



## Approches méthodologiques pour l'étude de la parole expressive en politique :

Analyses perceptives d'extraits d'un débat conflictuel, le conseil  
municipal de Montreuil en 2013

*Methodological approaches to the phonetic study of conflict-ridden political  
debates: Perceptive analysis of expressive speech extracted from Montreuil's City  
Council in 2013*

**Charlotte Kouklia**

---

Revue.org

**Éditeur**

Publications de l'Université de Provence

**Édition électronique**

URL : <http://tipa.revues.org/1690>

DOI : 10.4000/tipa.1690

ISSN : 2264-7082

**Référence électronique**

Charlotte Kouklia, « Approches méthodologiques pour l'étude de la parole expressive en politique : », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 32 | 2016, mis en ligne le 27 janvier 2017, consulté le 11 mai 2017. URL : <http://tipa.revues.org/1690> ; DOI : 10.4000/tipa.1690

---

Ce document a été généré automatiquement le 11 mai 2017.

© Travaux interdisciplinaires sur la parole et le langage

---

# Approches méthodologiques pour l'étude de la parole expressive en politique :

Analyses perceptives d'extraits d'un débat conflictuel, le conseil municipal de Montreuil en 2013

*Methodological approaches to the phonetic study of conflict-ridden political debates: Perceptive analysis of expressive speech extracted from Montreuil's City Council in 2013*

Charlotte Kouklia

---

## Préambule

- 1 Lieu de tensions fortement médiatisées, le débat politique pose la question du rapport de force dans une interaction parfois conflictuelle. La présente étude s'interroge sur le rôle que jouent les caractéristiques expressives de parole dans une telle situation de communication. Tout en insistant sur la complexité de mener une étude phonétique, tant acoustique que perceptive, de la parole politique, cet article propose un ensemble d'approches et de méthodologies complémentaires permettant de poser les jalons d'une étude phonétique appliquée à la parole naturelle.

## Présentation des tensions, enjeux et stratégies inhérentes au débat politique

- 2 Cette étude se base sur des débats issus d'archives télévisées de séances du conseil municipal de Montreuil<sup>2</sup> en 2013, durant les derniers mois de mandat de Maire de Dominique Voynet (Europe Écologie Les Verts). Au cours des séances sélectionnées dans

cette étude, Dominique Voynet est encore pressentie comme potentiellement en campagne pour un second mandat<sup>3</sup>. 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 entre la majorité et les partis d'opposition un ensemble d'affrontements verbaux et, plus rarement, physiques. La durée des séances, de 6 h en moyenne, rend également compte des fortes tensions entre la majorité et les partis d'opposition<sup>4</sup> ; le conseil démarrant à 19 h pour ne pas finir avant 1 h voire 2 h du matin.

- 3 La teneur hautement conflictuelle des échanges et l'accroissement des débats fleuves sont notamment liés à l'approche des élections municipales. Les enjeux, à l'échelle de la ville et les retombées médiatiques à l'échelle nationale, sont analogues dans une moindre mesure à ce qui peut être observé durant les élections présidentielles. Comme l'illustrent les recherches de Marion Ballet (2012), les périodes électorales sont des moments particulièrement propices à la médiatisation des personnalités politiques, principalement via une forme d'appel aux émotions, notamment en suscitant chez les électeurs des émotions telles que la peur, l'espoir, la compassion ou l'indignation. Les partis d'opposition ont souvent recours à une forme d'appel à l'indignation, par exemple sur le plan des affaires administratives et sociales. De telles stratégies politiques opèrent dans le but de susciter une colère fédératrice, une envie de changement et sont particulièrement mises en place par les partis d'opposition (Ballet, 2012).
- 4 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, 2015a). L'*ethos*, la construction de l'image de soi, correspond à un pan fondamental de la communication politique car il coïncide avec l'image du politicien telle qu'elle est véhiculée par celui-ci pour convaincre. Il fait partie des trois outils de l'argumentation politique au même titre que le *logos* (argumentation par la raison) et le *pathos* (argumentation par les émotions). Et si, pour Patrick Charaudeau (2015b), dans le débat politique, il est encore plus question de persuader que d'argumenter, il n'est pas étonnant que l'*ethos*, « moyen de persuasion le plus puissant » selon Aristote<sup>5</sup>, marque particulièrement les paroles politiques.
- 5 Dans la conception traditionnelle, pour convaincre et avoir des chances d'être élu, un politicien se doit de véhiculer l'image d'une personne morale, bienveillante et compétente. Pour Isabella Poggi et Francesca D'Errico (2010), le politicien devra également apparaître comme ayant plus de pouvoir que l'autre dans une situation donnée. Cet ajout de la dominance aux dimensions de l'*ethos* s'explique, selon ces chercheuses en psychologie sociale, par le fait que, malgré son honnêteté et ses compétences, on donnera difficilement sa voix à un politicien qui n'arrive pas à prendre l'ascendant sur l'autre, notamment dans un débat. Dans leur étude sur les signaux de dominance, Poggi & D'Errico (2010) indiquent qu'ils seraient, entre autres, constitués d'une visible maîtrise des émotions (contrôle de soi) et d'agressivité plus ou moins tacite (mépris, sarcasme, irritation, indignation). La dominance que l'on attribuera alors perceptivement au locuteur dépendra de l'image que l'on va se faire de lui, et de celle que l'on va se faire de son interlocuteur. Ainsi, comme celle-ci se construit dans un rapport de comparaison interactif, il n'est pas rare de voir les interlocuteurs essayer d'entacher l'image de leur opposant (Poggi *et al.*, 2011). En effet, une stratégie de dominance très courante et efficace, lorsqu'elle est menée à bien, consiste à tenter de discréditer l'opposant, l'empêcher d'améliorer son image de façon plus ou moins implicite, en jouant

sur l'équilibre entre les mots employés, la façon de les dire (expressivité vocale), et la gestuelle qui les accompagne.

## Quel rôle attribuer à l'expressivité vocale ? Propositions méthodologiques pour l'étude phonétique de la parole dans le débat politique

- 6 Malgré la place importante qu'occupe la voix dans la fonction expressive du langage (Jakobson, 1963/81), celle-ci n'est, à notre sens, qu'encore trop peu analysée dans l'étude de la parole politique. Si l'on compte un nombre conséquent d'analyses sur le plan verbal (lexical, pragmatique, syntaxique) et non-verbal, on peut néanmoins regretter que les études effectuées sur le plan de l'expressivité vocale politique ne soient pas plus nombreuses. Une telle lacune peut, entre autres, s'expliquer par le fait qu'étudier la parole expressive produite en contexte pose des problèmes d'ordre méthodologique et expérimental. Ces problèmes sont en grande partie dus au fait que les possibilités de contrôle et de calibration du corpus se réduisent fortement lorsque l'on passe d'une étude menée en laboratoire à une étude menée sur de la parole naturelle. Une étape de contrôle, néanmoins toujours réalisable sur de telles données, consiste à valider et évaluer perceptivement le corpus de travail.
- 7 Il semble que cette étape d'évaluation s'avère d'autant plus cruciale que les informations perceptives obtenues peuvent contribuer à la construction d'un cadre indispensable à l'analyse acoustique des données<sup>6</sup>. Les propositions méthodologiques présentées ci-dessous seront donc mises à l'épreuve sous forme de tests de perception. Amorcer l'analyse phonétique de l'expressivité vocale dans le débat politique par sa composante perceptive semble particulièrement justifié du fait que, comme le souligne Amossi (1999) : « la notion d'*ethos* renvoie à des choses très différentes selon qu'on considère le point de vue du locuteur ou celui du destinataire : l'*ethos* visé n'est pas nécessairement l'*ethos* produit. »
- 8 **Propositions d'analyses du poids respectif de l'expressivité du contenu sémantique et des indices paralinguistiques**
- 9 La façon de parler de tout individu participe à son expressivité et apporte des informations importantes sur son état émotionnel et physique, son attitude à l'égard de l'autre, ses intentions, sa personnalité. Ces informations, véhiculées par un ensemble de caractéristiques vocales, faciales et gestuelles particulières, représentent ce que l'on appelle le message secondaire, paralinguistique, en complémentarité du message primaire, linguistique (Fónagy, 1971). En parallèle au traitement de l'information linguistique, les informations du message secondaire sont décodées par l'auditeur, qu'il s'agisse de signes involontaires (indices), comme une voix tremblante résultant d'une peur ou d'un stress, ou de signes volontaires (signaux) comme une attitude tendre ou méprisante. Que l'on ait affaire à des expressions brutes (émotions vives) ou contrôlées (attitudes, modalités), l'émergence d'expressivité vocale est intimement liée au contexte dans lequel celle-ci survient. Dans une situation de communication réelle, la distinction entre émotion brute et attitude contrôlée ne constitue pas une barrière étanche. Il est tout à fait possible de passer du brut au contrôlé (coup de rage progressivement contenu se transformant peu à peu en irritation) ou du contrôlé au brut (irritation de moins en moins contenue jusqu'à laisser échapper une forme très vive de colère).

- 10 Lorsque l'on souhaite étudier la composante vocale de la parole expressive d'un individu dans une situation de communication réelle, il est impératif de prendre en compte le fait que le message linguistique (contenu sémantique) est souvent lui aussi porteur d'expressivité. Lorsque l'on travaille sur de la parole expressive produite en contexte réel de communication, il est impossible de contrôler les phrases produites comme on le ferait dans l'élaboration d'une étude en laboratoire, pour lesquelles on privilégiera un corpus constitué de phrases dites « sémantiquement neutres » pour s'assurer que l'expressivité ne puisse être exprimée qu'au niveau vocal (cf. « le papier est sur la table », « j'entre dans le bâtiment », « il est huit heures »). Il y a toutefois une précaution possible, aisément conduite, qui consiste à mettre en place, en parallèle à tout test d'écoute, un test de perception de la transcription orthographique (alors présentée seule) des stimuli originaux sur lequel on effectue les évaluations perceptives (Gendrot, 2004 ; Audibert, 2008). Cette méthodologie sera appliquée dans la première partie des analyses perceptives de la présente étude (voir Analyses perceptives 1).
- 11 **Propositions pour l'analyse de l'*ethos* perçu**
- 12 L'*ethos* voulu et exprimé par le locuteur (*ethos* discursif), qu'il soit véhiculé de manière explicite (affirmations) ou implicite (sous-entendus), dépend toujours de la façon dont il va être perçu par les autres (Maingueneau, 1999). Outre l'influence de la situation de communication et l'impact de la comparaison possible avec son adversaire, l'*ethos* dépend, et ce de façon décisive, de l'éventuelle représentation préalable que l'on a de l'orateur (*ethos* préalable) (Charaudeau, 2015a ; Donot & Emediato, 2015). Aussi, Maingueneau (1999) souligne que plus les locuteurs sont présents sur la scène médiatique, plus ils sont directement associés à un type d'*ethos* que toute nouvelle déclaration de leur part viendra renforcer ou modifier. De ce fait, la mise en place de tests d'écoute, sur une population d'auditeurs ne connaissant pas les orateurs analysés pourra permettre d'étudier directement la perception de l'*ethos* discursif des locuteurs sans s'inquiéter de l'influence d'*ethos* préalables. L'analyse perceptive de l'*ethos* discursif par des auditeurs dits « naïfs » pourrait ensuite être éventuellement comparée à celle d'auditeurs dits « experts » (politologues, phonéticiens, analystes du discours, etc.).
- 13 Dans la présente étude, l'évaluation « naïve » a été comparée à l'auto-évaluation par les locuteurs d'extraits de leur propre parole. Cette approche, lorsqu'elle est possible, compte tenu de la difficulté de pouvoir conduire des entretiens avec des personnalités publiques aux emplois du temps surchargés, peut entre autres permettre de comparer *ethos* voulu et *ethos* perçu pour chaque extrait du corpus traité (voir Analyses perceptives partie 2).

## Le corpus CMM : sélection et traitement des données issues du conseil municipal de Montreuil (06-2013/11-2013)

### Présentation du déroulement des séances

- 14 Chaque séance du conseil municipal a un ordre du jour constitué à cette période de plus de 50 délibérations à voter. Après quelques formalités administratives, la Maire déclare ouverte la séance. Un membre de la majorité présente la première délibération à l'ordre du jour. Les membres de l'opposition peuvent demander la parole pour plus d'informations ou pour faire une remarque, puis éventuellement un membre de la

majorité, la Maire, leur répond et une fois le débat terminé fait procéder l'assemblée à un vote et les conseillers passent à la délibération suivante dans l'ordre du jour. 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.

- 15 Depuis juin 2013, un dispositif a été mis en place pour réguler les tours de parole durant le conseil municipal. Géré à partir d'un poste informatique confié à la Maire, étant aussi bien modératrice du débat qu'intervenante, le système ne permet d'ouvrir qu'un micro à la fois, selon l'ordre dans lequel les locuteurs doivent ou veulent intervenir. Ce système de régulation des tours de parole permet d'éviter les cas de chevauchements de parole, si fréquents dans les débats politiques qu'ils font obstacles à un grand nombre d'analyses acoustiques. Les données étudiées ici sont donc issues d'une interaction différée, où les interlocuteurs ne dialoguent pas spontanément mais doivent attendre leur tour pour s'exprimer. Notons que les intervenants peuvent parfois essayer de se couper la parole mais l'acoustique de la salle des fêtes dans laquelle a lieu le conseil municipal rend très difficile l'écoute de ceux-ci s'ils ne sont pas amplifiés par un microphone.

### Sélection des locuteurs

- 16 La présente étude se concentre sur la parole de 5 locuteurs : la Maire, Dominique Voynet (EELV) et 4 de ses opposants : Gaylord Le Chequer (RGC/PG), Frédéric Molossi (PS), Jean-Jacques Serey (FASE/PCF) et Alexandre Tuaille (RSM). Outre leurs interventions fréquentes et échanges récurrents et animés avec la Maire, ces locuteurs ont été sélectionnés pour leurs styles de parole dynamiques, s'écartant autant 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).

### Protocole méthodologique et traitement des données

- 17 Six séances du conseil municipal ayant eu lieu entre le 4 juin et le 21 novembre 2013 ont été analysées. À partir de 24 heures d'archives segmentées par locuteur, nous avons effectué une segmentation par locuteur puis une première annotation manuelle des tours de parole à la recherche de l'émergence d'une certaine expressivité vocale. Nous avons ensuite effectué une première sélection de 40 extraits audio<sup>7</sup> courts par locuteur. Les extraits ont été découpés de manière à éviter la présence d'indices verbaux de type injures, insultes, ainsi que pour supprimer des indices contextuels tels que les noms des interlocuteurs. Exemples de segments supprimés : « *Comme vient de le dire mon collègue Frédéric Molossi, (...)* » ; « *Oui merci Madame Voynet, (...)* ». Lorsque cela impliquait de tronquer une partie d'une phrase, les coupures ont majoritairement été effectuées à des frontières majeures entre groupes intono-syntaxiques (Delattre, 1961), et parfois à des frontières mineures. Une telle préparation des données a également été pensée de manière à obtenir des stimuli adéquats pour la passation de tests de perception. Les phrases, qu'elles soient courtes ou tronquées ne dépassent pas les 12 secondes (durée moyenne : 8,7 secondes), durée moyenne après laquelle le taux d'informations contextuelles disponibles commence à devenir trop haut pour se fier aux résultats de l'analyse perceptive (Morange & Candea, 2010).

- 18 Un ensemble de 125 stimuli, 25 pour chacun des 5 locuteurs, a ensuite été sélectionné après une évaluation perceptive de l'hostilité et du destinataire présumé du message (voir Analyses perceptives partie 1). Ce corpus réduit a pu servir de matériel de travail au cours d'entretiens individuels dirigés avec chacun des 5 locuteurs. Les entretiens comprenaient, entre autres, une tâche d'auto-évaluation de leurs productions (échelles dimensionnelles de valence, excitation et contrôle). Les résultats de cette auto-évaluation ont été comparés à la perception de 24 auditeurs francophones natifs (voir Analyses perceptives partie 2).

## Analyses perceptives partie 1 : impact du contenu sémantique sur la perception de l'hostilité et du destinataire présumé d'extraits de parole expressive

### Stimuli, protocole et consignes

- 19 Une première campagne d'analyses a été mise en place afin d'obtenir des premières informations perceptives sur chacun des 200 stimuli du corpus (40 extraits x 5 locuteurs) : le taux d'hostilité perçue, et le destinataire présumé (voir consignes ci-dessous). 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). La transcription simplifiée a été effectuée pour faciliter la lecture et s'assurer de l'intelligibilité des extraits lors des évaluations perceptives basées sur l'écrit : ajouts de marques de ponctuation et suppression d'un ensemble de marqueurs d'oralité (« euh » d'hésitation, amorces de mots répétitions, auto-corrections) (Blanche-Benveniste, 2010).
- 20 Au total, 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*. Les stimuli des deux tests ont été présentés dans un ordre aléatoire, avec possibilité de réécouter une fois chaque extrait dans le test audio.
- 21 **Des consignes similaires dans les deux versions du test (test audio et test écrit)**
- 22 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 hostilité (0), une légère hostilité (1), une forte hostilité (2).
  - 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)\*\*.

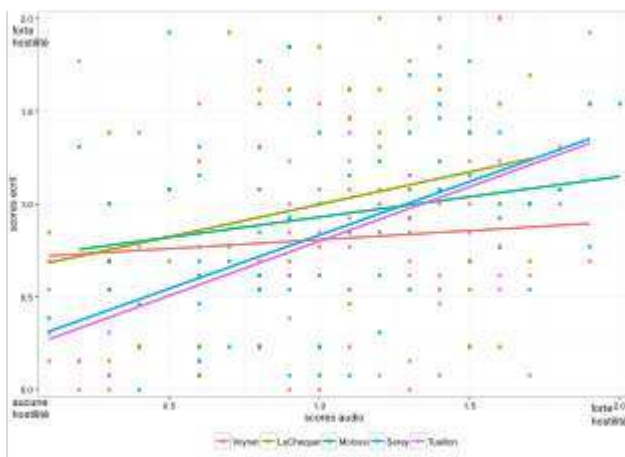
\* *Synonymes* : inimitié, animosité, malveillance ; *antonymes* : bienveillance, sympathie.

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

## Analyses des résultats obtenus dans l'évaluation de l'hostilité perçue

- 23 Tous locuteurs confondus, une corrélation positive a été calculée entre scores écrits et scores audio ( $\text{corr} = .646$ ) et on observe une répartition semblable entre les deux conditions (voir Figure 1 ci-dessous). Quelle que soit la condition (audio vs écrit), les stimuli sont globalement perçus comme exprimant une légère hostilité. Les scores moyens (audio :  $m=0.95$ ,  $\sigma=0.02$  ; écrit :  $m=0.88$ ,  $\sigma=0.05$ ) indiquent néanmoins un degré moindre d'hostilité lorsque les stimuli sont présentés sous leur forme retranscrite seulement.

Figure 1. Corrélation entre scores d'hostilité perçue, recueillis dans le test de perception à l'oral (abscisse) et à l'écrit (ordonnée) pour les stimuli de chaque locuteur



- 24 Bien que la différence entre scores moyens soit significative entre écrit et audio, à l'échelle du locuteur on ne trouve de différence entre scores par condition que pour Le Chequer (HSD de Tukey,  $p < .05$ ) et Serey (HSD de Tukey,  $p < .01$ ). Ces deux locuteurs sont également ceux pour qui l'on trouve les corrélations les plus significatives entre scores écrit et scores audio (Le Chequer :  $\text{cor} = .703$  ; SEREY :  $\text{cor} = .743$ ), voir Figure 1 ci-dessus.
- 25 Ceci laisse à penser qu'une part non négligeable de l'hostilité véhiculée dans leurs stimuli passerait par leur parole. Les stimuli de Serey sont ceux pour lesquels on observe la différence la plus notable entre scores écrits et scores audio, indiquant ici que ses stimuli sont majoritairement perçus comme plus hostiles à l'oral qu'à l'écrit. L'analyse post-hoc (HSD de Tukey) permet d'observer qu'à l'écrit, les extraits de Serey sont d'ailleurs perçus comme significativement moins hostiles que ceux de Voynet ( $p = .0372$ ). Dans les résultats du test écrit, ce sont effectivement les stimuli de Voynet qui obtiennent le score le plus haut d'hostilité perçue (voir Figure 2 ci-dessous) avec 30 % de stimuli perçus en moyenne comme « fortement hostiles ». Au contraire, Serey obtient le plus haut taux de stimuli perçus comme n'exprimant aucune hostilité à l'écrit (40 %).

Exemple 1 - Voynet : « Et c'est ces gens là qui nous donnent des leçons sur comment s'occuper de la ville et des Murs à pêches ? »

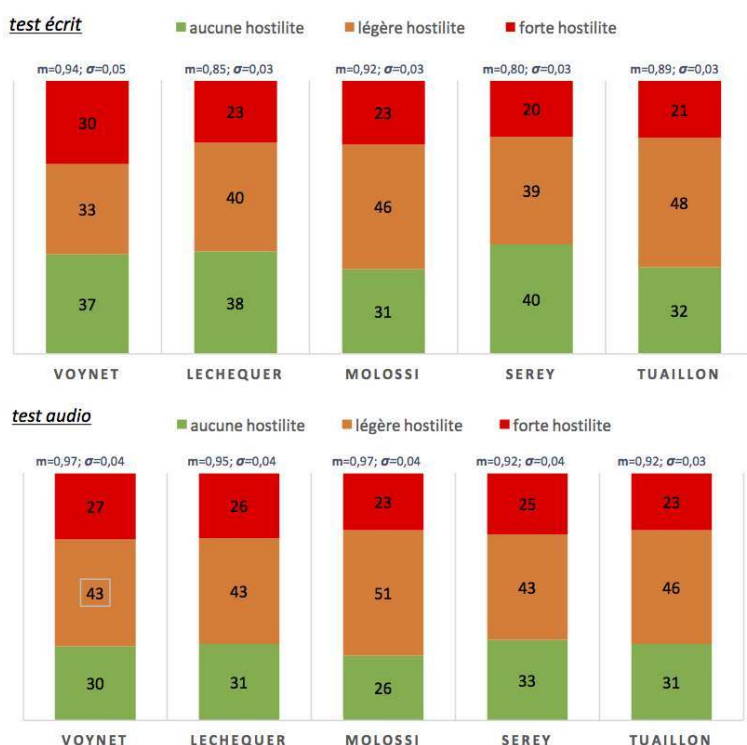
score moyen d'hostilité perçue à l'écrit :  $1.89/2$  ; oral :  $1.85/2$



Exemple 2 – Serey : « 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, (...) »  
score moyen d'hostilité perçue à l'écrit : 0/2 ; oral : 1.46/2

- 26 Les différences significatives entre les scores d'hostilité attribués à l'ensemble des stimuli de VOYNET et SEREY ne s'observent que lorsque ceux-ci sont évalués à partir de leur transcription orthographique seule (HSD de Tukey,  $p = .03$ ). Dans le test d'écoute, les résultats de l'ANOVA [ $F(2,4) = 0.4, p = 0.7$ ] n'indiquent pas de différence significative interlocuteurs, tous étant perçus comme « légèrement hostiles ».

Figure 2. Taux d'occurrences (pourcentages) par catégorie d'évaluation de l'hostilité perçue à l'écrit et à l'oral



Le score moyen (m) et l'écart-type(σ) sont également donnés pour chaque locuteur. Pour rappel, la catégorie « aucune hostilité » équivaut à un score de 0, « légère hostilité » 1, et « forte hostilité » 2.

## Analyses de la perception du destinataire présumé du message

- 27 On observe une forte corrélation positive entre les scores obtenus dans l'évaluation audio et ceux de l'évaluation de l'écrit ( $cor = .789$ ). Plus les stimuli sont perçus comme s'adressant à un interlocuteur particulier à l'oral, plus ils le sont à l'écrit (voir Figure 3).

Figure 3. Corrélations entre scores moyens d'estimation du destinataire pour écrit et audio

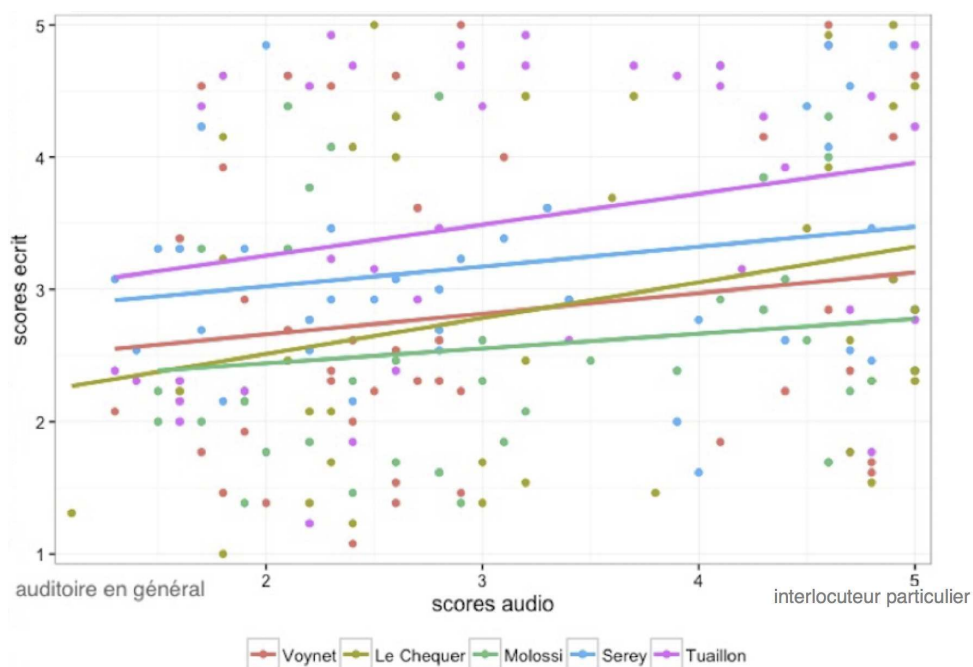


Figure 4. Taux d'occurrences (pourcentages) par catégorie proposée d'annotation perceptive du destinataire présumé du message de chaque stimuli dans le test audio (en haut) et le test écrit (en bas)

**test audio** A qui semble s'adresser le locuteur?

■ 1 : à l'auditoire ■ 2 ■ 3 ■ 4 ■ 5 : à un interlocuteur particulier

Tuillon	21	7	11	11	50	
Serey	30		11	19	13	27
Molossi	31		11	20	14	24
LeChequer	21	13	16	12	38	
Voynet	32	10	16	10	31	

**test écrit** A qui semble s'adresser le locuteur ?

■ 1 : à l'auditoire ■ 2 ■ 3 ■ 4 ■ 5 : à un interlocuteur particulier

Tuillon	15	15	13	14	43	
Serey	21	15	18	16	30	
Molossi	33		24	13	15	15
LeChequer	32	18	11	12	28	
Voynet	33	16	13	13	25	

28 La comparaison avec les scores obtenus, lorsque les stimuli sont présentés sous leur forme écrite seulement, nous indique que les scores moyens sont légèrement mais significativement plus bas, donc réfèrent à moins d'extraits perçus comme s'adressant un

interlocuteur particulier. Tuailleon obtient comme à l'oral le score le plus haut, par contre Le Chequer obtient, lui, des scores plus bas, donc cela laisse à penser qu'une partie des informations permettant de savoir s'il s'adresse ou non à quelqu'un de particulier serait portée plus par les indices paralinguistiques que par le contenu sémantique.

- 29 Dans la tendance opposée, bien qu'en général les scores qu'obtiennent ses extraits ne soient pas significativement plus bas que ceux des autres locuteurs, Voynet obtient le plus haut taux d'occurrences perçues en moyenne comme s'adressant à l'auditoire en général avec 32 %. Enfin, notons que Molossi et Serey, semblant s'adresser plus généralement à l'auditoire sont ceux qui reçoivent les plus hauts taux de stimuli pour lesquels l'auditeur n'arrive pas à déterminer à qui ils s'adressent (voir taux d'occurrences de la valeur 3 dans la Figure 3).

## Corrélations entre hostilité perçue et estimation du destinataire à l'oral et à l'écrit

- 30 A 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 un interlocuteur particulier (oral :  $r = .621$  ; écrit :  $r = .619$ ) ; voir Figures 5 et 6 ci-dessous. Sans surprise, 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 5. Corrélations entre scores d'hostilité perçue et scores d'estimation du destinataire (test audio)

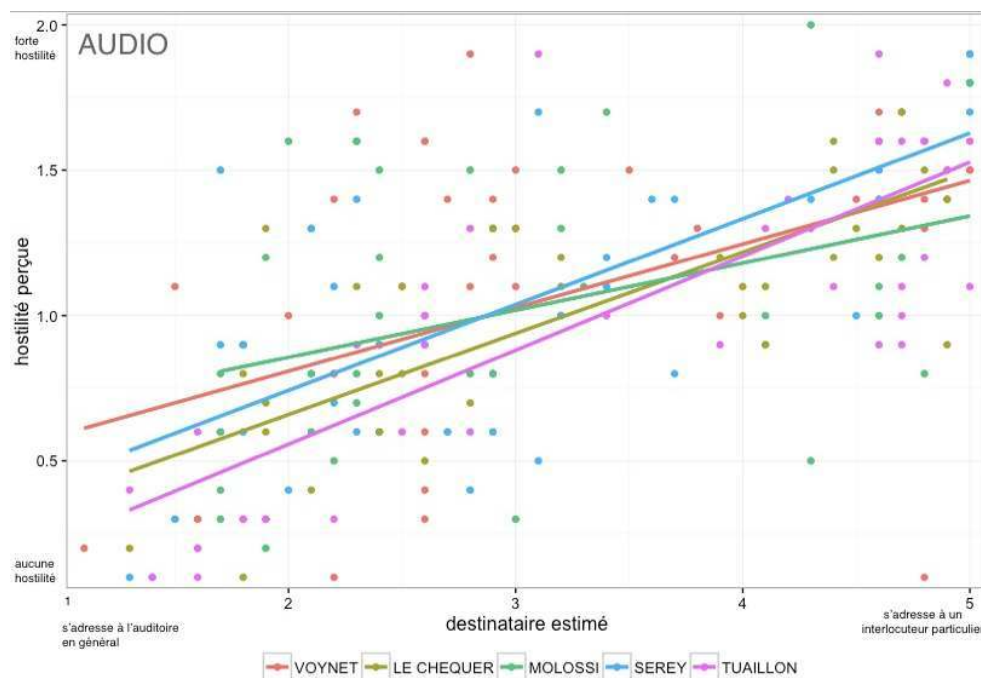
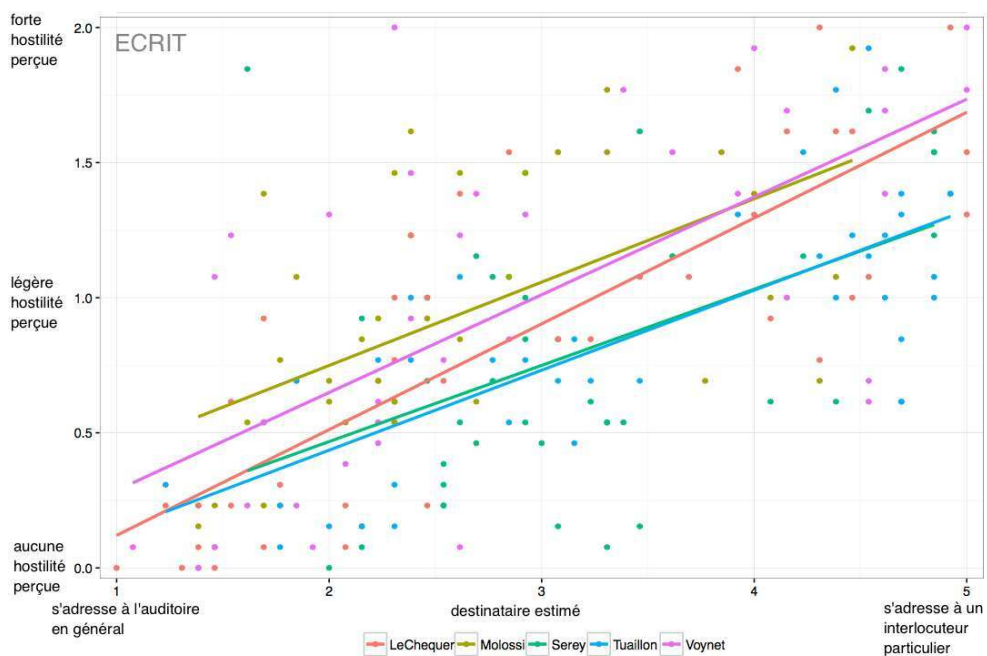


Figure 6. Corrélations entre scores d'hostilité perçue et d'estimation du destinataire (test écrit)



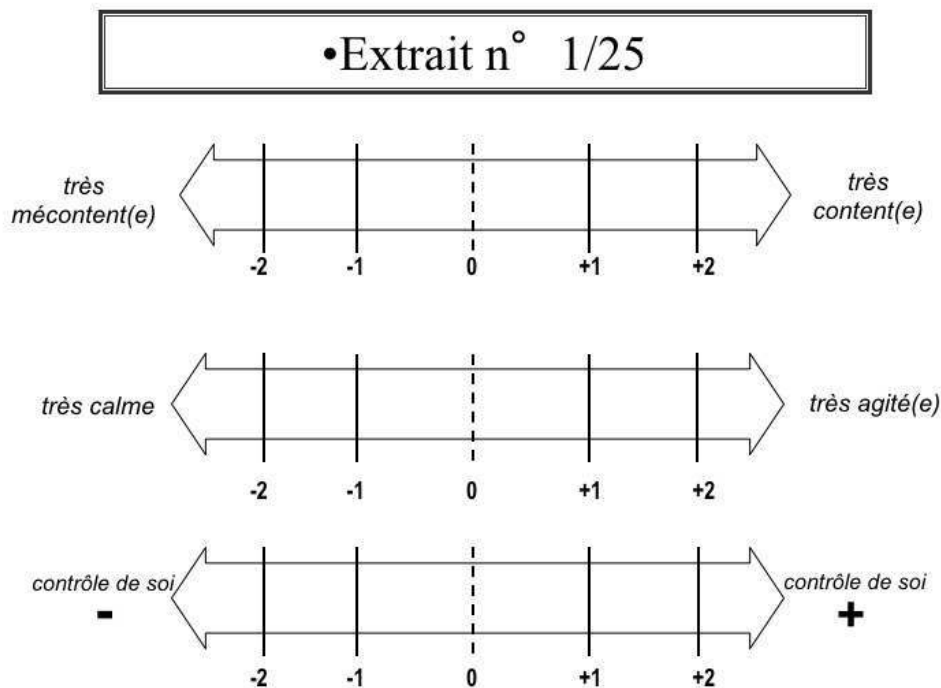
## Analyses perceptives partie 2 : Les profils émotionnels

- 31 Outre la description des émotions par des catégories fermées, classées par émotions brutes et leurs dérivées plus nuancées (Darwin, 1965 ; Plutchik, 1980 ; Ekman, 1992), celles-ci peuvent être décrites par les dimensions de valence, excitation (*arousal*) et dominance/pouvoir (Osgood, 1966 ; Banse & Scherer, 1996). La valence représente la façon dont un individu évalue un événement donné et décrit sa réaction d'une forme négative (aversion, mécontentement) à positive (attirance, contentement). La dimension d'excitation englobe différents aspects relatifs au degré d'activation émotionnelle, et indique le degré d'excitation, allant de l'apathie à l'agitation totale. Dans les modèles tridimensionnels, de plus en plus défendus dans la littérature (Russel & Mehrabian, 1977 ; Fontaine, Scherer, Roesch & Ellsworth, 2007), la troisième dimension correspond à celle de la dominance/pouvoir par rapport à autrui dans une situation donnée, proche de la dimension de contrôle de la situation. Cette troisième dimension est également corrélée avec la dimension de contrôle de soi (Banse & Scherer 1996 ; Pereira, 2000).
- 32 L'analyse dimensionnelle permet d'obtenir un profil émotionnel pour chaque stimulus soumis à évaluation perceptive. De nombreuses études ont été menées tant en psychologie qu'en phonétique sur les caractéristiques expressives propres aux descriptions dimensionnelles des émotions. Une telle approche permet également de décrire de façon objectivable des différences entre des types d'affects appartenant à un même sous-ensemble ; par exemple l'irritation se différencie d'une forme plus expansive comme la rage, par un moindre degré d'activation émotionnelle (Banse & Scherer, 1996).

## Protocole expérimental et stimuli

- 33 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. La passation du test d'écoute s'est déroulée de manière vive et harmonieuse avec tous les locuteurs et ceux-ci ont spontanément tous déclaré se rappeler la situation de communication correspondante, pour la grande majorité des extraits entendus. 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 7. Grille dimensionnelle telle qu'elle a été présentée aux politiciens durant la tâche d'auto-évaluation de leurs stimuli



- 34 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/très content), d'excitation (très calme/très agité), et de contrôle de soi. Pour la dimension de contrôle de soi, il leur a été précisé qu'un score de -2 correspondait à 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.
- 35 En parallèle à cela, 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. Afin d'évaluer le risque potentiel d'

*ethos* préalable (voir Introduction), il était notamment demandé aux participants de préciser en fin de test s'ils pensaient avoir reconnu des locuteurs. Les auditeurs, étudiants en première et deuxième année de Licence de Sciences du Langage, ont participé à l'étude sur la base du volontariat. Tous ont précisé qu'ils n'avaient reconnu aucun locuteur. Sur l'interface du site en ligne, les participants dits « naïfs » ont eu accès à une échelle à 7 points, les résultats obtenus ont donc 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.

## Résultats de l'analyse dimensionnelle (auto-évaluation et auditeurs « naïfs »)

- 36 Dans les résultats de l'évaluation par les auditeurs « naïfs » tous les locuteurs sont perçus comme plutôt mécontents, et plutôt dominants. Ces résultats sont à l'image de situation de communication dont les extraits proviennent, un débat politique tendu et houleux, peu propice aux réactions positives. De plus, le faible taux de scores extrêmes dans toutes les dimensions (voir Tableau 1) s'explique aisément du fait qu'il s'agit ici d'une situation de communication relativement formelle, filmée, et de ce fait requérant plus de contrôle des productions que lors d'un échange privé.

Tableau 1. Moyennes des scores obtenus par locuteur pour chaque dimension dans l'auto-évaluation et l'évaluation par 24 auditeurs « naïfs » des 125 stimuli

	valence		excitation		Contrôle/ dominance	
	auto-eval.	« naïfs »	auto-eval.	« naïfs »	auto-eval.	« naïfs »
Voynet	1.08	-0.70	0.44	0.12	0.92	0.59
Le Chequer	-0.08	-0.66	-0.48	-0.26	0.28	0.55
Molossi	-0.84	-0.84	-0.12	0.09	-0.2	0.62
Serey	-0.67	-0.68	0.42	0.04	1.08	0.46
Tuailon	-0.36	-0.44	0.64	0.02	0.76	0.58

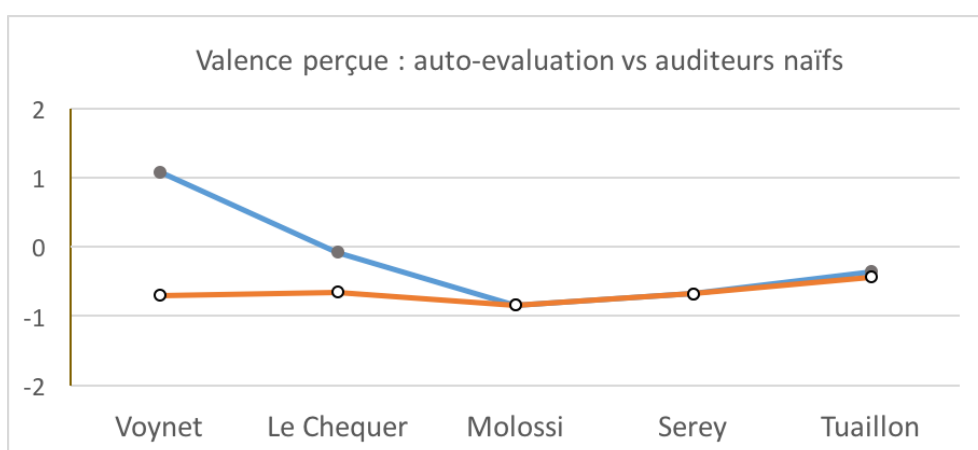
Tableau 2. Résultats des ANOVAS à 2 facteurs pour chaque dimension (modalité\*locuteurs)

	Valence	Excitation	Contrôle-Dominance
ANOVAmodalité (auto-eval/naïfs)	F(1,3) =26.8 p<.0001 $\eta^2$ =10 %	F(1,3) =2.24 p =0.13, $\eta^2$ <.01 %	F(1,3) =0.004 p =0.94, $\eta^2$ <.01 %

ANOVA locuteurs	F(4,3) =13.4 p<.0005, $\eta^2$ =18 %	F(4,3) =4.88 p<.001, $\eta^2$ =7 %	F(4,3) =5.15 p<.0005, $\eta^2$ =8 %
ANOVA modalite*locuteurs	F(4,3) =12.6 p<.0001, $\eta^2$ =17 %	F(4,3) =2.06 p =0.05, $\eta^2$ =3 %	F(4,3) =4.2 p<.0001, $\eta^2$ =11 %

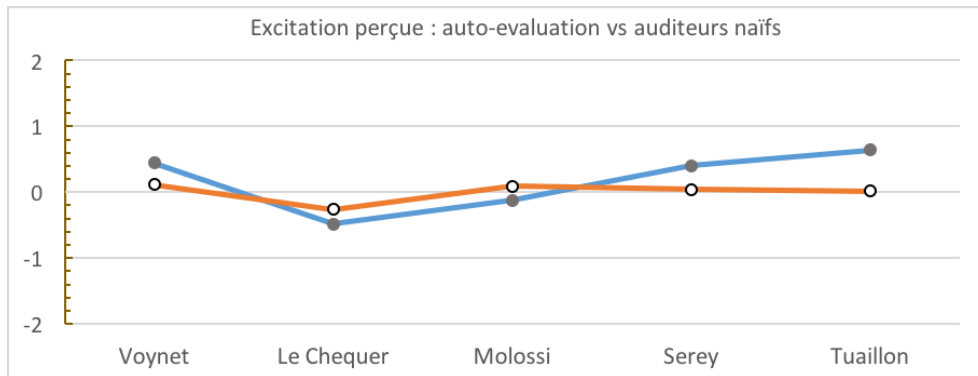
- 37 La valence est la seule dimension pour laquelle une différence significative est attestée entre scores de l'auto-évaluation et scores attribués par les participants naïfs. Une influence significative de l'interaction entre les facteurs modalité/locuteur a toutefois été observée dans toutes les dimensions ( $p < .05$ ) voir Tableau 2.

Figure 8. Scores moyens de valence perçue, comparés entre auto-évaluation par chaque locuteur (courbe bleue) et évaluation par 24 auditeurs (courbe orange)



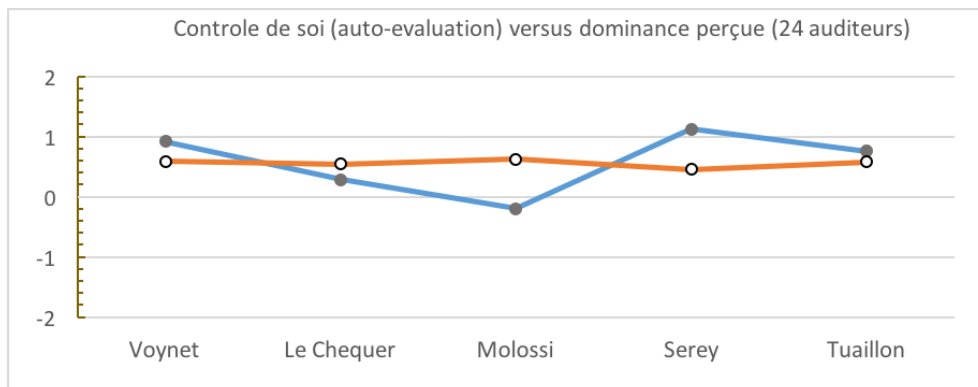
- 38 Dominique Voynet est la seule locutrice obtenant un score moyen de valence positif (76 % de scores « 1 » équivalent à « plutôt content-e ») dans l'auto-évaluation de ses stimuli. Un tel score est significativement plus élevé que celui obtenu par les 4 locuteurs de l'opposition (HSD de Tukey,  $p < .00001$ ) et elle est de plus la seule dont les scores diffèrent significativement entre son auto-évaluation et l'évaluation par les locuteurs naïfs (HSD de Tukey,  $p < .00001$ ).
- 39 Sur l'ensemble de ses stimuli, Voynet ne s'évalue jamais comme mécontente. Ses stimuli pourront être auto-évalués comme neutres, ou signifiant que la locutrice est « plutôt contente » (case sélectionnée dans 76 % des cas) ou « très contente » mais jamais « plutôt mécontente » ou « très mécontente ». Dans l'opposition, 3 opposants sur 4 se déclarent plutôt mécontents et cochent toutes les cases négatives au moins une fois ; Le Chequer ne se déclare jamais « très mécontent », tout au plus parfois « plutôt mécontent ». Par contre, si, dans leur auto-évaluation, Voynet et Le Chequer se sont attribués des scores de valence positifs ou neutres, ils ont toutefois été perçus comme légèrement mécontents par les auditeurs naïfs. L'analyse ANOVA des scores de l'évaluation par les auditeurs naïfs indiquent que tous les locuteurs sont perçus comme légèrement mécontents.

Figure 9. Scores moyens d'excitation perçue, comparés entre auto-évaluation (trait bleu) et évaluation par 24 auditeurs pour chaque locuteur



- 40 Aucune différence significative inter-locuteurs n'est observée dans l'évaluation par les auditeurs. Tous les locuteurs obtiennent des scores proches de zéro, bien que Voynet, Serey et Tuailleon se déclarent tous trois « plutôt tendus » dans plus de 52 % des cas. Le profil d'excitation attribué par les auditeurs naïfs est également le même pour Le Chequer que pour les autres locuteurs, bien que celui-ci se soit déclaré « plutôt pas agité » dans 51 % des cas et dont le score moyen d'excitation auto-évaluée était inférieur à tous les autres locuteurs (HSD de Tukey,  $p < .00001$  pour toutes les comparaisons).

Figure 10. Scores moyens de contrôle de soi (auto-évaluation ; trait bleu) et dominance perçue par 24 auditeurs (trait orange) pour chaque locuteur



- 41 La figure 10 montre que, malgré la corrélation attendue, un taux de contrôle de soi élevé n'entraîne pas forcément un taux aussi élevé de dominance perçue. Dans l'auto-évaluation, 3 locuteurs sur 5 estiment en moyenne fournir un léger effort de contrôle de soi, et sont perçus comme légèrement dominants par les auditeurs naïfs. Le Chequer et Molossi sont perçus par les auditeurs comme légèrement dominants et ce, malgré le fait que ce dernier s'attribue un score zéro de contrôle de soi dans 41 % des cas. Le taux homogène de contrôle de soi et de dominance perçue est cohérent avec la volonté d'apparaître calme et dominant pour s'assurer l'ascendant dans un débat (Poggi, D'Errico & Vincze, 2011).



## Conclusion

- 42 L'évaluation parallèle des transcriptions orthographiques isolées et des extraits audio originaux a permis d'obtenir des informations sur les poids respectifs du contenu sémantique et des indices paralinguistiques dans l'expression de l'hostilité et l'indication du destinataire du message. L'analyse des scores d'hostilité perçue à l'écrit a montré que le contenu sémantique portait déjà à lui seul des indices d'hostilité dans certains stimuli.
- 43 L'analyse de la perception du destinataire estimé du message, quant à elle, nous a permis d'obtenir quelques informations premières sur l'orientation de l'activité discursive dont sont extraits chacun de nos stimuli. Charaudeau (2015b) distingue différents types d'activités discursives de la joute oratoire que constitue l'échange dans un débat politique allant de celles orientées vers l'adversaire, où « l'activité consiste à tenter de le disqualifier au regard de ce que pourraient être sa légitimité et sa crédibilité » (Charaudeau, 2015b, p. 111), ce qui n'est pas sans rappeler les actes discréditant la dimension de compétence dans l'*ethos*, à celles orientées vers l'auditoire en général, au sein desquelles l'orateur peut par exemple se positionner comme porte-parole de certaines catégories sociales.
- 44 Enfin, concernant l'évaluation dimensionnelle, l'ensemble des productions a été évalué par les auditeurs « naïfs » comme exprimant une valence légèrement négative (« plutôt mécontent »), un degré neutre d'excitation (ni tendu ni calme) et un degré de dominance légèrement positif. Cette homogénéité ne rend pas compte d'indications particulières fournies par les locuteurs eux-mêmes : les locuteurs s'étant déclarés plus calmes ou moins mécontents que les autres n'ont pas été perçus différemment des autres dans l'évaluation par les auditeurs « naïfs ». De telles informations confirment un décalage entre *ethos* voulu et *ethos* perçu.

## Discussion

### Vers l'analyse acoustique de la parole expressive en politique

- 45 Analyser la parole expressive équivaut, d'une certaine façon, à quantifier l'écart de ce type de production par rapport à une norme, la parole « non expressive », non marquée (Fónagy, 1983 ; Léon, 2005). De ce fait, l'analyse de la parole expressive en politique exigera un ensemble de connaissances préalables sur la parole politique, sur ce qui fait le style « parole politique » en lui-même. Certaines études phonétiques ont d'ores et déjà permis d'en connaître ses caractéristiques principales sur des paramètres tels que le rythme syllabique et les mouvements mélodiques (Lucci, 1980 ; Fónagy, 1983 ; Bruce & Touati, 1992 ; Léon, 2005 ; Martin, 2009 ; Guaitella, 2014) ou bien les pauses comme marqueurs de pouvoir (Duez, 1991 ; Bechet *et al.*, 2013). On pourra se demander ce qu'il advient des pauses silencieuses, connues pour être marqueurs de pouvoir selon Duez (1991), dans l'interaction conflictuelle. On pensera notamment à la récente étude de Béchet *et al.* (2013), décrivant les pauses silencieuses dans la parole de François Hollande au cours de son débat en face-à-face avec Nicolas Sarkozy en 2012. Cette étude, jouissant d'une collaboration entre phonéticiens et analystes du discours, fait état d'un nombre conséquent de pauses démarcatives alliées à un débit de parole relativement lent.

- 46 L'analyse acoustique des caractéristiques acoustiques des extraits du corpus se fera donc d'une part en comparaison aux caractéristiques généralement attendues en parole politique, et d'autre part au regard des informations perceptives désormais disponibles. L'ensemble des tests de perception menés et analysés a permis d'obtenir une évaluation perceptible quantifiable du degré d'hostilité, de mécontentement, de tension, de contrôle de soi, et de dominance. Les informations perceptives recueillies dans notre étude vont pouvoir servir de repères conséquents dans l'analyse des caractéristiques de parole des 5 locuteurs du corpus.
- 47 Les paramètres acoustiques étudiés (F0, intensité relative intra-énoncé, durée et rythme syllabiques, durée et type de pauses) seront analysés au regard de leur corrélation avec les scores obtenus au sein des différents tests perceptifs et les profils émotionnels obtenus *via* l'approche dimensionnelle (valence, excitation, contrôle/dominance). Il est par exemple généralement rapporté qu'une forte tension et un effort articulatoire et respiratoire résultant d'un haut degré d'excitation (joie ou colère vive) se traduit par un accroissement de la moyenne et de la plage de fréquence fondamentale de la voix (F0) et un débit de parole élevé (nombre de syllabes par secondes). Au contraire, un faible degré d'excitation corrélé à une valence négative, comme c'est le cas pour des expressions de déprime et d'abattement, donne lieu à une F0 plus basse que la moyenne et une plage de variation plus réduite (Banse & Scherer, 1986 ; Fónagy, 1983).

### De la nécessité d'une approche pluridisciplinaire

- 48 Outre l'impact psychologique de la tension inhérente à une telle situation de communication, conflictuelle et publique, les interventions au sein d'un débat sont également régies par une nécessité de construction et de conservation d'une image de soi (*ethos*) favorable, en véhiculant des signaux de dominance, de compétence, de bienveillance ; ou en s'attaquant à l'une de ces trois dimensions de l'*ethos* de l'opposant dans le débat. Les théories et observations de Poggi *et al.* (2011), au sujet des actes cherchant à discréditer l'image de l'autre, indiquent des stratégies expressives différentes selon la dimension visée. Selon leurs observations, si l'on cherche à faire apparaître son opposant comme malveillant, il est fréquent voire profitable d'appuyer son propos (message linguistique) d'indices prosodiques et/ou gestuels signalant une forte activation émotionnelle, une forte implication personnelle. Au contraire, lorsque l'on s'attaque à la dominance de son opposant, trop habiller son propos de gestes entraînerait le risque d'un effet « boomerang », le locuteur pouvant alors passer pour intraitable, dur et sans pitié. S'il apparaît comme juste dans l'imaginaire collectif de s'indigner de la malhonnêteté, la malveillance ou l'incompétence, il est un peu plus questionnable, et donc un peu moins attendu, de mettre en cause la dominance d'un acteur politique en affichant une image de lui impuissant ou inconséquent dans le débat.
- 49 Au vu de telles observations, une annotation par types d'attaques à l'*ethos* de l'opposant se justifierait par le fait qu'elle pourrait nous apporter des informations complémentaires pour appréhender l'analyse acoustique d'extraits de parole déjà marqués au niveau de leur contenu sémantique et/ou d'extraits de parole aux caractéristiques expressives relativement discrètes. Notons à ce propos que si l'usage premier d'analyses perceptives comparatives oral/écrit pour le phonéticien consiste à pondérer l'effet perceptif potentiel du contenu sémantique, lorsque l'on analyse de la parole naturelle (produite hors laboratoire), il est impossible pour autant de s'abstraire de celui-ci pour effectuer des

analyses acoustiques. Comme le montrent les données de la présente étude, l'hostilité exprimée par le contenu sémantique et par les indices paralinguistiques sont très difficilement dissociables en situation de communication. L'analyste du discours pourrait, quant à lui, utiliser ces informations, non plus comme des obstacles mais comme des indices supplémentaires pour aller plus loin sur l'impact perceptif de certains éléments propres au contenu sémantique (choix lexicaux, structuraux) et les réintégrer à une analyse par tours de parole.

- 50 En somme, la constitution d'une équipe pluridisciplinaire semble indispensable pour analyser finement la complexité de ce qui fait et sous-tend les rapports de force dans le débat politique. Charaudeau (2015b) insiste d'ailleurs à ce sujet sur la nécessité de créer des passerelles pour faire communiquer les disciplines malgré leurs enjeux et terminologies différents. À notre sens, la pertinence de recherches sur l'expressivité de la parole politique ne peut se trouver que dans la pluridisciplinarité, et notamment dans le dialogue entre analyse du discours, sciences du langage, et psychologie sociale.

---

## BIBLIOGRAPHIE

- Ballet, M. (2012) *Peur, espoir, compassion, indignation : l'appel aux émotions dans les campagnes présidentielles (1981-2007)*, Paris : Dalloz.
- Banse, R., & K. R. Scherer (1996) Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70, 3, p. 614-636.
- Blanche-Benveniste, C. (2010) *Approches de la langue parlée en français*, Gap/Paris : Ophrys.
- Bruce, G., & P. Touati (1992) On the analysis of prosody in spontaneous speech with exemplification from Swedish and French, *Speech Communication*, 1, 4, p. 453-458.
- Béchet, M., Sandré, M., Hirsch, F., Richard, A., Marsac, F., & R. Sock (2014) De l'utilisation de la pause silencieuse dans le débat politique télévisé. Le cas de François Hollande. *Mots. Les langages du politique*, 103, p. 23-38.
- Charaudeau, P. (2015a) Le charisme comme condition du leadership politique. *Revue française des sciences de l'information et de la communication*, 7.
- Charaudeau, P. (2015b). Le débat présidentiel. Un combat de mots. Une victoire aux points. *Langage et société*, 151, p. 109-129.
- Darwin, C. (1872/1965) *The Expression of the Emotions in Man and Animals*. Chicago/London : University of Chicago Press.
- Duez, D. (1991) *La pause dans la parole de l'homme politique*. Paris : Éd. du Centre national de la recherche scientifique.
- Ekman, P. (1992) An argument for basic emotions. *Cognition and Emotion*, 6, 3-4, p. 169-200.
- Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & P. C. Ellsworth (2007b) The World of Emotions is not Two-Dimensional. *Psychological Science*, 18, 12, p. 1050-1057.
- Fónagy, I. (1983) *La vive voix : essais de psycho-phonétique*. Paris : Payot.

- Gendrot, C. (2004) Influence de la qualité de la voix sur la perception de quatre émotions simulées : une étude perceptive et physiologique, *Revue PARole*, 1, 13, p. 1-18.
- Jakobson, R. (1963/81) *Essais de linguistique générale : Les fondations du langage*, Paris : Minuit.
- Lucci, V. (1980) L'accent didactique, *Studia Phonetica*, Montréal, 15, p. 107-121.
- Léon, P. (2005) *Précis de phonostylistique : parole et expressivité*, Paris : A. Colin.
- Maingueneau, D. (1999) Ethos, scénographie, incorporation, in Ruth A. (ed.), *Images de soi dans le discours. La Construction de l'éthos*, Lausanne : Delachaux et Niestlé, p. 75-102.
- Morange, S., & M. Candea (2010) Aux frontières de l'écoute. Durée des échantillons et choix des auditeurs : deux variables déterminantes dans la construction des tests de perception. In D. Delomier, M.-A. Morel (eds.) *Frontières, du linguistique au sémiotique*, Limoges : Lambert-Lucas, p. 79-96.
- Plutchik, R., & H. Kellerman (2013) *Theories of Emotion*. Academic Press.
- Poggi, I., & F. D'Errico (2010) Dominance Signals in Debates. In Salah A. A. et al. (eds.), *Human Behavior Understanding*, Berlin/Heidelberg: Springer, p. 163-174.
- Poggi, I., D'Errico, F., Vincze, L. (2011) Discrediting moves in political debates. F. Ricci et al. (eds.), *Proceedings of Second International Workshop on User Models for Motivational Systems: the Affective and the Rational Routes to Persuasion (UMMS 2011)*, Berlin/Heidelberg: Springer, p. 84-99.
- Russell, J. A., & Mehrabian, A. (1977) Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11, 3, p. 273-294.

## NOTES

1. Acoustic analysis not presented in this study.
2. Les séances du conseil municipal de Montreuil sont retransmises en direct sur une chaîne locale (TVM Est Parisien) et les archives libres sont ensuite disponibles sur le site internet de la municipalité.
3. 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.
4. En 2013 à Montreuil, tous les partis minoritaires sont de gauche, il n'y a alors aucun parti de droite.
5. Aristote, 1967, *Rhétorique*, Paris, Les Belles Lettres, trad. M. Dufour.
6. Analyse non présentée dans cet article.
7. L'expressivité n'est pas étudiée au niveau gestuel dans cette recherche. Au regard des analyses possibles entre gestes et prosodie (Guaïtella, 2014), un effort particulier a été fourni pour maintenir possible l'alignement des annotations effectuées avec les pistes vidéos, rendant disponible ce corpus à d'éventuelles recherches multimodales.

---

## RÉSUMÉS

La présente étude porte sur l'expression de l'hostilité au sein du débat politique et de ses rapports de force et propose un ensemble d'approches méthodologiques adaptées à l'étude de la composante vocale de l'expressivité en situation de communication conflictuelle.

Un corpus a été constitué à partir d'extraits de parole de la Maire de Montreuil (D. Voynet) et de 4 de ses opposants durant le conseil municipal en 2013. Les données ont été analysées via la combinaison d'entretiens individuels et de tests de perception. Une première évaluation (20 auditeurs) a permis de pondérer l'influence du contenu sémantique (transcriptions orthographiques) et des indices paralinguistiques (extraits audio) dans la perception de l'hostilité et dans l'estimation du destinataire du message. Durant les entretiens, les 5 locuteurs ont évalué leurs propres stimuli sur trois dimensions (content/mécontent ; calme/agité ; contrôle(-/+)). Cette auto-évaluation a été comparée à une évaluation similaire par 22 auditeurs naïfs, qui ont perçu tous les locuteurs comme plutôt mécontents et dans le contrôle. La légère agitation ressentie par l'ensemble des locuteurs n'a pas été perçue par les auditeurs naïfs, de même que D. Voynet, s'étant déclarée "plutôt contente" a été perçue comme aussi mécontente que les autres locuteurs.

Usually, politicians have to give the image of a moral, benevolent and competent individual in order to be convincing and increase their chance to be elected. For contemporary psychosociologists (Poggi & D'Errico, 2010), politicians should also appear to be dominant, especially in political debates, where it becomes an essential quality. Dominance signals may include a strong displaying of self-control and more or less tacit aggressive expressions such as contempt, sarcasm, irritation, and indignation. Questions are raised about the role of vocal expressiveness in such challenges and situations. This study shows a comprehensive set of recommendations to address the methodological issues facing the phonetic analysis of political speech, beginning with perceptive evaluations.

Yet voice occupies an important part of the expressive function of language (Jakobson, 1963/81), phonetic studies on political speech are, in our view, still less done compared to lexical, pragmatic and even nonverbal studies. Such a gap may be explained by the fact that expressive natural speech, produced within a particular context, raises problems in experimental protocol of the study. Despite methodological difficulties, perceptual evaluation is an essential prerequisite step in the definition and validation of the data that will be submitted to acoustical analysis<sup>1</sup>. In addition, perceptual evaluation is particularly important in the study of political speech as intended *ethos* does not equal perceived *ethos* (Amossi, 1999). It is also imperative to take into account that in a natural setting, expressivity conveyed by the linguistic message (i.e. semantic content) and by paralinguistic cues most of the time can hardly be dissociated. A possible and easily conducted additional precaution could therefore be to conduct a "stand-alone" perception test of the orthographic transcriptions of the original sound stimuli which are analyzed in the auditory perception tests.

This study focuses on political debates extracted from video archives of Montreuil's city council sessions in 2013, during the last months of the mayor's term, Dominique Voynet (Europe Ecologie Les Verts). At that time, and since a couple years already, sessions were marked by tensions of politically charged climate of the city. Five speakers were selected and extract of their speeches

were submitted to perceptive evaluation (naïve speakers) and individual interviews: the mayor and four members of the opposition parties (Gaylord Le Chequer, Frédéric Molossi, Jean-Jacques Serey, Alexandre Tuillon).

A first perceptual evaluation campaign was launched on 200 stimuli (40 audio fragments x 5 speakers) concerning the perceived hostility (from none to strong) and the estimated addressee (from the audience to a specific interlocutor). Two perception tests were conducted simultaneously with 11 French participants for each: one regarding the original spoken data (audio stimuli condition), and the other regarding the orthographic transcriptions of the stimuli only (written form condition). The results of this first evaluation show that in both conditions, speakers' extracts were widely perceived as expressive a light hostility. The scores obtained in the written form condition show that semantic content is already conveying the expression of hostility by itself.

This first evaluation enabled the creation of a balanced subset of 125 stimuli (25 audio stimuli for 5 speakers) expressing variable degrees of hostility. This smaller corpus served as working material for individual interview conducted with each of the speakers. Interviews included particularly a self-evaluation task of their own stimuli. Speakers were given the task to describe their own stimuli through three 5-points Likert scales representing the dimensions of Valence (very dissatisfied/very satisfied), Arousal (very calm/very agitated), and Control (none/absolute). The information given by the politicians were then compared to a similar perceptual evaluation conducted online with 24 native French listeners (16 women, 8 men, average age= 22 years old). In the grid presented to the naïve listeners, the control dimension has been replaced by dominance (does not dominate the situation at all / absolutely dominates the situation). Such a difference is justified by the fact that it was felt appropriate to question politicians on their degree of control over their emotions in parallel and at the same time to gain information on how they were perceived by naïve listeners in terms of control of the situation.

The results of the naïve dimensional evaluation show a rather homogeneous profile for all five speakers' extracts as they were all perceived as expressing light dissatisfaction, a neutral degree of arousal (neither calm nor agitated) and a light form of dominance. Comparison with the self-evaluation illustrates some discrepancies between intended and perceived *ethos*: even though Dominique Voinet was the only speaker declaring being mostly expressing a light form of satisfaction, she was perceived as much unsatisfied as her opponents. Also, the light agitation related by the large majority of the speakers was perceived by the listeners in the Arousal evaluation, offering some evidence of the effectiveness of the politicians' self-control.

As a whole, the combination of the different perception tests conducted and analyzed enabled us to obtain a quantifiable description of the degree of hostility, dissatisfaction, agitation, control and dominance perceived. The collected information will serve as a guide for further acoustic analysis of the data. The acoustic parameters (F0, intensity, spectral configuration, duration, syllabic rhythm, pause duration) should be analyzed in the light of the perceptual scores and dimensional profiles obtained.

Constituting a multidisciplinary team seems to be quite essential for obtaining the most fine-grained analysis possible of what underlies the power relations in an interaction such as a political debate. Indeed, as argued by Charaudeau (2005), to approach this issue, it is much necessary to work on creating a gateway between disciplines despite their different terminologies and theories. In our view, the relevance of further research on vocal expressivity in political debates can only be found in multidisciplinary and by supporting a dialogue between different disciplines such as discourse analysis, speech sciences and social psychology.

## INDEX

**Mots-clés** : hostilité, débat politique, parole expressive, dominance, maîtrise des émotions

**Keywords** : hostility, political debates, expressive speech, dominance, emotion regulation

## AUTEUR

**CHARLOTTE KOUKLIA**

[charlotte.kouklia@univ-paris3.fr](mailto:charlotte.kouklia@univ-paris3.fr)

Laboratoire de Phonétique de Phonologie (LPP) CNRS/UMR7018, Université Sorbonne Nouvelle