

# Dissociation de la protrusion et de l'arrondissement dans la production des consonnes labialisées de l'anglais

Martine TODA<sup>1</sup>, Shinji MAEDA<sup>2</sup>, Andreas J. CARLEN<sup>3</sup>, Lyes MEFTAHI<sup>3</sup>

<sup>1</sup> ILPGA, Université Paris III & CNRS UMR 7018 – 19, rue des Bernardins, 75005 Paris, France

<sup>2</sup> ENST & URA820/CNRS – 46 rue Barrault 75634 Paris cedex 13

<sup>3</sup> SyncMagic, 74bis, rue des Archives 75003 Paris

Tél.: +33 (0)1 44 32 05 70 - Fax: +33 (0)1 44 32 05 73 Mél : toda@tsi.enst.fr

## ABSTRACT

No formal distinction is usually made between lip rounding and protrusion in articulatory description of English phonemes. Our study shows that in spite of the poor contribution of lips in phonological contrast, there is two lip rounding/protrusion patterns. These findings can be related to acoustical mechanisms of the labialised consonants. Labial approximant /w/ has a low F2 (Helmholtz resonance) that requires both strong rounding and protrusion, such as found in our data, while palatoalveolar fricatives (quarter-wave length resonance) shows a relatively wide lip aperture but an important protrusion that could contribute to lower their global spectrum to accurate the contrast with other sibilants.

## 1. INTRODUCTION

L'arrondissement et la protrusion des lèvres sont deux notions qui sont souvent utilisées comme synonymes dans la description du système phonologique de l'anglais. Toutefois, pour le français, [Ben91] ont montré que le geste labial de la chuintante /ʒ/ (aperture horizontale faible mais aperture verticale forte) n'était pas le même que celui de la voyelle arrondie /y/ (aperture verticale et horizontale faibles).

L'objectif de notre étude est de déterminer la forme des lèvres des phonèmes de l'anglais américain, en vue d'une modélisation trois-dimensions (3D) réaliste. Dans cette communication, nous nous concentrerons cependant sur les consonnes connues comme étant labialisées (arrondies et/ou protrusées ; [Oli93], [Ste98]) non fermées (non bilabiales et non labiodentales). Il s'agit des approximantes labiale /w/ et rhotique /ɹ/ et des fricatives et affriquées post-alvéolaires /ʃ/, /ʒ/, /tʃ/, et /dʒ/ (chuintantes). Ces phonèmes comptent tous un indice acoustique de fréquence basse (F2 bas pour /w/, F3 bas pour /ɹ/, et fréquence de bruit relativement basses des chuintantes en contraste avec les sifflantes) où la labialité est susceptible d'y jouer un rôle.

Bien que l'arrondissement et/ou la protrusion occupent un rôle minime dans le système phonologique de l'anglais, il ressort de notre étude que ces deux « gestes » ne sont pas concomitants dans la réalisation des consonnes. Le locuteur AC est en effet capable d'un usage différencié de

deux gestes labiaux, et cette différenciation pourrait être liée au mécanisme acoustique différent entre ces phonèmes.

## 2. MÉTHODE

### 2.1 Corpus, locuteur et technique d'acquisition

Un locuteur américain AC\* a prononcé le corpus suivant :

1. 79 logatomes des formes VCV (accent sur la deuxième syllabe) et VC, où le contexte vocalique V (symétrique pour la structure VCV) était composé de trois timbres vocaliques /a/, /i/ et /u/, et où C était chacune des consonnes de l'anglais /p, b, m, f, v, w, ɹ, ʃ, ʒ, tʃ, dʒ, s, z, ts, dz, t, d, θ, ð, k, g, n, l, j/ pour la structure /VCV/ et /ŋ, ɹ, ɹ/ pour la structure VC (/uzu/ et /u.ɹ/ sont lacunaires).
2. Logatomes de la forme /hVd/ avec 20 voyelles, diphtongues et voyelles rétroflexes.
3. Deux phrases courtes lues à trois vitesses, « normale », « rapide » et « lente », où il était laissé au sujet de choisir la vitesse d'élocution qu'il jugeait correspondre à ces critères.

Toutes les parties de ce corpus, d'une durée totale d'environ 75 secondes, ont donné lieu à une seule répétition et ont servi à l'analyse statistique (§2.2).

Les données articulatoires ont été obtenues grâce à la technique de *capture de mouvements*. Le locuteur a été filmé à l'aide de 6 caméras rapides (120Hz) à infrarouge. 61 marqueurs phosphorescents étaient placés sur le visage et illuminés avec une lumière infrarouge. Les données vidéo des marqueurs ont ensuite été converties en coordonnées 3D semi-automatiquement. Le son a été enregistré pendant le tournage et synchronisé à l'aide d'un clap.

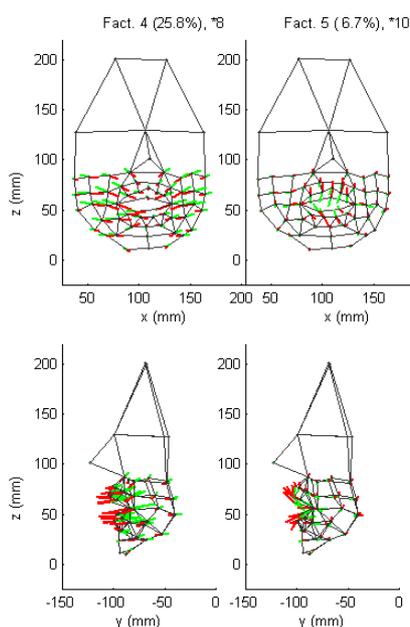
### 2.2 Analyse de facteurs arbitraires (AFA) et analyse en composantes principales (ACP)

Cette étape de l'analyse est exposée plus en détail dans la communication de Maeda *et al.* (JEP 2002). Toutes les coordonnées ont d'abord été alignées sur les 6 marqueurs posés sur le front et le nez (les parties les moins mobiles

du visage), pour éliminer le mouvement de la tête de l'analyse.

Afin d'isoler les mouvements liés à la mandibule, une analyse de facteurs arbitraires a été opérée sur un point positionné sur le menton. Les trois premiers facteurs ainsi extraits, expliquant respectivement 31,4%, 12,9% et 10,5% de la variance, correspondent au mouvement dans les dimensions verticale, transversale (antéro-postérieur) et horizontale (droite-gauche) de la mandibule.

Une analyse en composantes principales a été menée sur le résidu. Les deux premiers facteurs ainsi extraits expliquent chacun 25,8% et 6,7% de la variance totale et sont entièrement indépendants de la mâchoire. Le premier des deux – notre quatrième facteur – correspond à un mouvement de la totalité de la partie basse du visage centré autour des lèvres qui dessinent un mouvement de protrusion accompagné d'arrondissement dans le sens positif, et de rétraction et d'ouverture dans les valeurs négatives (fig. 1, gauche). Le cinquième facteur, quant à lui, correspond à un mouvement concentré surtout dans la région des lèvres, qui forment une protrusion accompagnée d'ouverture dans le sens positif et une fermeture des lèvres dans le sens négatif (fig. 1, droite). A partir du sixième facteur la variance expliquée n'atteint pas 5% de la var. totale et les facteurs d'ordre supérieure ont donc été ignorés.



**Figure 1 :** Représentation graphique des vecteurs associés à chaque nœud du maillage pour les facteurs 4 (gauche) et 5 (droite) issus d'une analyse en composantes principales après extraction des mouvements corrélés avec la mâchoire, de face (haut) et de profil (bas). Les valeurs ont été amplifiées respectivement de 8 et 10 fois afin d'être bien visibles.

### 3. RÉSULTATS

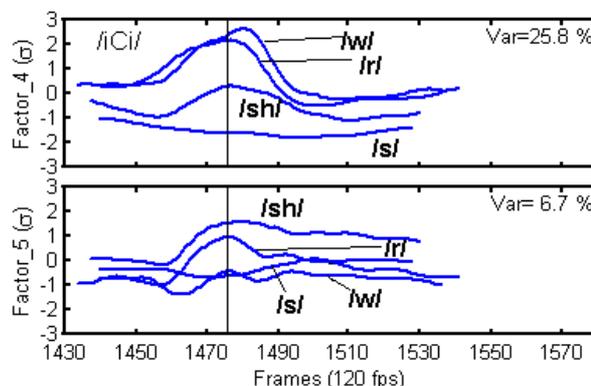
Nous avons mesuré la valeur des facteurs au milieu de la consonne (fig. 2).

La forme des lèvres est principalement déterminée par les facteurs 1 (facteur extrinsèque), 4 et 5 (facteurs intrinsèques – fig. 1).

En ce qui concerne le premier facteur, toutes les consonnes considérées ne semblent pas présenter de

variation notable\*\* (fig. 3, gauche) sinon le fait que la tendance générale est à une fermeture accrue (valeur du facteur 1 plus élevée) en contexte /u/.

En revanche, les facteurs 4 et 5 semblent permettre la discrimination de deux ou trois groupes.

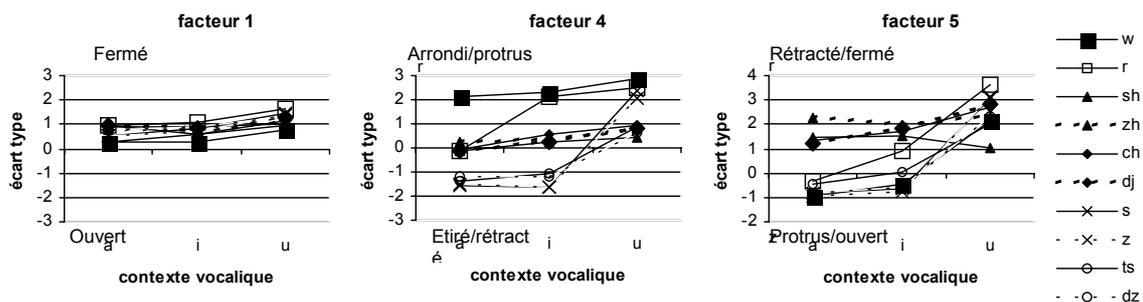


**Figure 2 :** Variation des facteurs 4 et 5 (valeur exprimée en écart type) au cours du temps pour les consonnes /w, r, sh, s/, en contexte vocalique /i/, alignées au centre de la consonne

Pour le facteur 4 (fig. 3, centre), on peut identifier trois groupes dans les contextes /a/ et /i/ : 1. /w/ (et /ɹ/ en contexte /i/) – lèvres protruses et arrondies ; 2. les chuintantes (et /ɹ/ en contexte /a/) – lèvres moyennement protruses et arrondies ; et 3. les sifflantes – lèvres étirées et rétractées. En contexte /u/, on distingue deux groupes : 1. /w, ɹ, s, z/ – lèvres protruses et arrondies ; et 2. les chuintantes et /ts, dz/ – lèvres moyennement protruses et arrondies. L'approximante /ɹ/ a un comportement non consistant, en se rangeant dans le groupe des chuintantes pour l'occurrence du contexte vocalique /a/ et avec /w/ dans le contexte /i/.

Le facteur 5 (fig. 3, droite) permet, quant à lui, de distinguer deux groupes de consonnes en contexte /a/ et /i/ : 1. les chuintantes – lèvres protruses et ouvertes ; et 2. /w/ et les sifflantes – lèvres moyennement protruses et moyennement ouvertes. En contexte /u/, toutes les consonnes sauf /z/ forment un groupe – lèvres protruses. Ici encore, l'approximante /ɹ/ a un comportement indépendant en s'approchant de /w/ et des sifflantes en contexte /a/ et se situant entre les deux groupes en contexte /i/.

En conclusion, si /ɹ/ présente un comportement non consistant, /w/ d'une part et les chuintantes d'autre part ont un comportement systématiquement opposé : /w/ est protrus et arrondi (facteur 4 élevé) mais moyennement ouvert et protrus (facteur 5 moyen), tandis que les chuintantes ont un facteur 4 moyen (lèvres moyennement protruses et arrondies), et un facteur 5 élevé (lèvres protruses et ouvertes).



**Figure 3 :** Valeur des facteurs 1, 4 et 5 au centre de la consonne. Les consonnes représentées sont /w, r, ʃ, ʒ, tʃ, dʒ/ et /s, z, ts, dz/ pour comparaison, dans trois contextes vocaux : /a, i, et u/.

#### 4. DISCUSSION : ACOUSTIQUE DES CONSONNES LABIALISÉES

Afin d'interpréter les gestes observés chez notre locuteur dans la réalisation des consonnes labialisées, nous allons nous pencher sur le mécanisme acoustique sous-jacent de trois groupes de consonnes, présentés en tableau 1.

**Tableau 1 :** Acoustique des consonnes anglaises labialisées.

	<i>indice acoustique caractéristique</i>	<i>résonance impliquée</i>	<i>geste permettant d'abaisser cette résonance</i>
/w/ [Ste98]	F2 bas (F1 et F2 groupés)	Résonance de Helmholtz de la cavité antérieure	recul de la langue, arrondissement et protrusion labiale
/ʃ, ʒ, tʃ, dʒ/ [Ste98]	sommet d'énergie dans les fréquences relativement basses	Résonance de la cavité antérieure à la constriction	protrusion labiale
/ɹ/ <b>rétroflexe</b> [Ste98]	F3 très bas	Paire pôle-zéro (résonance quart d'onde) due à la cavité branchante sublinguale	recul de la langue
/ɹ/ <b>dorso-palatal avec constriction labiale large</b> [Esp00]		Paire pôle-zéro (résonance quart d'onde) due à la configuration de la langue (cavités latérales – résonances quart d'onde)	recul de la langue, protrusion labiale
/ɹ/ <b>dorso-palatal avec constriction labiale étroite</b> [Esp00]		Résonance de Helmholtz de la cavité antérieure	recul de la langue, protrusion labiale et arrondissement labial

**L'approximante labiale /w/** L'approximante labiale se caractérise par son second formant (F2) très bas qui vient former avec le premier formant une concentration d'énergie dans les basses fréquences (et qui abaissent l'amplitude de F3). Cet indice acoustique la distingue des autres sonantes. /w/ peut être modélisé par deux cavités de Helmholtz [Ste98], une première formée par la cavité pharyngale et la constriction dorsale, et une deuxième par la cavité antérieure à cette constriction orale et la constriction labiale. La constriction labiale est donc une condition nécessaire à la réalisation de /w/ et non un geste destiné à renforcer des indices acoustiques présents.

Si la deuxième résonance de Helmholtz (qui donne lieu au F2) est bien associée à la cavité antérieure [Ste98], alors : 1. le recul de la constriction orale (ayant comme conséquence d'agrandir le volume de la cavité orale) ; 2. la fermeture encore plus prononcée des lèvres, si elle est possible (ce qui se traduit par un arrondissement) ; ou encore 3. l'allongement des lèvres (la protrusion) auront

tous trois comme effet d'abaisser la fréquence de cette résonance, en favorisant par là le rapprochement et le renforcement mutuel des deux premiers formants.

Nos données montrent que le locuteur a produit à la fois un arrondissement et une protrusion importants, ce qui va dans le sens des prédictions émanant de l'acoustique.

**Les chuintantes /ʃ, ʒ, tʃ, dʒ/** Les fricatives dentales, du point de vue statique, se distinguent principalement par leur distribution spectrale. Les sifflantes présentent un spectre dont le sommet d'amplitude se situe au-delà de 3700Hz, alors que pour les chuintantes il se situe entre 1700 et 4500 Hz [Oli93]. Cette différence spectrale peut être expliquée par la position de la constriction, qui est plus postérieure pour les chuintantes que les sifflantes. La cavité de résonance (située en avant de la constriction) est donc plus longue pour les chuintantes que les sifflantes [Ste98], ce qui explique une fréquence de résonance plus basse pour les chuintantes\*\*\*. Si on tient

compte des lèvres, la protrusion, qui allonge la cavité de résonance, aurait comme conséquence d'abaisser la fréquence de cette dernière. La protrusion viendrait donc *renforcer* un indice acoustique déjà présent.

Là encore, nos données indiquent que le locuteur aurait pu avoir orienté son geste labial en fonction du mécanisme acoustique des chuintantes, en produisant une protrusion non accompagnée d'arrondissement prononcé tel qu'observé pour la consonne /w/.

**L'approximante rhotique /ɹ/** Enfin, l'approximante rhotique /ɹ/ a déjà été décrite comme pouvant donner lieu à une variation articuloire importante [Gun99], mais cette variation, liée au contexte, permettrait une réalisation acoustique consistante : un F3 très bas qui la distingue de tous les autres phonèmes. D'après [Esp00], on pourrait distinguer traditionnellement deux types de /ɹ/ du point de vue articuloire : le /ɹ/ *rétroflexe* et le /ɹ/ que nous appellerons *dorso-palatal* (« *bunched* »).

La configuration du /ɹ/ rétroflexe pourrait être comparée à deux cavités de Helmholtz [Ste98], donnant F1 et F2. La formation d'une cavité branchante sublinguale due au geste rétroflexe introduirait une paire pôle-zéro dans la zone de fréquences habituellement occupée par le deuxième formant. C'est cet élément qui donnerait lieu à un F3 particulièrement bas.

La configuration dorso-palatale, quant à elle, reviendrait au même mécanisme que celui du /ɹ/ rétroflexe, à savoir la création d'une paire pôle-zéro dans la région du deuxième formant s'ajoutant aux deux premiers formants résultant de deux résonances de Helmholtz. [Esp00]. Si la constriction labiale n'est pas étroite, la cavité en avant de la constriction palatale formerait une résonance quart d'onde, mais la forme particulière de la langue créerait l'équivalent de cavités latérales, ce qui générerait une paire pôle-zéro. La protrusion labiale aurait donc comme effet, dans ce cas, d'allonger ces cavités latérales, abaissant ainsi leur résonance. Au contraire, si la constriction labiale est suffisamment étroite, cette constriction formerait, avec la cavité orale antérieure, un résonateur de Helmholtz. La protrusion labiale aurait, là encore, comme effet d'abaisser cette résonance. Dans tous les cas, la protrusion labiale aurait donc comme effet d'accentuer le F3 bas, contribuant ainsi à l'iconicité du /ɹ/.

Nos données sur /ɹ/ présentent une instabilité. Cette dernière pourrait être liée à la coarticulation (influence du contexte vocalique) et au choix différencié des types de /ɹ/ [Esp00], mais notre étude ne nous permet pas d'en savoir plus à cause du faible nombre de répétitions et de l'absence de données simultanées sur la langue.

## CONCLUSION

Dans cette étude, nous avons observé une différenciation des gestes d'arrondissement et de protrusion pour les

consonnes labialisées de l'anglais. Cette différenciation, du moins en ce qui concerne les chuintantes et /w/, sont en accord avec les prédictions émanant de leur mécanisme acoustique. Cette différenciation indique aussi l'existence de (au moins) deux unités fonctionnelles labiales, dans une langue pour laquelle la labialité occupe un rôle phonologique marginal. Il serait intéressant d'approfondir cette étude sur des langues tel que le français ou le suédois (opposition de labialité pour les voyelles), ou l'éthiopien (série de consonnes labialisées).

## REMERCIEMENTS

Nous remercions Rémi Brun d'*Attitude-Studio* pour l'acquisition des données articuloires. Cette étude a été menée dans le cadre du projet SAALSA (programme PRIAMM no. 7215/17/00 du CNC).

## BIBLIOGRAPHIE

- [Ben91] Benoît, C., Abry, C. et Boë, L.-J. (1991), "The effects of context on labiality in French", *Proc. 2<sup>nd</sup> Eurospeech Conference*, Genes, Italie, vol. 1, pp. 153-156.
- [Esp00] Espy-Wilson, C. Y., Boyce, S. E., Jackson, M., Narayanan, S et Alwan, A. (2000), "Acoustic modeling of American English /r/", *J. Ac. S. Am.* 108 (1), pp. 343-356.
- [Fle90] Fletcher, S. G. et Newman, D. G. (1991), "[s] and [ʃ] as a function of linguopalatal contact place and sibilant groove width", *J. Ac. S. Am.* 89 (2), pp. 850-858.
- [Gun99] Guenther, F. H., Espy-Wilson, C. Y., Boyce, S. E., Matthies, M. L., Zandipour, Majid et Perkell, J. S. (1999), "Articulatory tradeoffs reduce acoustic variability during American English /r/ production", *J. Ac. S. Am.* 105 (5), 2854-2865.
- [Jon00] Jongman, A., Wayland, R. et Wong, S. (2000), "Acoustic characteristics of English fricatives", *J. Ac. S. Am.* 108 (3), pp. 1252-1263.
- [Oli93] Olive, J. P., Greenwood, A. et Coleman, J. (1993), *Acoustics of American English Speech*, N. Y.: Springer.
- [Ste98] Stevens, K. N. (1998), *Acoustic Phonetics*, Cambridge, MA. : MIT press.

---

\* Ce locuteur est installé en France depuis 5 ans au moment de l'enregistrement. Son bilinguisme a peut-être pu influencer nos résultats.

\*\* Nous n'avons pas été en mesure de mener de test statistique, à cause du très faible nombre de répétitions.

\*\*\* D'après [Fle90], l'aire de constriction pourrait aussi participer à la différence spectrale des sifflantes et des chuintantes.