donner lieu à des caractéristiques notamment prosodiques très différentes, pour ne pas dire contradictoires.

II.1. De l'émotion brute universelle à l'approche fonctionnelle socioconstructiviste

II.1.1. L'origine universelle des émotions fondamentales

Charles Darwin, biologiste naturaliste anglais du 19^{ème} siècle est un des premiers chercheurs à théoriser l'émotion, ses origines et ses manifestations à travers les espèces. Pour Darwin (1872), les émotions ne sont pas spécifiques à l'espèce humaine mais universelles et s'expliquent par les mécanismes de la sélection naturelle. Selon cette conception, on peut considérer, à titre d'exemples non exclusifs, la colère comme réponse combative à un danger, la peur comme réponse par la fuite, la tristesse comme indication de la souffrance, le dégoût comme protection corporelle, etc. À partir de ses analyses, Darwin démontrera par une comparaison entre expressions émotionnelles animales et humaines des caractéristiques expressives fondamentales ou primaires liées aux principaux instincts de survie des espèces.

Les émotions primaires relèveraient des instincts de survie des espèces. Les émotions dites de base, ou fondamentales selon Paul Ekman (1992) à qui l'on doit une contribution majeure aux théories évolutionnistes de Darwin (1872) sont au nombre de six : **joie, tristesse, colère, peur, surprise, et le dégoût**¹. Au travers des études de Ekman (1992), ces six émotions, lorsque produites par des comédiens et prises en photographie sont perçues avec des taux de reconnaissance très élevés quelles que soit la langue et la culture des participants aux évaluations perceptives, conférant leur caractère universel. Ce caractère universel confère aux émotions fondamentales le statut d'émotions primaires. Si le nombre précis d'émotions considérées comme primaires varie selon les études (voir revue Cornelius, 1996), les critères pour les considérer comme telles reposent toujours selon une conception évolutionniste héritée de Darwin (1872). Pour Robert Plutchik (1980), elles sont notamment de l'ordre de huit ; ce dernier ajoute aux catégories déjà considérées par Darwin celles de **l'excitation** et de **l'attirance** (Plutchik, 1980).

Dans l'optique évolutionniste, les nombreuses autres émotions potentiellement exprimables découlent des émotions primaires (Ekman, 1992) ; par exemple l'anxiété étant affiliée à la peur, elle en est une émotion secondaire. Plutchik (1980) a également précisé que certaines émotions considérées comme secondaires peuvent être constituées d'un mélange entre deux catégories émotives fondamentales : par exemple, le mépris, qui serait un mélange de colère et de dégoût.

¹ Cette liste d'émotions représente ce que Cornelius (2000) appelle les *Big Six*, les six grandes catégories d'émotions considérées comme basiques ou primaires.

La faible description des émotions secondaires dans la théorie évolutionniste peut s'expliquer par le fait que celle-ci s'intéresse principalement à établir une taxonomie des émotions majeures, fondamentales à travers le grand nombre de réactions émotionnelles pouvant être rencontrées dans les interactions humaines. À partir de cette conception, les seuls critères de description des émotions secondaires sont qu'elles ne semblent pas être directement la manifestation d'un réflexe de survie, expliquant la non universalité de leurs expressions et le fait que celles-ci ne seraient reconnues qu'au sein de chacune des cultures où elles émergent.

II.1.2. Vers une considération de la perception et de l'évaluation émotionnelle

Au-delà du consensus sur le caractère immédiat et non conscientisé de la réaction émotionnelle dans sa manifestation primaire, certains chercheurs contemporains de Darwin, comme les psychologues danois Carl Lange et américain William James, vont dès lors insister sur l'influence de la proprio-perception de cette réaction.

En effet, c'est en s'écartant de la posture évolutionniste sans pour autant s'y opposer, que les travaux de James (1884) et Lange (1885) proposent une approche différente des origines des émotions, plaçant la perception de l'individu au cœur de la catégorisation émotionnelle des réactions physiologiques ressenties : si la réaction physiologique peut être attribuée aux différentes fonctions de survie des espèces, ce qui définit l'émotion, c'est la perception et le ressenti de l'individu face à l'ampleur de ses propres réactions (James, 1884 ; Lange, 1885).

Fort de la volonté d'appuyer son propos, ce dernier ira même jusqu'à affirmer : « On se sent triste parce qu'on pleure, en colère parce que l'on porte un coup, apeuré parce que l'on tremble » (James, 1884 ; p.190, traduction personnelle). À travers l'effet quelque peu provocateur de telles déclarations, James insiste sur ce qu'il considère être une nécessité fondamentale : il faut distinguer entre réaction physiologique et perception d'une émotion ; cette dernière sera évaluée selon lui à partir du niveau d'excitation/activation physiologique (*arousal*) ressenti.

II.1.2.1. L'approche cognitiviste et la théorie de l'évaluation (*appraisal*)

L'importance de l'aspect perceptif dans la catégorisation émotionnelle effectuée par James (1884) et Lange (1885) se retrouve également dans les fondements des théories de la psychologue américaine Magda Arnold (1960).

Toutefois, pour cette dernière, la perception de l'individu n'est pas essentielle du fait de son ressenti mais parce qu'elle englobe un type précis d'évaluation (*appraisal*) que celui-ci effectue d'une situation donnée. Selon Arnold (1960), ce traitement résulte en premier lieu de l'évaluation de la valence d'un événement comme plaisant (valence positive) ou déplaisant (valence négative). La réaction d'attraction ou de répulsion que l'événement suscite de manière directe et non réflexive chez l'individu entraînera chez ce dernier un ensemble spécifique de réactions. Un autre apport considérable de la théorie de l'évaluation (*appraisal*) réside dans la prise en compte de variations d'activation émotionnelle (*arousal*) selon l'évaluation de l'événement donné. En effet, en plus de déterminer si celui-ci produit un effet d'attraction ou de répulsion (valence), la théorie de l'évaluation permet également de déterminer l'ampleur de l'activation émotionnelle, allant de faible, voire inexistante, à très intense. De même que chez James (1884), on retrouve la notion d'excitation (*arousal*) (1884) que celui-ci ne considérait qu'en tant que réaction physiologique ne prenant sa valeur émotionnelle qu'après perception par l'individu.

Outre l'évaluation d'un événement à travers ses caractéristiques de valence ou d'activation émotionnelle, la théorie de l'évaluation d'Arnold (1960) et l'ensemble des études élaborées à partir de cette approche prévoient un nombre variable de facteurs (« composantes ») d'évaluation : effort demandé, degré de responsabilité, contrôle de soi demandé, etc.². Chaque émotion est donc considérée dans cette approche comme résultante d'une combinaison spécifique de ses composantes d'évaluation. De ce fait, la théorie de l'évaluation permet également une approche dynamique de l'émotion, du fait que tout changement de configuration de ses différentes composantes peut entraîner un changement d'émotion : par exemple, le passage d'un événement évalué comme légèrement déplaisant à très déplaisant modifie l'émotion qui en résulte, passant alors de l'irritation (catégorie secondaire de la classe « colère ») à la rage (catégorie primaire de la colère).

² Voir notamment Scherer (1999) pour la présentation d'un modèle componentiel prenant également en compte l'évaluation de la nouveauté de l'événement, et Scherer (2005) pour une revue des études suivant la théorie de l'*appraisal*.

Les théories d'Arnold (1960) et les principes de l'approche cognitive ne remettent pas en cause la forte activation physiologique résultante d'une évaluation pouvant être affiliée à une classe émotionnelle. Par contre, selon les principes de la théorie de l'évaluation (Arnold, 1960), l'émotion primaire ne peut pas pour autant être considérée comme universelle puisque l'évaluation dont elle provient est effectuée par l'individu à partir d'un faisceau de facteurs allant de son expérience personnelle au contexte social spécifique dans lequel l'événement évalué survient et les normes et principes de la société dans laquelle celui-ci évolue.

L'apport de la théorie de l'évaluation est fondamental parce qu'il introduit l'importance du contexte social et culturel des individus dans leurs évaluations des événements. Ainsi Arnold (1960) permet de tracer un pont entre une conception physiologique de l'émotion et son ancrage sociétal, qui sera ensuite mis en avant par des psycho-sociologues comme John Averill (1980) ; voir section suivante.

II.1.2.2. De l'approche catégorielle à l'approche dimensionnelle

La théorie de l'*appraisal* (Arnold, 1960 ; Scherer, 1999), au travers de ses différentes composantes, permet notamment une description plus fine des types d'émotions dites secondaires. En effet, dans la théorie de l'*appraisal*, ce qui distingue l'émotion primaire d'une émotion secondaire n'est pas tant l'universalité de ses manifestations physiques mais principalement son degré d'activation émotionnelle (excitation, ou en anglais *arousal*).

De plus, de par leur nature graduelle, présentant un continuum entre deux polarités (positif-négatif, fort-faible, etc.), les composantes principales de la théorie de l'*appraisal* peuvent permettre un complément à la description des émotions en catégories discrètes (joie, peur, tristesse, etc.). Dans une telle approche, il est possible de situer les émotions dans un espace bidimensionnel à partir de leurs configurations croisées de valence et d'excitation (*arousal*) : on y situera par exemple la colère en tant qu'émotion fondamentale à un niveau de valence minimal et à un niveau d'excitation maximale.

Bien qu'aisément modélisée, la pertinence d'une approche bidimensionnelle (valence, excitation) n'aura de cesse d'être remise en question, notamment du fait de son aspect réducteur (Russell & Mehrabian, 1977 ; Fontaine & Scherer, 2007). En effet, si celle-ci ne permet pas de distinguer la peur de la colère en tant qu'émotion de base, la distinction entre émotions dites secondaires à travers ce modèle est encore plus difficile.

De plus, si, comme nous l'avons vu, les composantes de l'*appraisal* ne se limitent pas à la valence et l'excitation (activation émotionnelle) pour pouvoir fournir la description la plus fine possible des émotions, Scherer insistera sur le fait qu'une représentation bidimensionnelle ne peut être que caduque (Fontaine & Scherer, 2007).

Dans la majorité des modèles tridimensionnels (valence, excitation, dominance ; voir notamment Russell & Mehrabian, 1977), la troisième dimension, dont la terminologie précise varie à travers les modèles, correspond à celle de la dominance, également appelée pouvoir ou puissance, ou puissance-pouvoir selon les études, et qui permet de considérer les émotions selon l'évaluation de la position dominante ou non (soumission) de l'individu face à la survenue du stimulus.

La dimension de dominance étant centrale dans la communication politique ainsi que dans toute interaction de nature compétitive (Parisi & CastelFranchi, 1975 ; Poggi & D'Errico, 2010), nous reviendrons sur sa définition dans la partie II.2 et expliciterons l'usage qu'il en sera fait dans les analyses perceptives menées dans la présente étude.

II.1.3. L'approche socioconstructiviste

Si, comme on l'a vu, l'approche cognitiviste suggère une prise en compte de l'impact du contexte social et culturel sur notre évaluation émotionnelle des événements, l'approche socioconstructiviste place celui-ci au cœur de la constitution des émotions : plus qu'un facteur à prendre en compte, l'ancrage des émotions dans la société serait à l'origine de leurs évaluations et de leurs expressions. La conception socioconstructiviste des émotions n'exclut pas l'idée d'une réalité physiologique, mais réfute principalement l'appréhension du phénomène de l'émotion en se basant exclusivement sur la réaction physiologique à un stimulus (Averill, 1982)³. Pour James Averill (1980), socio-psychologue américain, les émotions se construisent à partir des normes et attentes sociétales et leurs manifestations s'expliquent à travers celles-ci.

« Les émotions ne sont pas seulement des vestiges de notre histoire phylogénétique, et ne peuvent pas non plus être expliquées en des termes strictement physiologiques. Au contraire, elles sont des constructions sociales, et ne peuvent être comprises dans leur entièreté qu'à un niveau d'analyse sociale. » (Averill, 1980, p. 309⁴)

Par exemple, bien que la colère sous sa forme expansive soit considérée systématiquement comme une émotion fondamentale dans toutes les études se basant sur les théories darwiniennes mais également jamesiennes et cognitivistes, dans l'approche socioconstructiviste, la colère toute aussi vive qu'elle soit, est construite socialement. Pour Averill (1980), la colère, plus qu'une manifestation d'une réaction physiologique primaire, correspond à un jugement moral : celle-ci provient de notre évaluation d'un événement ou d'une action comme injuste et/ou survenue à partir d'actes de malveillance intentionnelle.

Si la colère occupe une place fondamentale dans la justification d'une conception socioconstructiviste de l'émotion, c'est principalement du fait qu'Averill (1980, 1983) conçoit son expression comme une contrepartie socialement acceptable de l'agression physique, dans une société dont l'une des règles fondamentales et fondatrices est la prohibition de l'acte de justice par soi-même.

³ De la même façon, comme le souligne Cornelius (2000), certains disciples de Darwin comme Ekman (1972) reconnaissent l'influence du social, notamment sur la régulation des émotions.

⁴ Traduction personnelle.

Averill (1983) insiste sur l'idée que quel que soit l'impact d'une réaction physiologique par rapport à un stimulus déplaisant, la réponse émotionnelle et son expression s'effectuent selon son adéquation par rapport à notre système de normes culturelles, et plus encore selon la fonction qu'elle pourra revêtir dans l'interaction en société.

II.1.4. Vers une approche fonctionnaliste de l'émotion

La conception socioconstructiviste de l'émotion comme ancrée dans un ensemble d'appréciations culturelles de son origine à son expression, est reprise dans une volonté d'être modélisée, notamment informatiquement, dès la fin des années 1970 par des psychologues comme Cristiano Castelfranchi en Italie ou Robert Abelson aux Etats-Unis, qui ira jusqu'à proposer ce qu'il appelle un *modèle fonctionnaliste de l'émotion* (Abelson, 1979).

Abelson (1979, 1983) et Castelfranchi (Parisi et Castelfranchi, 1976 ; Castelfranchi & Paglieri, 2007) ne placent toutefois pas l'émotion au centre de leurs modèles : pour eux, ce sont les buts des individus et les normes et attentes de la société dans laquelle ils évoluent qui déterminent leurs interactions. Dans cette approche, l'émotion pourra par contre être considérée comme la résultante d'une évaluation par les individus de l'accomplissement potentiel/effectif de leurs buts (ex : joie), ou au contraire de l'impossibilité de les accomplir (ex : frustration).

Nous présentons ici tout particulièrement le modèle des buts et croyances de Parisi et Castelfranchi (1976). Dans ce modèle, si les croyances d'une société ont une importance non négligeable quant à l'évaluation que se font les individus d'une action donnée (Abelson, 1979), ce sont les enjeux personnels et/ou sociétaux qui fondent les interactions sociales. Les principes du modèle des buts et croyances sont résumés comme tels, les psycho-sociologues Isabella Poggi et Francesca D'Errico (2010) utilisant également celui-ci :

« Pour atteindre ses buts, un système met en place un plan d'action organisé hiérarchiquement en exploitant à la fois des ressources internes (ses capacités d'actions et ses convictions) et des ressources externes (conditions sociétales, ressources matérielles, implications d'autres personnes dans la poursuite du but). » Poggi & D'Errico, 2010, p. 165, traduction personnelle.

L'utilisation ici du mot *système* s'explique par le fait que les principes du modèle des buts et croyances sont pensés pour être opérants tant pour décrire les interactions interpersonnelles

que la communication homme/machine et servir à faire évoluer les technologies en intelligence artificielle. En ce sens, l'approche de Parisi & Castelfranchi (1976) comme d'Abelson (1982) peut être considérée comme socio-cognitiviste⁵.

Dans cette conception des interactions sociales, ce qui va compter alors, c'est d'être en mesure (« *power of* ») d'atteindre un but. Et lorsqu'un individu n'est pas en mesure d'atteindre un but par lui-même, notamment du fait d'un manque de ressources internes ou externes, il dépend de tierces personnes pour atteindre son but (Poggi & D'Errico, 2010). Ce rapport de dépendance peut concerner des enjeux d'importances très variables allant d'un individu A demandant son chemin dans la rue à un individu B, à une personne politique A réussissant à convaincre un électeur B de voter pour lui et de ce fait d'atteindre son but, être élu.

L'intérêt majeur du modèle consiste alors à analyser et prévoir les stratégies opérantes pour que l'individu « dépendant » obtienne l'adhésion d'une autre personne à ses buts personnels. C'est sur ce point qu'entrent en jeu les normes et croyances culturelles, car elles seront en partie garantes de l'évaluation de l'attitude de l'individu A par l'individu B pour déterminer s'il adhérera à ses buts.

⁵ Dans l'importance accordée aux normes culturelles à travers lesquelles les situations sont perçues et évaluées, on retrouve certains principes de la théorie de l'*appraisal* (Arnold, 1960). De façon plus générale, il s'agit là, comme dans l'ensemble du courant cognitiviste, d'envisager un phénomène relativement au traitement qu'en fait l'être humain dans le but notamment de pouvoir le modéliser.

II.2. Comprendre les enjeux de la communication politique

II.2.1. La communication politique, une affaire d'image de soi (*ethos*)

Comme on vient de le voir dans la section précédente, l'approche sociocognitive de Castelfranchi et collègues (1976) accorde une place centrale aux buts des individus dans la caractérisation de leurs actions en société. De ce fait, si on considère la communication politique à travers le modèle des buts et croyances (Castelfranchi, 2003; Poggi & D'Errico, 2010), on voit que pour atteindre leurs buts électoraux, les personnes politiques doivent parvenir à véhiculer une image favorable pour s'assurer de l'évaluation positive des électeurs et surtout de leur adhésion.

Cette importance de l'image de soi dans la communication n'est pas sans rappeler l'un des trois concepts fondateurs de la rhétorique selon Aristote (-343, 1991 pour la présente édition) : l'*ethos*, la construction de l'image de soi pour convaincre, le *logos* (convaincre par la raison), le *pathos* (convaincre par les émotions). Pour Aristote (-343, 1991), « l'*ethos* est le moyen de persuasion le plus puissant ».

Dans la conception aristotélicienne de l'*ethos* et ainsi que dans son usage contemporain en analyse du discours politique, pour convaincre et avoir des chances d'être élu, une personne politique se doit de véhiculer l'image d'une personne morale, bienveillante et compétente (Amossi, 1999 ; Charaudeau, 2015). A travers l'approche sociocognitiviste, l'*ethos* s'agrémenté également d'une image de dominance, que la personne politique doit savoir signaler, notamment en cas de compétitivité électorale (Castelfranchi, 2003 ; Poggi & D'Errico, 2010).

Nous décrivons ici les dimensions principales de l'*ethos* à travers la théorie des buts et croyances (Castelfranchi, 2003 ; Poggi & D'Errico, 2010) c'est-à-dire en considérant en quoi chacune d'elle peut contribuer à la réalisation des buts de la personne politique⁶ :

- **La moralité et la bienveillance** : Si l'image morale véhiculée par une personne politique permet de la présenter comme étant en adéquation avec les normes culturelles, la communication d'une image bienveillante et honnête permet quant à elle d'instaurer un rapport de confiance avec l'électeur, qui pourra notamment espérer voir à travers

⁶ Pour une description plus orientée rhétorique et analyse du discours, nous renvoyons aux ouvrages et articles de Amossi (1999, 2008), Maingueneau (1999) et la revue de la littérature proposée par Donot et Emediato (2015).

son adhésion à la personne politique ou à son parti ses propres buts défendus (réduction des impôts, égalités salariales, etc.).

- **La compétence** : Une autre qualité à mettre en avant dans la construction de son image par la personne politique est sa compétence, permettant de se positionner comme à même (*in power of*) de réaliser les buts de l'électeur de par ses connaissances, ses capacités d'analyses.
- **La dominance** : Une image dominante renforce alors la confiance de l'électeur dans les capacités de la personne politique à pouvoir marquer son autorité, notamment en étant capable de défendre, en cas de conflit, les valeurs de son parti. L'image dominante permet également de favoriser la perception d'une personne politique comme capable d'être à la tête d'une fonction importante dans la société.

Pour Poggi et D'Errico (2010), dans le cadre du débat politique, la dimension de dominance prime sur toutes les autres dimensions de l'*ethos*, sous-entendant que malgré son honnêteté et ses compétences apparentes, un politicien qui n'arrive pas à être perçu comme ayant l'ascendant sur son opposant aura du mal à obtenir l'adhésion de l'électeur et ainsi réaliser ses buts électoraux. Cette conception des dimensions de bienveillance et de compétence comme secondaires à la dimension de dominance, comme n'étant finalement pas indispensables pour qu'une personne politique parvienne à véhiculer une image favorable et atteindre ses buts électoraux, se retrouve dans la conception du débat politique de Charaudeau (2015) qui insiste pour dire qu'on peut y apparaître tout à fait convaincant et dominant en tenant un raisonnement erroné ou malhonnête.

Dans l'approche socio-cognitiviste, l'importance cruciale de la dominance dans la communication politique réside dans la nature compétitive des buts électoraux et le fait que lorsqu'une personne politique est perçue comme dominante, elle est perçue comme ayant potentiellement *plus de pouvoir* qu'une autre. Poggi & D'Errico (2010) insistent sur la nécessité de pouvoir distinguer les notions de statut (pouvoir hiérarchique), de pouvoir (pouvoir d'influencer l'autre) et de dominance (*plus de pouvoir*). Selon elles, ce qui différencie la dominance d'autres notions relatives de pouvoir, c'est que celles-ci ne reposent pas tant sur des critères factuels que sur la perception des électeurs ; ceci pouvant expliquer en partie l'ampleur des enjeux d'un débat politique pour la construction d'une image (*ethos*) favorable comme le souligne notamment Dominique Maingueneau (1999).

II.2.2 Signaux de dominance et agressivité dans le débat politique

Bien qu'on vienne de considérer son importance pour renforcer la dimension de dominance de l'ethos, véhiculer une image de calme et de maîtrise de soi n'est pas toujours aisé dans une situation comme le débat politique. En effet, la nature même du débat politique est conflictuelle et donne lieu à de fréquentes attaques entre opposants (Burger et collègues, 2011 ; Poggi & D'Errico, 2010 ; Charaudeau, 2015 ; Poggi et collègues, 2011). On pensera notamment à l'exploitation stratégique par Nicolas Sarkozy de la colère explicitement exprimée par Ségolène Royal pour discréditer celle-ci et s'attaquer à sa dominance (« Il faut être calme pour être président ») ; voir au sujet de ce débat l'analyse conversationnelle multimodale effectuée par De Chanay et collègues (2011).

Pour autant, pour Poggi & D'Errico (2010), que celle-ci soit indice d'une excitation émotionnelle ou non, l'agressivité fait partie des signaux de dominance dans le débat. Toutefois, puisqu'il faut également, comme on vient de le voir, pour les participants tenir compte de l'impératif premier de véhiculer autant que possible une image de soi calme, l'expression d'attitudes plus tacites qu'explicités serait privilégiée (mépris, sarcasme, ironie sarcastique) ; Poggi & D'Errico (2010).

D'ordre général, dans une optique socio-cognitiviste où le but fondamental des locuteurs est de gagner en pouvoir hiérarchique en véhiculant une image favorable à l'adhésion des électeurs, bien que l'agressivité puisse en quelque sorte signaler une forme de dominance⁷, elle ne doit pas pour autant porter préjudice aux autres dimensions de l'*ethos* des locuteurs.

Si comme on vient de le voir une expression de colère fortement activée émotionnellement (colère chaude : coup de rage, indignation forte ; Banse & Scherer, 1996) peut nuire indirectement à la perception de la dominance en faisant apparaître le locuteur comme perdant son calme, elle peut risquer de la même façon de le faire passer pour malveillant ou immoral si celui-ci s'attaque trop vivement et directement à son opposant (insultes, menaces et expressions multimodales de colère chaude).

⁷ Pour rappel, dans une approche dimensionnelle des émotions (Banse & Scherer, 1996), la colère et ses sous-catégories se situent dans une position plus dominante que par exemple la tristesse ou la peur.

II.2.3 L'attaque à l'image de l'autre, stratégie de dominance dans le débat

Une stratégie courante dans le débat politique consiste à performer des actes visant à discréditer (*discrediting moves*⁸) l'image de l'opposant (Poggi et collègues, 2011). Les actes discréditants peuvent être explicites (accusations directes, insultes) ou implicites (indifférence feinte, ironie) et viser personnellement l'adversaire ou seulement ses choix et actions. Surtout, le principe de ces actes repose sur le fait que dans un débat, l'image véhiculée par les personnes politiques est perçue par les spectateurs en comparaison permanente à celle de l'opposant. De ce fait, la performance d'actes discréditants est une façon indirecte de construire son image comme meilleure que celle de l'opposant, en visant à discréditer celui-ci au niveau de chaque dimension de son *ethos* (bienveillance, compétence, dominance).

Les actes discréditants peuvent chercher à mettre à mal l'image de l'opposant au niveau de sa bienveillance, en faisant état de sa malhonnêteté ou de son attitude malveillante, de sa compétence, en le faisant passer pour ignorant ou stupide, mais aussi mettre en jeu sa dominance, en le ridiculisant et en renvoyant une image de lui impuissante et inconséquente (voir tableau ci-dessous).

DIMENSION VISÉE	SOUS-ENTENDUS SUR L'OPPOSANT
BIENVEILLANCE	Immoral, malveillant, malhonnête
COMPÉTENCE	Mal informé, stupide, incompetent
DOMINANCE	Impuissant, ridicule, facilement décontenancé

Tableau II—1 : Tableau reprenant la typologie d'actes discriminants de Poggi et collègues (2011) selon la dimension de l'ethos visée par l'attaquant chez son opposant et sous-entendus équivalents pour les dimensions de bienveillance, compétence et dominance.

⁸ Le terme de *discrediting move* pourrait également se traduire comme « geste discréditant » ou « tactique » discréditante. Nous avons préféré le traduire en « acte » discréditant pour plus de clarté et de cohérence par rapport à son inscription dans une approche du débat en buts et croyances (Poggi et collègues, 2011) où ses participants mènent des actions dans le but de gagner l'adhésion des électeurs.

À travers la performance de ces actes, quelle que soit la dimension visée dans l'image de l'autre, la personne politique peut espérer par effet de comparaison apparaître ultimement comme plus dominante que son adversaire. Toutefois, à travers l'analyse qualitative de débats politiques italiens, Poggi et collègues (2011) remarquent que l'effet favorable pour l'attaquant dépend conjointement de la dimension qu'il cherche à atteindre chez son opposant et du degré d'expressivité de l'attaque.

Selon les observations de Poggi et collègues (2011), si l'on cherche à faire apparaître son opposant comme malveillant, il apparaît profitable d'appuyer son propos d'une expressivité physique et vocale signalant une forte activation émotionnelle, une forte implication personnelle⁹. Au contraire, s'attaquer à la dominance de l'opposant à travers une attitude ouvertement hostile comporterait le risque d'un effet « boomerang » ; l'attaquant risquant alors de nuire à sa propre image en passant pour cruel ou injuste aux yeux des électeurs. Ces tendances opposées s'expliquent par le fait qu'il apparaît juste et normal dans l'imaginaire collectif de s'indigner de la malhonnêteté, la malveillance ou l'incompétence. Il est en revanche un peu plus questionnable et un peu moins attendu de mettre en cause la dominance d'un acteur politique en affichant une image de lui comme impuissant ou inconséquent dans le débat.

⁹ Cette observation n'est pas sans rappeler l'usage d'une forme de colère à visée fédératrice par les partis d'opposition lors des campagnes électorales, incitant les électeurs à s'indigner des injustices commises à leur égard par le gouvernement sortant (Ballet, 2012). Bien que cette tactique résulte plus d'une forme d'appel aux émotions (*pathos*), nous pouvons tout aussi bien considérer ici l'effet d'une posture de colère renvoyant principalement l'image (*ethos*) d'une personne politique bienveillante, soucieuse du traitement démocratique des citoyens.

II.3. Débat politique et expressivité vocale : Caractéristiques de parole attendues

II.3.1. Introduction à l'approche phonétique

Avant de présenter ici un ensemble de caractéristiques pouvant être attendues dans l'étude phonétique d'extraits de débats, il s'agit dans un premier temps de déterminer quels liens peuvent être établis entre les théories de l'émotion et de la communication politique abordées en ce début de chapitre et celles de la composante de parole de l'expressivité.

À en croire Véronique Aubergé (2002), le phénomène de la parole, qui plus est expressive, est intrinsèquement lié à tous les aspects théoriques de l'émotion que nous venons d'aborder (physiologique, cognitif, conventionnel, social, motivé) :

« La parole, dans toutes ses émotions, est à la fois un ensemble de signaux de l'appareil sensori-moteur et des événements du langage. Produite pour/comme être perçue (perception-action), communicant les codes dans un système inter-personnel, elle intègre la complexité de la machine biologique, du système cognitif du langage et celle de l'interaction motivée. »

Aubergé (2002, p. 263)

Mais comment l'expressivité peut être envisagée spécifiquement sur le plan phonétique ? Pour le linguiste, phonéticien et psychanalyste hongrois Iván Fónagy, la parole est constituée d'un **double encodage**¹⁰ (Fónagy, 1971) : Lorsqu'un individu s'exprime, en parallèle à l'information transmise à travers un code *linguistique*, sa façon de parler, à travers un second code *paralinguistique*¹¹, apporte un ensemble vaste et hétérogène d'informations à son sujet : son état émotionnel et physique, son rapport à l'autre, ses intentions, sa personnalité, son genre, son âge, sa profession, son origine socioculturelle, etc.

Selon Fónagy (1983, 2006), pour que les caractéristiques du code paralinguistique soient identifiées par les interlocuteurs, celles-ci doivent constituer un **geste-écart** du locuteur par rapport à la norme à travers laquelle sa parole est perçue. À partir de cette notion, on considère

¹⁰ Si le modèle du double encodage que l'on doit à Fónagy (1971) est particulièrement adapté à l'analyse phonétique appliquée des caractéristiques vocales expressives, nécessitant de pouvoir s'abstraire conceptuellement de la part d'expressivité pouvant être véhiculée par les mots employés, celui-ci trouve sa genèse dans une réflexion théorique instaurée en 1944 par le linguiste suisse Charles Bally (1944) pour qui tout énoncé peut être considéré à travers son contenu propositionnel, le *dictum*, et le jugement produit par le locuteur sur ce contenu, le *modus*.

¹¹ Les informations paralinguistiques sont alors décodées par les interlocuteurs et/ou auditeurs, qu'elles soient véhiculées intentionnellement, tel un signal (ton menaçant), ou non intentionnellement, tel un indice (voix tremblante résultant d'un stress non contrôlé).

généralement les caractéristiques expressives comme geste-écart par rapport à une norme non expressivement marquée. Dans le cadre de notre étude, la parole expressive des locuteurs va donc être à considérer comme geste-écart par rapport aux caractéristiques générales du style de parole politique en français.

II.3.2. Le phonostyle de la parole politique et les enjeux du débat

La parole politique est en effet un style de parole à part entière, au même titre que celui propre au journaliste d'informations, au reporter sportif, au prêtre, au conférencier, etc. (Fónagy, 1983 ; Léon, 1993). En français, les caractéristiques principales du style sonore ou *phonostyle* politique sont de l'ordre d'un débit articulatoire lent, de temps de pauses élevés, et, sur le plan mélodique, un registre de F0 stable et bas (Lucci, 1980 ; Duez, 1991 ; Léon, 1993 ; Bardiaux, 2010 ; Simon et collègues, 2010)¹². Les différentes études susmentionnées font également état de la production fréquente d'accents initiaux, pouvant revêtir différentes fonctions allant d'une fonction *didactique* (Lucci, 1980 ; Léon, 1993), contribuant à renforcer l'intelligibilité du propos et faciliter le décodage par l'auditeur, à une fonction plus *impressive*, visant à attirer l'attention de l'auditeur sur certains éléments du discours par le biais d'accents dits de focalisation. Les accents initiaux peuvent également être la marque d'insistance ou d'emphase (Carton, 1974 ; Léon, 1993 ; Guaitella, 2014), indiquant alors l'implication émotionnelle du locuteur, et revêtant ainsi une fonction plus *expressive*.

Comme l'a souligné Vincent Lucci (1980, 1983), si l'on peut retrouver de telles caractéristiques dans l'ensemble des styles de parole argumentative publique en français, la particularité de la parole politique résiderait dans la récurrence, voire le systématisme avec lequel ces caractéristiques, notamment accentuelles, émergent. Elles peuvent également former des patrons prosodiques réguliers, permettant aux locuteurs de se distinguer d'autres phonostyles jusqu'à parfois représenter plus encore qu'un style de parole politique général, un style de parole idiolectal et emblématique. Ainsi, Danielle Duez (1991, 1999) a montré que celui du général De Gaulle était constitué de groupes rythmiques ternaires portés par de nombreux accents initiaux. Plus récemment, le style de parole de Nicolas Sarkozy a été analysé

¹² Nous notons ici que comme pour tout autre type de locuteurs, les caractéristiques des locuteurs politiques sont elles aussi soumises à variation selon les spécificités de la situation de communication et le degré de formalité de l'interaction (Koch & Oesterreicher, 2008). Des études récentes ont notamment montré que le débit de parole est plus élevé en condition de débat que d'allocation publique (Bardiaux, 2010) et que les caractéristiques de registre de F0 peuvent être amenées à varier selon le degré d'implication du locuteur (Simon et collègues, 2010). Nous revenons sur ce phénomène en section II.3.3 sur les caractéristiques de parole agressive.

comme caractérisé par des contours mélodiques descendants aux frontières de continuation majeure (Martin, 2009), déviants par rapport aux normes prosodiques du français et facilitant d'autant plus la perception d'un statut particulier : celui de personne politique.

II.3.2.1. Synthèse de la problématique

L'approche socio-cognitiviste du débat politique proposée par Poggi et collègues (2010, 2011) permet de prévoir chez les personnes politiques une émergence de signaux de dominance et d'hostilité pour assurer la perception de leur ascendance durant les débats¹³.

Compte tenu de tels objectifs pour les locuteurs, nous pouvons faire l'hypothèse de caractéristiques de voix et de parole ayant, en parallèle à celles indexant leur statut de personne politique :

- Une fonction impressive (Fónagy, 1983, 2006 ; Léon, 1993), c'est-à-dire visant à produire un effet chez l'interlocuteur et/ou l'auditeur, ici favorisant l'attribution au locuteur d'un caractère dominant.
- Une fonction expressive (Jakobson, 1963), avec des caractéristiques résultantes à l'expression et la régulation d'attitudes notamment agressives.

Pour fournir une description aussi large que possible des caractéristiques de parole pouvant être attendues pour remplir ces fonctions dans le débat, nous présenterons dans un premier temps un ensemble d'informations sur la parole politique et les signaux de dominance (**II.3.2**). Dans un second temps nous décrirons les caractéristiques relatives à l'expression vocale de l'hostilité (**II.3.3**) pour laquelle nous partirons des hypothèses de Poggi & D'Errico (2010) concernant la favorisation d'expressions agressives contrôlées, qui nous permettront de replacer dans un contexte plus large la littérature phonétique sur ce type d'attitudes.

¹³ Nous notons toutefois que ces observations n'ont été effectuées que sur des extraits de débats entre personnes politiques contemporaines dans des sociétés occidentales. En effet, l'expression de l'hostilité, qui plus est en situation de débat, pourrait être perçue de façon différente selon les sociétés et cultures considérées.

II.3.3. Parole politique et dominance, de part et d'autre de la pause silencieuse

Dans cette section, nous abordons les signaux de dominance dans la parole politique afin d'identifier les principales informations sur les caractéristiques prosodiques pouvant être attendues dans notre corpus d'étude présenté au chapitre suivant.

En effet, au sein même du phonostyle politique, on retrouve des caractéristiques susceptibles d'être interprétées comme marqueurs de dominance par les auditeurs, telle qu'une F0 moyenne basse (Morton, 1977 ; 1984) ou des temps de pauses élevés (Duez, 1991, 1999).

Sur le plan mélodique, les rapports entre hauteur de F0 et perception de la dominance ont été conceptualisés par John Morton (1977) et John Ohala (1984) à partir d'études éthologiques (rapports de force et de séduction entre espèces animales) et transposables à la communication humaine. Pour ces chercheurs, on effectue inconsciemment en perception des rapports entre caractéristiques physiologiques (tailles des cavités pharyngales et buccales, taille et épaisseur des plis vocaux) et caractéristiques fréquentielles (hauteur de la fréquence fondamentale, hauteurs des fréquences de résonance). D'après cette théorie dite du *code fréquentiel* (Ohala, 1984), les locuteurs peuvent alors effectuer volontairement des modifications de leurs caractéristiques fréquentielles moyennes dans le but de produire chez l'interlocuteur la perception d'un locuteur de physionomie plus imposante et qui serait donc dominant en cas de rapport de force.

Au niveau temporel, les recherches de Duez (1991, 1999), portant quant à elle directement sur la parole politique, ont permis de démontrer un rapport symbolique entre pouvoir et silence à partir d'analyses de corpus de parole politique dans différentes situations de communication (élocution publique, entretien amical). Duez (1991, 1999) a notamment montré que le nombre et la durée des pauses silencieuses est corrélée au statut hiérarchique des locuteurs, la conclusion principale de ses études étant que la pause est un marqueur symbolique du pouvoir.

À partir des recherches de Duez (1991), nous émettons ici l'hypothèse que la pause silencieuse peut également être utilisée comme signal de dominance, quel que soit la position hiérarchique du locuteur et nous faisons le choix d'étudier la variation des caractéristiques prosodiques mélodiques et temporelles de la parole politique à partir de la pause silencieuse.

II.3.3.1. La pause : du marqueur de pouvoir au signal de dominance

Si, selon Duez (1991), la pause silencieuse symbolise par sa plus longue durée et sa plus grande fréquence d'occurrence le pouvoir hiérarchique plus important du locuteur, c'est entre autres parce qu'elle correspond à une manifestation de la supériorité du locuteur à l'égard de ses interlocuteurs, pouvant se permettre d'effectuer de longues pauses sans même craindre d'être interrompu.

En effet, outre le fait que l'insertion fréquente de pauses silencieuses permet au locuteur de faciliter l'encodage du discours par l'auditeur, les pauses silencieuses en tant que marqueur de pouvoir hiérarchique et surtout en tant que signal de dominance¹⁴ permet au locuteur de signaler son ascendance dans le débat par le maintien de son tour de parole malgré les silences qu'il y intègre. En effet, dans la gestion interactive des tours de parole, une frontière prosodique entre deux unités syntaxiques majeures (pauses structurantes) marquée par une pause, peut typiquement être un moment acceptable et socialement adapté pour interrompre l'interlocuteur et amorcer un autre tour de parole (Sacks et collègues, 1978). De ce fait, les observations de Duez (1991) sur la durée très élevée des pauses structurantes dans les discours d'investiture des dirigeants politiques sont une bonne démonstration du fait que les locuteurs ont dans cette situation tellement de pouvoir qu'ils ne risquent en aucun cas d'être interrompus.

Toutefois, le contexte d'élocution publique étant par nature plus formel et plus monologique que celui du débat politique (Koch & Oesterreicher, 2008), comment peuvent faire les locuteurs pour utiliser ce type de pauses pour signaler leur dominance sans pour autant risquer de mettre à mal celle-ci en se faisant interrompre par leur opposant ?

Sacks et collègues (1978) font état dans leur célèbre synthèse du fonctionnement du système de tours de parole d'une stratégie intéressante consistant, pour éviter d'offrir à l'interlocuteur une chance de l'interrompre à la place de choix que constitue la frontière entre unités syntaxiques majeures, à déplacer la production de leurs pauses juste après celle-ci. Ce phénomène constitue ce que Gaëlle Ferré (2004) appellera ensuite « pauses démarcatives déplacées », qui contrairement aux pauses démarcatives ayant pour fonction de structurer le discours et de guider son interprétation par l'auditeur (fonction structurante), auront une fonction paralinguistique, pour une gestion pragmatique des tours de parole.

¹⁴ La distinction faite ici ne renie en aucun cas le fait que les conclusions de Danielle Duez tendent directement vers l'hypothèse d'une utilisation de la pause en tant que signal de dominance, mais par contre puisque pouvoir hiérarchique et dominance sont deux phénomènes distincts comme nous l'avons vu dans la section précédente, nous insistons sur ce point. En d'autres termes, nous émettons l'hypothèse que la pause silencieuse peut être employée quelle que soit le degré de pouvoir hiérarchique.

Enfin, il est nécessaire ici de prendre en compte le fait que lorsque Duez (1991) considère la pause silencieuse comme marqueur de pouvoir, elle la considère à travers sa fonction impressive, visant à signaler la supériorité hiérarchique (ou symbolique) du locuteur, tout en reconnaissant dans le cas de pauses situées entre unités syntaxiques majeures le fait que celles-ci puissent également avoir une fonction structurante. Les fonctions des pauses et a fortiori de la prosodie ne s'opposent pas les unes aux autres du point de vue de Duez (2007) :

« On a en fait une véritable hiérarchie des fonctions selon le contexte situationnel. La fonction structurale et contrastive est omniprésente dans tout message oral, mais elle est particulièrement prédominante dans la parole lue, la communication scientifique, dans la parole dite « neutre » ; dans les échanges spontanés, les conversations, elle peut s'effacer au profit de la fonction pragmatique. Dans les messages publics, c'est la fonction attitudinale qui passe au premier plan. La superposition des différentes fonctions s'actualise par des variations de l'organisation temporelle et fréquentielle qui sont perçues, identifiées et interprétées par l'auditeur au cours de l'acte de communication. » Duez (2007, p.187)

II.3.3.2. Fonctions des pauses et caractéristiques prosodiques générales

À travers sa revue de la littérature, Duez (1991) détermine trois principaux types de pauses : la pause d'hésitation, la pause de démarcation, et la pause de focalisation. La *pause d'hésitation* aurait pour fonction de fournir un temps d'accès supplémentaire au lexique en cas d'hésitation, la *pause de démarcation* une fonction de structuration pour faciliter le décodage. La *pause de focalisation* quant à elle est présente pour attirer l'attention sur ce qui va suivre et peut avoir tout autant une fonction impressive (focus) pour susciter chez l'auditeur une hiérarchisation sémantique de l'élément focalisé ou une fonction expressive (emphase), attirant plutôt l'attention de l'auditeur sur l'implication émotionnelle du locuteur à l'égard de son propos. Cette distinction a été notamment présentée habilement au sujet de la parole politique par Guaïtella (2014).

Cette typologie sera notamment reprise par Ferré (2004), qui y rajoute le cas des pauses de démarcation déplacées à fonction pragmatique de maintien du tour de parole. Cette typologie sera suivie par Béchet et collègues (2013) pour en analyser d'éventuelles caractéristiques distinctives au sein des pauses silencieuses dans le débat Hollande-Sarkozy de l'entre-deux tour des présidentielles en 2012. La figure ci-dessous présente les caractéristiques prosodiques attendues de part et d'autre des pauses silencieuses selon Ferré (2004) :

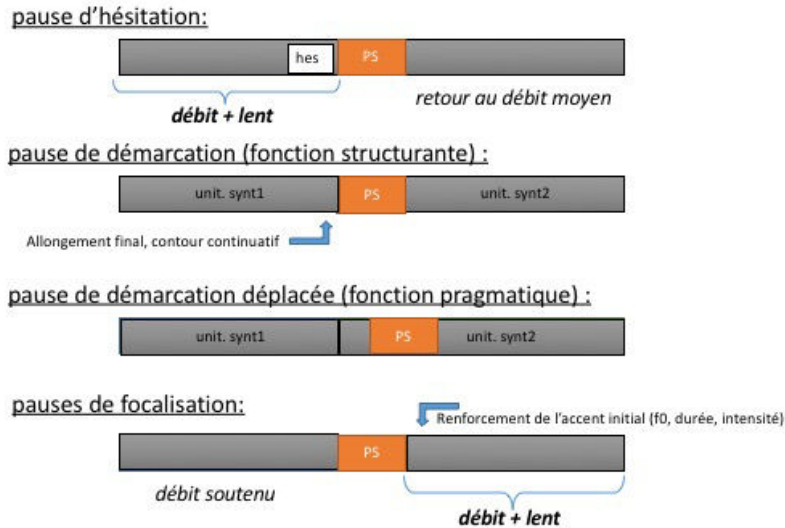


Figure 1 : schématisation personnelle des observations de Ferré (2004) concernant les caractéristiques prosodiques de part et d'autre des pauses silencieuses (PS= pause silencieuse) selon que celles-ci aient principalement une fonction d'hésitation, de structuration, de maintien du tour de parole (pragmatique) ou de focalisation. Lorsque non-spécifiés les blocs gris peuvent correspondre tout autant à des unités syntaxiques qu'à des unités prosodiques incongruentes avec la syntaxe.

Pour Ferré (2004), selon la fonction de la pause silencieuse, on peut trouver des caractéristiques prosodiques différentes, notamment de part et d'autre des pauses. Si Duez (1991) avait déjà insisté sur les rapports entre fonction des pauses et structuration prosodico-syntaxique, les pauses à usage paralinguistique pouvant être à l'origine d'unités prosodiques majeures incongruentes avec la structure syntaxique de l'énoncé, l'apport de Ferré (2004) consiste à s'attacher à chercher et décrire les corrélats prosodiques des différents types fonctionnels de pauses.

Pour elle, on observe notamment, et cela sera confirmé par Béchet et collègues (2013) dans la parole politique en débat, des profils contrastifs temporels différents selon les types de pauses :

- Dans le cas de *pauses d'hésitation*, que Duez (1991, 1999) souligne être plus rares dans la parole politique que dans d'autres styles de parole, on observe un ralentissement pré-pausal du débit articulatoire du fait de la recherche de mots, l'insertion éventuelle de marques sonores d'hésitation, et un retour au débit articulatoire moyen après la pause. Concernant la localisation syntaxique, celle-ci est très variable, d'où sa non spécification dans notre figure.
- Dans le cas de *pauses à fonction structurante*, on observe un débit articulatoire similaire de part et d'autre des pauses silencieuses, une congruence avec la syntaxe (les pauses structurantes sont parfois également appelées pauses syntaxiques comme le souligne

Candea, 2000), et la production d'autres caractéristiques relatives à la production de frontières prosodiques tel qu'un allongement final et un contour continuatif.

- Dans le cas de *pauses à fonction pragmatique* de maintien de tour de parole (pauses de démarcation déplacées), on observe comme mentionné plus haut l'insertion d'une pause silencieuse pas tout à fait à la frontière entre deux unités syntaxiques majeures mais juste après le premier mot de l'unité syntaxique suivante¹⁵.
- Dans le cas de *pauses de focalisation*, on observe toujours selon Ferré (2004), un fort contraste de débit articulatoire de part et d'autre de la pause ; le débit étant soutenu avant celle-ci puis ralenti après celle-ci pour contribuer, conjointement aux phénomènes de renforcements initiaux relatifs à la production d'un focus à la mise en lumière de l'élément de discours focalisé.

II.3.3.3. La pause comme marqueur subsidiaire du style didactique

Enfin, cas non prévu par Ferré du fait qu'elle n'étudie pas la parole politique, notons que la pause peut aussi venir précéder la production d'un accent « didactique » (Lucci, 1980) et Léon (1971, 2005) parlera alors pour ces cas d'une *pause de joncture externe*.

Pour Vincent Lucci (1980), la prise de parole politique fait partie des styles de parole relatifs au discours argumentatif public au même titre, par exemple, que la conférence notamment scientifique. Lucci (1980) fait état d'une production fréquente d'accents primaires qu'il nomme accents didactiques car exerçant pour lui une mise en lumière d'éléments clés du discours.

Cependant, outre une fonction pouvant sembler être analogue à la fonction de focalisation, les accents didactiques de par leur fréquence d'occurrences revêtraient surtout une fonction identificatrice, contribuant à indexer un style de parole politique tout en symbolisant l'implication des locuteurs politiques à l'égard de leur discours.

Si l'accent didactique s'éloigne de l'accent de focalisation, voire d'accents d'insistance pouvant avoir une fonction expressive, c'est selon Lucci (1980) et Léon (1971, 2005) parce qu'ils permettent, outre un effet de scansion dans le discours, le découpage régulier du flux de parole en unités restreintes, de durées isochroniques, que la combinaison de ces différents types d'accentuation primaire facteurs est particulièrement typique du style de parole politique¹⁶.

¹⁵ Pour ce cas précis de configuration, Ferré (2004) n'indique pas de profils prosodiques particuliers ; aussi nous prendrons soin de décrire les caractéristiques observées.

¹⁶ Nous notons à ce propos que Degand et Simon (2009, 2011) font état pour la parole politique d'insertions régulières et fréquentes de pauses silencieuses pour marquer des frontières syntaxiques, même intermédiaires, ces dernières étant pourtant sensées être moins marquées prosodiquement que les frontières syntaxiques majeures (Delattre, 1966). Ce type précis de structuration prosodico-syntaxique particulier, que Degand et Simon considèrent comme « orientée syntaxe » s'explique selon elles

En somme, l'accent didactique est différent d'un phénomène de focalisation à proprement parler car sa fonction ne relève pas tant d'une fonction impulsive ponctuelle que d'une fonction identificatrice, plus stable. On s'interroge alors sur la nature de ses corrélats prosodiques et notamment des caractéristiques temporelles de part et d'autre des pauses pré-accent didactique.

également par un effet de style didactique, dans le but, en segmentant régulièrement le discours et ce de façon congruente avec la structure syntaxique, de favoriser l'intelligibilité et l'encodage par l'auditeur.

II.3.4. Les caractéristiques phonétiques des attitudes agressives contrôlées

Si, comme nous l'avons vu, l'émergence d'agressivité dans la parole politique en situation de débat peut grandement s'expliquer par une approche socioconstructiviste et socio-cognitiviste (Castelfranchi, 2003 ; Poggi & D'Errico, 2010 ; Poggi et collègues, 2011), qui prévoit même une favorisation d'attitudes agressives contrôlées (irritation, sarcasme, mépris), la description précise de telles caractéristiques vocales nécessite le recours à des cadres théoriques pluriels issus des différents courants présentés en II.1. De plus, quel que soit le courant théorique qui l'aborde, que ce concept soit pertinent ou non pour lui, la colère dans sa forme non fondamentale, non primaire, est complexe à décrire au niveau de caractéristiques phonétiques qui lui soient propres (Scherer, 2003).

Dans l'optique de pouvoir fournir des informations directement analysables et applicables à une étude de corpus de parole politique expressive (donc produite hors laboratoire), nous nous restreindrons ici exclusivement à la présentation des caractéristiques prosodiques de fréquence fondamentale (F0), durée, et pauses silencieuses et renvoyons à Kouklia (2012) pour une revue des caractéristiques de parole agressive incluant également les paramètres de qualité de voix, d'intensité ainsi que les caractéristiques spectrales consonantiques et vocaliques¹⁷. Nous n'hésiterons toutefois pas lorsque nécessaire à utiliser à titre d'exemple explicatif des informations sur ces autres paramètres.

II.3.4.1. Colère chaude versus colère froide : de Darwin à Scherer

La **colère** en tant qu'émotion fondamentale (Ekman, 1992) est en fait une colère expansive, communément appelée **colère chaude** (Banse & Scherer, 1996). Cette forme de colère s'oppose à une variante plus retenue, la **colère froide**, alors considérée dans une approche évolutionniste comme émotion secondaire¹⁸. En effet, on observe des caractéristiques acoustiques moins soulignées pour les sous-catégories faiblement activées par rapport à celles

¹⁷ Un corpus de débats issus d'enregistrements audio-visuels comme celui que nous allons analyser ne permet pas du fait de la qualité variable du signal acoustique d'effectuer des analyses fiables d'intensité ni de qualité de voix, et les analyses segmentales spectrales peuvent également être impactées.

¹⁸ D'un point de vue terminologique, on trouve une distinction « colère chaude » / « colère froide » (Banse et Scherer, 1996), qui est la plus répandue dans la littérature. On note également une distinction entre « colère de frustration » et « colère d'agression » (Frick, 1986), ainsi qu'une opposition entre « colère contenue » et « colère expansive » (Fónagy, 1983).

de leur variante fortement activée émotionnellement, relatives donc à un fort degré d'*excitation* sur le plan dimensionnel (Scherer, 1986 ; Fontaine et Scherer, 2007).

Les caractéristiques prosodiques majeures de la **colère chaude** sont les suivantes :

- Accroissement de la fréquence fondamentale (F0) moyenne et de sa plage de variation¹⁹ due à une variété de facteurs allant d'une plus forte tension laryngée à un débit d'air sortant plus important, tous relatifs à une forte activation émotionnelle (Scherer, 2003).
- Contours mélodiques fortement descendants dont l'ampleur est permise par l'accroissement général de la plage de variation de F0 (Scherer, 2003).
- Débit articulatoire élevé et fortement irrégulier et réduction de la proportion de silence pouvant aller jusqu'à une absence de pauses (Scherer, 2003).
- Rythme de durées syllabiques perturbé par rapport à la norme isochronique du français (Léon, 1973) dû à l'alternance irrégulière de séquences de syllabes produites avec un rythme *staccato* et d'allongements syllabiques exagérés (Fónagy, 1983).
- Une insertion irrégulière de pauses silencieuses de durée courte et majoritairement incongruentes avec les frontières syntaxiques (Léon, 1973).

Si les caractéristiques que nous venons de lister ici pour la **colère chaude** présentent un fort consensus à travers des études et des revues de la littérature²⁰, il n'en est pas de même pour les caractéristiques relatives à la **colère froide** pour lesquelles on trouve des caractéristiques plus hétérogènes et parfois contradictoires, bien que sur ce point Banse et Scherer (1996) insistent sur le fait que certaines études ne spécifient pas assez quel type de colère a été enregistré.

En effet, dans la conception cognitive et componentielle de Scherer (2003), la colère froide (irritation) se distingue la colère chaude (rage) de par son degré plus faible d'**activation émotionnelle**. Scherer (2003) explique que d'ordre général, les caractéristiques phonétiques relatives à l'expression de colère froide, si elles sont moins aisément discernables, sont tout simplement moins marquées car relatives à une moindre activation ; par exemple cela peut selon lui expliquer le fait que l'on retrouve dans plusieurs études des mentions d'une plage de variation légèrement plus réduite dans les expressions de colère froide que pour la colère chaude.

¹⁹ La plage de variation de la fréquence fondamentale (F0) correspond à l'étendue des valeurs de F0 (F0 max – F0 min) et peut-être calculée en Hertz ou en demi-tons.

²⁰ Nous renvoyons prioritairement pour celles-ci à Scherer (2003) et Juslin & Laukka (2003).

Si cette distinction permet de prévoir, pour l'expression de colère froide, l'observation de caractéristiques prosodiques expressives variables et moindres à celles attendues pour l'expression de colère chaude, il nous semble important de ne pas écarter l'éventualité qu'il y ait parmi ces expressions des traits spécifiques, ne pouvant pas seulement s'expliquer par l'atténuation des caractéristiques de la colère chaude.

II.3.4.2. L'apport « psychophonétique » de Fónagy

Dans l'approche psychanalytique du néanmoins linguiste et phonéticien Ivan Fónagy (1983), les différentes sous-catégories d'émotions ne se distinguent pas tant par le degré d'activation émotionnelle que par leur base pulsionnelle. Fónagy considère les pulsions telles qu'on peut les aborder en psychanalyse comme l'explication première des particularités d'une expression émotive, voire de toute expression vocale.

Cette conception héritée de la psychanalyse permet notamment dans le cas de la colère de prévoir certaines caractéristiques comme étant particulièrement spécifiques à la colère froide, retenue, car Fónagy (1983) lui attribue un caractère à part, du fait de sa tendance à la rétention.

A partir des théories freudiennes, la haine (colère froide) et la rage (colère chaude) sont toutes deux des variantes de la pulsion sadique-anale, structure pulsionnelle prégénitale. Fónagy (1983) fait alors la différence entre la colère chaude, forme expansive basée sur une pulsion d'éjection, et la haine, caractérisée par une tendance pulsionnelle à la rétention.

À partir de cette base pulsionnelle, Fónagy (1983), décrit des caractéristiques phonétiques expressives comme spécifiques à la colère froide :

- les plis vocaux se tendent et la durée des instants glottiques s'accroît de façon à maintenir la glotte fermée, qui en accroissant la pression sous-glottique retrouve ainsi son statut de sphincter.

- Les occlusives sourdes sont renforcées dans leur temps et leur surface d'occlusion, pour accentuer la sensation de rétention et imiter la fermeture des sphincters. Parmi elles, les occlusives vélares sont particulièrement postériorisées pour faire écho, en se rapprochant de la glotte, à la fermeture brusque de celle-ci en cas de dégoût.

- La durée relative des voyelles par rapport aux consonnes se réduit, les sons non voisés sont favorisés au détriment des voyelles et des sonantes.

Si de telles caractéristiques relatives d'une part à la qualité de voix et d'autre part à des variations segmentales ne sont pas traitées dans cette étude, nous décrivons ces phénomènes

car il permettent de comprendre l'origine d'un postulat général fait par Fónagy (1983) au sujet des expressions de type colère froide : outre le silence inhérent à la tenue des occlusives sourdes, les pauses silencieuses pourraient, lorsque combinées avec des caractéristiques mentionnées ci-dessous, contribuer à la caractérisation du style « retenu » de la colère froide, et se démarquer ainsi de la colère chaude.

Enfin, il convient de garder à l'esprit que la catégorie « colère froide », si elle est majoritairement considérée dans les études à travers l'attitude d'irritation (Scherer, 2003) et ici avec Fónagy (1983) à travers la haine, peut englober également des attitudes comme entre autres l'**ironie sarcastique**. Une revue de la littérature succincte concernant son émergence sur le plan phonétique indique que la variante agressive de l'**ironie sarcastique** est courante dans les contextes d'interaction conflictuelle (Anolli, 2002 ; Bryant et Fox Tree, 2002 ; Nauke et Braun, 2011).²¹

²¹ Cette considération a également été faite pour le cas du débat politique, notamment par De Chanay & Kerbrat-Orrechioni (2006), Poggi & D'Errico (2010) et Charaudeau (2015).

Arrivés au terme de cette vue d'ensemble des caractéristiques vocales des attitudes agressives, il est utile de préciser ici qu'une partie conséquente des caractéristiques que nous venons de présenter sont issues d'études portant sur des corpus de parole expressive actée, c'est-à-dire produite par des comédiens en laboratoire (Scherer, 2003 ; Juslin & Laukka, 2003) et non sur des corpus de parole expressive naturelle, tels notamment ceux constitués de débats politiques.

Si les revues de la littérature font état d'un nombre croissant d'études en parole non actée (induction en laboratoire, corpus de parole naturelle), la prépondérance, encore à ce jour, de protocoles axés sur la simulation de parole expressive par des comédiens s'explique par le fait que ceux-ci permettent un contrôle maximal des conditions d'acquisition de données ainsi que d'effectuer des comparaisons systématiques sur un contenu lexical comparable *toutes choses égales par ailleurs*, garanties de la robustesse des analyses phonétiques perceptives et productives effectuées ensuite sur le corpus.

Malgré ces avantages méthodologiques, la parole expressive actée doit être considérée avec précaution quant à sa représentabilité par rapport à de la parole expressive naturelle. Si la distinction simulé-naturel n'est pas toujours effectuée par les auditeurs (Audibert et collègues, 2008), la parole simulée comporte des caractéristiques perceptives et productives qui ne se retrouvent pas nécessairement dans la parole expressive naturelle). Il a notamment été montré que les attitudes expressives sont plus aisément distinguées les unes des autres en parole simulée qu'en parole spontanée, mais également que la parole expressive naturelle est moins marquée par une variabilité mélodique que la parole simulée (Jurgens et collègues, 2011 ; Audibert et collègues, 2010).

L'étude de parole expressive spontanée produite en situation écologique comporte quant à elle des obstacles pour l'analyse robuste des données tels que la qualité variable du signal, et l'impossible non contrôle et non comparabilité du contenu lexical et phonétique produit. Ce dernier point nous semble essentiel. À notre sens, le fait que les caractéristiques expressives de la parole politique aient si peu été étudiées pourrait résider dans la difficulté d'effectuer des analyses systématiques sur de la parole produite hors laboratoire.

Nous présentons alors à partir de la page suivante dans le chapitre II l'ensemble des précautions et ajouts méthodologiques qui ont été apportés à notre étude d'un corpus d'extraits de parole politique pour permettre la mise en place d'un protocole d'analyse phonétique expérimentale des caractéristiques de parole expressive sur un corpus d'extraits de débats politiques.

III. Protocole de recueil et traitement des données

Dans ce chapitre, nous présentons les choix qui ont été effectués dans le but d'obtenir un corpus d'extraits de parole expressive issue de débats politiques.

Nous avons pris pour objet d'étude les séances télévisées du conseil municipal de Montreuil en 2013. À cette période, pour réguler les échanges vifs et fortement conflictuels, un système de régulation des tours de parole a été installé. Cette configuration particulière a permis l'acquisition d'extraits de parole non superposée et rendant de ce fait possibles des analyses phonétiques acoustiques. Une segmentation en tours de parole a été effectuée sur 24 heures d'archives ainsi qu'une annotation d'extraits porteurs d'expressivité vocale de degrés variés, notamment de l'ordre d'attitudes agressives contrôlées. Une première sélection de 200 extraits (40*5 locuteurs, durée moyenne : 8,6 secondes) a été soumise à validation perceptive par 24 participants francophones natifs (voir en IV.3 pour le détail du protocole et les résultats).

À partir des résultats de cette première analyse perceptive, un sous-ensemble de 125 stimuli (25 stimuli par locuteurs) a été sélectionné (**corpus original**) et a fait l'objet d'une relecture par chaque locuteur de ses propres extraits transcrits orthographiquement (**corpus contrôle**) dans le cadre d'entretiens individuels. Les locuteurs ont reçu pour consigne de fournir une lecture « neutre » des extraits, de la façon la plus calme et intelligible possible, à la manière d'une dictée. Cette condition de contrôle rend possible des analyses comparatives entre extraits originaux et leurs relectures. Les données ont été segmentées en phonèmes, syllabes et mots. Une annotation syntaxique à deux niveaux (unités majeures et intermédiaires) a été effectuée et nous décrivons ici les méthodes suivies. Cette annotation syntaxique a permis, après annotation prosodique (voir Chapitre V) une quantification de l'écart entre corpus original et relecture en termes de structuration prosodico-syntaxique des extraits (voir Chapitre VII).

III.1. Description du conseil municipal de Montreuil en 2013

III.1.1. Un contexte politique tendu

Notre étude se base sur des débats issus d'archives télévisées de séances du conseil municipal de Montreuil en 2013, durant les derniers mois de mandat de Maire de Dominique Voynet (Europe Écologie Les Verts). Les séances du conseil municipal de Montreuil sont retransmises en direct sur une chaîne locale (TVM Est Parisien) et les archives libres sont ensuite consultables et téléchargeables sur le site internet de la municipalité.

Au cours des séances sélectionnées dans cette étude, Dominique Voynet est encore pressentie comme potentiellement en campagne pour un second mandat¹. Pendant cette période, et depuis quelques années déjà, le climat politique est très tendu et le déroulé des séances du conseil est houleux. De nombreux articles de presse relatent régulièrement un ensemble d'affrontements verbaux et, plus rarement, physiques entre la majorité et les partis d'opposition. La durée des séances, de 6h en moyenne, rend également compte des fortes tensions entre la majorité et les partis d'opposition ; le conseil démarrant à 19h pour ne pas finir avant 1h voire 2h du matin.

III.1.2. Une interaction régulée

Depuis juin 2013, un dispositif a été mis en place pour réguler les tours de parole durant le conseil municipal et de ce fait réduire les risques de chevauchements de parole. Géré à partir d'un poste informatique confié à la Maire, étant aussi bien modératrice du débat qu'intervenante, le système de régulation des tours de parole ne permet d'ouvrir qu'un micro à la fois, selon l'ordre dans lequel les locuteurs doivent ou veulent intervenir. Lors d'échanges particulièrement animés, les intervenants essaient parfois de parler fort hors micro pour se couper mutuellement la parole mais l'acoustique de la salle dans laquelle a lieu le conseil municipal rend très difficile l'écoute de ceux-ci s'ils ne sont pas amplifiés par un microphone.

L'instauration de ce système de gestion des tours de parole donne lieu à des interactions en différé, au cours desquelles les interlocuteurs ne dialoguent pas spontanément mais doivent attendre leur tour pour s'exprimer.

¹ Le 25 novembre 2013, Dominique Voynet annonce par déclaration envoyée à l'AFP qu'elle ne se représentera pas une seconde fois aux élections municipales. Dans sa déclaration, elle fait état d'une « dégradation de la vie politique », et qui plus est du climat politique montreuillois.

Toutefois, malgré cet effet différé, les locuteurs semblent s'efforcer à restituer un degré vif d'implication lorsqu'enfin on leur donne la parole, parfois plus d'une demi-heure après l'avoir demandée².

III.1.3. Déroutement type d'une séance

Le conseil municipal de Montreuil en 2013 est constitué de 50 conseillers municipaux. Chaque séance du conseil municipal a un ordre du jour constitué à cette période de plus de 40 délibérations à voter. Le déroulement des séances du conseil répond à un ensemble de règles et obligations administratives. Nous présentons ici le déroulement type d'une séance du conseil en cette période.

Après avoir fait l'appel des conseillers municipaux présents et des procurations, après avoir précisé les amendements à l'ordre du jour et les modifications du compte rendu de la séance précédente, la Maire déclare ouverte la séance. Un conseiller municipal élu de la majorité présente la première délibération à l'ordre du jour. Les membres de l'opposition peuvent alors demander la parole pour plus d'informations ou pour faire une remarque³, ce à quoi la majorité peut répondre, soit en la personne de la Maire, soit de l'élu qui avait présenté la délibération et qui leur répond. Une fois le débat clos l'assemblée vote et les conseillers passent à la délibération suivante dans l'ordre du jour.

La durée des délibérations est variable et dépend de l'importance des faits à voter ainsi que de leur caractère potentiellement polémique. Ainsi, si certaines délibérations sont votées en moins d'une minute, d'autres peuvent engager des débats pouvant dépasser 1h30 de durée (vidéo-protection, migrants, etc.).

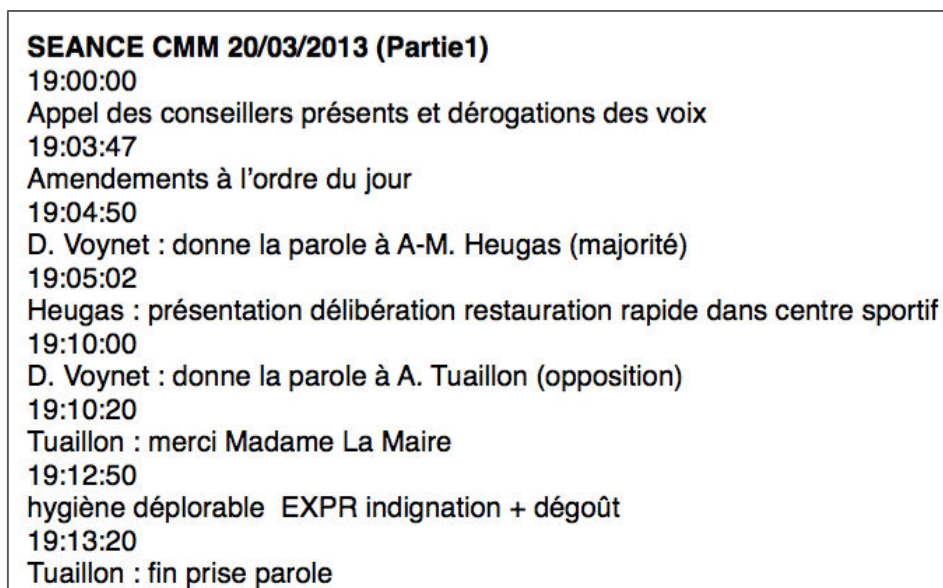
² Le temps d'attente peut être très long dans le cas de délibérations donnant en lieu à un grand nombre de réactions de la part de l'opposition, de par leur thématique polémique, ou même la nature de leur mise en place (délibération ajoutée à l'ordre du jour tardivement, documents non rendus publics à temps).

³ Lors de leurs interventions, les membres de l'opposition s'adressent parfois directement à l'intervenant de la majorité qui avait présenté la délibération, parfois à la Maire, d'autres fois à l'auditoire en général.

III.2. Des archives brutes à la constitution du corpus et sa condition de contrôle

III.2.1. L'annotation large des séances étudiées (24h d'archives)

Six séances du conseil municipal ayant eu lieu entre le 4 juin et le 21 novembre 2013 ont été étudiées, soit une durée totale de 24h. Dans un premier temps, nous avons pris des notes en assistant aux séances du conseil municipal parmi les spectateurs, puis nous avons annoté le fichier d'archive audio-visuelles de chaque séance étudiée⁴. Cette première étape nous a permis de nous familiariser avec le déroulement des séances et les conditions d'enregistrement (variations de niveaux des micros, influence du bruit de fond). Les séances du conseil municipal étant entre autres constituées de présentations de projets, d'interventions de représentants citoyens et de moments de délibérations et de votes, nous avons cherché dans un premier temps à isoler les temps de débats à proprement parler. Nos notes, constituées de remarques sur les grands temps du débat et les moments d'émergence d'expressivité, ont été prises sur un petit ordinateur portable. À l'aide d'un raccourci clavier stipulant les temps, nous avons pu situer nos notes au sein d'une frise temporelle (voir capture d'écran ci-dessous).



SEANCE CMM 20/03/2013 (Partie1)
19:00:00
Appel des conseillers présents et dérogations des voix
19:03:47
Amendements à l'ordre du jour
19:04:50
D. Voynet : donne la parole à A-M. Heugas (majorité)
19:05:02
Heugas : présentation délibération restauration rapide dans centre sportif
19:10:00
D. Voynet : donne la parole à A. Tuillon (opposition)
19:10:20
Tuillon : merci Madame La Maire
19:12:50
hygiène déplorable EXPR indignation + dégoût
19:13:20
Tuillon : fin prise parole

Figure III-1 : Capture d'écran du début d'un fichier texte d'annotation du début d'une séance du conseil municipal. Les chiffres indiquent l'heure à laquelle se déroulent les événements annotés.

⁴ Les archives filmées du conseil municipal ne comportent pas d'éléments coupés au montage à l'exception des pauses (une à deux pauses par séances).

Une fois la séance passée et les archives filmées téléchargées sur le site de la Mairie, la frise temporelle mise en place lors de la prise de note nous a permis de convertir nos remarques en tires d'annotation sur le logiciel ELAN⁵. Ces tires étaient alignées aux fichiers vidéos⁶ grâce aux codes temporels, ce qui nous a permis de gagner un temps considérable au moment d'accéder à nouveau aux passages ayant attiré notre attention. Nous avons ensuite segmenté l'intégralité de chaque séance par délibération et par locuteurs.

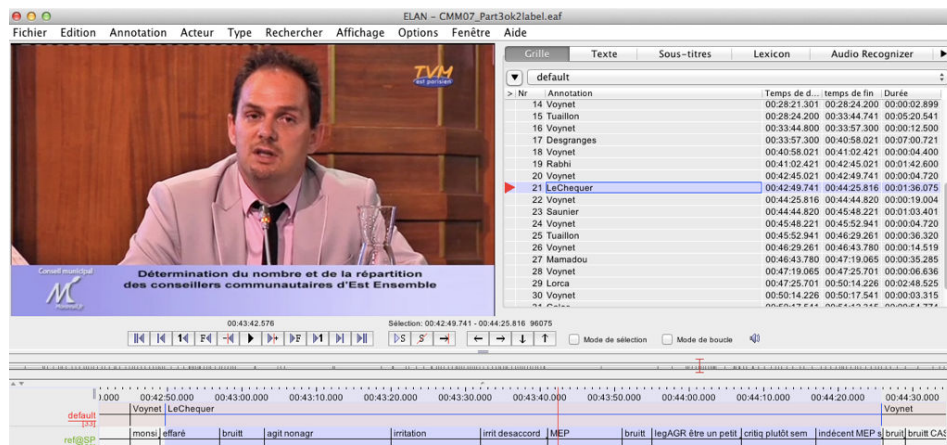


Figure III-2 : Capture d'écran d'un extrait de fichier d'annotation sur le logiciel ELAN et la segmentation en tours de parole par locuteurs. La tire du bas (titre vert clair) correspond aux notes textuelles brutes exportées puis corrigées manuellement.

⁵ <https://tla.mpi.nl/tools/tla-tools/elan/>

⁶ La conversion des notes textuelles sur ELAN a été effectuée grâce à un script élaboré par Nicolas Audibert.

III.2.2. La sélection des locuteurs

La présente étude se concentre sur la parole de cinq locuteurs : la Maire, Dominique Voynet (EELV) et quatre de ses opposants : Gaylord Le Chequer (RGC/PG), Frédéric Molossi (PS), Jean-Jacques Serey (FASE/PCF) et Alexandre Tuaille (RSM)⁷.

Outre leurs interventions fréquentes et échanges récurrents et animés avec la Maire, les quatre locuteurs masculins ont été sélectionnés pour leurs styles de parole dynamiques, s'écartant aussi souvent que possible de la parole lue pour aller vers un style proche de la parole spontanée ou pseudo-spontanée (improvisations autour d'un texte préparé) et ainsi permettre l'émergence de caractéristiques vocales expressives⁸.

Ces critères de sélection sont plus restrictifs qu'ils n'y paraissent, notamment du fait que malgré le nombre total élevé de conseillers municipaux (50), le nombre de locuteurs participants régulièrement dans les moments de débats à proprement parler est beaucoup plus réduit⁹. Si nous avions au départ prévu un nombre légèrement plus élevé de locuteurs étudiés, nous avons néanmoins été contraints de le réduire durant la première étape d'annotation large (détaillée ci-dessous) du fait, pour certains locuteurs, d'un trop faible nombre d'interventions ne permettant pas d'extraire un nombre suffisamment varié de prises de parole.



Figure III-3 : Photographies des locuteurs étudiés durant le conseil municipal de Montreuil (93100) en 2013 : Dominique Voynet (EELV) et quatre membres de l'opposition : Frédéric Molossi (PS), Gaylord Le Chequer (RGC/PG), Alexandre Tuaille (RSM) et Jean-Jacques Serey (PCF).

⁷ RGC : Rassemblement Gauche Citoyenne ; PS : Parti Socialiste, FASE/PCF : Parti Communiste Français ; RSM : Renouveau Socialiste Montreuillois. En 2013 à Montreuil, tous les partis minoritaires sont de gauche, il n'y a alors aucun parti de droite.

⁸ Sur le même principe, certains locuteurs bien que récurrents dans le débat n'ont produit durant leurs remarques que des interventions constituées de parole lue et n'ont donc pas été retenus.

⁹ Les séances sont également entre autres constituées de moments de présentations de projets et de votes de délibérations, lors desquels le recours à la lecture de texte préparés est majoritaire.

III.2.2.1. Sélection des extraits et protocole de segmentation

Une fois les locuteurs sélectionnés, nous avons alors pour chacun d'entre eux annoté un ensemble d'une centaine d'extraits sur lesquels nous percevons qualitativement l'émergence d'expressivité de degré variable.

Le type d'expressivité le plus récurrent parmi nos annotations consiste en une variété d'attitudes agressives plus ou moins contrôlées : irritation, sarcasme, ironie sarcastique, indignation¹⁰. Nous avons également fait en sorte d'intégrer pour chaque locuteur des extraits de parole non lue pour lesquels nous ne percevons pas d'hostilité particulière.

Nous avons alors appliqué un ensemble de critères restrictifs (listés ci-dessous) pour extraire de ces extraits des portions sonores courtes pouvant servir de stimuli dans des analyses perceptives¹¹. Le but de cette démarche était d'obtenir des extraits de 6 à 12 secondes, assez courts pour éviter un trop fort recours des auditeurs aux informations apportées par le contexte (Morange et Candea, 2010), mais néanmoins constitués d'extraits non seulement de mots ou groupes de mots isolés. Sur ce dernier point, pour des questions d'intelligibilité des extraits, nous avons autant que possible essayé de préserver des séquences syntaxiques cohérentes (unité syntaxique intermédiaire ; Tanguy et collègues, 2012) et privilégier des frontières de phrase¹², de continuation majeure entre deux unités syntaxiques majeures, ou de continuation mineure entre deux unités syntaxiques intermédiaires ; voir tableau ci-dessous.

	FRONT. PHRASE	FRONT. SYNT. MAJEURE	FRONT. SYNT. INTERM.
VOYNET	15	19	6
LE CHEQUER	10	24	6
MOLOSSI	10	21	9
SEREY	12	20	8
TUAILLON	14	19	7

Tableau III—1 : Nombre de stimuli tronqués s'achevant à une frontière de phrase, d'unité syntaxique majeure, ou d'unité syntaxique intermédiaire.

¹⁰ Voir Chapitre II pour le détail sur les attitudes agressives.

¹¹ Cette tâche a été effectuée sur Praat à partir d'exportations des annotations effectuées au préalable sur ELAN.

¹² Ensemble d'unités syntaxiques à modalité assertive, injonctive ou interrogative ; Riegel et collègues (1994, p.103).

D'un point de vue qualitatif, les critères de restriction sont les suivants :

- Absence de parole superposée
- Absence bruit de fond trop audible tels que des applaudissements ou des huées.
- Pas de mention directe du nom de l'interlocuteur auquel s'adresse le locuteur
- Pas d'insultes ou injures directes.

L'exemple ci-dessous correspond à un extrait de 7.29 secondes dont le début a été tronqué pour éviter la mention directe du nom de la Maire et dont la durée pré-segmentation courte était de 16 secondes.

Exemple III-1 : Tuailon, stimulus n°6 (seule la portion de texte en gras correspond au stimulus).

Vous avez raison Madame Voynet **vous aviez dit en conseil municipal du 6 juin que vous alliez organiser un débat en conseil municipal suivant sur la tranquillité publique** et vous estimiez que c'était une chose acquise.

En suivant cette méthodologie, nous sommes parvenus à constituer un ensemble de 200 stimuli (40 par locuteurs) respectant ces différents critères. Cet ensemble a été soumis à évaluation perceptive de l'hostilité exprimée (aucune, légère, faible) et du destinataire présumé du message (auditoire ou interlocuteur particulier) ; voir Chapitre III Analyses perceptives, étude n°1 pour le détail du protocole d'évaluation et des résultats.

III.3. L'enregistrement de la condition de contrôle (corpus relecture)

III.3.1. Préparation des données

De manière à permettre au sein d'entretiens individuels de durée maximale de 2h, de faire effectuer à chaque locuteur une tâche de perception (auto-évaluation des stimuli) et une de tâche de production (enregistrement), nous avons réduit le corpus à un sous-ensemble de 25 stimuli par locuteurs soit 125 stimuli au total.

La sélection des 25 stimuli par locuteurs a été effectuée au regard des résultats de l'évaluation perceptive n°1 en se basant sur les scores d'hostilité perçue de façon à conserver pour chaque locuteur un panel de stimuli allant d'aucune hostilité perçue à une forte hostilité perçue (voir partie IV.3 pour présentation du protocole et des résultats).

III.3.1.1. Transcription orthographique simplifiée et étendue des extraits

Pour des raisons d'intelligibilité et pour faciliter la tâche de lecture pour les locuteurs, les transcriptions des extraits à relire ont été simplifiées et étendues.

À partir de transcriptions orthographiques effectuées pour servir de support à la première passe de segmentation automatique des données sur Praat avec le plug-in *EasyAlign* (Goldman, 2010)¹³, nous avons effectué un ensemble de simplifications pour permettre une lecture fluide des extraits en suivant les recommandations de Claire Blanche-Benveniste (2010) : les modifications effectuées sur les transcriptions correspondent à la suppression des marques de « disflue » (hésitations, répétitions et révisions) et à l'ajout de ponctuation (majuscules, virgules en frontières d'unités syntaxiques majeures et points en fin de phrases).

¹³ Au sein de ces transcriptions, les hésitations audibles, répétitions et auto-corrrections ont été systématiquement retranscrites orthographiquement de façon à assurer un fonctionnement optimal de l'alignement forcé.

- **Transcription fine** : oui euh on on a nous avons déjà eu comme le dit la d'ailleurs la délibération ce débat lors euh de l'un de nos de l'une de nos précédentes séances
- **Transcription simplifiée** : Oui, nous avons déjà eu, comme le dit la d'ailleurs la délibération, ce débat lors de l'une de nos précédentes séances.

Dans les cas de stimuli constitués de phrases inachevées, tronquées au niveau d'une frontière de continuation majeure ou mineure, nous avons étendu la transcription jusqu'à la frontière syntaxique suivante pour permettre aux locuteurs de produire la relecture de l'extrait de la façon la plus naturelle et aisée possible. L'insertion des transcriptions dans un contexte syntaxique plus large a permis aux locuteurs une lecture fluide et cohérente des transcriptions, ensuite redécoupés sur Praat pour correspondre aux stimuli originaux.

Exemple III-3 : Tuailon stimulus n°1 - transcription simplifiée et augmentée (seule la partie en caractère gras correspond au stimulus extrait).

Dans ce cas-là manifestement vous n'avez pas été écoutée, **ce qui pose alors la question de la responsabilité, au sein de votre administration générale, et au sein de votre propre équipe**, notamment en la personne de la Maire Adjointe déléguée à la tranquillité publique.

De façon à ne pas risquer de troubler la lecture des locuteurs, nous n'avons pas indiqué sur les polycopiés qui leur étaient confiés de distinction entre texte correspondant au stimulus tronqué et extension de la transcription comme on peut par contre le voir sur l'exemple ci-dessus¹⁴.

¹⁴ L'ensemble des transcriptions simplifiées pour relecture est consultable en annexes.

III.3.2. Lieu et déroulement général des entretiens :

Nous avons rencontré les cinq locuteurs du corpus lors d'entretiens individuels entre décembre 2013 et mars 2014. La passation des entretiens s'est effectuée dans une pièce calme, dans un bureau municipal ou au domicile des locuteurs. Les enregistrements ont été effectués sur ordinateur portable à l'aide d'un micro-casque (AKG C520) relié à une carte son M-AUDIO (M-TRACK II).

Nous présentons tout d'abord le déroulement des entretiens individuels dans leur totalité avant de détailler les points relatifs à l'enregistrement de la condition de contrôle.

- 1) rappel des consignes, installation des hauts parleurs pour le test d'écoute
- 2) lecture et signature des attestations d'autorisation d'utilisation des enregistrements
- 3) test d'écoute (25 stimuli par locuteurs)
- 4) installation du micro-casque, discussion informelle
- 5) distribution des transcriptions et explicitation des consignes
- 5) calibration du microphone
- 6) enregistrement

Figure III-4 : Liste des étapes des entretiens individuels effectués avec les locuteurs.

Nous avons choisi de faire passer le test d'écoute avant le test de production afin de favoriser une réaction spontanée de la part des locuteurs lors de la tâche d'auto-évaluation perceptive. L'écoute préalable des stimuli originaux n'a à notre sens pas troublé la tâche de production car le but premier des locuteurs était de fournir une lecture neutre, de s'abstraire du poids des mots employés et de la remémoration du contexte de production.

III.3.3. Consignes et passation de la tâche de relecture

Voici les consignes telles qu'elles ont été lues à chaque locuteur :

« Le but de cet enregistrement est de fournir une lecture calme et claire et n'exprimant aucune attitude particulière des extraits qui vont suivre. Nous vous conseillons de vous abstraire autant que possible de la signification de votre message initial, de l'état dans lequel vous étiez, et de votre but à ce moment précis. Nous allons vous demander de lire ces extraits de la manière la plus désincarnée possible, à la manière d'une dictée : intelligible, fluide, neutre, en pensant à bien rendre cohérente la structure de vos phrases¹⁵. »

La spécification dans les consignes de produire la lecture des extraits à la manière d'une dictée se justifie par le fait que sans consigne plus précise, une tâche de lecture « neutre » peut être interprétée et donc réalisée différemment d'un locuteur à un autre (ton monotone, voix robotique, etc.). Ce choix s'est d'autant plus justifié du fait que nous avons affaire à des professionnels de la parole publique, plus ou moins aguerris de la lecture oralisée de documents. Pour certains locuteurs, l'analogie avec une tâche de lecture type dictée semble également avoir été un élément facilitant pour évaluer eux-mêmes la qualité des lectures effectuées. Certains locuteurs se sont même déclarés amusés par la lecture neutre d'extraits d'interventions qui étaient très vives ou ironiques durant le conseil, néanmoins cet amusement induit ne nous a pas semblé avoir d'incidence sur leurs productions.

La calibration du niveau d'entrée du microphone a été effectuée en prenant pour support le premier extrait de transcription à produire. Cette étape a également permis d'effectuer une première vérification de la compréhension des consignes et de fournir des précisions lorsque cela nous a paru nécessaire, notamment dans des cas de locuteur produisant l'ensemble de leur lecture en hyper-articulant ou avec un débit de parole excessivement lent.

Une fois la calibration du micro et les premiers essais effectués, nous avons commencé l'enregistrement et incité les locuteurs à conserver un débit de parole relativement homogène tout au long de l'enregistrement¹⁶. Compte tenu de la longueur des phrases à produire, et de la durée imposée des entretiens, nous n'avons pas pu envisager de faire produire une répétition systématique de chaque énoncé.

¹⁵ Nous avons dans certains cas dû préciser qu'une lecture de type « dictée » ne signifiait pas pour autant dire à haute voix les marques de ponctuation, ni répéter des groupes de mots.

¹⁶ Nous avons en effet prévenu les locuteurs en cas de ralentissements du débit de parole dû à la monotonie de la tâche de lecture ou au contraire à son accélération, due à la volonté de terminer la tâche au plus vite.

On compte néanmoins quelques cas pour lesquels les locuteurs ont demandé spontanément à recommencer la lecture d'un extrait, n'étant pas satisfaits de leur production. Nous avons également demandé aux locuteurs de réitérer la lecture d'un extrait si celui-ci comportait une erreur (omission ou inversion d'ordre de mots).

III.3.3.1. Extraits relus rejetés du corpus final

Au total, neuf extraits ont dû être rejetés du corpus : deux extraits pour cause d'erreurs de notre part dans la simplification des transcriptions orthographiques, et sept pour cause d'erreurs de lecture passées inaperçues durant les sessions d'enregistrement (ajout ou omission de mots).

Le cas le plus fréquent d'erreur de lecture concerne l'ajout de particules de négation, absentes de la transcription et pourtant restituées lors de la lecture sans que nous nous en apercevions. Malgré toute notre attention, de tels phénomènes sont passés inaperçus pour le locuteur comme pour nous qui lisons silencieusement le texte en même temps qu'eux. Ce phénomène a été notamment déjà observé par Claire Blanche-Benveniste (2010) qui l'explique par l'influence du code de l'écrit sur la production de l'oral, amenant le lecteur à modifier intuitivement le texte en voulant éviter de produire un énoncé qui risquerait d'être jugé incorrect.

	EXTRAITS APPARIABLES	LECTURES ERRONÉES	TRANSCR. ERRONÉES
VOYNET	24	0	1
LE CHEQUER	23	2	0
MOLOSSI	22	3	0
SEREY	22	2	1
TUAILLON	25	0	0
Total	116	7	2

Tableau III—2 : détail du nombre d'énoncés relus appariables au corpus original, de lectures erronées et de transcriptions erronées sur les 125 extraits enregistrés durant la tâche de lecture.

Comme le résume le tableau III—3 ci-dessus, 116 extraits relus sur les 125 enregistrés vont pouvoir être comparés aux productions originales.

III.4. Segmentation et annotation lexico-syntaxique des données

Suite à l'enregistrement de la condition « relecture », nous avons donc obtenu 116 paires de d'extraits de parole en condition originale et en relecture. Dans cette partie, nous allons présenter en détail les étapes de traitement et d'annotation des données qui ont permis ensuite d'effectuer des analyses phonétiques comparatives entre les deux conditions, notamment en comparant la structuration syntaxique des énoncés (annotées en suivant les critères explicités en III.4.2) et les structurations prosodiques¹⁷ selon la condition du corpus (annotées suivant les critères explicités en V.1).

III.4.1. Segmentation et annotation segmentale

III.4.1.1. Phonèmes, syllabes et mots

Chaque extrait segmenté a été transcrit orthographiquement avec une transcription fine faisant état des pauses remplies audibles telles que les « euh », les auto-corrrections et les répétitions. Une segmentation automatique en mot, syllabe, et phone, a été réalisé sur nos données grâce au plugin *EasyAlign* de Praat (outil d'alignement forcé créé en 2004 par Jean-Philippe Goldman), puis une correction manuelle a été effectuée. D'autres outils tels que SPPAS (Bigi et Hirst, 2012) existent de nos jours, cependant nous avons préféré utiliser *EasyAlign* car nous avons appris suite à de nombreuses années d'utilisation à connaître ses failles et savoir repérer efficacement les erreurs potentielles.

Sans surprise, les stimuli issus de la condition « relecture » ont nécessité moins de temps de correction car les productions étaient plus canoniques qu'en parole spontanée. Le corpus original a demandé plus de corrections. En effet, *EasyAlign* fonctionne moins bien sur la parole spontanée, qui plus est expressive, que sur la parole lue : les algorithmes mis au point pour la segmentation automatique reposent entre autres sur des seuils limites de durée de segments. Avec les corpus de parole expressive, il peut arriver que certains segments subissent de fortes modifications notamment de durée et ceci entraîne des erreurs de segmentation. De telles erreurs peuvent également survenir en cas de débit articulatoire extrêmement élevé, induisant une réduction du temps de durée moyenne des syllabes inaccentuées faisant passer celles-ci en deçà du seuil minimal fixé par l'algorithme. Outre ce type d'erreurs dues à des valeurs extrêmes

¹⁷ En effet, si la structuration syntaxique des énoncés est la même dans le corpus original et le corpus relecture, la structuration prosodique quant à elle peut différer et donner lieu à différentes frontières prosodico-syntaxiques pour chaque extrait selon la condition.

sortant des seuils fixés par les algorithmes d'*EasyAlign*, tout mot subissant une réduction comparativement à sa phonétisation canonique incluse dans le dictionnaire de prononciation utilisé a demandé une correction manuelle accrue (élisions et/ou insertions des schwas, assimilations de nasalité).

III.4.1.2. Annotation des marques de « disfluence »

Pour plus de clarté et de comparabilité, nous utilisons ici le terme « disfluence » cependant nous le maintenons entre guillemets pour signifier notre distanciation quant à la définition première de celui-ci. En effet, depuis quelques décennies, le terme de « disfluence » est de plus en plus discuté dans les études en parole non pathologique car la production de ce type de phénomènes ne produirait pas systématiquement d'interruption dans le flux de parole (Candea, 2017). Certains chercheurs contemporains vont même jusqu'à parler de marques de « (dis)fluence » (Grosman, 2015), sous-entendant que la production de pauses remplies puisse être au contraire une façon de maintenir la fluence du discours. D'autres chercheurs comme Maria Candea (2000, 2017) préfèrent à la notion de marque de « disfluence » celle de « marque du travail de formulation » (ou TdF) en partie parce qu'une telle dénomination ne sous-entend pas d'impact négatif sur la fluidité du discours.

L'annotation des marques de « disfluence » telle qu'elle a été effectuée dans la présente étude se limite à trois phénomènes : les pauses remplies de type « euh », les répétitions (« *de de de la municipalité* ») et les révisions (« *on a nous avons vu* ») ; voir tableau ci-dessous.

	VOYNET	LE CHEQUER	MOLOSSI	SEREY	TUAILLON
Pauses remplies (« euh »)	19	31	18	5	57
Répétitions	4	6	5	3	8
Révisions	10	1	6	2	5

Tableau III—3 : Nombre de marques de pauses remplies (« euh »), de répétitions, et de révisions par locuteurs dans le corpus original.

Lors de la transcription fine du corpus, un seuil limite de deux réécoutes a été fixé pour la transcription des pauses remplies de type « euh ». Cette restriction a été mise en place de façon à éviter de coder comme pause remplie des schwas épenthétiques, plus courts et moins aisément perceptibles (Candea, 2000).

Les répétitions annotées comme telles peuvent concerner des mots isolés (majoritairement des articles et des connecteurs) mais également parfois des groupes entiers de mots.

Conformément aux recommandations du manuel ESTER 2 (Galliano et collègues, 2009) ayant également servi de repère pour l'annotation du corpus ETAPE (Gravier et collègues, 2012) constitué de parole journalistique et politique, nous avons veillé à ne pas annoter comme répétitions les cas de répétitions de groupes de mots semblant être produites consciemment par les locuteurs à des fins stylistiques ; voir exemple ci-dessous.

Exemple III-4 : Le Chequer, stimulus n°1 (transcription orthographique simplifiée).

Vous avez voulu faire de ce sujet <i>un clivage</i> , un clivage politique entre vous et votre prédécesseur.
--

III.4.2. Annotation syntaxique

La méthodologie d'annotation syntaxique adoptée dans notre étude est analogue à celle appliquée par Degand et Simon (2009, 2011) et suit les recommandations regroupées dans le protocole de codage de Tanguy et collègues (2012) que nous avons suivi (sauf cas contraires explicités plus bas).

Nous avons effectué cette annotation dans le but de la confronter à une annotation des frontières prosodiques majeures pour ultimement comparer la réalisation prosodico-syntaxique des énoncés spontanés et lus, notamment à travers l'analyse de la distribution des pauses silencieuses selon leur position dans la structure syntaxique¹⁸. De plus, du fait du caractère hétérogène et qui plus est complexe des phrases du corpus¹⁹, il nous a paru indispensable de pouvoir avoir accès à la structure syntaxique des énoncés pour éclairer certaines analyses phonétiques par la suite.

Si nous avons suivi les recommandations du manuel de Tanguy et collègues (2012) pour effectuer l'annotation syntaxique de notre corpus, nous précisons toutefois que cette annotation est effectuée dans notre étude dans le but d'apparier les deux conditions de notre corpus et d'effectuer une analyse de la réalisation prosodique des unités syntaxiques. Nous n'effectuons

¹⁸ Voir Chapitre V pour le protocole d'annotation prosodique et descriptions des résultats, et chapitre VII pour la confrontation entre frontières syntaxiques et prosodiques.

¹⁹ Nous insistons sur ce point de justification de l'étape d'annotation syntaxique sur notre corpus car celle-ci n'est pas autant nécessaire dans l'analyse de corpus de parole expressive de laboratoire, où les structures de phrases sont majoritairement moins complexes voire inexistantes (mots isolés).

pas dans cette thèse d'analyse des unités syntaxiques à proprement parler. L'annotation syntaxique appliquée ici suit cette structuration à deux niveaux bien qu'elle prenne l'unité réactionnelle (UR) comme unité majeure, annotée dans l'exemple ci-dessous entre crochets et la séquence syntaxique comme unité intermédiaire, annotée entre parenthèses.

Exemple III-5 : Tuailon, stimulus n°1 composé d'une unité syntaxique majeure et 5 unités syntaxiques intermédiaires

[(Ce qui pose)_{SV} (alors)_{SADv} (la question de la responsabilité)_{SO} (au sein de votre administration générale)_{SRd} (et au sein de votre propre équipe)_{SRd}]^{urv}

III.4.2.1. Découpage en unités syntaxiques majeures

L'unité syntaxique majeure considérée dans le manuel de Tanguy et collègues (2012) est l'unité réactionnelle. Cette unité est quasiment équivalente à la proposition verbale (on parlera alors d'unité réactionnelle verbale, « urv »²⁰) mais elle est plus flexible en ce sens qu'elle peut également accepter les cas pour lesquels le noyau verbal est absent mais déductibles du contexte (unités réactionnelles elliptiques, « ure ») et les unités régies par un noyau non verbal (unités réactionnelles averbales, « ura »). À cette typologie nous avons ajouté deux suffixes :

- un suffixe « F » permettant de spécifier les cas d'unités réactionnelles marquant la fin d'une phrase syntaxique (urvF, ureF, uraF) de façon à pouvoir distinguer ultérieurement les frontières de fin de phrase de frontières d'unités syntaxiques majeures non terminales.

- un suffixe « S » permettant de signaler la présence d'une proposition subordonnée à l'intérieur de l'unité réactionnelle annotée (urvS).

²⁰ Les acronymes relatifs au niveau d'annotation syntaxique majeure sont notés en caractères minuscules (Tanguy et collègues, 2012).

III.4.2.2. Découpage en unités syntaxiques intermédiaires

Les unités intermédiaires composant chaque unité syntaxique majeure sont annotées à partir des catégories suivantes :

- Le syntagme verbal (SV), le syntagme sujet (SS) et le syntagme objet (SO),
- Le syntagme régi (SR) avec une spécification si le syntagme se trouve à droite (SRd) ou à gauche (SRg) du noyau de l'unité réactionnelle.
- Le syntagme adverbial (SADv) et le syntagme adjectival (SADj) sont également utilisés.
- Les marqueurs de discours (MD) et le syntagme adjectif (a) ; (ad) si celui-ci se trouve à droite, (ag) à gauche de l'élément qu'il complète dans l'unité réactionnelle.

III.4.2.2.1. Le cas particulier des subordonnées

Dans le manuel d'annotation de Tanguy et collègues (2012) les subordonnées sont à considérer et annoter comme de simples séquences syntaxiques comme on peut l'observer dans l'exemple ci-dessous directement repris du manuel²¹.

Exemple III-6 : Annotation selon le manuel de Tanguy et collègues (2012) d'une unité réactionnelle verbale contenant une subordonnée à fonction objet. Exemple issu du manuel (p. 4).

[(je pense)_{SV} (que votre éditeur il préfère vous avoir que pas vous avoir)_{SO}]^{urV}

Nous avons fait le choix de spécifier les séquences constituant les propositions subordonnées dans notre corpus. Nous justifions ce choix d'une part par la fréquence avérée de ce type de constructions dans la parole politique, et d'autre part du fait que celles-ci contiennent généralement un grand nombre de mots (Blanche-Benveniste, 2010). En effet, Claire Blanche-Benveniste (2010) souligne le fait que lorsque l'on argumente, les constructions syntaxiques à subordinations, voire à subordinations enchâssées, sont fréquentes. Nous avons également procédé à ce découpage supplémentaire de manière à pouvoir rendre compte plus finement des phénomènes prosodiques (accentuation, allongements, contours mélodiques, éventuelles insertions de pauses silencieuses) qui pourraient être liés à des formes particulières de structurations prosodiques sur ce type de phénomènes syntaxiques, nous avons fait le choix de spécifier les séquences syntaxiques elles-mêmes contenues dans les subordonnées.

²¹ Selon cette conceptualisation de l'annotation syntaxique, se limiter à une annotation syntaxique à deux niveaux implique de ne pas effectuer de découpage syntaxique à l'intérieur de subordonnées même si celles-ci sont elles-mêmes constituées de séquences sujet, verbe, objet (etc.) car le découpage en séquences doit se faire « uniquement par rapport au verbe recteur de premier niveau syntaxique » (Tanguy et collègues, 2012, p. 5).

III.4.2.3. Décompte des unités syntaxiques annotées

On compte au total 228 unités syntaxiques majeures (unités *SyntMaj*). Locuteur par locuteur, on décompte au total 37 unités *SyntMaj* pour Le Chequer, 44 unités *SyntMaj* pour Voynet et Tuillon, 51 pour Serey et 52 pour Molossi. On compte au total 690 unités syntaxiques intermédiaires. Locuteur par locuteur, le locuteur pour lequel on décompte le plus petit nombre est Serey (92). Les autres locuteurs produisent entre 141 (Le Chequer) et 166 unités syntaxiques intermédiaires (Tuillon).

Tableau III—4 : Nombre d'unités syntaxiques majeures (*SyntMaj*) et intermédiaires (*SyntInt*) annotées

	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
SyntMaj	44	37	52	51	44
SyntInt	147	141	142	92	166

III.4.2.4. Alignement de l'annotation syntaxique

L'annotation syntaxique de chaque extrait du corpus a ensuite été alignée avec les enregistrements originaux et relus à travers deux nouvelles tires dans les fichiers d'annotation au format TextGrid correspondant aux fichiers sons des extraits. Cette étape d'alignement prend en compte les variations de réalisation prosodique de ces unités majeures et intermédiaires du fait de l'éventuelle insertion de pauses silencieuses. L'analyse de la réalisation prosodique des énoncés n'est toutefois effectuée que dans un second temps, après annotation des frontières prosodiques majeures ; la tâche d'annotation prosodique étant effectuée selon les recommandations de Degand et Simon (2009, 2011) de façon totalement indépendante de la tâche d'annotation syntaxique.

Après présentation dans le chapitre suivant de la méthodologie et des résultats des analyses perceptives (**IV. Perception de l'hostilité et profils dimensionnels**), permettant notamment de consulter en **IV.3** les étapes de validation perceptive du corpus, nous présentons dans le chapitre **V** les critères et résultats de l'annotation prosodique et enfin dans le chapitre **VI** la confrontation des frontières syntaxiques et prosodiques majeures entre extraits originaux et relus.

IV. Perception de l'hostilité et profils dimensionnels

Le **Chapitre IV** présente deux études perceptives : **IV.3** Perception de l'hostilité exprimée et du destinataire présumé (interlocuteur particulier vs. auditoire) et **IV.4** : profils émotionnels tridimensionnels (valence, excitation, et dominance). Nous exposons pour chaque étude une présentation du protocole expérimental mis en place et les résultats obtenus, puis nous effectuons l'interprétation conjointe des résultats des deux études (**IV.5**).

La première étude perceptive (**III.1**), visant à sélectionner les extraits, a été menée sur 200 stimuli (40*5 locuteurs, durée moyenne : 8.6 secondes) ; elle a été conduite en parallèle dans deux conditions : un test de perception audio et un test de perception écrite (transcriptions orthographiques). 24 participants francophones natifs (12 par condition) ont évalué l'hostilité perçue et si les locuteurs semblent s'adresser à l'auditoire en général ou à un interlocuteur particulier. Dans les deux tests, les stimuli ont été perçus, tous locuteurs confondus, comme exprimant en moyenne une légère hostilité. Les scores d'hostilité perçue sont plus élevés en condition audio qu'à l'écrit, toutefois la différence n'est significative que pour trois locuteurs sur cinq. L'analyse comparée des scores obtenus à l'audio et à l'écrit a permis de quantifier pour chaque stimulus l'écart entre l'expressivité véhiculée par le contenu sémantique et celle ajoutée par les caractéristiques vocales des locuteurs.

La seconde étude perceptive (**IV.4**) consiste en la comparaison des résultats d'une auto-évaluation par les locuteurs eux-mêmes (25 extraits*5 locuteurs) de leurs productions et de la perception de 24 auditeurs francophones, sous la forme de profils émotionnels constitués des dimensions émotionnelles de valence (ici content/mécontent-e), excitation (ici calme/agité) et dominance/contrôle (faible/fort-e). En parallèle à l'auto-évaluation, un test similaire a été conduit avec 24 auditeurs francophones natifs. La comparaison systématique des résultats fait état d'écarts significatifs entre les informations apportées par les locuteurs dans leur auto-évaluation et les profils dimensionnels attribués par les auditeurs naïfs. Les différences interlocuteurs qui étaient significatives dans l'auto-évaluation ne se retrouvent pas dans la perception des auditeurs naïfs qui font état de profils homogènes exprimant un léger mécontentement, un degré neutre d'excitation et une légère dominance. Enfin, la comparaison des résultats des deux études (**IV.4**) a montré des corrélations significatives, tous locuteurs confondus, entre hostilité perçue, mécontentement, excitation et dominance perçue. L'hostilité et la dominance perçue sont significativement corrélées pour tous les locuteurs sauf Voynet, l'hostilité et l'excitation pour tous sauf Le Chequer.

IV.1. Enjeux de l'analyse perceptive et choix méthodologiques

Comme nous l'avons vu dans le chapitre 1, les enjeux pour les participants d'un débat politique influent sur leur façon de se présenter, leur *ethos*, car pour convaincre, ceux-ci doivent présenter une image favorable tout en véhiculant une attitude agressive contrôlée (Poggi & D'Errico, 2010).

Amorcer l'analyse phonétique de l'expressivité vocale dans le débat politique par sa composante perceptive nous semble donc particulièrement justifié, du fait que l'image que cherche à renvoyer le locuteur, son *ethos* voulu, dépend toujours de la façon dont il va être perçu par les autres, son *ethos* perçu (Amossi, 1999 ; Maingueneau, 1999). Une précaution de fond s'impose toutefois concernant un éventuel *ethos* préalable (Maingueneau, 1999) : plus les locuteurs sont présents sur la scène médiatique, plus ils sont directement associés à un type d'*ethos* avant même toute nouvelle déclaration de leur part. Cela signifie dans le cas d'une étude perceptive de parole politique la nécessité de prendre en compte si les participants connaissent les locuteurs ou non.

En phonétique expérimentale, les analyses perceptives représentent une étape indispensable à l'étude d'un corpus, dont les deux enjeux principaux sont la **validation** et l'**évaluation perceptive**. Outre le fait que dans le cas précis de cette étude cette démarche a pour but de contribuer à l'analyse de l'*ethos* voulu et de l'*ethos* perçu (**évaluation perceptive**), la pertinence et la fiabilité des analyses effectuées repose en grande partie sur la **validation perceptive** des phénomènes que l'on souhaite étudier.

La **validation perceptive** permet en premier lieu d'objectiver la sélection sur critères qualitatifs à travers laquelle le corpus a été constitué en tant qu'ensemble d'extraits exprimant, selon nous, des attitudes agressives de degré d'activation émotionnelle et de contrôle variables, grâce à une validation perceptive du corpus menée au travers de tests de perception.

Cependant, la validation perceptive d'un corpus d'extraits de débats politiques, même constitué d'extraits courts, de productions isolées de locuteurs, sans chevauchements de parole, est une étape complexe et nécessite une adaptation du protocole expérimental classique du fait que l'expressivité peut également être véhiculée au niveau lexical et syntaxique : ce pourquoi nous proposons dans un premier temps une double analyse perceptive croisée entre extraits audio et supports écrits.

IV.2. Principes pour la présentation des analyses statistiques

Nous présentons ici notre méthodologie d'analyse statistique et nos choix pour la représentation graphique des résultats présentés ; méthodologie qui sera en tous points similaire dans les chapitres d'analyses prosodiques.

IV.2.1.1. Méthodologie d'analyses statistiques

Notre méthodologie d'analyse statistique comprend trois étapes : les deux premières permettent de circonscrire l'ensemble des informations que nous souhaitons connaître pour chaque variable, la troisième permet de comparer les différentes variables analysées en perception et en production (analyses prosodiques) :

1) Analyses comparatives inter-conditions :

Nous comparons d'abord les moyennes de valeurs dans les extraits originaux et relus en effectuant des **tests t** appariés tous locuteurs confondus, puis locuteur par locuteur. Pour faciliter la lecture, les décimales sont marquées par un point de façon à être non confondues avec les différentes informations séparées par des virgules. Aussi, pour éviter le cumul de parenthèses imbriquées, les degrés de liberté sont présentés entre crochets.

2) Analyses comparatives inter-locuteurs au sein de chaque condition :

Nous effectuons ensuite au sein de chaque condition des **ANOVA** avec le locuteur comme facteur et en cas d'influence significative du locuteur nous détaillons les différences inter-locuteurs significatives avec pour test post-hoc l'HSD de Tukey.

3) Analyses comparatives inter-variables :

Pour comparer ensuite les rapports entre les différentes variables mesurées et analysées, nous effectuons des **corrélations** de Spearman, car celles-ci sont plus adaptées que celles de Pearson aux comparaisons entre variables incluant des variables d'échelle ordinale, tels que nos scores perceptifs issus d'échelles de Likert.

IV.2.1.2. Méthodologie de représentation visuelle des résultats

Le mode de représentation graphique privilégié des analyses comparatives présentées dans cette étude correspond à une superposition de deux représentations graphiques de la distribution d'un jeu de données : la boîte à moustaches et le diagramme en violon.

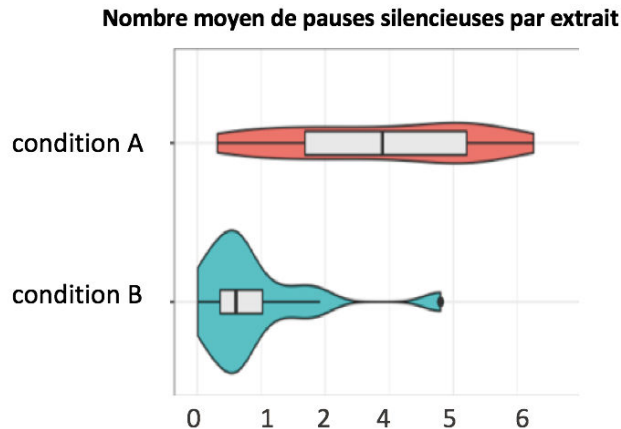


Figure IV-1 : Exemple de présentation en diagrammes en violons et boîtes à moustaches.

Nous présentons dans la figure ci-dessus un exemple fictif représentant une comparaison de distribution de données entre deux conditions pour faciliter la compréhension de la description des graphiques :

- Dans une représentation en boîte à moustaches, la barre verticale représente la valeur médiane, la boîte représente l'écart interquartile (soit les 50% de la distribution situés autour de la médiane), et la longueur des lignes extérieures à la boîte appelées « moustaches » indique la plage de données qui est comprise dans 1,5 fois l'écart interquartile.
- Un diagramme en violon représente la courbe de densité de la distribution des données, présentée de manière symétrique de chaque côté de la boîte pour plus de lisibilité et lui conférant ainsi parfois une forme de violon. Il constitue de plus un complément opportun à la boîte à moustache, car il permet de considérer visuellement les tendances de distribution impossibles à considérer à la lecture seule d'un graphique type boîte à moustache (distributions bimodales, asymétriques).

Enfin, nous précisons que si, pour une meilleure compréhension, nous avons opté dans la présentation de l'exemple fictif ci-dessus pour une visualisation horizontale des graphiques, ceux-ci seront généralement présentés à la verticale.

Plus précisément, lorsque nous présenterons les distributions de valeurs moyennes par stimulus, les graphiques seront systématiquement présentés à la verticale, et ne seront présentés à l'horizontale que lors de présentations de distributions de valeurs non moyennées.

IV.3. Étude n°1 : Hostilité perçue et estimation du destinataire du message

Le corpus étant composé d'extraits sélectionnés sur la base de leurs degrés variables d'hostilité¹ exprimée, il est nécessaire de confronter la sélection effectuée à la perception d'auditeurs dits « naïfs »², de façon à obtenir une validation perceptive du corpus et une quantification de l'hostilité exprimée dans les extraits.

Par une approche croisée en perception audio et écrite (transcriptions orthographiques présentées seules), l'enjeu de notre protocole expérimental est de pouvoir quantifier l'influence mutuelle du contenu sémantique et des caractéristiques vocales dans la perception de l'hostilité et l'estimation du destinataire du message (« à qui semblent s'adresser les locuteurs ? » - auditoire dans sa globalité ou interlocuteur particulier).

IV.3.1. Protocole méthodologique

Deux tests perceptifs ont été mis en place en parallèle : l'un portant sur la perception du corpus à l'oral (extraits audio originaux) et l'autre à l'écrit (transcriptions orthographiques simplifiées). Comme nous l'avons explicité dans le chapitre précédant (voir III.3.1), la transcription simplifiée a été effectuée pour faciliter la lecture et s'assurer de l'intelligibilité des extraits lors des évaluations perceptives (Blanche-Benveniste, 2010).

22 personnes francophones natives (11 participants par tests, âge moyen : 26 ans) ont participé aux tests. Le test écrit (8 femmes, 3 hommes) a été mis en place sous la forme d'un formulaire *Google* et le test audio (7 femmes, 4 hommes) à l'aide d'un test *Praat*.

¹ Nous avons préféré, pour la conduction des tests de perception, le terme d'hostilité plutôt que d'agressivité car nous avons remarqué dans une étude perceptive antérieure, portant également sur un ensemble d'attitudes agressives que l'agressivité perçue était systématiquement corrélée à l'activation émotionnelle (Koukolia, 2012). Pour pouvoir rendre compte de variations fines et tacites d'expressivité dans un corpus de parole politique, nous avons souhaité pour cette étude éviter de référer directement au terme d'hostilité avec les participants des tests. De plus, le terme d'hostilité, associé au fait de s'opposer (« être hostile à ») et pouvant être considéré sous une forme tacite (« hostilité latente ») nous a paru propice à la description d'extraits de débats politiques ; voir consignes données aux auditeurs page suivante.

² Dans un protocole expérimental impliquant une validation perceptive, la notion de participants « naïfs » renvoie au fait que ceux-ci effectuent la tâche sans aucune autre information que les consignes données (pas de connaissances des hypothèses de recherche du domaine, du protocole expérimental habituel, d'informations contextuelles, etc.). Dans le cas de notre étude, cela signifie également que ceux-ci ne connaissent pas les personnes politiques dont ils écoutent des extraits de parole, ne risquant pas alors d'être influencés par l'image qu'ils se seraient fait d'elles au préalable.

Les stimuli du test audio ont été présentés dans un ordre aléatoire, avec possibilité de réécouter une fois chaque extrait, ceux du test écrit dans un ordre pseudo-aléatoire³.

Dans le test audio, nous avons également demandé aux participants de spécifier dans une question en post-test s'ils avaient reconnu des locuteurs parmi ceux écoutés. Si nous avons craint que Dominique Voynet soit reconnue au détriment des autres locuteurs compte tenu de sa plus grande notoriété, aucun participant n'a déclaré reconnaître de locuteur particulier.

Les consignes, présentées ci-dessous, ont été les mêmes pour les deux conditions du test.

Le test suivant contient des extraits de parole issus d'un conseil municipal.

L'ensemble des extraits comporte des annonces, remarques, interactions spontanées effectuées par différents locuteurs.

Pour chaque extrait, il vous sera demandé de spécifier :

- le degré d'hostilité* exprimée : aucune, légère, ou forte.

- à qui le locuteur semble s'adresser sur une échelle de 1 à 5 allant de l'auditoire dans sa globalité (1) à un interlocuteur particulier (5) **.

*** Synonymes : inimitié, animosité, malveillance ; antonymes : bienveillance, sympathie.**

**** si cette information ne peut être déterminée, vous pouvez cocher le niveau 3 de l'échelle."**

Figure IV-2 : consignes du test de perception n°1 (similaires pour la version audio et la version écrite)

³ La plateforme *Google* ne permettant pas à ce moment de présenter les stimuli écrits dans un ordre aléatoire complet, nous avons pallié ce problème en créant 6 versions différentes du test comportant chacune un ordre aléatoire différent, de façon à ce qu'au maximum 2 participants par tests évaluent les stimuli dans le même ordre.

IV.3.2. Résultats de l'étude de l'hostilité perçue

Les résultats de l'analyse perceptive de l'hostilité sont présentés à la fois de façon catégorielle (aucune, légère, forte) ainsi qu'en considérant ces valeurs sur une échelle ordinale allant de 0 (aucune) à 2 (forte).

Nous présentons ici les résultats principaux concernant l'hostilité perçue en condition audio et écrite, mais également dans la différence d'hostilité perçue entre les deux conditions, à travers *dHost*, variable constituée à partir de mesures de différence entre score d'hostilité perçue à l'oral et à l'écrit pour chaque stimulus ($dHost = \text{score audio} - \text{score écrit}$). Nous présentons ensuite les rapports entre scores audio et écrits au travers de corrélations.

IV.3.2.1. Hostilité perçue : tendances générales

En condition audio, tous locuteurs confondus, la moyenne des scores d'hostilité perçue est de 1.11/2 (écart-type : 0.5), et passe à 0.88/2 (écart-type : 0.6) lorsque les stimuli sont présentés en condition écrite seule.

Tableau IV—1 valeurs principales des scores d'hostilité perçue à l'audio et à l'écrit et des valeurs de *dHost*, différence entre scores audio et écrit pour chaque stimulus (moyenne et écart-type, minimum, maximum). Correspondance des scores d'hostilité perçue : 0 = aucune hostilité, 1 = faible hostilité, 2 = forte hostilité.

	MOYENNE (ÉCART-TYPE)			SCORE MINIMUM		SCORE MAXIMUM	
	AUDIO	ÉCRIT	dHost	AUDIO	ÉCRIT	AUDIO	ÉCRIT
VOYNET	1.08 (0.54)	0.88 (0.69)	0.2 (0.60)	0.11	0	1.89	2
LE CHEQUER	1.14 (0.44)	0.88 (0.69)	0.25 (0.48)	0.22	0	1.67	2
MOLOSSI	1.18 (0.49)	0.98 (0.56)	0.19 (0.47)	0.22	0.15	2	1.92
SEREY	1.1 (0.54)	0.75 (0.57)	0.34 (0.37)	0.09	0	2	1.85
TUAILLON	1.06 (0.52)	0.86 (0.48)	0.2 (0.46)	0.09	0.08	2	1.92
Tous loc.	1.11 (0.51)	0.87 (0.60)	0.23 (0.47)	0.15	0.05	1.91	1.94

Comme l'indique le tableau ci-dessus, tous locuteurs confondus, la valeur moyenne de $dHost$ est de 0.23 (écart-type moyen : 0.47). L'interprétation des valeurs de la variable $dHost$ est la suivante : plus la valeur de $dHost$ est positive, plus elle réfère à des cas pour lesquels les stimuli sont perçus comme plus hostiles à l'oral qu'à l'écrit. Plus la valeur est proche de zéro et plus elle réfère à des stimuli perçus similairement quelle que soit leur présentation (audio vs écrit). Enfin, plus la valeur de $dHost$ est négative, plus elle réfère à des cas de stimuli perçus comme plus hostiles en condition écrite qu'en condition audio. Cette valeur de $dHost$ proche de zéro pour tous les locuteurs indique donc la faible différence moyenne entre scores obtenus à l'oral et à l'écrit pour les stimuli.

Tous locuteurs confondus, les stimuli sont en moyenne perçus comme exprimant moins d'hostilité lorsque présentés en condition écrite qu'en condition audio ($t[198]=5.5$, $p < .001$, $\eta^2=5\%$). Toutefois, comme l'indique la figure ci-dessous les scores d'hostilité perçue sont significativement plus hauts que ceux perçus à l'écrit seulement pour trois locuteurs sur cinq : Serey ($t[37]=4.5$, $p < .001$, $\eta^2=9\%$), Le Chequer ($t[39]=2.9$, $p < .01$, $\eta^2=5\%$), et Tuailon ($t[39]=2.9$, $p < .01$, $\eta^2=4\%$) ; voir tableau ci-dessus pour les valeurs moyennes.

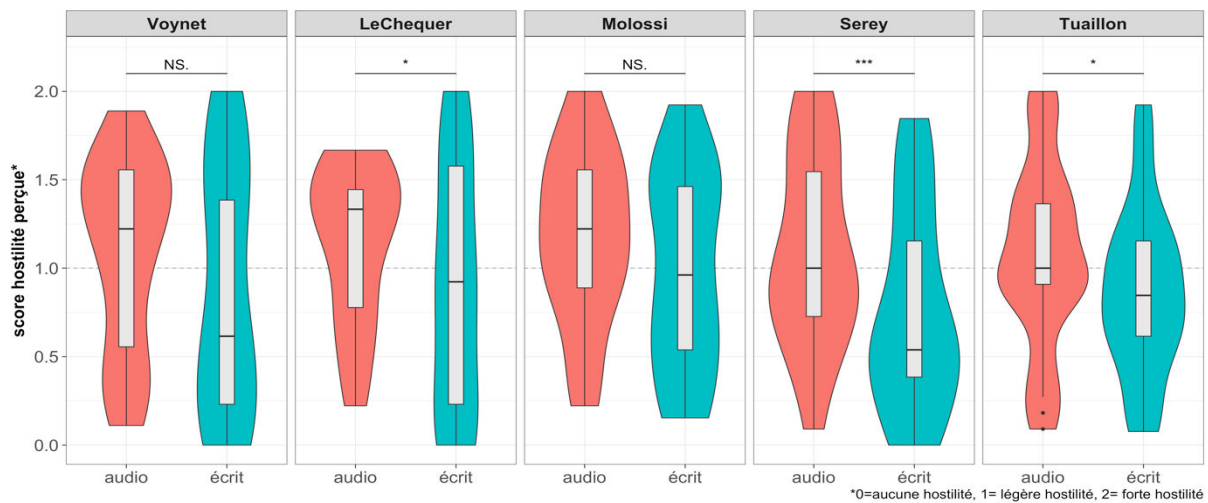


Figure IV-3 Diagrammes en violons et boîtes à moustaches des scores d'hostilité perçue dans le test audio et dans le test écrit sur 200 stimuli (40 x 5 locuteurs). Les étoiles représentent la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif. Correspondance des scores d'hostilité perçue : 0 = aucune hostilité, 1 = faible hostilité, 2 = forte hostilité.

IV.3.2.2. Rapports entre hostilité perçue à l'oral, à l'écrit et dHost

Tableau IV—2 corrélations de Spearman entre scores d'hostilité perçue dans le test audio et le test écrit, dHost (audio-écrit) et scores audio, dHost et scores écrits. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

	Hostilité audio / hostilité écrit	dHost / hostilité audio	dHost / hostilité écrit
VOYNET	$\rho = 0.505 (**)$	$\rho = 0.188 (NS)$	$\rho = -0.691 (***)$
LE CHEQUER	$\rho = 0.71 (***)$	$\rho = -0.211 (NS)$	$\rho = -0.759 (***)$
MOLOSSI	$\rho = 0.616 (**)$	$\rho = 0.319 (NS)$	$\rho = -0.508 (**)$
SEREY	$\rho = 0.733 (***)$	$\rho = 0.145 (NS)$	$\rho = -0.524 (**)$
TUAILLON	$\rho = 0.542 (**)$	$\rho = 0.563 (**)$	$\rho = -0.318 (NS)$
<i>Tous loc.</i>	$\rho = 0.628 (***)$	$\rho = 0.208 (*)$	$\rho = -0.575 (***)$

Les scores d'hostilité perçue à l'oral sont fortement inter-corrélés avec ceux du test écrit ($\rho = .628$, $p < .001$). Comme le montrent le tableau ci-dessus et la figure ci-dessous, cette corrélation est significative pour tous les locuteurs : plus leurs stimuli sont perçus comme hostiles à l'oral plus ils le sont également à l'écrit.

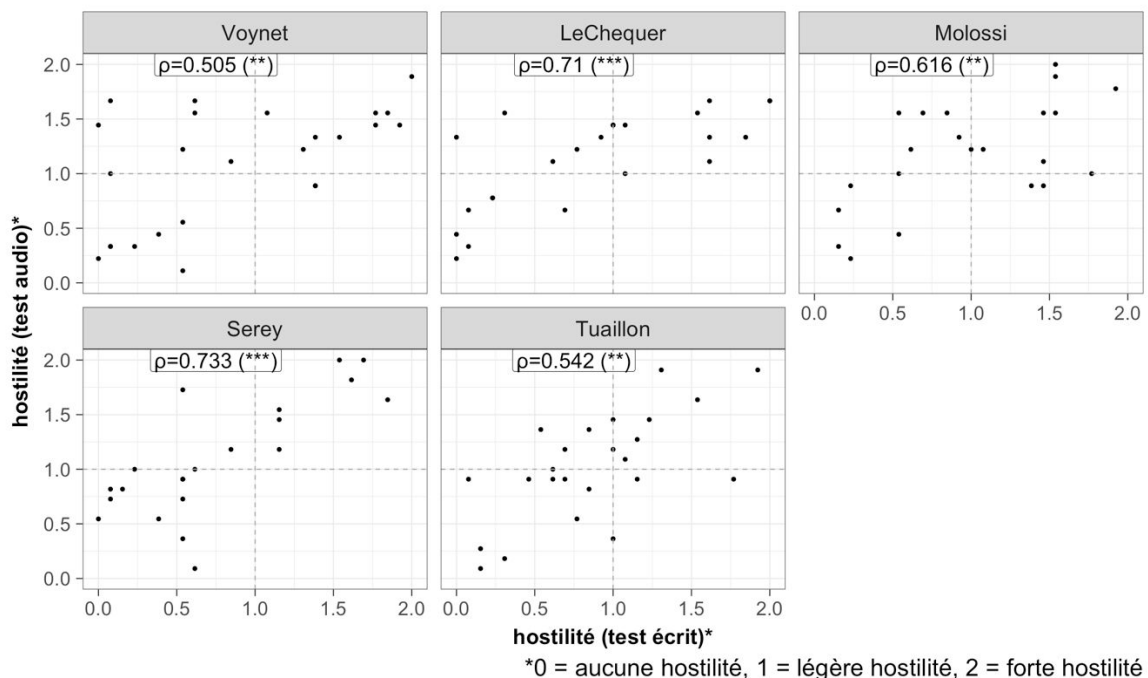


Figure IV-4 Corrélations entre scores d'hostilité perçue dans les test audio et écrit. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Tous locuteurs confondus, on calcule une corrélation négative significative entre valeurs de $dHost$ et hostilité perçue à l'écrit ($\rho = -.575$, $p < .001$). Les stimuli pour lesquels on trouve les valeurs de $dHost$ ⁴ les plus élevées sont ceux qui sont perçus comme les moins hostiles en condition écrite. Cette corrélation est significative pour tous les locuteurs sauf Tuillon.

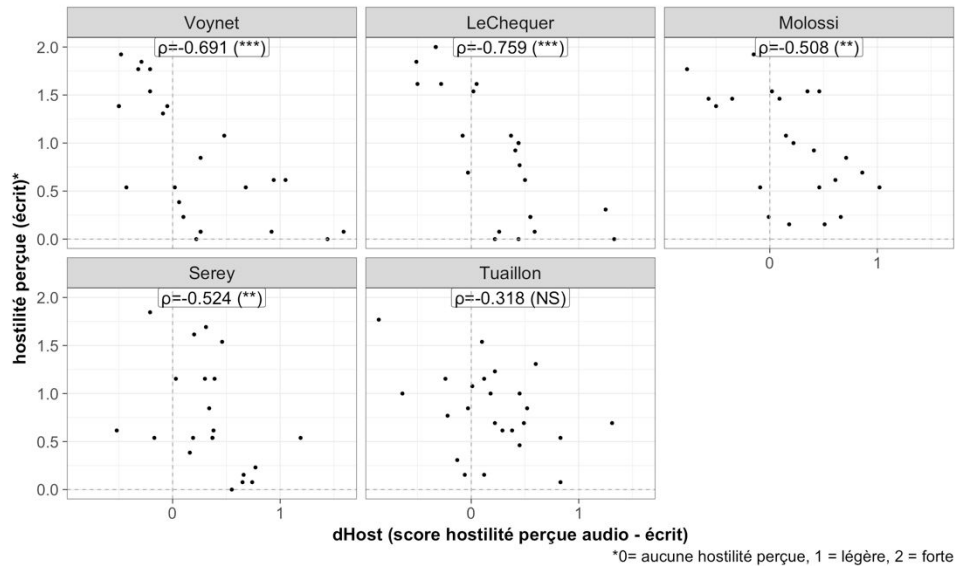


Figure IV-5 Corrélations entre hostilité perçue dans le test écrit et $dHost$, mesure de différence entre scores audio et scores écrits. Les étoiles représentent la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Tuillon est également le seul locuteur pour lequel on trouve une corrélation significative entre $dHost$ et scores d'hostilité perçue à l'oral ($\rho = -.563$, $p < .01$) : ses stimuli perçus comme les plus hostiles à l'oral sont ceux pour lesquels on trouve les valeurs de $dHost$ les plus élevées.

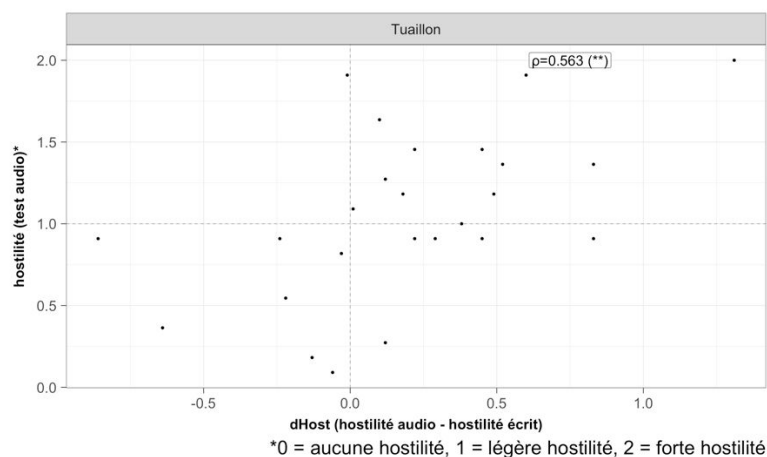


Figure IV-6 Corrélations entre hostilité perçue (audio) et $dHost$ pour Tuillon. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

⁴ Plus les valeurs de $dHost$ sont élevées plus il y a de forte différence entre hostilité perçue à l'oral et hostilité perçue à l'écrit.

IV.3.3. Résultats de l'estimation du destinataire supposé

Nous avons souhaité dans cette première étude perceptive, non seulement évaluer l'hostilité perçue, mais aussi quantifier l'impact des indices linguistiques et paralinguistiques dans l'estimation par les auditeurs d'à qui s'adressent les locuteurs : s'ils semblent plutôt s'adresser à l'auditoire dans sa globalité (score de 1/5) ou à un interlocuteur particulier (score de 5/5).

Comme pour l'hostilité perçue, nous présentons dans un premier temps les caractéristiques générales en condition audio, écrite, et à travers une mesure de différence entre scores audio et écrits, ici *dDest* ($dDest = \text{score audio} - \text{score écrit}$), puis les rapports entre estimation du destinataire à l'oral et à l'écrit à travers des corrélations.

IV.3.3.1. Caractéristiques générales

Tableau IV—3 valeurs principales des scores d'estimation d'à qui s'adressent les locuteurs en condition audio écrit et des valeurs de *dDest*, différence entre scores audio et écrit pour chaque stimulus (moyenne et écart-type, minimum, maximum). Correspondance des scores : 1=certainement à l'auditoire dans sa globalité, 3= indéterminé, 5=certainement à interlocuteur particulier.

	MOYENNE (ÉCART-TYPE)			SCORE MINIMUM		SCORE MAXIMUM	
	AUDIO	ÉCRIT	dDest	AUDIO	ÉCRIT	AUDIO	ÉCRIT
VOYNET	2.9 (1.05)	2.68 (1.19)	0.2 (0.60)	1.11	1.08	4.89	5
LE CHEQUER	3.46 (1.24)	3 (1.4)	0.25 (0.48)	1.33	1	4.89	5
MOLOSSI	2.91 (0.95)	2.59 (0.84)	0.19 (0.47)	1.78	1.38	5	4.46
SEREY	3.06 (1.18)	3.24 (0.91)	0.34 (0.37)	1.36	1.62	5	4.85
TUAILLON	3.65 (1.22)	3.7 (1.09)	0.2 (0.46)	1.36	1.77	5	4.85
Tous loc.	3.20 (1.13)	3.04 (1.08)	0.15 (0.72)	1.39	1.37	4.96	4.83

Tous locuteurs confondus, on ne trouve pas de différence significative entre scores audio et écrits ($t[115]=-2.1$, $p=0.3$, $\eta^2=0\%$) : le score moyen audio est de 3.20/5 (écart-type : 1.13) et à l'écrit de 3.04/5 (écart-type : 1.08).

La valeur moyenne de *dDest*, variable de différence entre score d'estimation du destinataire en condition audio et en condition écrite, est proche de zéro (0.15, écart-type : 0.72) ce qui

signifie, comme pour toutes les autres variables de différence de la présente étude, peu de différence entre les deux conditions.

Les comparaisons inter-condition par locuteur indiquent que Le Chequer ($t[22]=-3.4$, $p<.01$, $\eta^2=3\%$) et Molossi ($t[21]=-3.9$, $p<.001$, $\eta^2=3\%$) sont en moyenne perçus comme s'adressant plus directement à un interlocuteur particulier en condition audio qu'en condition écrite, comme on peut l'observer sur la figure ci-dessous.

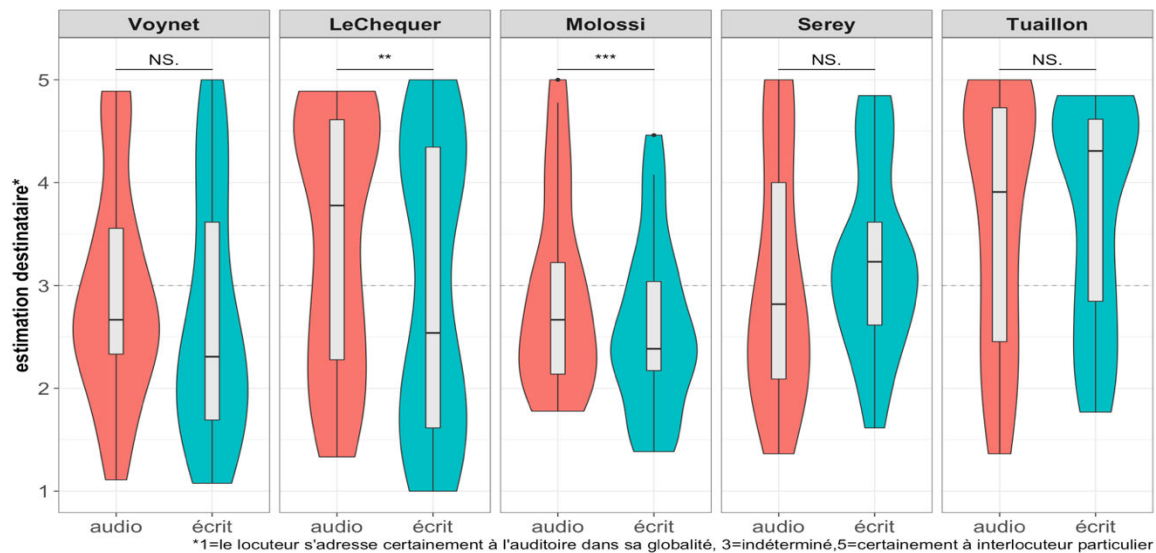


Figure IV-7 Diagrammes en violons et boîtes à moustaches des scores d'estimation du destinataire dans le test audio et dans le test écrit. Les étoiles représentent la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

Dans la condition audio, on trouve une influence significative des locuteurs sur la variation des scores ($F[4,111]=3.91$, $p<.01$; $\eta^2=12\%$) : les scores moyens obtenus par Tuailleon sont significativement supérieurs à ceux de Molossi ($p<.001$) et de Voynet ($p<.01$).

En condition audio, les stimuli perçus comme adressés à l'auditoire dans sa globalité représentent un tiers des productions de Voynet, Serey et Molossi. Les stimuli de Tuailleon et Le Chequer, deux des principaux adversaires de Voynet, sont majoritairement perçus comme s'adressant à un interlocuteur particulier. Cette catégorie perceptive (« s'adresse certainement à un interlocuteur particulier ») représente 50% des productions de Tuailleon, et 38% des productions de Le Chequer, contre 25% en condition écrite. Notons également que Tuailleon et Le Chequer sont les deux locuteurs pour lesquels on trouve les valeurs de $dDest$ les plus hautes, c'est-à-dire de cas de stimuli pour lesquels il y a une forte différence entre score d'estimation du destinataire présumé à l'audio et à l'écrit.

Voici ci-dessous les deux exemples correspondant aux valeurs maximales absolues de $dDest$ sur l'ensemble du corpus. Nous spécifions aussi pour information les scores également attribués à ces stimuli pour la dimension d'hostilité.

Exemple 1 : Le Chequer, stimulus n°13, transcription simplifiée

« Nous vous avons proposé d'organiser une table ronde avec l'ensemble des élus, et du conseil municipal, les conseillers généraux, régionaux, le député, tout le monde (...) »

Estimation destinataire* audio = 4.33/5, destinataire écrit = 1.77/5, $dDest = 2.56$.

Perception hostilité** audio = 1.55/2, hostilité écrit = 0.30/2, $dHost = 1.25$.

*1 = auditoire, 3 = indéterminé, 5 = interlocuteur ;

**0 = aucune hostilité, 1 = légère, 2 = forte

Exemple 2 : Tuailon, stimulus n°11, transcription simplifiée.

« Vous avez fait le choix de ne pas jouer l'hégémonie et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble afin de ne pas effrayer nos partenaires. »

Estimation destinataire* audio = 4.64/5, destinataire écrit = 1.76/5, $dDest = 2.86$.

Perception hostilité** audio = 0.90/2, hostilité écrit = 0.07/2, $dHost = 0.83$.

*1 = auditoire, 3 = indéterminé, 5 = interlocuteur ;

**0 = aucune hostilité, 1 = légère, 2 = forte

Pour ces deux exemples, nous voyons que les scores perceptifs obtenus en condition audio et en condition écrite concernant l'estimation du destinataire diffèrent fortement.

Présentés sous leur forme retranscrite seule (condition écrite), les deux extraits ci-dessus ont été perçus par les participants comme étant adressés à l'auditoire en général, alors qu'en perception audio ils ont été perçus comme adressés à un interlocuteur particulier.

Notons toutefois que si l'extrait de Le Chequer (stimulus n°13) est également perçu comme beaucoup plus hostile en condition audio qu'en condition écrite, cette tendance est moindre pour l'extrait de Tuailon (stimulus n°11).

IV.3.3.2. Rapports entre estimation du destinataire audio, écrite et *dDest*

Tous locuteurs confondus, on observe une corrélation significative pour tous les locuteurs entre scores audio et scores écrits dans l'estimation du destinataire ($\rho = .749$, $p < .001$).

Tableau IV—4 corrélations de Spearman entre scores d'estimation du destinataire en condition audio et écrite, *dDest* et scores audio, *dDest* et scores écrits. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	audio / écrit	dDest / audio	dDest / écrit
VOYNET	$\rho = 0.697 (***)$	$\rho = 0.231 (\text{NS})$	$\rho = -0.469 (**)$
LE CHEQUER	$\rho = 0.85 (***)$	$\rho = -0.014 (\text{NS})$	$\rho = -0.496 (**)$
MOLOSSI	$\rho = 0.88 (***)$	$\rho = 0.459 (*)$	$\rho = 0.07 (\text{NS})$
SEREY	$\rho = 0.719 (***)$	$\rho = 0.674 (**)$	$\rho = 0.047 (\text{NS})$
TUAILLON	$\rho = 0.65 (***)$	$\rho = 0.581 (**)$	$\rho = -0.162 (\text{NS})$
<i>Tous loc.</i>	$\rho = 0.749 (***)$	$\rho = 0.316 (**)$	$\rho = -0.333 (**)$

Tous locuteurs confondus, il y a une corrélation significative entre *dDest* (mesure de différence entre scores du test audio et scores test écrit) et scores audio : plus il y a de différence entre audio et écrit, plus les stimuli sont perçus comme s'adressant à un interlocuteur particulier à l'écrit. Comme on peut le constater dans le tableau récapitulatif des corrélations ci-dessus, Voynet et Le Chequer sont à la fois les deux seuls locuteurs pour lesquels cette corrélation n'est pas significative, et les deux seuls locuteurs pour lesquels on trouve une corrélation entre *dDest* et les scores écrits : pour ces deux locuteurs, plus les stimuli sont perçus comme s'adressant à un interlocuteur particulier en condition audio, moins il y a de différence de perception du destinataire présumé entre test audio et test écrit.

IV.3.4. Corrélations entre hostilité perçue et estimation du destinataire

À l'oral comme à l'écrit, on observe une corrélation positive significative entre l'hostilité perçue et la perception d'un locuteur s'adressant à un interlocuteur particulier ou en général à l'auditoire (oral : $\rho = .596$, $p < .001$; écrit : $\rho = .660$, $p < .001$) : quelle que soit la modalité de présentation des stimuli, plus les stimuli sont perçus comme hostiles, plus les locuteurs sont perçus comme s'adressant à un interlocuteur particulier.

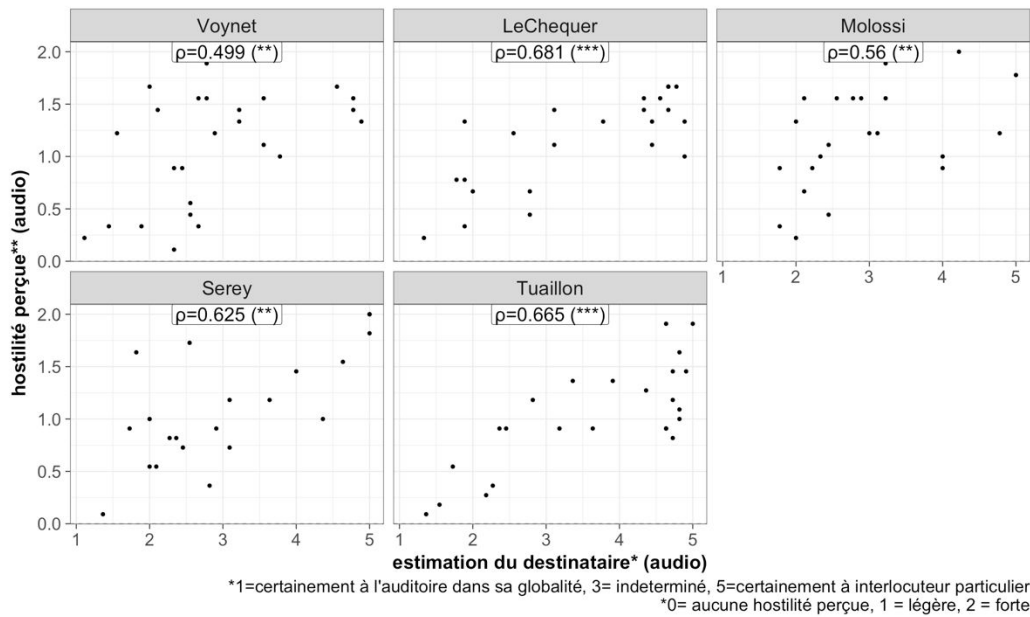


Figure IV-8 Corrélations entre hostilité perçue et estimation du destinataire au sein du test audio. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

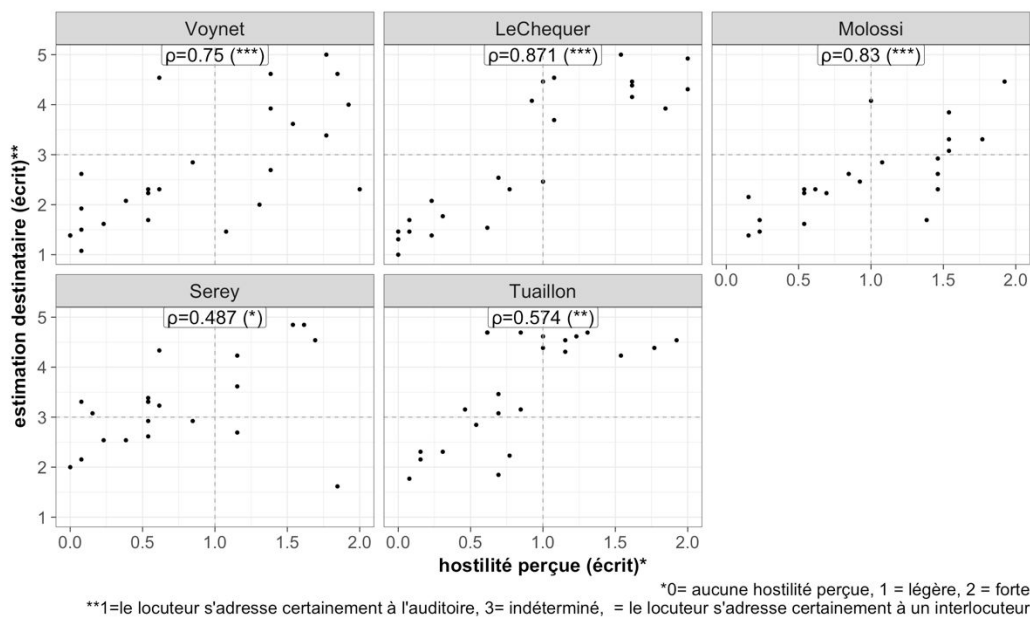


Figure IV-9 Corrélations entre hostilité perçue et estimation du destinataire au sein du test écrit. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

IV.4. Étude n°2 : Profils émotionnels (valence, excitation et contrôle/dominance)

IV.4.1. Protocole méthodologique de l'étude n°2

Durant les entretiens individuels, les 5 locuteurs du corpus ont accepté de réécouter des extraits de leur propre parole et de les évaluer sur des grilles d'évaluation fournies sur support papier. Les stimuli audio (25 par locuteur) leur ont été présentés dans un ordre aléatoire *via* des hauts parleurs situés devant eux. Bien que chaque extrait puisse être réécouté une seconde fois, les locuteurs ont généralement fait leur choix à partir de la première écoute et se sont arrêtés à celui-ci.

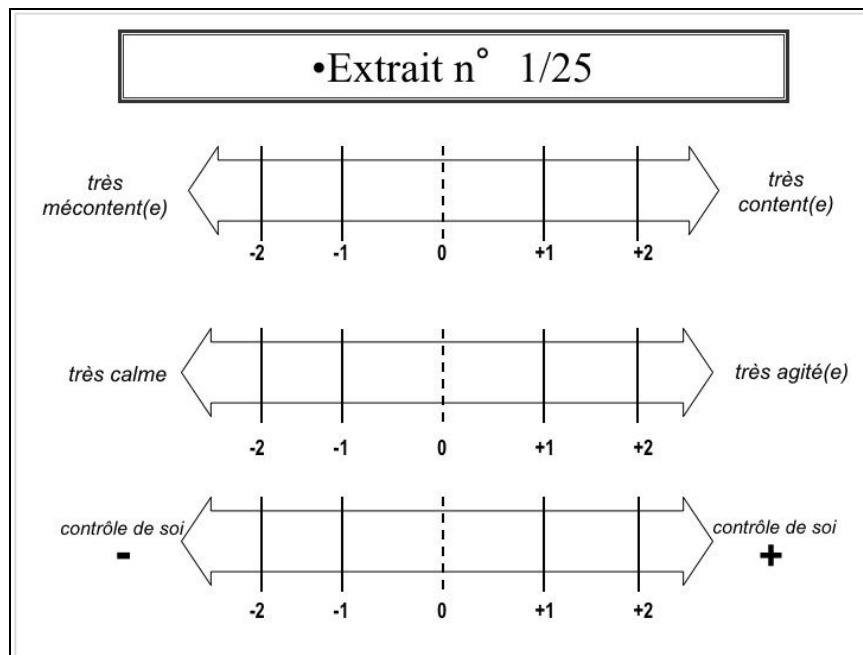


Figure IV-10 : Grille dimensionnelle telle qu'elle a été présentée aux cinq locuteurs du corpus durant la tâche d'auto-évaluation de leurs stimuli.

Les locuteurs ont eu la tâche de décrire leurs propres stimuli à travers 3 échelles de Likert à 5 points représentant les dimensions de valence (très mécontent-e/très content-e), d'excitation (très calme/très agité-e), et de contrôle de soi (minimal/maximal). Pour la dimension de contrôle de soi, il leur a été précisé qu'un score de -2 correspondait à un degré de contrôle de soi minimal, telle une perte totale du contrôle de ses émotions, un score de +2 correspondait à un degré de contrôle de soi maximal, que celui-ci résulte d'un effort particulier de la part du locuteur ou non. Le degré zéro présent dans chaque échelle a été spécifié comme l'indication d'un niveau « neutre », par exemple « ni content ni mécontent », pouvant également indiquer qu'il n'y a « rien à signaler dans cet extrait » concernant telle ou telle dimension.

En parallèle à l'auto-évaluation, 24 participants (16 femmes, 8 hommes, âge moyen : 22 ans) ont évalué les 125 stimuli (25 stimuli par locuteur) à l'aide d'une interface en ligne possédant une grille analogue à celle fournie dans l'auto-évaluation. Les auditeurs, étudiants en première et deuxième année de Licence de Sciences du Langage, ont participé à l'étude sur la base du volontariat. Sur l'interface du site en ligne, les participants ont eu accès à une échelle à 7 points, les résultats obtenus ont ensuite été convertis pour pouvoir être comparés à ceux de l'auto-évaluation. Dans la grille présentée aux participants, la dimension de contrôle de soi a été remplacée par celle de la dominance perçue. Il nous a semblé plus adéquat de questionner les politiciens sur leur degré de contrôle de soi, et de tester en parallèle comment ils étaient perçus en termes de contrôle de la situation (dominant *versus* soumis) dans une situation donnée. Afin d'évaluer le risque potentiel d'*ethos* préalable⁵, il a été notamment demandé aux participants de préciser en fin de test s'ils pensaient avoir reconnu des locuteurs. Tous les auditeurs ont précisé qu'ils n'avaient reconnu aucun locuteur.

⁵ Pour rappel, la perception des personnes politiques est fortement influencée par leur connaissance au préalable par les auditeurs et spectateurs (Charaudeau, 2015). Cet *ethos* préalable pourrait constituer un biais à l'évaluation perceptive des extraits de parole des locuteurs si ceux-ci étaient reconnus par les auditeurs. Or, parmi les participants de nos analyses perceptives, personne n'a reconnu de locuteurs.

IV.4.2. Résultats pour la valence perçue et auto-attribuée

Nous présentons dans cette partie les scores obtenus pour la dimension de valence en auto-évaluation et en perception. Nous présentons également les valeurs moyennes de *dVal*, mesure de différence stimulus par stimulus de l'écart entre scores auto-évalués et scores perceptifs. Pour faciliter l'interprétation de cette variable, nous rappelons que plus les valeurs de *dVal* sont hautes, plus elles vont correspondre à des stimuli pour lesquels les locuteurs se sont auto-évalués comme moins mécontents qu'ils n'ont été perçus. Au contraire, plus les valeurs de *dExc* sont basses plus elles vont renvoyer à des cas de stimuli pour lesquels les locuteurs ont été perçus comme plus contents que ce qu'ils avaient déclaré dans leur auto-évaluation.

IV.4.2.1. Tendances générales

Tous locuteurs confondus, les stimuli ont été en moyenne perçus comme exprimant une valence légèrement négative, avec des scores proches du degré neutre (0) dans l'auto-évaluation et proches du degré « légèrement mécontent » (+1) en perception, dans l'évaluation par les auditeurs naïfs. Le score moyen de valence auto-évaluée est de -0.17 (écart-type : 0.81) contre -0.69 (écart-type : 0.67).

	MOYENNE (ÉCART-TYPE)			SCORE MINIMUM		SCORE MAXIMUM	
	AUTO	PERC	Diff.	AUTO	PERC	AUTO	PERC
VOYNET	0.88 (0.69)	1.08 (0.49)	1.78 (0.81)	0	-1.68	2	0.78
LE CHEQUER	0.88 (0.69)	-0.09 (0.85)	0.56 (0.91)	-1	-1.63	1	0.16
MOLOSSI	0.98 (0.56)	-0.91 (0.81)	-0.01 (0.65)	-2	-1.95	1	0.06
SEREY	0.75 (0.57)	-0.57 (0.6)	0.15 (0.55)	-2	-1.4	0	0.21
TUAILLON	0.86 (0.48)	-0.36 (1.32)	0.08 (1.1)	-2	-1.53	2	0.76
Tous loc.	-0.17 (0.81)	-0.69 (0.57)	0.51 (0.80)	-1.4	-1.63	1.2	0.40

Tableau IV—5 valeurs principales des scores de valence auto-évaluée (AUTO) par les locuteurs du corpus et perçue (PERC) par 24 auditeurs naïfs (moyenne et écart-type, minimum, maximum). Pour rappel : -2 = très mécontent-e, 0 = neutre, 2 = très content-e.

Les scores de valence perçue sont significativement plus bas que ceux que se sont attribués en moyenne les locuteurs durant la tâche d'auto-évaluation ($t[115]=5.4, p<.001, \eta^2=11\%$), ce qui signifie que les stimuli des locuteurs sont perçus comme légèrement plus mécontents que le degré déclaré durant l'auto-évaluation. De plus, nous précisons que si certains locuteurs peuvent aller jusqu'à s'auto-attribuer des scores de 2 (« très content-e ») à certains de leurs stimuli, le score maximum de valence perçue reste inférieur à 1 pour tous les locuteurs.

Tous les locuteurs sont en moyenne perçus comme exprimant un léger mécontentement (valence négative), sans différence interlocuteur ($F[4,111]=1.82, p=0.13; \eta^2=6\%$).

On ne trouve de différence significative entre auto-évaluation et perception « naïfs » que pour Voynet ($t[24]=11, p<.001, \eta^2=71\%$) et Le Chequer ($t[22]=3, p<.01, \eta^2=16\%$) ; comme on peut l'observer dans la figure ci-dessous.

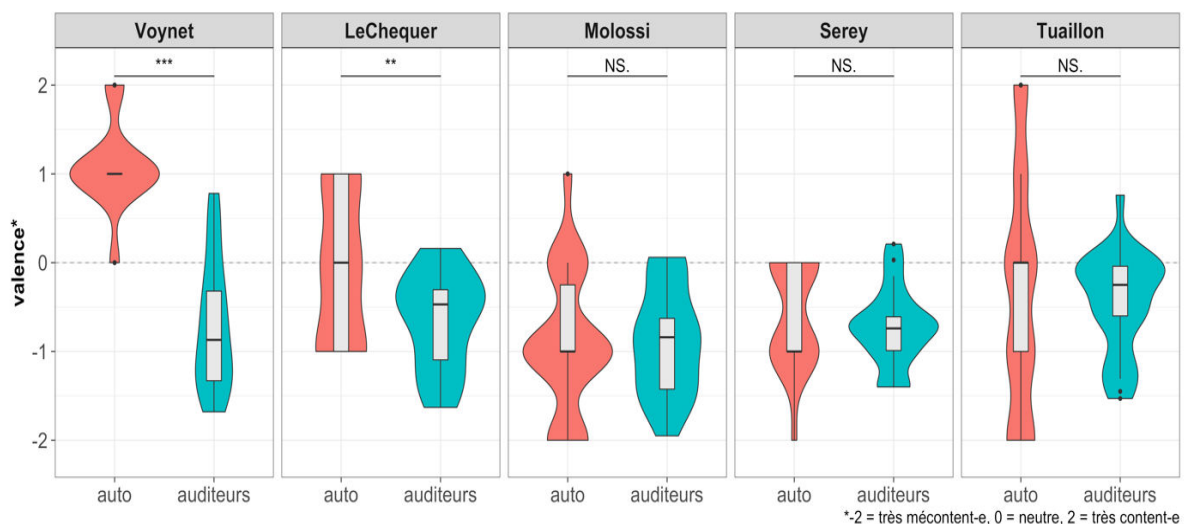


Figure IV-11 Comparaisons entre scores de valence auto-évaluée et de valence perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

La différence significative de scores selon le locuteur dans les résultats de l'auto-évaluation ($F[4,111]=18.22, p<.001; \eta^2=4\%$) semble essentiellement due aux profils particuliers de Voynet et Le Chequer. Voynet attribue à ses stimuli en moyenne l'expression d'une valence plus positive (« plutôt contente ») que tous les autres locuteurs ($p<.001$ pour toutes les comparaisons post-hoc). Elle est la seule locutrice avec un score moyen positif de valence auto-évaluée (1.08), équivalent au niveau « plutôt content-e » de l'échelle, ce niveau représentant 76% de ses évaluations. Le Chequer quant à lui s'attribue en moyenne des scores relatifs à une valence neutre (« ni content-e ni mécontent-e »).

Nous notons également que son taux de stimuli auto-évalués avec une valence positive (+1, « plutôt content » ou +2, « très content ») est de 30%, ce qui est significativement plus élevé que tous les autres locuteurs de l'opposition. Les trois autres locuteurs, pour lesquels on n'observe pas de différence entre scores issus de l'auto-évaluation et du test de perception, s'attribuent des scores de valence négative. Molossi, Serey et Tuaille sont en effet à la fois perçus et évalués comme légèrement mécontents.

Ni Voynet, ni Le Chequer ne conservent les particularités de leurs profils de valence auto-évalués dans les scores obtenus au sein de l'évaluation par les auditeurs naïfs, où tous les locuteurs se voient attribuer sans distinction des scores moyens, correspondant à une valence légèrement négative. Dans l'évaluation « naïfs », Tuaille est le locuteur dont le score moyen de valence est le moins négatif (-0.44), bien que celui-ci ne soit significativement supérieur qu'à celui de Molossi (-0.9) [$t(42.2)=2,65$, $p<.01$].

IV.4.2.2. Rapports entre scores de valence auto-évaluée, perçue et *dVal*

En plus de l'analyse des différences entre scores moyens issus de l'auto-évaluation et scores perceptifs (24 auditeurs), nous avons décidé d'analyser les rapports entre valence auto-évaluée et valence perçue à travers des corrélations de Spearman entre scores issus de l'auto-évaluation, de l'évaluation par les auditeurs naïfs mais également avec *dVal*, la mesure de différence entre les deux scores pour chaque stimulus. En effet, selon nous, l'analyse des corrélations entre scores issus des différents tests (auto-évaluation, perception) et mesures de l'écart entre scores auto-évalués et perçus (ici *dVal* pour la dimension de valence) peut nous apporter un éclairage sur les raisons de l'écart entre auto-évaluation et perception⁶ : les différences les plus fortes sont-elles plus corrélées aux tendances en perception ou en auto-évaluation ?

Tous locuteurs confondus, les scores de valence auto-évaluée et de valence perçue sont significativement corrélés ($\rho = .303$, $p<.01$) : plus les stimuli des locuteurs sont auto-évalués comme exprimant une valence négative, plus ils sont perçus comme tels.

⁶ Dans la même logique d'analyse, nous avons déjà étudié les rapports entre hostilité perçue à l'oral, à l'écrit, et *dHost* mesure de différence entre scores audio et écrits.

Toutefois, cette tendance n'est pas significative pour Voynet ($\rho = .231$) ni Le Chequer ($\rho = .276$), qui sont les deux seuls locuteurs pour lesquels nous ne trouvons pas de différence significative entre scores auto-évalués et perçus.

Tableau IV—6 corrélations de Spearman entre scores de valence auto-évaluée et perçue, *dVal* et scores auto-évalués, *dVal* et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de *p* des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	Auto-éval. / perception	dVal / perception	dVal / auto-éval.
VOYNET	$\rho = 0.231$ (NS)	$\rho = -0.792$ (***)	$\rho = 0.336$ (NS)
LE CHEQUER	$\rho = 0.276$ (NS)	$\rho = -0.282$ (NS)	$\rho = 0.807$ (***)
MOLOSSI	$\rho = 0.601$ (**)	$\rho = -0.055$ (NS)	$\rho = 0.724$ (***)
SEREY	$\rho = 0.428$ (*)	$\rho = -0.249$ (NS)	$\rho = 0.736$ (***)
TUAILLON	$\rho = 0.575$ (**)	$\rho = 0.12$ (NS)	$\rho = 0.848$ (***)
<i>Tous loc.</i>	$\rho = 0.303$ (**)	$\rho = -0.206$ (*)	$\rho = 0.850$ (***)

Concernant *dVal*, tous locuteurs confondus, nous voyons que cette variable est beaucoup plus fortement corrélée aux scores de valence auto-évaluée ($\rho = 0.850$, $p < .001$) qu'aux scores de valence perçue ($\rho = -0.206$, $p < .05$). Ces corrélations signifient que les stimuli pour lesquels les locuteurs ont renseigné le plus haut niveau de valence dans leur auto-évaluation sont ceux pour lesquels on observe l'écart le plus important avec les scores obtenus en perception, et donc les valeurs de *dVal* les plus grandes. En d'autres termes, les plus fortes divergences de scores entre perception et auto-évaluation correspondent aux cas de stimuli pour lesquels les locuteurs s'auto-évaluent comme contents (valence positive).

Les scores de valence perçue ne sont significativement corrélés à *dVal* que pour une locutrice sur cinq, Voynet ($\rho = -0.792$, $p < .001$) : plus celle-ci elle est perçue comme mécontente, plus il y a d'écart entre auto-évaluation et perception dans ses scores ; cette tendance étant cohérente avec le fait que Voynet est la seule locutrice à avoir en moyenne auto-évalué ses stimuli avec des valeurs de valence positive.

IV.4.3. Résultats pour l'excitation perçue et auto-évaluée

Nous présentons dans cette partie les scores obtenus pour la dimension d'excitation en auto-évaluation et en perception. Nous présentons également les valeurs moyennes de $dExc$, mesure de différence stimulus par stimulus de l'écart entre scores auto-évalués et scores perceptifs. Pour faciliter l'interprétation de cette variable, nous rappelons que plus les valeurs de $dExc$ sont hautes, plus elles vont correspondre à des stimuli pour lesquels les locuteurs se sont auto-évalués comme plus agités qu'ils n'ont été perçus. Au contraire, plus les valeurs de $dExc$ sont basses plus elles vont renvoyer à des cas de stimuli pour lesquels les locuteurs ont été perçus comme plus agités que ce qu'ils avaient déclaré dans leur auto-évaluation.

IV.4.3.1. Excitation : tendances générales

Les stimuli de tous locuteurs sont globalement perçus et auto-évalués comme exprimant un degré d'excitation neutre (« ni calme ni agité ») avec une moyenne de 0.55 (écart-type : 0.85) dans l'auto-évaluation et de 0.58 (0.47) dans la perception par les auditeurs, les locuteurs étant perçus comme légèrement moins agités que ce qu'ils déclarent eux-mêmes.

	MOYENNE (ÉCART-TYPE)			SCORE MINIMUM		SCORE MAXIMUM	
	AUTO	PERC	Diff.	AUTO	PERC	AUTO	PERC
VOYNET	0.92 (0.49)	0.59 (0.49)	0.32 (1.07)	0	-0.3	0	1.63
LE CHEQUER	0.26 (0.81)	0.54 (0.65)	-0.17 (0.74)	-1	-0.82	-1	1.42
MOLOSSI	-0.32 (0.84)	0.68 (0.42)	-0.35 (1.3)	-2	-0.57	-2	1.32
SEREY	1.14 (0.65)	0.49 (0.43)	0.29 (1.25)	0	-0.17	0	1.28
TUAILLON	0.76 (1.48)	0.58 (0.35)	0.62 (1.63)	-2	-0.05	-2	1.36
Tous loc.	0.55 (0.85)	0.58 (0.47)	0.14 (1.20)	-1.00	-0.38	-1.00	1.40

Tableau IV—7 valeurs principales des scores d'excitation auto-évaluée par les locuteurs du corpus (AUTO) et perçue (PERC) par 24 auditeurs naïfs (moyenne et écart-type, minimum, maximum). Pour rappel : -2 = très calme, 0 = neutre, 2 = très agité-e.

Tous locuteurs confondus, on ne trouve pas de différence significative entre scores auto-évalués et scores perçus pour cette dimension ($t[115]=1.3$, $p=0.2$, $\eta^2=1\%$).

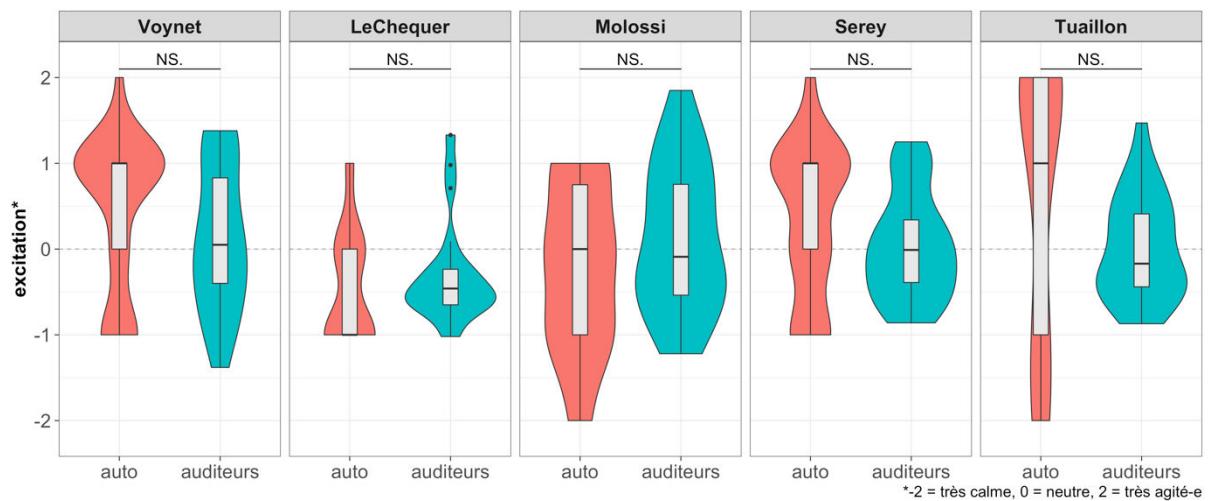


Figure IV-12 : Comparaisons entre scores d'excitation auto-évaluée et d'excitation perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Au sein de l'auto-évaluation, on observe une influence significative du locuteur sur les scores ($F[4,111]=4.55$, $p < .001$, $\eta^2=14\%$). Les tests post-hoc indiquent que Le Chequer s'attribue en moyenne des scores d'excitation plus bas que Tuailon ($p < .001$) et Voynet ($p < .05$), et que les scores moyens d'excitation auto-évaluée de Tuailon sont supérieurs à ceux de Molossi ($p < .05$). Les profils auto-évalués d'excitation de Tuailon et Le Chequer sont en effet très différents : Tuailon est le seul locuteur avec une majorité de stimuli auto-évalués comme exprimant un état « très agité » ; 48% contre seulement 0 à 5% pour les autres locuteurs. Quant à Le Chequer, 57% de ses stimuli auto-évalués comme « plutôt calme » (score -1), 35% comme d'excitation « neutre » (score 0).

Comme pour la valence, on n'observe aucune différence inter-locuteurs dans les scores d'excitation attribués par les auditeurs naïfs ($F[4,111]=1.43$, $p=0.23$; $\eta^2=5\%$).

IV.4.3.2. Corrélations entre excitation auto-évaluée, excitation perçue et $dExc$

Comme on peut le constater dans le tableau ci-dessous, on ne trouve aucune corrélation significative entre scores d'excitation auto-évaluée et scores d'excitation perçue. Pour aucun locuteur, les stimuli auto-évalués comme les plus agités ne sont perçus comme tels. Cette tendance est très différente de celle observée pour la valence.

Tableau IV—8 corrélations de Spearman entre scores d'excitation auto-évaluée et perçue, *dExc* (mesures de différence entre auto-évaluation et perception) et scores auto-évalués, *dExc* et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	Auto-éval. / perception	<i>dExc</i> / perception	<i>dExc</i> / auto-éval.
VOYNET	$\rho = 0.261$ (NS)	$\rho = -0.571$ (**)	$\rho = 0.593$ (**)
LE CHEQUER	$\rho = 0.091$ (NS)	$\rho = -0.597$ (**)	$\rho = 0.673$ (***)
MOLOSSI	$\rho = -0.007$ (NS)	$\rho = -0.6$ (**)	$\rho = 0.793$ (***)
SEREY	$\rho = -0.365$ (NS)	$\rho = -0.718$ (***)	$\rho = 0.879$ (***)
TUAILLON	$\rho = 0.203$ (NS)	$\rho = -0.281$ (NS)	$\rho = 0.845$ (***)
<i>Tous loc.</i>	$\rho = 0.303$ (**)	$\rho = -0.39$ (***)	$\rho = 0.842$ (***)

Concernant *dExc*, tous locuteurs confondus, nous voyons que cette variable est plus fortement corrélée aux scores d'excitation auto-évalués ($\rho = 0.842$, $p < .001$) qu'aux scores d'excitation perçue ($\rho = -0.390$, $p < .001$). Ces corrélations signifient que les stimuli pour lesquels les locuteurs ont renseigné le plus haut niveau d'excitation dans leur auto-évaluation sont ceux pour lesquels on observe l'écart le plus important avec les scores obtenus en perception, et donc les valeurs de *dExc* les plus grandes. Dans une moindre mesure, pour tous les locuteurs sauf Tuailleon, plus les locuteurs sont perçus comme calmes (excitation minimale), plus il y a également d'écart entre scores auto-évalués et scores perçus. L'absence de corrélation significative pour Tuailleon signifie que pour ce locuteur, les stimuli pour lesquels il est perçu comme le plus calme ne sont pas nécessairement ceux pour lesquels on trouve l'écart le plus important entre excitation auto-évaluée et excitation perçue.

IV.4.4. Résultats pour le contrôle de soi auto-évalué et la dominance perçue

Dans cette section, nous présentons les résultats obtenus pour la troisième dimension testée. Pour rappel, nous avons fait le choix de tester deux dimensions différentes mais analogues entre la tâche d'auto-évaluation effectuée par les locuteurs et le test perceptif effectué par les auditeurs naïfs (voir section IV.4.1 page 68 pour une justification de ce choix). Les locuteurs ont en effet auto-évalué leurs stimuli à travers une échelle de contrôle de soi, et les auditeurs ont évalué les locuteurs à travers une échelle de dominance perçue. Les résultats sont analysés ici en suivant la même méthodologie que pour la valence et l'excitation : présentation et comparaison des tendances principales dans les résultats de l'auto-évaluation et de la perception, puis analyse des rapports entre scores auto-évalués, scores perçus et mesure de différence ; ici $dDom$ (scores de contrôle auto-évalué – scores de dominance perçue).

IV.4.4.1. Contrôle de soi et dominance perçue : tendances principales

	MOYENNE (ÉCART-TYPE)			SCORE MINIMUM		SCORE MAXIMUM	
	AUTO	PERC	Diff.	AUTO	PERC	AUTO	PERC
VOYNET	0.92 (0.49)	0.59 (0.49)	0.33 (0.65)	0	-0.3	2	1.63
LE CHEQUER	0.26 (0.81)	0.54 (0.65)	-0.28 (0.72)	-1	-0.82	1	1.42
MOLOSSI	-0.32 (0.84)	0.68 (0.42)	-1 (0.99)	-2	-0.57	1	1.32
SEREY	1.14 (0.65)	0.49 (0.43)	0.65 (0.43)	0	-0.17	2	1.28
TUAILLON	0.76 (1.48)	0.58 (0.35)	0.18 (1.68)	-2	-0.05	2	1.36
<i>Tous loc.</i>	0.55 (0.85)	0.58 (0.47)	-0.02 (0.90)	-1	-0.382	1.6	1.402

Tableau IV—9 valeurs principales des scores de contrôle de soi auto-évalué par les locuteurs du corpus et dominance perçue par 24 auditeurs naïfs : moyennes et écart-types, et scores minimum et maximum. Pour rappel : -2 = contrôle/dominance min, 0 = neutre, 2 = contrôle/dominance max.

Tous locuteurs confondus, ceux-ci sont perçus comme légèrement dominants et se déclarent légèrement dans le contrôle. On observe des valeurs moyennes proches les unes des autres dans

les deux conditions : le score moyen de contrôle de soi auto-évalué est de 0.55 (écart-type : 0.85) contre 0.58 pour la dominance perçue (écart-type : 0.47).

Les valeurs moyennes de *dDom*, mesure de différence entre scores auto-évalués et perçus sont proches de zéro (moyenne : -0.02, écart-type : 0.90). Nous précisons ici que de la même façon que pour *dVal* et *dExc*, les valeurs de *dDom* s'interprètent comme telles : une valeur proche de zéro équivaut à une différence minimale entre contrôle de soi et dominance perçue, une valeur positive équivaut à des cas de stimuli pour lesquels les locuteurs se déclarent plus dans le contrôle qu'ils ne sont perçus comme dominants.

Le score moyen de contrôle auto-évalué le plus élevé est relatif à Serey (1.14), suivi par Voynet (0.92) ; ces deux locuteurs étant les seuls à n'avoir utilisé que des scores de contrôle de soi positifs pour évaluer leurs stimuli. Les stimuli auto-évalués avec un niveau de léger contrôle de soi représentent 76% des stimuli de Voynet et 48% des stimuli de Serey.

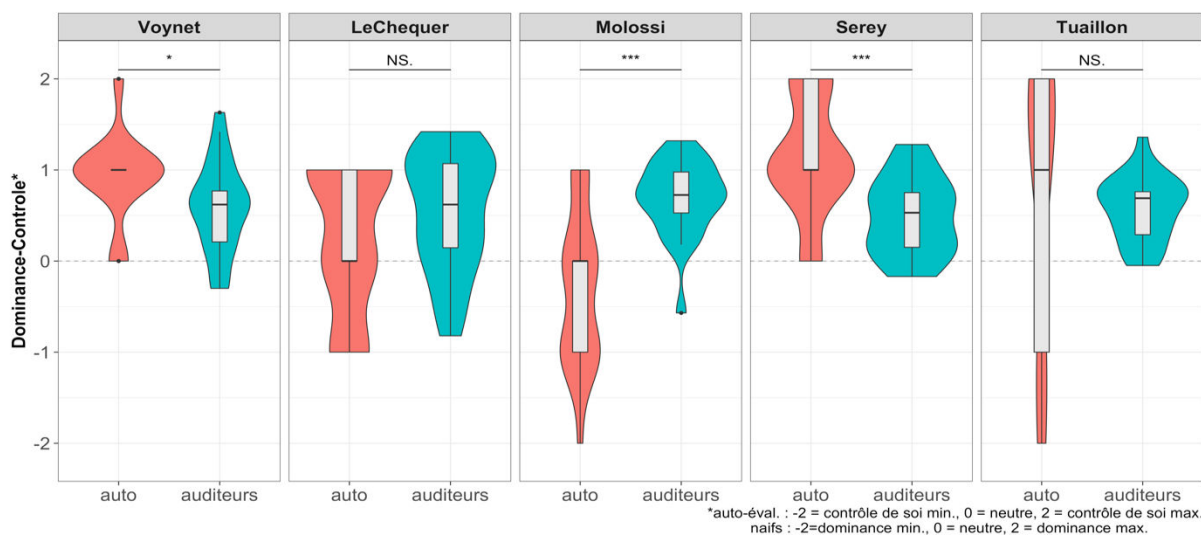


Figure IV-13 Comparaisons entre scores de contrôle de soi auto-évalué et de dominance perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Si, tous locuteurs confondus, on n'observe pas de différence significative entre contrôle de soi et dominance perçue tous locuteurs confondus ($t[115] = -0.2$, $p = 0.9$, $\eta^2 = 0\%$), les scores diffèrent significativement pour trois locuteurs sur cinq : Serey ($t[20] = 6.9$, $p < .001$, $\eta^2 = 30\%$), Molossi ($t[21] = -4.7$, $p < .001$, $\eta^2 = 45\%$) et Voynet ($t[24] = 2.6$, $p < .01$, $\eta^2 = 11\%$).

Il n'y a donc pas de différence significative entre contrôle de soi et dominance perçue pour Le Chequer et Tuailon, bien que pour ce dernier la variance de ses scores auto-évalués soit significativement plus élevée que ses scores perceptifs ($F[24] = 17.5$, $p < .001$, $\eta^2 = 27\%$), comme on peut l'observer sur la figure ci-dessus.

On trouve une influence significative du locuteur au sein des scores de contrôle auto-évalué ($F[4,111]=8.77, p<.001; \eta^2=24\%$) : Molossi s'attribue en moyenne des scores de contrôle de soi plus négatifs que ceux de Voynet et Tuailon ($p<.001$ pour les deux comparaisons). En effet, 41% des stimuli de ce locuteur ont été auto-évalués avec un niveau -1 de contrôle de soi, contre 1 à 20% pour les autres locuteurs.

Au sein de l'évaluation par les auditeurs naïfs, tous les locuteurs ont été uniformément perçus comme légèrement dominants, sans différence significative entre interlocuteurs ($F[4,111]=0.48, p=0.75; \eta^2=2\%$).

IV.4.4.2. Rapports entre contrôle de soi auto-évalué, dominance perçue et *dDom*

Tous locuteurs confondus, la corrélation entre scores de contrôle de soi auto-évalué et de dominance perçue n'est pas significative ($\rho = 0.027$), toutefois cette corrélation très faible n'est pas tant à mettre sur le compte d'une absence totale de corrélation mais plutôt de corrélations de directions différentes selon les locuteurs. On trouve une corrélation significative entre auto-évaluation du contrôle de soi et perception de la dominance pour trois locuteurs : Serey ($\rho = 0.790, p<.001$) et Le Chequer ($\rho = 0.494, p<.01$), pour qui celle-ci est positive, et Tuailon ($\rho = -0.444, p<.05$) pour qui est elle négative.

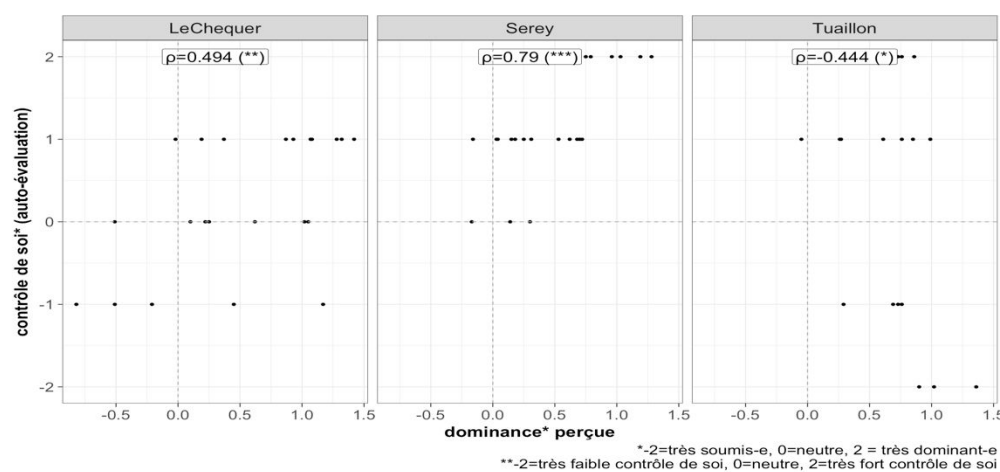


Figure IV-14 Corrélations entre contrôle de soi auto-évalué et dominance perçue (auditeurs naïfs) pour Le Chequer, Serey et Tuailon. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

Comme l'illustre la figure ci-dessus, pour Serey et Le Chequer, plus ces deux locuteurs se déclarent comme exerçant un fort contrôle sur leurs émotions, plus ils sont perçus comme dominants par les auditeurs naïfs. Pour Tuailon ce sont au contraire les stimuli pour lesquels ce locuteur se déclare exercer un contrôle de soi minimal qui sont perçus comme plus dominants.

Tableau IV—10 corrélations de Spearman entre scores de contrôle de soi auto-évalué et de dominance perçue, *dDom* (mesures de différence entre auto-évaluation et perception) et scores auto-évalués, *dDom* et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de *p* des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $n.s = p > .05$, non significatif.

	Auto-éval. / perception	dDom / perception	dDom / auto-éval.
VOYNET	$\rho = 0.098$ (NS)	$\rho = 0.627$ (***)	$\rho = -0.682$ (***)
LE CHEQUER	$\rho = 0.494$ (**)	$\rho = 0.648$ (***)	$\rho = -0.275$ (NS)
MOLOSSI	$\rho = -0.162$ (NS)	$\rho = 0.909$ (***)	$\rho = -0.476$ (*)
SEREY	$\rho = 0.79$ (***)	$\rho = 0.711$ (***)	$\rho = 0.162$ (NS)
TUAILLON	$\rho = -0.444$ (*)	$\rho = 0.943$ (***)	$\rho = -0.674$ (***)
<i>Tous loc.</i>	$\rho = 0.027$ (NS)	$\rho = 0.874$ (***)	$\rho = -0.413$ (***)

Concernant *dDom*, tous locuteurs confondus, nous voyons que cette variable est plus fortement corrélée aux scores de dominance ($\rho = 0.874$, $p < .001$) qu'aux scores de contrôle de soi auto-évalué ($\rho = -0.413$, $p < .001$). Ces corrélations signifient que les stimuli pour lesquels les locuteurs sont perçus comme les plus dominants sont ceux pour lesquels on observe l'écart le plus important avec les scores de contrôle de soi auto-évalué, et donc les valeurs de *dDom* les plus grandes. Dans une moindre mesure, pour trois locuteurs sur cinq (Voynet, Molossi et Tuaille), plus les locuteurs déclarent un fort contrôle de soi plus il y a d'écart entre scores auto-évalués et scores perçus.

IV.4.5. Inter-corrélations entre dimensions perçues (auditeurs « naïfs »)

Nous présentons ici les rapports entre scores perceptifs des trois dimensions de valence, excitation et dominance (auditeurs « naïfs »).

Tableau IV—11 corrélations de Spearman entre scores de valence, excitation et dominance perçue. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, n.s = $p > .05$, non significatif.

	valence / excitation	valence / dominance	dom. / excitation
VOYNET	$\rho = -0.609 (**)$	$\rho = -0.314 (NS)$	$\rho = 0.411 (*)$
LE CHEQUER	$\rho = -0.345 (NS)$	$\rho = -0.652 (***)$	$\rho = 0.408 (*)$
MOLOSSI	$\rho = -0.734 (***)$	$\rho = -0.85 (***)$	$\rho = 0.625 (**)$
SEREY	$\rho = -0.629 (**)$	$\rho = -0.732 (***)$	$\rho = 0.531 (**)$
TUAILLON	$\rho = -0.616 (**)$	$\rho = -0.454 (*)$	$\rho = 0.538 (**)$
<i>Tous loc.</i>	$\rho = -0.579 (***)$	$\rho = -0.578 (***)$	$\rho = 0.464 (***)$

IV.4.5.1. Valence perçue et excitation perçue

Tous locuteurs confondus, la valence perçue est significativement corrélée à l'excitation perçue ($\rho = -0.579$, $p < .001$) : plus les stimuli sont perçus comme exprimant un fort mécontentement, plus ils sont également perçus comme exprimant une forte excitation. Le Chequer est le seul locuteur pour lequel la corrélation entre ces deux dimensions n'est pas significative ($\rho = -0.345$) alors que celui-ci n'est toutefois pas perçu comme plus calme ni moins hostile que les autres. Cela signifie qu'il est perçu comme tout autant mécontent que les autres locuteurs sans pour autant que cela soit significativement corrélé à l'excitation perçue.

IV.4.5.2. Dominance perçue et valence perçue

Tous locuteurs confondus, la dominance perçue est significativement corrélée à la valence perçue ($\rho = -.578$, $p < .001$). Pour les quatre membres de l'opposition, les stimuli perçus comme exprimant le plus fort mécontentement sont aussi ceux qui sont perçus comme les plus dominants avec des coefficients de corrélations pouvant monter jusqu'à $\rho = -.851$ pour Molossi. Cette corrélation n'est par contre pas significative pour Voynet ($\rho = -.314$).

IV.4.5.3. Dominance perçue et excitation perçue

Tous locuteurs confondus, les scores d'excitation et de dominance perçue sont significativement corrélés ($\rho = .464, p < .001$). Les stimuli pour lesquels les locuteurs sont perçus comme les plus agités sont perçus comme dominants.

Nous présentons la figure ci-dessous représentant en nuages de points les corrélations par locuteurs pour permettre de visualiser les tendances suivantes : si l'on trouve pour l'ensemble des locuteurs de nombreux cas de stimuli perçus comme exprimant du calme et de la dominance, on ne trouve au contraire aucun stimulus perçus comme exprimant à la fois une forte excitation et une faible dominance.

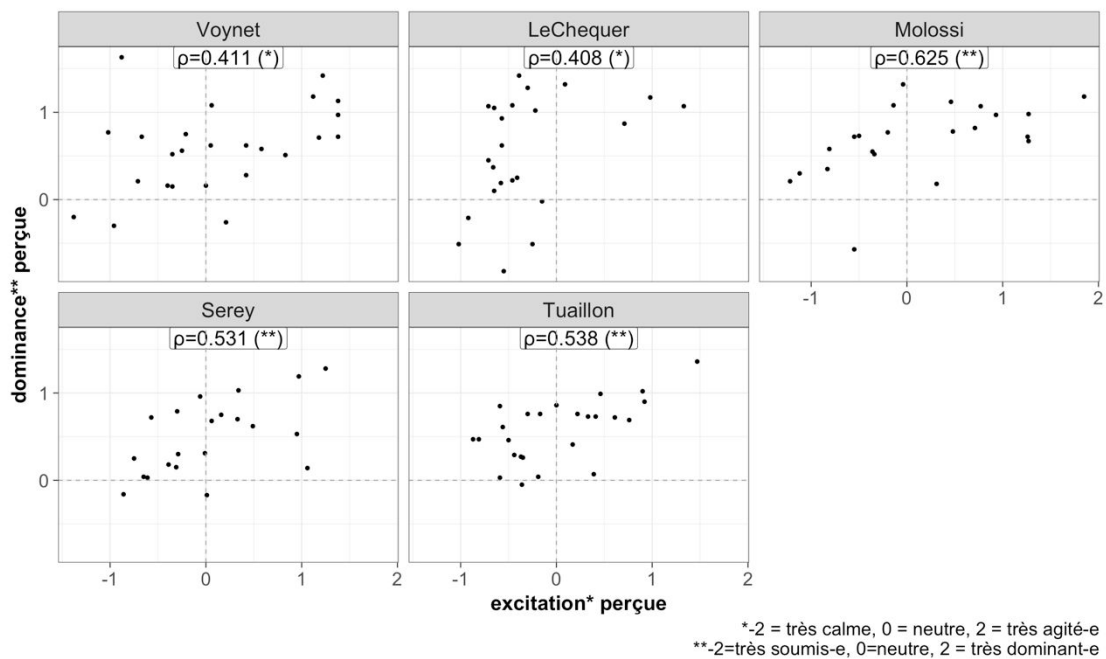


Figure IV-15 Corrélations entre excitation perçue et dominance perçue (auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

IV.4.6. Corrélations entre dimensions auto-évaluées

Si les dimensions perçues sont globalement inter-corrélées, il n'en est pas de même lorsque l'on considère les rapports entre dimensions auto-évaluées.

Tableau IV—12 corrélations de Spearman entre scores de valence, excitation et dominance perçue. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : p<.001 = ***, p<.01 = **, p<.05 = *, n.s = p>.05, non significatif.

	valence / excitation	valence / contrôle	contrôle / excit.
VOYNET	$\rho = -0.43$ (*)	$\rho = -0.118$ (NS)	$\rho = 0.308$ (NS)
LE CHEQUER	$\rho = -0.223$ (NS)	$\rho = 0.119$ (NS)	$\rho = -0.362$ (NS)
MOLOSSI	$\rho = 0.054$ (NS)	$\rho = 0.387$ (NS)	$\rho = -0.128$ (NS)
SEREY	$\rho = -0.039$ (NS)	$\rho = -0.414$ (NS)	$\rho = -0.496$ (*)
TUAILLON	$\rho = -0.556$ (**)	$\rho = 0.296$ (NS)	$\rho = -0.168$ (NS)
Tous loc.	$\rho = -0.178$ (*)	$\rho = 0.209$ (*)	$\rho = 0.019$ (NS)

Comme le tableau ci-dessus l'indique, on calcule seulement trois corrélations significatives entre scores auto-évalués des dimensions de valence, excitation et contrôle de soi :

- Deux corrélations négatives entre valence et excitation auto-évaluées pour Voynet ($\rho = -.450$, $p < .05$) et Tuailleon ($\rho = -.615$, $p < .01$), seuls locuteurs pour lesquels les stimuli auto-évalués comme étant les plus agités sont également ceux pour lesquels ils déclarent être les plus mécontents.
- Une corrélation négative entre excitation et contrôle auto-évalué uniquement pour Serey ($-.496$, $p < .05$). Les stimuli pour lesquels ce locuteur s'attribue les plus hauts scores d'excitation sont ceux pour lesquels les scores de contrôle de soi sont les plus bas, ce qui peut alors s'interpréter comme une perte de contrôle renseignée par celui-ci.

IV.4.7. Corrélations entre mesures de différences (*dVal*, *dExc*, *dDom*)

Enfin, nous avons calculé des corrélations entre les différentes mesures de différences issues de notre seconde campagne d'évaluation perceptive. Les seules variables de différence inter-corrélées significativement sont *dVal* et *dExc*. Pour trois locuteurs sur cinq on observe une corrélation significative négative entre *dExc* et *dVal* (Voynet, $\rho = -.404$, $p < .05$; Le Chequer $\rho = -.513$, $p < .05$; et Tuailleon, $\rho = -.665$, $p < .001$). Plus ces locuteurs se déclarent plus agités qu'ils ne sont perçus, moins il y a de différence entre valence auto-évaluée et perçue.

IV.5. Analyse des rapports entre hostilité perçue et profils dimensionnels

Dans cette partie nous étudions les rapports entre perception de l'hostilité et les dimensions des profils émotionnels (valence, excitation et contrôle de soi / dominance) à travers les corrélations entre scores d'hostilité perçue et toutes les autres variables perceptives : scores dimensionnels perçus, auto-évalués, et mesures de différence entre auto-évaluation et perception pour chaque dimension (*dVal*, *dExc*, *dDom*).

IV.5.1. Rapports entre hostilité et valence

Tous locuteurs confondus, la valence perçue (auditeurs naïfs) est significativement corrélée avec l'hostilité perçue ($\rho = -.709$, $p < .001$) : plus ceux-ci sont perçus comme mécontents, plus ils sont également perçus comme hostiles.

Tableau IV—13 corrélations de Spearman entre hostilité perçue, valence perçue, valence auto-évaluée et *dVal*, mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent le degré de significativité relatif à la valeur de *p* des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	hostilité et valence perçue	hostilité et valence auto-évaluée	hostilité et <i>dVal</i>
VOYNET	$\rho = -0.661 (***)$	$\rho = -0.35 (\text{NS})$	$\rho = 0.505 (**)$
LE CHEQUER	$\rho = -0.607 (**)$	$\rho = -0.016 (\text{NS})$	$\rho = 0.451 (*)$
MOLOSSI	$\rho = -0.839 (***)$	$\rho = -0.64 (**)$	$\rho = -0.063 (\text{NS})$
SEREY	$\rho = -0.827 (***)$	$\rho = -0.481 (*)$	$\rho = 0.064 (\text{NS})$
TUAILLON	$\rho = -0.597 (**)$	$\rho = -0.286 (\text{NS})$	$\rho = 0.054 (\text{NS})$
<i>Tous loc.</i>	$\rho = -0.709 (***)$	$\rho = -0.242 (**)$	$\rho = 0.103 (\text{NS})$

Les corrélations entre hostilité perçue et valence auto-évaluée ne sont significatives que pour deux locuteurs sur cinq : Molossi ($\rho = -.640$, $p < .001$) et Serey ($\rho = .481$, $p < .05$), qui sont les deux locuteurs pour lesquels on trouve le moins de différences entre scores auto-évalués et scores perceptifs, mais également le moins de stimuli auto-évalués avec une valence supérieure à zéro.

Quant à *dVal*, celle-ci n'est corrélée à l'hostilité perçue que pour Voynet ($\rho = -.505$, $p < .01$) et Le Chequer ($\rho = -.451$, $p < .05$). Pour ces deux locuteurs, qui étaient les seuls pour lesquels on trouve une différence significative entre scores moyens de valence auto-évaluée et perçue, les stimuli perçus comme les plus hostiles sont ceux pour lesquels ils se déclarent moins mécontents qu'ils ne le sont perçus.

IV.5.2. Rapports entre hostilité et excitation

Tous locuteurs confondus, l'hostilité perçue est corrélée à l'excitation perçue ($\rho = .558$, $p < .001$) et non significativement corrélée à l'excitation auto-évaluée ($\rho = .082$).

Tableau IV—14 corrélations de Spearman entre hostilité perçue, excitation perçue, excitation auto-évaluée et *dExc* mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent la valeur de p des corrélations calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	hostilité et excitation perçue	hostilité et excitation auto-évaluée	hostilité et <i>dExc</i>
VOYNET	$\rho = 0.521 (**)$	$\rho = 0.557 (**)$	$\rho = 0.075 (\text{NS})$
LE CHEQUER	$\rho = 0.302 (\text{NS})$	$\rho = -0.221 (\text{NS})$	$\rho = -0.398 (\text{NS})$
MOLOSSI	$\rho = 0.758 (***)$	$\rho = 0.189 (\text{NS})$	$\rho = -0.281 (\text{NS})$
SEREY	$\rho = 0.642 (**)$	$\rho = -0.384 (\text{NS})$	$\rho = -0.541 (**)$
TUAILLON	$\rho = 0.704 (***)$	$\rho = 0.153 (\text{NS})$	$\rho = -0.224 (\text{NS})$
<i>Tous loc.</i>	$\rho = 0.562 (***)$	$\rho = 0.082 (\text{NS})$	$\rho = -0.233 (**)$

Pour tous les locuteurs sauf Le Chequer, plus les stimuli sont perçus comme hostiles plus ils sont également perçus comme exprimant une forte excitation. En effet, pour ce locuteur, ses stimuli perçus comme les plus hostiles ne sont pas essentiellement ceux qui sont perçus comme ayant le plus haut niveau d'excitation perçue ; rappelons également qu'il était le seul locuteur pour lequel la valence perçue n'était pas non plus corrélée à l'excitation perçue.

On ne trouve de corrélation significative entre excitation auto-évaluée et hostilité perçue que pour une locutrice, Voynet ($\rho = .646$, $p < .001$). Pour cette locutrice, les stimuli perçus comme exprimant le plus d'hostilité sont également ceux pour lesquels Voynet s'auto-attribue les plus hauts scores d'excitation.

Enfin, concernant les corrélations entre hostilité perçue et *dExc*, variable regroupant les mesures d'écarts entre excitation auto-évaluée et perçue, elles ne sont significatives que pour un locuteur, Serey ($\rho = -.541$, $p < .01$), pour qui les stimuli perçus comme exprimant le plus d'hostilité sont ceux pour lesquels il s'est déclaré plus calme qu'il n'a été perçu.

IV.5.3. Corrélations entre hostilité, contrôle de soi et dominance

Tous locuteurs confondus, on trouve une corrélation positive significative entre hostilité et dominance perçue ($\rho = .573$, $p < .001$), indiquant que plus les stimuli des locuteurs sont perçus comme hostiles, plus ils sont également perçus comme dominants. L'hostilité est corrélée à la dominance perçue pour tous les locuteurs sauf Voynet ($\rho = .309$, $p = 0.1$).

Tableau IV—15 corrélations de Spearman entre hostilité perçue, valence perçue, valence auto-évaluée et dVal, mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

	hostilité et dominance perçue	hostilité et contrôle auto-évalué	hostilité et dDom
VOYNET	$\rho = 0.319$ (NS)	$\rho = 0.457$ (*)	$\rho = 0.073$ (NS)
LE CHEQUER	$\rho = 0.704$ (***)	$\rho = 0.616$ (**)	$\rho = 0.046$ (NS)
MOLOSSI	$\rho = 0.717$ (***)	$\rho = -0.132$ (NS)	$\rho = -0.293$ (NS)
SEREY	$\rho = 0.75$ (***)	$\rho = 0.625$ (**)	$\rho = 0.152$ (NS)
TUAILLON	$\rho = 0.388$ (*)	$\rho = -0.506$ (**)	$\rho = -0.499$ (**)
<i>Tous loc.</i>	$\rho = 0.573$ (***)	$\rho = 0.047$ (NS)	$\rho = -0.178$ (*)

La corrélation entre hostilité perçue et contrôle auto-évalué est significative pour tous les locuteurs sauf Molossi. Pour trois locuteurs sur cinq, plus les locuteurs sont perçus comme hostiles plus ils déclarent exercer un fort contrôle sur leurs productions (Serey $\rho = .640$, $p < .01$; Le Chequer $\rho = .615$, $p < .01$; Voynet $\rho = .446$, $p < .05$).

Pour ces trois locuteurs, plus ceux-ci déclarent un fort contrôle de soi plus ils sont également perçus comme hostiles. On trouve la tendance contraire pour Tuailon, seul locuteur pour lequel les stimuli perçus comme les plus hostiles sont ceux pour lesquels son degré de contrôle de soi auto-évalué est minimal ($\rho = -.506$, $p < .01$).

Tuailon est également le seul locuteur pour lequel on trouve une corrélation significative entre hostilité perçue et $dDom$ ⁷, mesure de différence entre contrôle de soi auto-évalué et dominance perçue. Comme on peut le voir sur la figure ci-dessous, ses stimuli perçus comme les moins hostiles sont également ceux pour lesquels il s'auto-attribue un degré de contrôle de soi plus élevé que son score de dominance perçue.

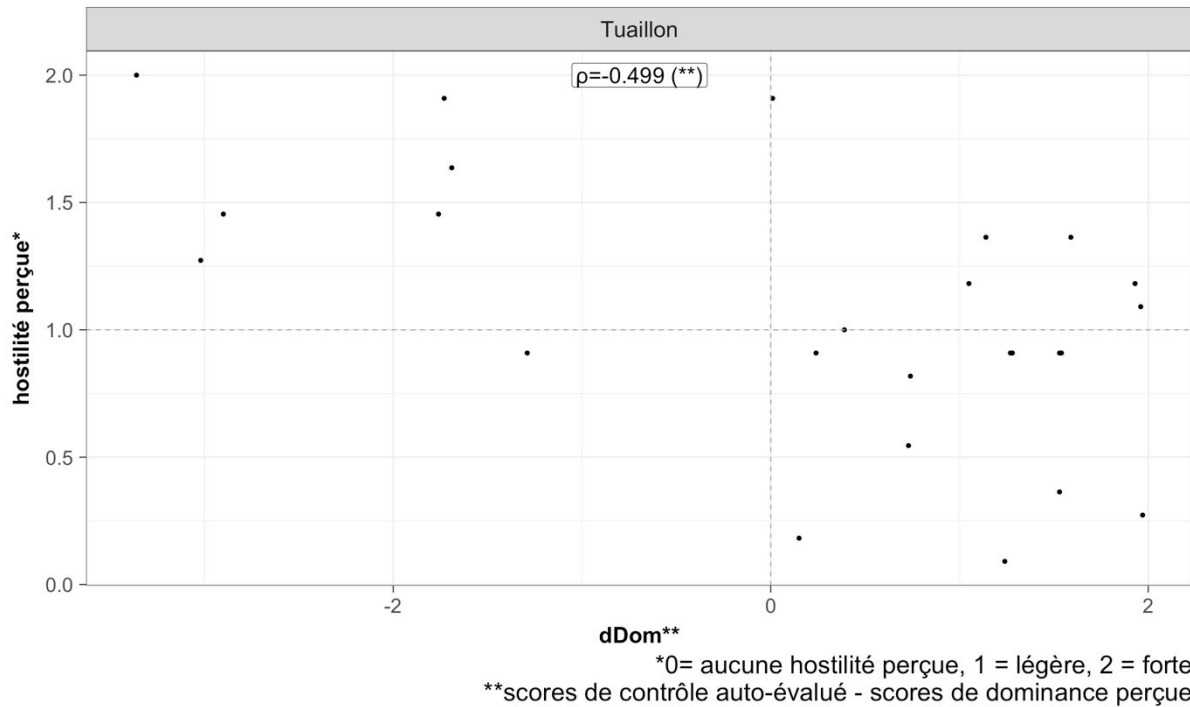


Figure IV-16 : Corrélations entre hostilité perçue (audio) et $dDom$, mesure de différence entre scores de contrôle de soi auto-évalué et scores de dominance perçue pour Tuailon. Les étoiles représentent la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

⁷ Pour rappel, les valeurs les plus hautes de $dDom$ équivalent à des cas de stimuli pour lesquels les locuteurs déclarent un contrôle de soi maximal et ne sont perçus pas perçus comme spécialement dominants ou soumis. De même les valeurs les plus basses de $dDom$ équivalent à des cas de stimuli pour lesquels les locuteurs obtiennent des scores de dominance perçue plus élevés que leurs scores de contrôle de soi auto-évalués.

V. Délimitation des unités d'analyse par les frontières prosodiques majeures

Dans un premier temps, après avoir explicité les critères d'annotation appliqués (V.1), nous présentons les résultats sur le corpus original et sa contrepartie relue de l'annotation semi-automatique des frontières prosodiques majeures (V.2) et de l'annotation qualitative des contours mélodiques aux frontières (V.3). L'annotation des frontières prosodiques majeures a été effectuée à l'échelle de la syllabe, avec pour critères la présence d'une pause silencieuse de durée supérieure à 140ms et/ou un allongement relatif supérieur à trois fois les syllabes avoisinantes (marquage par la durée), et/ou une saillance mélodique d'empan de 10 demi-tons minimum (marquage par F0). 411 frontières prosodiques majeures ont été annotées dans le corpus original contre 382 dans le corpus relu. En ce qui concerne les frontières prosodiques, les résultats sont les suivants :

- Tous locuteurs confondus, le corpus original comporte plus de frontières prosodiques marquées par des pauses silencieuses que le corpus relecture pour quatre locuteurs sur cinq. Comparativement au corpus relecture, il comporte moins de frontières prosodiques marquées par la F0 (trois locuteurs sur cinq) ainsi que par la durée (un locuteur, Le Chequer).

- Les frontières marquées par la durée ne représentent dans les deux conditions du corpus que 16 à 20% des frontières prosodiques, à l'exception de Voynet pour qui elles en représentent 40%.

- Les comparaisons original-relecture à l'échelle du stimulus indiquent un nombre plus élevé de frontières prosodiques marquées uniquement par l'insertion d'une pause silencieuse pour quatre locuteurs sur cinq. Voynet est la seule locutrice à conserver un nombre similaire de frontières majeures marquées par l'ensemble des paramètres prosodiques (pause + F0 + durée) entre corpus original et relecture.

L'annotation des contours mélodiques a été effectuée de façon manuelle et qualitative à chaque frontière prosodique majeure. Les résultats sont les suivants :

- Dans le corpus original, selon les locuteurs, les contours montants représentent 24 à 30% des contours annotés, les contours de focalisation 16 à 20%, les contours plats 10 à 20%, les contours en cloche 8 à 10% et les contours terminaux 2 à 10%.

- Les comparaisons systématiques effectuées à l'échelle de l'extrait indiquent que le corpus original se distingue du corpus relecture par un nombre inférieur de contours montants pour un locuteur, et par un nombre supérieur de contours en cloche (trois locuteurs), de contours plats (trois locuteurs sur cinq), de contours descendants (deux locuteurs sur cinq) et de contours de focalisation (un locuteur).

V.1. Critères pour l'annotation des frontières prosodiques majeures

V.1.1. Protocole d'annotation des frontières

Nous présentons ici le protocole mis en place pour effectuer, à l'échelle de la syllabe, une annotation en frontières prosodiques majeures dont les principes généraux suivent les critères de Mertens et Simon (2013) à travers la prise en compte de trois critères :

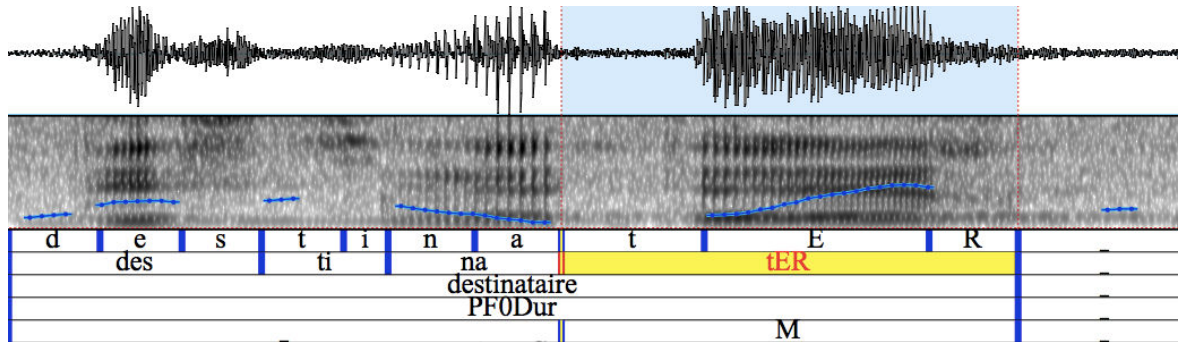
- La pause silencieuse (P) : syllabe suivie par une pause silencieuse d'une durée supérieure à 140 ms¹ ; dans le cadre de cette étude nous adopterons alors le terme de pause « frontière ».
- La fréquence fondamentale (F0) : syllabe sur laquelle on mesure au niveau de la voyelle une variation mélodique supérieure ou égale à 10 demi-tons ou une hauteur moyenne de F0 supérieure de 10 demi-tons à la moyenne des deux syllabes précédentes et de la syllabe suivante (Syll-2/Syll+1)² ; conformément aux critères appliqués par Simon et Degand (2011) et Mertens et Simon (2013).
- La durée syllabique (Dur) : syllabes d'une durée relative supérieure ou égale à 3 fois la moyenne des syllabes environnantes sur une fenêtre d'analyse asymétrique de deux syllabes avant et une syllabe après la syllabe analysée. Cette procédure d'analyse a été également effectuée conformément aux critères appliqués par Simon et Degand (2011) et Mertens et Simon (2013), à l'exception du fait que nous avons reconsidéré dans un second temps les cas d'allongements dus à l'insertion de schwas épenthétiques (voir partie V.1.2 ci-dessous).

Pour clarifier cette méthode d'annotation, nous présentons à la page suivante un exemple d'annotation de frontière prosodique majeure selon les critères susmentionnés.

¹ La pause silencieuse est le seul paramètre pour lequel nous avons adapté le seuil fixe de 200 millisecondes utilisé par Mertens et Simon (2013), de façon à réduire le risque de ne pas prendre en compte certains cas de pauses intra-syntaxiques (ou non structurantes) connues pour être de durée réduite (Candea, 2000 ; Ferré 2004), et dont l'importance paralinguistique est néanmoins importante (Sacks, Schlegloff et Jefferson, 1978 ; Duez, 1991 ; Ferré, 2004 ; Béchet, 2013). De plus, tous types de pauses confondues, Grosman et Simon (2018) ont récemment montré que parmi 14 styles de parole analysés, la parole extraite de débats politiques est celle pour laquelle on mesure les durées de pauses silencieuses les plus basses et la variabilité la plus réduite, se situant ainsi à l'extrême opposé des caractéristiques observées pour la parole politique publique. La question du seuil de durée minimal des pauses silencieuses sera reprise dans la discussion.

² Fenêtre d'analyse adaptée en cas de syllabes suivies d'une pause ou précédées de seulement une syllabe et non deux, voir exemple page 4.

Nous avons implémenté ces critères dans un script d'analyses créé sur Praat et nous avons ensuite vérifié et corrigé manuellement les annotations obtenues. La présence d'une pause silencieuse « frontière » a été codée « P », celle d'un allongement syllabique « Dur », et celle d'un contour de F0 proéminent « F0 ». Nous avons alors annoté chaque frontière prosodique majeure selon le « profil » multiparamétrique des caractéristiques obtenues.



Exemple V-1 : Spectrogramme (échelle de 0 à 5000 Hz), courbe mélodique (échelle 75 à 400 Hz) et tirs d'annotation en phones, syllabes, mots, et frontière prosodique majeure, et contour mélodique du mot « destinataire » (extrait du stimulus n°13 Tuailleon_stimulus1_orig) dont la syllabe finale donne lieu à une annotation de frontière prosodique majeure de type « PF0Dur » (cumul des trois paramètres annotés). Transcription simplifiée de l'extrait entier : « et puis vous avez été aussi forcément, comme nous tous, destinataires de mails et puis d'informations, de blogs (...) ».

Dans l'exemple ci-dessus, la syllabe [tɛʀ] ([tɛʀ] en notation SAMPA dans la tire d'annotation Praat) a été annotée comme frontière prosodique de type PF0Dur, c'est-à-dire comme remplissant cumulativement les trois critères de pause silencieuse, de F0 et de durée (syllabique) pour l'annotation prosodique majeure en même temps :

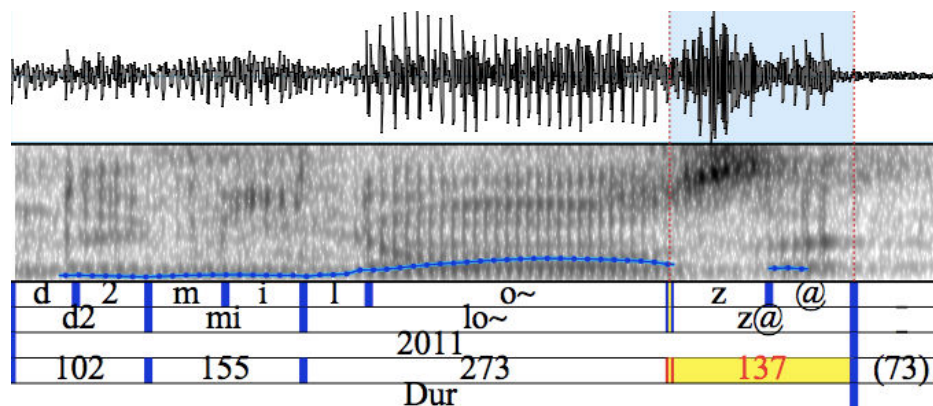
- Pause (P) : ici la pause silencieuse suivant la syllabe dure 156 ms, et est donc considérée comme pause dite « frontière ».
- F0 : sur l'ensemble de la durée de la voyelle [ɛ] on calcule une valeur de F0 minimale de 114 Hz, et maximale de 205 Hz. Le calcul de la plage de variation de F0 à partir de ces valeurs équivaut à une valeur de 10.15 demi-tons³.
- Durée : la syllabe [tɛʀ] durant 382 ms et la moyenne des durées des deux syllabes précédentes ([ti] 104 ms + [na] 143 ms) étant de 123 ms, le calcul de l'allongement relatif est de 3.10 (382 ms / 123 ms), signifiant que la syllabe [tɛʀ] dure un peu plus de trois fois plus longtemps que la moyenne des deux syllabes précédentes.

³ Le calcul implémenté dans le script Praat d'annotation correspond à $12 * \log_2(F0_{min}/F0_{max})$.

V.1.2. Adaptation de l'annotation dans les cas de schwas épenthétiques

Durant la phase de correction manuelle de la segmentation automatique en phones, syllabes et mots, nous avons remarqué plusieurs cas de schwas épenthétiques, notamment pré-pausaux en plus des pauses remplies de type « euh » dit d'hésitation préalablement annotées et segmentées (voir III.4.2.1 annotation des « disfluences »).

Lors de la correction manuelle de l'annotation des frontières prosodiques majeures, si conformément aux recommandations de Degand et collègues (2014) nous avons exclu les cas de « disfluences » de l'annotation des frontières prosodiques majeures, nous avons néanmoins été amené à reconsidérer la présence de ces schwas épenthétique du fait des erreurs d'annotation automatique qu'ils provoquaient. Nous présentons, pour expliciter ce cas et décrire notre procédure d'annotation adaptée, l'exemple ci-dessous du mot « 2011 » terminé par un schwa épenthétique dans le stimulus n°5 produit par Tuailleon (corpus original).



Exemple V-2 : Spectrogramme, courbe mélodique (échelle 75 à 400 Hz) et tirs d'annotation en phones, syllabes, mots, et durées syllabiques du mot « 2011 » terminé par un schwa épenthétique (extrait du stimulus n°5 de Tuailleon) dont la syllabe finale [zə] analysée ne donnerait lieu à aucune annotation de frontière prosodique majeure. Cependant, si l'on considère la « macro-syllabe » [lɔzə] incluant le schwa épenthétique comme dans notre étude, celle-ci donne par contre lieu à une annotation de frontière prosodique marquée par la durée « Dur ». Transcription simplifiée de l'extrait entier : « moi je trouve quand même qu'il y a quelques différences entre le rapport interne que vous aviez présenté en 2011, et ce rapport (...) ».

Durant l'annotation semi-automatique, avant correction, aucune frontière prosodique n'était annotée à partir de la syllabe finale [zə] du mot « 2011 » de durée de 137 millisecondes. Pour éviter ces cas de non-considération de frontière, erronés selon nous, nous avons décidé dans cette étude, de la même façon qu'Avanzi (2011) faisant état de problèmes analogues, de considérer la « macro-syllabe »⁴ [lɔzə] d'une durée totale de 410 millisecondes pour effectuer

⁴ Les cas de schwas épenthétiques entraînant une vérification spécifique pour l'annotation prosodique ont été inventoriés puis traités de façon semi-automatique avec un script *Praat*. Nous avons fait le choix d'annoter ces phénomènes à part et de ne pas effectuer directement de modification dans la tire syllabe, car la question de l'impact du schwa (épenthétique ou non) sur la syllabation nécessite des recherches approfondies, centrées sur ce point, sur un type de corpus plus large ou plus équilibré.

l'analyse, et calculant alors un allongement relatif de 3.2 par rapport à la moyenne des deux syllabes précédentes [dø] (102 millisecondes) et [mi] (155 millisecondes).

Enfin nous précisons avant d'amorcer la présentation synthétique des résultats que les raisons de nos adaptations personnelles (durée des pauses silencieuses, prises en compte des schwas épenthétiques pour les phénomènes d'allongement) ainsi que les limites des seuils respectés pour la durée et la F0 seront discutées dans le chapitre IX final.

V.1.3. Typologie des frontières prosodiques annotées

Conformément aux principes appliqués et détaillés ci-dessus, nous avons alors annoté chaque frontière prosodique majeure selon le « profil » multiparamétrique des caractéristiques obtenues. Il en résulte une typologie à sept niveaux que nous détaillons ci-dessous. :

1. « PF0Dur » : frontière prosodique majeure marquée par la combinaison des trois paramètres : pause silencieuse, F0 et durée.
2. « PF0 » : frontière prosodique majeure marquée par une pause silencieuse et F0.
3. « PDur » : frontière prosodique majeure marquée par l'insertion d'une pause silencieuse et d'un allongement relatif supérieur à 3.
4. « P » : frontière prosodique majeure marquée uniquement par l'insertion d'une pause silencieuse.
5. « F0Dur » : frontière prosodique majeure marquée par F0 et la durée, mais pas par une pause silencieuse.
6. « F0 » : frontière prosodique majeure marquée uniquement par une variation de F0.
7. « Dur » : frontière prosodique majeure marquée uniquement par un allongement relatif supérieure à 3.

V.2. Résultats de l'annotation prosodique paramètre par paramètre

Nous présentons ici les résultats de l'annotation prosodique en fonction des pauses silencieuses, de l'allongement syllabique final et des variations de F0. Après présentation du décompte général des frontières prosodiques majeures annotées, nous présentons les résultats en détaillant paramètre par paramètre à l'échelle du corpus et en moyenne par stimulus, puis nous donnons le détail des différentes configurations multiparamétriques annotées.

V.2.1. Frontières prosodiques majeures annotées

À travers la présence d'au moins un des trois paramètres selon les critères susmentionnés, 411 frontières prosodiques majeures (et donc autant d'unités prosodiques majeures) ont été annotées dans le corpus original et 382 dans le corpus relecture. Nous détaillons ici le nombre par locuteur et par condition, en distinguant frontières prosodiques majeures marquées ou non marquées, entre autres paramètres, par une pause silencieuse (durée supérieure à 140 ms) ; voir section V.2.2 ci-dessous pour la présentation des résultats concernant spécifiquement les frontières marquées par ce paramètre.

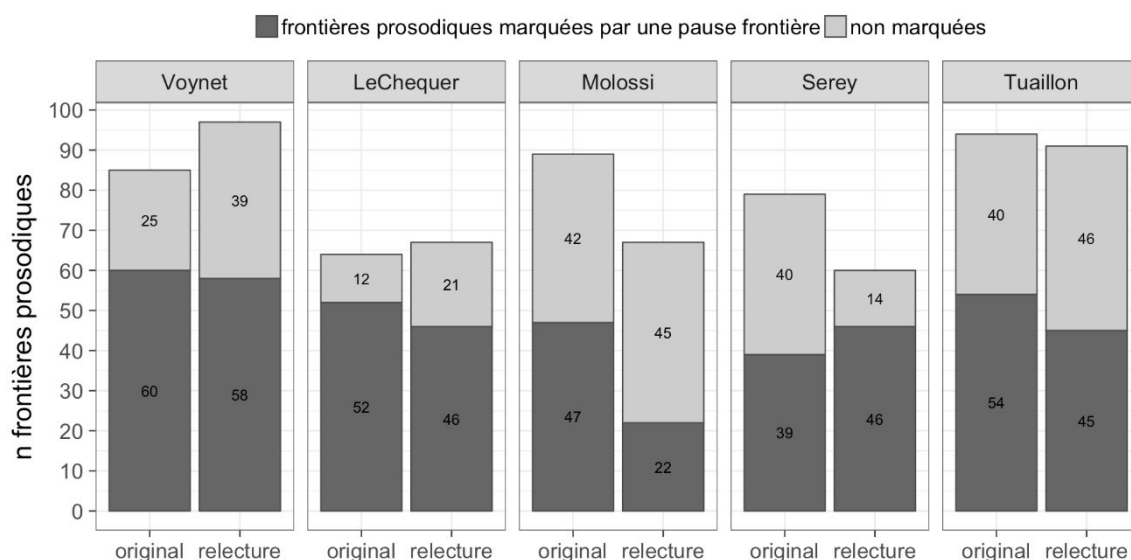


Figure V-1 : Décompte du nombre de frontières prosodiques majeures annotées par locuteur et par condition avec spécification de la présence ou non d'une pause « frontières » (durée > 140 ms) parmi les critères ayant justifié l'annotation.

Si on considère à présent le nombre de frontières prosodiques majeures par stimulus (voir Figure V-2 : ci-dessous), les comparaisons statistiques montrent que si tous locuteurs confondus il n'y a pas de différence significative entre condition originale et relecture, Molossi ($t[21]=2.1$, $p<.05$, $\eta^2=17\%$) et Serey ($t[22]=2.2$, $p<.05$, $\eta^2=18\%$) produisent significativement plus de frontières prosodiques majeures au sein de leurs extraits originaux que de leurs extraits relus.

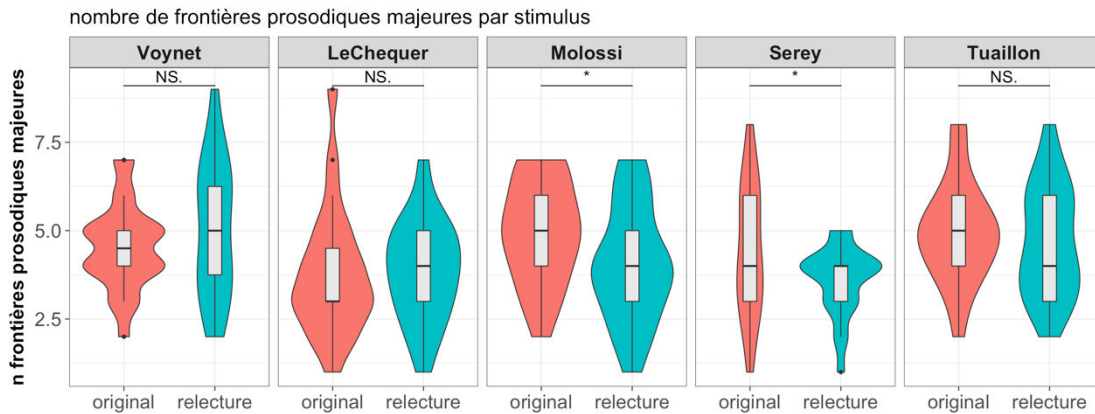


Figure V-2 :Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques majeures par stimulus (tous paramètres confondus) par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : p<.001 = ***, p<.01 = **, p<.05 = *, p>.05 = NS, non significatif.

V.2.2. Frontières prosodiques marquées par des pauses

On compte au total 252 frontières prosodiques majeures marquées entre autres paramètres par une pause silencieuse de durée supérieure à 140 ms dans le corpus original contre 218 en condition de relecture. Dans le corpus original, le décompte total de frontières prosodiques marquées par des pauses par locuteur va de 39 pour Serey à 60 pour Voynet. Le décompte précis par locuteur est consultable sur la Figure V-1 ci-dessus. Ce décompte équivaut également au nombre de pauses « frontières » analysables dans le corpus⁵. Nous présentons ci-dessous la proportion que représentent les frontières prosodiques marquées, entre autres paramètres, par une pause silencieuse par rapport à l'ensemble des frontières prosodiques annotées pour chaque locuteur.

Dans le corpus original, Le Chequer est le locuteur pour lequel les frontières marquées par une pause silencieuse représentent la proportion la plus importante parmi l'ensemble des frontières prosodiques majeures annotées : 81%, contre 69% dans le corpus relecture. Molossi est le locuteur pour lequel on trouve la différence original-relecture la plus marquée, la proportion de ses frontières prosodiques majeures annotées comme marquées par une pause « frontière » passant de 53% (47 frontières) à 33% (22) entre la condition originale et la condition relecture.

⁵ Si durant l'annotation prosodique nous avons tenu compte de la présence ou non d'une pause en fin de stimulus, celles-ci sont exclues de l'analyse car pas audibles par les auditeurs, car tous les stimuli ont été tronqués juste avant. Pour information, le nombre de stimuli dont l'unité prosodique finale est marquée par une pause (bien que non analysable ici) est de 110 dans le corpus original et de 105 dans le corpus relecture.

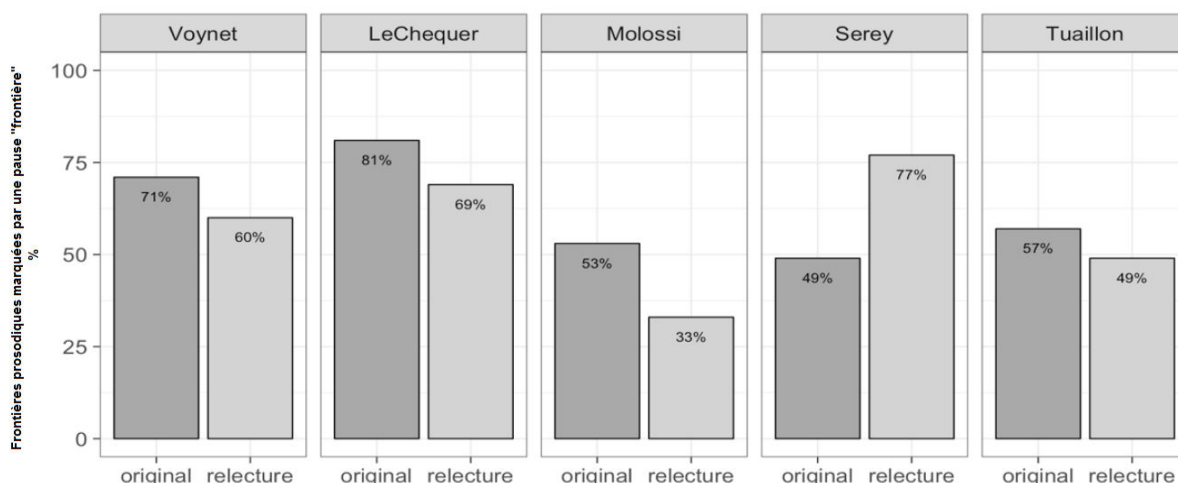


Figure V-3 : Proportions de frontières prosodiques majeures marquées par une pause « frontière » (durée > 140 ms) pour chaque locuteur dans le corpus original et la relecture.

Si l'on considère à présent le nombre de pauses « frontières » par stimulus et que l'on compare les deux conditions comme l'illustre la figure ci-dessous, on constate que Molossi est le seul locuteur pour lequel il y a significativement moins de pauses dans la condition relecture ($t[21]=4.3$, $p<.001$, $\eta^2=47\%$).

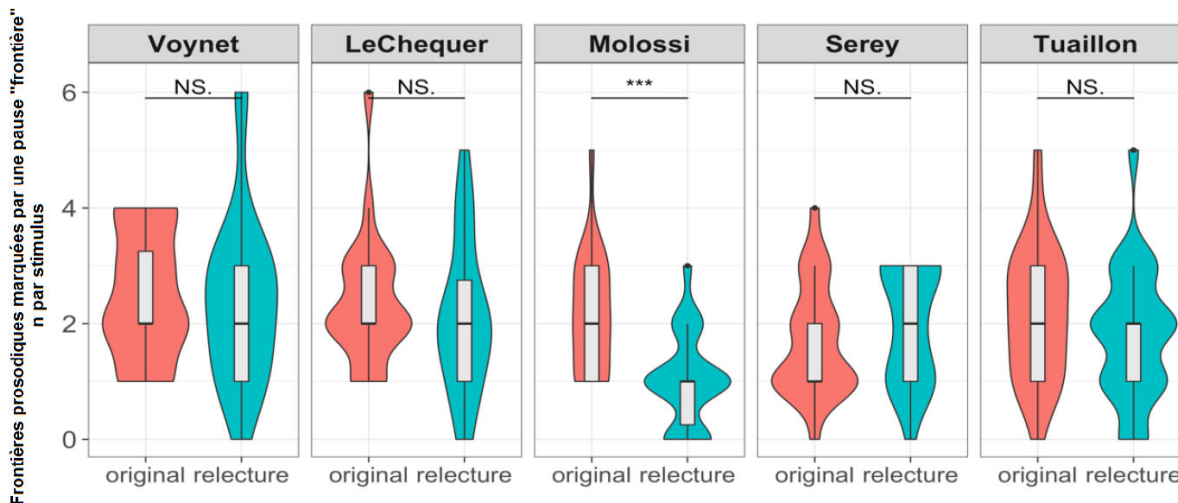


Figure V-4 : Diagrammes en violon et boîtes à moustaches du nombre de frontières prosodiques majeures marquées entre autres par l'insertion d'une pause « frontières » (durée > 140 ms) par stimulus dans le corpus original et la relecture. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p<.001$, ** = $p<.01$, * = $p<.05$, ns = $p>.05$.

V.2.3. Frontières marquées seulement par la présence d'une pause

Nous présentons ici le nombre de frontières prosodiques annotées seulement du fait de la présence d'une pause silencieuse (>140ms). Il s'agit donc des cas de frontières prosodiques

pour lesquelles ni une variation de F0, ni une durée supérieure au seuil fixé n'a été annotée, le seul paramètre ayant permis l'annotation étant la présence d'une pause.

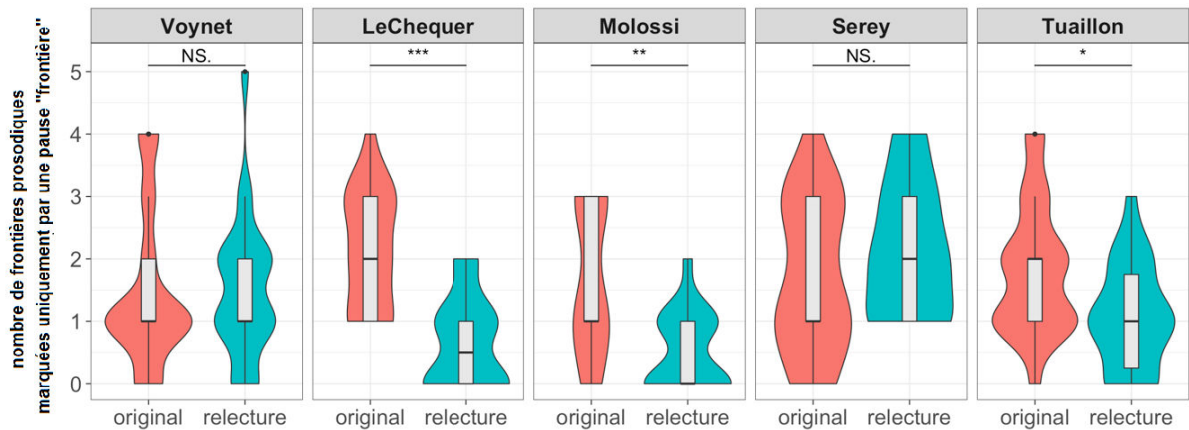


Figure V-5 : Diagrammes en violon et boîtes à moustaches du nombre de frontières prosodiques majeures marquées uniquement par l'insertion d'une pause « frontières » (durée > 140 ms) par stimulus dans le corpus original et la relecture. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, ns = $p > .05$.

Tous locuteurs confondus, il y a significativement plus de frontières de ce type dans le corpus original qu'en relecture concernant le nombre moyen par stimulus ($t[108]=3.7$, $p < .001$, $\eta^2=11\%$). Cette différence est significative pour Le Chequer ($t[21]=5.9$, $p < .001$, $\eta^2=63\%$), Molossi ($t[17]=3.8$, $p < .001$, $\eta^2=46\%$) et Tuailleon ($t[21]=2.6$, $p < .01$, $\eta^2=25\%$).

V.2.4. Frontières prosodiques marquées par F0

Nous présentons à présent le nombre de frontières prosodiques annotées comme marquées par F0. Pour rappel, les critères pour considérer le marquage mélodique d'une frontière prosodique sont les suivants : soit une montée mélodique supérieure ou égale à 10 demi-tons mesurée sur la durée de la voyelle, soit une hauteur moyenne de F0 supérieure de 10 demi-tons à la moyenne des deux syllabes précédentes et de la syllabe suivante.

On répertorie un total de 270 frontières prosodiques annotées comme marquées par F0 dans le corpus original contre 304 dans le corpus relecture.

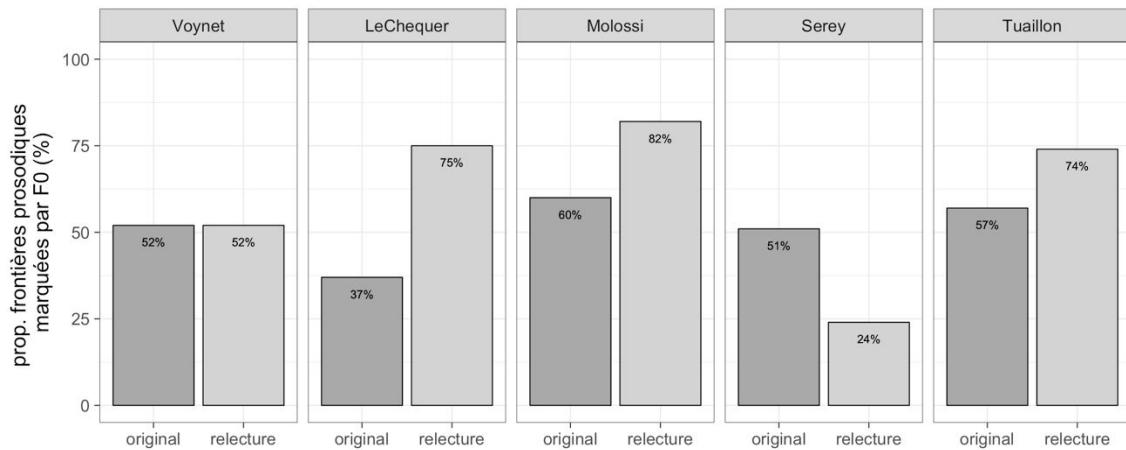


Figure V-6 : Proportions de frontières prosodiques majeures marquées, entre autres paramètres, par F0 pour chaque locuteur dans le corpus original et la relecture.

Comme on peut l'observer dans la figure ci-dessus, dans le corpus original, Le Chequer est le locuteur pour lequel les frontières prosodiques marquées par F0 sont les plus minoritaires par rapport à l'ensemble de ses frontières prosodiques produites (37% contre 51 à 60% pour les autres locuteurs). Dans le corpus relecture, Serey est le locuteur pour lequel on trouve la proportion la plus faible de frontières prosodiques majeures par F0 (24% contre 52 à 82% pour les autres locuteurs).

À l'échelle du stimulus, ces deux tendances particulières se retrouvent. Le Chequer produit significativement moins de frontières marquées par F0 dans le corpus original qu'en relecture ($t[21]=-3.6$, $p<.01$, $\eta^2=39\%$), et Serey significativement plus ($t[20]=3.3$, $p<.01$, $\eta^2=35\%$). L'ANOVA et les tests post-hoc effectués sur le corpus original ($F[4,108]=3.66$, $p<.01$; $\eta^2=12\%$) indiquent que Le Chequer produit en moyenne moins de frontières prosodiques marquées par F0 dans le corpus original que Tuillon ($p<.001$) et Molossi ($p<.01$).

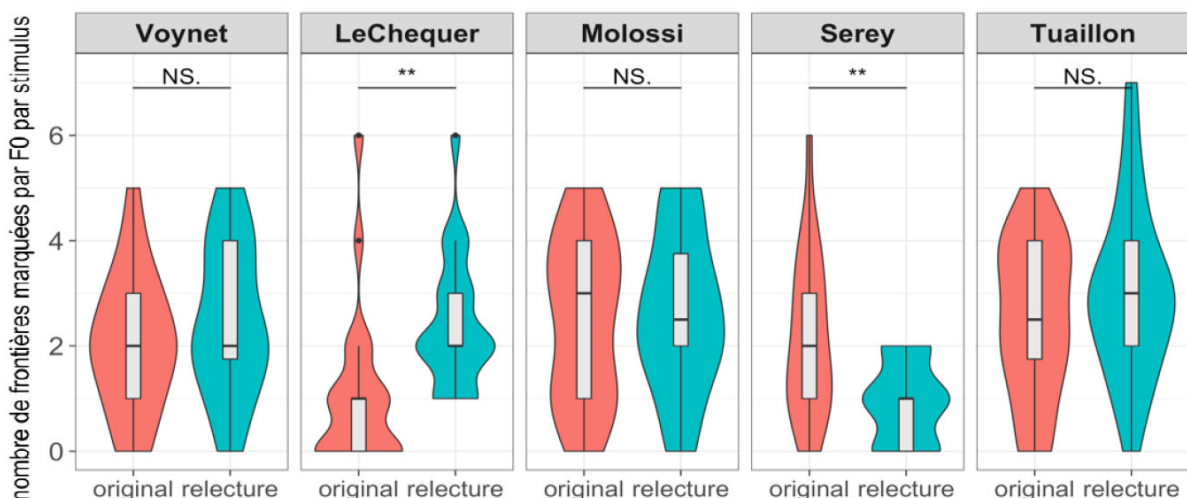


Figure V-7 : Nombre de frontières prosodiques marquées entre autres par F0 par stimulus selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p<.001$, ** = $p<.01$, * = $p<.05$, ns = $p>.05$.

V.2.5. Frontières prosodiques marquées par la durée

On compte un total de 120 frontières prosodiques annotées comme marquées par la durée dans le corpus original contre 143 dans le corpus relecture. Voynet est la locutrice pour laquelle on compte le plus haut nombre et la proportion la plus importante de frontières marquées par la durée dans le corpus original : 44 frontières « durée », soit 40% du total de ses frontières prosodiques annotées. Pour les autres locuteurs, le taux de frontières prosodiques marquées par la durée varie de 16% (Serey) à 21% des frontières prosodiques annotées (Tuillon).

Le Chequer est le seul locuteur pour lequel on trouve significativement plus de frontières marquées par la durée en condition originale qu'en relecture ($t[18]=-2.3$, $p<.05$, $\eta^2=23\%$). On observe également au sein du corpus original une différence significative entre locuteurs ($F[4,77]=3.41$, $p=0.01$; $\eta^2=15\%$), et les tests post-hoc indiquent que Voynet produit significativement plus de frontières marquées par la durée que Le Chequer ($p<.01$) et Tuillon ($p<.05$).

V.2.6. Typologie multiparamétrique des frontières prosodiques majeures

Pour tous les locuteurs, le type de frontière prosodique le plus répandu correspond aux frontières annotées « P », c'est-à-dire aux frontières annotées comme majeures sur le plan prosodique uniquement du fait de la présence d'une pause « frontière » et non des paramètres de durée et/ou F0 ; voir Figure V-8 page suivante.

Pour Dominique Voynet, on trouve également un nombre élevé de frontières annotées « PF0Dur » c'est-à-dire pour lesquelles les trois paramètres sont présents pour constituer la frontière (Pause + F0 + durée). Ce type précis de frontière est minoritaire pour tous les locuteurs sauf Voynet dans le corpus original : Voynet en produit 20 contre 4 à 7 seulement pour les autres locuteurs. En effet, pour les quatre autres locuteurs le nombre de frontières prosodiques « PF0Dur » est très réduit dans le corpus original. Cette particularité dans les caractéristiques prosodiques de Voynet pourrait s'expliquer par son rôle de Maire, ainsi que de médiatrice des débats, pouvant entraîner un degré de formalité plus élevé dans certaines de ces informations et éventuellement des caractéristiques de parole plus proches de la parole lue que de la parole spontanée (Koch et Oesterreicher, 2008).

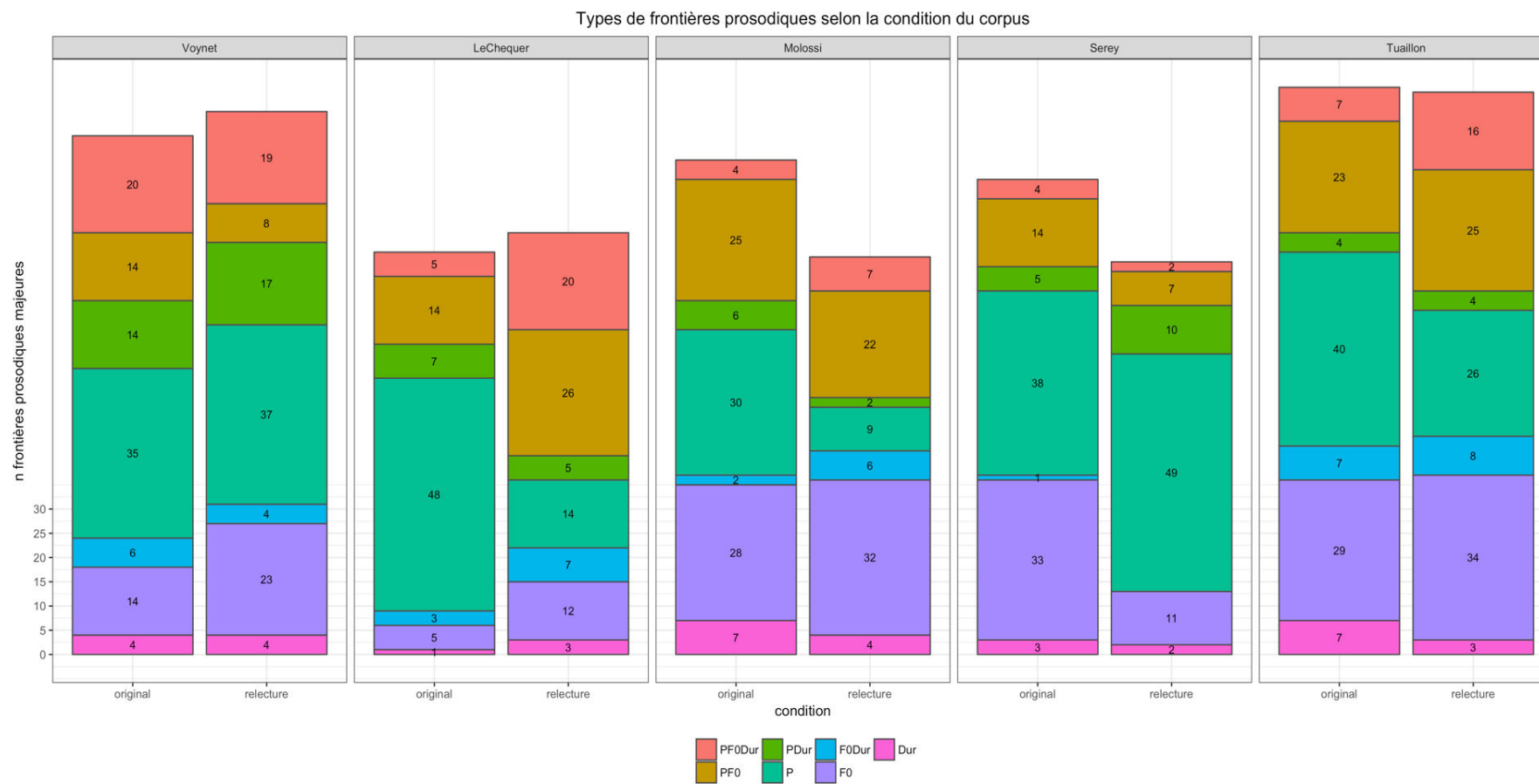


Figure V-8 : Décompte des différentes configurations multiparamétriques pour les frontières prosodiques majeures.

V.3. Annotation des contours mélodiques aux frontières prosodiques majeures

Dans cette partie nous présentons les résultats de l'annotation manuelle des contours mélodiques qui a été effectuée à chaque frontière prosodique annotée comme majeure.

À partir de la terminologie proposée ici par Degand et collègues (2014), nous avons annoté les contours mélodiques à chaque frontière prosodique majeure à travers les catégories suivantes : contour montant « M », contour plat « P », contour descendant (« D », considéré comme terminal « T » dans la terminologie de Degand et collègues) et contour de focalisation « F ». Nous avons ajouté à ces quatre types de contours deux autres catégories : une pour les contours en cloches (« C »), et une pour les cas de syllabes où la F0 ne peut être détectée (« noF0 »).

Nous présentons ici les caractéristiques principales pour chaque type de contours en partant des plus montants aux plus descendants :

- Le contour montant « M » tel qu'il est envisagé dans la terminologie de Degand et collègues correspond aux contours de continuation (Delattre, 1966). La présente annotation ne permet pas en l'état de différencier contours de continuation mineure ou majeure. Suivant le protocole de Degand et collègues, sont annotées avec un contour montant toutes syllabes à partir desquelles on considère « une montée intra-syllabique ou un intervalle positif par rapport à la syllabe précédente » (Degand et collègues, 2014).
- Le contour en cloche (« C ») correspond à une courbe mélodique montante puis descendante de forme arrondie.
- Le contour plat (« P ») est un contour plat qui correspond aux contours dits « suspens » (« S ») dans la terminologie de Degand et al (2014) et renvoie à la définition suivante : « plat, parfois allongé, sans écart mélodique par rapport à la syllabe qui précède » (Degand et collègues, 2014 p. 2615). Ce dernier point est essentiel car un contour relativement plat mais de hauteur moyenne supérieure à la syllabe précédente sera plutôt considéré comme appartenant à la catégorie « M » qu'à la catégorie « P ».
- Le contour descendant (« D ») correspond aux courbes mélodiques descendantes englobées dans la catégorie contour terminal (« T ») chez Degand et collègues (2014). Nous avons fait le choix d'adapter la terminologie d'annotation au sujet des contours descendants de façon à pouvoir prendre en compte le caractère descendant de certains contours sans pour autant que ceux-ci soient produits en fin de phrase déclarative comme les contours dits terminaux (Degand et collègues, 2014) ou de finalité (Delattre, 1966).

Les contours descendants ont donc été annotés, au même titre que les contours montants, sans condition de localisation dans la structure syntaxique.

- Le contour focus (« F ») a été annoté dans un second temps, une fois les catégories précédentes de contours annotés, au regard de l'annotation lexicale et syntaxique. En cas de frontières prosodiques majeures localisées sur des syllabes non finales de mots (ou de groupes de mots), les contours initialement annotés comme appartenant à l'une des catégories susmentionnées (généralement « D » ou « M ») ont été ré-annotés comme « F ».
- Les contours situés en frontières d'unités réactionnelles finales de phrases ont été ré-annotés comme contours terminaux (« T ») et sont majoritairement constitués de contours initialement annotés comme descendants (« D ») ou dévoisés (« noF0 »).

Nous précisons ici que contrairement aux contours montants, descendants, plats et en cloche, l'annotation des contours de focalisation (« F ») et des contours terminaux (« T ») fait appel à une évaluation basée sur des paramètres extra-prosodiques. Dans ces deux cas en effet, il ne s'agit pas d'évaluer simplement les caractéristiques acoustiques, et la forme du contour, mais son interaction avec la structure syntaxique : un contour descendant situé en fin de phrase sera considéré comme terminal, un contour descendant situé à l'intérieur même d'une séquence syntaxique (unité syntaxique intermédiaire) sera considéré comme contour de focalisation.

V.3.1. Présentation générale des contours mélodiques annotés

Nous présentons dans un premier temps le total des contours mélodiques annotés sur l'ensemble du corpus, puis nous comparons les caractéristiques pour chaque locuteur et chaque condition par stimulus.

Tous locuteurs confondus, on trouve dans le corpus original une majorité de contours montants, suivis de contours plats, descendants, et en cloche, dont le détail peut être consulté Figure V-9 page suivante.

Dans le corpus original, les contours annotés comme montants peuvent aller jusqu'à représenter un tiers des contours des frontières prosodiques majeures produites par des locuteurs : 24% pour Voynet contre 28 à 30% pour les autres locuteurs. En condition de relecture, cette proportion va de 21 à 53% des contours produits.

Les contours en cloche ne représentent que 8 à 10% des contours réalisés aux frontières prosodiques majeures dans le corpus original contre 1 à 6% des contours réalisés dans le corpus contrôle. Ces proportions dans le corpus relecture équivalent à seulement 1 à 9 contours en cloche en condition de lecture selon les locuteurs. Ce type de contour est peu courant en parole lue. Voynet est la locutrice qui produit le plus de contours en cloche dans le corpus original, à savoir 18, soit 10% des contours qu'elle produit.

Les contours annotés comme plats sans différence de hauteur moyenne de F0 avec la syllabe précédente, représentent 10 à 20% des contours produits par les locuteurs dans le corpus original. En condition de relecture, la proportion va de 4% pour Molossi à 12% pour Voynet.

Les contours de type focus représentent 16% à 20% des contours produits aux frontières prosodiques majeures dans le corpus original contre 5 à 14% dans le corpus relecture.

Réalisation des contours mélodiques sur l'ensemble du corpus

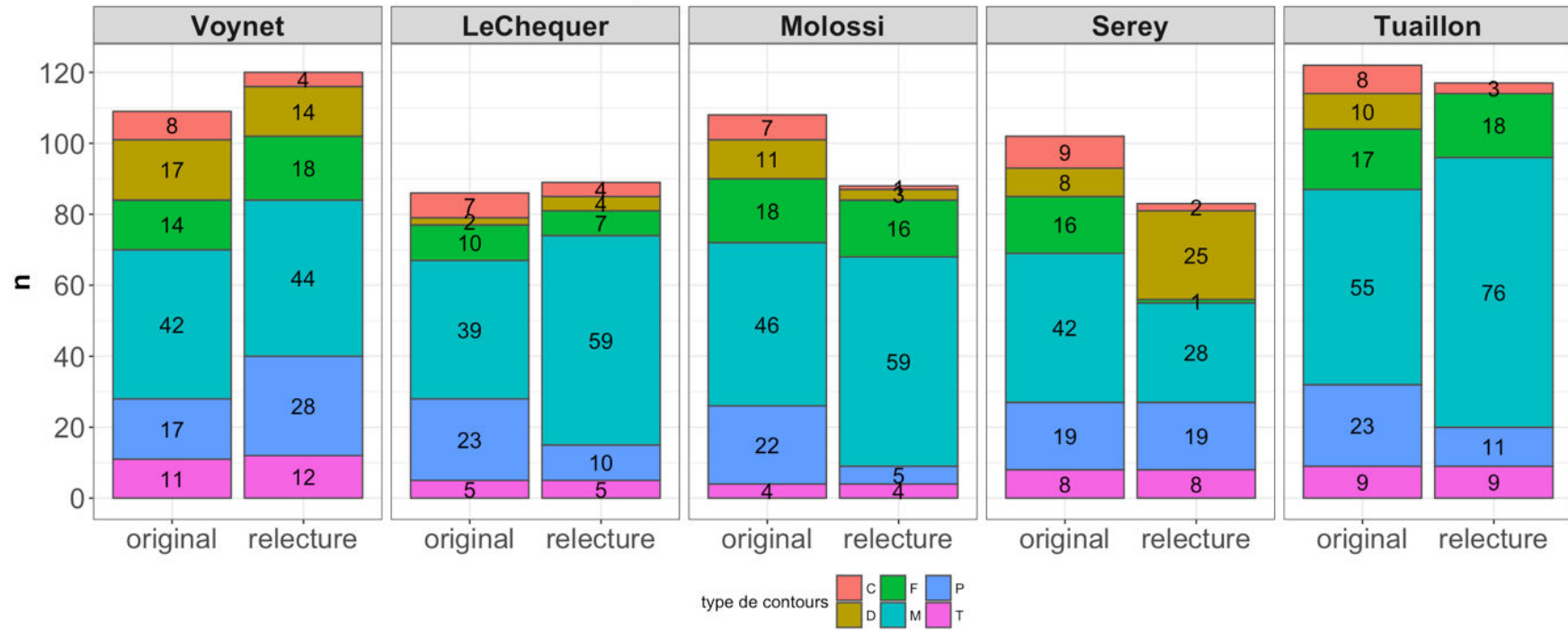


Figure V-9 : Nombre total de contours en cloche (C), de type focus (F), plats (P), descendants (D), montants (M) et terminaux (T) par locuteur et par condition.

V.3.2. Nombre de contours montants par stimulus :

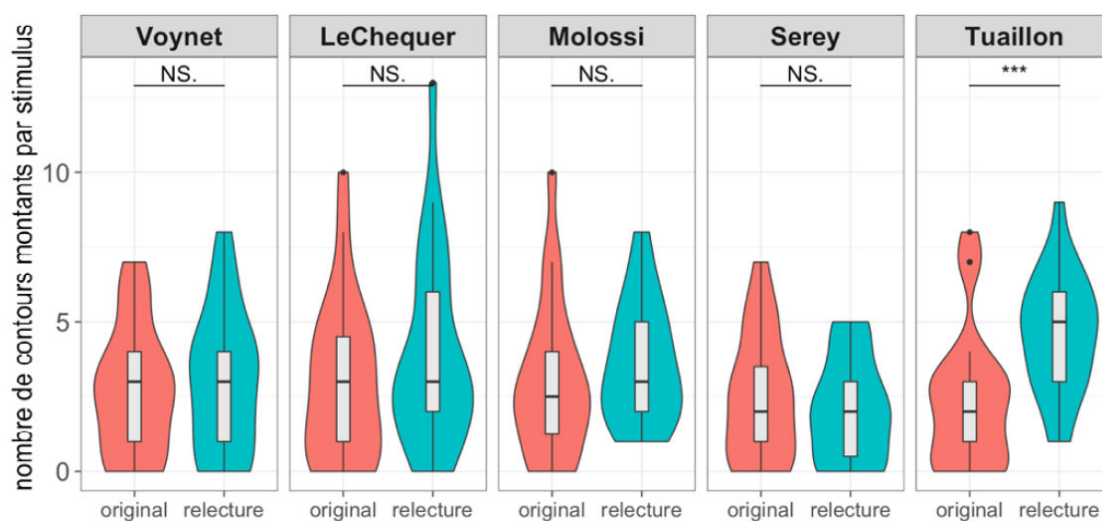


Figure V-10 : Diagrammes en violons et boîtes à moustaches représentant le nombre de contours mélodiques annotés comme montants par stimulus (tous paramètres confondus) par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

Tous locuteurs confondus, le nombre de contours montant par stimulus est en moyenne supérieur dans la condition relecture par rapport aux extraits originaux ($t[116] = -2.7$, $p < .01$, $\eta^2 = 6\%$). Si cette tendance peut s'observer pour Molossi et Le Chequer (voir figure ci-dessus), Tuailon est le seul locuteur pour lequel la différence est significative, et pour lequel on trouve significativement moins de contours montants par stimulus dans ses extraits originaux que dans leurs contreparties relues ($t[24] = -3.8$, $p < .001$, $\eta^2 = 37\%$).

V.3.3. Nombre de contours en cloche

Concernant les contours annotés comme en cloche, nous n'observons pas de différence significative inter-condition ni de différence inter-locuteurs au sein de chaque condition.

V.3.4. Nombre de contours type focus

Les contours de type focus⁶ représentent 16% à 20% des contours produits aux frontières prosodiques majeures dans le corpus original contre 5 à 14% dans le corpus relecture.

⁶ Pour rappel l'annotation des contours de type focus constitue une annotation large de phénomènes de proéminences prosodiques (mélodiques et/ou temporelles) sur des syllabes non finales de mots : les contours de type focus annotés dans la présente étude peuvent être montants comme descendants.

Si l'on considère à présent leur répartition par stimulus, on voit que le nombre de contours de type focus par locuteur n'est significativement supérieur en condition originale (plutôt qu'en condition de relecture) que pour un locuteur sur cinq : Serey ($t[22]=3.5$, $p<.01$, $\eta^2=36\%$).

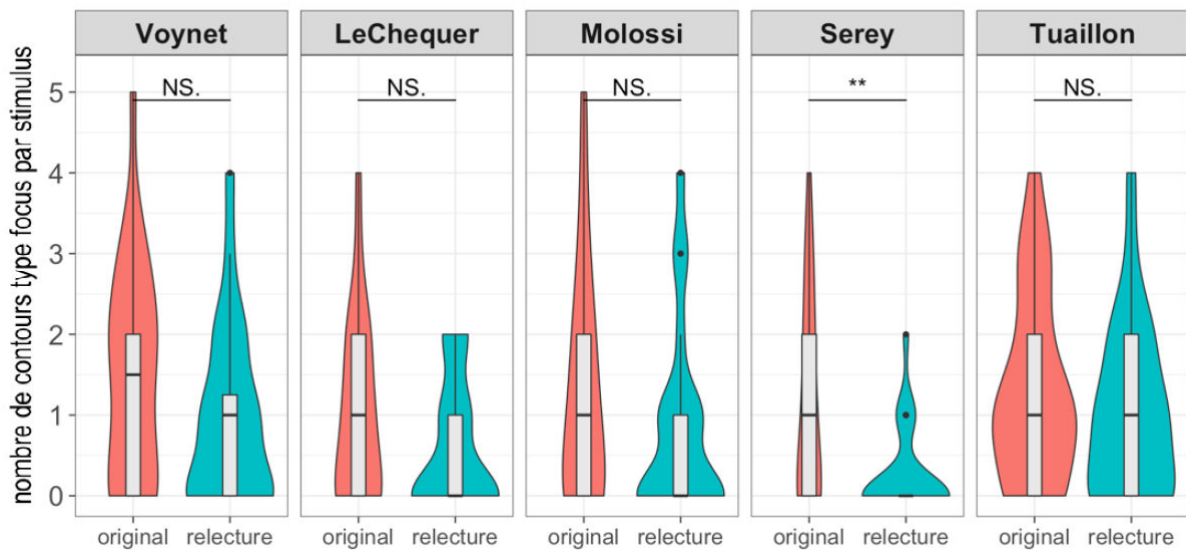


Figure V-11 : Diagrammes en violons et boîtes à moustaches du nombre de contours mélodiques de type focus par stimulus pour chaque locuteur selon la condition. Les étoiles au-dessus des crochets correspondent aux valeurs de p des tests t effectués pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= $p<.05$, **= $p<.01$, ***= $p<.001$, NS.= non significatif).

V.3.5. Nombre de contours plats

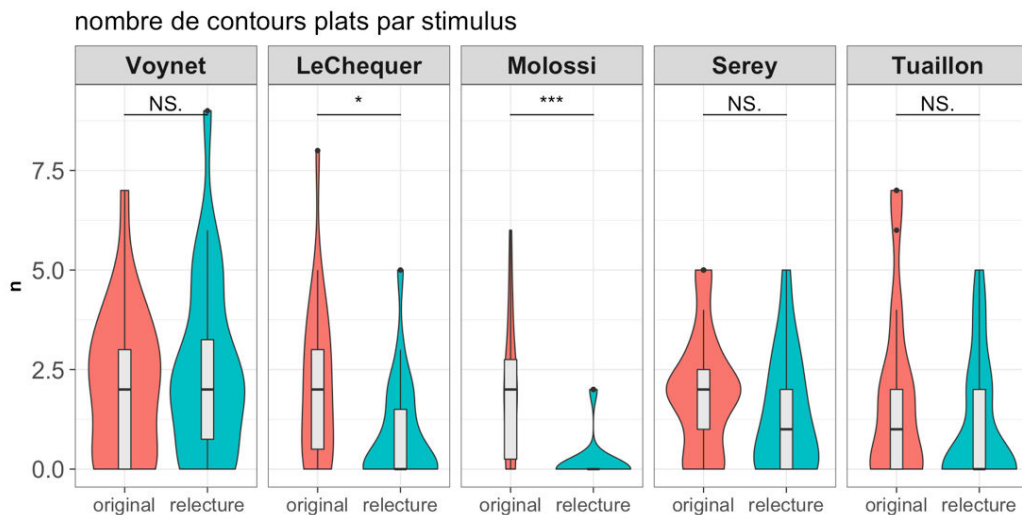


Figure V-12 : Diagrammes en violons et boîtes à moustaches du nombre de contours mélodiques annotés comme plats par stimulus pour chaque locuteur selon la condition. Les étoiles au-dessus des crochets correspondent aux valeurs de p des tests t effectués pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= $p<.05$, **= $p<.01$, ***= $p<.001$, NS.= non significatif).

Comme on peut l'observer sur la figure ci-dessus, le nombre de contours plats est significativement supérieur dans les stimuli originaux au nombre calculé pour les stimuli relus pour deux locuteurs sur cinq : Le Chequer ($t[22]=2.7$, $p<.01$, $\eta^2=25\%$) et Molossi ($t[21]=4.6$,

$p < .001$, $\eta^2 = 50\%$). Si l'on compare les caractéristiques de locuteurs au sein de chaque condition, on constate qu'il n'y a aucune variation significative inter-locuteurs dans le corpus original ($F[4,112] = 0.11$, $p = 0.98$; $\eta^2 = 0\%$). En condition de relecture, une seule différence inter-locuteurs est trouvée : Voynet produit significativement plus de contours plats que Molossi ($F[4,112] = 6.68$, $p = 0$; $\eta^2 = 19\%$, $p < .01$).

V.3.6. Nombre de contours descendants

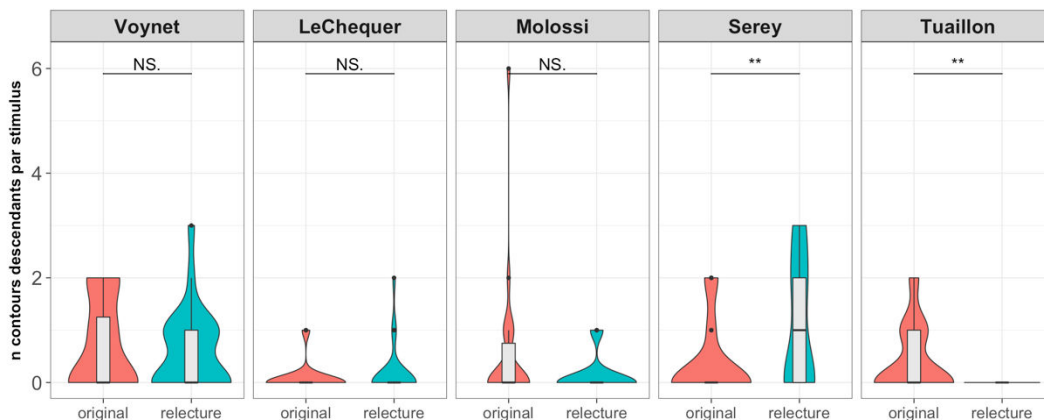


Figure V-13 : Diagrammes en violon et boîtes à moustache représentant le nombre de contours annotés comme descendants par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Concernant les contours descendants, à distinguer des contours terminaux⁷ (voir section suivante ci-dessous), on calcule une différence significative entre conditions pour seulement deux locuteurs sur cinq : Serey ($t[22] = -2.9$, $p < .01$, $\eta^2 = 28\%$), produisant significativement plus de contours descendants en condition de relecture, et Tuailon ($t[24] = 3.1$, $p < .01$, $\eta^2 = 29\%$), ce dernier ne produisant pas de contours descendants en condition de relecture.

V.3.7. Nombre de contours terminaux

Concernant les contours annotés comme terminaux, c'est-à-dire situés en frontière d'unité syntaxique majeure finale de phrase, nous n'observons pas de différence significative inter-condition ni de différence inter-locuteurs au sein de chaque condition.

⁷ Pour rappel, les contours descendants localisés en fin de phrases ont été ré-annotés comme terminaux et les contours descendants produits en syllabes non finales de mots ont été ré-annotés en contours de type focus.

VI. Analyses comparées entre extraits originaux et relus : pauses, durées et F0

Le chapitre VI présente les caractéristiques prosodiques des locuteurs à travers un ensemble de variables relatives aux pauses silencieuses (VI.2), au débit et variations de durées segmentales (VI.3), ainsi qu'aux caractéristiques mélodiques (VI.4) du corpus original et de ses différences avec le corpus relecture. Une dernière section présente les inter-corrélations significatives entre variables temporelles et mélodiques dans le corpus original (V.5).

Les caractéristiques prosodiques ont été analysées pour chaque variable à travers le calcul des valeurs moyennes et d'écart-type à l'échelle de chaque unité prosodique majeure et des différences de valeurs moyennes entre unités consécutives, en particulier les variations d'un bloc inter-pausal à l'autre au sein des extraits.

Les résultats de cette analyse sont les suivants :

- **La durée relative des pauses** par rapport à la durée totale des extraits est significativement supérieure à celle en relecture pour un locuteur (Serey) et significativement inférieure pour un locuteur (Le Chequer). La proportion totale de silence par extrait n'est supérieure aux caractéristiques en relecture que pour un locuteur, Molossi.

- **Les moyennes de débit articulatoire** sont significativement supérieures en condition originale pour Le Chequer et Tuillon. L'écart-type du débit articulatoire est supérieur à celui de la condition relecture pour Voynet et Serey. La durée moyenne des noyaux vocaliques de syllabes pré-pausales est significativement supérieure à la relecture pour Voynet et significativement inférieure pour Le Chequer et Tuillon.

- **Au niveau mélodique**, les caractéristiques principales des locuteurs en comparaison au corpus relecture sont une élévation de la F0 moyenne pour tous les locuteurs sauf Tuillon, une réduction de la plage de variation de F0 (trois locuteurs), et une plus forte variabilité de plage de variation d'un bloc inter-pausal à l'autre (deux locuteurs sur cinq).

- **Les corrélations** entre variables temporelles indiquent une corrélation négative significative entre débit articulatoire et variabilité de durée des syllabes (IVP) pour tous les locuteurs. Pour trois locuteurs sur cinq, l'accroissement du débit articulatoire est également corrélé à la réduction de la durée des noyaux vocaliques des syllabes finales pré-pausales. Au niveau mélodique, la F0 moyenne n'est significativement corrélée à la plage de variation que pour deux locuteurs sur cinq. Enfin, nous ne trouvons pas de corrélation significative entre F0 moyenne et débit articulatoire.

VI.1. Enjeux et principes méthodologiques pour l'analyse prosodique

À la suite de l'annotation prosodique effectuée, nous pouvons à présent effectuer des analyses prosodiques en prenant pour unité de référence le bloc inter-pausal¹. Nous présentons ici les caractéristiques prosodiques mélodiques et temporelles des locuteurs dans le corpus original, tout en les comparant à celles mesurées dans la condition de contrôle (corpus relecture). Nous présentons d'abord les caractéristiques temporelles des pauses silencieuses, puis les caractéristiques de débit articulatoire et des variations de durée, ainsi que les caractéristiques mélodiques principales. L'objectif des chapitres d'analyses **VI** à **VII** est double :

- Présenter les caractéristiques prosodiques générales des locuteurs dans le corpus original et pouvoir situer celles-ci par rapport à celles mesurées sur le corpus relecture.
- Constituer un jeu de variables prosodiques pour comparer celles-ci aux variables perceptibles dans le **Chapitre VIII** pour déterminer leurs corrélations avec les scores perceptifs issus du chapitre **IV**.

Du fait de ces deux objectifs, dans les trois chapitres qui vont suivre, les différentes caractéristiques prosodiques étudiées pourront être parfois présentées dans un premier temps selon leur configuration générale au sein du corpus, mais elles seront toujours ramenées à une moyenne par stimulus, afin de faire de cette caractéristique analysée une variable ensuite utilisable en comparaison avec les variables issues de la perception.

¹ Nous considérons ici comme bloc inter-pausal toute unité prosodique majeure ayant été annotée du fait de la présence d'une pause « frontières » (durée minimale 140 ms).

VI.1.1. Présentation de la méthodologie pour les statistiques

Outre l'analyse des pauses silencieuses, toutes les caractéristiques temporelles et mélodiques sont analysées au travers de trois différentes statistiques : la moyenne par stimulus, l'écart-type de la moyenne par stimulus et la différence moyenne d'un bloc inter-pausal à l'autre par stimulus² (différence inter-blocs).

- La **valeur moyenne**, calculée à l'échelle du bloc inter-pausal puis moyennée à l'échelle du stimulus.
- L'**écart-type** de la moyenne est également présenté à l'échelle du stimulus.
- La **différence inter-blocs**, correspondant à la mesure de différence de valeurs moyennes entre blocs inter-pausaux consécutifs. L'objectif de ce dernier calcul est de permettre une mesure des variations de caractéristiques prosodiques de part et d'autre des pauses silencieuses.

Pour comprendre le principe de cette dernière mesure, nous présentons le schéma ci-dessous qui prend pour exemple le calcul du débit articulaire à partir d'une paire d'extraits fictifs (original, relecture) chacun constitués de deux blocs inter-pausaux³.

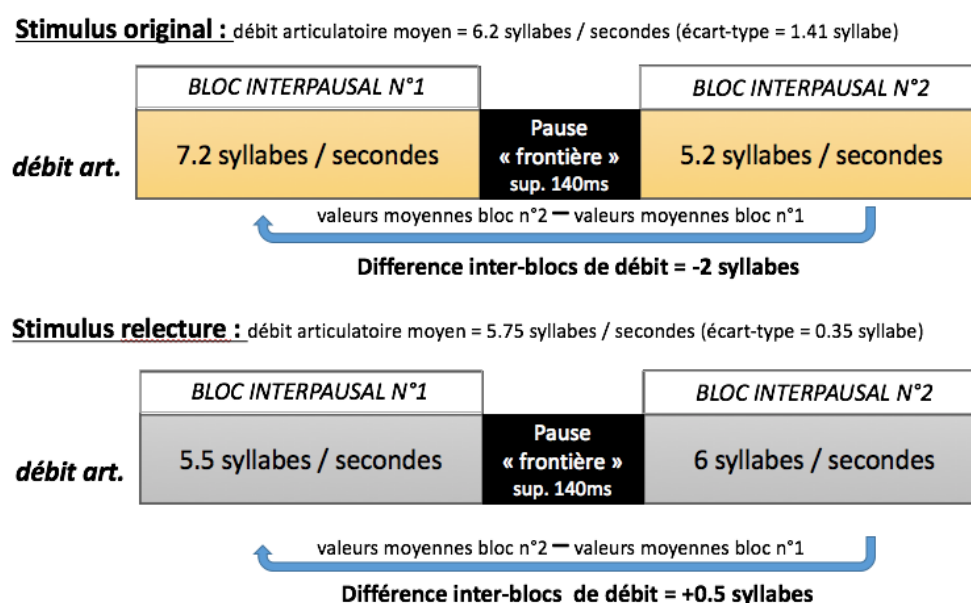


Figure VI-1 : Schéma explicatif de la mesure de différence inter-blocs sur une paire d'extraits fictifs constitués chacun de deux blocs inter-pausaux pour lesquels sont spécifiées les valeurs par bloc de débit articulaire (n syllabes / secondes).

² Ces trois différentes statistiques sont analysées au sein du corpus original, de la relecture, puis une mesure de différence original-relecture est enfin calculée en vue d'être utilisée comme variable supplémentaire dans les analyses comparatives entre perception et production.

³ Cette mesure va également être utilisée pour comparer de bloc à bloc des valeurs de variables mélodiques telles que la fréquence fondamentale (F0) moyenne ou sa plage de variation.

Dans la figure ci-dessus, on peut voir que si le débit moyen par stimulus est relativement proche d'une condition à l'autre (original = 6.2 syllabes par secondes, relecture = 5.75 syllabes par secondes), l'écart-type est plus réduit en relecture (0.35 syllabes contre 1.4 en original). Quant à la mesure de différence inter-blocs (original = -2, relecture = 0.5), elle indique un ralentissement du débit articulatoire post pausal moyen⁴ de 2 syllabes dans l'extrait original par rapport au bloc précédent. Cette mesure est également présentée lorsque pertinente, en valeur absolue : on parlera alors différence inter-blocs absolue, pour rendre compte de la variabilité générale entre blocs consécutifs (original = 2, relecture = 0.5) sans spécifier l'ordre des blocs du début à la fin de chaque stimulus. En effet, dans le cas de différences inter-blocs irrégulières (par exemple dans le cas du débit, d'accélération suivies de décélérations), ces différences dont la valeur sera alternativement positive et négative peuvent se compenser une fois moyennées à l'échelle d'un extrait complet. La mesure dite « absolue » de différence inter-blocs permet alors de rendre compte de cette irrégularité. Le lecteur pourra notamment se référer à la section VI.4.2.2 pour une analyse comparée de ces deux mesures (différence inter-blocs, différence inter-blocs absolue) sur la F0 moyenne.

⁴ Dans cet exemple fictif la moyenne peut aisément être déduite de la figure, car chaque extrait ne comporte que deux blocs inter-pausaux, mais il peut y en avoir jusqu'à 7 par stimulus dans le présent corpus, et leur nombre et durée peut bien sûr être amené à varier selon la condition (original, relecture).

VI.2. Analyses des pauses silencieuses

VI.2.1. Nombre et durées des pauses silencieuses

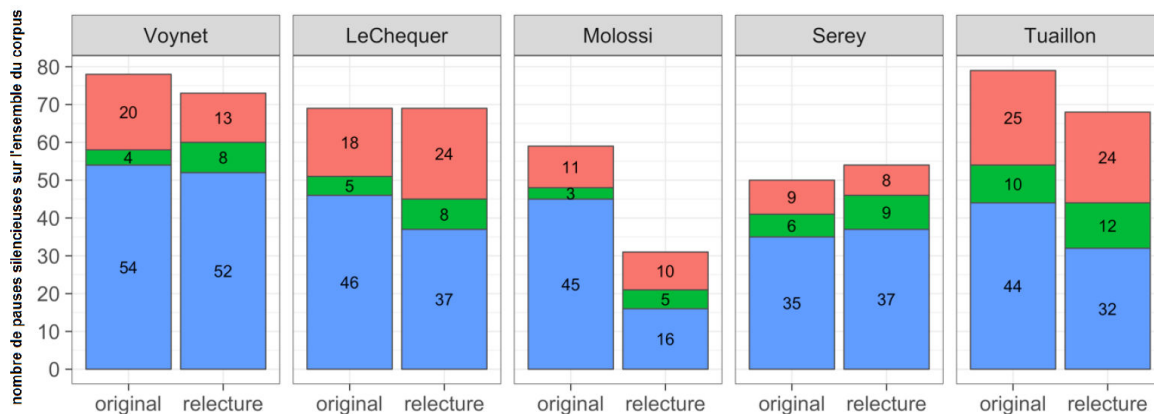


Figure VI-2 : nombre total de pauses silencieuses par locuteur selon la condition du corpus et leurs durées : rouge = pauses de durées inférieures à 140 millisecondes. Code couleur : vert = durées entre 140 et 200 millisecondes, bleu = durées supérieures à 200 millisecondes.

On compte dans le corpus original, tous locuteurs et toutes durées confondues, 335 pauses silencieuses contre 295 dans le corpus relecture. Au sein du corpus original, Voynet est la locutrice qui produit le plus de pauses (78, dont 54 pauses « frontières » de durée supérieure à 140 ms) et Serey celui qui en produit le moins (50, dont 41 pauses « frontières »). Le locuteur qui produit le plus de pauses de durée inférieure à 140 millisecondes est Tuaille (25). Ce locuteur est également celui pour lequel on mesure le débit articulaire le plus élevé⁵.

Tous locuteurs confondus, les durées relevées des pauses silencieuses sur l'ensemble du corpus vont de 64 à 1650 millisecondes ; voir Figure VI-3 page suivante. La durée moyenne des pauses silencieuses dans le corpus original va de 263 ms pour Tuaille (écart-type : 168 ms) à 431 pour Serey (écart-type : 313 ms), ce dernier étant celui pour lequel on trouve la durée la plus longue d'une pause silencieuse à 1650 ms.

Dans l'optique d'utiliser ces mesures en comparaison avec la perception dans le chapitre VIII, si l'on considère la durée moyenne par stimulus des pauses silencieuses, celle-ci va de 274 ms pour Tuaille à 496 ms pour Serey. Comme on peut le constater sur la figure ci-dessus, Serey est le seul locuteur pour lequel on calcule une différence significative entre conditions ($t[21]=3.1, p<.01, \eta^2=31\%$). L'ANOVA et les tests post-hoc effectués sur les valeurs issues du corpus original indiquent également que les durées moyennes des pauses silencieuses par

⁵ Voir section VI.5.1 pour une analyse détaillée des corrélations entre débit articulaire et nombre et durée des pauses silencieuses.

stimulus de Serey sont significativement supérieures à celles de tous les locuteurs sauf Molossi ($F[4,111]=7.68$, $p=0$; $\eta^2=22\%$, $p<.05$ pour toutes les comparaisons significatives).

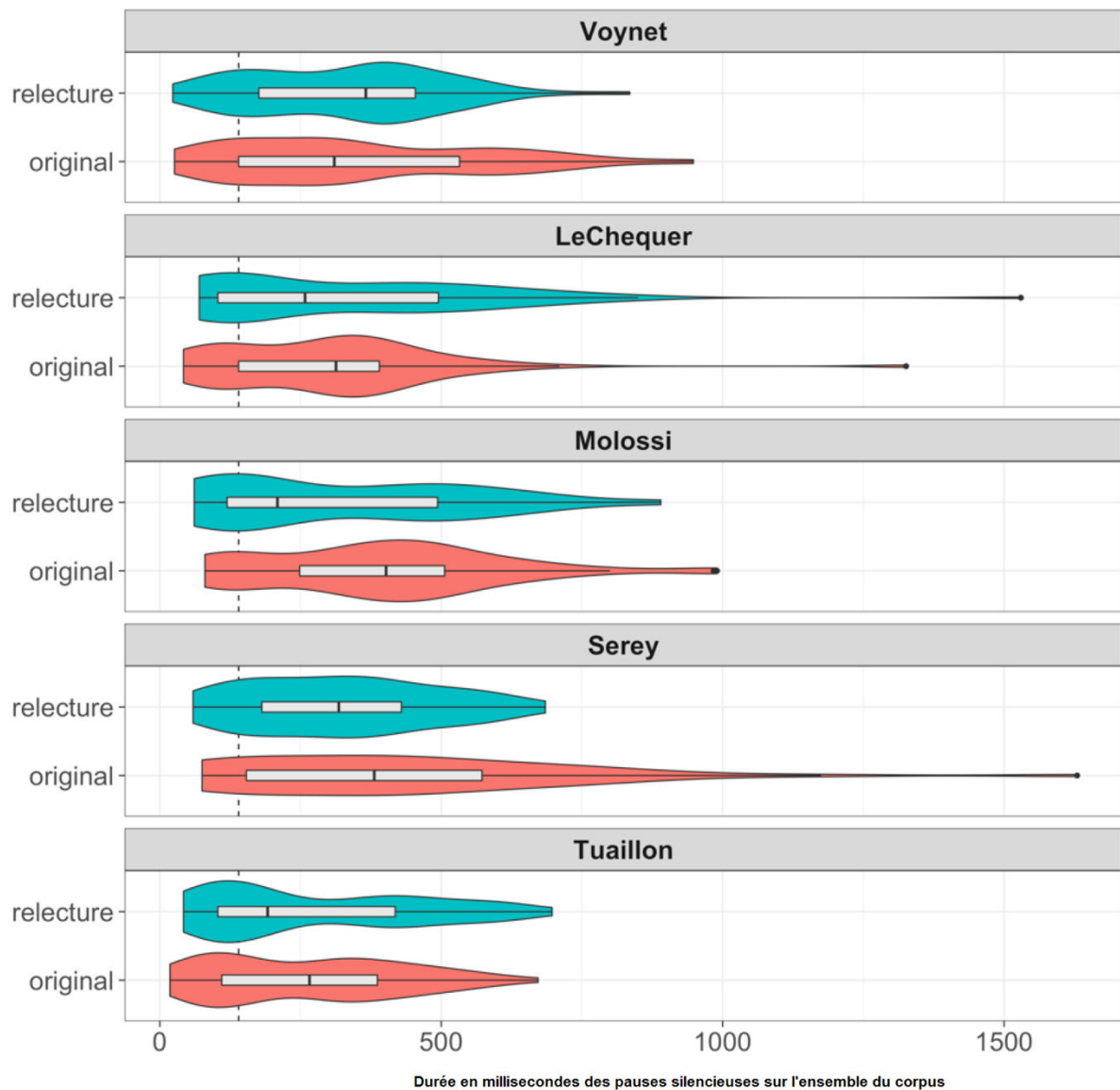


Figure VI-3 : Distribution des durées des pauses silencieuses sur l'ensemble du corpus par locuteur et par condition. La barre horizontale représente le seuil de durée de 200 millisecondes majoritairement employé dans les études pour considérer une pause silencieuse comme marqueur de frontière prosodique.

VI.2.2. Proportion de silence par stimulus

Nous avons calculé la proportion de silence par stimulus (durées cumulées des pauses / temps total de chaque extrait).

Dans le corpus original, le temps de silence représente en moyenne 10 à 14% de la durée des extraits sans différence significative entre locuteurs ($F[4,110]=1.45$, $p=0.22$; $\eta^2=5\%$). Molossi est le seul locuteur pour lequel on calcule une différence significative entre proportion de silence par stimulus en condition originale et en relecture ($t[21]=4.4$, $p<.001$, $\eta^2=48\%$).

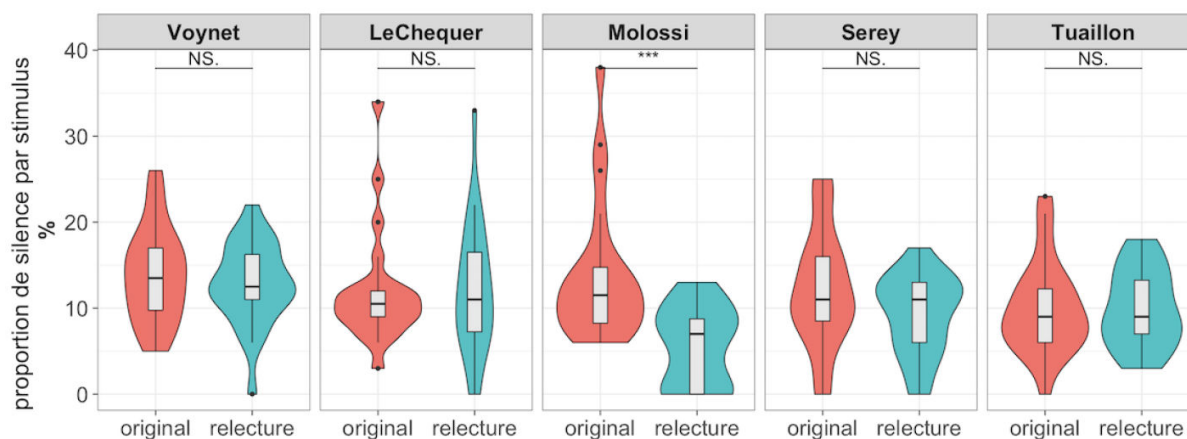


Figure VI-4 : Diagrammes en violons et boîtes à moustaches représentant la proportion de silence (durées cumulées des pauses / durée totale du stimulus) pour le corpus original et la relecture pour chaque locuteur. Les étoiles représentent le niveau de significativité obtenu en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p<.001$, ** = $p<.01$, * = $p<.05$, ns = $p>.05$.

Selon nous, la majeure partie des différences original-relecture significatives concernant les variables temporelles et mesures relatives aux pauses silencieuses de Molossi sont influencées par le fait que Molossi est le locuteur pour lequel on compte la plus forte différence du nombre de pauses silencieuses entre original et relecture. On compte au total 11 stimuli pour lesquels le temps de pause est équivalent à zéro, dont plus de la moitié est produite par Molossi en condition de lecture. En effet, Molossi produit 6 stimuli sur 25 sans aucun temps de pause en relecture.

VI.2.3. Durée relative des pauses silencieuses de type « frontière »

En plus de la mesure de proportion de silence présentée dans la section précédente 0, nous avons également calculé la durée relative moyenne (%) de chaque pause « frontières » une à une au sein des stimuli (durée pause / durée totale du stimulus).

Dans le corpus original, les pauses « frontières » représentent en moyenne 4 à 7% de la durée des extraits dans lesquels elles se trouvent. Dans le corpus original, le locuteur pour lequel on

trouve les pauses silencieuses « frontières » ayant la durée relative moyenne la plus longue est Serey, avec une durée moyenne relative de quasiment 8% et des durées maximales allant jusqu'à 20%. L'ANOVA effectuée sur le corpus original indique un effet significatif du locuteur ($F[4,105]=3.61$, $p<.01$) : la durée relative moyenne des pauses de Serey est significativement supérieure à celle de Tuailleon, Le Chequer et Voynet ($p<.001$ pour toutes comparaisons post-hoc). De plus, comme on peut l'observer sur la figure ci-dessous, la durée relative moyenne des pauses de Serey dans le corpus original est significativement supérieure à celle des pauses qu'il produit dans la contrepartie lue de ses stimuli ($t[20]=3.9$, $p<.001$, $\eta^2=43\%$).

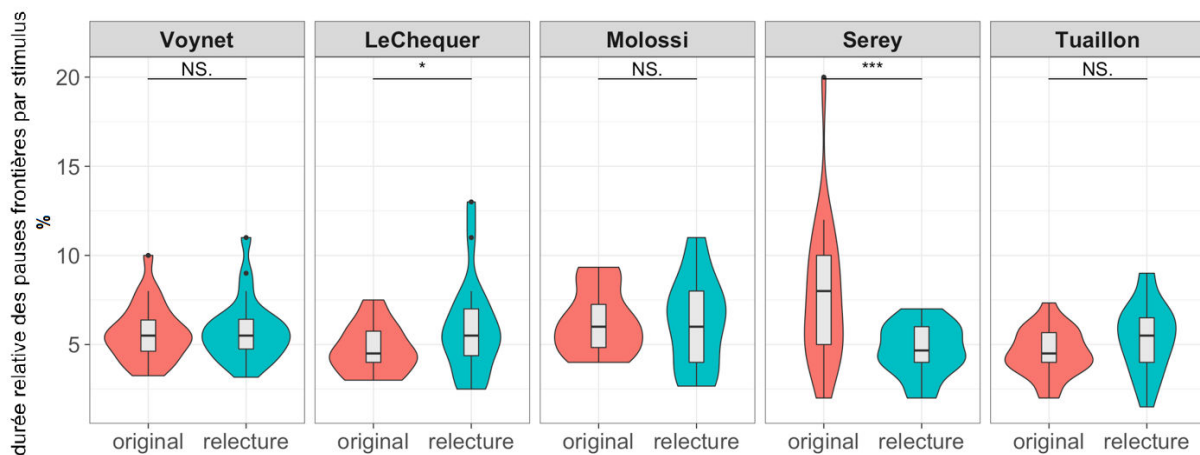


Figure VI-5 : Diagrammes en violons et boîtes à moustaches représentant la durée relative en pourcentage des pauses silencieuses « frontières » par rapport à la durée totale des stimuli selon la condition du corpus. Les étoiles représentent le niveau de significativité obtenu en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p<.001$, ** = $p<.01$, * = $p<.05$, ns = $p>.05$.

On trouve une différence significative entre conditions également pour Le Chequer ($t[18]=-2.2$, $p<.05$, $\eta^2=21\%$), bien que pour celui-ci la tendance soit inverse : la durée relative moyenne de ses pauses étant plus basse dans la version originale de ses stimuli qu'en relecture (original : 4.71%, écart-type : 1.35% ; relecture : 5.89%, écart-type : 2.70%). Le Chequer fait partie des locuteurs pour lesquels la durée relative moyenne des pauses silencieuses dans le corpus original est la plus basse au même titre que Tuailleon (moyenne : 4.57%, écart-type : 1.31%).

VI.3. Analyses du débit et du rythme syllabique

VI.3.1. Débit articuloire moyen (syllabes)

Les valeurs de débit articuloire moyen calculé au sein du corpus original vont de 5 syllabes par secondes (écart-type : 0.47 syllabes) pour Serey à 6 syllabes par secondes (écart-type : 0.75 syllabes) pour Tuailon. Le débit articuloire de Tuailon est significativement supérieur à celui de Serey et Molossi (ANOVA : $F[4,108]=3.39$, $p=0.01$; $\eta^2=11\%$; $p<.05$ pour les deux comparaisons).

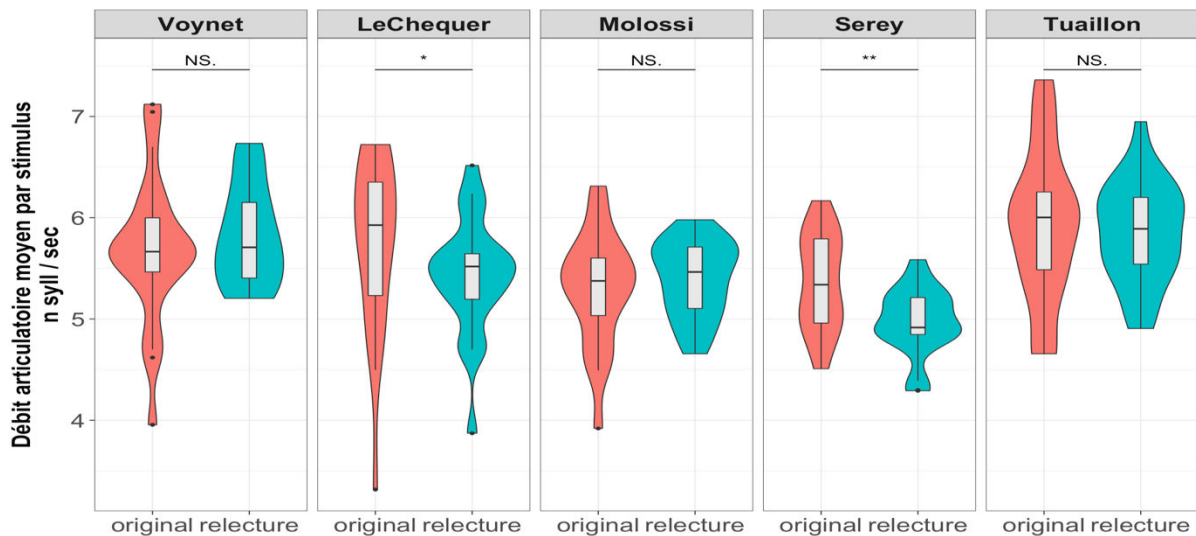


Figure VI-6 : Diagrammes en violons et boîtes à moustaches représentant le débit articuloire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur. Les étoiles correspondent au degré de significativité des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif)

Le débit articuloire moyen entre corpus original et relecture ne diffère que pour Serey ($t[21]=3$, $p<.01$, $\eta^2=30\%$) et Le Chequer ($t[21]=2.6$, $p<.01$, $\eta^2=24\%$), pour lesquels on trouve un débit articuloire moyen plus rapide en condition originale qu'en condition de relecture.

Au sein du corpus relecture, on note également des différences de débit articuloire selon le locuteur ($F[4,111]=14.61$, $p<.001$; $\eta^2=34\%$) : Serey a un débit articuloire moyen plus bas que tous les autres locuteurs ($p<.05$ pour toutes les comparaisons post-hoc). Inversement, Voynet et Tuailon ont produit leurs relectures avec un débit articuloire significativement plus élevé que tous les autres locuteurs ($p<.05$ pour toutes les comparaisons).

VI.3.2. Variations de débit articulaire

Nous présentons à présent un ensemble de mesures de variation du débit articulaire : l'écart-type, l'index de variabilité paire à paire (IVP) de la durée syllabique, la variation moyenne de débit articulaire d'un bloc inter-pausal à l'autre, la durée des voyelles finales de blocs. Pour l'ensemble de ces mesures, nous renvoyons à la section VI.5.1 pour une analyse des corrélations avec le débit articulaire moyen.

VI.3.2.1. Écart-type du débit articulaire

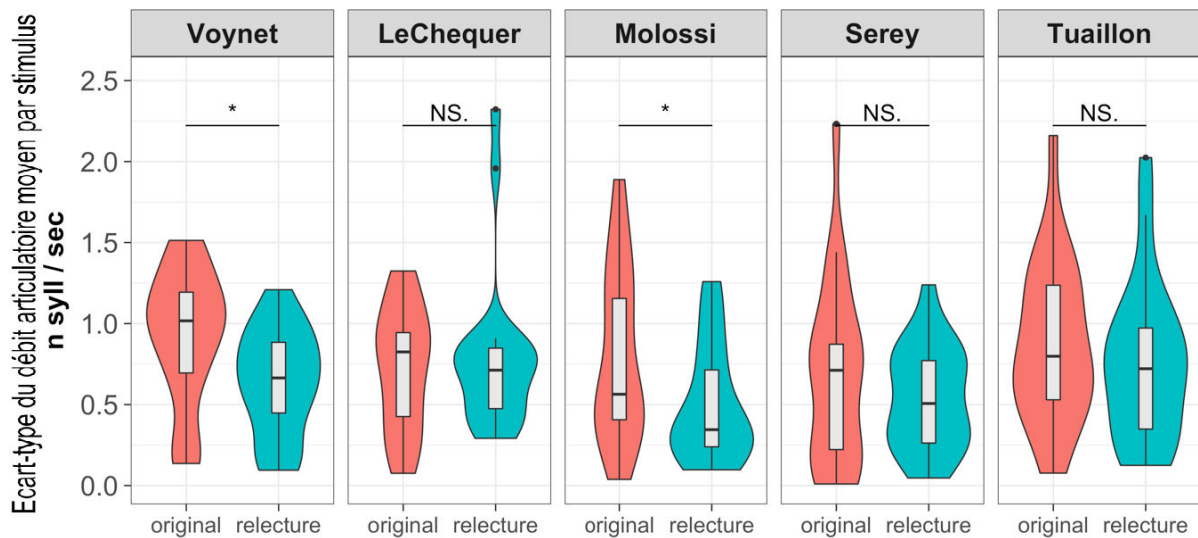


Figure VI-7 : Diagrammes en violons et boîtes à moustaches représentant l'écart-type du débit articulaire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur. Les étoiles correspondent au degré de significativité des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif)

Les valeurs moyennes d'écart-type du débit articulaire dans le corpus original vont de 0.72 syllabes pour Le Chequer à 0.92 pour Voynet. On ne calcule pas au sein du corpus original de différences significatives d'un locuteur à un autre.

Si tous locuteurs confondus on trouve une différence significative d'écart-type du débit articulaire entre conditions ($t[94]=2.8$, $p<.01$, $\eta^2=8\%$), la différence n'est significative que pour deux locuteurs sur cinq, Molossi ($t[13]=2.6$, $p<.05$, $\eta^2=34\%$) et Voynet ($t[22]=2.7$, $p<.01$, $\eta^2=24\%$), pour qui l'écart-type du débit articulaire est plus élevé dans le corpus original qu'en condition de relecture.

VI.3.2.2. Index de variabilité paire à paire de la durée syllabique (IVP)

L'index de variabilité paire à paire (IVP ou en anglais PVI pour « Pairwise variability index ») est une mesure mise en place par Grabe & Low (2002) pour fournir une mesure plus fine des variations locales de débit articulaire de durée syllabique. Plus la mesure d'IVP de durée syllabique est basse, moins il y a de variation de durée d'une syllabe à une autre. Inversement, plus les valeurs de l'IVP sont hautes, plus elles sont relatives à de fortes et fréquentes différences de durée d'une syllabe à l'autre.

La moyenne de l'IVP de durée syllabique tous locuteurs confondus est la même dans les deux conditions du corpus (original : 0.9, écart-type : 0.03 ; relecture : 0.9, écart-type : 0.03).

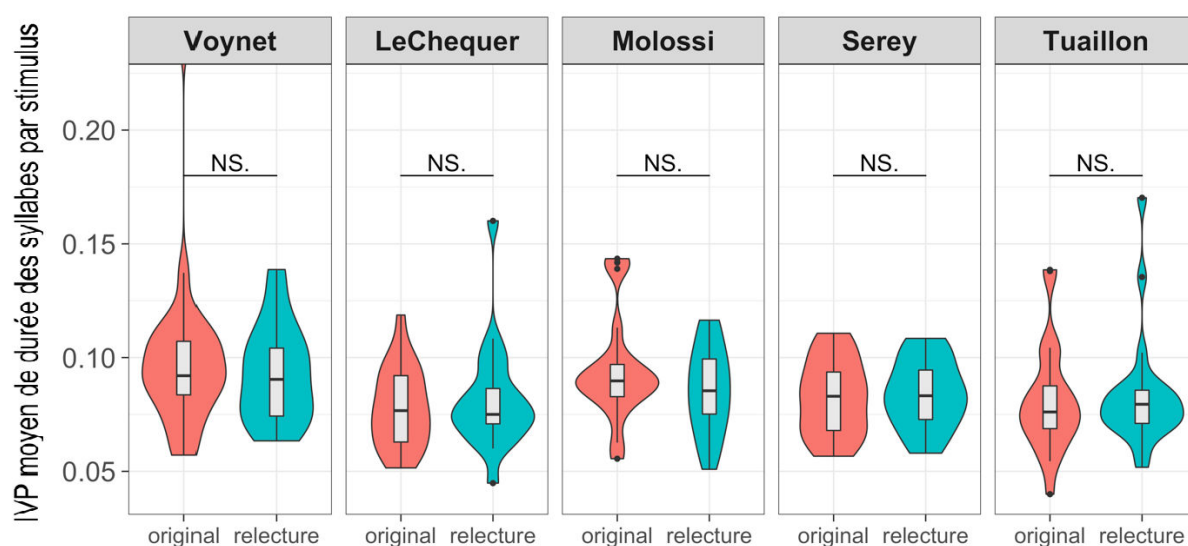


Figure VI-8 : Diagrammes en violons et boîtes à moustaches de l'écart-type du débit articulaire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif).

Dans le corpus original, Voynet est la locutrice pour laquelle on trouve les valeurs moyennes de l'IVP de durée syllabique les plus hautes (moyenne : 1.0, écart-type : 0.04) et Le Chequer et Tuailleon les plus basses avec pour tous les deux un IVP moyen de 0.8 (écart-type : 0.02).

On trouve une influence significative du facteur locuteur sur les mesures ($F[4,102]=3.17$, $p<.05$; $\eta^2=11\%$). L'IVP de durée syllabique est significativement plus faible pour Le Chequer (moyenne : 0.8, écart-type : 0.02) que pour Voynet en condition audio ($p<.05$).

VI.3.2.3. Différence inter-blocs du débit articulaire moyen

Nous présentons ici une analyse des variations de débit articulaire de part et d'autre des pauses silencieuses, soit d'un bloc inter-pausal à l'autre. Cette mesure peut être considérée comme analogue à l'IVP de durée syllabique, sauf que dans le cas présent, nous mesurons la variabilité de débit articulaire de bloc inter-pausal à bloc inter-pausal, et donc de part et d'autre des pauses silencieuses ; voir pour rappel section VI.1.1 pour une description avec exemple des principes de cette mesure.

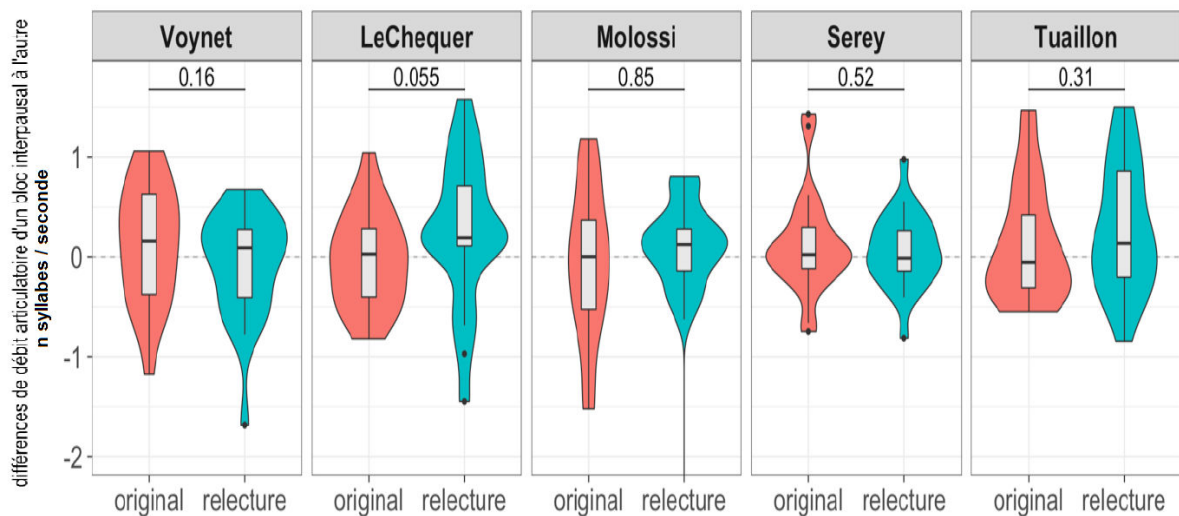


Figure VI-9 : Différences moyenne de débit articulaire (n syll / secondes) d'un bloc inter-pausal à l'autre pour chaque locuteur selon la condition. L'axe vertical représente la différence de débit articulaire en nombre de syllabe par secondes d'un bloc inter-pausal à l'autre. Les valeurs correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus.

On ne calcule aucune différence inter-conditions ni inter-locuteurs sur cette mesure.

VI.3.2.4. Durées des voyelles des syllabes finales pré-pausales

Nous avons calculé la durée des voyelles de syllabes finales de blocs inter-pausaux. Dans le corpus original, la durée moyenne des voyelles de syllabes finales va de 111 millisecondes (écart-type : 23 ms) pour Tuailleon à 163 millisecondes pour Voynet (écart-type : 39 ms), et 166 millisecondes pour Molossi (écart-type : 33 ms). L'ANOVA effectuée sur ces durées originales indique une influence significative du facteur locuteur ($F[4,104]=10.2$, $p<.01$; $\eta^2=28\%$) : les valeurs moyennes de Voynet et de Molossi sont significativement supérieures à celles Le Chequer, Serey et Tuailleon ($p<.001$ pour toutes les comparaisons). On ne trouve pas de différence significative de durée entre ces trois derniers.

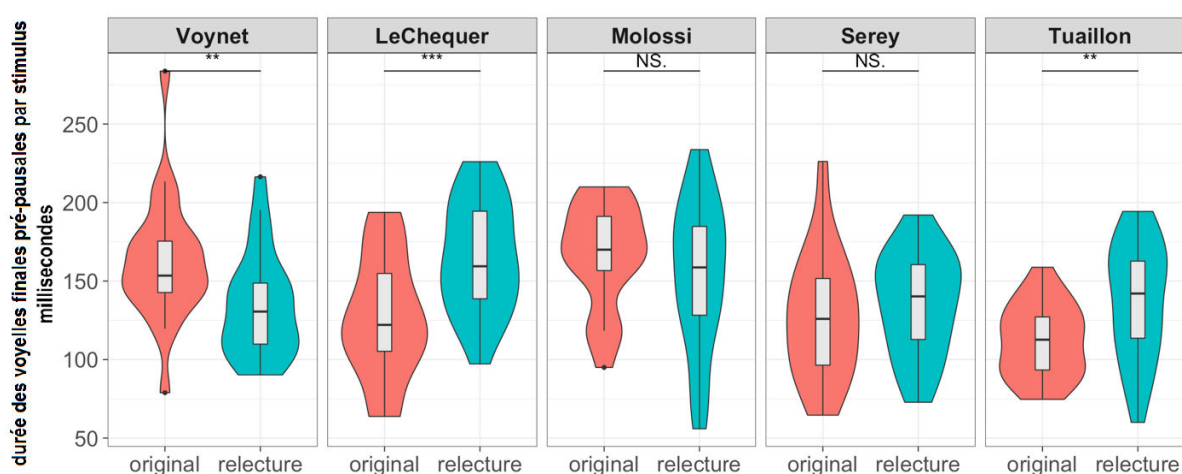


Figure VI-10 : Diagrammes en violon et boîtes à moustache représentant en millisecondes les valeurs de durée des voyelles situées dans les syllabes finales pré-pausales en condition originale et en relecture pour chaque locuteur. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre stimuli originaux et stimuli relus (* = $p<.05$, ** = $p<.01$, *** = $p<.001$, NS. = non significatif)

Comme on peut l'observer sur la figure ci-dessus, on calcule une différence significative de la durée des voyelles noyaux de syllabes finales pré-pausales pour trois locuteurs sur cinq : Le Chequer ($t[20]=-4.8$, $p<.001$, $\eta^2=54\%$), Tuailleon ($t[23]=-3.2$, $p<.01$, $\eta^2=31\%$) et Voynet ($t[22]=2.9$, $p<.01$, $\eta^2=28\%$). Pour Voynet, les voyelles de syllabes finales sont en moyenne plus longues dans le corpus original qu'en relecture, où la moyenne des mesures passe à 130 ms (écart-type : 33 ms).

Pour Le Chequer et Tuailleon, locuteurs pour lesquels on trouve en moyenne les durées les plus courtes de syllabes finales dans le corpus original, c'est le contraire : les durées moyennes des voyelles de syllabes finales pré-pausales sont significativement plus longues en condition de relecture que dans le corpus original.

VI.3.2.5. Allongement syllabique relatif final

Nous présentons dans cette partie les résultats de la mesure de l'allongement syllabique relatif final tel qu'il a été utilisé pour déterminer les frontières prosodiques majeures marquées par le paramètre de durée. Pour rappel le calcul de l'allongement syllabique correspond à la comparaison entre la durée de la syllabe analysée et la moyenne des deux syllabes précédentes, ou des deux syllabes précédentes et de la syllabe suivante, lorsque la syllabe analysée n'est pas suivie d'une pause. A partir de ce calcul, durant l'annotation des frontières prosodiques majeures, selon les critères de Mertens et Simon (2013), nous n'avons considéré le marquage des frontières par la durée syllabique que les cas d'allongement relatif supérieur ou égal à 3.

Lorsque l'on observe l'ensemble des valeurs d'allongement syllabique relatif moyen à l'échelle du stimulus, on constate que la moyenne des valeurs est fortement inférieure au seuil de 3 correspondant à une frontière prosodique marquée par la durée, illustrée notamment par la ligne en pointillés dans la figure ci-dessous.

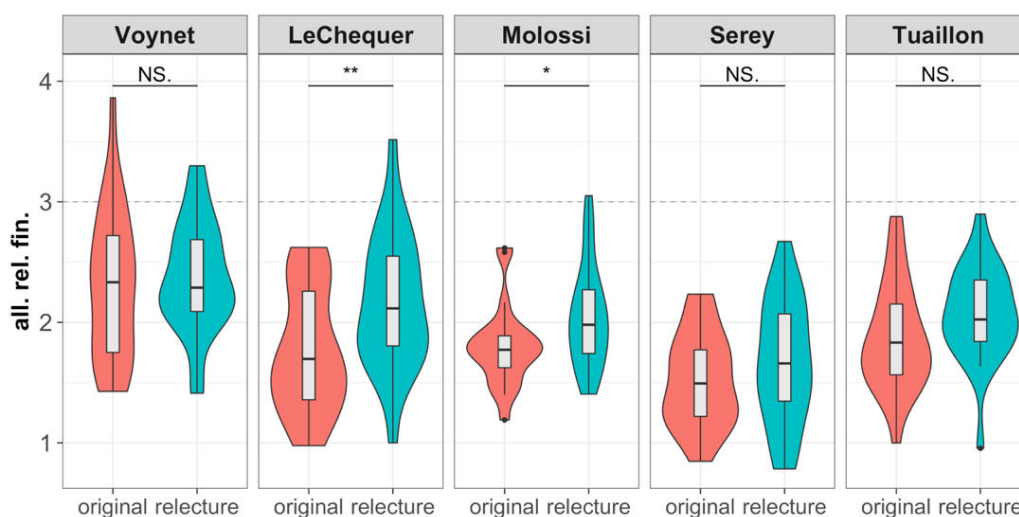


Figure VI-11 : Valeurs moyennes par stimulus de l'allongement syllabique relatif final d'unité prosodique majeure par locuteur selon la condition (all. rel. fin. : durée syllabe / moyenne syllabes avoisinantes). La ligne horizontale en pointillés représente le seuil d'allongement relatif final de 3 permettant l'annotation de frontières prosodiques majeures. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (* = $p < .05$, ** = $p < .01$, *** = $p < .001$, NS. = non significatif)

Tous locuteurs confondus, on calcule une différence significative d'allongement relatif final moyen par stimulus entre les conditions ($t[116] = -4$, $p < .001$, $\eta^2 = 12\%$): celui-ci est en moyenne plus élevé en condition de relecture. Si cette tendance s'observe pour tous les locuteurs, la différence n'est toutefois significative que pour deux locuteurs sur cinq : Le Chequer ($t[22] = -1.3$, $p = 0.2$, $\eta^2 = 7\%$) et Serey ($t[21] = 2.4$, $p < .05$, $\eta^2 = 22\%$).

VI.4. Analyses de la fréquence fondamentale (F0)

VI.4.1. Fréquence fondamentale (F0) moyenne

Les valeurs de fréquence fondamentale moyenne (F0) présentées dans la figure ci-dessous ont été calculées en faisant la moyenne par bloc inter-pausal puis par stimulus des mesures de F0 effectuées dans le tiers milieu des voyelles sur l'ensemble du corpus.

En condition de lecture calme, la F0 moyenne de Voynet, seule femme parmi les locuteurs étudiés, est de 244 Hertz (écart-type : 32) et celle des locuteurs masculins va de 112 Hz pour Le Chequer à 130 Hz pour Serey. L'ANOVA et les tests post-hoc effectués sur les valeurs moyennes en condition de relecture indiquent que la F0 de Voynet est significativement plus haute que celle de ses interlocuteurs masculins ($p < .001$ pour tous les comparaisons post-hoc). Les tests post-hoc indiquent également que la F0 moyenne de Le Chequer est significativement inférieure à celle de Serey et Tuailleon ($p < .001$ pour les deux comparaisons), Le Chequer étant le locuteur pour lequel on mesure la F0 moyenne la plus basse dans les deux conditions du corpus.

Dans le corpus original la fréquence fondamentale (F0) moyenne par locuteur est significativement plus haute dans les extraits originaux que dans la relecture ($t[115]=8.6$, $p < .001$, $\eta^2=39\%$).

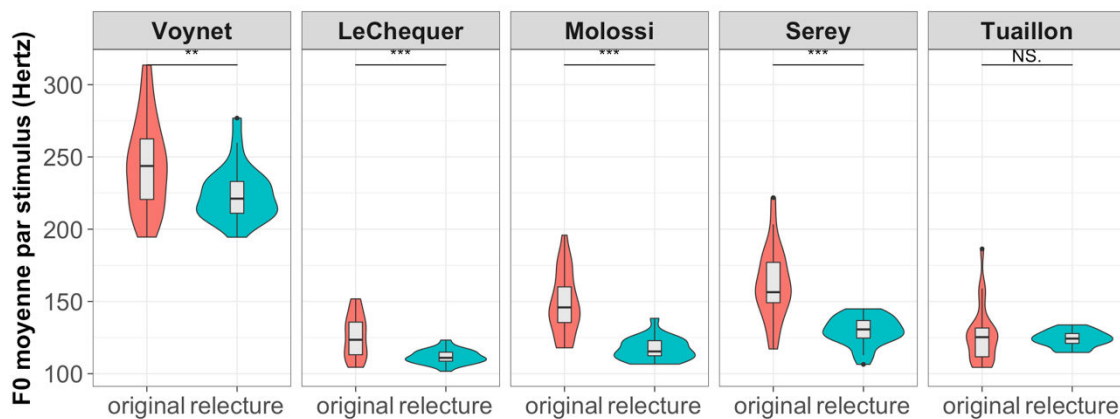


Figure VI-12 : Diagrammes en violon et boîtes à moustaches représentant la F0 moyenne (Hertz) par stimulus selon le locuteur et la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (* = p < .05, ** = p < .01, *** = p < .001, NS. = non significatif)

Tuailleon est le seul locuteur pour lequel on ne trouve pas de différence significative entre les deux conditions ($t[23]=0.6$, $p=0.6$, $\eta^2=2\%$). Sa F0 moyenne est de 127 Hertz en condition originale contre 124 Hertz en relecture, soit une différence de seulement 0.19 demi-tons entre les deux conditions ; voir détails dans le tableau ci-dessous.

	ORIGINAL (Hz)	RELECTURE (Hz)	DIFFÉRENCE (demi-tons)
VOYNET	244 (32)	225 (19)	1.36 (2.13)
LE CHEQUER	126 (15)	112 (5)	1.95 (2.13)
MOLOSSI	149 (22)	117 (8)	4 (2.43)
SEREY	162 (25)	130 (10)	3.71 (2.48)
TUAILLON	127 (20)	124 (5)	0.19 (2.65)

Tableau VI—1 Moyenne de la fréquence fondamentale (F0) moyenne en Hertz dans le corpus original et la relecture, et moyenne des mesures de différence en demi-tons. L'écart-type est présenté entre parenthèses.

Dans le corpus original, l'influence du facteur locuteur est significative ($F[4,111]=103.93$, $p<.01$; $\eta^2=79\%$). On mesure pour Voynet, seule locutrice du corpus sinon constitué de locuteurs, une F0 moyenne de 244 Hertz (écart-type : 32 Hz). En condition de lecture, Le Chequer est le locuteur pour lequel on calcule la F0 moyenne la plus basse (112 Hz, écart-type : 15 Hz), significativement plus basse que celle de Tuailon (124Hz, écart-type : 4 Hz) et Serey (165, écart-type : 20 Hz) ; $p<.001$ pour les deux comparaisons. Dans le corpus original, en condition de débat, Le Chequer est toujours le locuteur pour lequel on mesure la F0 moyenne la plus basse (126 Hz), mais cette fois-ci il n'y pas de différence significative avec la F0 moyenne de Tuailon. Dans le corpus original, Le Chequer et Tuailon ont une F0 moyenne significativement plus basse que celle de Molossi et Serey ($p<.001$ pour toutes les comparaisons).

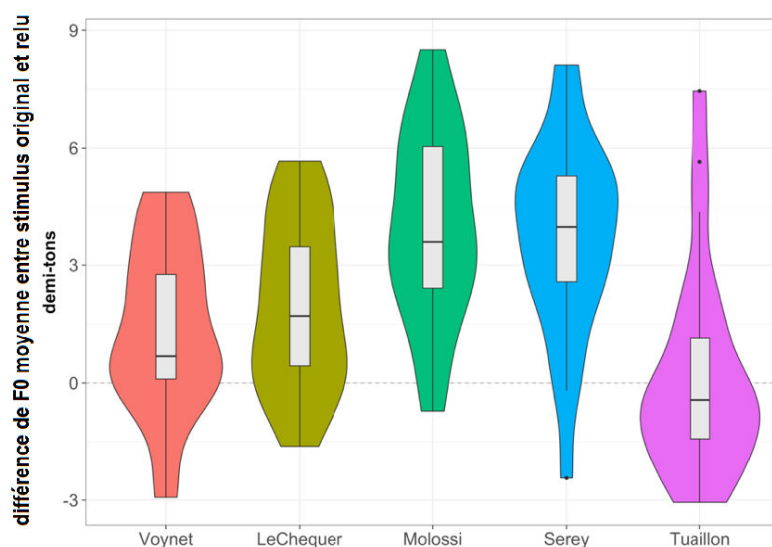


Figure VI-13 : Diagrammes en violon et boîtes à moustache des mesures de différence original-relecture de F0 moyenne par stimulus présentées en demi-tons pour chaque locuteur

Comme la figure ci-dessus permet de le visualiser, les différences les plus marquées de F0 moyenne entre stimuli originaux et relus sont trouvées pour Molossi et Serey. On notera également que la médiane (trait vertical dans la boîte à moustache) des valeurs de F0 moyenne de Tuailon se trouve sous zéro, ainsi que le mode de sa distribution, ce qui signifie que pour ce locuteur, la majorité de ses stimuli relus sont produits en condition relecture avec une F0 moyenne inférieure au corpus original.

L'ANOVA effectuée sur les mesures de différences en demi-tons confirme une influence significative du locuteur ($F[4,111]=10.61$, $p<.01$; $\eta^2=28\%$) et les tests post-hoc indiquent que les mesures de différence de F0 entre stimuli originaux et relus de Molossi sont significativement supérieures à celles de Voynet ($p<.001$), Tuailon ($p<.01$) et Le Chequer ($p<.05$). Les mesures de différence de F0 moyenne de Serey entre original et relecture sont quant à elles supérieures à celles de Tuailon ($p<.001$) et Voynet ($p<.01$).

VI.4.2. Variations de la F0 moyenne

Pour étudier la variation de la F0 moyenne nous présentons dans un premier temps les valeurs d'écart-type par stimulus, puis les valeurs de variation de F0 moyenne d'un bloc inter-pausal à l'autre.

VI.4.2.1. Écart-type de F0 en demi-tons

Les valeurs d'écart-type de F0, comme toutes mesures de variation de F0, sont calculées en demi-tons pour les rendre directement comparables entre locuteurs.

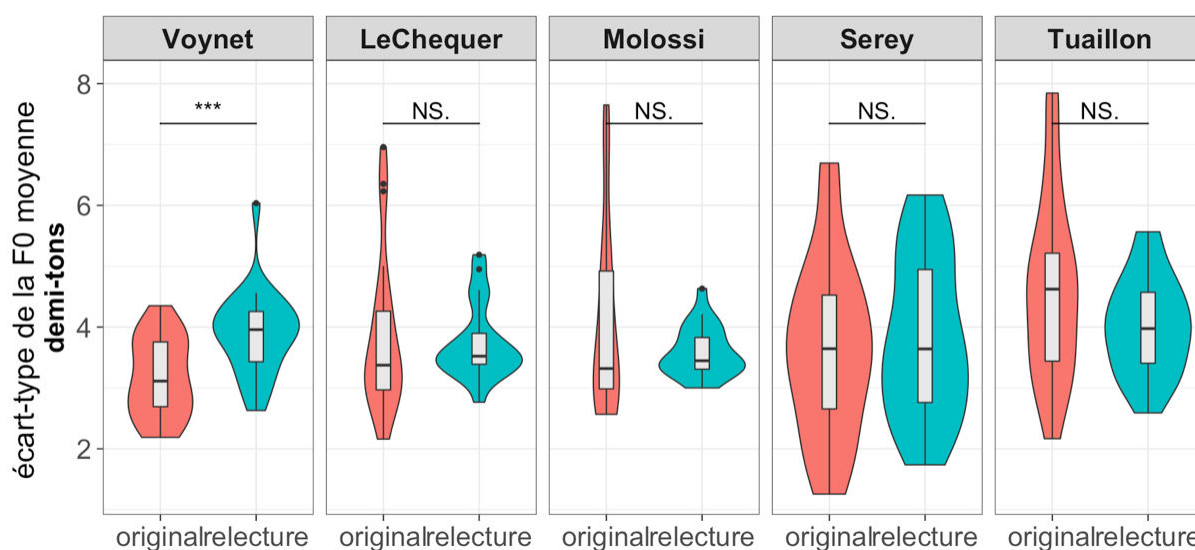


Figure VI-14 : Diagrammes en violon et boîtes à moustaches représentant l'écart-type de la F0 moyenne en demi-tons par stimulus selon le locuteur et la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif)

Comme on peut l'observer sur la figure ci-dessus, Voynet est la seule locutrice pour laquelle on trouve une différence significative entre conditions : l'écart-type de ses valeurs de F0 moyenne est significativement plus réduit dans les stimuli originaux que dans leur contrepartie relue ($t[23]=-3.9$, $p<.001$, $\eta^2=40\%$).

Au sein du corpus original, l'écart-type moyen de F0 va de 3,2 demi-tons pour Voynet à 4,62 demi-tons pour Tuailleon. L'ANOVA effectuée indique une influence significative du facteur locuteur sur les valeurs d'écart-type de F0 ($F[4,111]=4.03$, $p<.01$; $\eta^2=13\%$) et les tests post-hoc montrent que l'écart-type moyen de Tuailleon est significativement plus réduit que celui de Voynet ($p<.001$).

VI.4.2.2. Différence inter-blocs de F0 moyenne

Nous avons mesuré les variations de F0 moyenne d'un bloc inter-pausal à l'autre et nous présentons ici les valeurs moyennes absolues puis brutes par stimulus ; voir pour rappel la description et un exemple de ces mesures en page 110.

Nous présentons dans un premier temps les valeurs absolues de différence de F0 moyenne de bloc à bloc, mesure comparable à l'écart-type, puis les mesures brutes de manière à pouvoir étudier d'éventuelles dynamiques d'un bloc à l'autre, sur le déroulé du stimulus. Cette mesure s'était avérée pertinente dans l'étude du débit articulaire parce qu'elle nous permettait de confronter les observations de Ferré (2004) par rapport à nos données, de regarder si de part et d'autre des pauses il y avait des patrons spécifiques de contraste de débit articulaire. Dans le cas de la F0 moyenne, étudier la variation d'un bloc à l'autre sur ce genre de paramètre peut permettre d'avoir une visualisation générale de l'évolution de la hauteur moyenne de F0 au sein du stimulus. En parole lue non expressive on s'attend à un abaissement graduel de la F0 moyenne au sein des phrases, traduit par des valeurs de différences inter-blocs négatives.

La variation absolue de F0 moyenne inter-blocs est en moyenne tous locuteurs confondus de 0.9 demi-tons dans les deux conditions du corpus. Quelle que soit la condition du corpus on ne trouve pas de différence significative inter-locuteurs. La variation absolue de F0 moyenne inter-blocs ne diffère significativement entre conditions que pour un locuteur, Serey, pour qui on trouve une variation plus réduite en condition de lecture qu'en condition originale ($t[22]=2.4$, $p<.05$, $\eta^2=20\%$).

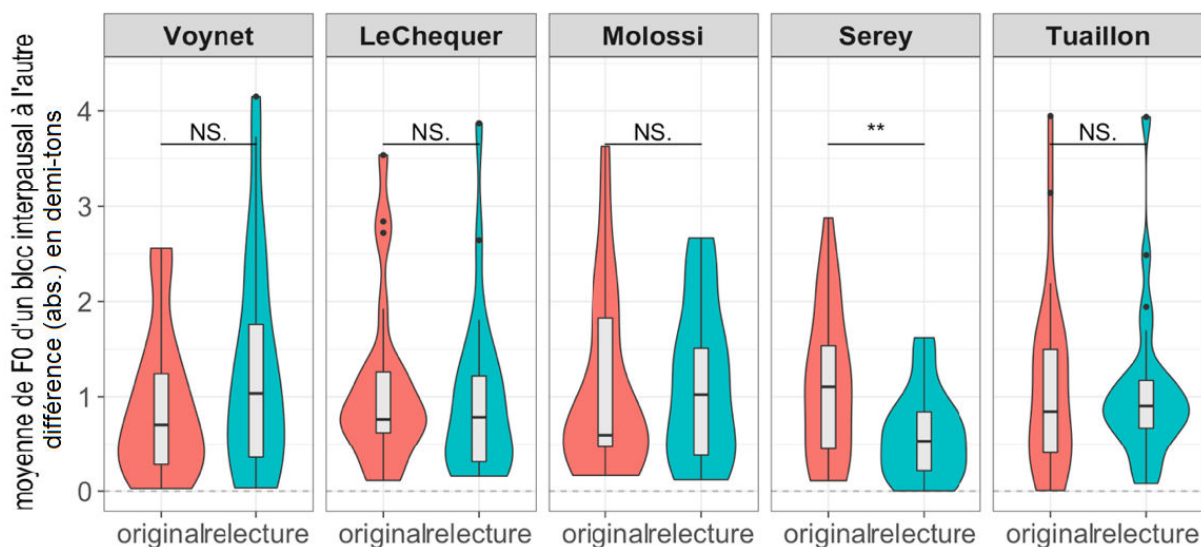


Figure VI-15 : Diagrammes en violon et boîtes à moustaches représentant les différences absolues de F0 moyenne d'un bloc inter-pausal à l'autre (demi-tons) par stimulus selon le locuteur et la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif)

La comparaison des valeurs brutes de variation de F0 moyenne d'un bloc inter-pausal à l'autre tous locuteurs confondus indique une différence significative ($t[116]=3.5$, $p<.001$, $\eta^2=9\%$) : les différences bloc à bloc dans le corpus relecture sont légèrement négatives (moyenne générale : -0.70 , écart-type : 1.13) alors que celles-ci sont quasiment nulles dans le corpus original (moyenne générale : -0.70 , écart-type : 1.13). Les valeurs négatives au sein du corpus relecture peuvent être interprétées comme une indication que la F0 moyenne baisse d'un bloc inter-pausal à l'autre. Dans le corpus relecture, on constate pour tous les locuteurs sauf Serey, un abaissement de la F0 moyenne d'un bloc à inter-pausal à l'autre.

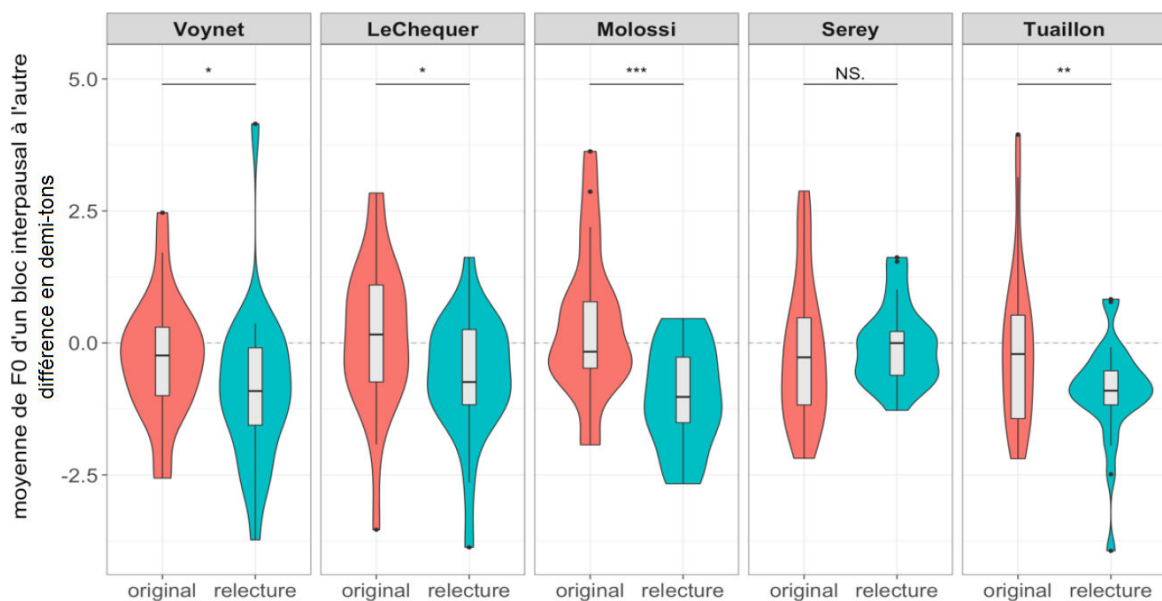


Figure VI-16 : Diagrammes en violon et boîtes à moustaches représentant les différences brutes de F0 moyenne d'un bloc inter-pausal à l'autre (demi-tons) par stimulus selon le locuteur et la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= p<.05, ** = p<.01, *** = p<.001, NS. = non significatif)

Comme on peut l'observer sur la figure ci-dessus, la différence est significative pour tous les locuteurs sauf pour Serey ($t[22]=-0.4$, $p=0.7$, $\eta^2=1\%$). Serey est le seul locuteur pour lequel la variation moyenne de plage de F0 est en moyenne nulle, et non légèrement négative.

VI.4.3. Plage de variation de la fréquence fondamentale (F0 min – F0 max)

Nous présentons ici les valeurs moyennes par stimulus de la plage de variation de F0 qui correspond à l'étendue de ses valeurs. Comme pour les autres variables présentées ici, nous avons moyenné la plage de F0 par bloc inter-pausal à l'échelle du stimulus.

En condition de relecture, la plage de variation de F0 moyenne par stimulus tous locuteurs confondus est de 19 demi-tons. Les valeurs moyennes par locuteur vont de 15 demi-tons (écart-type : 2.6 demi-tons) pour Molossi à 22 demi-tons (écart-type : 6.6 demi-tons) pour Tuailleon. L'ANOVA effectuée sur les valeurs issues des extraits lus (corpus relecture) indique une influence du locuteur ($F[4,111]=5.31, p<.01; \eta^2=16\%$) : la plage de variation de Molossi est significativement plus réduite que celle de Serey et Tuailleon ($p<.001$ pour toutes les comparaisons post-hoc).

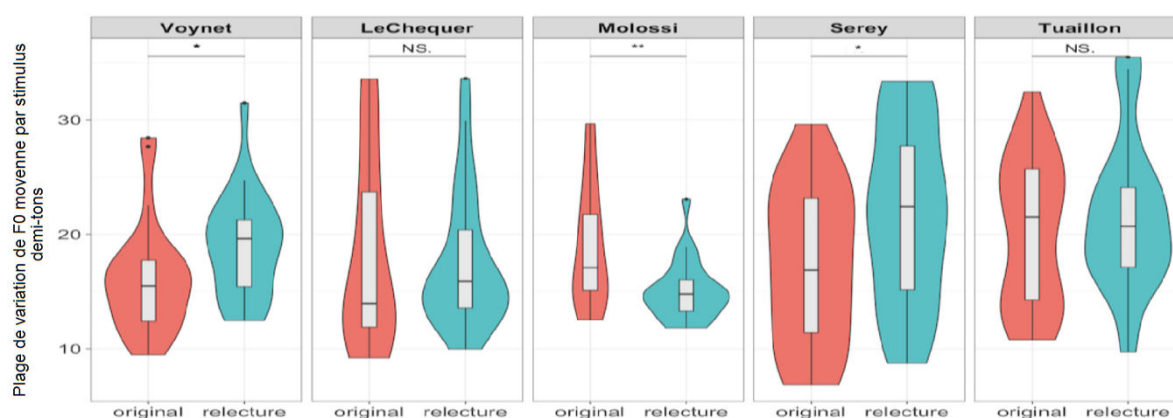


Figure VI-17 : Diagrammes en violons et boîtes à moustaches représentant la plage de variation moyenne par locuteur selon la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (*= $p<.05$, ** = $p<.01$, *** = $p<.001$, NS. = non significatif)

Pour Voynet ($t[23]=-2.3, p<.05, \eta^2=19\%$) et Serey ($t[22]=-2, p<.05, \eta^2=16\%$), la plage de F0 est en moyenne plus réduite dans les stimuli originaux que dans les stimuli relus. Pour Molossi c'est le contraire ($t[21]=3.4, p<.01, \eta^2=36\%$). Molossi est le seul locuteur qui produit ses extraits avec une plage de variation de F0 plus étendue en condition originale qu'en relecture.

VI.4.3.1. Écart-type de la plage de variation de F0

L'écart-type de la plage de variation de F0 est de 6,12 demi-tons dans le corpus original et 6,20 demi-tons en condition de relecture. Dans le corpus original, on ne trouve pas de différence significative inter-locuteurs. Dans le corpus relecture, l'écart-type moyen de plage de F0 de Molossi est de 2,28 demi-tons et significativement plus faible que celui de Serey et Tuailleon

(ANOVA : $F[4,95]=6.51$, $p<.01$; $\eta^2=22\%$, $p<.001$ pour les deux comparaisons post-hoc). L'écart-type de la plage de variation en lecture de Le Chequer est également plus faible que celui de Tuailleon ($p<.01$).

Molossi est le seul locuteur pour lequel on calcule une différence significative d'écart-type moyen de la plage de variation de F0 par stimulus entre corpus original et relecture ($t[14]=4$, $p<.001$, $\eta^2=54\%$).

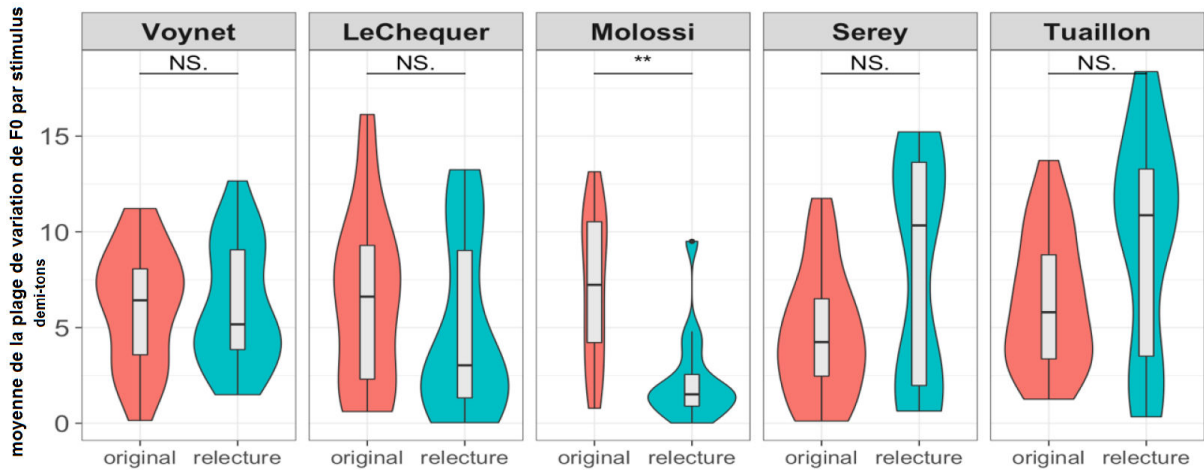


Figure VI-18: Diagrammes en violon et boîtes à moustaches représentant l'écart-type de la plage de variation moyenne de F0 en demi-tons par stimulus selon le locuteur et la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (* = $p<.05$, ** = $p<.01$, *** = $p<.001$, NS. = non significatif)

VI.4.3.2. Différences inter-blocs d'étendue de F0

La différence inter-blocs inter-pausaux de plage de F0 est en moyenne tous locuteurs confondus de 7.75 demi-tons dans le corpus original et de 6.23 demi-tons en condition de relecture. Tous locuteurs confondus, il n'y a pas de différence significative entre ces valeurs. On calcule une différence significative pour deux locuteurs sur cinq seulement : pour Molossi ($t[20]=7.5$, $p<.001$, $\eta^2=74\%$) et Le Chequer ($t[22]=2.4$, $p<.05$, $\eta^2=20\%$), la variation de plage de F0 de bloc à bloc dans les extraits originaux est significativement supérieure à celle calculée dans les extraits relus.

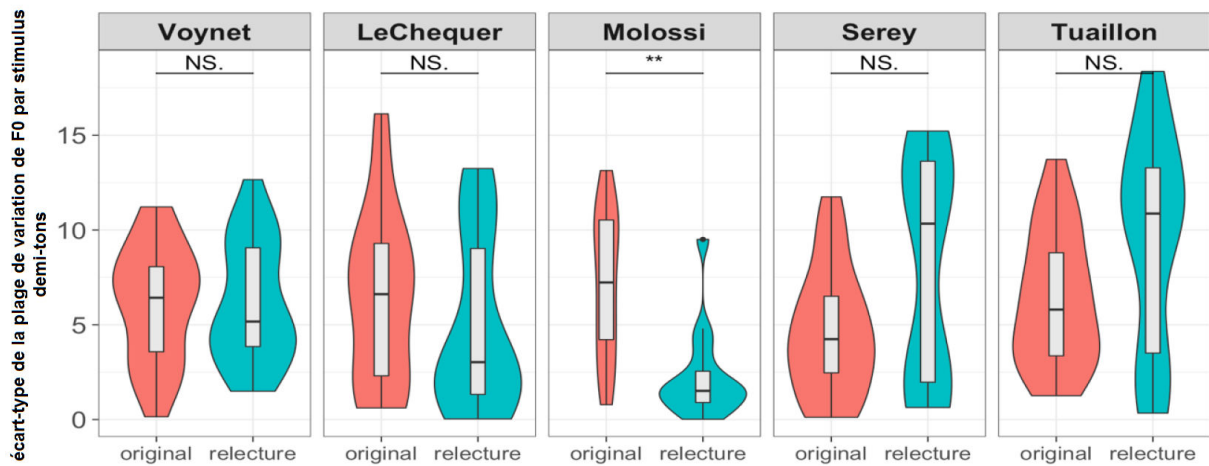


Figure VI-19 : Diagrammes en violon et boîtes à moustaches représentant les différences absolues d'étendue de la F0 de d'un bloc inter-pausal à l'autre moyenne par locuteur selon la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (* = p<.05, ** = p<.01, *** = p<.001, NS. = non significatif).

Si l'on considère à présent la différence brute de plage de F0 bloc à bloc, on observe des profils relativement hétérogènes parmi les locuteurs. Pour Voynet, la différence est en moyenne négative dans les deux conditions ce qui signifie que, d'un bloc inter-pausal à l'autre durant le déroulé de la portion de parole étudiée, la plage de variation de F0 tend à être plus réduite ; autrement dit que l'étendue de F0 a tendance à se réduire progressivement entre le début et la fin de l'énoncé.

Le seul locuteur pour lequel on calcule une différence significative entre conditions est Molossi ($t[20]=-2.2$, $p<.05$, $\eta^2=19\%$) : on mesure au sein de ses stimuli originaux des valeurs moyennes plus négatives que dans leur contrepartie relue.

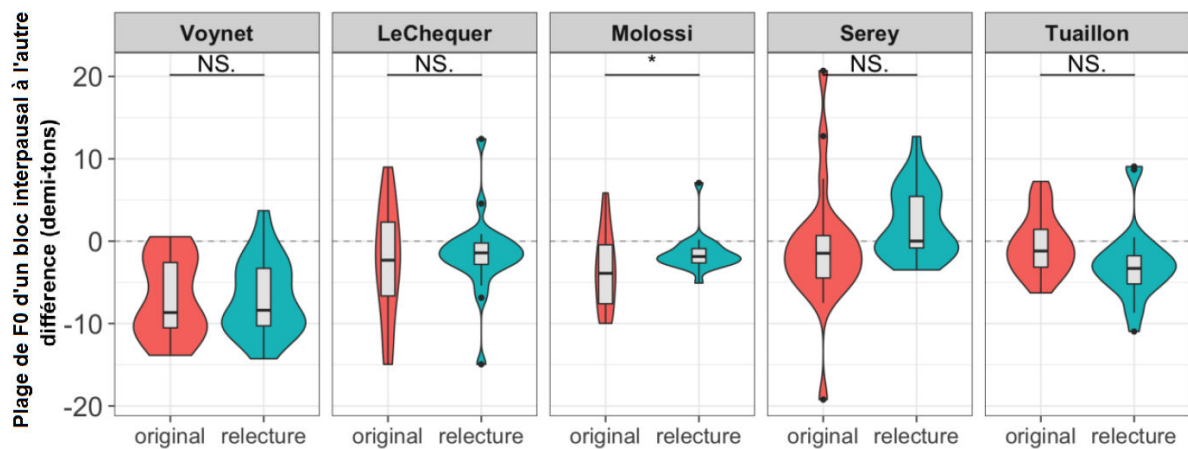


Figure VI-20 : Diagrammes en violon et boîtes à moustaches représentant les différences brutes de plage de variation de F0 d'un bloc inter-pausal à l'autre moyenne par locuteur selon la condition. Les étoiles correspondent au degré de significativité des valeurs p des tests t appariés calculés pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus (* = $p<.05$, ** = $p<.01$, *** = $p<.001$, NS. = non significatif).

VI.5. Inter-corrélations entre variables mélodiques et temporelles

Les corrélations présentées concernent exclusivement les caractéristiques mesurées dans le corpus original puisque c'est celui-ci qui est confronté à la perception des auditeurs. Nous présentons dans un premier temps les corrélations entre variables temporelles, puis entre variables mélodiques et enfin entre variables mélodiques et temporelles.

VI.5.1. Inter-corrélations entre variables temporelles

VI.5.1.1. Débit articuloire et pauses silencieuses

On ne trouve aucune corrélation significative dans le corpus original entre mesures de débit articuloire (nombre de syllabes par secondes sur temps de parole) et durées moyennes des pauses brutes comme relatives (durée pause / durée stimulus). On ne calcule pas non plus de corrélation significative entre débit articuloire et nombre de pauses silencieuses par stimulus, ni entre débit articuloire et proportion de silence par stimulus.

VI.5.1.2. Débit articuloire moyen et écart-type

On ne calcule aucune corrélation significative entre mesures de débit articuloire moyen et son écart-type.

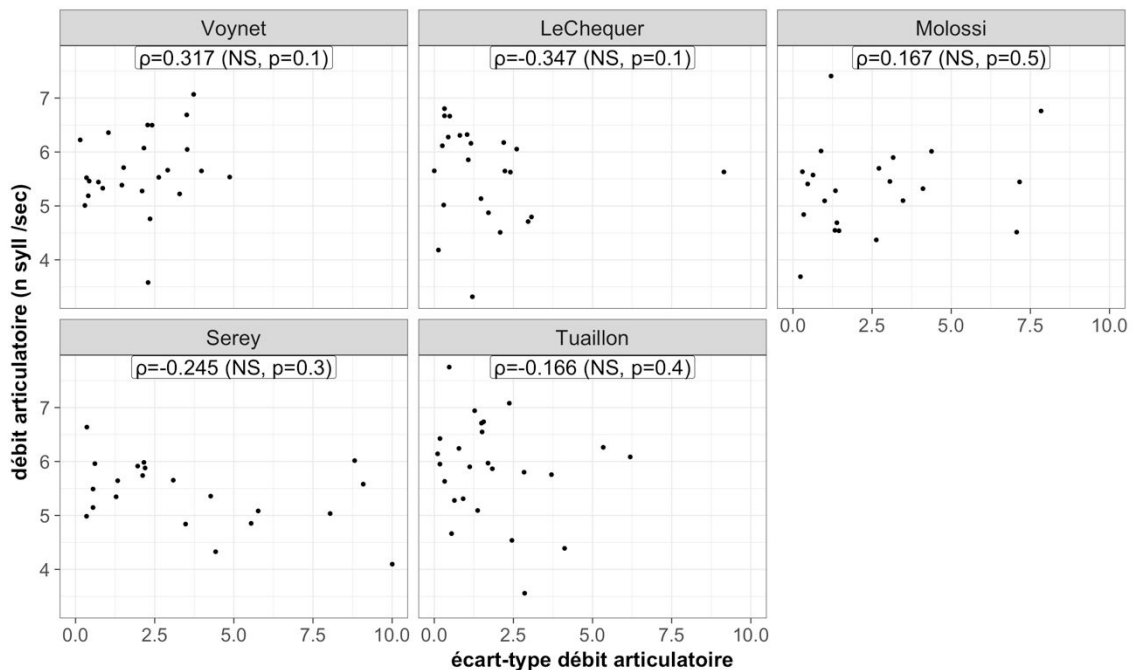


Figure VI-21 : Corrélations entre le débit articuloire moyen (n syllabes / secondes) et son écart-type. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VI.5.1.3. Débit articuloire et IVP de durée des syllabes

Tous locuteurs confondus, on calcule une corrélation significative négative entre débit articuloire et l'index de variabilité paire à paire (IVP) de la durée des syllabes dans le corpus original ($\rho = -.632$, $p < .01$). Pour l'ensemble des locuteurs du corpus, plus le débit articuloire est élevé, plus l'IVP moyen de durée syllabique par bloc inter-pausal est réduit.

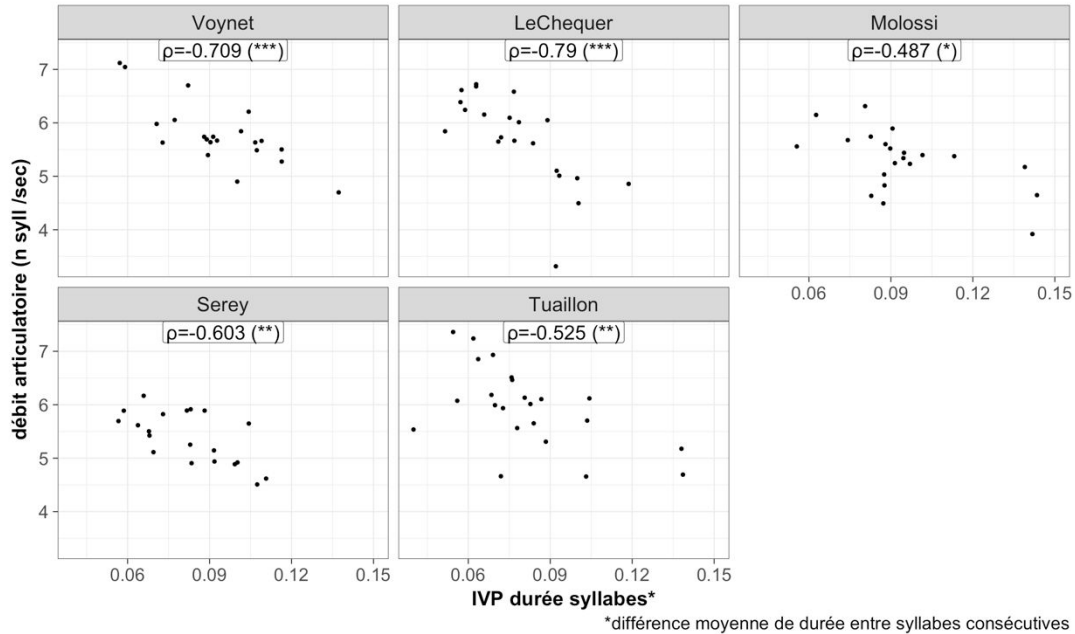


Figure VI-22 : Corrélations entre débit articuloire moyen et index de variabilité paire à paire de la durée des syllabes (IVP). Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VI.5.1.4. Débit articuloire et nombre de frontières marquées par la durée

Si l'on analyse les rapports entre débit articuloire et marquage temporel des frontières prosodiques, on trouve une corrélation significative pour un seul locuteur : Le Chequer ($\rho = -.439$, $p < .05$). Les stimuli de Le Chequer produits avec un débit articuloire maximalement élevé sont ceux pour lesquels on ne compte aucune frontière prosodique marquée par un allongement final, les stimuli comportant la présence de ce type de frontières prosodiques étant produits avec un débit articuloire plus bas.

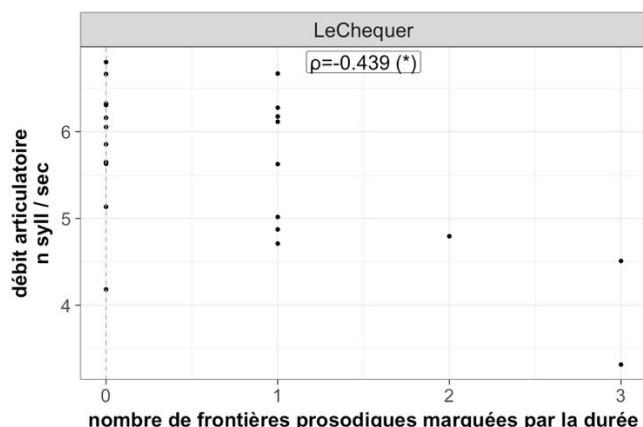


Figure VI-23 : Corrélations entre débit articulaire et nombre de frontières prosodiques marquées par la durée (allongement syllabique final relatif > 3). Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VI.5.1.5. Débit articulaire et durée des voyelles finales pré-pausales

La durée des voyelles finales est significativement corrélée au débit articulaire pour trois locuteurs sur cinq : Le Chequer ($\rho = -.531$, $p < .01$), Serey ($\rho = -.574$, $p < .01$), et Tuailon ($\rho = -.511$, $p < .01$). Pour ces trois locuteurs, plus le débit articulaire moyen est élevé, plus les voyelles finales pré-pausales sont courtes. Il est intéressant de remarquer que malgré le fait que la mesure de durée des voyelles finales soit brute et non relative au débit, la corrélation n'est pas significative pour tous les locuteurs. On n'observe pas de corrélation entre durée des voyelles finales et débit articulaire pour Voynet ($\rho = -.215$, $p = 0.3$) et Molossi ($\rho = -.305$, $p = 0.2$), ce qui veut dire que pour ces deux locuteurs la durée des voyelles finales semble relativement indépendante du débit articulaire.

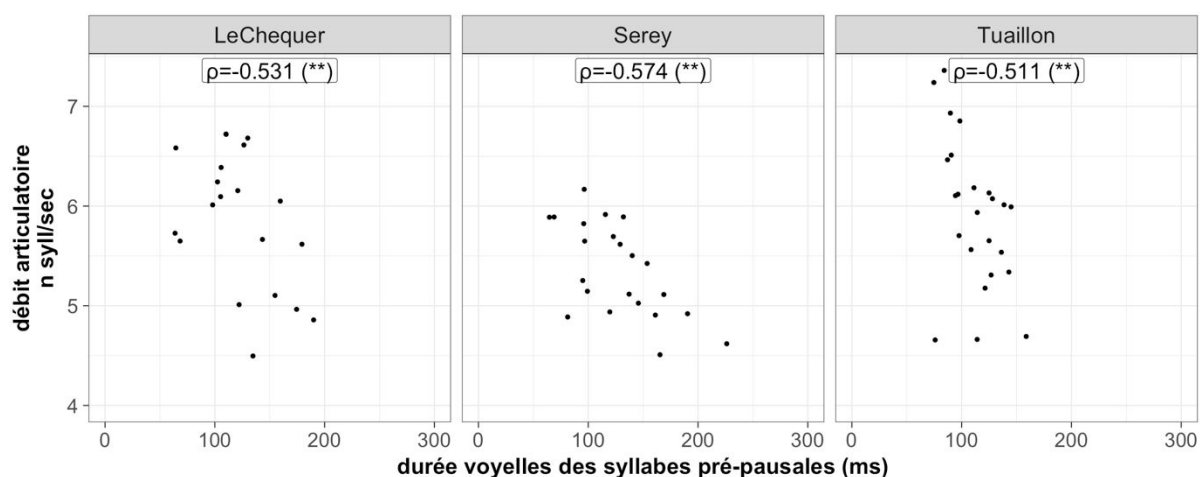


Figure VI-24 : Corrélations entre débit articulaire moyen (n syll /sec) et durée moyenne des voyelles des syllabes pré-pausales (fin de blocs inter-pausaux). Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VI.5.1.6. Débit articuloire moyen et différences inter-blocs

Les variations de débit articuloire de part et d'autre des pauses, soit d'un bloc inter-pausal à un autre, sont corrélées au débit articuloire moyen pour deux locuteurs sur trois : Le Chequer ($\rho=-.400$, $p<.05$) et Molossi ($\rho=-.556$, $p<.01$). Pour ces deux locuteurs, les extraits pour lesquels on mesure le débit articuloire moyen le plus élevé sont également des extraits au sein desquels on trouve en moyenne un débit articuloire légèrement plus faible après une pause silencieuse qu'avant celle-ci.

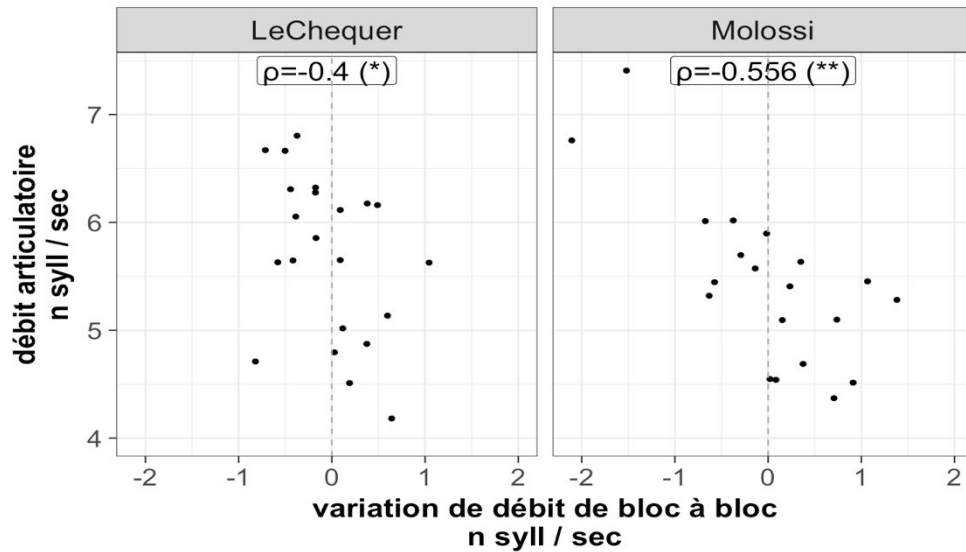


Figure VI-25 : Corrélations entre débit articuloire moyen et variations de bloc à bloc. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

VI.6. Inter-corrélations entre variables mélodiques

VI.6.1. Moyenne (Hertz) et plage de variation (demi-tons) de F0

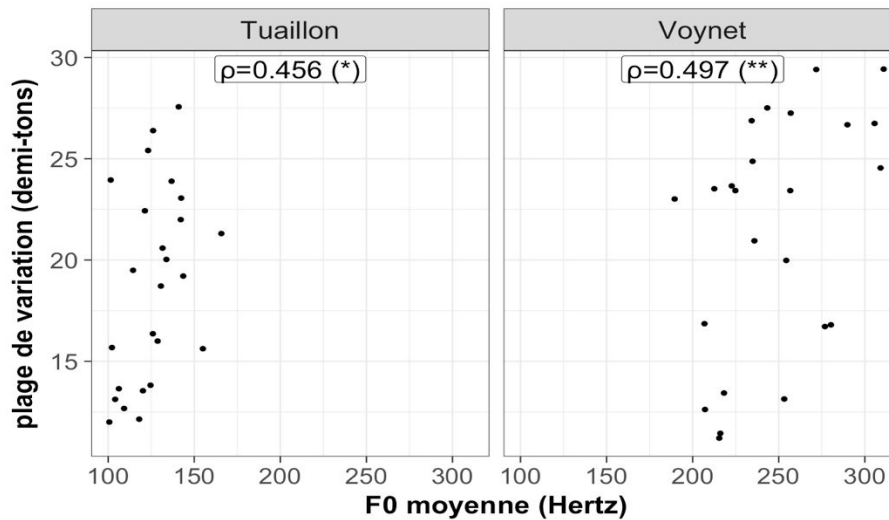


Figure VI-26 : Inter-corrélations entre F0 moyenne (Hertz) et plage de variation de F0 (demi-tons) pour Voynet et Tuailon. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

Nous calculons une corrélation significative entre F0 moyenne et plage de variation pour deux locuteurs sur cinq : Pour Voynet ($\rho = .497$, $p < .01$) et Tuailon ($\rho = .456$, $p < .05$), les stimuli sur lesquels on mesure la F0 moyenne la plus élevée sont également ceux pour lesquels on mesure la plage de variation de F0 la plus étendue. Cette corrélation positive n'est pas significative pour les trois autres locuteurs : Le Chequer ($\rho = 0.334$, $p = 0.09$), Molossi ($\rho = 0.232$, $p = 0.3$) et Serey ($\rho = 0.222$, $p = 0.3$).

VI.7. Inter-corrélations entre variables mélodiques et temporelles

Dans cette partie nous présentons par ordre de significativité les corrélations entre variables temporelles et mélodiques principales.

VI.7.1. Frontières marquées par la durée et plage de variation de F0

Si le nombre de frontières marquées par la durée n'est corrélé pour aucun locuteur à la F0 moyenne mesurée par stimulus, il est significativement corrélé à la plage de variation pour deux locuteurs sur cinq : Serey ($\rho=.695$, $p<.01$) et Voynet ($\rho=.392$, $p<.05$), pour qui les stimuli au sein desquels on décompte le plus haut nombre de frontières prosodiques majeures marquée par la durée sont également ceux pour lesquels la plage de F0 mesurée est la plus étendue. Cette tendance est particulièrement marquée pour Serey, pour lequel on voit sur la figure ci-dessous que les stimuli dans lesquels il ne produit pas de marquage temporel des frontières prosodiques sont ceux pour lesquels on mesure la plage de variation de F0 la plus réduite.

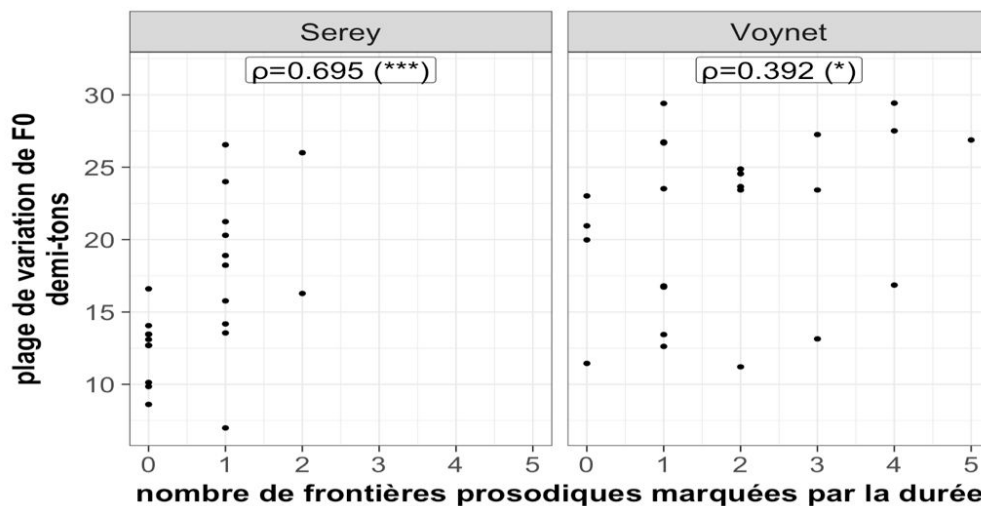


Figure VI-27 : Corrélations entre plage de F0 et nombre de frontières prosodiques marquées par la durée pour Serey et Voynet. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = \text{NS}$, non significatif.

VI.7.2. Débit articulatoire et F0 moyenne

Le débit articulatoire et la F0 moyenne ne sont corrélés que pour un locuteur sur cinq, Le Chequer, seul locuteur pour lequel les extraits produits avec le débit articulatoire le plus élevé sont également ceux pour lesquels on mesure une F0 moyenne maximale ($\rho=.521$, $p<.01$).

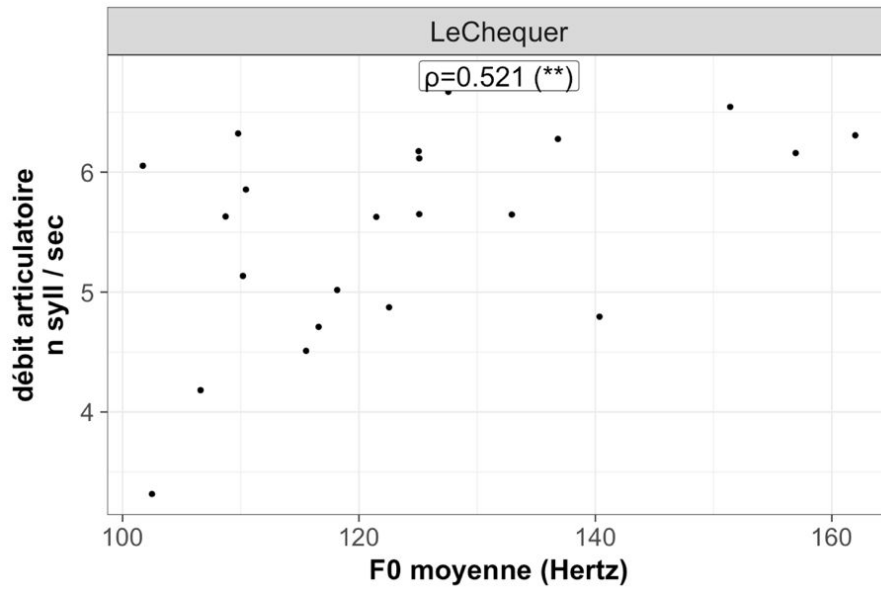


Figure VI-28 : Inter-corrélations entre débit articulaire (n syllabes / secondes sur temps de parole) et F0 moyenne (Hertz) par stimulus pour Le Chequer. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VII. Confrontation entre structurations prosodique et syntaxique

Dans ce chapitre, nous analysons les rapports entre prosodie et syntaxe à travers une analyse croisée des frontières prosodiques et des frontières syntaxiques annotées puis nous cherchons à déterminer si certaines caractéristiques prosodiques fines sont directement influencées par le type de frontières prosodico-syntaxiques présentes dans le corpus original et sa contrepartie relue.

Les frontières prosodiques ont été analysées selon leur localisation dans la structure syntaxique des énoncés et catégorisées comme suit, selon trois types : frontières prosodiques inter-unités syntaxiques majeures (*SyntMaj*), inter-unités syntaxiques intermédiaires (*SyntInt*), intra-unités syntaxiques intermédiaires (*intra-synt*).

Les résultats sont les suivants :

- Dans le corpus original comme dans sa contrepartie relue, le type de frontière prosodico-syntaxique le plus prépondérant correspond aux frontières inter-unités syntaxiques intermédiaires, qui représentent à 35 à 42% des frontières produites dans le corpus original contre 37 à 49% dans le corpus relecture. Si le nombre moyen par stimulus de frontières prosodiques inter-unités syntaxiques majeures ne varie pas significativement entre conditions du corpus, on compte significativement plus de frontières prosodiques inter-unités syntaxiques intermédiaires pour trois locuteurs sur cinq et de frontières prosodiques intra-syntaxiques pour tous les locuteurs.

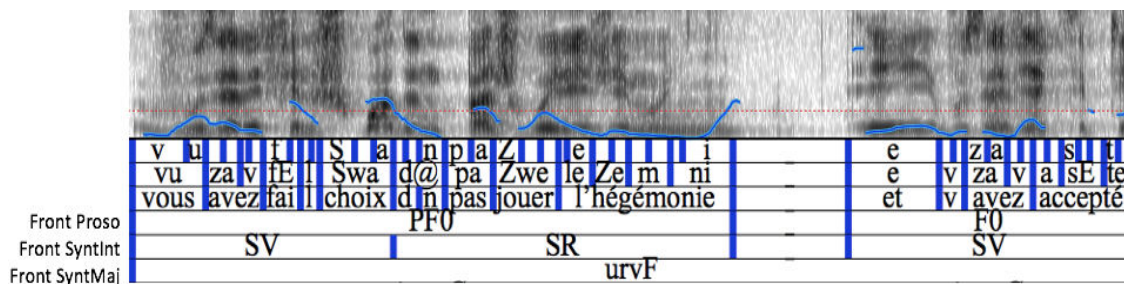
- L'analyse de l'influence du facteur « localisation syntaxique » sur les caractéristiques prosodiques originales préalablement mesurées indique que les frontières prosodiques intra-syntaxiques se distinguent des autres types de frontières prosodico-syntaxiques par des pauses silencieuses de durées brutes courtes, des mesures d'allongements syllabiques finaux pré-pausaux réduits, ainsi que de fortes différences de plage de variation de F0 de part et d'autre des pauses silencieuses.

VII.1. Typologie des frontières prosodico-syntaxiques analysées

Conformément aux recommandations de Degand & Simon (2009, 2011), nous avons annoté parallèlement les frontières syntaxiques (majeures et intermédiaires¹) et les frontières prosodiques majeures dans les extraits originaux et relus².

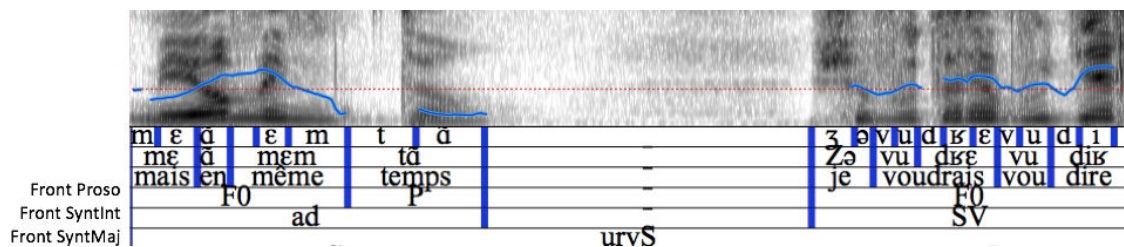
Nous avons ensuite analysé la localisation des frontières prosodiques dans la structure syntaxique en adoptant la typologie suivante :

1. **Frontière prosodique inter-unités syntaxiques majeures (*SyntMaj*)** : la frontière prosodique est située entre deux unités syntaxiques majeures, position équivalente à une frontière « inter-proposition » dans la terminologie de Duez (1991).



Exemple VII-1 : Exemple de frontière prosodique majeure marquée par une pause « frontière » de 405 millisecondes située entre deux unités syntaxiques majeures, extrait du stimulus n°11 de Tuailon en condition originale. Transcription simplifiée de l'extrait entier : « Vous avez fait le choix de ne pas jouer l'hégémonie, et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble afin de ne pas effrayer nos partenaires. ».

2. **Frontière prosodique inter-unités syntaxiques intermédiaires (*SyntInt*)** : la frontière prosodique est située entre deux unités syntaxiques intermédiaires, position équivalente à une frontière « inter-constituant » dans la terminologie de Duez (1991).

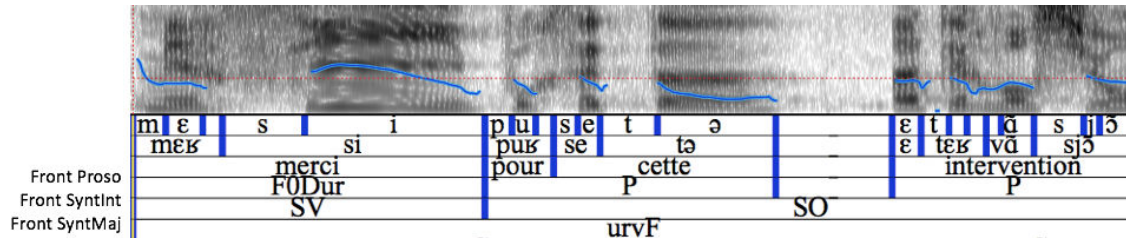


Exemple VII-2 : Exemple de frontière prosodique majeure marquée par une pause « frontière » située entre deux unités syntaxiques intermédiaires, extrait du stimulus n°1 de Voynet en condition de relecture. Transcription simplifiée de l'extrait entier : « mais en même temps, je voudrais vous dire qu'il n'est pas de tradition qu'on étale la vie personnelle des agents dans cette enceinte ».

¹ Voir partie III.4.2 pour le descriptif des critères pour l'annotation syntaxique.

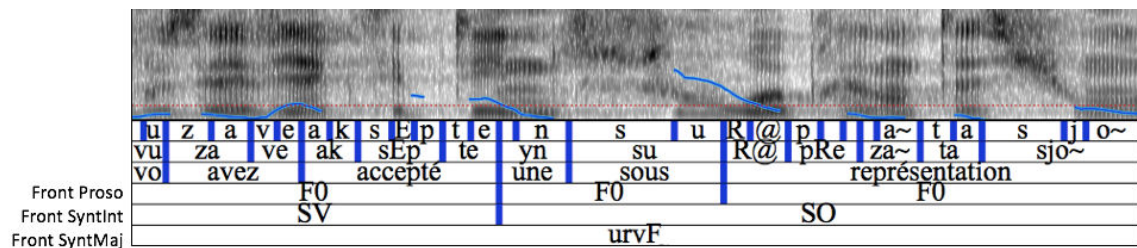
² Voir partie V.1 pour le descriptif des critères pour l'annotation prosodique.

3. Frontière prosodique intra-unités syntaxiques intermédiaires (*intra-synt*) : la frontière prosodique est située à l'intérieur même d'une unité syntaxique intermédiaire, position équivalente à une frontière « intra-constituant » dans la terminologie de Duez (1991).



Exemple VII-3 : Exemple de frontière prosodique majeure intra-syntaxique marquée par une pause « frontière » de 389 millisecondes située entre « cette » et « intervention » unités syntaxiques majeures, extrait du stimulus n°24 de Voynet en condition originale. Transcription simplifiée de l'extrait entier : « Merci pour cette intervention équilibrée, et de bonne foi, et qui ne sue pas l'ambiance électorale. »

4. Proéminences intra-unités lexicales (*intra-mot*)³ : la frontière prosodique est située en syllabe non finale de mot ou de groupe de mots, et correspond principalement à divers cas d'accents initiaux ayant donné lieu à l'annotation prosodique d'une saillance mélodique et/ou temporelle⁴.



Exemple VII-4 : Exemple de frontière prosodique majeure marquée par une frontière prosodique de type « F0 » sur la syllabe [su] du préfixe dans le mot « sous-représentation ». Extrait du stimulus n°11 de Tuailon en condition originale. Transcription simplifiée de l'extrait entier : « Vous avez fait le choix de ne pas jouer l'hégémonie, et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble afin de ne pas effrayer nos partenaires. »

³ Cette quatrième catégorie n'apparaît pas lorsqu'on concentre l'analyse de la structuration prosodico-syntaxique sur les pauses silencieuses.

⁴ Nous renvoyons pour une description des différents cas d'accents initiaux pouvant être rencontrés à la partie II.4.2 et pour une discussion de l'analyse de ce type de phénomène dans le cadre d'une étude sur la structuration prosodico-syntaxique et la parole expressive au chapitre IX.

VII.2. Présentation des résultats

Dans cette section nous présentons la répartition des frontières prosodiques majeures annotées en fonction de leur position dans la structuration syntaxique des énoncés. Nous présentons les tendances générales au sein du corpus original et comparons celles-ci à la contrepartie relue des extraits (corpus relecture). Dans un premier temps, nous présentons une vue d'ensemble de la répartition des frontières prosodiques selon la syntaxe dans le corpus, puis nous présentons les occurrences par stimulus selon la localisation syntaxique. Enfin, après une précision spécifique sur la répartition des pauses « frontières »⁵, nous étudions et présentons les corrélations significatives observées entre localisation syntaxique et caractéristiques prosodiques analysées dans le chapitre VI.

VII.2.1. Répartition générale des frontières prosodico-syntaxiques

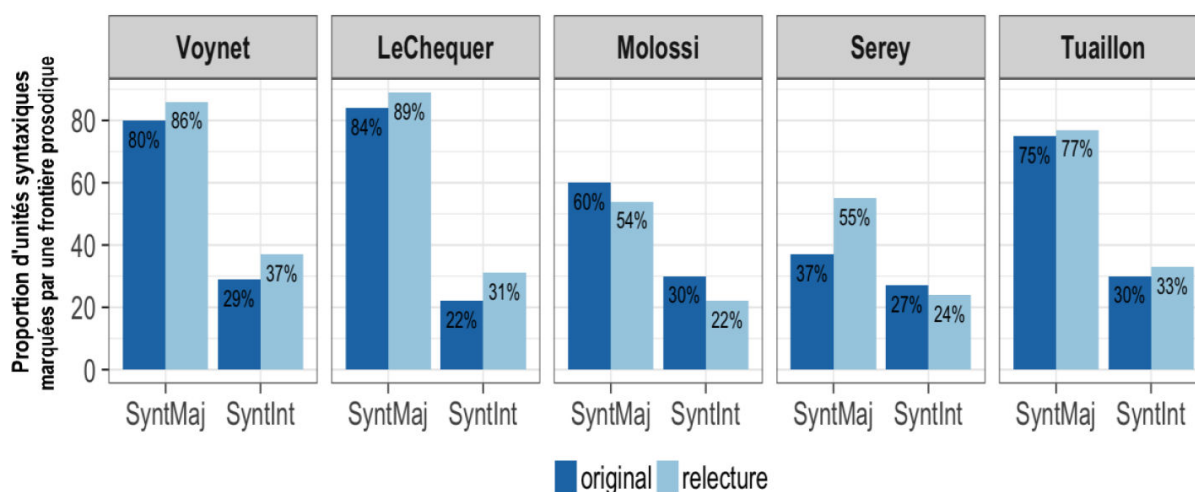


Figure VII-1 : Proportion d'unités syntaxiques majeures (SyntMaj) et intermédiaires (SyntInt) marquées par une frontière prosodique par rapport au total d'unités syntaxiques majeures et intermédiaires annotées.

Dans le corpus original, pour quatre locuteurs sur cinq, 60 à 84% des unités *SyntMaj* sont marquées par une frontière prosodique. La proportion est plus basse pour un locuteur, Serey pour qui seulement 37% des unités *SyntMaj* sont marquées prosodiquement. Dans le corpus original relecture, la proportion d'unités *SyntMaj* marquées prosodiquement peut monter jusqu'à 89% pour Le Chequer et 86% pour Voynet.

⁵ Les frontières prosodiques peuvent être annotées comme majeures sans la présence d'une pause silencieuse dans certains cas de saillances mélodiques et/ou temporelles pertinentes (frontières « F0Dur », « F0 » et « Dur »). Compte tenu de l'importance de l'analyse des pauses silencieuses dans notre étude, nous analysons distinctement la position des pauses.

A l'exception de Serey pour qui seulement 14% des unités syntaxiques intermédiaires produites sont terminées par une frontière prosodique majeure, les unités syntaxiques intermédiaires des autres locuteurs sont terminées par des frontières prosodiques majeures dans 19 à 20% des cas.

Nous présentons à présent la répartition des frontières prosodiques selon leur localisation dans la structure des énoncés ; voir figure suivante ci-dessous.

Répartition des frontières prosodiques dans la structure syntaxique

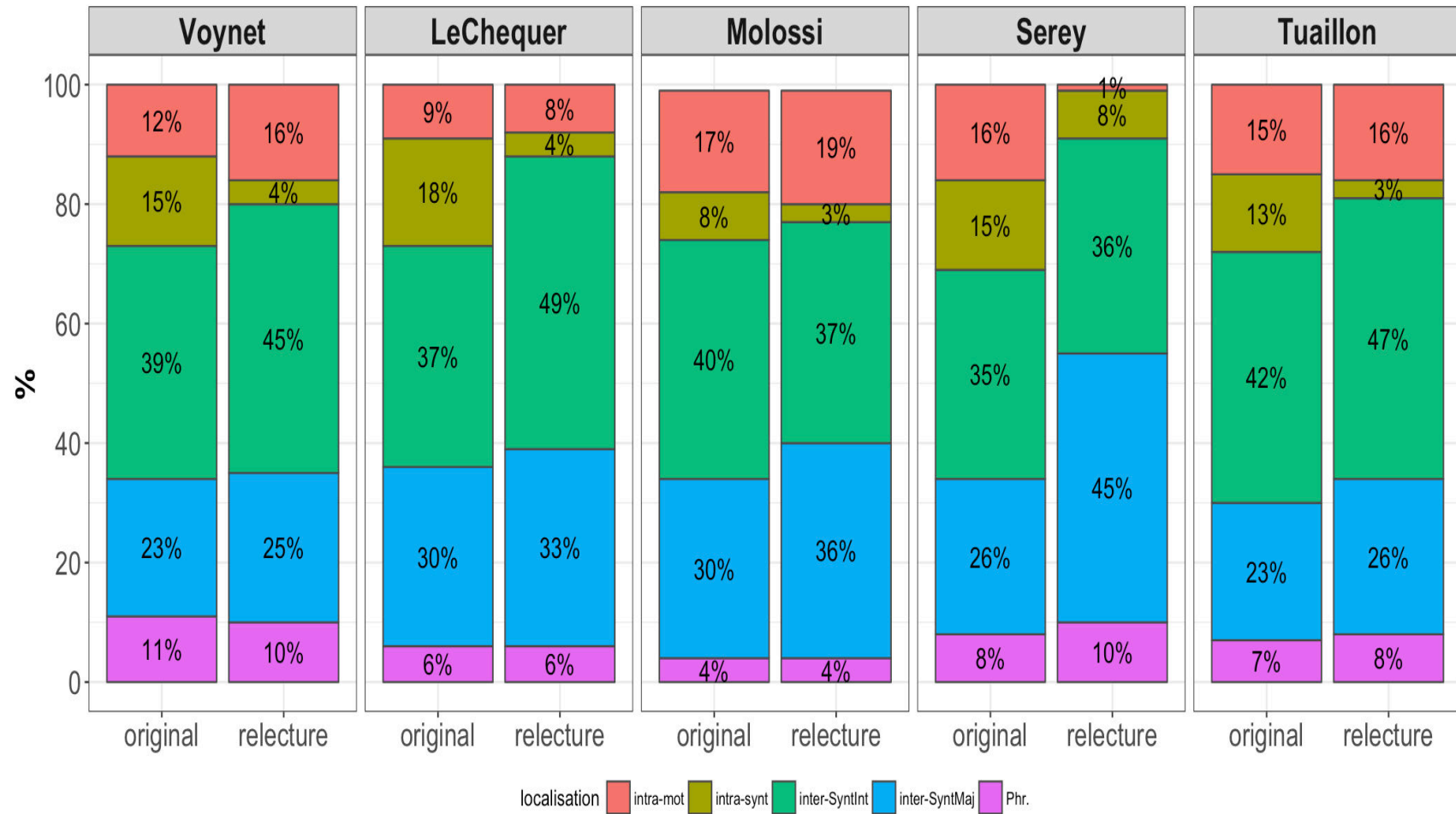


Figure VII-2 : Répartition des frontières prosodiques selon leur localisation dans la structure des énoncés.

VII.3. Distributions par stimulus

VII.3.1. Frontières prosodiques inter-unités syntaxiques majeures

Dans le corpus original comme dans la relecture, on compte en moyenne par stimulus 1 à 2 frontières prosodiques majeures situées en frontières d'unités syntaxiques majeures (inter-SyntMaj). On ne calcule aucune différence significative du nombre moyen de frontières prosodiques inter-SyntMaj entre conditions.

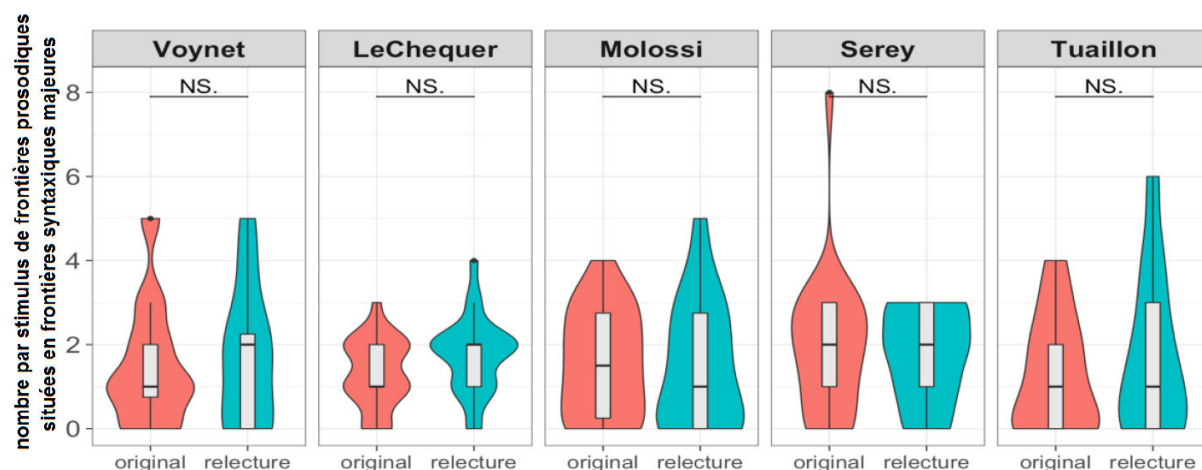


Figure VII-3 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques situées en frontières d'unités syntaxiques majeures selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, $p > .05$ = NS, non significatif.

VII.3.2. Frontières prosodiques inter-unités syntaxiques intermédiaires

On compte 1 à 3 frontières prosodiques inter-unités syntaxiques intermédiaires par stimulus dans chacune des conditions du corpus.

Comme on peut l'observer sur la figure-ci dessous, on calcule une différence significative du nombre de frontières prosodiques inter-unités syntaxiques intermédiaires pour trois locuteurs sur cinq : Molossi ($t[18]=2.7$, $p < .05$, $\eta^2=29\%$), Voynet ($t[23]=-2.4$, $p < .05$, $\eta^2=20\%$) et Serey ($t[22]=-2.2$, $p < .05$, $\eta^2=17\%$). Voynet et Serey produisent en moyenne plus de frontières prosodiques en position inter-unités syntaxiques intermédiaires dans la version originale de leurs stimuli que dans leurs relectures. Pour Molossi la tendance est inverse : on décompte significativement plus de frontières de ce type dans les stimuli originaux [que dans les stimuli relus](#).

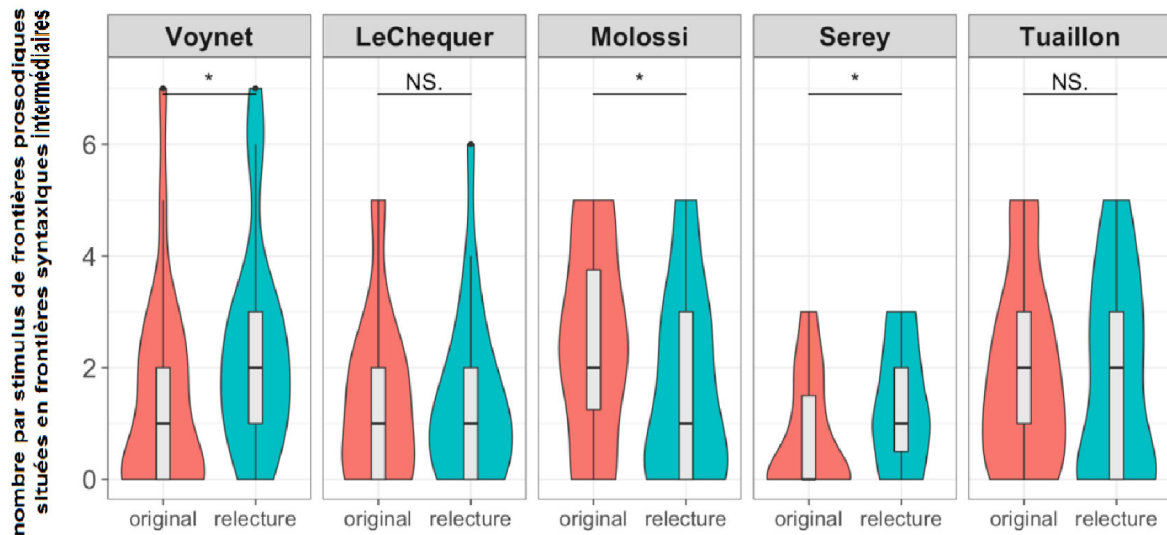


Figure VII-4 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques situées en frontières d'unités syntaxiques majeures selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, $p > .05$ = NS, non significatif.

On trouve une différence inter-locuteurs significative dans le corpus original ($F[4,112]=3.32$, $p=0.01$; $\eta^2=11\%$) : le nombre moyen par stimulus pour Molossi est significativement supérieur à celui de Serey ($p < .001$). Au sujet de ce dernier, nous rappelons pour faciliter l'interprétation du faible nombre de frontières de ce type que Serey est le seul locuteur dont les frontières prosodiques majeures se situent majoritairement entre unités syntaxiques majeures et non entre unités syntaxiques intermédiaires. Dans le corpus original, les frontières inter-unités syntaxiques majeures représentent 44% du total des frontières prosodiques majeures produites par Serey, contre seulement 26 à 34% pour les autres locuteurs. Au contraire, les frontières prosodiques inter-unités syntaxiques intermédiaires ne représentent que 18% des frontières prosodiques produites par Serey, contre 31 à 48% des frontières prosodiques par les autres locuteurs.

VII.3.3. Frontières prosodiques intra-unités syntaxiques

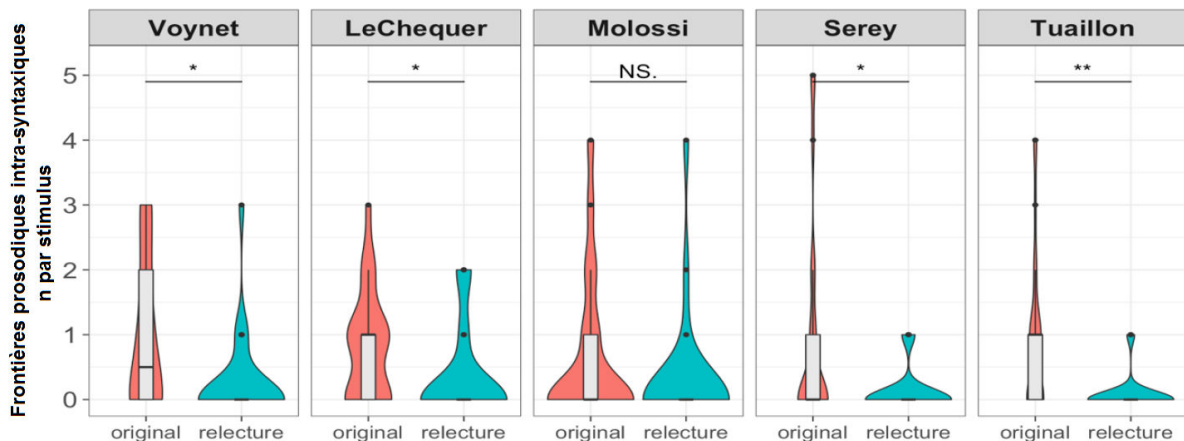


Figure VII-5 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques intra unités syntaxiques selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, $p > .05$ = NS, non significatif.

Les frontières prosodiques intra-syntaxiques sont les frontières pour lesquelles on observe le plus de différences entre corpus original et relecture ainsi qu'entre locuteurs au sein de chaque condition. Pour rappel, dans le corpus original, le nombre total par locuteur de frontières prosodiques intra-syntaxiques va de 13 à 23 contre seulement 2 à 8 dans le corpus relecture. Dans le corpus original, les frontières prosodiques intra-syntaxiques représentent 19 à 21% de la totalité des frontières prosodiques produites par tous les locuteurs, à l'exception de Molossi, pour qui elles ne représentent que 13%.

Tous locuteurs confondus, on calcule une différence significative entre corpus original et relecture concernant le nombre de frontières prosodiques intra-syntaxiques par stimulus ($t[116]=5.6$, $p < .001$, $\eta^2=21\%$). En effet, comme on peut l'observer pour tous les locuteurs sauf Molossi, le nombre moyen de ce type de frontières est supérieur en condition originale.

VII.3.4. Proéminences prosodiques intra-mots

Les proéminences prosodiques intra-mots, situées en syllabe non finale de mot, représentent 2 à 8% des frontières prosodiques majeures annotées dans le corpus.

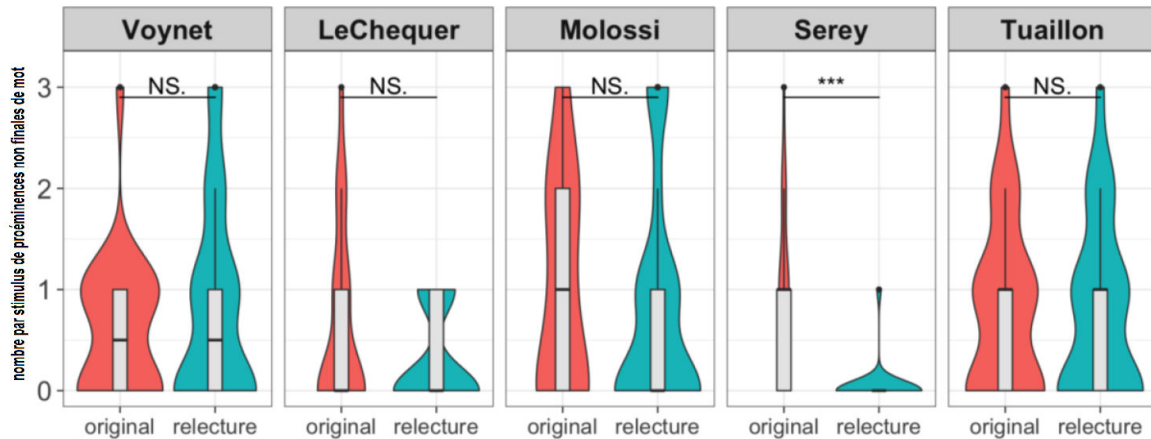


Figure VII-6 : Diagrammes en violons et boîtes à moustaches représentant le nombre de proéminences prosodiques intra mots syntaxiques selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = p<.001, ** = p<.01, * = p<.05, p>.05 = NS, non significatif.

À l'échelle du stimulus, on ne trouve de différence significative entre conditions que pour Serey ($t[22]=4.2$, $p<.001$, $\eta^2=45\%$), comme on peut l'observer sur la figure ci-dessus. L'absence de différence significative entre conditions pour les autres locuteurs est essentiellement due selon nous au fait que l'annotation de ce type de frontières ne permet pas de dissocier les cas de proéminence syllabique initiale de type accent didactique de ceux de type accents d'insistance. Nous reviendrons sur ce problème dans le chapitre IX (Discussion).

VII.4. Localisation des pauses « frontières » dans la structure syntaxique

VII.4.1. Pauses « frontières » inter-unités syntaxiques majeures

De même que le nombre de frontières prosodiques majeures situées entre unités syntaxiques majeures, le nombre de pauses « frontières » ne diffère pas entre stimuli originaux et relus.

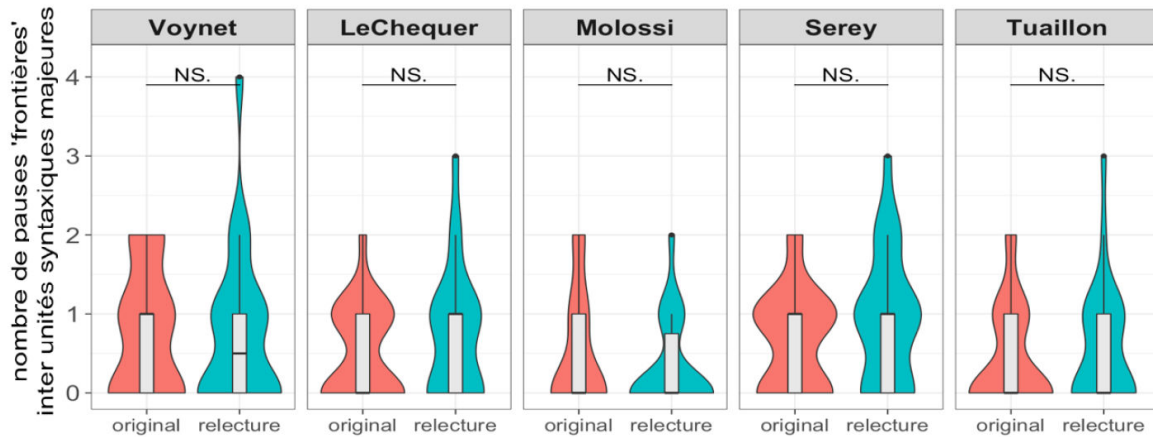


Figure VII-7 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » situées en frontières d'unités syntaxiques majeures selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, ns = $p > .05$

VII.4.2. Pauses « frontières » inter-unités syntaxiques intermédiaires

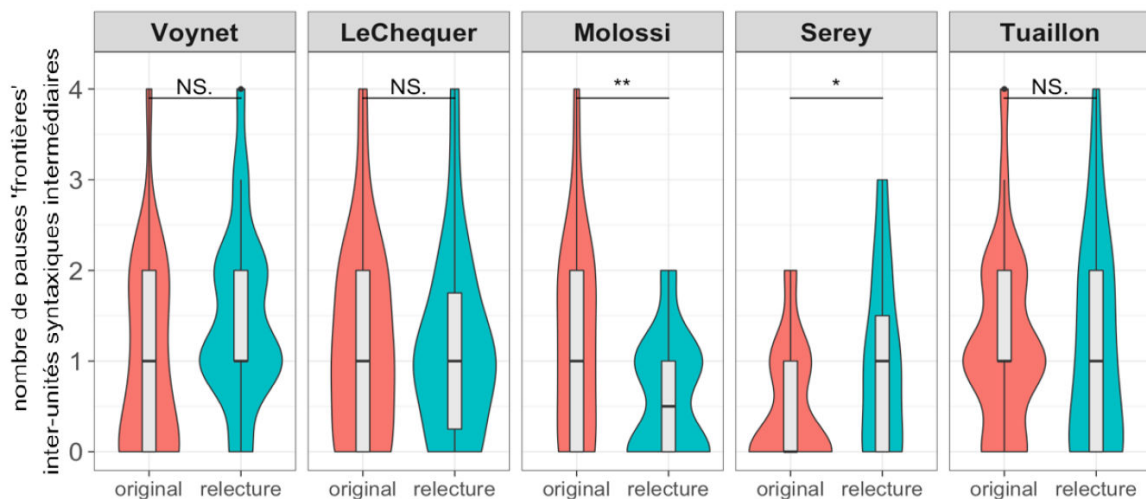


Figure VII-8 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » situées en frontières d'unités syntaxiques intermédiaires selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p < .001$, ** = $p < .01$, * = $p < .05$, ns = $p > .05$.

Si l'on considère non pas le nombre de frontières prosodiques majeures en frontières inter-unités syntaxiques mais le nombre de pauses silencieuses « frontières », écartant donc les cas de frontières prosodiques majeures marquées uniquement par F0 ou la durée, on constate comme on peut le voir sur la figure ci-dessus que la différence observée entre corpus original et relecture pour les frontières prosodiques dans leur globalité ne se retrouve que pour Molossi

($t[21]=2.8$, $p<.01$, $\eta^2=28\%$) et Serey ($t[22]=-2.4$, $p<.05$, $\eta^2=21\%$), mais pas pour Voynet ($t[23]=-1.8$, $p=0.1$, $\eta^2=12\%$). Rappelons que cette dernière produisait significativement moins de frontières prosodiques inter-unités syntaxiques majeures en condition originale qu'en relecture mais pas pour autant significativement moins de pauses frontières à cette localisation syntaxique.

VII.4.3. Pauses « frontières » intra-syntaxiques par stimulus

Il convient tout particulièrement pour cette localisation syntaxique d'analyser non pas le nombre de frontières prosodiques tous paramètres confondus mais les cas d'insertion de pauses silencieuses « frontières » intra-unités syntaxiques intermédiaires.

Comme on peut le voir dans la **Erreur ! Nous n'avons pas trouvé la source du renvoi.**, dans le corpus original, les pauses « frontières » intra-syntaxiques représentent près d'un tiers du total des pauses « frontières » produites dans cette condition pour tous les locuteurs sauf Molossi, pour qui elles ne représentent que 11% des pauses produites.

Dans le corpus relecture, on ne compte aucune pause « frontière » intra-syntaxique pour Molossi. Pour les autres locuteurs, le nombre total varie entre 2 et 5, ce qui représente au maximum 11% des pauses produites par Le Chequer, et 9% des pauses de Voynet.

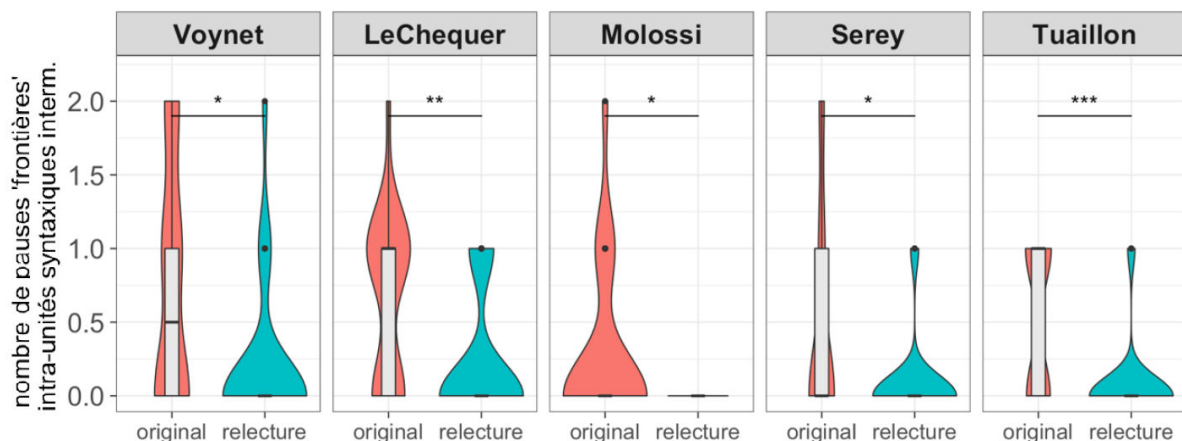


Figure VII-9 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » intra-unités syntaxiques intermédiaires selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : *** = $p<.001$, ** = $p<.01$, * = $p<.05$, ns = $p>.05$.

On constate que les différences entre stimuli originaux et relus sont encore plus marquées lorsque l'on considère le nombre de pauses « frontières » que le nombre de frontières prosodiques tous paramètres confondus. La différence la plus significative entre stimuli originaux et relecture est trouvée pour Tuailleon ($t[24]=3.8$, $p<.001$, $\eta^2=37\%$).

Répartition des pauses 'frontières' dans la structure syntaxique

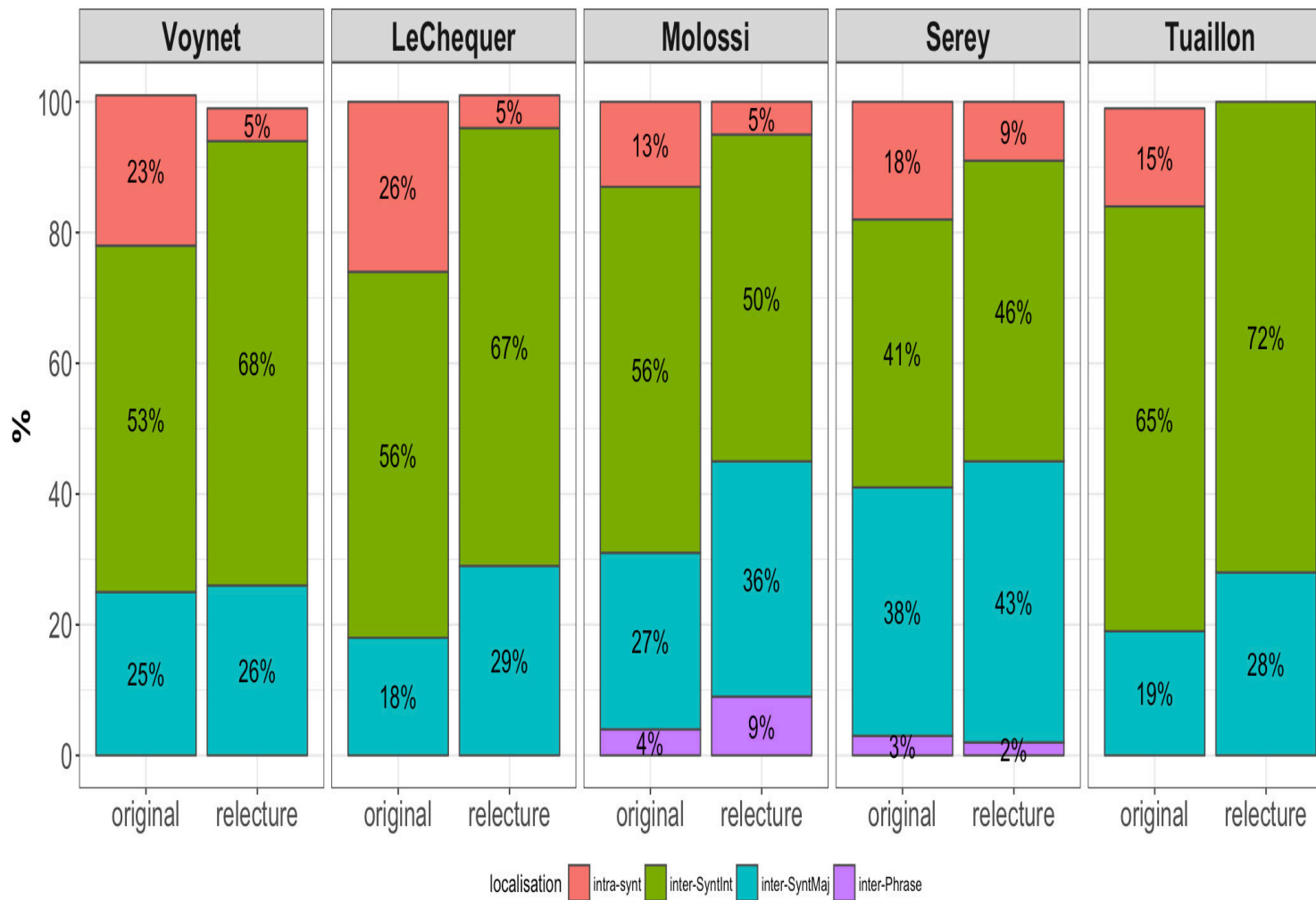


Figure VII-10 : Proportions de pauses « frontières » selon leur position dans la structure syntaxique

VII.5. Caractéristiques prosodiques selon la position syntaxique des frontières

Dans cette partie, nous présentons les caractéristiques prosodiques pour lesquelles nous observons une variation due au type prosodico-syntaxique de frontières (inter-unités syntaxiques majeures, intermédiaires, intra-syntaxiques). Dans cette partie nous analysons les caractéristiques des variables issues de l'analyse prosodique à travers des ANOVA à deux facteurs : locuteur et localisation syntaxique des frontières. Nous présentons les cas pour lesquels on trouve une variation significative des valeurs selon la localisation syntaxique des frontières.

VII.5.1. Caractéristiques temporelles

VII.5.1.1. Durée brute des pauses silencieuses

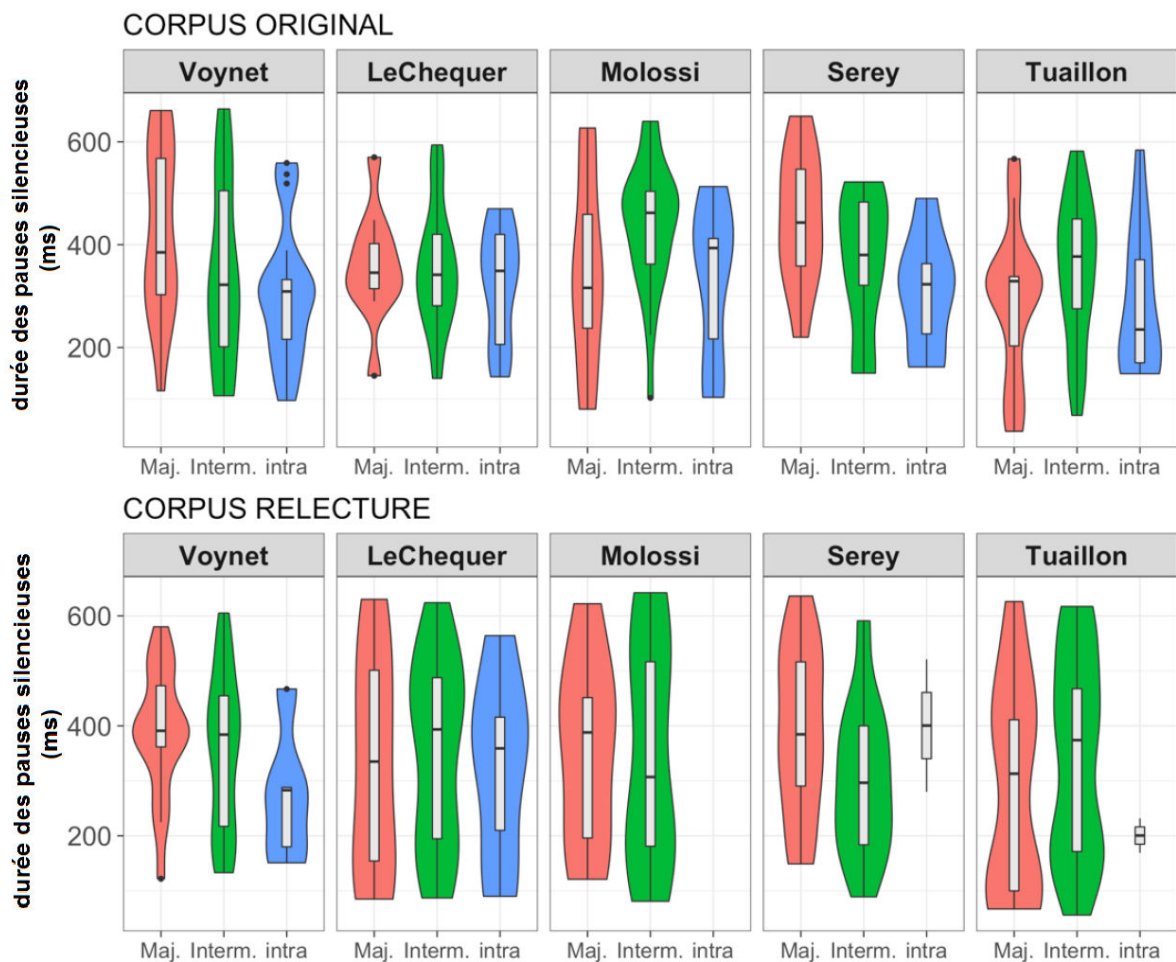


Figure VII-11 : Durée des pauses silencieuses (ms) selon leur position dans la structure syntaxique des énoncés. Maj : inter-unités syntaxiques majeures, Interm. : inter-unités syntaxiques intermédiaires, intra : intra-unités syntaxiques

L'ANOVA effectuée sur le corpus original indique une influence significative de la position syntaxique sur la durée brute des pauses silencieuses post frontières ($F[2,216]=4.05$, $p<0.05$; $\eta^2=4\%$) : tous locuteurs confondus, les pauses silencieuses intra-syntaxiques sont significativement plus courtes que celles situées en frontières inter-unités syntaxiques majeures et intermédiaires ($p<.05$ pour les deux comparaisons).

Cette tendance se retrouve particulièrement pour Voynet et Serey comme on peut l'observer sur la figure ci-dessous. Pour Molossi et Tuillon seules les pauses silencieuses inter-unités syntaxiques intermédiaires sont de durées supérieures aux pauses intra-syntaxiques. Pour Le Chequer, les pauses silencieuses sont de durée moyenne similaire quelles que soient leur position dans la structure syntaxique.

VII.5.1.2. Allongement final relatif pré-pausal

L'ANOVA effectuée sur le corpus original indique une influence significative de la position syntaxique sur l'allongement relatif syllabique final ($F[2,242]=4.83$, $p<0.01$; $\eta^2=4\%$) : tous locuteurs confondus, l'allongement relatif final pré-pausal le plus réduit est mesuré pour les frontières prosodiques intra-syntaxiques, significativement inférieur à celui des frontières prosodiques inter-unités syntaxiques intermédiaires (Tukey HSD : $p<.001$).

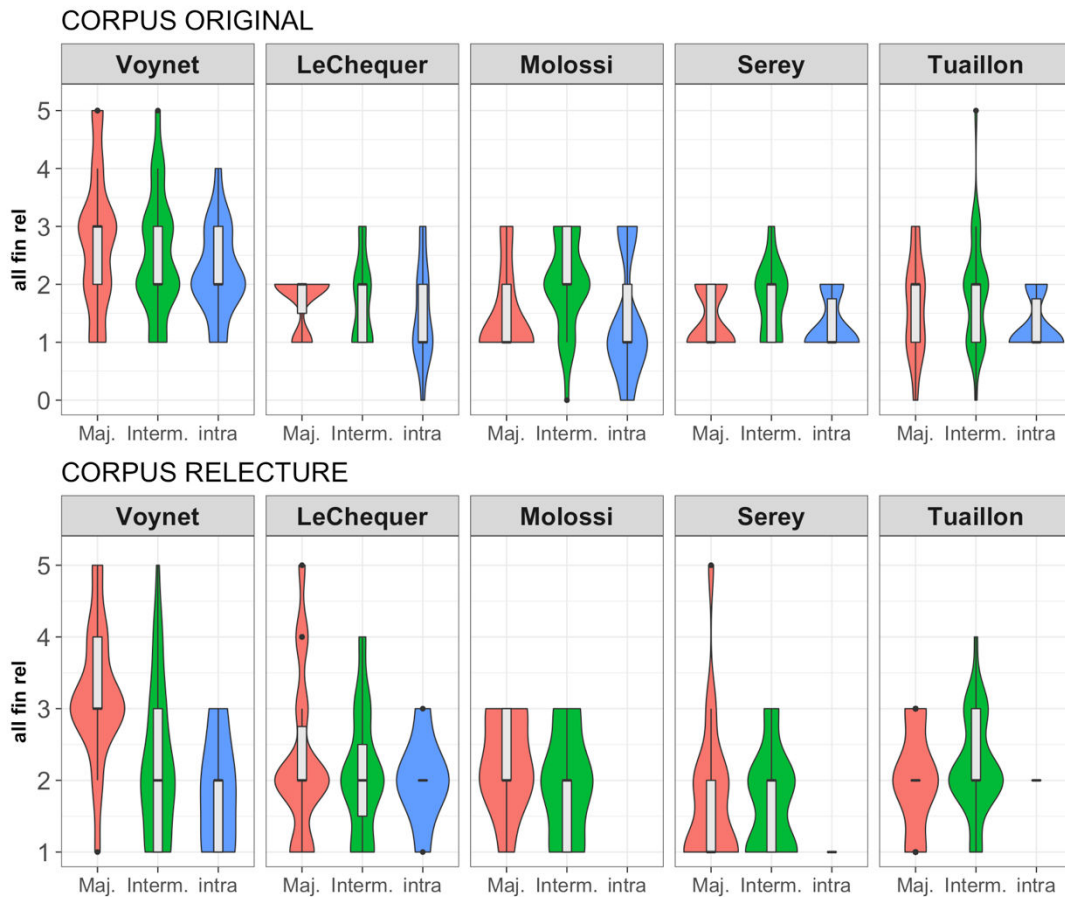


Figure VII-12 : Allongement final pré-pausal relatif (syllabique) selon la position des frontières prosodiques position dans la structure syntaxique des énoncés. Maj : inter-unités syntaxiques majeures, Interm. : inter-unités syntaxiques intermédiaires, intra : intra-unités syntaxiques.

VII.5.2. Caractéristiques mélodiques

VII.5.2.1. Plage de variation de F0

Si nous n'observons aucune influence significative de la position syntaxique des frontières prosodiques sur la F0 moyenne, il n'en est pas de même pour la plage de F0.

L'ANOVA effectuée sur le corpus original indique une influence significative de la position syntaxique sur la plage de variation de F0 des unités prosodiques pré-pausales ($F[2,400]=5.15$, $p<.01$; $\eta^2=3\%$).

Si on ne calcule pas de différence significative entre plages de variations moyennes d'unités prosodiques situées en frontières syntaxiques majeures ou intermédiaires, les unités prosodiques s'achevant à l'intérieur même d'unités syntaxiques, situées donc en position intra-syntaxique, sont produites avec une plage de variation significativement plus réduite qu'en position inter unités syntaxiques intermédiaires ($p<.001$).

Bien que le nombre de valeurs soit trop faible pour pouvoir effectuer ici des statistiques pour chaque locuteur, cette tendance peut s'observer qualitativement chez Voynet, Le Chequer et Tuillon.

VIII. Corrélats acoustiques de l'hostilité, de la dominance, de la valence et de l'excitation

Dans le **Chapitre VIII**, les résultats des analyses perceptives et prosodiques sont mis en relation dans le but d'estimer l'influence de caractéristiques prosodiques spécifiques sur la perception des locuteurs par les auditeurs. Nous présentons les corrélats significatifs de l'hostilité, de la valence, de l'excitation et de la dominance perçues. Au niveau perceptif, les variables utilisées sont les scores issus des évaluations perceptives. Au niveau acoustique, les variables sont issues de l'annotation prosodico-syntaxique ainsi que de l'analyse des caractéristiques temporelles et mélodiques moyennes par extraits et des différences mesurées entre unités prosodiques consécutives. Pour chaque variable prosodique, les valeurs originales et les valeurs d'écart par rapport aux extraits relus sont analysées.

Si le chapitre **IV** nous avait permis de constater que les locuteurs ont tous été perçus de façon homogène comme légèrement hostiles, légèrement mécontents et légèrement dominants, les corrélats prosodiques de ces dimensions se révèlent très variables selon les locuteurs. Parmi les paramètres les plus représentés dans les corrélations significatives on notera que :

- **L'accroissement de la F0 moyenne** est corrélé à l'hostilité pour deux locuteurs sur trois, au mécontentement pour trois locuteurs sur cinq, à l'excitation et à la dominance pour tous les locuteurs sauf Tuillon.

- La dominance perçue est corrélée au nombre de **contours descendants** pour Voynet et en cloche pour Serey. Pour ces deux locuteurs elle est également corrélée à la **réduction de plage de variation de F0** d'un bloc inter-pausal à l'autre.

- **La durée des pauses silencieuses** n'est positivement corrélée à la dominance perçue pour aucun locuteur, on calcule même une corrélation négative pour Le Chequer.

- **Le débit articulatoire** est corrélé à l'hostilité perçue pour deux locuteurs sur cinq : pour Voynet les stimuli perçus comme les plus hostiles sont ceux pour lesquels son débit articulatoire est le plus lent, pour Tuillon le plus rapide. Tuillon est également le seul locuteur pour lequel la dominance est positivement corrélée au débit articulatoire.

Enfin, si l'insertion de pauses silencieuses intra-syntaxiques n'est pas directement corrélée à l'hostilité et la dominance perçues, on trouve pour ces deux dimensions des corrélats prosodiques que nous avons mis en évidence comme étant propres à ce type de structuration prosodico-syntaxique (contrastes de débit articulatoire et de plage de variation de part et d'autre de celles-ci).

VIII.1. Présentation de la méthodologie d'analyse

Dans ce chapitre, nous présentons pour chaque dimension perçue (hostilité, valence, excitation, dominance) les corrélats acoustiques des locuteurs du corpus original. Nous explicitons ici notre méthodologie d'analyse ainsi que les variables perceptives et productives confrontées au sein de cette analyse.

Pour ce faire, nous confrontons les scores perceptifs obtenus dans les études du chapitre IV aux caractéristiques mesurées en termes de structuration prosodique majeure (voir chapitre V) mais également de structuration prosodico-syntaxique (voir chapitre VII) ainsi qu'aux caractéristiques de F0, des pauses silencieuses, et du débit articulatoire et des variations de durées syllabiques mesurées (voir chapitre VI). Pour chacun de ces sous-ensembles des résultats, nous confrontons à la perception d'une part les caractéristiques mesurées dans le corpus original, et d'autre part les mesures d'écart entre caractéristiques originales et relues pour chaque stimulus (différence original – relecture).

Nous ne présentons ici que les corrélations significatives. Compte tenu de l'hétérogénéité des tendances déjà observées entre les locuteurs au niveau des analyses acoustiques, nous avons ici veillé à considérer que la non significativité d'une corrélation tous locuteurs confondus peut tout autant être due à une absence de corrélation pour chacun des locuteurs qu'à des corrélations de directions opposées d'un locuteur à un autre.

VIII.2. Analyses des corrélats de l'hostilité perçue

Nous présentons ici les corrélats dans la parole de l'hostilité perçue en partant des corrélations avec des caractéristiques multiparamétriques des frontières prosodiques majeures, puis de celles avec les caractéristiques de F0 et des variables temporelles mesurées, et enfin nous présentons les corrélations avec les différents types de structuration prosodique. Cet ordre de présentation sera le même dans la description des corrélats des autres dimensions perceptives.

VIII.2.1. Hostilité perçue et frontières prosodiques majeures

Nous présentons ici les corrélations entre hostilité perçue et le nombre de frontières prosodiques majeures annotées tous paramètres confondus puis selon l'annotation du marquage prosodique par une pause, F0, et la durée¹.

	<i>Gen.</i>	VOYNET	LE	MOLOSSI	SEREY	TUAILLON	CHEQUER
Tous param	orig	.164 (NS)	.219 (NS)	-.172 (NS)	.603 (**)	.202 (NS)	-0.04 (NS)
	diff	-.069 (NS)	.037 (NS)	-.476 (*)	.108 (NS)	.265 (NS)	-.164 (NS)
pause	orig	.053 (NS)	-.022 (NS)	-.037 (NS)	.313 (NS)	-.083 (NS)	.109 (NS)
	diff	-.117 (NS)	.066 (NS)	-.564 (**)	.086 (NS)	-.091 (NS)	-.122 (NS)
F0	orig	-.048 (NS)	-.134 (NS)	-.028 (NS)	.119 (NS)	.285 (NS)	.007 (NS)
	diff	-.151 (NS)	-.214 (NS)	-.237 (NS)	.128 (NS)	.341 (NS)	-.348 (NS)
durée	orig	-.016 (NS)	.201 (NS)	-.341 (NS)	.477 (NS)	.329 (NS)	-.312 (NS)
	diff	-.168 (NS)	.066 (NS)	-.446 (*)	.023 (NS)	.203 (NS)	-.265 (NS)

Tableau VIII—1 : Corrélations entre dominance perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par une pause, marquées par F0, et par la durée. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif. Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus.

Si l'on analyse les corrélations entre hostilité perçue et nombre et types de frontières prosodiques majeures, on ne constate aucune tendance significative tous locuteurs confondus.

¹ Pour rappel chaque critère est annoté de façon non exclusive, ceci signifie pour une frontière marquée par F0 par exemple que celle-ci peut également être marquée par la durée et/ou une pause.

Locuteur par locuteur, on observe toutefois des corrélations significatives pour Molossi et Le Chequer.

Pour Molossi, l'hostilité perçue est positivement corrélée au nombre de frontières prosodiques majeures tous paramètres confondus ($\rho=.603$, $p<.01$) : les stimuli pour lesquels ce locuteur est perçu comme hostile sont ceux pour lesquels il produit le plus frontières prosodiques majeures.

Pour Le Chequer, c'est au contraire la réduction du nombre de frontières prosodiques majeures ($\rho=-.476$, $p<.01$) et qui plus est de celles marquées par une pause « frontière » ($\rho=-.564$, $p<.01$) qui est corrélée à l'hostilité perçue. Nous précisons que pour ce dernier ces résultats sont ici obtenus en confrontant l'hostilité perçue aux mesures de différences entre nombres calculés dans ses extraits originaux et ses extraits relus : c'est lorsque celui-ci produit moins de frontières prosodiques majeures et moins de pauses « frontières » par stimulus en comparaison à ses caractéristiques en relecture qu'il est perçu comme le plus hostile.

De plus, pour Le Chequer l'hostilité perçue est également corrélée à une réduction du nombre de frontières prosodiques majeures marquées par la durée ($\rho=-.446$, $p<.05$), toujours ici en comparaison à ses caractéristiques mesurées en condition de relecture.

VIII.2.2. Hostilité perçue et caractéristiques de F0

L'hostilité perçue est corrélée à l'accroissement de F0 moyenne pour deux locuteurs sur cinq, Voynet ($\rho=.414$, $p<.05$) et Molossi ($\rho=.509$, $p<.01$). Comme l'illustre la figure ci-dessous, pour ces deux locuteurs, les stimuli perçus comme maximalement hostiles sont également ceux pour lesquels on mesure la F0 moyenne la plus haute.

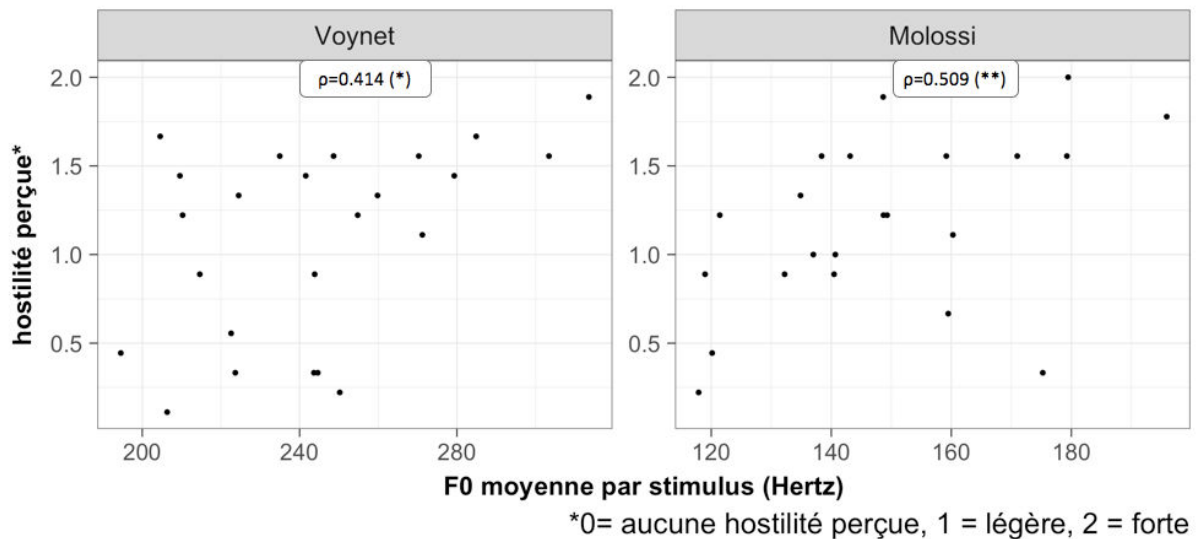


Figure VIII-1 Corrélations de Spearman entre hostilité perçue et F0 moyenne sur le corpus original (Hertz) pour Voynet et Molossi. Les étoiles représentent le taux de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

L'hostilité perçue n'est significativement corrélée à la plage de variation moyenne de F0 que pour les stimuli de Voynet ($\rho=.391$, $p<.05$).

Enfin, pour un locuteur, Serey, l'hostilité perçue est significativement corrélée aux différences de variations de plage de F0 d'un bloc inter-pausal à l'autre ($\rho=-.449$, $p<.05$). Pour ce locuteur, les stimuli perçus comme les plus hostiles sont ceux pour lesquels on mesure la plus forte réduction moyenne de la plage de variation de F0 d'un bloc inter-pausal au suivant.

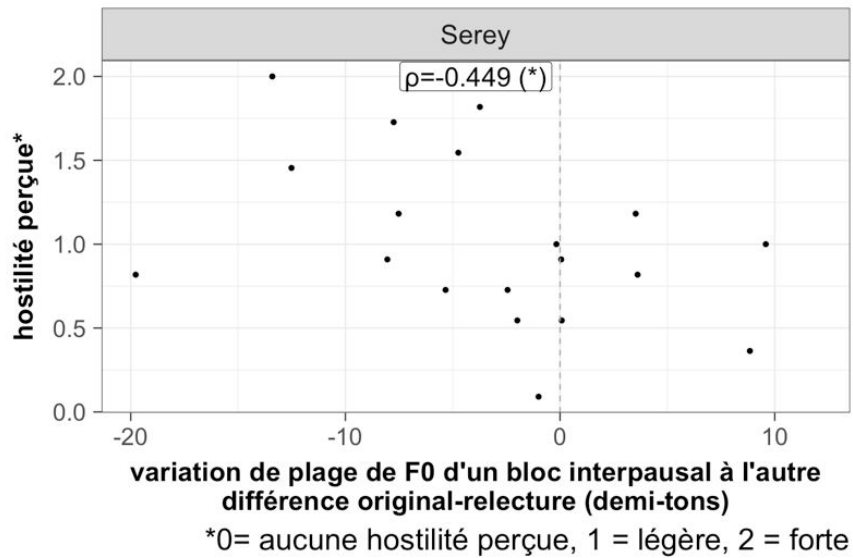


Figure VIII-2 : Corrélations entre hostilité perçue et différences original-relecture de la variation de plage de F0 d'un bloc inter-pausal à l'autre par stimulus pour Serey. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VIII.2.3. Hostilité perçue et caractéristiques temporelles

Concernant les variables temporelles issues de l'analyse prosodique, on trouve des corrélations significatives entre débit articulatoire et hostilité perçue pour Tuailleon et Voynet mais celles-ci sont de directions différentes : cette corrélation est positive pour Tuailleon ($\rho = .468$, $p < .05$) et négative pour Voynet ($\rho = -.400$, $p < .05$). Pour cette dernière, comme on peut l'observer sur la figure ci-dessous, la force de la corrélation est atténuée par les deux stimuli pour lesquels on mesure le débit articulatoire le plus élevé.

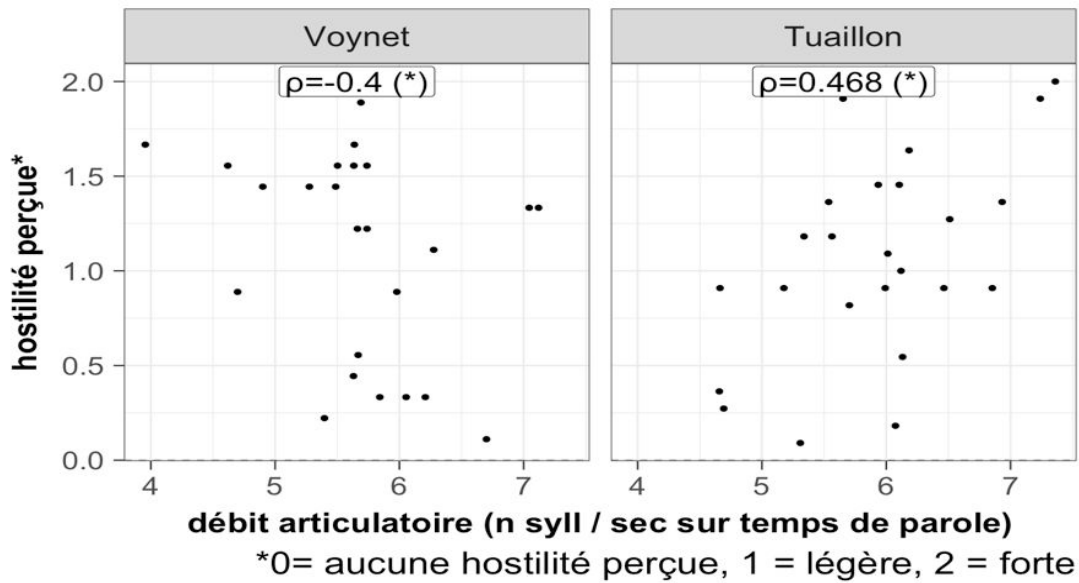


Figure VIII-3 : Corrélations entre hostilité perçue et débit articulaire pour Voynet et Tuaille. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Si l'on ne trouve pas de corrélation avec le débit articulaire moyen par stimulus pour Le Chequer, pour ce locuteur l'hostilité perçue est significativement corrélée avec l'écart-type de débit articulaire ($\rho = -.622$, $p < .01$).

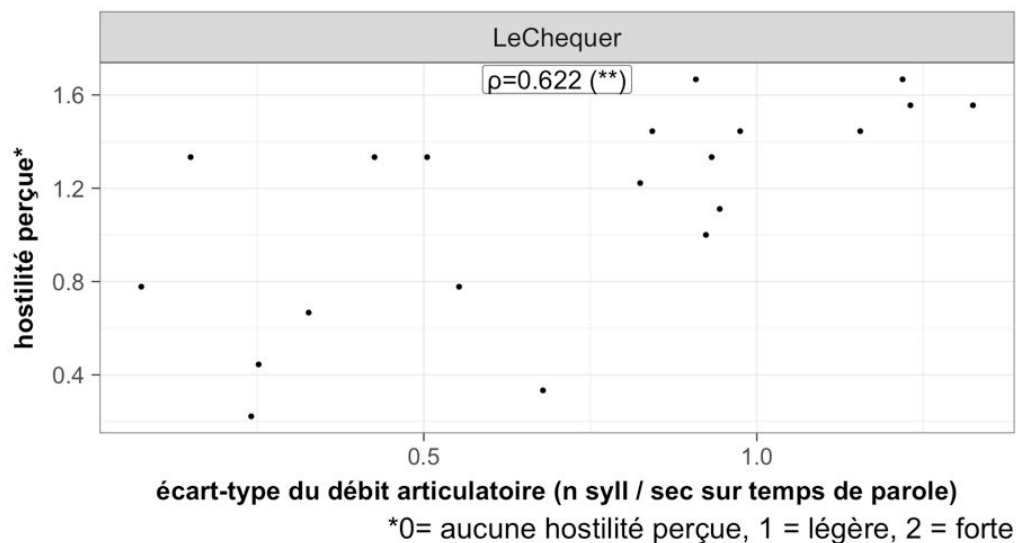


Figure VIII-4 : Corrélations entre hostilité perçue et écart-type du débit articulaire pour Le Chequer. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

Le Chequer est également le seul locuteur pour lequel on trouve une corrélation significative entre hostilité perçue et variation de débit articulaire de bloc à bloc inter-pausal ($\rho = -.421$, $p < .05$). Chez ce locuteur, les stimuli perçus comme exprimant le plus haut degré d'hostilité sont également ceux pour lesquels le débit articulaire pré-pausal est plus élevé que dans le bloc inter-pausal suivant, comme on peut l'observer sur la figure ci-dessous.

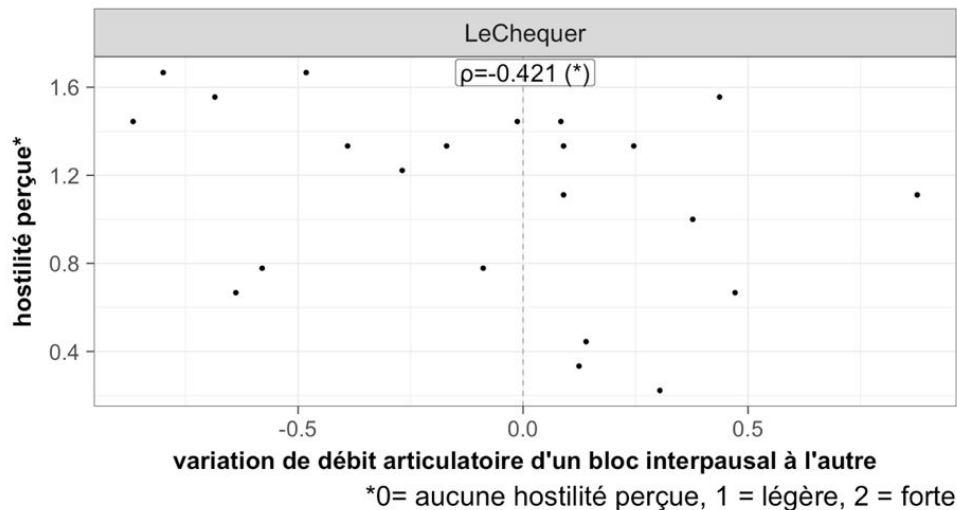


Figure VIII-5 : Corrélations entre hostilité perçue et variation de débit articulaire d'un bloc inter-pausal à l'autre. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

On trouve également pour Tuailon, et pour ce locuteur seulement, une corrélation significative entre réduction de l'index de variabilité paire à paire (IVP) de la durée des syllabes et accroissement de l'hostilité perçue ($\rho = -.425$, $p < .05$), comme l'illustre la figure ci-dessous.

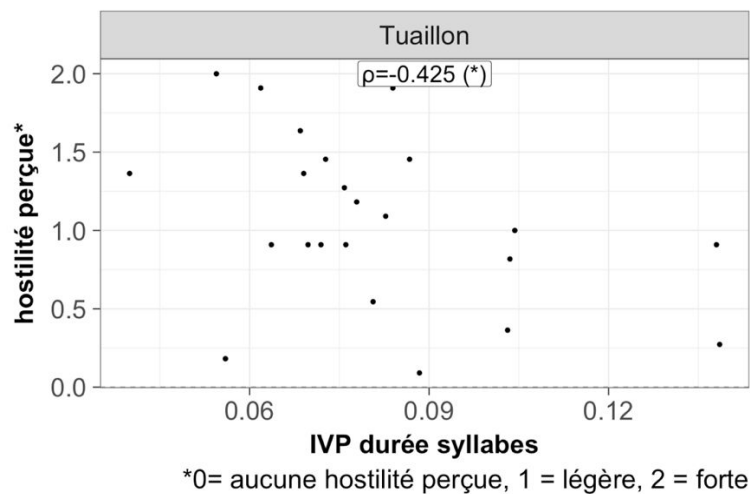


Figure VIII-6 Corrélations entre hostilité perçue et index de variabilité de durée syllabique (IVP) de durée syllabique pour Tuailon. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

On trouve également une corrélation avec l'hostilité perçue pour Molossi, mais pour celui-ci ce n'est pas l'IVP moyen des syllabes du corpus original qui est négativement corrélé à l'hostilité perçue mais la variation d'IVP d'un bloc inter-pausal à l'autre ($\rho = -.439$, $p < .05$).

Une telle corrélation demanderait une investigation précise des stimuli car celle-ci ne nous permet pas en l'état de déterminer dans quelle mesure les variations de débit sont importantes pour les stimuli jugés les plus hostiles : ce résultat nous indique seulement que ces variations de débit diffèrent peu d'un bloc inter-pausal à l'autre.

VIII.2.4. Hostilité perçue et structuration prosodico-syntaxique

		TOUS LOC.	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
UNIT SYNT	orig	.341 (**)	.464 (*)	.306 (NS)	.197 (NS)	.618 (**)	.147 (NS)
MAJ	diff	.136 (NS)	.253 (NS)	.17 (NS)	.154 (NS)	.373 (NS)	-.145 (NS)
UNIT SYNT INT.	orig	-.022 (NS)	.143 (NS)	-.449 (*)	.161 (NS)	-.158 (NS)	-.07 (NS)
	diff	.032 (NS)	.119 (NS)	-.523 (**)	.07 (NS)	.104 (NS)	.21 (NS)
INTRA UNIT SYNT	orig	-.02 (NS)	.059 (NS)	-.236 (NS)	.138 (NS)	-0.04 (NS)	-.535 (NS)
	diff	-0.27 (*)	.095 (NS)	-.528 (*)	-.093 (NS)	-.313 (NS)	-.492 (NS)

Tableau VIII—2 Corrélations de Spearman entre hostilité perçue et nombre de frontières situées entre unités syntaxiques majeures (UnitSyntMaj), intermédiaires (UnitSyntInt) et en position intra unités syntaxiques (IntraUnitSynt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus.

Concernant les rapports entre hostilité perçue et structuration prosodico-syntaxique, nous voyons que tous locuteurs confondus, l'hostilité perçue est significativement corrélée au nombre moyen de frontières prosodiques majeures situées entre unités syntaxiques majeures ($\rho = .341$, $p < .01$). Cette corrélation est significative pour deux locuteurs sur cinq : Serey ($\rho = .618$, $p < .01$) et Voynet ($\rho = .464$, $p < .05$). Cela signifie que plus ces deux locuteurs produisent de frontières prosodiques majeures congruentes avec les unités syntaxiques majeures, plus ils sont perçus comme hostiles².

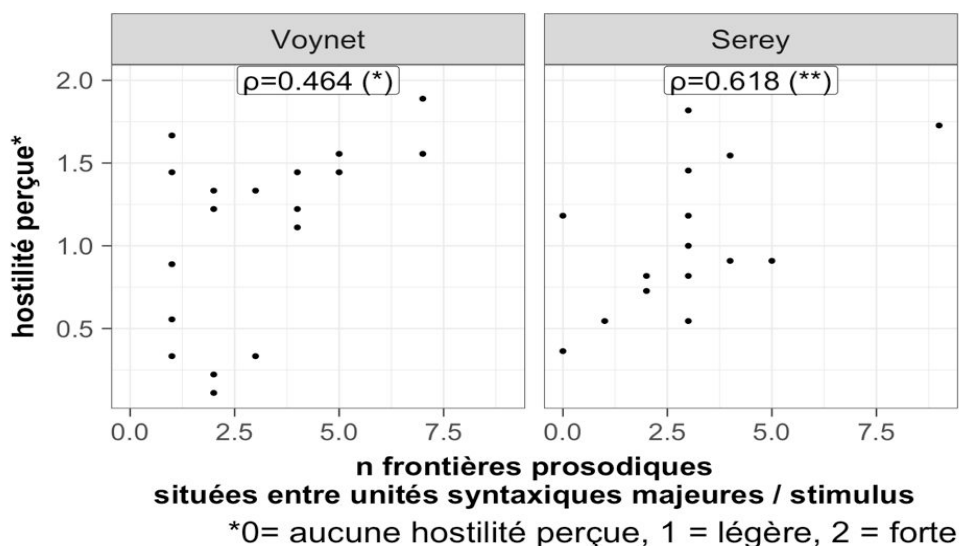


Figure VIII-7 : Corrélations entre hostilité perçue et nombre de frontières prosodiques situées entre unités syntaxiques majeures par stimulus pour Voynet et Serey. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

² Nous rappelons ici que Serey est le seul locuteur qui produit dans le corpus original des frontières prosodiques majoritairement situées entre unités syntaxiques majeures et non majoritairement entre unités syntaxiques intermédiaires comme les quatre autres locuteurs.

Les autres types de localisation syntaxique des frontières prosodiques ne semblent pas être corrélés à l'hostilité perçue. On calcule d'ailleurs une corrélation négative entre hostilité perçue et nombre de frontières prosodiques situées en position intra-syntaxiques pour un locuteur, Le Chequer. Pour ce locuteur, les stimuli perçus comme exprimant le plus d'hostilité sont ceux pour lesquels la différence de nombre de frontières prosodiques intra-syntaxiques entre original et relecture est la plus faible ($\rho = -.528$, $p < .05$). Rappelons que pour Le Chequer comme pour tous les autres locuteurs, on compte plus de frontières prosodiques situées en position intra-syntaxiques dans le corpus original qu'en relecture, et que le lien entre nombre de frontières intra- et inter-unités syntaxiques intermédiaires n'est pas direct.

VIII.3. Corrélats de la dominance perçue

VIII.3.1. Dominance perçue et frontières prosodiques majeures

Nous présentons ici les corrélations entre dominance perçue et le nombre de frontières prosodiques majeures annotées tous paramètres confondus puis selon l'annotation du marquage prosodique par une pause, F0, et la durée.

		TOUS LOC.	VOYNET	LE CHEQUER	MOLOSSI	SEREY	TUAILLON
<i>Tous param.</i>	orig	.134, (NS)	.11 (NS)	.026 (NS)	.592 (**)	.039 (NS)	-.141 (NS)
	diff	-.087 (NS)	-.032 (NS)	-.478 (*)	.263 (NS)	.081 (NS)	-.271 (NS)
pause	orig	.162 (NS)	.037 (NS)	.143 (NS)	.504 (**)	-.09 (NS)	.089 (NS)
	diff	-.03 (NS)	.025 (NS)	-.422 (*)	.423 (*)	-.154 (NS)	-.233 (NS)
F0	orig	.178 (NS)	.233 (NS)	.444 (*)	.25 (NS)	.182 (NS)	-.146 (NS)
	diff	-.055 (NS)	-.157 (NS)	.165 (NS)	.28 (NS)	.267 (NS)	-.442 (*)
durée	orig	-.031 (NS)	.271 (NS)	-.231 (NS)	-.046 (NS)	.513 (NS)	-.278 (NS)
	diff	-.249 (*)	.052 (NS)	-.54 (**)	-.256 (NS)	.285 (NS)	-.393 (NS)

Tableau VIII—3 : Corrélations entre dominance perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par une pause, marquées par F0, et par la durée. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif. Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus.

Tous paramètres et toutes localisations syntaxiques confondus, la dominance perçue est corrélée à un nombre élevé de frontières prosodiques majeures par stimulus pour Molossi ($\rho = -.592$, $p < .01$). Pour Le Chequer, on constate une corrélation contraire à travers les mesures de différence entre extraits originaux et relus : moins celui-ci produit de frontières prosodiques majeures comparé au nombre produit en lecture, plus celui-ci est perçu comme dominant ($\rho = -.478$, $p < .05$).

Les tendances observées pour ces deux locuteurs se retrouvent dans les corrélations entre dominance perçue et nombre de frontières marquées par une pause (pause « frontière »), et notamment dans la différence de nombre de celles-ci par stimulus entre productions originales et relues.

Comme on peut l'observer sur la figure ci-dessous, les stimuli évalués comme les plus dominants pour Molossi sont ceux pour lesquels il produit plus de pauses « frontières » qu'en relecture ($\rho=0.423$, $p<.05$), alors que pour Le Chequer c'est le contraire ($\rho=-0.422$, $p<.05$).

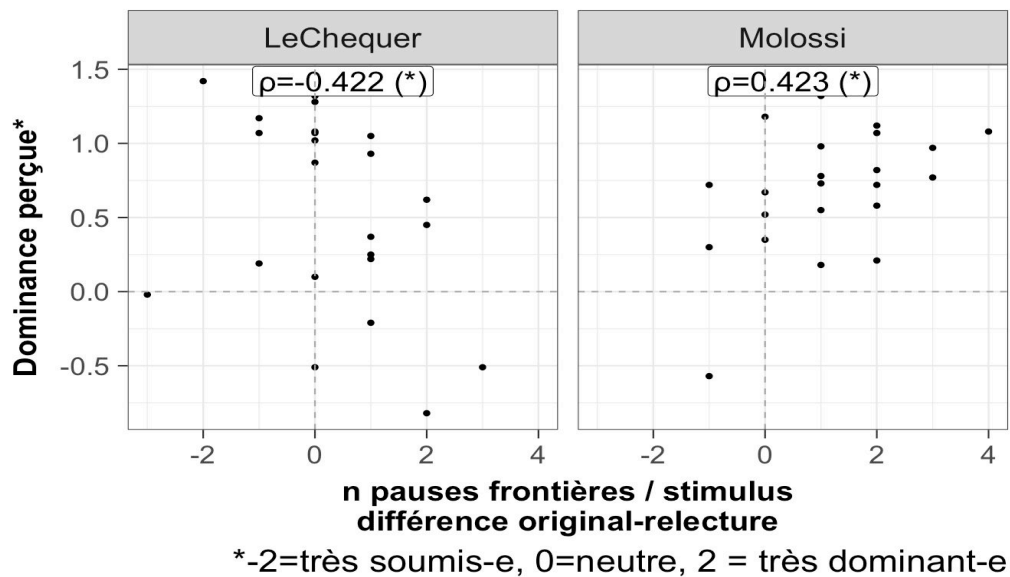


Figure VIII-8 : Corrélations entre dominance perçue et différence original-relecture du nombre de pauses « frontières » par stimulus pour Le Chequer et Molossi. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

La dominance perçue est également corrélée pour Le Chequer avec le nombre de frontières prosodiques marquées par F0 ($\rho=.444$, $p<.05$) : plus il en produit par stimulus, plus il est perçu comme dominant ; voir Figure VIII-9. De plus, pour ce locuteur les stimuli perçus comme les plus dominants sont également ceux pour lesquels il produit le moins de frontières marquées par la durée, en comparaison à la condition de contrôle relecture ($\rho=.540$, $p<.01$) ; voir Figure VIII-10.

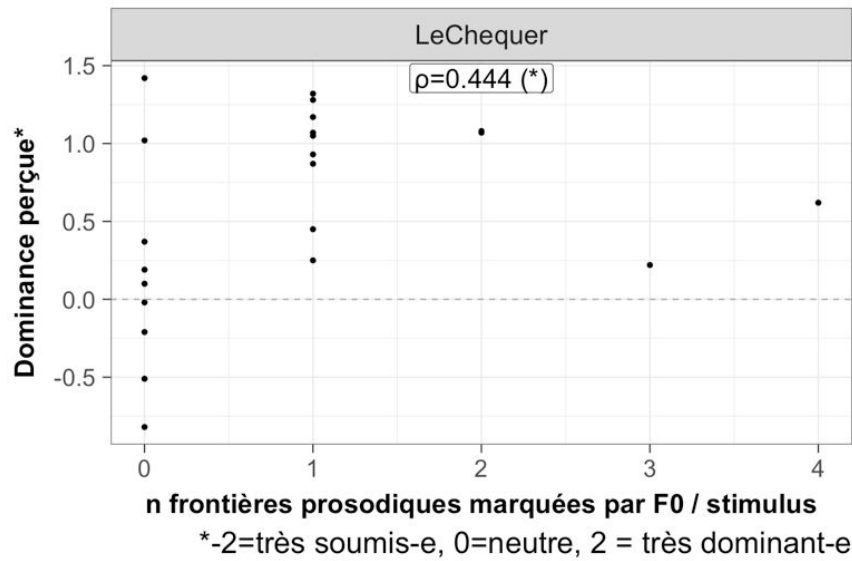


Figure VIII-9 : Corrélations entre dominance perçue et nombre de frontières prosodiques marquées par F0 pour Le Chequer. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

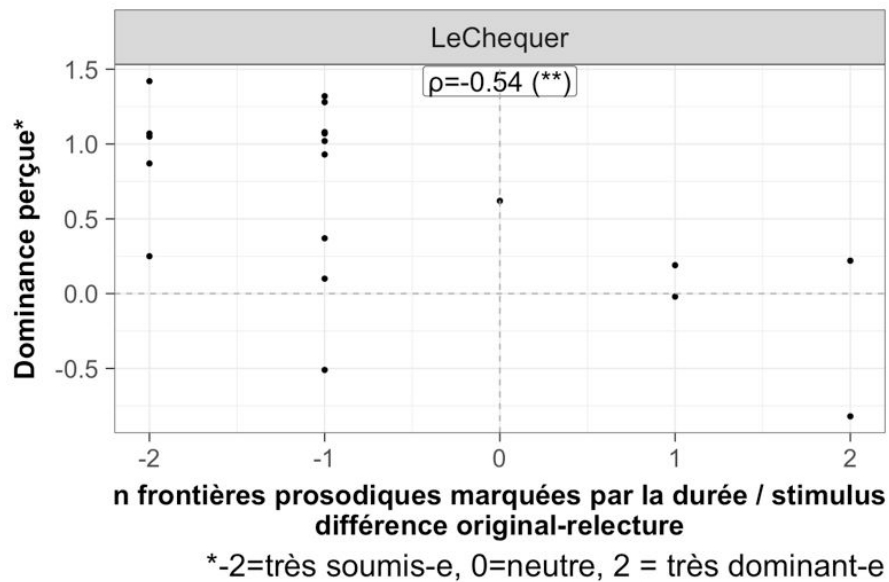


Figure VIII-10 : Corrélations entre dominance perçue et différence original-relecture du nombre de frontières prosodiques marquées par la durée par stimulus pour Le Chequer. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VIII.3.2. Dominance perçue et caractéristiques de F0

Comme pour la valence et l'excitation, la dominance perçue est significativement corrélée à un accroissement de F0 moyenne ($\rho=.211$, $p<.05$). Cette corrélation est significative pour tous les locuteurs sauf Tuailon comme l'illustre la figure ci-dessous.

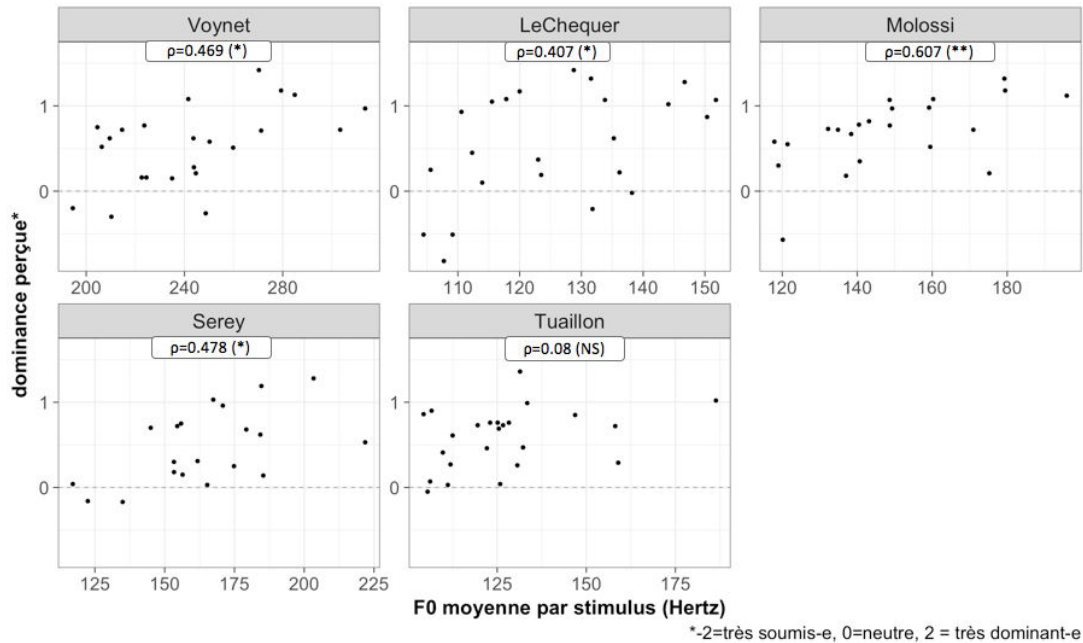


Figure VIII-11 : Corrélations entre dominance perçue et fréquence fondamentale moyenne F0 en Hertz ; échelle adaptée aux locuteurs pour plus de lisibilité. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, $p>.05 = NS$, non significatif.

VIII.3.3. Dominance perçue et contours mélodiques annotés

L'analyse des rapports entre dominance perçue et réalisation des contours mélodiques aux frontières prosodiques majeures indique une corrélation significative avec le nombre moyen de contours en cloche pour Serey ($\rho=.771$, $p<.05$) et de contours descendants pour Voynet ($\rho=.505$, $p<.05$).

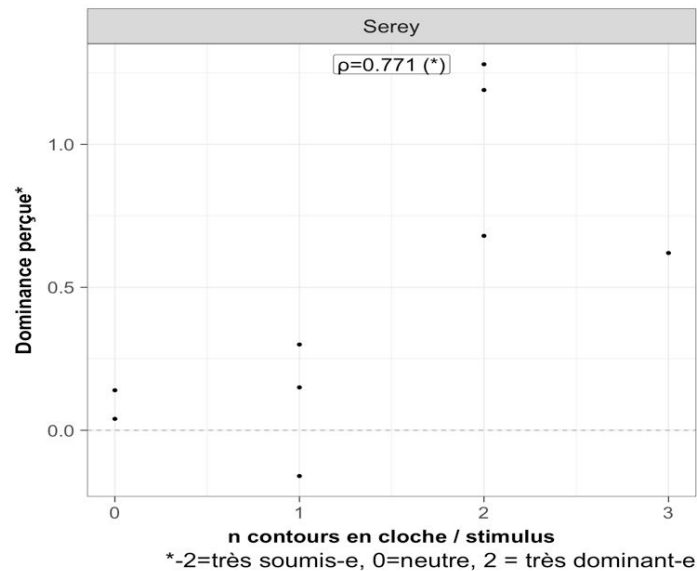


Figure VIII-12 : Corrélations entre dominance perçue et nombre moyen de contours mélodiques annotés comme en cloche pour Serey. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

Comme l'illustre la figure ci-dessous, en plus d'être corrélé à l'excitation perçue, le nombre moyen de contours mélodiques descendants dans les stimuli de Voynet est corrélé à la dominance perçue.

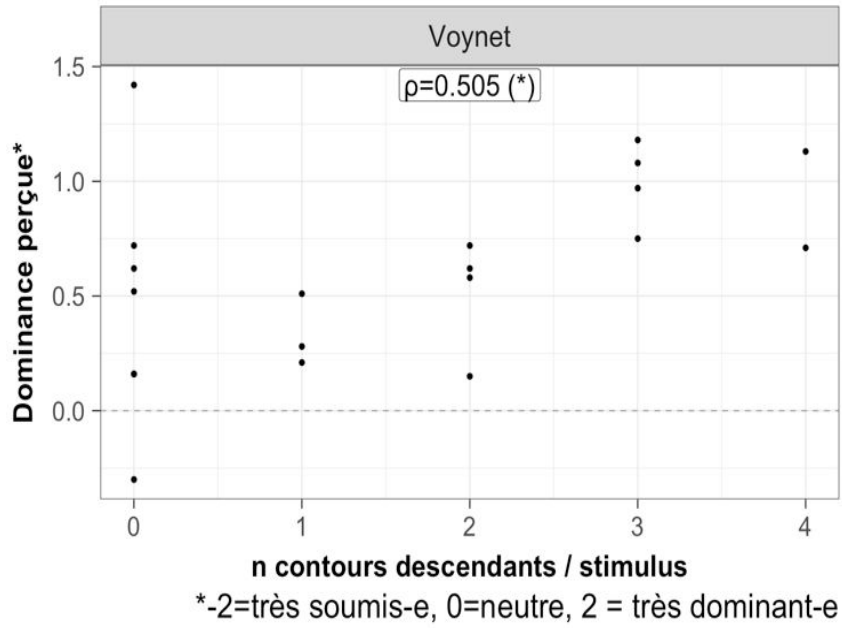


Figure VIII-13 : Corrélations entre dominance perçue et nombre moyen de contours mélodiques annotés comme descendants pour Voynet. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

VIII.3.4. Dominance perçue et caractéristiques temporelles

La dominance perçue est significativement corrélée au débit articulaire moyen pour un locuteur seulement : Tuailleon (original : $\rho=.612$, $p<.01$, différence original-relecture : $\rho=.655$, $p<.001$). On obtient également une corrélation significative entre dominance perçue et écart-type du débit articulaire original pour deux locuteurs : Voynet ($\rho=.643$, $p<.05$) et Le Chequer ($\rho=.641$, $p<.05$), voir figure ci-dessous.

Pour Voynet, les scores de dominance perçue de ses stimuli sont également corrélés à la durée de la voyelle des syllabes finales pré-pausales ($\rho=.427$, $p<.05$). Cette variable est également significativement corrélée à la dominance perçue pour Tuailleon mais, pour ce locuteur la corrélation est négative ($\rho=-.6$, $p<.01$). Il apparaît intéressant de rappeler ici que les mêmes tendances opposées entre Voynet et Tuailleon se retrouvent dans les rapports entre valence perçue et débit articulaire : pour Voynet les stimuli perçus comme les plus mécontents sont ceux pour lesquels le débit articulaire est le plus lent et pour Tuailleon le plus rapide. Pour rappel, la durée des voyelles finales pré-pausales est significativement corrélée au débit articulaire pour Tuailleon ($\rho=-.511$, $p<.01$) et non pour Voynet ($\rho=-.215$, $p=0.3$). Nous rappelons également que Voynet est la seule locutrice pour laquelle la proportion de frontières annotées comme marquées par la durée ne réduit pas entre corpus original et relecture.

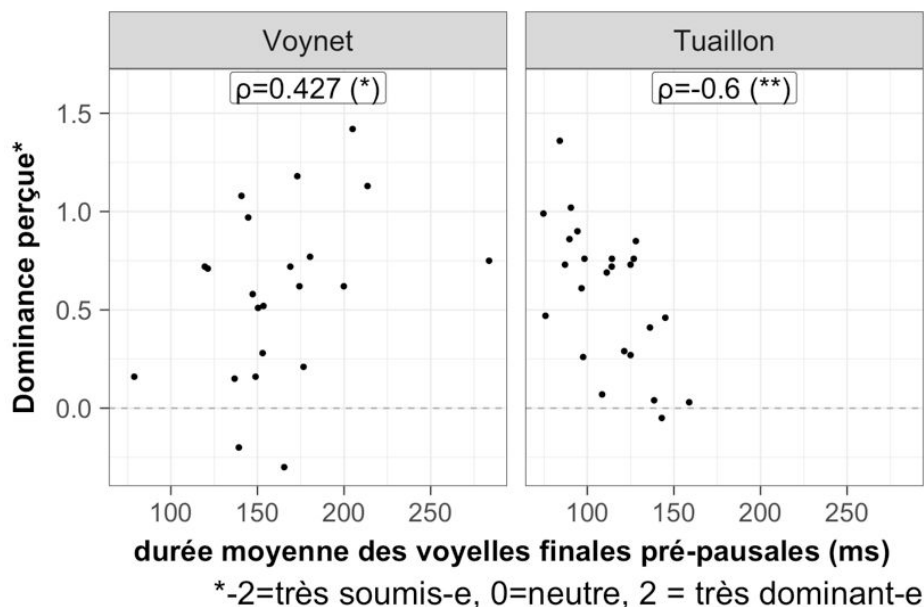


Figure VIII-14 : corrélations entre dominance perçue et durée des voyelles finales pré-pausales pour Voynet et Tuailleon. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

VIII.3.5. Dominance perçue et structuration prosodico-syntaxique

		TOUS LOC.	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
nFront <i>sMaj</i>	orig	.161 (NS)	.239 (NS)	.208 (NS)	.125 (NS)	.352 (NS)	-.079 (NS)
	diff	.135 (NS)	.306 (NS)	-.101 (NS)	.507 (*)	.288 (NS)	-.254 (NS)
nFront <i>sInt</i>	orig	.187 (NS)	.088 (NS)	-.175 (NS)	.355 (NS)	.116 (NS)	.236 (NS)
	diff	.113 (NS)	.11 (NS)	-.217 (NS)	.186 (NS)	0 (NS)	0 (NS)
nFront <i>intra-s.</i>	orig	-.335 (**)	-.382 (NS)	-.248 (NS)	-.327 (NS)	.53 (NS)	-.625 (*)
	diff	-.379 (**)	-.377 (NS)	-.564 (*)	-.408 (NS)	.58 (NS)	-.53 (NS)

Tableau VIII—4 Corrélations de Spearman entre dominance perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif.

Enfin, concernant la localisation syntaxique des frontières prosodiques (ici tous paramètres confondus), tous locuteurs confondus, ce sont les frontières prosodiques situées en position intra-syntaxique qui sont le plus corrélées à la dominance perçue, mais de façon négative (tous locuteurs confondus : $\rho = -.335$, $p < .01$). Pour Tuailon, moins ce dernier en produit par stimulus et plus il est perçu comme dominant ($\rho = -.625$, $p < .05$). La même tendance s'observe chez Le Chequer ($\rho = -.564$, $p < .05$) mais pour celui-ci à travers la mesure de différence original-relecture : les stimuli pour lesquels il produit aussi peu de frontières intra-syntaxiques qu'en lecture sont ceux qui sont perçus comme les plus dominants.

Nous constatons pour Le Chequer que la dominance perçue est également corrélée à un nombre plus élevé de frontières prosodiques situées entre unités syntaxiques majeures en condition originale qu'en condition de relecture ($\rho = .507$, $p < .05$).

VIII.5. Corrélats de la valence perçue

VIII.5.1. Valence perçue et frontières prosodiques majeures

		TOUS LOC.	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
Tous param.	orig	-.138 (NS)	-.442 (*)	-.197 (NS)	-.612 (**)	-.219 (NS)	-.313 (NS)
	diff	-.004 (NS)	.014 (NS)	.245 (NS)	-.148 (NS)	-.231 (NS)	-.036 (NS)
pause	orig	-.171 (NS)	-.184 (NS)	-.329 (NS)	-.455 (*)	.003 (NS)	-.094 (NS)
	diff	.045 (NS)	.12 (NS)	.276 (NS)	-.291 (NS)	.09 (NS)	-.002 (NS)
F0	orig	-.06 (NS)	.137 (NS)	-.241 (NS)	-.254 (NS)	-.11 (NS)	.078 (NS)
	diff	.002 (NS)	.214 (NS)	.173 (NS)	-.364 (NS)	-.161 (NS)	.188 (NS)
durée	orig	.025 (NS)	-.176 (NS)	.223 (NS)	-.023 (NS)	-.07 (NS)	.353 (NS)
	diff	.209 (NS)	.031 (NS)	.41 (NS)	.132 (NS)	.091 (NS)	.257 (NS)

Tableau VIII—5 : Corrélations entre valence perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par F0, et par la durée. Orig = décompte par stimulus dans le corpus original, diff. = différence stimulus par stimulus entre décompte dans le corpus original et dans la relecture. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif.

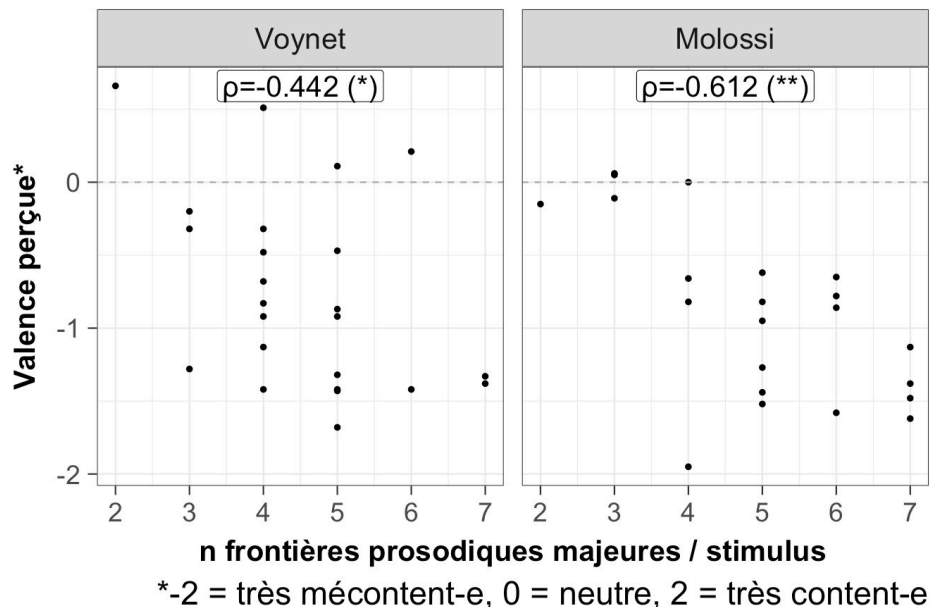


Figure VIII-15 : Corrélations entre valence perçue et nombre de frontières prosodiques majeures (tous paramètres confondus) pour Voynet et Molossi. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif.

En ce qui concerne la valence, comme on peut l'observer sur la figure ci-dessus, pour Voynet ($\rho = -.442$, $p < .05$) et pour Molossi ($\rho = -.612$, $p < .01$), plus ceux-ci produisent de frontières prosodiques (tous paramètres confondus) plus ils sont perçus comme mécontents. Pour Molossi cette corrélation se retrouve également spécifiquement pour les frontières prosodiques marquées par une pause ($\rho = -.455$, $p < .05$) ce qui signifie que plus celui-ci produit de pauses « frontières », plus il est perçu comme mécontent. Comme on peut l'observer dans le tableau ci-dessus, ceci est la seule corrélation significative entre valence perçue et nombre moyen par stimulus de frontières prosodiques marquées par une pause.

VIII.5.2. Valence perçue et structuration prosodico-syntaxique

		TOUS LOC.	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
nFront sMaj	orig	-.342 (**)	-.447 (NS)	-.277 (NS)	-.054 (NS)	-.561 (*)	-.148 (NS)
	diff	-.22 (*)	-.466 (*)	-.132 (NS)	-.354 (NS)	-.202 (NS)	.276 (NS)
nFront sInt	orig	-.083 (NS)	-.221 (NS)	.131 (NS)	-.266 (NS)	.029 (NS)	-.077 (NS)
	diff	-.085 (NS)	-.051 (NS)	.194 (NS)	-.148 (NS)	-.368 (NS)	-.202 (NS)
nFront intra-s.	orig	.272 (**)	.414 (*)	.142 (NS)	.03 (NS)	.403 (NS)	.191 (NS)
	diff	.336 (**)	.371 (NS)	.326 (NS)	.216 (NS)	.414 (NS)	.148 (NS)

Tableau VIII—6 Corrélations de Spearman entre valence perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

Tous locuteurs confondus, la valence perçue est négativement corrélée au nombre moyen de frontières prosodiques majeures situées entre unités syntaxiques majeures ($\rho = -.342$, $p < .01$) et positivement au nombre de frontières intra-syntaxiques ($\rho = .272$, $p < .01$). Ces corrélations générales s'expliquent par les corrélations observées pour Serey et Voynet.

Pour Serey, les stimuli pour lesquels il est perçu comme le plus mécontent sont ceux pour lesquels il produit le moins de frontières prosodiques entre unités syntaxiques majeures. Nous rappelons ici que ce locuteur est celui qui produit à l'échelle du corpus original le plus de frontières prosodico-syntaxiques de ce type. La corrélation est donc négative ($\rho = -.561$, $p < .05$).

Pour Voynet, sa production de frontières prosodiques intra-syntaxiques est positivement corrélée à la valence ($\rho = -.561$, $p < .05$), cela signifie que les stimuli pour lesquels elle produit le plus de frontières prosodico-syntaxiques de ce type sont ceux pour lesquels elle est perçue comme la plus contente, ou en d'autres termes comme la moins mécontente.

VIII.5.3. Valence perçue et caractéristiques de F0

Concernant les rapports entre valence perçue et F0, tous locuteurs confondus, plus ceux-ci sont perçus comme mécontents, plus leur F0 moyenne par stimulus est haute dans le corpus original et plus celle-ci diffère des mesures effectuées dans le corpus relecture (corpus original : $\rho = -0.329$, $p < .01$, différence : $\rho = -0.374$, $p < .01$). Locuteur par locuteur, la valence perçue est corrélée à la F0 moyenne pour trois locuteurs sur cinq : Voynet ($\rho = -0.621$, $p < .001$), Le Chequer ($\rho = -0.436$, $p < .05$) et Molossi ($\rho = -0.507$, $p < .01$). Pour ces trois locuteurs, le mécontentement perçu est corrélé à l'accroissement de leur F0 moyenne.

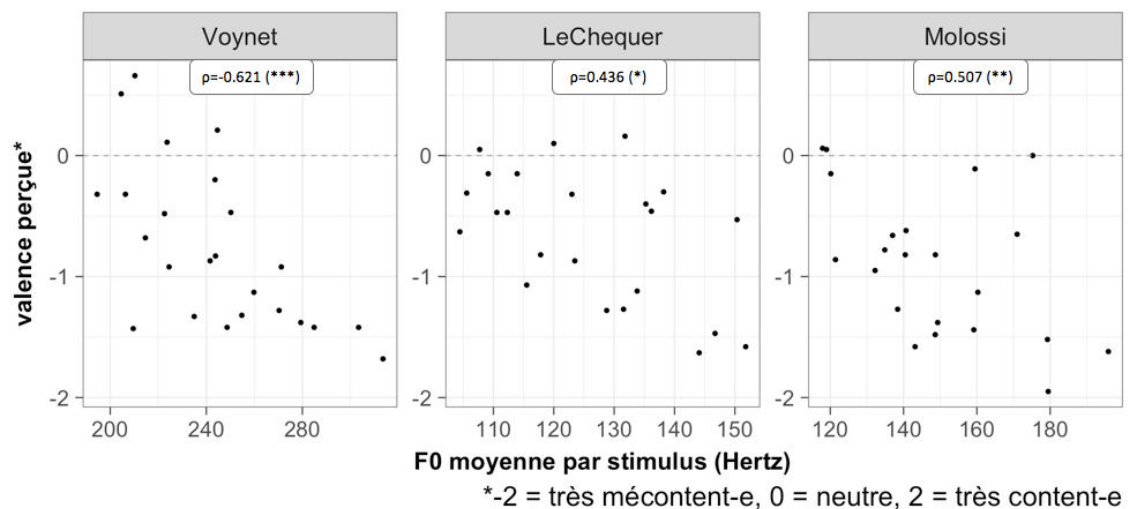


Figure VIII-16 : Corrélations de Spearman entre valence perçue et F0 moyenne sur le corpus original (Hertz) pour Voynet, Le Chequer et Molossi. Les étoiles représentent le taux de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

S'agissant des différences inter-blocs en valeur absolue de F0 moyenne dans le corpus original, on ne trouve de corrélation significative avec la valence que pour Voynet ($\rho = -0.753$, $p < .001$) et Le Chequer ($\rho = -0.527$, $p < .01$). Pour ces deux locuteurs, les stimuli perçus comme exprimant le plus de mécontentement (valence négative) sont ceux pour lesquels on trouve les plus fortes variations de F0 moyenne d'un bloc inter-pausal à l'autre dans les stimuli originaux.

VIII.5.4. Valence perçue et contours mélodiques annotés

L'analyse des rapports entre type de contours mélodiques nous indique seulement une corrélation significative entre valence perçue et nombre de contours en cloche, pour Serey ($\rho = -0.788$, $p < .01$). Plus Serey produit de contours en cloche plus il est perçu comme mécontent ; voir illustration ci-dessous.

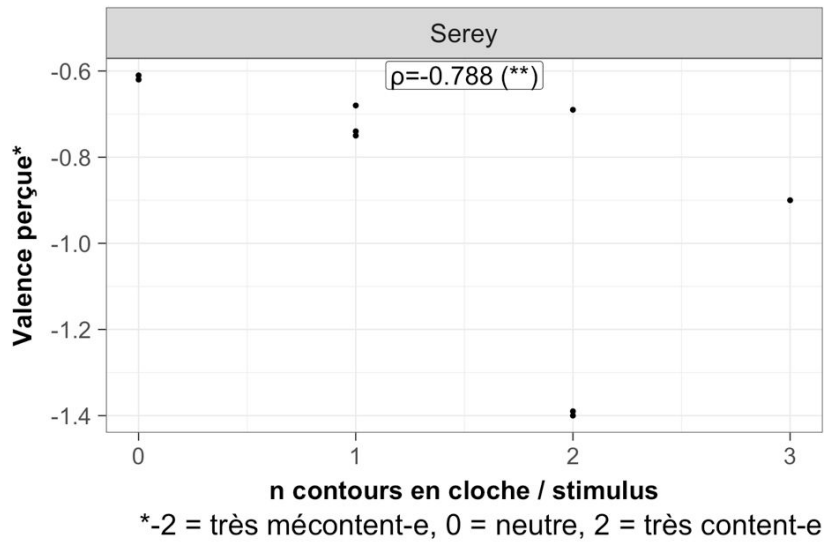


Figure VIII-17 : Corrélation entre valence perçue et contours manuellement annotés comme en cloche pour Serey. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

VIII.5.5. Corrélat temporels de la valence perçue

Tous locuteurs confondus, la valence perçue est corrélée avec la différence de débit articulatoire entre original et relecture ($\rho = -.181$, $p < .05$) bien qu'il n'y ait qu'un locuteur pour lequel on trouve de corrélation significative : pour Le Chequer ($\rho = -.455$, $p < .05$), les stimuli perçus comme exprimant le plus de mécontentement sont ceux pour lesquels son débit articulatoire est plus élevé qu'il ne l'est en condition de relecture.

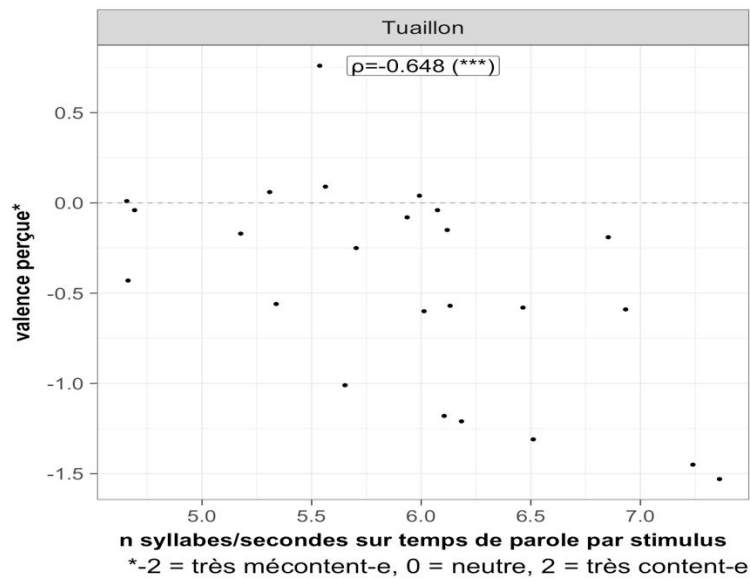


Figure VIII-18 : Corrélation entre valence perçue et débit articuloire moyen (n syllabes / secondes sur temps de parole par stimulus) pour Tuailon. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

Si on considère à présent les corrélations significatives uniquement locuteur par locuteur, la corrélation la plus forte concerne le débit articuloire moyen pour Tuailon ($\rho = -.648$, $p < .001$) ; cette corrélation est d'ailleurs la corrélation la plus forte trouvée pour la valence perçue tous paramètres confondus : plus ce locuteur est perçu comme mécontent, plus son débit articuloire est élevé. On observe également pour les stimuli de Tuailon une corrélation négative avec la durée des voyelles finales de blocs inter-pausaux, plus celles-ci sont courtes plus il est perçu comme mécontent ($\rho = .468$, $p < .05$).

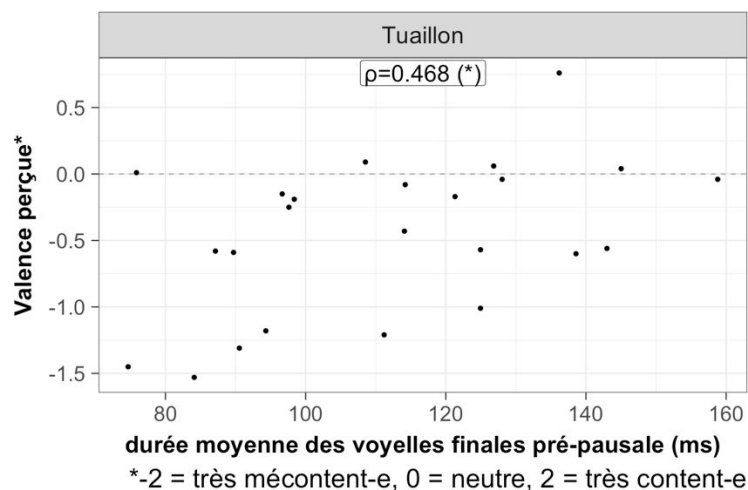


Figure VIII-19 : Corrélation entre valence perçue et durée moyenne des voyelles finales pré-pausales pour Tuailon. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

VIII.6. Corrélats de l'excitation perçue

VIII.6.1. Excitation perçue et frontières prosodiques majeures

		TOUS LOC.	VOYNET	LE CHEQUER	MOLOSSI	SEREY	TUAILLON
Tous param.	orig	-.138 (NS)	.092 (NS)	-.014 (NS)	.423 (*)	.345 (NS)	.056 (NS)
	diff	-.007 (NS)	-.197 (NS)	-.412 (*)	.083 (NS)	.396 (NS)	.083 (NS)
pause	orig	.031, (NS)	-.05 (NS)	.003 (NS)	.147 (NS)	-.079 (NS)	.145 (NS)
	diff	-.138 (NS)	-.228 (NS)	-.410 (*)	.001 (NS)	-.128 (NS)	.002 (NS)
F0	orig	-.002 (NS)	-.202 (NS)	.161 (NS)	.101 (NS)	-.202 (NS)	-.017 (NS)
	diff	-.057 (NS)	-.518 (**)	-.233 (NS)	.089 (NS)	-.083 (NS)	-.199 (NS)
durée	orig	.057, (NS)	.134 (NS)	-.029 (NS)	.267 (NS)	-.069 (NS)	-.329 (NS)
	diff	-.102 (NS)	-.116 (NS)	-.224 (NS)	-.057 (NS)	-.169 (NS)	-.237 (NS)

Tableau VIII—7 : Corrélations entre excitation perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par F0, et par la durée. Orig = décompte par stimulus dans le corpus original, diff. = différence stimulus par stimulus entre décompte dans le corpus original et dans la relecture. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif.

L'excitation est positivement corrélée au nombre de frontières prosodiques tous paramètres confondus pour Molossi ($\rho=.423$, $p<.05$). Pour Le Chequer la tendance est inverse. Comme on peut le visualiser sur la figure ci-dessous, les stimuli pour lesquels il est perçu comme le plus calme sont ceux au sein desquels il produit plus de frontières prosodiques majeures tous paramètres confondus qu'en condition de relecture ($\rho=-.412$, $p<.05$), cette corrélation se retrouvant également pour ce locuteur avec le nombre de frontières prosodiques marquées par une pause ($\rho=-.412$, $p<.01$).

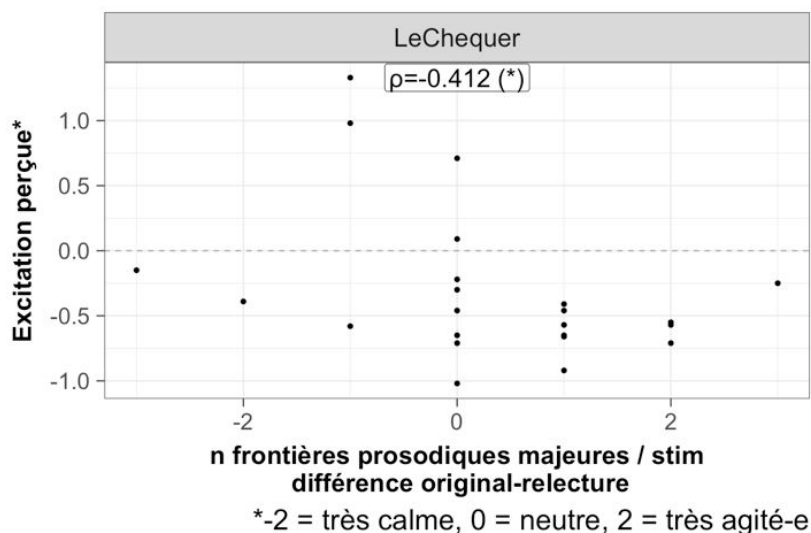


Figure VIII-20 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen par stimulus de frontières prosodiques majeures pour Le Chequer. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

Pour Voynet, les stimuli pour lesquels elle produit autant ou plus de frontières prosodiques marquées par F0 qu'en relecture sont ceux pour lesquels elle est perçue comme la plus calme ($\rho = -.518$, $p < .01$), et à l'inverse ceux pour lesquels elle en produit moins sont perçus comme exprimant le plus d'excitation.

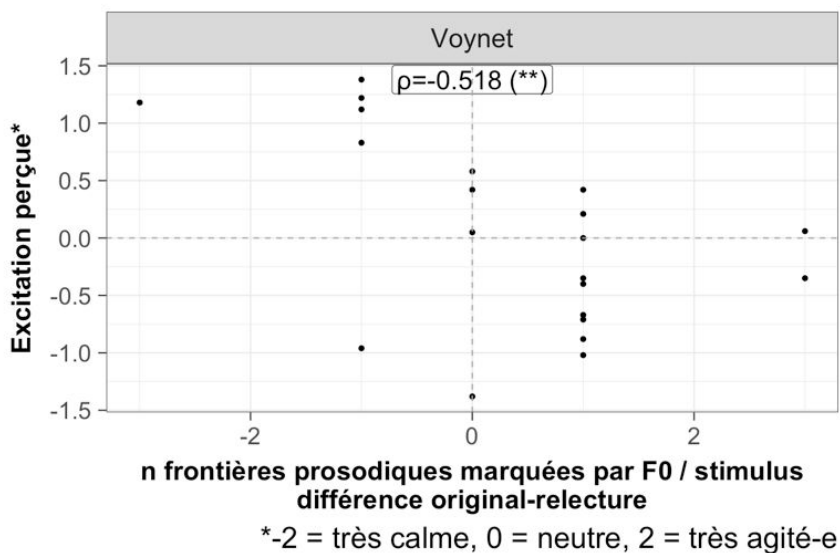


Figure VIII-21 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen par stimulus de frontières prosodiques majeures marquées par F0 pour Voynet. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.

VIII.6.2. Excitation perçue et caractéristiques de F0

Tous locuteurs confondus, les corrélats principaux de l'excitation perçue sont principalement d'ordre mélodique. L'excitation perçue est corrélée à l'accroissement de la F0 moyenne ($\rho = .306, p < .01$), et ceci est significatif pour tous les locuteurs sauf Tuailleon comme l'indique la figure ci-dessous.

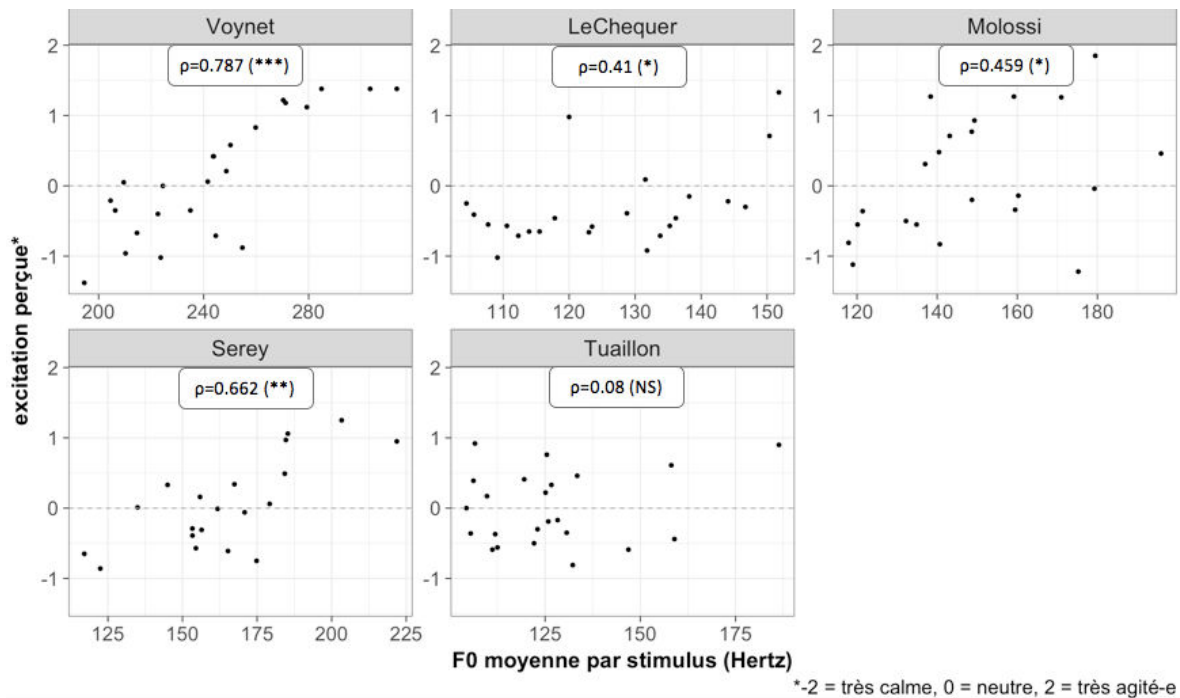


Figure VIII-22 : Corrélations entre excitation perçue et F0 moyenne (Hz). Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif.

On trouve également une corrélation significative avec les variations de plage de F0 d'un bloc inter-pausal à l'autre dans le corpus original ($\rho = -.311, p < .01$) : tous locuteurs confondus, les stimuli recevant les scores d'excitation perçue les plus élevés sont ceux pour lesquels on mesure le moins de variation de plage de F0 d'un bloc à l'autre.

VIII.6.3. Excitation perçue et contours mélodiques annotés

L'analyse des rapports entre excitation perçue et réalisation des contours mélodiques aux frontières prosodiques majeures indique une corrélation significative entre excitation perçue et nombre de contours descendants pour Voynet. Comme l'illustre la figure ci-dessous, plus Voynet produit en moyenne de contours descendants plus elle est perçue comme agitée ($\rho=.574$, $p<.01$).

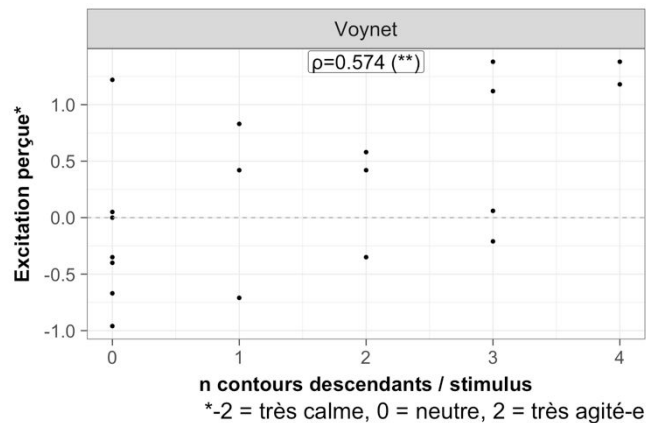


Figure VIII-23 : Corrélations entre excitation perçue et nombre moyen de contours annotés comme descendants dans le corpus original pour Voynet. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

Concernant Le Chequer, les stimuli pour lesquels il est perçue comme le plus agité sont ceux pour lesquels il produit le moins de contours montants en comparaison à la condition de contrôle ($\rho=-.451$, $p<.05$).

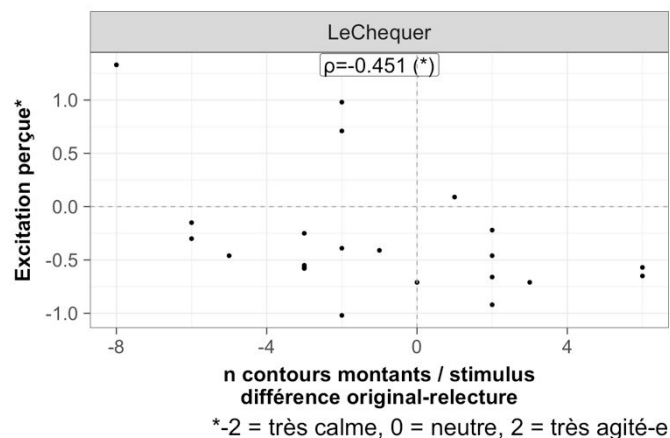


Figure VIII-24 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen de contours mélodiques montants par stimulus pour Le Chequer. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

VIII.6.4. Excitation perçue et caractéristiques temporelles

Concernant les rapports avec les variables temporelles, on remarque que l'excitation perçue est significativement corrélée avec le débit articulaire moyen ($\rho=.222$, $p<.05$), cette corrélation étant significative pour deux locuteurs sur cinq : Le Chequer ($\rho=.448$, $p<.05$) et Tuillon ($\rho=.434$, $p<.05$). Pour Le Chequer, l'excitation perçue est également corrélée à la différence de débit articulaire moyen entre original et relecture ($\rho=.504$, $p<.01$).

L'analyse des corrélations locuteur par locuteur indique que la durée de la voyelle des syllabes finales pré-pausales est également significativement corrélée à l'excitation perçue pour Le Chequer ($\rho=-.521$, $p<.01$) et Tuillon ($\rho=-.447$, $p<.05$). Comme l'indique la figure ci-dessous, plus les voyelles finales pré-pausales sont courtes plus ces deux locuteurs sont perçus comme agités.

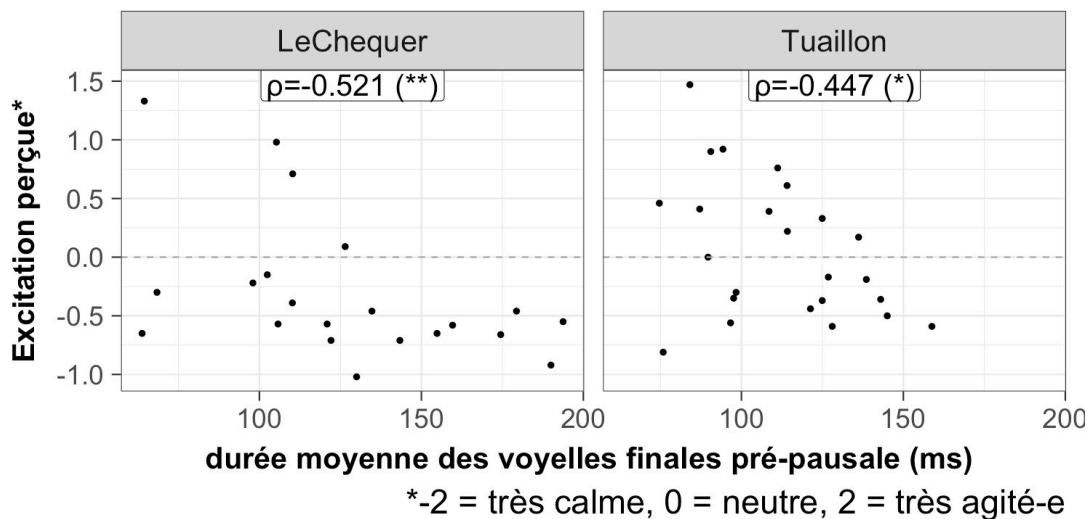


Figure VIII-25 : Corrélations entre excitation perçue et durée moyenne des voyelles des syllabes finales pré-pausales pour Le Chequer et Tuillon. Les étoiles représentent la valeur de p : $p<.001 = ***$, $p<.01 = **$, $p<.05 = *$, NS = non significatif.

VIII.6.5. Excitation perçue et structuration prosodico-syntaxique

		TOUS LOC.	VOYNET	LECHEQUER	MOLOSSI	SEREY	TUAILLON
nFront <i>sMaj</i>	orig	.232 (*)	.45 (NS)	-.262 (NS)	.234 (NS)	.303 (NS)	.348 (NS)
	diff	.127, (NS)	.345 (NS)	-.264 (NS)	.142 (NS)	.167 (NS)	-.033 (NS)
nFront <i>sInt</i>	orig	-.033, (NS)	-.129 (NS)	-.172 (NS)	-.013 (NS)	-.228 (NS)	.042 (NS)
	diff	-.065, (NS)	-.228 (NS)	-.58 (**)	.016 (NS)	.033 (NS)	.239 (NS)
nFront <i>intra-s.</i>	orig	-.081, (NS)	-.185 (NS)	.023 (NS)	-.248 (NS)	0 (NS)	-.383 (NS)
	diff	-.117, (NS)	-.207 (NS)	-.123 (NS)	-.355 (NS)	-.058 (NS)	-.291 (NS)

Tableau VIII—8 Corrélations de Spearman entre excitation perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. Les étoiles représentent la valeur de p : p<.001 = ***, p<.01 = **, p<.05 = *, NS = non significatif.

Comme on peut l'observer dans le tableau ci-dessus, la seule corrélation significative entre excitation perçue et localisation syntaxique des frontières prosodiques concerne Le Chequer ($\rho = -.580$, $p < .01$). Pour ce locuteur, moins il produit de frontières prosodiques entre unités syntaxiques intermédiaires par rapport à la contrepartie relue de ses stimuli plus il est perçu comme agité. Cette tendance permet indirectement de constater qu'il y a en moyenne une réduction du marquage prosodique des frontières syntaxiques intermédiaires pour les stimuli perçus comme exprimant le plus d'excitation chez Le Chequer.

IX. Discussion et Conclusion

Dans ce chapitre, à l'issue de ce travail de recherche doctoral, nous essayons de déterminer les apports et limites du protocole expérimental mis en place et de livrer quelques réflexions sur la continuation possible de l'analyse de l'expressivité vocale.

Nos trois objectifs complémentaires de départ, énoncés en introduction, étaient les suivants :

1. Analyser l'impact de l'expressivité vocale sur la perception de la dominance dans des extraits de parole politique en débat : nous discuterons donc principalement de la dimension perceptive de notre protocole (approche dimensionnelle et auto-évaluation).
2. Fournir un protocole méthodologique adapté à l'analyse phonétique appliquée de parole expressive naturelle : nous discuterons des apports et limites de notre condition de contrôle.
3. Apporter une continuité aux analyses de Danielle Duez (1991) sur les rapports entre pouvoir et pauses silencieuses, notamment au niveau perceptif et au niveau acoustique en analysant les corrélats mélodiques et temporels des pauses (débit articulatoire, allongements). Concernant ce troisième objectif, nous effectuerons un retour critique sur les analyses prosodiques effectuées. Nous nous efforcerons de fournir une synthèse de nos recommandations pour un protocole pluriel et pluridisciplinaire pour une analyse plus approfondie, notamment en perception, des fonctions paralinguistiques de la pause silencieuse.

IX.1. Expressivité vocale et perception des personnes politiques en débat

Hostilité, dominance perçue et leurs corrélats de parole : En ce qui concerne la question de l'influence de l'expressivité vocale sur la perception de la dominance, notre étude confirme les hypothèses de Poggi & D'Errico (2010) quant à l'influence d'expressions d'attitudes agressives contrôlées dans la communication d'une image de soi (*ethos*) favorable.

Dans notre étude, tous les locuteurs ont été perçus comme légèrement dominants. Les principaux corrélats de parole de la dominance et de l'hostilité perçue dans nos données nous rappellent en de nombreux points le ton « menaçant » décrit par Fónagy (1983). Selon Fónagy, le ton « menaçant » consiste en un débit articulatoire élevé, une faible variation de durées syllabiques et un faible allongement final, et des écarts marqués de la plage de variation de F0 de part et d'autre des pauses avec une prédilection pour des contours mélodiques plats. Dans notre étude, la plage de variation de F0 à la fois réduite et fortement variable de part et d'autre des pauses est le corrélat de la dominance et de l'hostilité le plus commun pour tous les locuteurs sauf Dominique Voynet.

En effet, Dominique Voynet est la seule locutrice pour laquelle la dominance perçue n'est pas significativement corrélée à l'hostilité perçue : pour les quatre autres locuteurs, les extraits dans lesquels ceux-ci sont perçus comme les plus hostiles sont également ceux pour lesquels ils sont perçus comme les plus dominants. Si notre corpus ne permet en aucun cas d'émettre d'hypothèses sur l'influence du genre du locuteur à ce sujet, la question reste ouverte : les différences entre Voynet et les quatre autres locuteurs seraient-ils plus à mettre sur le compte de son statut de femme ou de son statut de Maire ? Nous pouvons effectivement également considérer l'influence de la différence de rôles et d'enjeux entre Dominique Voynet et les quatre membres de l'opposition étudiés en prenant en compte le contexte préélectoral dans lequel se sont déroulés les débats¹.

La corrélation positive entre dominance et hostilité perçue pour les locuteurs masculins de notre corpus rappelle les conclusions de la thèse de Marion Ballet (2012)² selon laquelle exprimer sa colère serait une tactique favorable pour les membres des partis d'opposition, suscitant ainsi une indignation et une envie de changement de la part des électeurs. Dans

¹ Pour rappel, la période étudiée en 2013 se situe six mois avant la fin du mandat de Dominique Voynet, soit dans une période pré-électorale. Il s'agit de plus d'une période durant laquelle celle-ci n'a pas encore déclaré qu'elle ne se représenterait pas et est donc supposée candidate par ses opposants.

² Les travaux de Marion Ballet (2012) s'inscrivent dans la réflexion menée notamment par Philippe Braud (1996) et Christian Plantin (1998, 2011) sur la place de l'émotion dans la communication politique.

l'optique socio-cognitiviste que nous avons suivie (Poggi & D'Errico, 2010), nous pouvons considérer cette hypothèse comme une piste de plus pour expliquer l'efficacité de l'expression d'attitudes agressives comme démonstration d'une image engagée, impliquée socialement de la part des membres de l'opposition, pouvant alors bénéficier à la dimension de bienveillance de leur *ethos* à l'égard des électeurs.

Dans cette optique, nous pourrions expliquer que si l'hostilité n'est pas positivement corrélée à la dominance perçue de Dominique Voynet, c'est que l'impératif de celle-ci dans la construction de son *ethos* dans le débat est autre, pour ne pas dire opposé : marquer sa dominance notamment en restant calme face aux attaques. Si comme le soulignent Poggi & D'Errico (2010) une forme d'agressivité est exprimée, encore faut-il que celle-ci soit contrôlée. Comme nous l'avons vu, plusieurs des extraits de parole de Dominique Voynet perçus comme hostiles et néanmoins dominants se caractérisent par un débit articulatoire ralenti³ et un écart entre hostilité exprimée au niveau verbal et hostilité exprimée typique de l'ironie (Anolli, 2000), ce qui est en accord avec l'interprétation de ces extraits comme exprimant de l'agressivité contrôlée.

Vers une évaluation perceptive adaptée à la parole politique : l'auto-évaluation

L'opportunité que nous avons eue de mener une auto-évaluation des stimuli par les locuteurs du corpus en parallèle à une évaluation par des auditeurs « naïfs » nous a permis d'entrevoir selon les locuteurs, des rapports parfois discordants entre *ethos* voulu (l'image que l'on souhaite véhiculer) et *ethos* perçu⁴ (Amossi, 1999). Si nous observons tous locuteurs confondus une corrélation significative entre mécontentement auto-évalué et perçu, nous avons notamment pour certains locuteurs pu constater des incongruences significatives entre les deux évaluations : Voynet est perçue comme plus mécontente (valence négative) qu'elle ne s'auto-évalue, Le Chequer se déclare plus calme qu'il n'est effectivement perçu. Il est intéressant de noter que Voynet est la seule locutrice pour laquelle on trouve un écart significatif entre valence auto-évaluée et valence perçue, ce qui va selon nous dans le sens de la différence de stratégie entre Maire sortante et opposants en période électorale municipale : les quatre opposants n'ont aucun mal à se déclarer mécontents durant leurs auto-évaluations. Quant à Le Chequer, on rappellera que s'il n'est certes pas perçu comme plus calme que les

³ Nous rappelons ici que Voynet est la seule locutrice pour laquelle le débit articulatoire est corrélé négativement à l'hostilité perçue.

⁴ Nous avons néanmoins conscience du fait que l'auto-évaluation ne correspond pas seulement à l'*ethos* voulu, et qu'elle peut même en différer si l'on pense à un locuteur qui aurait voulu apparaître calme aux yeux des spectateurs sur le moment et qui s'auto-évaluerait néanmoins comme énervé.

autres locuteurs du corpus, il reste le seul locuteur pour lequel l'hostilité perçue n'est pas corrélée à l'excitation perçue.

Enfin, concernant l'auto-évaluation du contrôle de soi, nous avons vu que cette dimension entretient des rapports variables avec la dominance perçue selon les locuteurs. Le Chequer et Serey sont ainsi perçus comme plus dominants dans le cas des stimuli pour lesquels ils déclarent un contrôle de soi maximal. À l'inverse, pour Tuillon ce sont les stimuli pour lesquels il déclare un contrôle de soi minimal qui sont perçus comme les plus dominants, ce qui pourrait être interprété pour ce dernier comme une non-nécessité de contrôle. En revanche, Serey est le locuteur pour lequel on trouve la corrélation significative la plus forte entre contrôle de soi et excitation auto-évaluée, signifiant que lorsque celui-ci se déclare agité il considère néanmoins exercer un contrôle sur l'expression de ses émotions.

La participation des locuteurs à l'auto-évaluation a été une des étapes les plus enrichissantes de ce travail de recherche, et l'une des plus agréables à mener du fait de l'atmosphère joviale dans laquelle se sont déroulés les tests d'écoutes durant les entretiens⁵. Si nous craignons au départ une réticence et une gêne de la part des locuteurs⁶, ceux-ci se sont tous pliés à l'exercice avec curiosité avec, pour certains, l'impression exprimée de remplir un « quiz de personnalité ». Nous notons également que tous les locuteurs ont fréquemment indiqué reconnaître instantanément le contexte précis dont était extrait le stimulus écouté, surtout lorsque celui-ci correspondait aux moments les plus vifs du débat.

Si nous devons réitérer une telle tâche d'auto-évaluation ou simplement effectuer des recommandations pour la mise en place d'un protocole de ce type, a posteriori nous recommanderions alors d'ajouter deux échelles perceptives aux grilles d'auto-évaluation :

- Une échelle de **souvenance** du contexte précis dont est extrait le stimulus, allant de « je ne me rappelais même pas avoir dit ça » à « je me souviens parfaitement de ce moment et des termes employés »
- Une échelle de degré de **spontanéité** des productions allant de la lecture oralisée à des propos tenus dans l'émergence soudaine d'un état émotionnel⁷.

⁵ Nous revenons plus bas sur l'enregistrement de la condition de contrôle ; voir section 0.

⁶ Tous ont déclaré spontanément ne pas apprécier écouter leur propre voix et tous ont fait des remarques sur leur façon de parler auto-perçue comme criblée de pauses remplies, incorrecte syntaxiquement. En réponse à cela nous nous sommes permise de leur faire part des remarques de Blanche-Benveniste (2010) sur la syntaxe de l'oral pour leur permettre de relativiser leurs impressions.

⁷ Ce type d'échelles a notamment fait ses preuves dans la description et distinction des styles de parole selon la situation de communication proposée par Koch & Oesterreicher (2008), comme on peut le voir dans des études telles que celles de Bardiaux (2010) ou Degand et collègues (2014), et gagnerait selon nous à être plus utilisé en perception.

Si ces deux échelles nous semblent importantes à tester dans une auto-évaluation pour considérer leurs rapports avec les profils émotionnels, l'échelle de spontanéité pourrait de plus être confrontée directement à la perception d'auditeurs « naïfs ».

Vers une évaluation perceptive adaptée à la parole politique

Au niveau de la perception par les auditeurs naïfs, notre étude a montré l'utilité d'une évaluation dimensionnelle (valence, excitation, dominance) d'extraits de parole politique, notamment pour travailler ensuite sur une analyse des caractéristiques vocales et pour identifier des profils différents selon les locuteurs. Outre le profil particulier de Dominique Voynet par rapport aux autres locuteurs concernant les rapports entre dominance perçue et hostilité perçue, on peut également rappeler ici le profil de Gaylord Le Chequer, seul locuteur pour qui l'hostilité perçue n'est pas corrélée à l'excitation perçue, ce qui peut être interprété comme une façon particulièrement contrôlée de produire des attitudes agressives.

Toutefois, notre étude ne permet pas de distinguer les locuteurs sur le plan de la dominance perçue, tous ayant été perçus comme légèrement dominants par les auditeurs « naïfs ». Cette homogénéité pourrait s'expliquer en premier lieu selon nous par l'influence du style de parole politique, et son impératif, on l'a vu, d'apparaître essentiellement dominant⁸. Il pourrait être intéressant d'effectuer des analyses perceptives analogues à celles menées dans notre étude, mais incluant cette fois-ci directement les autres dimensions principales de l'*ethos* telles que la compétence et la bienveillance, ces deux dernières pouvant contribuer indirectement à la perception de la dominance. Nous n'avons pas trouvé d'études à ce jour questionnant les rapports entre dimensions émotionnelles de type valence, excitation et dimensions de bienveillance et de compétence.

De plus, concernant ces dimensions de bienveillance et de compétence, la question de leur évaluation perceptive se pose. Peuvent-elles s'évaluer de la même façon que la dominance et les dimensions principales utilisées en psychologie, pour lesquelles des extraits courts peuvent s'avérer suffisants, ou nécessitent-elles plus d'informations contextuelles et lexicales que ne permettent de fait une approche par test de perception sur stimuli constitués de phrases tronquées ? En d'autres termes, quelle peut être l'importance de l'expressivité vocale sur l'image de bienveillance et de compétence renvoyée par le locuteur ?

⁸ On pourra également penser à la nécessité pour les auditeurs, lorsqu'on traite de la question de la dominance en politique, de pouvoir comparer les locuteurs pour déterminer qui a « plus de pouvoir » que l'autre, ce qui demanderait un protocole expérimental très différent de celui-ci ; nous reviendrons sur ce point dans l'ultime section de cette discussion.

Expressivité perçue à l'oral et à l'écrit : comment expliquer leurs rapports ?

L'évaluation parallèle des transcriptions orthographiques isolées et des extraits audio originaux a permis d'obtenir des informations sur les poids respectifs des indices verbaux et des indices vocaux, notamment dans l'expression de l'hostilité.

Le fait que les locuteurs aient été perçus comme exprimant en moyenne une légère hostilité à l'écrit comme à l'oral confirme la multimodalité de l'expressivité en parole naturelle.

Avant de présenter en conclusion finale nos recommandations pour une étude appliquée des caractéristiques vocales expressives prenant en compte l'expressivité lexico-syntaxique⁹, nous nous interrogeons d'abord dans la présente section sur comment mieux appréhender la variabilité des rapports entre expressivité vocale et expressivité verbale et prévoir certaines tendances en parole politique.

Nous pourrions ici, dans la continuité de notre appui sur leur approche sociocognitive, suivre les hypothèses et observations de Poggi et collègues (2011) concernant un recours variable à l'expressivité paralinguistique selon le type d'attaque effectuée à l'encontre de l'opposant dans le débat. Pour préserver une image de soi favorable tout en nuisant à celle de l'opposant, les chercheuses prévoient une exploitation beaucoup plus tacite des caractéristiques vocales expressives si l'on s'attaque directement à sa dominance, par exemple en le ridiculisant, que si l'on s'attaque à sa bienveillance. Il s'agit notamment, comme on l'a vu dans l'optique de Ballet (2012), de convaincre l'auditoire que l'opposant traite injustement les citoyens et qu'il faut s'en indigner.

Si ces hypothèses se confirment, une annotation par type d'attaque à l'*ethos* pourra permettre de mieux prévoir le degré d'expressivité vocale véhiculée dans les extraits pour mieux en appréhender l'analyse en perception comme en production.

Aussi, si dans certains types d'attaques on s'attend à une expressivité vocale moindre, il apparaîtra d'autant plus nécessaire de pouvoir analyser en parallèle les autres niveaux par lesquels passe l'expression agressive (lexique, syntaxe, recours implicite au contexte, etc.). Pour tous ces aspects, une approche complémentaire et parallèle menée avec des chercheurs en analyse du discours pourrait s'avérer fructueuse.

⁹ Le parti pris qui sera le nôtre dans les conclusions de ce travail et nos ultimes recommandations est que l'analyse fine des rapports entre expressivité lexico-syntaxique d'une part et vocale de l'autre est la clé d'une meilleure description des caractéristiques vocales des attitudes contrôlées pour lesquelles on sait qu'un consensus à travers les études strictement phonétiques est difficile à observer (Scherer, 2003).

IX.2. Avantages et limites de la condition de contrôle (corpus relecture)

Si l'évidence de l'apport du corpus relecture dans l'acquisition de mesures comparées apparaît dans les chapitres d'analyse de la thèse, nous prenons ici le temps d'apporter quelques remarques sur les limites de l'utilisation d'une condition de contrôle lue et non spontanée, notamment dans la difficulté de validation perceptive de celle-ci.

Choix de la condition de contrôle

L'évaluation perceptive des caractéristiques vocales expressives des locuteurs du corpus selon l'hostilité, la valence, l'excitation et la dominance perçue a montré son utilité. La comparaison acoustique et prosodique entre le corpus spontané et sa relecture a permis de quantifier l'écart entre productions expressives et lecture oralisée présumée « neutre » pour chaque locuteur, sur un même matériel sémantico-syntaxique. L'intérêt fondamental de l'enregistrement de cette condition a consisté à pouvoir ensuite effectuer des comparaisons systématiques sur les unités prosodiques majeures et les unités syntaxiques sur des contenus similaires. De plus, les mesures de différences original-relecture effectuées sur les caractéristiques prosodiques et temporelles étudiées nous ont permis de dégager des corrélats significatifs de la perception de l'hostilité et de la dominance, à travers des contrastes d'étendue de la plage de variation de F0 et de débit articulatoire de part et d'autres des pauses silencieuses, très différents dans les extraits originaux et dans les extraits relus.

À notre sens, le **problème** soulevé par la question de la **condition de contrôle** réside du rapport contradictoire en phonétique entre **écologie des productions** analysées et possibilités d'investigation fine des caractéristiques de parole. En effet, une première question que nous nous sommes posée : **pourquoi ne pas prendre des extraits du corpus où les locuteurs ne sont pas perçus comme hostiles comme condition de contrôle ?**

Si nous insistons à ce sujet sur le fait que seul un contenu lexico-syntaxique similaire peut permettre une analyse comparative quantitative pertinente des caractéristiques segmentales et suprasegmentales, il n'a pas été chose aisée dans la sélection des extraits de notre corpus original de trouver des portions de parole à la fois non agressives et non lues. Lors de la sélection des locuteurs, beaucoup d'entre eux (rappelons qu'il y avait 50 conseillers municipaux dotés de micros durant les séances du conseil municipal étudiées en 2013) ont été écartés du fait qu'ils intervenaient uniquement en produisant des lectures déclamées de leurs interventions pré-écrites. Si rien n'assure pour tous les stimuli de notre corpus final que ceux-ci sont produits de façon absolument spontanée, les cinq locuteurs de la présente étude ont été

sélectionnés pour leur prédilection pour un discours de style spontané, des interventions fréquentes, ainsi que des rapports conflictuels (entre la Maire et les quatre opposants).

Ceci nous amène à une deuxième question complémentaire : **pourquoi comparer les productions expressives et la parole lue présumée « neutre » ?** Question à laquelle nous répondons qu'en prenant pour impératif d'avoir des contenus comparables, il est difficile de faire autrement sans multiplier les approches : à savoir, travailler d'abord sur des corpus de parole naturelle puis sur des reproductions ciblées en laboratoire, comme nous le préconiserons, entre autres, au terme de cette discussion.

Une expérience éclairante de « contextualisation » de la parole « neutre »

C'est justement à cause des biais perceptifs dus à l'évaluation de la parole présumée « neutre » et lue (Banziger et Scherer, 2007 ; Kouklia & Audibert, 2013), que nous avons fait le choix dans notre présente étude de dissocier ce qui serait de l'ordre d'une condition de contrôle en perception (scores d'hostilité perçue à l'écrit) d'une condition de contrôle en production (corpus relecture)¹⁰.

Pour Banziger & Scherer (2007), la parole présumée « neutre » ne peut être un référent fiable pour être comparée ensuite aux productions expressives du fait de la sous-spécification de sa nature en terme dimensionnel pour lesquels on ne peut que prévoir une valence neutre, et éventuellement un léger contrôle sur les productions car non affectées par une quelconque réaction ou posture émotionnelle¹¹.

Partant de la prise de position de Banziger & Scherer (2007), nous avons décidé de consacrer une courte étude, en parallèle à nos recherches doctorales, consacrée à l'analyse perceptive de la parole dite « neutre ». Nous avons cherché à établir si la nature de la parole dite « neutre », généralement lue dans les études, ne pouvait pas influencer la spécification de celle-ci comme légèrement contrôlée (Kouklia & Audibert, 2013). Nous avons participé à l'enregistrement par des comédiens d'un corpus de parole expressive simulée en laboratoire¹² comprenant l'expression de l'anxiété, de l'irritation, du contentement et de l'abattement. Ce corpus incluait ces productions d'affects de degrés d'activation émotionnelle variables et une condition de contrôle « neutre » :

- Une version lue, typique des études en parole expressive de laboratoire, au cours de laquelle les comédiens ont relu les différentes phrases sur lesquelles ils ont travaillé en parole expressive (« *il est huit heures.* », « *j'entre dans le bâtiment.* », « *le papier est sur table.* ») avec pour seule consigne de le faire « de façon neutre » et de marquer des temps de pause entre chaque phrase pour éviter un éventuel effet de liste.

¹⁰ Ce choix n'exclut pas la pertinence de la mise en place d'une validation perceptive du corpus « relecture » que nous avons enregistré avec les locuteurs mais s'explique du fait des obstacles de fond et de forme que nous avons rencontrés et que nous décrivons ici.

¹¹ Banziger & Scherer (2007) iront jusqu'à remplacer la condition « neutre » dans leurs expérimentations par une condition de léger contentement pour lesquels il est possible de mieux spécifier les caractéristiques attendues avec une valence, une excitation et une dominance légèrement positives.

¹² Le corpus en question était constitué de variations expressives sur des voyelles tenues produites isolément pour pouvoir ensuite être testées dans une expérience neuro-phonétique de perception des affects chez des patients bipolaires (Carminati et collègues, 2018). Pour faciliter la tâche de production aux comédiens participants à l'étude, nous les avons fait travailler sur un ensemble de phrases et de scénarios inspirés de l'étude de Fonagy et Bérard (1972), que nous avons extraites pour les besoins de notre étude.

- Une version dialoguée, pour laquelle ceux-ci avaient pour tâche de produire les phrases en répondant à des questions formulées par les expérimentatrices (« *Quelle heure est-il ?* », « *Où es-tu ?* », « *Où est le papier ?* ») avec pour consigne d'acteur de répondre comme s'ils avaient simplement pour but de fournir une réponse à la personne, sans avis ni affect particulier à l'égard de la situation.

Nous avons ensuite fait passer un test de perception dimensionnelle (valence, excitation, dominance) à 23 auditeurs « naïfs ». Le corpus comprenant a) l'ensemble des expressions d'attitudes contrôlées (anxiété, irritation, contentement, abattement) produites par les comédiens et b) et c) les deux conditions « neutres » (« neutre lue », « neutre dialoguée »). Les résultats de l'analyse perceptive ont montré que

- la version « neutre dialoguée » a été perçue comme neutre sur les trois dimensions de valence, excitation, dominance.
- La condition « neutre lue » a été perçue comme neutre sur le plan de la valence et de l'excitation. Au niveau de la dimension de dominance, elle a été perçue comme exprimant une légère soumission, sans différence significative avec l'expression d'abattement (tristesse faiblement activée), variante expressive perçue comme la plus soumise du corpus.

Comment expliquer alors que la condition « neutre lue » ait été perçue comme soumise ? Nous pouvons expliquer ce biais perceptif en l'appréhendant dans l'optique socio-cognitiviste (Castelfranchi, 2003 ; Poggi et collègues, 2007, 2010, 2011) : si les individus s'expriment toujours pour accomplir des buts personnels et interpersonnels selon les règles de la situation de communication et les croyances de leurs cultures, le risque encouru de faire évaluer l'expressivité d'extraits de parole lue « neutre » à des auditeurs « naïfs » est que ceux-ci tentent plus ou moins inconsciemment de reconstituer une raison, un but derrière la production « neutre lue » du locuteur, alors qu'il n'y en a pas. Il nous apparaît surprenant, mais cohérent à travers cette conception qu'il puisse suffire de poser des questions à des comédiens, de faire passer leur but de « je lis la phrase parce que l'expérimentateur me le demande » à « je joue une situation dans laquelle je réponds à une question d'ordre strictement informationnel » pour que leurs productions « neutres » soient déjà mieux perçues comme telles.

Cependant, si le recours au dialogue joué a été très simple à mettre en place dans cette étude passée (Kouklia & Audibert, 2013), la réplique d'une telle méthodologie pour la condition de contrôle de notre corpus d'extraits de débats nous a paru beaucoup plus douteuse.

Si Enos et Hirschberg (2006), comme Banziger et Scherer (2007), insistent sur le fait que la parole expressive actée ne peut maintenir sa pertinence dans le domaine de l'analyse de la parole expressive que tant qu'un travail assidu de direction d'acteur est effectué, nous avons fait le choix résolu avec nos locuteurs politiques de ne pas essayer de mettre en place de productions expressives actées, quand bien même celles-ci se limiteraient à la production d'une condition « neutre ».

Renonçant à un quelconque artefact d'écologie de production du présent corpus « relecture »¹³ de la présente thèse, nous avons enregistré celui-ci avec les locuteurs durant des entretiens individuels en pleine conscience des intérêts et limites d'une telle condition de contrôle.

Des difficultés de la validation perceptive de la condition de contrôle

Une validation perceptive est une nécessité avant toute analyse phonétique. Afin d'étudier la faisabilité d'une évaluation perceptive combinant enregistrements originaux et relecture, nous avons effectué un pré-test avec 5 auditeurs « naïfs » en adoptant un protocole de discrimination dimensionnelle de paires de stimuli originaux versus relus¹⁴. Dans ce pré-test, les stimuli étaient joués consécutivement et les auditeurs devaient considérer sur trois échelles de Likert à 5 points si dans le second extrait le locuteur avait l'air plus ou moins mécontent, agité et dominant que dans le premier. Les réponses des participants du pré-test se sont avérées problématiques à interpréter : la question post-test sur les indices exploités par les auditeurs pour différencier les productions originales des productions relues au sein de chaque paire, tous les auditeurs ont fait état d'un bruit de fond présent lors des débats. Nous avons donc ajouté du bruit de fond distinctif dans les stimuli du corpus relecture et refait passer l'expérience. Trois auditeurs sur cinq ont déclaré spontanément avoir eu du mal à se souvenir de ce qu'ils avaient perçu du premier extrait de la paire écoutée. Ceci peut aisément s'expliquer par le fait que la durée des paires concaténées pouvait aller jusqu'à 25 secondes ; un test de discrimination typique en phonétique comportant généralement des extraits beaucoup plus courts (voyelles, mots isolés).

Pour évaluer l'expressivité véhiculée par la condition de contrôle, nous avons fait le choix de faire évaluer par un panel de sujets naïfs les transcriptions orthographiques, qui nous a permis de quantifier le poids sémantique des énoncés. Nous n'avons donc pas testé directement le corpus relecture en perception, et ne disposons donc que de l'évaluation de la

¹³ Il n'est pas possible de refaire des enregistrements contrôle dans la salle des fêtes de la Mairie, ni de demander aux locuteurs de rejouer seuls « les scènes » desquelles étaient extraits leurs stimuli.

transcription écrite plutôt que d'une validation perceptive spécifique à celui-ci. Aussi, d'un point de vue strictement formel, nous ne pouvons assurer que le corpus relecture est inexpressif par rapport au corpus original qu'en nous appuyant que sur des éléments de distinction situationnelle entre les deux conditions¹⁵ :

- Dans le **corpus original**, les locuteurs s'expriment dans un style de parole spontanée, avec des buts et enjeux électoraux et interactionnels forts, dans un contexte public, avec une distance élevée entre eux et le public justifiant l'installation de hauts parleurs, et dans des états émotionnels variables¹⁶.
- Dans le **corpus relecture**, les locuteurs produisent une lecture oralisée d'extraits de leur propre parole, avec pour but de répondre à notre demande de relecture, avec la consigne de ne pas exprimer d'expressivité particulière et de lire à la manière d'une dictée, dans une pièce calme avec un micro-casque.

Nous recommandons à tout chercheur souhaitant mettre en place un protocole similaire de faire produire aux locuteurs un enregistrement supplémentaire avec pour support un texte similaire pour tous. Un tel enregistrement aurait pu servir à tester ce que des auditeurs perçoivent de la voix des locuteurs à partir d'un même support textuel, et ainsi tester une certaine caractérogie vocale : certains locuteurs pouvant être perçus comme plus enjoués ou dynamiques que d'autres sur un contenu sémantique similaire (Fónagy, 1983). Il aurait pu aussi servir de base à des analyses systématiques fines comparatives des comportements des locuteurs en condition de lecture. Le résultat de ces comparaisons aurait pu s'avérer précieux dans l'analyse comparée de leurs caractéristiques de structuration prosodico-syntaxique qui a suivi entre corpus original et relecture, que nous allons à présent discuter.

¹⁵ Les considérations suivantes sont inspirées du modèle de Koch & Oesterreicher (2008) dont une traduction synthétique à visée pédagogique a été fournie par Modicom (2015) pour promouvoir leur grille de description situationnelle pour les chercheurs non germanophones. Nous précisons que leur modèle se veut non exhaustif et peut inclure selon les spécificités de la situation de communication des échelles de description variées tant que celles-ci illustrent d'une façon ou d'une autre une distinction graduelle entre deux polarités : le domaine de l'oral (spontané, direct, personnel, investissement émotionnel, etc.) et le domaine de l'écrit (formel, pré-écrit, public, pas d'investissement émotionnel etc.). Ces grilles sont couramment utilisées pour distinguer les styles de parole.

¹⁶ Deux locuteurs sur cinq ont déclaré qualitativement (durant l'entretien mais hors de l'évaluation dimensionnelle) avoir été soumis à une tension émotionnelle très accrue pendant toutes les séances de conseil étudiées ici.

IX.3. Arguments pour une analyse multifactorielle de la pause silencieuse

Au sortir des analyses prosodiques de cette recherche, nos hypothèses sur l'exploitation des pauses silencieuses en tant que signal de dominance (Duez, 1991) ne sont que partiellement confirmées. Dans notre corpus, la pause n'est pas directement corrélée à la dominance. En revanche, nous retrouvons son impact dans les variations prosodiques de part et d'autre des pauses, notamment au travers de forts contrastes de plage de variation de F0 pré et post-pausaux.

Dans cette partie de la discussion, notre objectif est d'inciter à une méthodologie d'analyse de la pause silencieuse permettant d'analyser plus finement les rapports entre ses fonctions, sa durée, sa localisation dans la structure syntaxique, et ses corrélats prosodiques en contexte gauche et droit.

Conformément aux conclusions qui étaient déjà celles de Duez (1991) et qui sont énoncées en synthèse de la récente revue de la littérature sur les pauses effectuées par Grosman et collègues (2018), les rapports entre caractéristiques et fonctions des pauses sont complexes : les rapports entre localisation syntaxique et fonctions paralinguistiques sont non exclusifs. Pour appuyer ce point de vue, nous posons au regard des caractéristiques observées quantitativement et qualitativement dans nos données, les jalons d'une approche cumulative des critères d'analyses des pauses pour contribuer à la détermination de leurs fonctions. Enfin, nous proposons un ensemble d'approches plurielles et pluridisciplinaires pour pouvoir notamment traiter plus finement la question de la perception de la fonction paralinguistique des pauses.

Pour une actualisation des critères d'annotation des frontières prosodiques

Vers une remise en question d'un seuil de durée minimal

Pour des raisons de comparabilité avec la majorité des études portant sur la pause et les phénomènes de structuration prosodique et syntaxique majeures, nous avons conservé dans cette étude un **seuil de durée limite** (140 ms) pour appréhender le phénomène de la pause silencieuse. Toutefois il nous semble impératif pour notre domaine de recherche de pouvoir proposer une approche plus dynamique de la durée des pauses. Duez (1991) dès ses premières recherches, avait insisté sur l'influence du débit articulatoire et des caractéristiques temporelles générales de chaque locuteur, et de telles précautions nous semblent nécessaires à mettre en place.

Si nos observations gagneraient à être confrontées à d'autres corpus de parole politique en situation de débat, la conclusion à laquelle nous aboutissons d'ores et déjà à l'issue de notre recherche est que le seuil de durée minimale des pauses, qui est à 200 ms dans la majorité des études (voir la revue de Grosman et collègues, 2018) n'est pas toujours adapté à la parole spontanée et entraîne le risque d'une non-considération d'un nombre considérable de pauses intra-syntaxiques. Ce seuil a souvent été choisi de façon pratique, pour que les pauses puissent être repérées automatiquement par programmes informatiques et non confondues avec le silence des occlusives sourdes, dont la durée n'excède que rarement 200 ms. En réalité, une pause peut être de durée bien moindre, et quelques millisecondes de glottalisation peuvent être perçues comme une pause. Les pauses de focalisation, localement juste avant le mot à mettre en valeur peuvent également être de durée très courte, bien en deçà de 50 millisecondes, et clairement perçues comme des pauses (Vaissière, 2018, communication personnelle).

La définition de critères robustes permettant d'appréhender de façon plus dynamique des seuils pertinents pour les paramètres mélodiques et temporels engagés dans la production des frontières prosodiques est tout sauf triviale. Pour aboutir à cela, nous recommandons la mise en place d'études mettant en rapport analyses acoustiques et perception des fonctions prosodiques correspondantes sur des corpus constitués dans le but précis de pouvoir effectuer des comparaisons systématiques selon la configuration de chaque critère d'annotation prosodique testé (pause silencieuse, F0, durée).

En concentrant nos recherches sur des unités définies autour de la pause silencieuse, nous avons principalement analysé les caractéristiques générales mélodiques et rythmiques (durées et débits syllabiques) en les appréhendant comme des corrélats des pauses silencieuses. L'émergence d'accents variés, allant d'accents didactiques (Lucci, 1980) à des renforcements initiaux dus à l'émergence d'hostilité (Fónagy, 1983) en passant par des accents de focalisation (Léon, 1971), n'a pas pu être étudiée à sa juste valeur du fait d'une analyse des caractéristiques mélodiques et temporelles pré-pausales et non post-pausales : l'annotation des frontières prosodiques que nous avons effectuée se concentrait principalement sur les caractéristiques en contexte gauche des pauses silencieuses et nous avons annoté les frontières majeures seulement.

Nous n'avons pas fait une étude séparée des phénomènes de renforcements initiaux. Un renforcement initial (ou *initial rise*) est souvent réalisé au début d'un mot lexical, en dehors de tout accent d'insistance et il est très présent en parole lue (Vaissière, 1974). La présence d'un renforcement initial impacte également les caractéristiques segmentales, mélodiques mais également spectrales au sein des syllabes affectées. En effet, comme le souligne Vaissière (2010), l'accentuation initiale en début de mot, renforcée ou non par un type accent d'insistance, peut notamment se traduire par un saut de F0 entre la syllabe précédant le début de mot et la syllabe initiale. Sur le plan articulatoire, pour la consonne initiale on peut également prévoir un contact accru entre la langue et le palais ainsi qu'un dévoisement, phénomène d'ailleurs attesté par Fónagy (1983) comme caractéristique d'expression d'attitudes agressives (voir partie II.3.3.2 page 30). Un renforcement initial régulier est très présent dans la parole adressée au public, où il devient un corrélat d'un certain style. La force comparée des renforcements initiaux en début de mots successifs (et la présence d'une pause éventuelle) dans un énoncé permet de dériver une structure informative de cet énoncé. Quant à elle, l'ampleur comparée de la montée de continuation, des allongements finaux et éventuellement des pauses en fin d'unités permettent essentiellement de dériver une structure « syntaxique » de ce même énoncé (Vaissière et Michaud, 2006).

Bien qu'elle n'ait pas été ici directement analysée, l'influence des accents initiaux est toutefois présente en filigrane dans toutes nos analyses, notamment à travers la mesure des différences de durée ou encore de fréquence fondamentale de part et d'autre des frontières considérées.

En d'autres termes, il apparaît évident à l'issue de notre étude que l'approche complémentaire prioritaire à apporter à l'analyse des pauses au travers des frontières prosodico-syntaxiques majeures doit consister en une analyse approfondie du contexte droit de celles-ci. Lorsqu'il est précédé d'une pause, le renforcement initial, qu'il ait une fonction identificatrice indexant le style politique (accent didactique), impressive (accent de focalisation attirant l'attention du locuteur sur un élément du discours), ou pragmatique (maintien forcé du tour de parole) est un indice non négligeable dans l'appréhension des fonctions des pauses, qui accompagnent souvent le renforcement initial.

Dans l'approche que nous avons adoptée, nous sommes principalement partis de la localisation des pauses en rapport avec la syntaxe. L'étude des pauses doit être faite également en fonction de la présence de renforcements initiaux dans le mot suivant ou précédent la pause. En effet, nous avons pu observer qualitativement l'occurrence de pauses silencieuses que l'on pourrait qualifier de focalisation postérieure, c'est à dire congruentes avec la structure syntaxique, néanmoins suivant directement la production d'un mot avec accent de focus, comme pour renforcer ce focus grâce au silence suivant.

Pour une approche cumulative et multifactorielle

Suite aux observations ayant et n'ayant pu être effectuées au sein de la présente étude, nous considérons que si les caractéristiques des pauses prises isolément ne peuvent prédire précisément leur fonction (nous pensons ici notamment à la durée mais également à la position dans la structure syntaxique), cela pourrait être un ensemble de caractéristiques. Par exemple, comme nous avons pu l'observer qualitativement dans notre étude, les pauses intra-syntaxiques peuvent tout aussi bien correspondre à des cas de pauses de focalisation ou de pauses de jonction externe précédant un accent didactique, voire même à des cas de pauses dites d'hésitation. Quels sont les critères pertinents pour qui différencier ces différents phénomènes ? Dans le cas de l'exemple que nous venons de citer, cela pourrait être une combinaison de localisation syntaxique, de présence ou absence de proéminences post-pausales, de différences de débit articulatoire, etc.

Nous suggérons ici l'intérêt d'une approche multifactorielle reposant sur l'annotation indépendante puis cumulée d'un ensemble hétérogène de facteurs et caractéristiques allant des contraintes physiologiques aux stratégies locutoires expressives des locuteurs incluant notamment pour chaque pause :

- Leur durée brute et leur durée relative à la durée de l'extrait traité, mais également à la durée du bloc précédent

- La forme des contours pré-pausaux (indiquant la montée de continuation ou la chute finale) et post-pausaux
- La comparaison de la F0 moyenne et de la plage de variation de F0 pré et post-pausales, les changements de registres
- La comparaison du débit articulatoire syllabique et de l'indice de variabilité paire à paire des durées pré et post-pausales
- La localisation syntaxique fine de toutes les pauses
- La présence d'indices audibles de respiration ou de déglutition, de coup de glotte et de glottalisation, de schwas épenthétiques pré-pausaux
- Le repérage de pauses remplies ou non remplies dites d'hésitation
- La présence d'accents primaires mélodiques pré ou post-pausaux.

IX.4. Vers une approche plurielle et pluridisciplinaire de l'expressivité vocale

Au sortir de cette recherche, nous insistons sur la nécessité impérieuse d'effectuer des analyses comparatives fiables et ciblées, que seul un corpus de laboratoire peut permettre. Si nous avons fait état de ses limites, nous insistons néanmoins sur le fait que l'accès à la condition de contrôle (relecture) nous a permis d'obtenir des informations cruciales sur les écarts produits par les locuteurs dans les extraits où ils ont été perçus comme dominants et/ou hostiles, comparativement à leurs propres caractéristiques de parole en lecture.

Toutefois, de par la nature complexe et hétérogène des énoncés constituant un corpus de parole naturelle comme celui que nous avons analysé ici, ni les analyses perceptives, ni les analyses prosodiques ne nous permettent, en l'état, d'établir des hypothèses fiables sur les liens directs entre caractéristiques prosodiques dues à l'insertion de pauses et perception de valeurs paralinguistiques, comme l'expression de l'hostilité ou de la dominance.

La clé du progrès réside dans une approche **plurielle**, alliant étude de corpus phonétique de parole spontanée et phonétique de laboratoire, et étude **pluridisciplinaire**. Il s'agit de favoriser la communication entre approches *descendantes*, telles que celles proposées en psycho-sociologie pour appréhender les enjeux qui sous-tendent les interactions, et nos méthodes, *ascendantes*, pour analyser la réalité perceptive et acoustique des phénomènes expressifs.

Pour une analyse phonétique plurielle

Arrivés au terme de cette étude, nous insistons sur l'importance d'une approche plurielle, et la nécessité d'appréhender de manière complémentaire et non dichotomique les différences entre parole naturelle et parole de laboratoire pour une approche plus fine des caractéristiques vocales expressives. Dans le cas de la parole politique et de l'impact de l'expressivité sur l'image transmise par les locuteurs, nous considérons que la pause silencieuse est un paramètre important à confronter avec sa perception, qu'il faut appréhender en prenant mieux en compte ses contextes droits et gauches afin de rendre compte de sa présence due aux phénomènes d'accentuations paralinguistiques.

Par une telle approche plurielle, notre mode d'investigation de la parole expressive, notamment politique, pourrait comprendre :

- 1) Une approche quantitative et qualitative sur un corpus de parole naturelle de taille réduite tel que celui que nous avons construit.
- 2) Une analyse quantitative sur des corpus de parole naturelle plus vastes préalablement annotés, comme le corpus ETAPE¹⁷ (Gravier et collègues, 2012) contenant plusieurs heures de parole politique en débat, ou encore le corpus LOCAS-F (Degand et collègues, 2014).
- 3) Une analyse ciblée en laboratoire permettant par la construction de corpus contrôlés, l'investigation perceptive de l'influence des différents paramètres des pauses silencieuses dans la perception de la parole expressive des locuteurs. Notons qu'il est également facile de manipuler la présence et la longueur des pauses d'un corpus lu ou spontané et de juger de l'effet perceptif obtenu.

¹⁷ Nous pensons ici également aux travaux menés par Wu et collègues (2016) sur l'analyse du schwa au travers notamment de ce corpus.

Vers une approche pluridisciplinaire

L'analyse perceptive et productive de l'expressivité vocale dans le débat politique, et plus généralement dans tout contexte d'interaction, exige selon nous un dialogue actif entre disciplines, favorisant tout à la fois l'accroissement de nos connaissances conceptuelles des règles et enjeux de la situation de communication et des analyses fines et contrôlées. Un travail de recherche s'impose tant dans une optique *Top-Down*, par une meilleure appréhension des buts implicites et tactiques des locuteurs durant dans un débat, que dans une approche *Bottom-Up*, en approfondissant dans le cadre d'études ciblées en laboratoire les observations parfois qualitatives et relatives que nous avons pu déjà obtenir à partir de données issues du terrain. Nous estimons particulièrement nécessaire, à ce sujet l'élaboration en laboratoire de corpus permettant l'analyse comparative systématique de l'influence de la réalisation des frontières prosodiques (rapports syntaxe/prosodie, caractéristiques mélodiques et temporelles) sur la perception attitudinale.

À notre sens, seul un dynamisme pluridisciplinaire permettra de passer outre les obstacles méthodologiques actuels pour permettre une appréhension pertinente de l'expressivité vocale et une analyse plus fine des caractéristiques phonétiques relatives à la parole issue de situations d'interaction naturelle.

IX.5. Principaux apports de cette étude

Selon nous, les contributions les plus marquantes du travail de recherches doctorales que nous avons mené se situent aux niveaux suivants.

L'ensemble de notre protocole a permis la mise en place d'analyses phonétiques appliquées à un corpus de parole naturelle. L'ajout d'une condition de contrôle a permis d'effectuer des comparaisons systématiques (généralement non réalisables sur un corpus de parole naturelle produite hors laboratoire) et de permettre ainsi une analyse plus fine de la variation expressive au niveau segmental et suprasegmental d'extraits de parole issus d'une interaction réelle. Outre le fait qu'elles ont permis une objectivation et une validation des données en termes d'hostilité, de dominance, d'excitation et de valence, l'ensemble des analyses perceptives mises en place ont été utilisées comme des clés pour l'analyse acoustique des données. L'ajout d'une condition écrite à l'analyse de l'hostilité perçue a permis une quantification de l'expressivité inhérente au contenu sémantique et celle ajoutée par les indices paralinguistiques. La comparaison des profils émotionnels auto-évalués et perçus par les auditeurs naïfs a permis quant à elle de quantifier l'écart entre *ethos* voulu (image voulant être véhiculée par la personnalité politique) et *ethos* perçu (Amossi, 1999).

Par une approche combinée entre analyses perceptives dimensionnelles et analyses prosodiques, nous avons pu dégager un ensemble de corrélats de la perception de l'hostilité et de la dominance dans des extraits de parole politique en débat : débit articulatoire plus variable, raccourcissement des allongements pré-pausaux, plage réduite de F0 et/ou variations fortes de part et d'autre des pauses silencieuses. Les corrélats de l'hostilité et de la dominance que nous avons mis en évidence dans les extraits analysés se rapprochent tout autant de caractéristiques attendues pour la production d'attitudes agressives contrôlées (Fónagy, 1983 ; Scherer, 2003) que de caractéristiques relevées dans le cas d'insertions de pauses de focalisation ou de maintien du tour de parole, dans la parole spontanée et/ou politique (Ferré, 2004 ; Béchet, 2013).

IX.	Discussion et Conclusion.....	185
IX.1.	Expressivité vocale et perception des personnes politiques en débat.....	186
IX.2.	Avantages et limites de la condition de contrôle (corpus relecture).....	191
IX.3.	Arguments pour une analyse multifactorielle de la pause silencieuse.....	197
IX.4.	Vers une approche plurielle et pluridisciplinaire de l'expressivité vocale	202
IX.5.	Principaux apports de cette étude	205

IX. Discussion et Conclusion

Dans ce chapitre, à l'issue de ce travail de recherche doctoral, nous essayons de déterminer les apports et limites du protocole expérimental mis en place et de livrer quelques réflexions sur la continuation possible de l'analyse de l'expressivité vocale.

Nos trois objectifs complémentaires de départ, énoncés en introduction, étaient les suivants :

1. Analyser l'impact de l'expressivité vocale sur la perception de la dominance dans des extraits de parole politique en débat : nous discuterons donc principalement de la dimension perceptive de notre protocole (approche dimensionnelle et auto-évaluation).
2. Fournir un protocole méthodologique adapté à l'analyse phonétique appliquée de parole expressive naturelle : nous discuterons des apports et limites de notre condition de contrôle.
3. Apporter une continuité aux analyses de Danielle Duez (1991) sur les rapports entre pouvoir et pauses silencieuses, notamment au niveau perceptif et au niveau acoustique en analysant les corrélats mélodiques et temporels des pauses (débit articulatoire, allongements). Concernant ce troisième objectif, nous effectuerons un retour critique sur les analyses prosodiques effectuées. Nous nous efforcerons de fournir une synthèse de nos recommandations pour un protocole pluriel et pluridisciplinaire pour une analyse plus approfondie, notamment en perception, des fonctions paralinguistiques de la pause silencieuse.

IX.1. Expressivité vocale et perception des personnes politiques en débat

Hostilité, dominance perçue et leurs corrélats de parole : En ce qui concerne la question de l'influence de l'expressivité vocale sur la perception de la dominance, notre étude confirme les hypothèses de Poggi & D'Errico (2010) quant à l'influence d'expressions d'attitudes agressives contrôlées dans la communication d'une image de soi (*ethos*) favorable.

Dans notre étude, tous les locuteurs ont été perçus comme légèrement dominants. Les principaux corrélats de parole de la dominance et de l'hostilité perçue dans nos données nous rappellent en de nombreux points le ton « menaçant » décrit par Fónagy (1983). Selon Fónagy, le ton « menaçant » consiste en un débit articulatoire élevé, une faible variation de durées syllabiques et un faible allongement final, et des écarts marqués de la plage de variation de F0 de part et d'autre des pauses avec une prédilection pour des contours mélodiques plats. Dans notre étude, la plage de variation de F0 à la fois réduite et fortement variable de part et d'autre des pauses est le corrélat de la dominance et de l'hostilité le plus commun pour tous les locuteurs sauf Dominique Voynet.

En effet, Dominique Voynet est la seule locutrice pour laquelle la dominance perçue n'est pas significativement corrélée à l'hostilité perçue : pour les quatre autres locuteurs, les extraits dans lesquels ceux-ci sont perçus comme les plus hostiles sont également ceux pour lesquels ils sont perçus comme les plus dominants. Si notre corpus ne permet en aucun cas d'émettre d'hypothèses sur l'influence du genre du locuteur à ce sujet, la question reste ouverte : les différences entre Voynet et les quatre autres locuteurs seraient-ils plus à mettre sur le compte de son statut de femme ou de son statut de Maire ? Nous pouvons effectivement également considérer l'influence de la différence de rôles et d'enjeux entre Dominique Voynet et les quatre membres de l'opposition étudiés en prenant en compte le contexte préélectoral dans lequel se sont déroulés les débats¹.

La corrélation positive entre dominance et hostilité perçue pour les locuteurs masculins de notre corpus rappelle les conclusions de la thèse de Marion Ballet (2012)² selon laquelle exprimer sa colère serait une tactique favorable pour les membres des partis d'opposition, suscitant ainsi une indignation et une envie de changement de la part des électeurs. Dans

¹ Pour rappel, la période étudiée en 2013 se situe six mois avant la fin du mandat de Dominique Voynet, soit dans une période pré-électorale. Il s'agit de plus d'une période durant laquelle celle-ci n'a pas encore déclaré qu'elle ne se représenterait pas et est donc supposée candidate par ses opposants.

² Les travaux de Marion Ballet (2012) s'inscrivent dans la réflexion menée notamment par Philippe Braud (1996) et Christian Plantin (1998, 2011) sur la place de l'émotion dans la communication politique.

l'optique socio-cognitiviste que nous avons suivie (Poggi & D'Errico, 2010), nous pouvons considérer cette hypothèse comme une piste de plus pour expliquer l'efficacité de l'expression d'attitudes agressives comme démonstration d'une image engagée, impliquée socialement de la part des membres de l'opposition, pouvant alors bénéficier à la dimension de bienveillance de leur *ethos* à l'égard des électeurs.

Dans cette optique, nous pourrions expliquer que si l'hostilité n'est pas positivement corrélée à la dominance perçue de Dominique Voynet, c'est que l'impératif de celle-ci dans la construction de son *ethos* dans le débat est autre, pour ne pas dire opposé : marquer sa dominance notamment en restant calme face aux attaques. Si comme le soulignent Poggi & D'Errico (2010) une forme d'agressivité est exprimée, encore faut-il que celle-ci soit contrôlée. Comme nous l'avons vu, plusieurs des extraits de parole de Dominique Voynet perçus comme hostiles et néanmoins dominants se caractérisent par un débit articulatoire ralenti³ et un écart entre hostilité exprimée au niveau verbal et hostilité exprimée typique de l'ironie (Anolli, 2000), ce qui est en accord avec l'interprétation de ces extraits comme exprimant de l'agressivité contrôlée.

Vers une évaluation perceptive adaptée à la parole politique : l'auto-évaluation

L'opportunité que nous avons eue de mener une auto-évaluation des stimuli par les locuteurs du corpus en parallèle à une évaluation par des auditeurs « naïfs » nous a permis d'entrevoir selon les locuteurs, des rapports parfois discordants entre *ethos* voulu (l'image que l'on souhaite véhiculer) et *ethos* perçu⁴ (Amossi, 1999). Si nous observons tous locuteurs confondus une corrélation significative entre mécontentement auto-évalué et perçu, nous avons notamment pour certains locuteurs pu constater des incongruences significatives entre les deux évaluations : Voynet est perçue comme plus mécontente (valence négative) qu'elle ne s'auto-évalue, Le Chequer se déclare plus calme qu'il n'est effectivement perçu. Il est intéressant de noter que Voynet est la seule locutrice pour laquelle on trouve un écart significatif entre valence auto-évaluée et valence perçue, ce qui va selon nous dans le sens de la différence de stratégie entre Maire sortante et opposants en période électorale municipale : les quatre opposants n'ont aucun mal à se déclarer mécontents durant leurs auto-évaluations. Quant à Le Chequer, on

³ Nous rappelons ici que Voynet est la seule locutrice pour laquelle le débit articulatoire est corrélé négativement à l'hostilité perçue.

⁴ Nous avons néanmoins conscience du fait que l'auto-évaluation ne correspond pas seulement à l'*ethos* voulu, et qu'elle peut même en différer si l'on pense à un locuteur qui aurait voulu apparaître calme aux yeux des spectateurs sur le moment et qui s'auto-évaluerait néanmoins comme énervé.

rappellera que s'il n'est certes pas perçu comme plus calme que les autres locuteurs du corpus, il reste le seul locuteur pour lequel l'hostilité perçue n'est pas corrélée à l'excitation perçue.

Enfin, concernant l'auto-évaluation du contrôle de soi, nous avons vu que cette dimension entretient des rapports variables avec la dominance perçue selon les locuteurs. Le Chequer et Serey sont ainsi perçus comme plus dominants dans le cas des stimuli pour lesquels ils déclarent un contrôle de soi maximal. À l'inverse, pour Tuailon ce sont les stimuli pour lesquels il déclare un contrôle de soi minimal qui sont perçus comme les plus dominants, ce qui pourrait être interprété pour ce dernier comme une non-nécessité de contrôle. En revanche, Serey est le locuteur pour lequel on trouve la corrélation significative la plus forte entre contrôle de soi et excitation auto-évaluée, signifiant que lorsque celui-ci se déclare agité il considère néanmoins exercer un contrôle sur l'expression de ses émotions.

La participation des locuteurs à l'auto-évaluation a été une des étapes les plus enrichissantes de ce travail de recherche, et l'une des plus agréables à mener du fait de l'atmosphère joviale dans laquelle se sont déroulés les tests d'écoutes durant les entretiens⁵. Si nous craignons au départ une réticence et une gêne de la part des locuteurs⁶, ceux-ci se sont tous pliés à l'exercice avec curiosité avec, pour certains, l'impression exprimée de remplir un « quiz de personnalité ». Nous notons également que tous les locuteurs ont fréquemment indiqué reconnaître instantanément le contexte précis dont était extrait le stimulus écouté, surtout lorsque celui-ci correspondait aux moments les plus vifs du débat.

Si nous devions réitérer une telle tâche d'auto-évaluation ou simplement effectuer des recommandations pour la mise en place d'un protocole de ce type, a posteriori nous recommanderions alors d'ajouter deux échelles perceptives aux grilles d'auto-évaluation :

- Une échelle de **souvenance** du contexte précis dont est extrait le stimulus, allant de « je ne me rappelais même pas avoir dit ça » à « je me souviens parfaitement de ce moment et des termes employés »
- Une échelle de degré de **spontanéité** des productions allant de la lecture oralisée à des propos tenus dans l'émergence soudaine d'un état émotionnel⁷.

⁵ Nous revenons plus bas sur l'enregistrement de la condition de contrôle ; voir section 0.

⁶ Tous ont déclaré spontanément ne pas apprécier écouter leur propre voix et tous ont fait des remarques sur leur façon de parler auto-perçue comme criblée de pauses remplies, incorrecte syntaxiquement. En réponse à cela nous nous sommes permise de leur faire part des remarques de Blanche-Benveniste (2010) sur la syntaxe de l'oral pour leur permettre de relativiser leurs impressions.

⁷ Ce type d'échelles a notamment fait ses preuves dans la description et distinction des styles de parole selon la situation de communication proposée par Koch & Oesterreicher (2008), comme on peut le voir dans des études telles que celles de Bardiaux (2010) ou Degand et collègues (2014), et gagnerait selon nous à être plus utilisé en perception.

Si ces deux échelles nous semblent importantes à tester dans une auto-évaluation pour considérer leurs rapports avec les profils émotionnels, l'échelle de spontanéité pourrait de plus être confrontée directement à la perception d'auditeurs « naïfs ».

Vers une évaluation perceptive adaptée à la parole politique

Au niveau de la perception par les auditeurs naïfs, notre étude a montré l'utilité d'une évaluation dimensionnelle (valence, excitation, dominance) d'extraits de parole politique, notamment pour travailler ensuite sur une analyse des caractéristiques vocales et pour identifier des profils différents selon les locuteurs. Outre le profil particulier de Dominique Voynet par rapport aux autres locuteurs concernant les rapports entre dominance perçue et hostilité perçue, on peut également rappeler ici le profil de Gaylord Le Chequer, seul locuteur pour qui l'hostilité perçue n'est pas corrélée à l'excitation perçue, ce qui peut être interprété comme une façon particulièrement contrôlée de produire des attitudes agressives.

Toutefois, notre étude ne permet pas de distinguer les locuteurs sur le plan de la dominance perçue, tous ayant été perçus comme légèrement dominants par les auditeurs « naïfs ». Cette homogénéité pourrait s'expliquer en premier lieu selon nous par l'influence du style de parole politique, et son impératif, on l'a vu, d'apparaître essentiellement dominant⁸. Il pourrait être intéressant d'effectuer des analyses perceptives analogues à celles menées dans notre étude, mais incluant cette fois-ci directement les autres dimensions principales de l'*ethos* telles que la compétence et la bienveillance, ces deux dernières pouvant contribuer indirectement à la perception de la dominance. Nous n'avons pas trouvé d'études à ce jour questionnant les rapports entre dimensions émotionnelles de type valence, excitation et dimensions de bienveillance et de compétence.

De plus, concernant ces dimensions de bienveillance et de compétence, la question de leur évaluation perceptive se pose. Peuvent-elles s'évaluer de la même façon que la dominance et les dimensions principales utilisées en psychologie, pour lesquelles des extraits courts peuvent s'avérer suffisants, ou nécessitent-elles plus d'informations contextuelles et lexicales que ne permettent de fait une approche par test de perception sur stimuli constitués de phrases tronquées ? En d'autres termes, quelle peut être l'importance de l'expressivité vocale sur l'image de bienveillance et de compétence renvoyée par le locuteur ?

⁸ On pourra également penser à la nécessité pour les auditeurs, lorsqu'on traite de la question de la dominance en politique, de pouvoir comparer les locuteurs pour déterminer qui a « plus de pouvoir » que l'autre, ce qui demanderait un protocole expérimental très différent de celui-ci ; nous reviendrons sur ce point dans l'ultime section de cette discussion.

Expressivité perçue à l'oral et à l'écrit : comment expliquer leurs rapports ?

L'évaluation parallèle des transcriptions orthographiques isolées et des extraits audio originaux a permis d'obtenir des informations sur les poids respectifs des indices verbaux et des indices vocaux, notamment dans l'expression de l'hostilité.

Le fait que les locuteurs aient été perçus comme exprimant en moyenne une légère hostilité à l'écrit comme à l'oral confirme la multimodalité de l'expressivité en parole naturelle.

Avant de présenter en conclusion finale nos recommandations pour une étude appliquée des caractéristiques vocales expressives prenant en compte l'expressivité lexico-syntaxique⁹, nous nous interrogeons d'abord dans la présente section sur comment mieux appréhender la variabilité des rapports entre expressivité vocale et expressivité verbale et prévoir certaines tendances en parole politique.

Nous pourrions ici, dans la continuité de notre appui sur leur approche sociocognitive, suivre les hypothèses et observations de Poggi et collègues (2011) concernant un recours variable à l'expressivité paralinguistique selon le type d'attaque effectuée à l'encontre de l'opposant dans le débat. Pour préserver une image de soi favorable tout en nuisant à celle de l'opposant, les chercheuses prévoient une exploitation beaucoup plus tacite des caractéristiques vocales expressives si l'on s'attaque directement à sa dominance, par exemple en le ridiculisant, que si l'on s'attaque à sa bienveillance. Il s'agit notamment, comme on l'a vu dans l'optique de Ballet (2012), de convaincre l'auditoire que l'opposant traite injustement les citoyens et qu'il faut s'en indigner.

Si ces hypothèses se confirment, une annotation par type d'attaque à l'*ethos* pourra permettre de mieux prévoir le degré d'expressivité vocale véhiculée dans les extraits pour mieux en appréhender l'analyse en perception comme en production.

Aussi, si dans certains types d'attaques on s'attend à une expressivité vocale moindre, il apparaîtra d'autant plus nécessaire de pouvoir analyser en parallèle les autres niveaux par lesquels passe l'expression agressive (lexique, syntaxe, recours implicite au contexte, etc.). Pour tous ces aspects, une approche complémentaire et parallèle menée avec des chercheurs en analyse du discours pourrait s'avérer fructueuse.

⁹ Le parti pris qui sera le nôtre dans les conclusions de ce travail et nos ultimes recommandations est que l'analyse fine des rapports entre expressivité lexico-syntaxique d'une part et vocale de l'autre est la clé d'une meilleure description des caractéristiques vocales des attitudes contrôlées pour lesquelles on sait qu'un consensus à travers les études strictement phonétiques est difficile à observer (Scherer, 2003).

IX.2. Avantages et limites de la condition de contrôle (corpus relecture)

Si l'évidence de l'apport du corpus relecture dans l'acquisition de mesures comparées apparaît dans les chapitres d'analyse de la thèse, nous prenons ici le temps d'apporter quelques remarques sur les limites de l'utilisation d'une condition de contrôle lue et non spontanée, notamment dans la difficulté de validation perceptive de celle-ci.

Choix de la condition de contrôle

L'évaluation perceptive des caractéristiques vocales expressives des locuteurs du corpus selon l'hostilité, la valence, l'excitation et la dominance perçue a montré son utilité. La comparaison acoustique et prosodique entre le corpus spontané et sa relecture a permis de quantifier l'écart entre productions expressives et lecture oralisée présumée « neutre » pour chaque locuteur, sur un même matériel sémantico-syntaxique. L'intérêt fondamental de l'enregistrement de cette condition a consisté à pouvoir ensuite effectuer des comparaisons systématiques sur les unités prosodiques majeures et les unités syntaxiques sur des contenus similaires. De plus, les mesures de différences original-relecture effectuées sur les caractéristiques prosodiques et temporelles étudiées nous ont permis de dégager des corrélats significatifs de la perception de l'hostilité et de la dominance, à travers des contrastes d'étendue de la plage de variation de F0 et de débit articulatoire de part et d'autres des pauses silencieuses, très différents dans les extraits originaux et dans les extraits relus.

À notre sens, le **problème** soulevé par la question de la **condition de contrôle** réside du rapport contradictoire en phonétique entre **écologie des productions** analysées et possibilités d'investigation fine des caractéristiques de parole. En effet, une première question que nous nous sommes posée : **pourquoi ne pas prendre des extraits du corpus où les locuteurs ne sont pas perçus comme hostiles comme condition de contrôle ?**

Si nous insistons à ce sujet sur le fait que seul un contenu lexico-syntaxique similaire peut permettre une analyse comparative quantitative pertinente des caractéristiques segmentales et suprasegmentales, il n'a pas été chose aisée dans la sélection des extraits de notre corpus original de trouver des portions de parole à la fois non agressives et non lues. Lors de la sélection des locuteurs, beaucoup d'entre eux (rappelons qu'il y avait 50 conseillers municipaux dotés de micros durant les séances du conseil municipal étudiées en 2013) ont été écartés du fait qu'ils intervenaient uniquement en produisant des lectures déclamées de leurs interventions pré-écrites. Si rien n'assure pour tous les stimuli de notre corpus final que ceux-ci sont produits de façon absolument spontanée, les cinq locuteurs de la présente étude ont été

sélectionnés pour leur prédilection pour un discours de style spontané, des interventions fréquentes, ainsi que des rapports conflictuels (entre la Maire et les quatre opposants).

Ceci nous amène à une deuxième question complémentaire : **pourquoi comparer les productions expressives et la parole lue présumée « neutre »** ? Question à laquelle nous répondons qu'en prenant pour impératif d'avoir des contenus comparables, il est difficile de faire autrement sans multiplier les approches : à savoir, travailler d'abord sur des corpus de parole naturelle puis sur des reproductions ciblées en laboratoire, comme nous le préconiserons, entre autres, au terme de cette discussion.

Une expérience éclairante de « contextualisation » de la parole « neutre »

C'est justement à cause des biais perceptifs dus à l'évaluation de la parole présumée « neutre » et lue (Banziger et Scherer, 2007 ; Kouklia & Audibert, 2013), que nous avons fait le choix dans notre présente étude de dissocier ce qui serait de l'ordre d'une condition de contrôle en perception (scores d'hostilité perçue à l'écrit) d'une condition de contrôle en production (corpus relecture)¹⁰.

Pour Banziger & Scherer (2007), la parole présumée « neutre » ne peut être un référent fiable pour être comparée ensuite aux productions expressives du fait de la sous-spécification de sa nature en terme dimensionnel pour lesquels on ne peut que prévoir une valence neutre, et éventuellement un léger contrôle sur les productions car non affectées par une quelconque réaction ou posture émotionnelle¹¹.

Partant de la prise de position de Banziger & Scherer (2007), nous avons décidé de consacrer une courte étude, en parallèle à nos recherches doctorales, consacrée à l'analyse perceptive de la parole dite « neutre ». Nous avons cherché à établir si la nature de la parole dite « neutre », généralement lue dans les études, ne pouvait pas influencer la spécification de celle-ci comme légèrement contrôlée (Kouklia & Audibert, 2013). Nous avons participé à l'enregistrement par des comédiens d'un corpus de parole expressive simulée en laboratoire¹² comprenant l'expression de l'anxiété, de l'irritation, du contentement et de l'abattement. Ce corpus incluait ces productions d'affects de degrés d'activation émotionnelle variables et une condition de contrôle « neutre » :

- Une version lue, typique des études en parole expressive de laboratoire, au cours de laquelle les comédiens ont relu les différentes phrases sur lesquelles ils ont travaillé en parole expressive (« *il est huit heures.* », « *j'entre dans le bâtiment.* », « *le papier est sur table.* ») avec pour seule consigne de le faire « de façon neutre » et de marquer des temps de pause entre chaque phrase pour éviter un éventuel effet de liste.

¹⁰ Ce choix n'exclut pas la pertinence de la mise en place d'une validation perceptive du corpus « relecture » que nous avons enregistré avec les locuteurs mais s'explique du fait des obstacles de fond et de forme que nous avons rencontrés et que nous décrivons ici.

¹¹ Banziger & Scherer (2007) iront jusqu'à remplacer la condition « neutre » dans leurs expérimentations par une condition de léger contentement pour lesquels il est possible de mieux spécifier les caractéristiques attendues avec une valence, une excitation et une dominance légèrement positives.

¹² Le corpus en question était constitué de variations expressives sur des voyelles tenues produites isolément pour pouvoir ensuite être testées dans une expérience neuro-phonétique de perception des affects chez des patients bipolaires (Carminati et collègues, 2018). Pour faciliter la tâche de production aux comédiens participants à l'étude, nous les avons fait travailler sur un ensemble de phrases et de scénarios inspirés de l'étude de Fonagy et Bérard (1972), que nous avons extraites pour les besoins de notre étude.

- Une version dialoguée, pour laquelle ceux-ci avaient pour tâche de produire les phrases en répondant à des questions formulées par les expérimentatrices (« *Quelle heure est-il ?* », « *Où es-tu ?* », « *Où est le papier ?* ») avec pour consigne d'acteur de répondre comme s'ils avaient simplement pour but de fournir une réponse à la personne, sans avis ni affect particulier à l'égard de la situation.

Nous avons ensuite fait passer un test de perception dimensionnelle (valence, excitation, dominance) à 23 auditeurs « naïfs ». Le corpus comprenant a) l'ensemble des expressions d'attitudes contrôlées (anxiété, irritation, contentement, abattement) produites par les comédiens et b) et c) les deux conditions « neutres » (« neutre lue », « neutre dialoguée »). Les résultats de l'analyse perceptive ont montré que

- la version « neutre dialoguée » a été perçue comme neutre sur les trois dimensions de valence, excitation, dominance.
- La condition « neutre lue » a été perçue comme neutre sur le plan de la valence et de l'excitation. Au niveau de la dimension de dominance, elle a été perçue comme exprimant une légère soumission, sans différence significative avec l'expression d'abattement (tristesse faiblement activée), variante expressive perçue comme la plus soumise du corpus.

Comment expliquer alors que la condition « neutre lue » ait été perçue comme soumise ? Nous pouvons expliquer ce biais perceptif en l'appréhendant dans l'optique socio-cognitiviste (Castelfranchi, 2003 ; Poggi et collègues, 2007, 2010, 2011) : si les individus s'expriment toujours pour accomplir des buts personnels et interpersonnels selon les règles de la situation de communication et les croyances de leurs cultures, le risque encouru de faire évaluer l'expressivité d'extraits de parole lue « neutre » à des auditeurs « naïfs » est que ceux-ci tentent plus ou moins inconsciemment de reconstituer une raison, un but derrière la production « neutre lue » du locuteur, alors qu'il n'y en a pas. Il nous apparaît surprenant, mais cohérent à travers cette conception qu'il puisse suffire de poser des questions à des comédiens, de faire passer leur but de « je lis la phrase parce que l'expérimentateur me le demande » à « je joue une situation dans laquelle je réponds à une question d'ordre strictement informationnel » pour que leurs productions « neutres » soient déjà mieux perçues comme telles.

Cependant, si le recours au dialogue joué a été très simple à mettre en place dans cette étude passée (Kouklia & Audibert, 2013), la réplique d'une telle méthodologie pour la condition de contrôle de notre corpus d'extraits de débats nous a paru beaucoup plus douteuse.

Si Enos et Hirschberg (2006), comme Banziger et Scherer (2007), insistent sur le fait que la parole expressive actée ne peut maintenir sa pertinence dans le domaine de l'analyse de la

parole expressive que tant qu'un travail assidu de direction d'acteur est effectué, nous avons fait le choix résolu avec nos locuteurs politiques de ne pas essayer de mettre en place de productions expressives actées, quand bien même celles-ci se limiteraient à la production d'une condition « neutre ».

Renonçant à un quelconque artefact d'écologie de production du présent corpus « relecture »¹³ de la présente thèse, nous avons enregistré celui-ci avec les locuteurs durant des entretiens individuels en pleine conscience des intérêts et limites d'une telle condition de contrôle.

Des difficultés de la validation perceptive de la condition de contrôle

Une validation perceptive est une nécessité avant toute analyse phonétique. Afin d'étudier la faisabilité d'une évaluation perceptive combinant enregistrements originaux et relecture, nous avons effectué un pré-test avec 5 auditeurs « naïfs » en adoptant un protocole de discrimination dimensionnelle de paires de stimuli originaux versus relus¹⁴. Dans ce pré-test, les stimuli étaient joués consécutivement et les auditeurs devaient considérer sur trois échelles de Likert à 5 points si dans le second extrait le locuteur avait l'air plus ou moins mécontent, agité et dominant que dans le premier. Les réponses des participants du pré-test se sont avérées problématiques à interpréter : la question post-test sur les indices exploités par les auditeurs pour différencier les productions originales des productions relues au sein de chaque paire, tous les auditeurs ont fait état d'un bruit de fond présent lors des débats. Nous avons donc ajouté du bruit de fond distinctif dans les stimuli du corpus relecture et refait passer l'expérience. Trois auditeurs sur cinq ont déclaré spontanément avoir eu du mal à se souvenir de ce qu'ils avaient perçu du premier extrait de la paire écoutée. Ceci peut aisément s'expliquer par le fait que la durée des paires concaténées pouvait aller jusqu'à 25 secondes ; un test de discrimination typique en phonétique comportant généralement des extraits beaucoup plus courts (voyelles, mots isolés).

Pour évaluer l'expressivité véhiculée par la condition de contrôle, nous avons fait le choix de faire évaluer par un panel de sujets naïfs les transcriptions orthographiques, qui nous a permis de quantifier le poids sémantique des énoncés. Nous n'avons donc pas testé directement le corpus relecture en perception, et ne disposons donc que de l'évaluation de la transcription écrite plutôt que d'une validation perceptive spécifique à celui-ci. Aussi, d'un point de vue strictement formel, nous ne pouvons assurer que le corpus relecture est inexpressif par rapport

¹³ Il n'est pas possible de refaire des enregistrements contrôle dans la salle des fêtes de la Mairie, ni de demander aux locuteurs de rejouer seuls « les scènes » desquelles étaient extraits leurs stimuli.

au corpus original qu'en nous appuyant que sur des éléments de distinction situationnelle entre les deux conditions¹⁵ :

- Dans le **corpus original**, les locuteurs s'expriment dans un style de parole spontanée, avec des buts et enjeux électoraux et interactionnels forts, dans un contexte public, avec une distance élevée entre eux et le public justifiant l'installation de hauts parleurs, et dans des états émotionnels variables¹⁶.
- Dans le **corpus relecture**, les locuteurs produisent une lecture oralisée d'extraits de leur propre parole, avec pour but de répondre à notre demande de relecture, avec la consigne de ne pas exprimer d'expressivité particulière et de lire à la manière d'une dictée, dans une pièce calme avec un micro-casque.

Nous recommandons à tout chercheur souhaitant mettre en place un protocole similaire de faire produire aux locuteurs un enregistrement supplémentaire avec pour support un texte similaire pour tous. Un tel enregistrement aurait pu servir à tester ce que des auditeurs perçoivent de la voix des locuteurs à partir d'un même support textuel, et ainsi tester une certaine caractérogie vocale : certains locuteurs pouvant être perçus comme plus enjoués ou dynamiques que d'autres sur un contenu sémantique similaire (Fónagy, 1983). Il aurait pu aussi servir de base à des analyses systématiques fines comparatives des comportements des locuteurs en condition de lecture. Le résultat de ces comparaisons aurait pu s'avérer précieux dans l'analyse comparée de leurs caractéristiques de structuration prosodico-syntaxique qui a suivi entre corpus original et relecture, que nous allons à présent discuter.

¹⁵ Les considérations suivantes sont inspirées du modèle de Koch & Oesterreicher (2008) dont une traduction synthétique à visée pédagogique a été fournie par Modicom (2015) pour promouvoir leur grille de description situationnelle pour les chercheurs non germanophones. Nous précisons que leur modèle se veut non exhaustif et peut inclure selon les spécificités de la situation de communication des échelles de description variées tant que celles-ci illustrent d'une façon ou d'une autre une distinction graduelle entre deux polarités : le domaine de l'oral (spontané, direct, personnel, investissement émotionnel, etc.) et le domaine de l'écrit (formel, pré-écrit, public, pas d'investissement émotionnel etc.). Ces grilles sont couramment utilisées pour distinguer les styles de parole.

¹⁶ Deux locuteurs sur cinq ont déclaré qualitativement (durant l'entretien mais hors de l'évaluation dimensionnelle) avoir été soumis à une tension émotionnelle très accrue pendant toutes les séances du conseil étudiées ici.

IX.3. Arguments pour une analyse multifactorielle de la pause silencieuse

Au sortir des analyses prosodiques de cette recherche, nos hypothèses sur l'exploitation des pauses silencieuses en tant que signal de dominance (Duez, 1991) ne sont que partiellement confirmées. Dans notre corpus, la pause n'est pas directement corrélée à la dominance. En revanche, nous retrouvons son impact dans les variations prosodiques de part et d'autre des pauses, notamment au travers de forts contrastes de plage de variation de F0 pré et post-pausaux.

Dans cette partie de la discussion, notre objectif est d'inciter à une méthodologie d'analyse de la pause silencieuse permettant d'analyser plus finement les rapports entre ses fonctions, sa durée, sa localisation dans la structure syntaxique, et ses corrélats prosodiques en contexte gauche et droit.

Conformément aux conclusions qui étaient déjà celles de Duez (1991) et qui sont énoncées en synthèse de la récente revue de la littérature sur les pauses effectuées par Grosman et collègues (2018), les rapports entre caractéristiques et fonctions des pauses sont complexes : les rapports entre localisation syntaxique et fonctions paralinguistiques sont non exclusifs. Pour appuyer ce point de vue, nous posons au regard des caractéristiques observées quantitativement et qualitativement dans nos données, les jalons d'une approche cumulative des critères d'analyses des pauses pour contribuer à la détermination de leurs fonctions. Enfin, nous proposons un ensemble d'approches plurielles et pluridisciplinaires pour pouvoir notamment traiter plus finement la question de la perception de la fonction paralinguistique des pauses.

Pour une actualisation des critères d'annotation des frontières prosodiques

Vers une remise en question d'un seuil de durée minimal

Pour des raisons de comparabilité avec la majorité des études portant sur la pause et les phénomènes de structuration prosodique et syntaxique majeures, nous avons conservé dans cette étude un **seuil de durée limite** (140 ms) pour appréhender le phénomène de la pause silencieuse. Toutefois il nous semble impératif pour notre domaine de recherche de pouvoir proposer une approche plus dynamique de la durée des pauses. Duez (1991) dès ses premières recherches, avait insisté sur l'influence du débit articulatoire et des caractéristiques temporelles générales de chaque locuteur, et de telles précautions nous semblent nécessaires à mettre en place.

Si nos observations gagneraient à être confrontées à d'autres corpus de parole politique en situation de débat, la conclusion à laquelle nous aboutissons d'ores et déjà à l'issue de notre recherche est que le seuil de durée minimale des pauses, qui est à 200 ms dans la majorité des études (voir la revue de Grosman et collègues, 2018) n'est pas toujours adapté à la parole spontanée et entraîne le risque d'une non-considération d'un nombre considérable de pauses intra-syntaxiques. Ce seuil a souvent été choisi de façon pratique, pour que les pauses puissent être repérées automatiquement par programmes informatiques et non confondues avec le silence des occlusives sourdes, dont la durée n'excède que rarement 200 ms. En réalité, une pause peut être de durée bien moindre, et quelques millisecondes de glottalisation peuvent être perçues comme une pause. Les pauses de focalisation, localement juste avant le mot à mettre en valeur peuvent également être de durée très courte, bien en deçà de 50 millisecondes, et clairement perçues comme des pauses (Vaissière, 2018, communication personnelle).

La définition de critères robustes permettant d'appréhender de façon plus dynamique des seuils pertinents pour les paramètres mélodiques et temporels engagés dans la production des frontières prosodiques est tout sauf triviale. Pour aboutir à cela, nous recommandons la mise en place d'études mettant en rapport analyses acoustiques et perception des fonctions prosodiques correspondantes sur des corpus constitués dans le but précis de pouvoir effectuer des comparaisons systématiques selon la configuration de chaque critère d'annotation prosodique testé (pause silencieuse, F0, durée).

Des rapports non exclusifs entre fonctions et localisation syntaxique des pauses

En concentrant nos recherches sur des unités définies autour de la pause silencieuse, nous avons principalement analysé les caractéristiques générales mélodiques et rythmiques (durées et débits syllabiques) en les appréhendant comme des corrélats des pauses silencieuses. L'émergence d'accents variés, allant d'accents didactiques (Lucci, 1980) à des renforcements initiaux dus à l'émergence d'hostilité (Fónagy, 1983) en passant par des accents de focalisation (Léon, 1971), n'a pas pu être étudiée à sa juste valeur du fait d'une analyse des caractéristiques mélodiques et temporelles pré-pausales et non post-pausales : l'annotation des frontières prosodiques que nous avons effectuée se concentrait principalement sur les caractéristiques en contexte gauche des pauses silencieuses et nous avons annoté les frontières majeures seulement. Nous n'avons pas fait une étude séparée des phénomènes de renforcements initiaux, dont la présence impacte également les caractéristiques segmentales, mélodiques mais également spectrales au sein des syllabes affectées. En effet, comme le souligne Vaissière (2010), l'accentuation initiale renforcée type accent d'insistance peut notamment se traduire par un saut de F0 en début de mot mais également sur le plan articulatoire un contact accru entre la langue et le palais, ainsi que par le dévoisement des consonnes initiales ; phénomène d'ailleurs attesté par Fonagy (1983) comme caractéristiques expressive agressive (voir partie II.3.3.2 page 30).

Bien qu'elle n'ait pas été ici directement analysée, l'influence des accents initiaux est toutefois présente en filigrane dans toutes nos analyses, notamment à travers la mesure des différences de durée ou encore de fréquence fondamentale de part et d'autre des frontières considérées.

En d'autres termes, il apparaît évident à l'issue de notre étude que l'approche complémentaire prioritaire à apporter à l'analyse des pauses au travers des frontières prosodico-syntaxiques majeures doit consister en une analyse approfondie du contexte droit de celles-ci. Lorsqu'il est précédé d'une pause, le renforcement initial, qu'il ait une fonction identificatrice indexant le style politique (accent didactique), impressive (accent de focalisation attirant l'attention du locuteur sur un élément du discours), ou pragmatique (maintien forcé du tour de parole) est un indice non négligeable dans l'appréhension des fonctions des pauses, qui accompagnent souvent le renforcement initial.

Dans l'approche que nous avons adoptée, nous sommes principalement partis de la localisation des pauses en rapport avec la syntaxe. L'étude des pauses doit être faite également en fonction de la présence de renforcements initiaux dans le mot suivant ou précédent la pause. En effet, nous avons pu observer qualitativement l'occurrence de pauses silencieuses que l'on pourrait qualifier de focalisation postérieure, c'est à dire congruentes avec la structure

syntaxique, néanmoins suivant directement la production d'un mot avec accent de focus, comme pour renforcer ce focus grâce au silence suivant.

Pour une approche cumulative et multifactorielle

Suite aux observations ayant et n'ayant pu être effectuées au sein de la présente étude, nous considérons que si les caractéristiques des pauses prises isolément ne peuvent prédire précisément leur fonction (nous pensons ici notamment à la durée mais également à la position dans la structure syntaxique), cela pourrait être un ensemble de caractéristiques. Par exemple, comme nous avons pu l'observer qualitativement dans notre étude, les pauses intra-syntaxiques peuvent tout aussi bien correspondre à des cas de pauses de focalisation ou de pauses de jonction externe précédant un accent didactique, voire même à des cas de pauses dites d'hésitation. Quels sont les critères pertinents pour qui différencier ces différents phénomènes ? Dans le cas de l'exemple que nous venons de citer, cela pourrait être une combinaison de localisation syntaxique, de présence ou absence de proéminences post-pausales, de différences de débit articulatoire, etc.

Nous suggérons ici l'intérêt d'une approche multifactorielle reposant sur l'annotation indépendante puis cumulée d'un ensemble hétérogène de facteurs et caractéristiques allant des contraintes physiologiques aux stratégies locutoires expressives des locuteurs incluant notamment pour chaque pause :

- Leur durée brute et leur durée relative à la durée de l'extrait traité, mais également à la durée du bloc précédent
- La forme des contours pré-pausaux (indiquant la montée de continuation ou la chute finale) et post-pausaux
- La comparaison de la F0 moyenne et de la plage de variation de F0 pré et post-pausales, les changements de registres
- La comparaison du débit articulatoire syllabique et de l'indice de variabilité paire à paire des durées pré et post-pausales
- La localisation syntaxique fine de toutes les pauses
- La présence d'indices audibles de respiration ou de déglutition, de coup de glotte et de glottalisation, de schwas épenthétiques pré-pausaux
- Le repérage de pauses remplies ou non remplies dites d'hésitation
- La présence d'accents primaires mélodiques pré ou post-pausaux.

IX.4. Vers une approche plurielle et pluridisciplinaire de l'expressivité vocale

Au sortir de cette recherche, nous insistons sur la nécessité impérative d'effectuer des analyses comparatives fiables et ciblées, que seul un corpus de laboratoire peut permettre. Si nous avons fait état de ses limites, nous insistons néanmoins sur le fait que l'accès à la condition de contrôle (relecture) nous a permis d'obtenir des informations cruciales sur les écarts produits par les locuteurs dans les extraits où ils ont été perçus comme dominants et/ou hostiles, comparativement à leurs propres caractéristiques de parole en lecture.

Toutefois, de par la nature complexe et hétérogène des énoncés constituant un corpus de parole naturelle comme celui que nous avons analysé ici, ni les analyses perceptives, ni les analyses prosodiques ne nous permettent, en l'état, d'établir des hypothèses fiables sur les liens directs entre caractéristiques prosodiques dues à l'insertion de pauses et perception de valeurs paralinguistiques, comme l'expression de l'hostilité ou de la dominance.

La clé du progrès réside dans une approche **plurielle**, alliant étude de corpus phonétique de parole spontanée et phonétique de laboratoire, et étude **pluridisciplinaire**. Il s'agit de favoriser la communication entre approches *descendantes*, telles que celles proposées en psychosociologie pour appréhender les enjeux qui sous-tendent les interactions, et nos méthodes, *ascendantes*, pour analyser la réalité perceptive et acoustique des phénomènes expressifs.

Pour une analyse phonétique plurielle

Arrivés au terme de cette étude, nous insistons sur l'importance d'une approche plurielle, et la nécessité d'appréhender de manière complémentaire et non dichotomique les différences entre parole naturelle et parole de laboratoire pour une approche plus fine des caractéristiques vocales expressives. Dans le cas de la parole politique et de l'impact de l'expressivité sur l'image transmise par les locuteurs, nous considérons que la pause silencieuse est un paramètre important à confronter avec sa perception, qu'il faut appréhender en prenant mieux en compte ses contextes droits et gauches afin de rendre compte de sa présence due aux phénomènes d'accentuations paralinguistiques.

Par une telle approche plurielle, notre mode d'investigation de la parole expressive, notamment politique, pourrait comprendre :

- 1) Une approche quantitative et qualitative sur un corpus de parole naturelle de taille réduite tel que celui que nous avons construit.
- 2) Une analyse quantitative sur des corpus de parole naturelle plus vastes préalablement annotés, comme le corpus ETAPE¹⁷ (Gravier et collègues, 2012) contenant plusieurs heures de parole politique en débat, ou encore le corpus LOCAS-F (Degand et collègues, 2014).
- 3) Une analyse ciblée en laboratoire permettant par la construction de corpus contrôlés, l'investigation perceptive de l'influence des différents paramètres des pauses silencieuses dans la perception de la parole expressive des locuteurs. Notons qu'il est également facile de manipuler la présence et la longueur des pauses d'un corpus lu ou spontané et de juger de l'effet perceptif obtenu.

¹⁷ Nous pensons ici également aux travaux menés par Wu et collègues (2016) sur l'analyse du schwa au travers notamment de ce corpus.

Vers une approche pluridisciplinaire

L'analyse perceptive et productive de l'expressivité vocale dans le débat politique, et plus généralement dans tout contexte d'interaction, exige selon nous un dialogue actif entre disciplines, favorisant tout à la fois l'accroissement de nos connaissances conceptuelles des règles et enjeux de la situation de communication et des analyses fines et contrôlées. Un travail de recherche s'impose tant dans une optique *Top-Down*, par une meilleure appréhension des buts implicites et tactiques des locuteurs durant dans un débat, que dans une approche *Bottom-Up*, en approfondissant dans le cadre d'études ciblées en laboratoire les observations parfois qualitatives et relatives que nous avons pu déjà obtenir à partir de données issues du terrain. Nous estimons particulièrement nécessaire, à ce sujet l'élaboration en laboratoire de corpus permettant l'analyse comparative systématique de l'influence de la réalisation des frontières prosodiques (rapports syntaxe/prosodie, caractéristiques mélodiques et temporelles) sur la perception attitudinale.

À notre sens, seul un dynamisme pluridisciplinaire permettra de passer outre les obstacles méthodologiques actuels pour permettre une appréhension pertinente de l'expressivité vocale et une analyse plus fine des caractéristiques phonétiques relatives à la parole issue de situations d'interaction naturelle.

IX.5. Principaux apports de cette étude

Selon nous, les contributions les plus marquantes du travail de recherches doctorales que nous avons mené se situent aux niveaux suivants.

L'ensemble de notre protocole a permis la mise en place d'analyses phonétiques appliquées à un corpus de parole naturelle. L'ajout d'une condition de contrôle a permis d'effectuer des comparaisons systématiques (généralement non réalisables sur un corpus de parole naturelle produite hors laboratoire) et de permettre ainsi une analyse plus fine de la variation expressive au niveau segmental et suprasegmental d'extraits de parole issus d'une interaction réelle. Outre le fait qu'elles ont permis une objectivation et une validation des données en termes d'hostilité, de dominance, d'excitation et de valence, l'ensemble des analyses perceptives mises en place ont été utilisées comme des clés pour l'analyse acoustique des données. L'ajout d'une condition écrite à l'analyse de l'hostilité perçue a permis une quantification de l'expressivité inhérente au contenu sémantique et celle ajoutée par les indices paralinguistiques. La comparaison des profils émotionnels auto-évalués et perçus par les auditeurs naïfs a permis quant à elle de quantifier l'écart entre *ethos* voulu (image voulant être véhiculée par la personnalité politique) et *ethos* perçu (Amossi, 1999).

Par une approche combinée entre analyses perceptives dimensionnelles et analyses prosodiques, nous avons pu dégager un ensemble de corrélats de la perception de l'hostilité et de la dominance dans des extraits de parole politique en débat : débit articulatoire plus variable, raccourcissement des allongements pré-pausaux, plage réduite de F0 et/ou variations fortes de part et d'autre des pauses silencieuses. Les corrélats de l'hostilité et de la dominance que nous avons mis en évidence dans les extraits analysés se rapprochent tout autant de caractéristiques attendues pour la production d'attitudes agressives contrôlées (Fónagy, 1983 ; Scherer, 2003) que de caractéristiques relevées dans le cas d'insertions de pauses de focalisation ou de maintien du tour de parole, dans la parole spontanée et/ou politique (Ferré, 2004 ; Béchet, 2013).

IX.	Discussion et Conclusion.....	185
IX.1.	Expressivité vocale et perception des personnes politiques en débat.....	186
IX.2.	Avantages et limites de la condition de contrôle (corpus relecture).....	191
IX.3.	Arguments pour une analyse multifactorielle de la pause silencieuse.....	197
IX.4.	Vers une approche plurielle et pluridisciplinaire de l'expressivité vocale	201
IX.5.	Principaux apports de cette étude	204

Références bibliographiques

- Abelson, R. P. (1979). Differences Between Belief and Knowledge Systems. *Cognitive Science*, 3(4), 355-366.
- Abelson R. P. (1983). Whatever Became of Consistency Theory? *Personality and Social Psychology Bulletin*, 9(1), 37-54.
- Amossy, R. (1999). *Images de soi dans le discours : La construction de l'ethos*. Delachaux et Niestlé.
- Amossy, R. (2008). Argumentation et Analyse du discours : perspectives théoriques et découpages disciplinaires. *Argumentation et analyse du discours*, (1).
- Anolli, L., Ciceri, R., & Infantino, M. G. (2000). Irony as a game of implicitness: Acoustic profiles of ironic communication. *Journal of Psycholinguistic Research*, 29(3), 275–311.
- Aristote. (1991). *Rhétorique*. Le Livre de Poche.
- Arnold, M. B. (1960). *Emotion and personality*. Columbia University Press.
- Aubergé, V. (2002). Prosodie et émotion. *Actes des deuxiemes assises nationales du GdR I*, 3, 263–273.
- Audibert, N. (2008). *Prosodie de la parole expressive : dimensionnalité d'énoncés méthodologiquement contrôlés authentiques et actés*. Thèse de Doctorat. Institut National Polytechnique de Grenoble-INPG.
- Audibert, N., Aubergé, V., and Rilliard, A. (2008). *How we are not equally competent for discriminating acted from spontaneous expressive speech*, Speech Prosody 2008, Campinan.
- Audibert, N., Aubergé, V., and Rilliard, A. (2010). *Prosodic correlates of acted vs. spontaneous discrimination of expressive speech: a pilot study*, 5th International Conference on Speech, Chicago.
- Averill, J. R. (1980). A constructivist view of emotion. In R. Plutchik & H. Kellerman, *Theories of Emotion* (p. 305-339). Academic Press.

- Averill, J. R. (1983). Studies on anger and aggression: Implications for theories of emotion. *American Psychologist*, 38(11), 1145-1160.
- Ballet, M. (2012). *Peur, espoir, compassion, indignation : l'appel aux émotions dans les campagnes présidentielles (1981-2007)*. Dalloz.
- Bally, C. (1944). *Linguistique Générale et Linguistique*. Berna.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614-636.
- Bänziger, T., & Scherer, K. R. (2007). Using Actor Portrayals to Systematically Study Multimodal Emotion Expression: The GEMEP Corpus. In A. C. R. Paiva, R. Prada, & R. W. Picard (Éd.), *Affective Computing and Intelligent Interaction* (p. 476-487). Springer Berlin Heidelberg.
- Bardiaux, A. (2010). Comment parlent les hommes politiques ? Analyse prosodique de la parole politique publique : de la variation stylistique à la variation individuelle. *Recherches en Communication*, 32(32), 207–223.
- Béchet, M., Sandré, M., Hirsch, F., Richard, A., Marsac, F., & Sock, R. (2013). De l'utilisation de la pause silencieuse dans le débat politique télévisé. Le cas de François Hollande. *Mots. Les langages du politique*, 103, 23-38.
- Bigi, B., & Hirst, D. (2012). SPEECH PHONETIZATION ALIGNMENT AND SYLLABIFICATION (SPPAS): A TOOL FOR THE AUTOMATIC ANALYSIS OF SPEECH PROSODY. *Speech Prosody*, 19-22.
- Blanche-Benveniste C. (2002). Phrase et construction verbale. *Verbum XXIV-1*, 23-36.
- Blanche-Benveniste, C. (2010). *Approches de la langue parlée en français*. Paris, Ophrys.
- Braud, P. (1996). *L'émotion en politique: problèmes d'analyse*. Paris, Presses de la Fondation nationale des sciences politiques.
- Bryant, G. A., & Fox Tree, J. E. (2002). Recognizing verbal irony in spontaneous speech. *Metaphor and symbol*, 17(2), p.99–119.
- Bühler, K. (1934). *Sprachtheorie. [Theory of language.]*. Oxford, Fischer.

- Burger, M., Jacquin, J., & Micheli, R. (2011). *La parole politique en confrontation dans les médias*. Bruxelles, De Boeck.
- Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes dits d'hésitation en français oral spontané. Étude sur un corpus de récits en classe de français*. Thèse de doctorat, Université Paris III.
- Candea, M. (2017). *Pratiques de prononciation et enjeux sociaux. Approches post-variationnistes en sociophonétique du français de France*. Thèse d'HDR, Université Grenoble Alpes.
- Carminati, M., Fiori-Duharcourt, N., & Isel, F. (2018). Neurophysiological differentiation between preattentive and attentive processing of emotional expressions on French vowels. *Biological psychology*, 132, 55–63.
- Carton, F. (1974). *Introduction à la phonétique du français*. Bordas.
- Castelfranchi, C. (2003). The Micro-Macro Constitution of Power. *ProtoSociology*, 18, 208-269.
- Castelfranchi, C., & Paglieri, F. (2007). The role of beliefs in goal dynamics: prolegomena to a constructive theory of intentions. *Synthese*, 155(2), 237-263.
- Charaudeau, P. (2015). Le débat présidentiel. Un combat de mots. Une victoire aux points. *Langage et société*, (151), 109-129.
- Cornelius, R. R. (1996). *The science of emotion: Research and tradition in the psychology of emotions*. Prentice-Hall, Inc.
- Cornelius, R. R. (2000). Theoretical approaches to emotion. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*. Belfast.
- Darwin, C. (1872). *The origin of species by means of natural selection: or, the preservation of favoured races in the struggle for life and the descent of man and selection in relation to sex*. Modern library.
- De Chanay, H. C., Giaufret, A., & Kerbrat-Orecchioni, C. (2011). *La gestion interactive des émotions dans la communication politique à la télévision : quand les intervenants perdent leur calme*. De Boeck Supérieur.

- De Chanay, H. C., & Kerbrat-Orecchioni, C. (2006). 100 minutes pour convaincre : l'éthos en action de Nicolas Sarkozy. In *100 minutes pour convaincre : l'éthos en action de Nicolas Sarkozy* (p. 309–329). Acta Universitatis Stokholmiensis.
- Degand, L., & Simon, A. C. (2009). On identifying basic discourse units in speech: theoretical and empirical issues. *Discours. Revue de Linguistique, Psycholinguistique et Informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics*, (4).
- Degand, L., Martin, L. J., & Simon, A.-C. (2014). Unités discursives de base et leur périphérie gauche dans LOCAS-F, un corpus oral multigenres annoté. In *SHS Web of Conferences* (Vol. 8, p. 2613–2625). EDP Sciences.
- Delattre, P. (1966). Les Dix Intonations de base du français. *The French Review*, 40(1), 1-14.
- Donot, M., & Emediato, W. (2015). La construction de la figure des leaders. Ethos, identité et charisme en perspective comparée. *Revue française des sciences de l'information et de la communication* (7).
- Duez, D. (1991). *La pause dans la parole de l'homme politique*. Editions du Centre national de la recherche scientifique.
- Duez, D. (1999). La fonction symbolique des pauses dans la parole de l'homme politique. *Faits de langues*, 7(13), 91-97.
- Duez, D. (2007). Prosodie et rythme. in P. Auzou, *Les dysarthries*. Solal.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3-4), 169-200.
- Enos, F., & Hirschberg, J. (2006). A framework for eliciting emotional speech: Capitalizing on the actors process. In *First international workshop on emotion: Corpora for research on emotion and affect (international conference on language resources and evaluation (LREC 2006)* (pp. 6–10).
- Ferré, G. (2004). Les pauses intra-constituants en anglais spontané. In *XXVes Journées d'Étude sur la parole* (1), pp. 217-220. Fès, Maroc.
- Fónagy, I. (1971). Double Coding in Speech. *Semiotica*, 3(3), 189–222.

- Fónagy, I. (1983). *La vive voix: essais de psycho-phonétique*. Payot.
- Fónagy, I. (2006). *Dynamique et changement*. Peeters.
- Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The World of Emotions is not Two-Dimensional. *Psychological Science, 18*(12), 1050-1057.
- Frick, R. W. (1986). The prosodic expression of anger: Differentiating threat and frustration. *Aggressive Behavior, 12*(2), 121-128.
- Galliano, S., Gravier, G., & Chaubard, L. (2009). The ESTER 2 evaluation campaign for the rich transcription of French radio broadcasts. In *Tenth Annual Conference of the International Speech Communication Association*.
- Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A., & Galibert, O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. In *LREC - Eighth international conference on Language Resources and Evaluation*. Turkey.
- Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*, 515-546.
- Grosman, I. (2015). (Dis)Fluencies as cognitive cues for native speakers of French . In *Proceedings of Disfluencies in Spontaneous Speech Workshop*. Edinburgh.
- Grosman, I., Simon, A.-C., & Degand, L. (2018). Variation de la durée des pauses silencieuses : Impact de la syntaxe, du style de parole et des disfluences. *Langage, 211*(3), 13.
- Guaïtella, I. (2014). Proéminences et éminences : savoir-faire discursif, faillances et défaillances des hommes politiques. *TIPA. Travaux interdisciplinaires sur la parole et le langage* (30).
- Jakobson, R. (1963). Linguistique et poétique. *Essais de linguistique générale, 1*, 209–248.
- James, W. (1884). What is an Emotion? *Mind, 9*(34), 188-205.
- Jürgens, R., Hammerschmidt, K., & Fischer, J. (2011). Authentic and Play-Acted Vocal Emotion Expressions Reveal Acoustic Differences. *Frontiers in Psychology, 2*.
- 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.

- Koch, P., & Oesterreicher, W. (2008). Mündlichkeit und Schriftlichkeit von Texten. *Textlinguistik*, 15, 199–215.
- Kouklia, C., & Audibert, N. (2013). Expressivity conveyed by contextualized vs. non-contextualized “neutral” acted speech: which control condition for expressive speech modeling? In *First International Workshop on Affective Social Speech Signals*.
- Lange, C. G. (1885). The Mechanism of the emotions. in Rand, Benjamin (Ed.) (1912). *The Classical Psychologists* (pp. 672-684). Houghton Mifflin.
- Léon, P. R. (1993). *Précis de phonostylistique : parole et expressivité*. Nathan.
- Léon, P. R. (2005). *Phonétisme et prononciations du français*. A. Colin.
- Lucci, V. (1980). L’accent didactique. *Studia Phonetica Montréal*, 15, 107-121.
- Lucci, V. (1983). *Étude phonétique du français contemporain à travers la variation situationnelle (débit, rythme, accent, intonation, e muet, liaisons, phonèmes)*. Publications de l’Université des Langues et Lettres de Grenoble.
- Maingueneau, D. (1999). Ethos, scénographie, incorporation. In *Images de soi dans le discours. La construction de l’èthos*, R. Amossy, 75-100.
- Mertens, P., & Simon, A. C. (2013). Towards automatic detection of prosodic boundaries in spoken French. In *Proceedings of the Prosody-Discourse Interface Conference* (pp. 81–87).
- Modicom, P.Y. (2015). *L’opposition oralité-scripturalité dans l’analyse de DISCOURS/TEXTES : une introduction au programme de Koch & Oesterreicher*.
- Morange, S., & Candea, M. (2010). Aux frontières de l’écoute. Durée des échantillons et choix des auditeurs : deux variables déterminantes dans la construction des tests de perception. In *Frontières. Du linguistique au sémiotique*, 79-96. Lambert-Lucas.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111(981), 855–869.

- Nauke, A., & Braun, A. (2011). The production and perception of irony in short context-free utterances. In *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, China* (pp. 1450–1453).
- Ohala, J. J. (1984). An ethological perspective on common cross-language utilization of F₀ of voice. *Phonetica*, 41(1), 1–16.
- Parisi, D., & Castelfranchi, C. (1976). *The discourse as a hierarchy of goals*. Urbino, Italia: Centro Internazionale di Semiotica e di Linguistica, Università di Urbino.
- Plantin, C. (1998). Les raisons des émotions. *Forms of argumentative discourse per un'analisi linguistica dell'argomentare*, 3–50.
- Plantin, C. (2011). *Les bonnes raisons des émotions: principes et méthode pour l'étude du discours émotionné*. Berne, Peter Lang, Sciences pour la communication.
- Plutchik, R. (1980). *Emotion, a psychoevolutionary synthesis*. New York, NY: Harper & Row.
- Poggi, I., & D'Errico, D. (2010). Dominance Signals in Debates. In A. A. Salah, T. Gevers, N. Sebe, & A. Vinciarelli, *Human Behavior Understanding* (p. 163-174). Springer Berlin Heidelberg.
- Poggi, I., D'Errico, F., & Vincze, L. (2011). Discrediting moves in political debates. In *First International Workshop on Decision Making and Recommendation Acceptance Issues in Recommender Systems (DEMRA 2011)*.
- Riegel, M., Pellat, J. C., & Rioul, R. (1994). *Grammaire méthodique du français*. PUF.
- Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3), 273-294.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction* (p. 7–55). New York, Academic Press.
- Scherer, K. R. (1999). On the Sequential Nature of Appraisal Processes: Indirect Evidence from a Recognition Task. *Cognition and Emotion*, 13(6), 763-793.

- Scherer, K. R. (2005). Appraisal Theory. In *Handbook of Cognition and Emotion* (p. 637-663).
John Wiley & Sons.
- Simon, A-C., Auchlin, A., Avanzi, M. ; Goldman, J-P (2010). *Les phonostyles : une description prosodique des styles de parole en français*. In Abecassi, M. & G. Ledegen, *Les voix des Français. En parlant, en écrivant*. 71-88.
- Simon, A. C., & Degand, L. (2011). L'analyse en unités discursives de base : pourquoi et comment ? *Langue française*, 2, 45–59.
- Tanguy, N., Van Damme, T., Degand, L., & Simon, A.-C. (2012). *Projet FRFC " Périphérie gauche des unités de discours " - Protocole de codage syntaxique*.
- Trubetzkoy, N. (1949). *Principes de phonologie*. Klincksieck.
- Vaissière, J. (1974). On French prosody, *Quarterly Progress Report*, Research Laboratory of Electronics, Massachusetts Institute of Technology 114, p. 212-223.
- Vaissière, J., & Michaud, A. (2006). *Prosodic constituents in French: A data-driven approach in Prosody and syntax: Cross-linguistic perspectives*, I. Fónagy, Y. Kawaguchi and T. Moriguchi, J. Benjamins, p.47-63.
- Vaissière, J. (2010). *Le français, langue à frontières par excellence*.
- Wu, Y., Adda-Decker, M., & Fougeron, C. (2016). Rôle des contextes lexical et post-lexical dans la réalisation du schwa : apports du traitement automatique de grands corpus. In *31èmes Journées d'Etudes sur la Parole* (p. 633–641).

Table des exemples

Exemple III-1 : Tuailon, stimulus n°6 (seule la portion de texte en gras correspond au stimulus).....	40
Exemple III-2 : Molossi, stimulus n°7 : transcriptions fines pour segmentation et simplifiées pour relecture.....	42
Exemple III-3 : Tuailon, stimulus n°1 : transcription simplifiée et augmentée (seule la partie en caractère gras correspond au stimulus extrait).....	42
Exemple III-4 : Le Chequer, stimulus n°1 : transcription orthographique simplifiée.	48
Exemple III-5 : Tuailon, stimulus n°1 composé d'une unité syntaxique majeure et 5 unités syntaxiques intermédiaires	49
Exemple III-6 : Annotation selon le manuel de Tanguy et collègues (2012) d'une unité rectionnelle verbale contenant une subordonnée à fonction objet. Exemple issu du manuel (p. 4).....	50
Exemple IV-1 : Le Chequer, stimulus n°13, transcription simplifiée	65
Exemple IV-2 : Tuailon, stimulus n°11, transcription simplifiée.....	65
Exemple V-1 : Spectrogramme (échelle de 0 à 5000 Hz), courbe mélodique (échelle 75 à 400 Hz) et tires d'annotation en phones, syllabes, mots, et frontière prosodique majeure, et contour mélodique du mot « destinataire» (extrait du stimulus n°13 Tuailon_stimulus1_orig) dont la syllabe finale donne lieu à une annotation de frontière prosodique majeure de type « PF0Dur » (cumul des trois paramètres annotés). Transcription simplifiée de l'extrait entier : « et puis vous avez été aussi forcément, comme nous tous, destinataires de mails et puis d'informations, de blogs (...) ».....	90
Exemple V-2 : Spectrogramme, courbe mélodique (échelle 75 à 400 Hz) et tires d'annotation en phones, syllabes, mots, et durées syllabiques du mot « 2011» terminé par un schwa épenthétique (extrait du stimulus n°5 de Tuailon) dont la syllabe finale [zə] analysée ne donnerait lieu à aucune annotation de frontière prosodique majeure. Cependant, si l'on considère la « macro-syllabe » [lɔ̃zə] incluant le schwa épenthétique comme dans notre étude, celle-ci donne par contre lieu à une annotation de frontière prosodique marquée par la durée « Dur ». Transcription simplifiée de l'extrait entier : « moi je trouve quand même qu'il y a quelques différences entre le rapport interne que vous aviez présenté en 2011, et ce rapport (...) ».....	91
Exemple VII-1 : Exemple de frontière prosodique majeure marquée par une pause « frontière » de 405 millisecondes située entre deux unités syntaxiques majeures, extrait du stimulus	

- n°11 de Tuaille en condition originale. Transcription simplifiée de l'extrait entier : « Vous avez fait le choix de ne pas jouer l'hégémonie, et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble afin de ne pas effrayer nos partenaires. » 140
- Exemple VII-2 : Exemple de frontière prosodique majeure marquée par une pause « frontière » située entre deux unités syntaxiques intermédiaires , extrait du stimulus n°1 de Voynet en condition de relecture. Transcription simplifiée de l'extrait entier : « mais en même temps, je voudrais vous dire qu'il n'est pas de tradition qu'on étale la vie personnelle des agents dans cette enceinte » 140
- Exemple VII-3 : Exemple de frontière prosodique majeure intra-syntaxique marquée par une pause « frontière » de 389 millisecondes située entre « cette » et « intervention » unités syntaxiques majeures, extrait du stimulus n°24 de Voynet en condition originale. Transcription simplifiée de l'extrait entier : « Merci pour cette intervention équilibrée, et de bonne foi, et qui ne sue pas l'ambiance électorale. » 141
- Exemple VII-4 : Exemple de frontière prosodique majeure marquée par une frontière prosodique de type « F0 » sur la syllabe [su] du préfixe dans le mot « sous-représentation ». Extrait du stimulus n°11 de Tuaille en condition originale. Transcription simplifiée de l'extrait entier : « Vous avez fait le choix de ne pas jouer l'hégémonie, et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble afin de ne pas effrayer nos partenaires. » 141

Table des tableaux

Tableau II—1 : Tableau reprenant la typologie d'actes discriminants de Poggi et collègues (2011) selon la dimension de l'ethos visée par l'attaquant chez son opposant et sous-entendus équivalents pour les dimensions de bienveillance, compétence et dominance.	17
Tableau III-1 : Nombre de stimuli tronqués s'achevant à une frontière de phrase, d'unité syntaxique majeure, ou d'unité syntaxique intermédiaire.	39
Tableau III-2 : détail du nombre d'énoncés relus appariables au corpus original, de lectures erronées et de transcriptions erronées sur les 125 extraits enregistrés durant la tâche de lecture.	45
Tableau III-3 : Nombre de marques de pauses remplies (« euh »), de répétitions, et de révisions par locuteurs dans le corpus original.	47
Tableau III-4 : Nombre d'unités syntaxiques majeures (<i>SyntMaj</i>) et intermédiaires (<i>SyntInt</i>) annotées	51
Tableau IV—1 valeurs principales des scores d'hostilité perçue à l'audio et à l'écrit et des valeurs de dHost, différence entre scores audio et écrit pour chaque stimulus (moyenne et écart-type, minimum, maximum). Correspondance des scores d'hostilité perçue : 0 = aucune hostilité, 1 = faible hostilité, 2 = forte hostilité.	59
Tableau IV—2 corrélations de Spearman entre scores d'hostilité perçue dans le test audio et le test écrit, dHost (audio-écrit) et scores audio, dHost et scores écrits. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	61
Tableau IV—3 valeurs principales des scores d'estimation d'à qui s'adressent les locuteurs en condition audio écrit et des valeurs de dDest, différence entre scores audio et écrit pour chaque stimulus (moyenne et écart-type, minimum, maximum). Correspondance des scores : 1=certainement à l'auditoire dans sa globalité, 3= indéterminé, 5=certainement à interlocuteur particulier.	63
Tableau IV—4 corrélations de Spearman entre scores d'estimation du destinataire en condition audio et écrite, dDest et scores audio, dDest et scores écrits. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	66
Tableau IV—5 valeurs principales des scores de valence auto-évaluée (AUTO) par les locuteurs du corpus et perçue (PERC) par 24 auditeurs naïfs (moyenne et écart-type, minimum, maximum). Pour rappel : -2 = très mécontent-e, 0 = neutre, 2 = très content-e.	70

Tableau IV—6 corrélations de Spearman entre scores de valence auto-évaluée et perçue, dVal et scores auto-évalués, dVal et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	73
Tableau IV—7 valeurs principales des scores d'excitation auto-évaluée par les locuteurs du corpus (AUTO) et perçue (PERC) par 24 auditeurs naïfs (moyenne et écart-type, minimum, maximum). Pour rappel : -2 = très calme, 0 = neutre, 2 = très agité-e.	74
Tableau IV—8 corrélations de Spearman entre scores d'excitation auto-évaluée et perçue, dExc (mesures de différence entre auto-évaluation et perception) et scores auto-évalués, dExc et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	76
Tableau IV—9 valeurs principales des scores de contrôle de soi auto-évalué par les locuteurs du corpus et dominance perçue par 24 auditeurs naïfs : moyennes et écart-types, et scores minimum et maximum. Pour rappel : -2 = contrôle/dominance min, 0 = neutre, 2 = contrôle/dominance max.	77
Tableau IV—10 corrélations de Spearman entre scores de contrôle de soi auto-évalué et de dominance perçue, dDom (mesures de différence entre auto-évaluation et perception) et scores auto-évalués, dDom et scores perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $n.s = p > .05$, non significatif.	80
Tableau IV—11 corrélations de Spearman entre scores de valence, excitation et dominance perçue. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $n.s = p > .05$, non significatif.	81
Tableau IV—12 corrélations de Spearman entre scores de valence, excitation et dominance perçue. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $n.s = p > .05$, non significatif.	83
Tableau IV—13 corrélations de Spearman entre hostilité perçue, valence perçue, valence auto-évaluée et dVal, mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	84

Tableau IV—14 corrélations de Spearman entre hostilité perçue, excitation perçue, excitation auto-évaluée et dExc mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent la valeur de p des corrélations calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	85
Tableau IV—15 corrélations de Spearman entre hostilité perçue, valence perçue, valence auto-évaluée et dVal, mesure de différence entre scores auto-évalués et perçus. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	86
Tableau VI—1 Moyenne de la fréquence fondamentale (F0) moyenne en Hertz dans le corpus original et la relecture, et moyenne des mesures de différence en demi-tons. L'écart-type est présenté entre parenthèses.	123
Tableau VIII—1 : Corrélations entre dominance perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par une pause, marquées par F0, et par la durée. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif. Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. .	158
Tableau VIII—2 Corrélations de Spearman entre hostilité perçue et nombre de frontières situées entre unités syntaxiques majeures (UnitSyntMaj), intermédiaires (UnitSyntInt) et en position intra unités syntaxiques (IntraUnitSynt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus.	164
Tableau VIII—3 : Corrélations entre dominance perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par une pause, marquées par F0, et par la durée. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif. Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. .	166
Tableau VIII—4 Corrélations de Spearman entre dominance perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par	

stimulus. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif..... 173

Tableau VIII—5 : Corrélations entre valence perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par F0, et par la durée. Orig = décompte par stimulus dans le corpus original, diff. = différence stimulus par stimulus entre décompte dans le corpus original et dans la relecture. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif. 174

Tableau VIII—6 Corrélations de Spearman entre valence perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif..... 175

Tableau VIII—7 : Corrélations entre excitation perçue et nombre moyen par stimulus de frontières prosodiques tous paramètres confondus, marquées par F0, et par la durée. Orig = décompte par stimulus dans le corpus original, diff. = différence stimulus par stimulus entre décompte dans le corpus original et dans la relecture. Les valeurs affichées correspondent aux coefficients de corrélation de Spearman. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif. 179

Tableau VIII—8 Corrélations de Spearman entre excitation perçue et nombre de frontières prosodiques (tous paramètres confondus) situées entre unités syntaxiques majeures (nFront sMaj), intermédiaires (nFront sInt) et à l'intérieur même d'unités syntaxiques intermédiaires (nFront intra-synt). Orig = décompte moyen par stimulus dans le corpus original, diff. = différence entre décompte dans le corpus original et dans la relecture par stimulus. Les étoiles représentent la valeur de p : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, NS = non significatif..... 184

Table des figures

Figure II-1 : schématisation personnelle des observations de Ferré (2004) concernant les caractéristiques prosodiques de part et d'autre des pauses silencieuses (PS= pause silencieuse) selon que celles-ci aient principalement une fonction d'hésitation, de structuration, de maintien du tour de parole (pragmatique) ou de focalisation.....	29
Figure III-1 : Capture d'écran du début d'un fichier texte d'annotation du début d'une séance du conseil municipal. Les chiffres indiquent l'heure à laquelle se déroulent les événements annotés.....	36
Figure III-2 : Capture d'écran d'un extrait de fichier d'annotation sur le logiciel ELAN et la segmentation en tours de parole par locuteurs. La tire du bas (titre vert clair) correspond aux notes textuelles brutes exportées puis corrigées manuellement.....	37
Figure III-3 : Photographies des locuteurs étudiés durant le conseil municipal de Montreuil (93100) en 2013 : Dominique Voynet (EELV) et quatre membres de l'opposition : Frédéric Molossi (PS), Gaylord Le Chequer (RGC/PG), Alexandre Tuillon (RSM) et Jean-Jacques Serey (PCF).	38
Figure III-4 : Liste des étapes des entretiens individuels effectués avec les locuteurs.....	43
Figure IV-1 : Exemple de présentation en diagrammes en violons et boîtes à moustaches. ...	56
Figure IV-2 : consignes du test de perception n°1 (similaires pour la version audio et la version écrite).....	58
Figure IV-3 : Diagrammes en violons et boîtes à moustaches des scores d'hostilité perçue dans le test audio et dans le test écrit sur 200 stimuli (40 x 5 locuteurs). Les étoiles représentent la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif. Correspondance des scores d'hostilité perçue : 0 = aucune hostilité, 1 = faible hostilité, 2 = forte hostilité.....	60
Figure IV-4 : Corrélations entre scores d'hostilité perçue dans les test audio et écrit. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.....	61
Figure IV-5 : Corrélations entre hostilité perçue dans le test écrit et dHost, mesure de différence entre scores audio et scores écrits. Les étoiles représentent la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.....	62

Figure IV-6 : Corrélations entre hostilité perçue (audio) et dHost pour Tuailleon. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	62
Figure IV-7 : Diagrammes en violons et boîtes à moustaches des scores d'estimation du destinataire dans le test audio et dans le test écrit. Les étoiles représentent la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	64
Figure IV-8 : Corrélations entre hostilité perçue et estimation du destinataire au sein du test audio. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	67
Figure IV-9 : Corrélations entre hostilité perçue et estimation du destinataire au sein du test écrit. Les étoiles représentent la valeur de p de la corrélation de Spearman calculée : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	67
Figure IV-10 : Grille dimensionnelle telle qu'elle a été présentée aux cinq locuteurs du corpus durant la tâche d'auto-évaluation de leurs stimuli.....	68
Figure IV-11 : Comparaisons entre scores de valence auto-évaluée et de valence perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	71
Figure IV-12 : Comparaisons entre scores d'excitation auto-évaluée et d'excitation perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	75
Figure IV-13 : Comparaisons entre scores de contrôle de soi auto-évalué et de dominance perçue (24 auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des test t appariés effectués sur les stimuli de chaque locuteur : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	78
Figure IV-14 : Corrélations entre contrôle de soi auto-évalué et dominance perçue (auditeurs naïfs) pour Le Chequer, Serey et Tuailleon. Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.....	79
Figure IV-15 : Corrélations entre excitation perçue et dominance perçue (auditeurs naïfs). Les étoiles représentent le degré de significativité relatif à la valeur de p des corrélations de	

Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.

.....	82
Figure IV-16 : Corrélations entre hostilité perçue (audio) et dDom, mesure de différence entre scores de contrôle de soi auto-évalué et scores de dominance perçue pour Tuillon. Les étoiles représentent la valeur de p des corrélations de Spearman calculées : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.	87
Figure V-1 : Décompte du nombre de frontières prosodiques majeures annotées par locuteur et par condition avec spécification de la présence ou non d'une pause « frontières » (durée > 140 ms) parmi les critères ayant justifié l'annotation.....	93
Figure V-2 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques majeures par stimulus (tous paramètres confondus) par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = NS$, non significatif.....	94
Figure V-3 : Proportions de frontières prosodiques majeures marquées par une pause « frontière » (durée > 140 ms) pour chaque locuteur dans le corpus original et la relecture.....	95
Figure V-4 : Diagrammes en violon et boites à moustaches du nombre de frontières prosodiques majeures marquées entre autres par l'insertion d'une pause « frontières » (durée > 140 ms) par stimulus dans le corpus original et la relecture. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : $*** = p < .001$, $** = p < .01$, $* = p < .05$, ns = $p > .05$	95
Figure V-5 : Diagrammes en violon et boites à moustaches du nombre de frontières prosodiques majeures marquées uniquement par l'insertion d'une pause « frontières » (durée > 140 ms) par stimulus dans le corpus original et la relecture. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : $*** = p < .001$, $** = p < .01$, $* = p < .05$, ns = $p > .05$	96
Figure V-6 : Proportions de frontières prosodiques majeures marquées, entre autres paramètres, par F0 pour chaque locuteur dans le corpus original et la relecture.	97
Figure V-7 : Nombre de frontières prosodiques marquées entre autres par F0 par stimulus selon la condition du corpus. Les étoiles représentent le niveau de significativité de la valeur de p obtenue en effectuant un test t apparié pour chaque locuteur entre version originale du stimulus et relecture : $*** = p < .001$, $** = p < .01$, $* = p < .05$, ns = $p > .05$	97

Figure V-8 : Décompte des différentes configurations multiparamétriques pour les frontières prosodiques majeures.	100
Figure V-9 : Nombre total de contours en cloche (C), de type focus (F), plats (P), descendants (D), montants (M) et terminaux (T) par locuteur et par condition.	104
Figure V-10 : Diagrammes en violons et boîtes à moustaches représentant le nombre de contours mélodiques annotés comme montants par stimulus (tous paramètres confondus) par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.	105
Figure V-11 : Diagrammes en violons et boites à moustaches du nombre de contours mélodiques de type focus par stimulus pour chaque locuteur selon la condition. Les étoiles au-dessus des crochets correspondent aux valeurs de p des tests t effectués pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus ($* = p < .05$, $** = p < .01$, $*** = p < .001$, NS. = non significatif).	106
Figure V-12 : Diagrammes en violons et boites à moustaches du nombre de contours mélodiques annotés comme plats par stimulus pour chaque locuteur selon la condition. Les étoiles au-dessus des crochets correspondent aux valeurs de p des tests t effectués pour chaque locuteur entre valeurs des stimuli originaux et valeurs des stimuli relus ($* = p < .05$, $** = p < .01$, $*** = p < .001$, NS. = non significatif).	106
Figure V-13 : Diagrammes en violon et boites a moustache représentant le nombre de contours annotés comme descendants par locuteur et par condition. Les étoiles représentent la valeur de p des tests t appariés effectués sur les stimuli de chaque locuteur entre condition originale et relecture : $p < .001 = ***$, $p < .01 = **$, $p < .05 = *$, $p > .05 = \text{NS}$, non significatif.	107
Figure VI-1 : Schéma explicatif de la mesure de différence inter-blocs sur une paire d'extraits fictifs constitués chacun de deux blocs inter-pausaux pour lesquels sont spécifiés les valeurs par bloc de débit articulatoire (n syllabes / secondes).	128
Figure VI-3 : nombre total de pauses silencieuses par locuteur selon la condition du corpus et leurs durées.	131
Figure VI-4 : Distribution des durées des pauses silencieuses sur l'ensemble du corpus par locuteur et par condition.	132
Figure VI-5 : Diagrammes en violons et boites à moustaches représentant la proportion de silence (durées cumulées des pauses /durée totale du stimulus) pour le corpus original et la relecture pour chaque locuteur.	133

Figure VI-6 : Diagrammes en violons et boîtes à moustaches représentant la durée relative en pourcentage des pauses silencieuses « frontières » par rapport à la durée totale des stimuli selon la condition du corpus.	134
Figure VI-7 : Diagrammes en violons et boîtes à moustaches représentant le débit articulatoire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur.	135
Figure VI-8 : Diagrammes en violons et boîtes à moustaches représentant l'écart-type du débit articulatoire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur.	136
Figure VI-9 : Diagrammes en violons et boîtes à moustaches de l'écart-type du débit articulatoire (syllabes / secondes) moyen par stimulus selon la condition et le locuteur.	137
Figure VI-10 : Différences moyenne de débit articulatoire (n syll / secondes) d'un bloc inter-pausal à l'autre pour chaque locuteur selon la condition.	138
Figure VI-11 : Diagrammes en violon et boîtes à moustache représentant en millisecondes les valeurs de durée des voyelles situées dans les syllabes finales pré-pausales en condition originale et en relecture pour chaque locuteur.	139
Figure VI-12 : Valeurs moyennes par stimulus de l'allongement syllabique relatif final d'unité prosodique majeure par locuteur selon la condition (all. rel. fin. : durée syllabe / moyenne syllabes avoisinantes).	140
Figure VI-13 : Diagrammes en violon et boîtes à moustaches représentation la F0 moyenne (Hertz) par stimulus selon le locuteur et la condition.	141
Figure VI-14 : Diagrammes en violon et boîtes a moustache des mesures de différence original-relecture de F0 moyenne par stimulus présentées en demi-tons pour chaque locuteur..	142
Figure VI-15 : Diagrammes en violon et boîtes à moustaches représentation l'écart-type de la F0 moyenne en demi-tons par stimulus selon le locuteur et la condition.	144
Figure VI-16 : Diagrammes en violon et boîtes à moustaches représentant les différences absolues de F0 moyenne d'un bloc inter-pausal à l'autre (demi-tons) par stimulus selon le locuteur et la condition.	145
Figure VI-17 : Diagrammes en violon et boîtes à moustaches représentant les différences brutes de F0 moyenne d'un bloc inter-pausal à l'autre (demi-tons) par stimulus selon le locuteur et la condition.	146
Figure VI-18 : Diagrammes en violons et boîtes à moustaches représentant la plage de variation moyenne par locuteur selon la condition.	147

Figure VI-19: Diagrammes en violon et boîtes à moustaches représentant l'écart-type de la plage de variation moyenne de F0 en demi-tons par stimulus selon le locuteur et la condition.	148
Figure VI-20 : Diagrammes en violon et boîtes à moustaches représentant les différences absolues d'étendue de la F0 de d'un bloc inter-pausal à l'autre moyenne par locuteur selon la condition.	149
Figure VI-21 : Diagrammes en violon et boîtes à moustaches représentant les différences brutes d'étendue de la F0 plage de variation d'un bloc inter-pausal à l'autre moyenne par locuteur selon la condition.	150
Figure VI-22 : Corrélations entre débit articuloire moyen et index de variabilité paire à paire de la durée des syllabes (IVP)	152
Figure VI-23 : Corrélations entre débit articuloire et nombre de frontières prosodiques marquées par la durée (allongement syllabique final relatif > 3).	153
Figure VI-24 : Corrélations entre débit articuloire moyen (n syll /sec) et durée moyenne des voyelles des syllabes pré-pausales (fin de blocs inter-pausaux).....	153
Figure VI-25 : Corrélations entre débit articuloire moyen et variations de bloc à bloc.	154
Figure VI-26 : Inter-corrélations entre F0 moyenne (Hertz) et plage de variation de F0 (demi-tons) pour Voynet et Tuillon.....	155
Figure VI-27 : Corrélations entre plage de F0 et nombre de frontières prosodiques marquées par la durée pour Serey et Voynet.	157
Figure VI-28 : Inter-corrélations entre débit articuloire (n syllabes / secondes sur temps de parole) et F0 moyenne (Hertz) par stimulus pour Le Chequer.	158
Figure VII-1 : Proportion d'unités syntaxiques majeures (SyntMaj) et intermédiaires (SyntInt) marquées par une frontière prosodique par rapport au total d'unités syntaxiques majeures et intermédiaires annotées.	165
Figure VII-2 : Répartition des frontières prosodiques selon leur localisation dans la structure des énoncés.	167
Figure VII-3 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques situées en frontières d'unités syntaxiques majeures selon la condition du corpus.	169
Figure VII-4 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques situées en frontières d'unités syntaxiques majeures selon la condition du corpus.	170

Figure VII-5 : Diagrammes en violons et boîtes à moustaches représentant le nombre de frontières prosodiques intra unités syntaxiques selon la condition du corpus.....	171
Figure VII-6 : Diagrammes en violons et boîtes à moustaches représentant le nombre de proéminences prosodiques intra mots syntaxiques selon la condition du corpus.....	172
Figure VII-7 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » situées en frontières d'unités syntaxiques majeures selon la condition du corpus.	173
Figure VII-8 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » situées en frontières d'unités syntaxiques intermédiaires selon la condition du corpus.	173
Figure VII-9 : Diagrammes en violons et boîtes à moustaches représentant le nombre de pauses « frontières » intra-unités syntaxiques intermédiaires selon la condition du corpus.....	174
Figure VII-10 : Proportions de pauses « frontières » selon leur position dans la structure syntaxique.....	175
Figure VII-11 : Durée des pauses silencieuses (ms) selon leur position dans la structure syntaxique des énoncés.....	177
Figure VII-12 : Allongement final pré-pausal relatif (syllabique) selon la position des frontières prosodiques position dans la structure syntaxique des énoncés.	178
Figure VIII-1 Corrélations de Spearman entre hostilité perçue et F0 moyenne sur le corpus original (Hertz) pour Voynet et Molossi.	187
Figure VIII-2 : Corrélations entre hostilité perçue et différences original-relecture de la variation de plage de F0 d'un bloc inter-pausal à l'autre par stimulus pour Serey.	188
Figure VIII-3 : Corrélations entre hostilité perçue et débit articuloire pour Voynet et Tuailon.	189
Figure VIII-4 : Corrélations entre hostilité perçue et écart-type du débit articuloire pour Le Chequer.....	189
Figure VIII-5 : Corrélations entre hostilité perçue et variation de débit articuloire d'un bloc inter-pausal à l'autre.	190
Figure VIII-6 Corrélations entre hostilité perçue et index de variabilité de durée syllabique (IVP) de durée syllabique pour Tuailon.	190
Figure VIII-7 : Corrélations entre hostilité perçue et nombre de frontières prosodiques situées entre unités syntaxiques majeures par stimulus pour Voynet et Serey.....	192
Figure VIII-8 : Corrélations entre dominance perçue et différence original-relecture du nombre de pauses « frontières » par stimulus pour Le Chequer et Molossi.	194

Figure VIII-9 : Corrélations entre dominance perçue et nombre de frontières prosodiques marquées par F0 pour Le Chequer.	195
Figure VIII-10 : Corrélations entre dominance perçue et différence original-relecture du nombre de frontières prosodiques marquées par la durée par stimulus pour Le Chequer.....	195
Figure VIII-11 : Corrélations entre dominance perçue et fréquence fondamentale moyenne F0 en Hertz ; échelle adaptée aux locuteurs pour plus de lisibilité.	196
Figure VIII-12 : Corrélations entre dominance perçue et nombre moyen de contours mélodiques annotés comme en cloche pour Serey.	197
Figure VIII-13 : Corrélations entre dominance perçue et nombre moyen de contours mélodiques annotés comme descendants pour Voynet.	198
Figure VIII-14 : corrélations entre dominance perçue et durée des voyelles finales pré-pausales pour Voynet et Tuailleon.	199
Figure VIII-15 : Corrélations entre valence perçue et nombre de frontières prosodiques majeures (tous paramètres confondus) pour Voynet et Molossi.	201
Figure VIII-16 : Corrélations de Spearman entre valence perçue et F0 moyenne sur le corpus original (Hertz) pour Voynet, Le Chequer et Molossi.....	203
Figure VIII-17 : Corrélations entre valence perçue et contours manuellement annotés comme en cloche pour Serey.	204
Figure VIII-18 : Corrélations entre valence perçue et débit articulatoire moyen (n syllabes / secondes sur temps de parole par stimulus) pour Tuailleon.	205
Figure VIII-19 : Corrélations entre valence perçue et durée moyenne des voyelles finales pré-pausales pour Tuailleon.	205
Figure VIII-20 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen par stimulus de frontières prosodiques majeures pour Le Chequer.....	208
Figure VIII-21 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen par stimulus de frontières prosodiques majeures marquées par F0 pour Voynet.	208
Figure VIII-22 : Corrélations entre excitation perçue et F0 moyenne (Hz).	209
Figure VIII-23 : Corrélations entre excitation perçue et nombre moyen de contours annotés comme descendants dans le corpus original pour Voynet.	210
Figure VIII-24 : Corrélations entre excitation perçue et différence original-relecture du nombre moyen de contours mélodiques montants par stimulus pour Le Chequer.	210

Figure VIII-25 : Corrélations entre excitation perçue et durée moyenne des voyelles des syllabes finales pré-pausales pour Le Chequer et Tuillon.211

Annexes

Le tableau présenté dans les pages suivantes fournit des informations sur les stimuli du corpus et permet d'accéder par liens cliquables aux fichiers audio. Nous explicitons ici l'ensemble des informations présentes dans le tableau :

1. **Nom du locuteur et numéro de stimulus**
2. **Condition originale** : lien vers l'extrait audio original
3. **Condition relecture** : lien vers l'extrait audio relu
4. **Transcription orthographique simplifiée** : Transcription telle qu'elle a été présentée aux locuteurs durant la tâche de relecture. Seule la partie en caractères gras correspond au stimulus¹.
5. **Host** : score d'hostilité perçue obtenu dans le test audio
0= aucune hostilité perçue, 1 = légère hostilité perçue, 2 = forte hostilité perçue
6. **HostE** : score d'hostilité perçue obtenu dans le test écrit.
7. **dHost** : score d'hostilité perçue audio – score écrit
8. **Dest** : estimation du destinataire du message dans le test audio
1=le locuteur s'adresse certainement à l'auditoire dans sa globalité, 3= indéterminé, 5=certainement à interlocuteur particulier
9. **DestE** : estimation du destinataire du message dans le test écrit.
10. **dCtx** : score d'hostilité perçue audio – score écrit
11. **ValN** : score de valence perçue (-2= très mécontent-e, 0= neutre, 2= très content-e)
12. **ValA** : score de valence auto-évaluée
13. **dVal** : score de valence perçue– score auto-évalué
14. **ExcN** : score d'excitation perçue
-2 = très calme, 0 = neutre, 2 = très agité-e
15. **ExcA** : score d'excitation auto-évaluée
16. **dExc** : score d'excitation perçue– score auto-évalué
17. **Dom** : score de dominance perçue
-2=très soumis-e, 0=neutre, 2 = très dominant-e
18. **Ctrl** : score de contrôle de soi auto-évalué
-2=très faible contrôle de soi, 0=neutre, 2=très fort contrôle de soi
19. **dDom** : score de dominance perçue perçue– score de contrôle auto-évalué

¹ Pour rappel, le reste de la transcription a seulement été ajouté pour faciliter la tâche de relecture pour les locuteurs.

nom du stimulus	original	relecture	transcription simplifiée étendue	host	hostE	dHost	Dest	DestE	dDest	Dom	Ctrl	dDom	ValN	ValA	dVal	ExcN	ExcA	dExc
Voynet_stimulus1	http://ppp.ils	http://ppp.ils	Merci Monsieur Mamadou, je vais donner pour les questions de personnels la parole à Hélène Zeidenberg, mais en même temps, je voudrais vous dire qu'il n'est pas de tradition qu'on étale la vie personnelle des agents dans cette enceinte , ça se fait pas vraiment, simplement je souhaite dire ici (...)	0.89	1.38	-0.50	2.33	3.92	-1.59	0.72	0.00	-0.72	-0.68	1.00	1.68	-0.67	0.00	0.67
Voynet_stimulus2	http://ppp.ils	http://ppp.ils	Bah oui mais on fait comme d'habitude, je demande si quelqu'un veut intervenir, si quelqu'un veut intervenir on la présente, et sinon on passe vite.	1.11	0.85	0.26	3.56	2.85	0.71	0.71	1.00	0.29	-0.92	1.00	1.92	1.18	1.00	-0.18
Voynet_stimulus3	http://ppp.ils	http://ppp.ils	J'en ai noté quelques uns, je veux pas les répéter maintenant, mais ils sont formulés d'une façon insinuante et insultante, mettant en cause ma compétence, mon honnêteté, d'une façon qui fait passer les deux mots que j'ai prononcés comme une manifestation bien légitime , à cette heure, de fatigue et d'exaspération.	1.56	1.77	-0.21	2.67	3.38	-0.72	0.15	2.00	1.85	-1.33	0.00	1.33	-0.35	1.00	1.35
Voynet_stimulus4	http://ppp.ils	http://ppp.ils	Ah oui bah 3 milliards répartis entre les collectivités, ça risque de faire d'après nos calculs, sur deux exercices, pratiquement 10% de dotations globales de fonctionnement en moins , voilà, il faudra faire avec.	0.44	0.38	0.06	2.56	2.08	0.48	-0.20	1.00	1.20	-0.32	1.00	1.32	-1.38	-1.00	0.38
Voynet_stimulus5	http://ppp.ils	http://ppp.ils	Sur ce carrefour là, ça va en tout cas améliorer singulièrement la sécurité, cette école était une de celles qui étaient les plus préoccupantes, parce que le carrefour est complexe, et très difficile à lire pour des gamins.	0.11	0.54	-0.43	2.33	2.31	0.03	0.52	0.00	-0.52	-0.32	1.00	1.32	-0.35	1.00	1.35
Voynet_stimulus6	http://ppp.ils	http://ppp.ils	On a bien fait, puisqu'effectivement maintenant on est plutôt satisfaits de ne pas avoir perdu ce temps qui nous permettra effectivement de relancer la réalisation des cloisonnements intérieurs de l'école, qui est en panne depuis plusieurs semaines aujourd'hui.	0.33	0.23	0.10	1.89	1.62	0.27	0.21	1.00	0.79	0.21	2.00	1.79	-0.71	-1.00	-0.29
Voynet_stimulus7	http://ppp.ils	http://ppp.ils	Et donc l'idée c'était de mutualiser un certain nombre d'informations sur les moyens dont dispose la ville en matière de tranquillité, sur ses compétences , par rapport à la loi, et comment elle entend y faire face, (...)	1.67	0.08	1.59	2.00	1.50	0.50	1.13	1.00	-0.13	-1.42	1.00	2.42	1.38	-1.00	-2.38
Voynet_stimulus8	http://ppp.ils	http://ppp.ils	Je l'ai dit, je condamne sans détour les conditions d'intervention et le comportement du directeur de la tranquillité publique , lors de la soirée de la Halle Dufriche le 19 mai dernier, qui ont conduit, et c'était la première fois, à un dépôt de plainte contre lui.	1.22	1.31	-0.09	1.56	2.00	-0.44	1.63	2.00	0.37	-1.32	1.00	2.32	-0.88	1.00	1.88
Voynet_stimulus9	http://ppp.ils	http://ppp.ils	Merci Madame, je vous propose éventuellement de prolonger le vote de ce soir par un travail de fond, plus approfondi.	0.33	0.08	0.26	2.67	1.08	1.59	0.77	1.00	0.23	0.11	2.00	1.89	-1.02	-1.00	0.02
Voynet_stimulus10	http://ppp.ils	http://ppp.ils	Mais il faut que les gens ici connaissent la vérité : avant que le PLU ne soit adopté, le POS en vigueur permettait de construire sur cet îlot, avec une hauteur maximale de 18 mètres sur front de rue, de 15 mètres en fond de parcelle.	1.44	0.00	1.44	2.11	1.38	0.73	1.18	1.00	-0.18	-1.38	1.00	2.38	1.12	2.00	0.88
Voynet_stimulus11	http://ppp.ils	http://ppp.ils	Merci Monsieur Tuailon. Le jugement, au tribunal, ne sanctionne pas un passage en force, il sanctionne le fait qu'on s'est trompés sur la qualification des caravanes , qui parce qu'elles n'ont pas de roues doivent être qualifiées de constructions légères (...)	0.56	0.54	0.02	2.56	2.23	0.32	0.16	1.00	0.84	-0.48	1.00	1.48	-0.40	0.00	0.40
Voynet_stimulus12	http://ppp.ils	http://ppp.ils	Bien sûr Madame, y'a pas de problème d'avoir des débats mais un débat c'est préparé, c'est pas un cavalier sur n'importe quelle question qui arrive à l'ordre du jour et sur lesquelles on demande à engager une discussion.	1.33	1.54	-0.21	3.22	3.62	-0.39	0.51	1.00	0.49	-1.13	1.00	2.13	0.83	1.00	0.17
Voynet_stimulus13	http://ppp.ils	http://ppp.ils	Une juste représentation des minorités, une juste représentation de chacune des villes, ça n'est pas de la tambouille politicienne, ça n'a rien à voir avec les gesticulations constatées ici, ça s'appelle juste la démocratie , et le respect de la loi.	1.44	1.92	-0.48	3.22	4.00	-0.78	1.08	1.00	-0.08	-0.87	2.00	2.87	0.06	1.00	0.94
Voynet_stimulus14	http://ppp.ils	http://ppp.ils	(...) et puis évidemment ça suppose une discussion ensuite, une discussion qui ne se termine pas forcément ce soir, d'accord?	1.56	0.62	0.94	3.56	2.31	1.25	0.72	1.00	0.28	-1.42	1.00	2.42	1.38	1.00	-0.38
Voynet_stimulus15	http://ppp.ils	http://ppp.ils	Ça n'est pas juste un bâtiment dans lequel on aurait, pour pouvoir faire des spectacles et des animations en plein air, rasé des préfas insalubres , c'est un projet culturel, c'est un projet de quartier, dans lequel on accueille aussi bien des gens du Bas Montreuil que des gens des quartiers les plus proches.	0.89	1.38	-0.50	2.44	2.69	-0.25	0.28	0.00	-0.28	-0.83	1.00	1.83	0.42	0.00	-0.42
Voynet_stimulus16	http://ppp.ils	http://ppp.ils	Et puis il y avait aussi un super projet immobilier, 300 pavillons dans les murs à pêches, vous vous souvenez de ce projet ? Génial. Et c'est ces gens-là qui nous donnent des leçons, qui nous expliquent leur stratégie et leur ambition pour protéger les murs à pêches ?	1.89	2.00	-0.11	2.78	2.31	0.47	0.97	1.00	0.03	-1.68	1.00	2.68	1.38	1.00	-0.38
Voynet_stimulus17	http://ppp.ils	http://ppp.ils	C'est un programme qui se mène sur le rythme pluriannuel, qui va donc se poursuivre, avec le souci à la fois d'améliorer les conditions, je dirais sur le plan de la sécurité du confort et de la salubrité, mais aussi sur le plan sportif.	0.22	0.00	0.22	1.11	1.38	-0.27	0.58	0.00	-0.58	-0.47	1.00	1.47	0.58	-1.00	-1.58
Voynet_stimulus18	http://ppp.ils	http://ppp.ils	Il nous a semblé que ce n'était pas réhibitoire, que tout dans la vie de Madame Françoise Héritier illustrait sa hauteur de vue, sa capacité effectivement à intervenir auprès des plus jeunes comme des plus âgés pour combattre toujours des représentations , et des constructions qui pèsent sur le sort des jeunes qui grandissent parmi nous aujourd'hui (...)	0.33	0.08	0.26	1.44	1.92	-0.48	0.62	1.00	0.38	-0.20	2.00	2.20	0.42	-1.00	-1.42
Voynet_stimulus19	http://ppp.ils	http://ppp.ils	vous n'aviez pas bien compris que les ressources de la collectivité ne dépendaient plus aujourd'hui d'un lien direct avec le nombre d'entreprises et les richesses produites sur le territoire, je vous expliquerai un jour comment ça marche le calcul de la cotisation économique des entreprises, c'est un peu plus compliqué que ce que vous affirmiez.	1.33	1.38	-0.05	4.89	4.62	0.27	0.16	1.00	0.84	-0.92	1.00	1.92	0.00	1.00	1.00
Voynet_stimulus20	http://ppp.ils	http://ppp.ils	Mais juste pour vous dire peut être deux choses : puisque vous ne souhaitez pas que certains autour de cette table, et je ne l'ai pas fait, parlent à votre place, juste vous dire que je ne l'aime pas non plus, qu'on le fasse à ma place.	1.44	1.77	-0.32	4.78	5.00	-0.22	0.62	1.00	0.38	-1.43	1.00	2.43	0.05	1.00	0.95
Voynet_stimulus21	http://ppp.ils	http://ppp.ils	Il faudrait être absolument débile pour rétorquer, pour rétorquer comme vous l'avez dit, que je n'y connaîtrais rien en matière d'urbanisme.	1.56	1.85	-0.29	4.78	4.62	0.16	-0.26	1.00	1.26	-1.42	1.00	2.42	0.21	1.00	0.79
Voynet_stimulus22	http://ppp.ils	http://ppp.ils	Merci beaucoup. Donc je rappelle que ces expos se tiennent toujours, elles se tiennent à la bibliothèque; et que les classes viennent visiter les expositions et que tout va bien.	1.22	0.54	0.68	2.89	1.69	1.20	-0.30	1.00	1.30	0.66	0.00	-0.66	-0.96	1.00	1.96
Voynet_stimulus23	http://ppp.ils	http://ppp.ils	On respire à fond, on arrête ses conversations téléphoniques, les tweets, les SMS, l'avantage c'est que ça fait pas de bruit , et on vote avant la pause, pour pouvoir prendre cette pause et revenir détendus, décontractés, pour la deuxième partie du conseil municipal.	1.56	1.08	0.48	2.78	1.46	1.32	1.42	1.00	-0.42	-1.28	1.00	2.28	1.22	1.00	-0.22
Voynet_stimulus24	http://ppp.ils	http://ppp.ils	Merci pour cette intervention équilibrée, et de bonne foi, et qui ne sue pas l'ambiance électorale. Pierre Desgranges a la parole (...)	1.67	0.62	1.05	4.56	4.54	0.02	0.75	1.00	0.25	0.51	1.00	0.49	-0.21	1.00	1.21

nom du stimulus	original	relecture	transcription simplifiée étendue	host	hostE	dHost	Dest	DestE	dDest	Dom	Ctrl	dDom	ValN	ValA	dVal	ExcN	ExcA	dExc
LeChequer_stimulus1	http://ppp.il	http://ppp.il	(...) Vous avez voulu faire de ce sujet un clivage, un clivage politique entre vous, et votre prédécesseur, les choses sont très claires. (...)	1.56	1.54	0.02	4.56	5.00	-0.44	1.05	0.00	-1.05	-1.07	-1.00	0.07	-0.65	-1.00	-0.35
LeChequer_stimulus2	http://ppp.il	http://ppp.il	Donc moi je trouve que ce débat est indécent, et que vous feriez mieux de vous mettre dans un couloir, de régler ça entre vous, mais de pas prendre les montreuillois à témoin de ça.	1.33	1.85	-0.51	3.78	3.92	-0.15	1.02	0.00	-1.02	-1.63	1.00	2.63	-0.22	-1.00	-0.78
LeChequer_stimulus3	http://ppp.il	http://ppp.il	Donc, aujourd'hui, la justice a rendu son plaidoyer, et le directeur en question a été réhabilité.	0.33	0.08	0.26	1.89	1.46	0.43	-0.82	-1.00	-0.18	0.05	0.00	-0.05	-0.55	1.00	1.55
LeChequer_stimulus4	http://ppp.il	http://ppp.il	Bon, nous, on pensait vous avoir convertis au vœu, donc on vous a pas reproché d'en avoir déposé un, par contre effectivement à la lecture du projet qui a été posé, (...)	1.44	1.00	0.44	4.67	4.46	0.21	-0.02	1.00	1.02	-0.30	1.00	1.30	-0.15	-1.00	-0.85
LeChequer_stimulus5	http://ppp.il	http://ppp.il	Si je me fie à la conférence des présidents qui s'est tenue, nous abordons ce soir un débat relatif à l'expérimentation d'un système de vidéo protection, puis à l'autorisation donnée à la Maire de répondre à l'appel à projet, et ensuite (...)	0.22	0.00	0.22	1.33	1.00	0.33	-0.21	-1.00	-0.79	0.16	0.00	-0.16	-0.92	-1.00	-0.08
LeChequer_stimulus6	http://ppp.il	http://ppp.il	85% des caméras d'Amiens, sont en centre ville, donc un peu la même configuration, ou sur les mêmes problématiques que nous nous prétendons aborder avec votre délibération.	0.67	0.08	0.59	2.00	1.69	0.31	0.87	1.00	0.13	-0.53	-1.00	-0.47	0.71	-1.00	-1.71
LeChequer_stimulus7	http://ppp.il	http://ppp.il	C'est sur rue, un bâtiment à R+4, sur une parcelle qui fait un peu plus de trois mètres de large, et avec dans le fond de parcelle, plusieurs bâtiments successifs, qui vont venir directement en mitoyenneté avec un immeuble d'habitation, et le jardin d'une voisine.	0.44	0.00	0.44	2.78	1.46	1.32	0.22	0.00	-0.22	-0.46	0.00	0.46	-0.46	0.00	0.46
LeChequer_stimulus8	http://ppp.il	http://ppp.il	Il ne s'agit pas de la situation d'une personne, dans un pavillon, qui verrait son cadre de vie dégradé.	0.67	0.69	-0.03	2.78	2.54	0.24	-0.51	-1.00	-0.49	-0.63	-1.00	-0.37	-0.25	0.00	0.25
LeChequer_stimulus9	http://ppp.il	http://ppp.il	(...), et donc vous demander ce qu'il en est, et à quel moment comptez-vous aborder ce débat sereinement avec nous au sein du conseil municipal.	1.00	1.08	-0.08	4.89	4.54	0.35	0.45	-1.00	-1.45	-0.47	-1.00	-0.53	-0.71	0.00	0.71
LeChequer_stimulus10	http://ppp.il	http://ppp.il	Donc là il s'agit simplement, on va dire d'habiller les choses, mais en réalité ce que vous êtes en train de faire c'est de procéder à cette liquidation, en réalité ce que vous demandez aux administrateurs, c'est de procéder progressivement à un transfert, (...)	1.11	1.62	-0.50	4.44	4.38	0.06	1.08	1.00	-0.08	-0.82	0.00	0.82	-0.46	-1.00	-0.54
LeChequer_stimulus11	http://ppp.il	http://ppp.il	Sur une question d'ordre du jour et de déroulement de notre séance. Pour savoir de quoi on va parler exactement, parce que si je m'en fie à l'ordre du jour que nous avons dans notre dossier de conseil municipal (...)	0.78	0.23	0.55	1.89	1.38	0.50	-0.51	0.00	0.51	-0.15	0.00	0.15	-1.02	0.00	1.02
LeChequer_stimulus12	http://ppp.il	http://ppp.il	C'est pour savoir : Est-ce que vous croyez vraiment qu'on a l'intention de discuter et de débattre à onze heures, d'un document que l'on découvre ce soir ?	1.67	1.62	0.05	4.78	4.15	0.62	1.32	1.00	-0.32	-1.27	1.00	2.27	0.09	0.00	-0.09
LeChequer_stimulus13	http://ppp.il	http://ppp.il	Nous vous avons proposé d'organiser une table ronde avec l'ensemble des élus, et du conseil municipal, les conseillers généraux, régionaux, le député, tout le monde, pour que ensemble, à l'échelle de Montreuil, on mette de côté la question de la campagne électorale et (...)	1.56	0.31	1.25	4.33	1.77	2.56	1.07	1.00	-0.07	-1.58	-1.00	0.58	1.33	1.00	-0.33
LeChequer_stimulus14	http://ppp.il	http://ppp.il	(...) les associations de parents d'élèves du quartier concerné, font un travail formidable avec une partie des gamins qui vivent sur ce site, pour contribuer à leur intégration, mais eux-mêmes le disent (...)	1.33	0.00	1.33	1.89	1.31	0.58	1.17	-1.00	-2.17	0.10	-1.00	-1.10	0.98	0.00	-0.98
LeChequer_stimulus15	http://ppp.il	http://ppp.il	La communauté d'agglomération Plaine Commune, avait rajouté, dans son CDT, un volet social. C'est un choix politique qui avait été fait par la communauté d'agglomération de Plaine Commune.	0.78	0.23	0.55	1.78	2.08	-0.30	0.62	0.00	-0.62	-0.40	0.00	0.40	-0.57	0.00	0.57
LeChequer_stimulus16	http://ppp.il	http://ppp.il	On revient d'une pause que vous avez bien voulu nous accorder avec votre gentillesse mémorable, et que nous apprécions tous, et on découvre un document relatif à la politique municipale en matière de tranquillité publique, et donc (...)	1.33	0.92	0.41	4.89	4.08	0.81	0.25	0.00	-0.25	-0.31	1.00	1.31	-0.41	-1.00	-0.59
LeChequer_stimulus17	http://ppp.il	http://ppp.il	Tout d'abord quelques remarques préalables sur le fameux document mystère qui nous est arrivé sur la table, en fait le document qui serait destiné, si on reprend le titre, à (...)	1.11	0.62	0.50	3.11	1.54	1.57	0.10	0.00	-0.10	-0.15	1.00	1.15	-0.65	-1.00	-0.35
LeChequer_stimulus18	http://ppp.il	http://ppp.il	Moi j'avais essayé d'aborder mais évidemment, la façon dont vous avez choisi d'engager le débat avait vocation à créer cette polémique, cette agitation, "les gentils / les méchants", "ceux qui s'en occupent ceux / qui hurlent", et (...)	1.33	1.62	-0.28	4.44	4.46	-0.02	0.19	1.00	0.81	-0.87	-1.00	-0.13	-0.58	-1.00	-0.42
LeChequer_stimulus19	http://ppp.il	http://ppp.il	que vous souhaitiez le prendre comme cela, avec dédain, comme vous en avez l'habitude dans la façon dont vous traitez la population, c'est votre problème, mais sur cette délibération, une fois encore je vous le dit	1.67	2.00	-0.33	4.67	4.92	-0.26	1.28	1.00	-0.28	-1.47	-1.00	0.47	-0.30	-1.00	-0.70
LeChequer_stimulus20	http://ppp.il	http://ppp.il	Pour mettre fin à l'activité de cette SEM, qui devait certainement receler les pires horreurs que veut bien fantasmer notre collègue Gilles Robel.	1.44	1.00	0.44	3.11	2.46	0.65	0.37	1.00	0.63	-0.32	1.00	1.32	-0.66	-1.00	-0.34
LeChequer_stimulus21	http://ppp.il	http://ppp.il	De même qu'il est inutile à mon avis de vous embringer sur des histoires politico-politiciennes, si vous avez des problèmes avec le passé, sondez votre propre morale, pas la nôtre.	1.67	2.00	-0.33	4.67	4.31	0.36	1.42	1.00	-0.42	-1.28	-1.00	0.28	-0.39	-1.00	-0.61
LeChequer_stimulus22	http://ppp.il	http://ppp.il	En cette période pré-électorale, ça serait quand même dommage que nos concitoyens, qui ne nous suivent pas sur le conseil municipal mais sur twitter, ne puissent pas avoir les arguments pertinents de nos collègues.	1.22	0.77	0.45	2.56	2.31	0.25	0.93	1.00	0.07	-0.47	1.00	1.47	-0.57	-1.00	-0.43
LeChequer_stimulus23	http://ppp.il	http://ppp.il	Je vous rappelle une nouvelle fois, que nous étions vous et nous d'accord pour la réalisation d'une piscine dans les hauts de Montreuil, que cela faisait partie de nos engagements, et (...)	1.44	1.08	0.37	4.33	3.69	0.64	1.07	1.00	-0.07	-1.12	0.00	1.12	-0.71	0.00	0.71
Molossi_stimulus1	http://ppp.il	http://ppp.il	Non, mais il faut que les Montreuillois sachent, Madame, que vous nous proposez en perspective le développement de la police municipale, pourquoi pas mais qu'est-ce que ça veut dire? nous n'en savons rien. "Convention police nationale", qu'est-ce que cela veut dire ?	2.00	1.54	0.46	4.22	3.85	0.38	1.18	-1.00	-2.18	-1.95	-2.00	-0.05	1.85	-1.00	-2.85
Molossi_stimulus2	http://ppp.il	http://ppp.il	La promiscuité qu'a décrite ma collègue Alexie Loreca dans son propos, chacun d'entre nous peut la constater. Et comme nous nous pensons que l'essentiel du dispositif doit nous conduire au vivre ensemble et à l'intégration, vous nous excuserez parfois de vous alerter sur le fait que (...)	0.89	0.23	0.66	2.22	1.69	0.53	0.78	0.00	-0.78	-0.82	0.00	0.82	0.48	0.00	0.48
Molossi_stimulus3	http://ppp.il	http://ppp.il	Je rajoute qu'il manque un élément à mes yeux est fondamental, qui est la question de la prévention : pas un mot dans ce document par exemple sur les clubs de prévention, tout juste au passage on nous parle de la P. J.J., et on se félicite de vingt-neuf T.I.G dans notre ville. Qu'en est-il de la politique de la jeunesse (...)?	1.56	0.54	1.02	2.11	1.62	0.50	0.98	-1.00	-1.98	-1.44	-1.00	0.44	1.27	1.00	-0.27
Molossi_stimulus4	http://ppp.il	http://ppp.il	Je ne sais pas si c'est cela qui a choqué tel ou tel membre de leur direction ou du gouvernement, donc très sincèrement, c'est votre choix ça n'est certainement pas le notre, et en ce qui nous concerne comme élus locaux, comme habitants de cette ville, (...)	1.00	1.77	-0.77	4.00	3.31	0.69	0.35	0.00	-0.35	-0.62	0.00	0.62	-0.83	0.00	0.83

nom du stimulus	original	relecture	transcription simplifiée étendue	host	hostE	dHost	Dest	DestE	dDest	Dom	Ctrl	dDom	ValN	ValA	dVal	ExcN	ExcA	dExc
Molossi_stimulus5	http://ppp.ilp	http://ppp.ilp	Il a été fait référence à plusieurs reprises, sur un certain nombre de dysfonctionnements graves, pour ne pas les qualifier autrement, et qui auraient été commis, je crois même qu'une procédure devant les prudhommes a été engagée, et (...)	0.89	1.38	-0.50	1.78	1.69	0.09	0.73	-2.00	-2.73	-0.95	-2.00	-1.05	-0.50	-1.00	-0.50
Molossi_stimulus6	http://ppp.ilp	http://ppp.ilp	Oui, nous avons déjà eu, comme le dit d'ailleurs la délibération, ce débat lors de l'une de nos précédentes séances du conseil communautaire (...)	0.44	0.54	-0.09	2.44	2.23	0.21	-0.57	0.00	0.57	-0.15	0.00	0.15	-0.55	0.00	0.55
Molossi_stimulus7	http://ppp.ilp	http://ppp.ilp	pour vanter non seulement les mérites de la fondation de cette communauté d'agglomération, mais également la dynamique territoriale qu'elle visait à mettre en œuvre, notamment avec l'adoption de son contrat de développement territorial.	0.22	0.23	-0.01	2.00	1.46	0.54	0.58	0.00	-0.58	0.06	1.00	0.94	-0.81	-2.00	-1.19
Molossi_stimulus8	http://ppp.ilp	http://ppp.ilp	(...) mais pardonnez-nous, à chaque fois, à chaque fois, vous rembarrez, chaque argument, même constructif, n'est jamais pris en considération, et vous continuez, sans jamais vous interroger ou remettre en cause ou quoique ce soit, vous vous obstinez (...)	1.78	1.92	-0.15	5.00	4.46	0.54	1.12	1.00	-0.12	-1.62	-2.00	-0.38	0.46	-2.00	-2.46
Molossi_stimulus9	http://ppp.ilp	http://ppp.ilp	Oui, oui, je vous en prie, permettez que j'aie au terme de non intervention. Nous voterons naturellement cette expérimentation, nous la voterons parce que elle est voulue, et que (...)	1.22	1.00	0.22	4.78	4.08	0.70	0.97	1.00	0.03	-1.38	-1.00	0.38	0.93	-1.00	-1.93
Molossi_stimulus10	http://ppp.ilp	http://ppp.ilp	Donc je tenais à le dire parce que j'aimerais bien qu'on évite de parler en notre nom et en notre place.	1.56	1.54	0.02	2.78	3.08	-0.30	1.32	-1.00	-2.32	-1.52	-2.00	-0.48	-0.04	1.00	1.04
Molossi_stimulus11	http://ppp.ilp	http://ppp.ilp	(...) aujourd'hui se mettre en place c'était quand même le minimum qu'il aurait pu être envisagé de faire, j'en aurai donc fini en ayant tenu cinq minutes.	1.33	0.92	0.41	2.00	2.46	-0.46	0.72	1.00	0.28	-0.78	-1.00	-0.22	-0.55	1.00	1.55
Molossi_stimulus12	http://ppp.ilp	http://ppp.ilp	(...) je crois, avec le sens des responsabilités, et avec une forme de hauteur de vue, sur ce dossier, sur lequel nous n'avons jamais cédé à la caricature (...)	0.67	0.15	0.51	2.11	2.15	-0.04	0.52	-1.00	-1.52	-0.11	-1.00	-0.89	-0.34	-1.00	-0.66
Molossi_stimulus13	http://ppp.ilp	http://ppp.ilp	Nous sommes attachés comme beaucoup ont pu le dire ici, au monde associatif dans ce qu'il permet d'éducation à la citoyenneté, dans ce qu'il permet de travail social et d'irrigation de l'ensemble des quartiers de notre ville, (...)	0.33	0.15	0.18	1.78	1.38	0.39	0.21	0.00	-0.21	0.00	0.00	0.00	-1.22	0.00	1.22
Molossi_stimulus14	http://ppp.ilp	http://ppp.ilp	À partir de là vous comprendrez bien qu'en effet, sur ce dossier, pour ne pas en rajouter, nous avons décidé, pour notre part, de ne pas participer au vote.	1.22	1.08	0.15	3.00	2.85	0.15	0.55	-1.00	-1.55	-0.86	-1.00	-0.14	-0.36	0.00	0.36
Molossi_stimulus15	http://ppp.ilp	http://ppp.ilp	le taux de non titulaires, qui aujourd'hui est de 23%, alors peut-être que les informations sont mauvaises, mais simplement là je trouve quand même que 23%, très sincèrement, c'est un peu (...)	1.56	0.85	0.71	3.22	2.62	0.61	0.67	-1.00	-1.67	-1.27	-1.00	0.27	1.27	-1.00	-2.27
Molossi_stimulus16	http://ppp.ilp	http://ppp.ilp	Alors le "encore" est de trop, mais je vous remercie Madame la Maire de me donner la parole, bonsoir mes chers collègues, non juste deux remarques : la première est sans aucun sous entendu, le fait est, et ce sont des éléments factuels, que la facture finale du projet sera passée du simple au double.	0.89	1.46	-0.57	4.00	2.92	1.08	0.30	0.00	-0.30	0.05	-1.00	-1.05	-1.12	0.00	1.12
Molossi_stimulus17	http://ppp.ilp	http://ppp.ilp	Cela étant, comme il n'y a aucune analyse de ces données, et surtout aucun bilan, aucune mise en perspective de l'action municipale, sur l'ensemble des documents qui apparaissent là, vous comprendrez que (...)	1.11	1.46	-0.35	2.44	2.31	0.14	1.08	-1.00	-2.08	-1.13	-1.00	0.13	-0.14	-1.00	-0.86
Molossi_stimulus18	http://ppp.ilp	http://ppp.ilp	Encore une fois, pas de parti pris, une espèce de catalogue à la Prévert, de lieu commun hein, en nous décrivant quelles sont les prérogatives du Maire et du Préfet, comme si c'était un élément fondamental et constitutif d'une véritable politique en matière de tranquillité publique	1.56	1.46	0.09	2.56	2.62	-0.06	0.82	1.00	0.18	-1.58	-1.00	0.58	0.71	-1.00	-1.71
Molossi_stimulus19	http://ppp.ilp	http://ppp.ilp	Et enfin : "augmentation des points école", c'est à peu près d'ailleurs le seul élément, clairement c'est-à-dire "augmentation" suppose qu'on va en avoir plus et on ne serait que s'en féliciter.	1.56	0.69	0.86	2.89	2.23	0.66	0.72	-1.00	-1.72	-0.65	-2.00	-1.35	1.26	1.00	-0.26
Molossi_stimulus20	http://ppp.ilp	http://ppp.ilp	Non non, mais j'ai eu le sentiment, et je voudrais ici profiter de l'occasion qui m'est donnée pour lever toute ambiguïté, même si ni ma collègue ni moi-même, ça ne vous aura échappé, sommes des dirigeants du parti socialiste, nous avons malgré tout une forme de liberté de conscience (...)	1.00	0.54	0.46	2.33	2.31	0.03	0.18	0.00	-0.18	-0.66	-1.00	-0.34	0.31	0.00	-0.31
Molossi_stimulus21	http://ppp.ilp	http://ppp.ilp	J'aimerais tout bêtement savoir la chose suivante, y'a pas trente-six hypothèses : ou bien ces agents ont acheté sur leurs fonds propres et personnels lesdits équipements, ou bien alors l'employeur a autorisé l'achat de ces dits équipements,	1.89	1.54	0.35	3.22	3.31	-0.09	1.07	-1.00	-2.07	-1.48	-1.00	0.48	0.77	1.00	0.23
Molossi_stimulus22	http://ppp.ilp	http://ppp.ilp	J'aimerais savoir dans quel cas de figure nous nous situons de manière totalement précise sur cette question, parce que faire semblant de ne pas voir ce qui se voit, et ce que tout le monde avait vu, (...)	1.22	0.62	0.61	3.11	2.31	0.80	0.77	0.00	-0.77	-0.82	0.00	0.82	-0.20	1.00	1.20
Serey_stimulus1	http://ppp.ilp	http://ppp.ilp	Par exemple, je me souviens très bien qu'il y avait un commissaire de police, qui assurait de la médiation, de façon bénévole, c'est vrai que vous êtes passés après avec des gens qui l'assuraient en étant effectivement agent municipaux, bon.	1.00	0.62	0.38	4.36	4.33	0.03	-0.17	0.00	0.17	-0.63	-1.00	-0.37	0.01	2.00	1.99
Serey_stimulus2	http://ppp.ilp	http://ppp.ilp	Ensuite concernant la transparence, simplement : les deux administrateurs de l'office qui n'ont pas pu assister au séminaire n'ont pas eu les documents.	0.91	0.54	0.37	2.91	2.92	-0.01	0.30	0.00	-0.30	-0.75	0.00	0.75	-0.29	0.00	0.29
Serey_stimulus3	http://ppp.ilp	http://ppp.ilp	Contrairement à ce qu'il a dit, y'a bien du logement social qui est en vente dans la ville actuellement; depuis un an, rue de Rosny devant ACLPME, y'a un panneau quatre par trois : « logements à vendre ».	1.18	0.85	0.34	3.64	2.92	0.71	0.70	0.30	-0.77	-1.00	-0.23	0.33	0.00	-0.33	
Serey_stimulus4	http://ppp.ilp	http://ppp.ilp	Je dois dire que je suis un peu gêné par ce débat parce que il est dans un contexte politique national, qui date pas de ces derniers mois d'ailleurs, qui est bien plus ancien, qui tente à diminuer l'expression politique des syndicats, (...)	0.55	0.38	0.16	2.00	2.54	-0.54	-0.16	1.00	1.16	-0.68	0.00	0.68	-0.86	1.00	1.86
Serey_stimulus5	http://ppp.ilp	http://ppp.ilp	donc voilà, c'est un peu dans ce contexte là qu'est prise cette délibération à quelques mois des élections municipales	1.00	0.23	0.77	2.00	2.54	-0.54	0.18	1.00	0.82	-0.40	-1.00	-0.60	-0.39	1.00	1.39
Serey_stimulus6	http://ppp.ilp	http://ppp.ilp	Il a choisi de développer les CAP, les centres d'activités : CAP Gaillard, Beaune, Voltaire, Marceau, Saint-Antoine, ou encore une pépinière d'entreprise.	0.09	0.62	-0.52	1.36	3.23	-1.87	0.25	1.00	0.75	0.03	0.00	-0.03	-0.75	1.00	1.75
Serey_stimulus7	http://ppp.ilp	http://ppp.ilp	Bah il faut simplement rappeler, que ces agents municipaux existaient à l'époque, ils assuraient la police du stationnement, la police sur la voirie, et que pour les intégrer dans le statut de la fonction territoriale, (...)	0.82	0.15	0.66	2.36	3.08	-0.71	0.72	1.00	0.28	-0.39	0.00	0.39	-0.57	1.00	1.57
Serey_stimulus8	http://ppp.ilp	http://ppp.ilp	Ceci étant, moi ya quelque chose, j'ai jeté un petit coup d'oeil là rapidement, sur les propositions qu'étaient faites par un certain nombre d'élus d'Europe Ecologie Les Verts sur ces questions là, (...)	0.73	0.54	0.19	2.45	3.31	-0.85	0.14	0.00	-0.14	-0.62	0.00	0.62	1.06	1.00	-0.06
Serey_stimulus9	http://ppp.ilp	http://ppp.ilp	(...) aussi bien sur les murs de nos villes, alors que c'est une tradition française depuis très très longtemps ; qu'avoir y compris des affiches politiques qui aujourd'hui valent des milliers de francs, et qui (...)	0.82	0.08	0.74	2.27	3.31	-1.03	0.04	1.00	0.96	-0.61	-1.00	-0.39	-0.65	1.00	1.65

nom du stimulus	original	relecture	transcription simplifiée étendue	host	hostE	dHost	Dest	DestE	dDest	Dom	Ctrl	dDom	ValN	ValA	dVal	ExcN	ExcA	dExc
Serey_stimulus10	http://ppp.ilp	http://ppp.ilp	J'ai vraiment le sentiment que c'est à géométrie variable, parce que, en centre ville, sur la place de la Résistance, on en trouve beaucoup (...)	0.36	0.54	-0.17	2.82	3.31	-0.49	0.15	1.00	0.85	-0.74	-1.00	-0.26	-0.31	1.00	1.31
Serey_stimulus12	http://ppp.ilp	http://ppp.ilp	(...) mais on arrive ce soir donc à la création, à la mise en place de caméras, bon pourquoi pas, si c'est effectivement la solution pour assurer la tranquillité, mais (...)	1.73	0.54	1.19	2.55	3.38	-0.84	0.53	1.00	0.47	-0.85	0.00	0.85	0.95	1.00	0.05
Serey_stimulus14	http://ppp.ilp	http://ppp.ilp	Il y a deux mois, nous avons voté l'achat des locaux de la poste, rue de Stalingrad, pour une bonne opération, faire une école, et ce mois-ci un nouveau vote (...)	0.55	0.00	0.55	2.09	2.00	0.09	0.03	1.00	0.97	0.21	0.00	-0.21	-0.61	1.00	1.61
Serey_stimulus15	http://ppp.ilp	http://ppp.ilp	J'ai noté qu'à Lyon, par exemple, le président du groupe Europe Ecologie Les Verts a demandé à ce qu'il y ait une charte de déontologie avant l'installation de la vidéo-surveillance.	0.73	0.08	0.65	3.09	2.15	0.94	0.31	1.00	0.69	-0.15	0.00	0.15	-0.01	-1.00	-0.99
Serey_stimulus17	http://ppp.ilp	http://ppp.ilp	(...), mais il est vrai que vous n'êtes pas responsable, non, pour le PLU ce sont messieurs Martinez et Saunier les coupables, pour le reste dans votre réponse à la chambre régionale des comptes, la faute en revient soit à votre prédécesseur, soit (...)	1.55	1.15	0.39	4.64	4.23	0.41	0.96	2.00	1.04	-0.99	-1.00	-0.01	-0.06	-1.00	-0.94
Serey_stimulus18	http://ppp.ilp	http://ppp.ilp	Je me dis aujourd'hui mais comment va se comporter la personne qui va être derrière ces vidéos, comment ça va être utilisé? Si vous maîtrisez pas plus ce dispositif que vous avez maîtrisé votre directeur, permettez-moi quand même d'avoir quelques inquiétudes.	1.82	1.62	0.20	5.00	4.85	0.15	1.19	2.00	0.81	-1.39	0.00	1.39	0.97	-1.00	-1.97
Serey_stimulus19	http://ppp.ilp	http://ppp.ilp	Y'a une question qui a été posée, qu'est-ce qu'ils faisaient dans cette galerie à deux heures du matin, à la Halle Marcel Dufriche, hein, on aimerait bien avoir la réponse quand même ce soir. Or je comprends que pour certains ça soit très difficile le grand écart, pour Patrick Petitjean il s'est très bien exprimé hein d'ailleurs.	1.18	1.15	0.03	3.09	2.69	0.40	0.68	1.00	0.32	-0.69	-1.00	-0.31	0.06	0.00	-0.06
Serey_stimulus20	http://ppp.ilp	http://ppp.ilp	La cour des comptes, elle vous dit : "attention, la santé de l'office, peut être il faudrait regarder tous les ans, pour qu'on ait un bilan". On n'a rien, mais là ça vous gêne pas, là, comme par hasard. Voilà, deux poids deux mesures, c'est un peu gênant quand même dans vos propos.	2.00	1.69	0.31	5.00	4.54	0.46	0.75	2.00	1.25	-1.40	-2.00	-0.60	0.16	1.00	0.84
Serey_stimulus21	http://ppp.ilp	http://ppp.ilp	(...), par contre moi ce que j'aurais aimé quand même, c'est que vous nous parliez de vos voltigeurs là, de votre police mobile, parce que c'était quand même ça la question, il y a un mois, (...)	2.00	1.54	0.46	5.00	4.85	0.15	1.28	2.00	0.72	-1.40	-1.00	0.40	1.25	0.00	-1.25
Serey_stimulus22	http://ppp.ilp	http://ppp.ilp	Je me dis aujourd'hui mais comment va se comporter la personne qui va être derrière ces vidéos, comment ça va être utilisé? Si vous maîtrisez pas plus ce dispositif que vous avez maîtrisé votre directeur, permettez-moi quand même d'avoir quelques inquiétudes.	0.91	0.54	0.37	1.73	2.62	-0.89	0.62	1.00	0.38	-0.90	0.00	0.90	0.49	-1.00	-1.49
Serey_stimulus23	http://ppp.ilp	http://ppp.ilp	(...) y'a une question précise qui a été posée, et j'attends qu'on ait la réponse ce soir, elle a été posée plusieurs fois : sur quelle ligne budgétaire a été pratiqué l'achat des bombes lacrymogènes, ... (...)	1.45	1.15	0.30	4.00	3.62	0.38	1.03	2.00	0.97	-1.06	-1.00	0.06	0.34	0.00	-0.34
Tuailon_stimulus1	http://ppp.ilp	http://ppp.ilp	(...), que ces missions cessent et que tout s'arrête, et dans ce cas là manifestement vous n'avez pas été écoutée, ce qui pose alors la question de la responsabilité, au sein de votre administration générale, et au sein de votre propre équipe, notamment en la personne de la Maire Adjointe déléguée à la tranquillité publique.	1.27	1.15	0.12	4.36	4.31	0.06	1.02	-2.00	-3.02	-1.31	-1.00	0.31	0.90	-1.00	-1.90
Tuailon_stimulus2	http://ppp.ilp	http://ppp.ilp	Enfin sur la sécurité Madame la Maire, je suis désolé vous avez énoncé un certain nombre de choses, les caméras sur le parking, oui très bien quand elles fonctionnent, les renforts, oui très bien mais pas toujours placés là où il faut, (...)	1.91	1.31	0.60	5.00	4.69	0.31	0.99	1.00	0.01	-1.45	-2.00	-0.55	0.46	1.00	0.54
Tuailon_stimulus3	http://ppp.ilp	http://ppp.ilp	Que vous ne teniez pas compte de l'avis des habitants avant de lancer des projets, que vous fassiez ça en catimini en permanence, et bien ça c'est une erreur.	1.91	1.92	-0.01	4.64	4.54	0.10	0.73	-1.00	-1.73	-1.01	-2.00	-0.99	0.33	2.00	1.67
Tuailon_stimulus4	http://ppp.ilp	http://ppp.ilp	En revanche, durant votre intervention Madame Voynet, au conseil communautaire, à aucun moment, vous n'avez fait état de votre contestation concernant la sous-représentation de Montreuil.	1.09	1.08	0.01	4.82	4.85	-0.03	0.04	2.00	1.96	-0.60	-2.00	-1.40	-0.19	2.00	2.19
Tuailon_stimulus5	http://ppp.ilp	http://ppp.ilp	J'ai bien entendu les propos de Madame Zeidenberg, moi je trouve quand même que il y a quelques différences entre le rapport interne que vous aviez présenté en 2011, et ce rapport de la chambre régionale des comptes.	0.82	0.85	-0.03	4.73	4.69	0.03	0.26	1.00	0.74	-0.25	-1.00	-0.75	-0.35	2.00	2.35
Tuailon_stimulus6	http://ppp.ilp	http://ppp.ilp	Vous avez raison Madame Voynet, vous aviez dit en conseil municipal du 6 juin, que vous alliez organiser un débat en conseil municipal suivant, sur la tranquillité publique, et vous estimez que c'était une chose acquise.	0.91	0.62	0.29	4.64	4.69	-0.06	0.47	2.00	1.53	-0.08	0.00	0.08	-0.87	-2.00	-1.13
Tuailon_stimulus7	http://ppp.ilp	http://ppp.ilp	L'accompagnement des responsables associatifs, leur formation, leur information, est une démarche tout aussi prioritaire que le versement de la subvention.	0.09	0.15	-0.06	1.36	2.31	-0.94	0.76	2.00	1.24	0.06	0.00	-0.06	-0.17	2.00	2.17
Tuailon_stimulus8	http://ppp.ilp	http://ppp.ilp	C'est un vrai sujet, qui mérite une réflexion sur laquelle la chambre n'a pas à se prononcer, mais qui doit guider aussi cette réforme du COS dans un avenir proche.	0.18	0.31	-0.13	1.55	2.31	-0.76	0.85	1.00	0.15	-0.04	0.00	0.04	-0.59	1.00	1.59
Tuailon_stimulus9	http://ppp.ilp	http://ppp.ilp	Il y a trois mois vous aviez décidé de baptiser le chantier de l'avenue de la Résistance, de l'école avenue la Résistance : "Stéphane Hessel", et je m'aperçois que dans la délibération suivante, la 9, vous continuez à appeler cette école l'Ecole de la Résistance, donc il semblerait que même les dénominations que vous décidez, vous avez du mal à les retenir	1.00	0.62	0.38	4.82	4.69	0.13	0.61	1.00	0.39	-0.15	-1.00	-0.85	-0.56	1.00	1.56
Tuailon_stimulus10	http://ppp.ilp	http://ppp.ilp	Huit des neuf maires qui composent cette agglomération Est Ensemble se sont prononcés pour le statu quo.	0.36	1.00	-0.64	2.27	4.62	-2.34	0.47	2.00	1.53	0.01	-2.00	-2.01	-0.81	2.00	2.81
Tuailon_stimulus11	http://ppp.ilp	http://ppp.ilp	Vous avez fait le choix de ne pas jouer l'hégémonie, et vous avez accepté une sous-représentation de notre ville au sein d'Est Ensemble, afin de ne pas effrayer nos partenaires.	0.91	0.08	0.83	4.64	1.77	2.87	0.46	2.00	1.54	0.04	0.00	-0.04	-0.50	2.00	2.50
Tuailon_stimulus12	http://ppp.ilp	http://ppp.ilp	Donc nous sommes contents de le découvrir maintenant, merci pour la primauté démocratique de cette présentation, par contre ce qu'il faudrait savoir(...)	1.36	0.85	0.52	3.91	3.15	0.76	0.41	2.00	1.59	0.76	2.00	1.24	0.17	-2.00	-2.17
Tuailon_stimulus13	http://ppp.ilp	http://ppp.ilp	(...) je pense notamment à un article que nous avions écrit à ce sujet le 25 mai 2010, et puis vous avez été aussi fortement comme nous tous, destinataires de mails et puis d'informations, de blogs, de vidéos, etcetera, qui vous interpellaient.	0.91	0.46	0.45	3.64	3.15	0.48	0.29	-1.00	-1.29	-0.17	0.00	0.17	-0.44	-1.00	-0.56
Tuailon_stimulus14	http://ppp.ilp	http://ppp.ilp	(...) Et par ailleurs je rejoins ce qu'a dit Frédéric Molossi tout à l'heure, effectivement nous n'avons pas du tout eu le même son de cloche de votre passage le dimanche à la piscine, rappelons que vous y êtes passé en tant qu'élu de permanence, ce jour là c'est pour ça que vous avez été alertée, (...)	1.45	1.23	0.22	4.73	4.62	0.11	0.90	-2.00	-2.90	-1.18	-1.00	0.18	0.92	2.00	1.08

nom du stimulus	original	relecture	transcription simplifiée étendue	host	hostE	dHost	Dest	DestE	dDest	Dom	Ctrl	dDom	ValN	ValA	dVal	ExcN	ExcA	dExc
Tuailon_stimulus15	http://ppp.ilp	http://ppp.ilp	Vous parlez de l'argent du contribuable, vous avez raison, l'argent du contribuable il faut le respecter, je rappellerai simplement que autour de cette table, il y a trente élus de la majorité, et que les trente élus de la majorité ont tous une délégation.	2.00	0.69	1.31	3.64	1.85	1.79	1.36	-2.00	-3.36	-1.53	-2.00	-0.47	1.47	2.00	0.53
Tuailon_stimulus16	http://ppp.ilp	http://ppp.ilp	Oui simplement pour vous rappeler juste un principe de base, c'est que la politique et la démocratie ce sont aussi des principes et des méthodes, et la bonne information des conseillers municipaux, fussent-ils d'opposition, en fait partie de ces principes et de ces méthodes.	1.36	0.54	0.83	3.36	2.85	0.52	0.86	2.00	1.14	-0.59	-2.00	-1.41	0.00	2.00	2.00
Tuailon_stimulus17	http://ppp.ilp	http://ppp.ilp	Il y a moins d'écoliers montreuillois qui vont prendre des cours de tennis que par le passé, alors c'est pas un reproche mais c'est une alerte, il faudrait comprendre pourquoi, et tenter de remédier à la situation.	0.91	0.69	0.22	3.18	3.46	-0.28	0.73	2.00	1.27	-0.58	1.00	1.58	0.41	-2.00	-2.41
Tuailon_stimulus18	http://ppp.ilp	http://ppp.ilp	La ville de Montreuil aurait pu rendre un hommage particulier à l'un de nos concitoyens, je pense à Léon Zygel , qui est résistant, déporté, et qui a passé énormément de temps, des années et des années dans les écoles de Montreuil (...)	0.55	0.77	-0.22	1.73	2.23	-0.50	0.27	1.00	0.73	-0.57	0.00	0.57	-0.37	-1.00	-0.63
Tuailon_stimulus19	http://ppp.ilp	http://ppp.ilp	(...) qui est résistant, déporté, et qui a passé énormément de temps, des années et des années dans les écoles de Montreuil, pour faire la pédagogie du totalitarisme, de l'antisémitisme et de la xénophobie, ça aurait pu éventuellement avoir un sens, (...)	0.27	0.15	0.12	2.18	2.15	0.03	0.03	2.00	1.97	-0.04	0.00	0.04	-0.59	0.00	0.59
Tuailon_stimulus20	http://ppp.ilp	http://ppp.ilp	Oui, je serai très bref comme vous me l'avez demandé Madame la Maire, simplement pour apporter une correction à ce que vous venez de dire, la délibération que nous avons voté (...)	1.18	1.00	0.18	4.73	4.38	0.34	-0.05	1.00	1.05	-0.56	2.00	2.56	-0.36	-1.00	-0.64
Tuailon_stimulus21	http://ppp.ilp	http://ppp.ilp	Je voudrais pas que M. Cuffini appelle ce genre de budget un budget électoraliste, mais c'est vrai qu'on peut noter une sorte de concomitance avec ce petit programme de grands travaux ou grand programme de petits travaux je sais plus très bien, qui a été d'ailleurs présenté il y a plusieurs mois en conseils de quartier, dans différents conseils de quartier.	1.18	0.69	0.49	2.82	3.08	-0.26	0.07	2.00	1.93	0.09	2.00	1.91	0.39	-2.00	-2.39
Tuailon_stimulus22	http://ppp.ilp	http://ppp.ilp	Bref le six juin dernier vous disiez également à ce conseil municipal, que puisque vous n'aviez pu tenir ce conseil jusqu'à son terme faute de quorum au sein de votre majorité, vous le reconvoquez pour le treize juin, et puis il a fallu que je vous relance par mail, (...)	1.45	1.00	0.45	4.91	4.85	0.06	0.76	-1.00	-1.76	-0.08	1.00	1.08	0.22	2.00	1.78
Tuailon_stimulus23	http://ppp.ilp	http://ppp.ilp	Dans son rapport, la chambre régionale des comptes présente une analyse précise et surtout complètement dépassionnée hein, une analyse donc à mille lieues de la volonté d'instrumentalisation qui vous animait en 2011.	0.91	1.15	-0.24	2.36	4.54	-2.17	0.76	1.00	0.24	-0.19	-1.00	-0.81	-0.30	1.00	1.30
Tuailon_stimulus24	http://ppp.ilp	http://ppp.ilp	(...) Je pense notamment, entre autres, à des yoyos dans les versements de subventions, qui ne permettaient pas une programmation budgétaire sereine.	0.91	1.77	-0.86	2.45	4.38	-1.93	0.72	2.00	1.28	-0.43	1.00	1.43	0.61	2.00	1.39
Tuailon_stimulus25	http://ppp.ilp	http://ppp.ilp	(...) et que vous souhaitez désormais soumettre à l'avis des élus de ce conseil municipal, et bah très bien, parlons en mais encore une fois ne déformez pas les propos, y'a pas des gens qui veulent fuir le débat, simplement des gens qui veulent être un peu respectés, pour leurs qualités d'élus, vous nous reprochiez tout à l'heure, vous reprochiez de manière générale, (...)	1.64	1.54	0.10	4.82	4.23	0.59	0.69	-1.00	-1.69	-1.21	-1.00	0.21	0.76	2.00	1.24