Une analyse prosodique de la parole souriante : étude préliminaire

Caroline Émond

Université du Québec à Montréal
Département de linguistique et de didactique des langues, C. P. 8888, Succ. Centre-ville, Montréal (Québec) H3C 3P8
caroemond@hotmail.com
www.phonetique.uqam.ca

ABSTRACT

Smile is a visible expression and it has been shown that it's audible too (Tartter [9], Schröder & al. [8], Aubergé & Cathiard [1]). The aim of this study is to describe the prosodic correlates of smiled speech produced by 6 speakers of Quebec French. In order to elicitate smile, subjects were required to read sentences containing caricatures. This condition was compared to a neutral condition, in which no caricatures were present. Utterances were digitized and analysed with Praat and F0 measurements were extracted. Results show that F0 register tends to be lower in normal speech than in smiled speech. Because the global shapes of F0 movements are similar in the two kinds of speech, we claim that this difference is phonetic rather than phonologic.

1. INTRODUCTION

Le rire, comportement humain remarquable, est présent dans toutes les cultures (Trouvain [12]) et il nous permet d'exprimer des émotions (associées habituellement à la joie et au plaisir). Selon Mowrer & al. [6], on définit généralement le rire comme étant à la fois une vocalisation et un comportement (vocal non verbal). Ce comportement est étrange, difficile à analyser et il a reçu beaucoup d'attention d'un point de vue psychologique, souvent en rapport avec l'humour. Les mêmes auteurs signalent aussi que le rire est un réflexe émotionnel instinctif, une réponse à un stimulus (humour, chatouillements, etc.) et qu'il est une mine d'or d'informations communiquées à travers les aspects visuels et auditifs. Toutefois, on en connaît peu sur les caractéristiques acoustiques du rire. Mais qu'en est-il du sourire exactement? S'agit-il d'une sorte de rire, mais de moindre intensité ou doit-on le considérer dans une classe à part? Comme plusieurs auteurs l'ont démontré, il est perceptible, mais là encore, tout comme pour le rire, on ne connaît à peu près rien de ses caractéristiques acoustiques. Or, un interlocuteur est tout de même capable d'identifier le rire et le sourire dans la voix de son vis-à-vis. S'il est possible de percevoir ces expressions, il est alors possible de leur trouver des caractéristiques communes, des paramètres acoustiques permettant de les identifier.

2. PROBLÉMATIQUE ET CADRE THÉORIQUE

2.1. Problématique

Ce qui ajoute à la particularité du rire c'est que, pour au moins un de ses aspects, il est paradoxal, en ce sens qu'il est à la fois stéréotypé et idiosyncratique. Comme le souligne Provine [7], bien que « la plupart des gens rient […] de façon assez semblable » (p. 70), ils ne rient pas de façon identique. Pour Bachorowski & al. [2], dans leur imposante étude portant sur 1024 rires de 97 adultes, les paramètres acoustiques du rire sont trop variables et complexes pour qu'on puisse invoquer une stéréotypie du rire. Pour ces auteurs, les études antérieures ont accordé trop d'importance au caractère stéréotypé du rire et leurs résultats vont à l'encontre de cette assertion, car le rire apparaît plutôt comme un répertoire de sons ayant plusieurs sous-catégories. Nous sommes en accord à la fois avec Bacharowski & al. [2] et Provine [7]. Bien qu'il soit difficile d'analyser le rire en raison de sa complexité (Bacharowski & al. [2]), « si le rire ne comportait pas une certaine invariance, nous ne reconnaîtrions pas que les gens rient et le rire ne serait pas un signal social efficace » Provine [7] (p. 70). En fait, nous pensons que ce qui fait défaut dans le domaine du rire, c'est peut-être le manque d'une véritable typologie, une classification des genres qui faciliterait justement l'analyse de cette réalité complexe.

Le sourire, quant à lui, a reçu moins d'attention que le rire. Nous avons décidé d'appliquer au sourire ce que nous venons tout juste de décrire pour le rire (invariance vs stéréotypie et idiosyncrasie). Le sourire est une expression qu'il est possible de reconnaître visuellement et comme sa production implique un changement de la configuration du conduit vocal, par rapport à sa position neutre, il est fort probable dans ce cas, qu'il soit audible également (Tartter [9]). Il a été démontré qu'il est possible de percevoir le sourire du point de vue acoustique (Tartter [9], Schröder & al. [8]), mais son côté idiosyncratique n'a pas permis jusqu'à maintenant d'identifier les paramètres précis et invariants qui lui sont associés (Aubergé & Cathiard [1]). Notre question de recherche est donc la suivante : Quels sont les corrélats prosodiques du sourire produits par des locuteurs et des locutrices du français québécois permettant de l'identifier? Notre objectif est de décrire ces corrélats prosodiques du sourire afin de

voir s'il est possible de dégager des points communs qui permettraient de mieux le définir d'un point de vue acoustique.

2.2. Cadre théorique

Dans plusieurs langues, on aurait tendance à considérer le sourire comme le « petit frère » du rire (Trouvain [12]), et de ce fait, à supposer qu'il pourrait se trouver sur le même continuum que ce dernier. Les commentaires et résultats des tests de perception de Trouvain [12] suggèrent en bout de ligne le rejet de l'hypothèse d'un continuum entre le rire et le sourire même s'il peut arriver parfois que ceux-ci partagent certaines propriétés acoustiques.

Comment alors faire la distinction entre les deux? Nous nous appuyons sur l'étude de Trouvain [13], car les termes utilisés ainsi que leurs définitions permettent de saisir les différences fondamentales entre rire et sourire. La principale distinction résiderait dans le fait que la parole peut être articulée simultanément au sourire, tandis que le rire est séparé de l'articulation. En d'autres mots, on peut parler en même temps que l'on sourit, alors qu'il est impossible de rire et de parler en même temps. On peut donc dire que le sourire est une forme du rire synchrone à la parole (« *speech-synchronous form of laughter* ») qui se divise en deux catégories : « *speech-laugh* » et « *smiled speech* » (Trouvain [13]). Nous traduisons ces termes respectivement par *parole rieuse* et *parole souriante*.

Dans la parole souriante, il est possible d'entendre, de percevoir le sourire, soit sur quelques syllabes seulement, soit dans sa totalité. Dans la parole rieuse, seulement deux syllabes, en moyenne, sont affectées par la production d'une sorte de rire qui ressemble plus à une aspiration ou à un trémolo durant le processus de phonation (Trouvain, communication personnelle). Même s'il est possible d'étoffer ces concepts, de leur apporter plus de nuance, nous nous en tenons, pour le moment, à ces brèves définitions qui sont suffisamment explicites et qui rendent compte de la distinction essentielle à faire entre les deux types de sourires dans la parole. Ici, nous nous intéressons à la parole souriante.

3. MÉTHODOLOGIE

3.1. Constitution et présentation du corpus

Nous avons d'abord sélectionné 10 caricatures réalisées par Chapleau [3], [4] parues en 2003 et en 2004 dans le quotidien La Presse. Le choix des caricatures, en ce qui a trait à leur côté humoristique, s'est fait de façon informelle par accord interjuges. La sélection s'est effectuée sur la base des critères suivants : Les évènements illustrés dans les caricatures devaient être encore assez évocateurs. Afin que les participants évitent de jouer le ou les personnages représentés dans les caricatures, celles-ci ne devaient pas avoir de phylactères. Les titres devaient avoir entre 12 et 24 syllabes, car même s'il est possible de percevoir le sourire sur une seule syllabe comme l'a démontré Tartter [9], nous nous appuyons sur Trouvain [12] qui suggère que le sourire est un évènement à long terme alors que le rire serait un évènement à court terme.

Suivant les mêmes critères, nous avons composé 20 autres énoncés : 10 énoncés seuls et 10 énoncés présentés avec des images de type neutre. Les 10 énoncés des caricatures étaient présentés sans le dessin et avec le dessin (total = 40 énoncés). Nous avons constitué 6 versions du corpus (pour les 6 participants) où les énoncés apparaissaient dans un ordre aléatoire, afin d'éviter qu'un énoncé ait un effet sur l'énoncé qui suivrait. Les 10 premiers énoncés, tout comme dans Schröder & al. [8], devaient tous être des énoncés « neutres », c'est-à-dire sans caricatures, pour que les participants ne se doutent pas, au début du moins, du but de l'expérimentation. Pour celle-ci, nous avons utilisé le logiciel Power Point.

Participants et enregistrements

Trois femmes et 3 hommes, recrutés en milieu universitaire, âgés entre 22 et 34 ans, ayant le français québécois comme langue maternelle, ont pris part aux enregistrements.

Les enregistrements ont eu lieu dans une chambre sourde où les participants ont été filmés et enregistrés, à l'aide d'un ordinateur portable IBM C17302, d'un caméscope numérique Panasonic AG-DVC30, d'une enregistreuse numérique (DAT) TASCAM et d'un microphone dynamique Shure Beta 58A.

Les participants étaient assis devant l'ordinateur portable, le micro était à environ 30 cm d'eux et ils étaient filmés à environ 45°. Les consignes apparaissaient à l'écran de l'ordinateur et nous ajoutions qu'ils devaient s'imaginer être dans un endroit agréable (chez un ami en train de prendre un verre par exemple) afin de leur faire oublier, autant que faire se peut, l'environnement dans lequel ils se trouvaient. C'est la raison pour laquelle nous étions présente durant l'expérimentation. Comme le rire (et le sourire selon nous) a un caractère interactif, nous voulions créer une atmosphère conviviale où les participants ne sentiraient pas l'absurdité de la situation s'ils devaient rire seuls dans la chambre sourde d'un laboratoire de phonétique. Nous ne nous empêchions donc pas de sourire lorsque les participants souriaient eux-mêmes. Une phase d'entraînement précédait les enregistrements. Juste avant, ils devaient lire à haute voix le formulaire de consentement qu'ils avaient signé. Ceci constituait l'échantillon de F0 de la parole normale, en situation de lecture. Nous n'avons pas pris les 10 premiers énoncés neutres, car, au début, la nervosité de certains participants aurait pu modifier la F0. Ceux-ci devaient lire les 40 énoncés et pour passer à l'énoncé suivant, ils appuyaient eux-mêmes sur la

touche du clavier correspondant à la barre d'espacement. L'enregistrement durait entre 5 et 10 minutes.

3.2. Test de perception

Nous avons numérisé les données avec Adobe Premiere et segmenté les énoncés avec Goldwave. Pour la sélection du sous-corpus, nous avons repéré acoustiquement les énoncés qui semblaient avoir été prononcés avec un sourire et nous avons visionné l'enregistrement afin de sélectionner tous les énoncés prononcés avec l'expression faciale (étirement des lèvres) correspondant au sourire. Au total, 32 énoncés ont été retenus.

Un accord interjuges (3 femmes et 2 hommes) a permis de déterminer le deuxième sous-corpus. Pour ce test de perception, nous avons ajouté 12 énoncés neutres aux 32 « souriants » déjà retenus (total = 44). Les personnes entendaient les énoncés une seule fois et devaient dire s'il s'agissait d'un énoncé produit avec un sourire dans la voix ou de façon neutre. Pour l'analyse, nous n'avons conservé que les séquences perçues de façon dominante par les 6 auditeurs (incluant nous-même), soit 4 réponses identiques (et plus) sur 6. Notre analyse portera donc, en grande partie, sur 12 énoncés.

3.3. Analyse acoustique

Nous avons analysé nos données avec Praat, un logiciel de traitement de la parole. Les données ont été segmentées en syllabes. Les mesures de F0 ont été extraites à l'aide de l'algorithme d'autocorrélation, par pas de 10 ms. Une moyenne de F0 a été déterminée, pour chaque phrase. De plus, toutes les séquences ont été transcrites à l'aide du système de transcription ToBI (voir Thibault [10]), qui implique l'assignation de tons haut (H) et bas (B) en différents points de la courbe intonative. Cette transcription nous permet d'étudier l'inventaire phonologique prosodique des séquences, alors que les mesures instrumentales de F0 correspondent à l'implémentation phonétique de cette suite de tons.

4. RÉSULTATS ET DISCUSSION

4.1. Inventaire phonologique : tons haut (H) et bas (B)

Comme les 12 énoncés de notre deuxième sous-corpus, n'avaient pas tous leur pendant neutre, nous en avons sélectionné 6 pour lesquels le problème ne se présentait pas, soit 1 par locuteur (total =12). Une analyse qualitative des tons H et B ne nous a pas permis de dégager un patron commun qui permettrait de différencier les énoncés neutres des énoncés de parole souriante.

4.2. Réalisation phonétique : moyenne et étendue de F0

On peut voir aux figures 1 et 2 que la moyenne de F0 en parole normale est plus basse que la moyenne de F0 en parole souriante, sauf pour un énoncé chez les femmes et deux chez les hommes. Ce résultat est conforme à ce qu'ont rapporté Mowrer & al. [6] et Hirson [5], pour le rire et Tartter [9], pour le sourire. Pour ce qui est de l'étendue de F0, on constate, aux figures 3 et 4, qu'elle est plus grande en parole normale qu'en parole souriante, à l'exception de 2 énoncés chez les femmes.

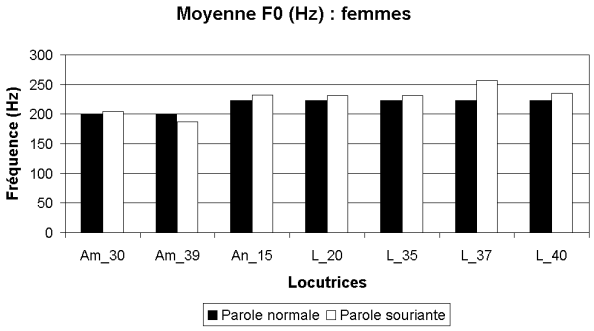

Figure 1 : Moyenne de F0 (Hz) pour les femmes, condition neutre et sourire

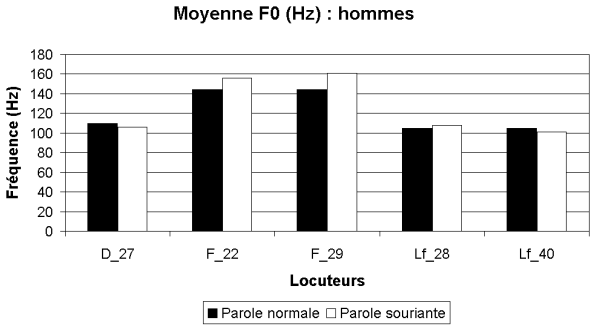

Figure 2 : Moyenne de F0 (Hz) pour les hommes, condition neutre et sourire

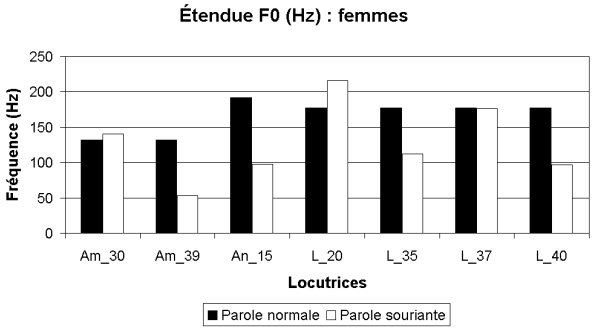

Figure 3 : Étendue de F0 (Hz) pour les femmes, condition neutre et sourire

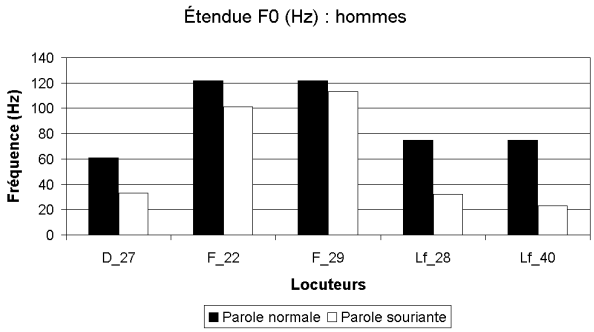

Figure 4 : Étendue de F0 (Hz) pour les hommes, condition neutre et sourire

Ces résultats nous rappellent ceux obtenus par Thibault [10] qui avait démontré que l'étendue de F0 est plus grande en situation de lecture qu'en parole spontanée. Une relation est possible avec nos résultats : l'étendue de F0 est plus grande en parole normale qu'en parole souriante. La lecture, chez Thibault [10], et la parole normale dans notre étude, correspondent aux conditions neutres (non marquées) ont une plus grande étendue; alors que la parole spontanée et la parole souriante, (conditions marquées), ont une plus petite étendue. Ceci suggèrerait un lien avec le style de parole, le style de discours et l'étendue de la F0. Ceci pourrait servir à un premier débroussaillage des contextes, des situations où apparaissent le rire et le sourire.

4.3. Limites de la recherche

La variation intra/inter sujets et certains impondérables nous empêche, pour l'instant, d'aboutir à des conclusions générales sur la parole souriante. Un des cas inattendus, pour tous les locuteurs, a été la propagation du sourire d'un énoncé *x* sur un énoncé *y* où les locuteurs, en lisant un énoncé neutre, continuaient de sourire en repensant à la caricature qu'ils venaient tout juste de voir. Conséquences : Il n'est pas possible d'analyser les mêmes énoncés pour tous les locuteurs et les énoncés des caricatures produits avec le sourire n'ont pas tous leur contrepartie neutre.

5. CONCLUSION

Nos données démontrent que le sourire est perceptible dans la parole. En plus du contenu linguistique lui-même, des informations paralinguistiques et extralinguistiques sont également transmises à travers les signaux de la parole (Traunmüller [11]). À ces signaux, il est impossible de soustraire les qualités individuelles et l'état d'un locuteur (sexe, âge, émotions, etc.). Lorsque le locuteur devient auditeur, il est intéressant de constater qu'il tient compte, inconsciemment, des qualités individuelles de son interlocuteur : « … in speech perception, listeners do not evaluate the acoustic cues directly but rather in a relational way, taking the personal properties of the speaker into account » (Traunmüller [11], p. 171) .

La moyenne et l'étendue de F0 offrent une piste intéressante si on les combine avec d'autres paramètres telles les situations de discours, les différents contextes. Nous avons remarqué que dans la parole souriante, il y a souvent des pauses, des hésitations, des faux départs, des reprises, etc. Ceci suggère que la facette pragmatique du sourire aurait un rôle non négligeable à jouer, ce qui pourrait nous aider éventuellement à bâtir une véritable typologie dans le domaine.

BIBLIOGRAPHIE

[1] V. Aubergé and M.-A. Cathiard. Can we hear the prosody of smile? *Speech Communication*, 40: 87-97, 2003.

[2] J.-A. Bachorowski,, M. J. Smoski and M. J. Owren. The acoustic features of human laughter. *Journal of the Acoustical Society of America*, 111 (3): 1582-1597, 2001.

[3] S. Chapleau. *L'année Chapleau 2003*. Boréal, Montréal, 2003.

[4] S. Chapleau. *L'année Chapleau 2004*. Boréal, Montréal, 2004.

[5] A. Hirson. Human Laughter – A Forensic Phonetic Perspective. In *Studies in Forensic Phonetics*, sous la dir. de A. Braun et J.-P. Köster, Wissenschaftlicher Verlag, Trier, 77-86, 1995.

[6] D. E. Mowrer, L. L. LaPointe and J. Case. Analysis of five acoustic correlates of laughter, *Journal of Nonverbal Behavior*, 11 (3): 191-199, 1987.

[7] R. Provine. *Le rire, sa vie, son œuvre*. Trad. de l'américain par J.-L. Fidel. Robert Laffont, Paris, 2003.

[8] M. Schröder, V. Aubergé and M.-A. Cathiard. Can we hear smile? In *Proc. Conf. on Spoken Language Processing*, volume 3, pages 559-562, 1998.

[9] V. C. Tartter. Happy talk: Perceptual and acoustic effects of smiling on speech, *Perception and Psychophysics*, 27 (1): 24-27, 1980.

[10] L. Thibault. Variations phonétiques et tonales en français québécois lu et spontané. Thèse de doctorat, Université du Québec à Montréal, 1998.

[11] H. Traunmüller. Conventional, Biological and Environmental Factors in Speech Communication: A Modulation Theory. *Phonetica*, 51: 170-183, 1994.

[12] J. Trouvain. Phonetic Aspects of "Speech-Laughs". In *Proc. of ORAGE, Orality and Gestuality Conference*, 634-639, 2001.

[13] J. Trouvain. Segmenting Phonetic Units in Laughter. In *Proc. of ICPhS*, 2793-2796, 2003.