

Etude de la structure formantique des voyelles produites par des locuteurs bègues en vitesses d'élocution normale et rapide

Fabrice Hirsch Véronique Ferbach-Hecker Florence Fauvet & Béatrice Vaxelaire

Institut de Phonétique de Strasbourg – E.A. 1339 – LiLPa -Composante Parole et Cognition - Université Marc Bloch
22, rue Descartes - 67084 Strasbourg
Tél : ++33(0) 88.41.73.64 - Mél : fabrice_hirsch@yahoo.fr

Abstract

The aim of this study is to analyse the steady—state portion of the first two formants (F1) and (F2) in the production of [pVp] sequences, containing vowels [i, a, u] pronounced in two speech rates (normal and fast) by groups of untreated and treated stutterers, and nonstutterer controls. Comparing data between treated or untreated stutterers and controls, a reduction of vowel space is observed for stutterers in a normal speaking rate. When speech rate increases, no reduction of vowel space is noticeable for untreated stutterers, contrary to treated stutterers and controls.

1. Introduction

L'objectif de cette investigation est d'étudier la structure formantique des voyelles produites par des bègues et d'anciens bègues, en comparant leurs données à celles de sujets de contrôle. En outre, nous étudierons également les effets de l'accélération de la vitesse d'élocution sur la structure formantique des voyelles, et cela pour les trois groupes de locuteurs.

Nous avons choisi d'analyser les voyelles orales [i, a, u] du français, à partir de données spectrographiques, dans deux conditions de vitesse d'élocution : normale et rapide. Le choix de ces trois voyelles est motivé par le fait qu'elles représentent les extrêmes du triangle vocalique du français. En les étudiant, il devient ainsi possible d'explorer les limites de l'espace vocalique maximal, puisque leurs productions reflèteraient les capacités articulatoires maximales du locuteur, lors de la réalisation de gestes vocaliques. Pour ce qui concerne des sujets ne souffrant d'aucune pathologie, de nombreuses études ([4] par exemple) ont montré qu'une augmentation de la vitesse d'élocution pouvait entraîner une compression des durées et une réduction de l'espace vocalique, c'est-à-dire une certaine centralisation des voyelles dans cet espace. Cependant, ce phénomène de centralisation n'a été observé que pour deux voyelles, en l'occurrence pour le [i] et pour le [u].

Qu'en sera-t-il pour des sujets bègues ou anciens bègues, sachant que la plupart des études [3] ont montré une centralisation du triangle vocalique déjà en vitesse d'élocution normale, pour les bègues (par rapport à des locuteurs de contrôle) ? De même, un travail plus récent de Blomgren *et al.* [1] confirme la réduction du triangle vocalique dans la parole fluente des bègues, bien qu'un travail de Prosek *et al.*[5] infirme de telles conclusions,

dans la mesure où leurs données ne présentent pas de centralisation vocalique, aussi bien en parole fluente qu'en parole disfluente.

L'intérêt de cette étude sera donc double : premièrement, il s'agira d'apporter des données supplémentaires afin de vérifier s'il y a ou non réduction du triangle vocalique dans la parole bègue ; deuxièmement de voir si un phénomène d'« undershoot » s'opère également pour ce type de locuteur en vitesse d'élocution rapide. Nos hypothèses majeures sont les suivantes : 1) on devrait observer un espace vocalique plus restreint en parole bègayée, par rapport à la parole non pathologique, étant donné la difficulté pour les bègues à gérer des faits coarticulatoires et en conséquence l'atteinte de « cibles » vocaliques adéquatement amples ; 2) il serait, en conséquence, difficile pour les locuteurs bègues de réduire davantage leur espace vocalique avec l'augmentation de la vitesse d'élocution, si celui-ci est déjà exigu.

2. Procédure expérimentale

2.1. Locuteurs et corpus

Neuf locuteurs adultes âgés de 25 à 30 ans, dont trois sans trouble de la parole, trois bègues et trois anciens bègues qui ont suivi une thérapie et qui ne présentent ni signe de bègayages ni stratégie d'évitement ont participé à cette étude. Tous avaient pour consigne de répéter à dix reprises des séquences [pVp] introduites dans des phrases porteuses, où [V] était soit [i], [a] ou [u]. Le corpus consistait donc à lire les phrases :

1. C'est une pipe ça.
2. C'est une pape ça
3. C'est une poupe ça.

Les répétitions ont été enregistrées avec un microphone Sennheiser e845 S relié à un portable PC (carte son RealTek AC97) et en utilisant Audacity (Fréquence d'échantillonnage : 44100 Hz – 16 bits) comme logiciel d'acquisition.

2.2. Mesures acoustiques et calcul de l'espace vocalique

Mesures acoustiques

Les données recueillies ont été analysées à l'aide du logiciel Praat©. Seules les séquences fluentes ont été retenues pour cette analyse. Les mesures de F1 et de F2 ont été prises au milieu de la structure formantique stable des voyelles.

Ces valeurs, qui représentent les résonances dans le conduit vocal, permettent de faire des inférences de la configuration du conduit vocal lors de la production d'une voyelle [6].

Calcul de l'espace vocalique

En plus de ces observations indirectes sur l'élévation et l'avancement de la langue dans la cavité buccale, l'aire du triangle vocalique sera également calculée [1]. Cette mesure, exprimée en Hz², permet d'obtenir une indication sur l'espace utilisé en vue de réaliser la distinction entre les voyelles. Il est à noter que l'aire des triangles vocaliques est obtenue par le calcul suivant :

$$\text{Aire} = \sqrt{P(P-a)(P-b)(P-c)}$$

où a, b et c représentent la distance entre les coordonnées de deux voyelles qui est quantifiée par :

$$\sqrt{(x_b - x_a)^2 + (y_b - y_a)^2} \quad (x \text{ étant la coordonnée de F1 d'une des trois voyelles et } y \text{ la coordonnée de F2}) \text{ et où } P \text{ est le résultat de : } (a + b + c) / 2$$

3. Résultats

Des analyses de variance (ANOVA) à 3 facteurs – *qualité vocalique* ([i, a, u]), *groupes de locuteurs* (bègues vs. locuteurs de contrôle vs. anciens bègues) et *vitesse d'élocution* (normale vs. rapide) - ont été effectuées sur les mesures de F1 et de F2. Pour la détermination de la signification statistique des *effets principaux*, ainsi que des *interactions*, seuls les résultats significatifs avec une probabilité de moins de 5% d'avoir été obtenus par chance ($p < 0.05$) seront retenus. Les trois effets principaux n'ont pas révélé de significativité statistique globale. Cependant, des analyses ANOVA du F2 ont été significatives pour le groupe de contrôle vs. les bègues et les anciens bègues vs. les bègues (quelles que soient la qualité vocalique et la vitesse d'élocution, $p < 0.05$). Des ANOVA séparées sur la vitesse d'élocution indiquent un effet significatif pour les anciens bègues et le groupe de contrôle (quelle que soit la qualité vocalique, $p < 0.05$).

3.1. Comparaison de la structure formantique en vitesse d'élocution normale

Structure formantique du [i]

La valeur moyenne de F1 a été mesurée à 218 Hz (E-T : 16 Hz) en moyenne, et celle de F2 à 2046 Hz (49 Hz) pour le locuteur de contrôle lorsqu'il répétait la voyelle [i]. Ce même son a été évalué avec une première zone de résonance à 273 Hz (7 Hz) et une deuxième à 1752 Hz (53) pour le locuteur bègue. Par conséquent, il est possible d'observer que la différence entre les deux productions provient surtout de F2, autrement dit de l'avancement de la langue dans la cavité buccale. Par inférence, il semblerait que le lieu d'articulation de la voyelle soit moins antérieur pour le locuteur bègue. En ce qui concerne le premier formant, les valeurs entre les deux locuteurs sont légèrement plus élevées pour B (218 Hz pour le locuteur de contrôle vs. 273 Hz pour le bègue), mais rappelons que l'effet principal était non significatif.

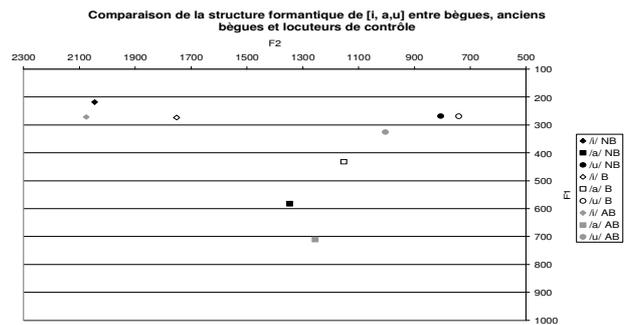


Figure 1 : comparaison des productions des voyelles [i, a, u] produites par un locuteur de contrôle (signes noirs), un locuteur bègue (signes vides) et un ancien bègue (signes gris).

Quant à l'étude de la structure formantique de la voyelle [i] chez l'ancien bègue, elle ne révèle pas de différences significatives avec les moyennes obtenues pour le locuteur de contrôle. En effet, l'opposition entre les valeurs F2 du bègue et du sujet non-pathologique ne semble pas pertinente, dans la mesure où le second formant a été quantifié à 2075 Hz (20 Hz) pour le sujet qui a suivi une thérapie (vs. 2046 Hz). Autrement dit, la position de la langue sur l'axe « antérieur – postérieur » semble identique pour le locuteur de contrôle et l'ancien bègue. Quant au premier formant, il est en moyenne de 272 Hz (9 Hz), c'est-à-dire qu'il est proche de ce qui a été constaté pour le bègue (273 Hz), et légèrement moins élevé, si on le compare à la moyenne obtenue pour le locuteur de contrôle (218 Hz).

Par conséquent, il est possible de conclure que la voyelle [i] est quasiment identique sur le plan qualitatif pour NB et pour AB et qu'elle est produite avec une position de la langue moins en avant pour B.

Structure formantique du [a]

Pour ce qui est du [a], réalisé par le sujet non-pathologique, on peut observer que le F1 est à 582 Hz (47 Hz) et le F2 à 1347 Hz (42 Hz). Ces valeurs ne sont pas celles prélevées pour le sujet bègue, tant pour le premier que pour le deuxième formant. En effet, si l'on prend le cas de F1, mesuré en moyenne à 432 Hz (26 Hz) pour le bègue, on peut noter une baisse sensible (même si l'effet principal n'avait pas été significatif) de cette valeur, ce qui traduirait le fait que la voyelle soit produite avec une ouverture plus petite pour cette catégorie de sujet, par rapport au locuteur de contrôle. Il en est de même pour la valeur moyenne de F2, qui est de 1153 Hz (25 Hz) pour le locuteur bègue, alors que ce même paramètre était de 1347 Hz pour le sujet de contrôle. Quant aux valeurs obtenues pour l'ancien bègue, elles sont différentes à la fois du locuteur de contrôle et du sujet bègue. En effet, les résultats montrent que F1 est plus élevé pour ce groupe de locuteur, puisqu'il a été quantifié à 710 Hz (14 Hz), tandis qu'il a été mesuré à 432 Hz pour B et à 582 Hz pour NB. De même, la valeur F2 se trouve à un niveau intermédiaire par rapport aux deux autres locuteurs, étant donné qu'elle

est de 1256 Hz (21 Hz) en moyenne (vs. 1347 Hz pour le locuteur NB et vs. 1153 Hz pour le locuteur B).

Structure formantique du [u]

Le premier formant de la voyelle [u] produite par le locuteur de contrôle a été observé à 268 Hz (23 Hz) en moyenne. Ce résultat est quasiment identique pour le sujet bègue, F1 ayant été mesuré à 269 Hz (11 Hz), et est moins élevé par rapport à l'ancien bègue, locuteur pour lequel la première zone de résonance était située à 325 Hz (8 Hz). C'est l'observation du deuxième formant qui révèle davantage de différences entre le locuteur de contrôle et l'ancien bègue. En effet, F2 a été mesuré à 805 Hz (50 Hz) pour le premier, alors qu'il est de 1004 Hz (32 Hz) pour le second. Il est important d'ajouter que la valeur de ce même paramètre chez le locuteur bègue est proche de celle de NB, étant donné qu'il a été mesuré à 741 Hz (50 Hz).

En conclusion, le [u] produit par le locuteur bègue est comparable à celui du sujet non-pathologique, si ce n'est qu'il est réalisé avec la langue placée légèrement plus en arrière. Quant aux réalisations de AB, elles suggèrent une antériorisation du [u] par rapport aux deux premiers locuteurs.

3.2. Effets de l'augmentation de la vitesse d'élocution selon le type de locuteurs

Conséquences de la vitesse d'élocution sur le [i]

L'augmentation de la vitesse d'élocution a pour effet d'accroître la valeur de F1 qui passe de 218 Hz à 253 Hz (14 Hz) lorsqu'il est demandé au locuteur de contrôle de parler plus rapidement. Parallèlement à cela, F2 diminue puisque la mesure pour la seconde zone de résonance atteint 1989 Hz (37 Hz) en vitesse d'élocution rapide, alors que ce même paramètre était quantifié à 2046 Hz lorsque le sujet de contrôle devait parler à vitesse d'élocution normale. Il en va de même pour AB, étant donné que F1 augmente très légèrement lorsque ce dernier parle plus rapidement, en passant de 272 Hz à 291 Hz (25 Hz). Quant au deuxième formant du [i], il diminue également puisqu'il était de 2075 Hz en vitesse d'élocution normale et qu'il se trouve à 2032 Hz (27 Hz) lorsque ce locuteur parle plus rapidement. Cette tendance ne se confirme pas pour le locuteur bègue. En effet, il est possible de constater une stabilité des valeurs de F1 (273 Hz en vitesse d'élocution normale vs. 263 Hz (14 Hz) en vitesse d'élocution rapide) et de F2 (1752 Hz vs. 1772 Hz (205 Hz)). Par conséquent, il n'y a pas de différence significative entre les productions de la voyelle [i] en vitesses d'élocution normale et rapide, chez les bègues.

Conséquences de la vitesse d'élocution sur le [a]

Nous avons pu observer plus haut que le F1 de la voyelle produite par le locuteur de contrôle était de 582 Hz et que le F2 était évalué à 1347 Hz. Lorsque la vitesse d'élocution augmente, c'est principalement la valeur du premier formant qui se trouve modifiée, étant donné qu'il passe de 582 Hz à 517 Hz (27 Hz). Quant à la deuxième

zone de résonance, elle reste stable (1347 Hz en vitesse d'élocution normale vs. 1296 Hz (43 Hz)). Peu d'évolutions ont été observées pour la même voyelle chez l'ancien bègue. Ainsi, la valeur de F1 reste stable en passant de 710 Hz à 681 Hz (29 Hz) lorsque le locuteur a pour consigne d'accélérer la vitesse d'élocution, tout comme la moyenne pour F2, puisqu'elle est de 1256 Hz en vitesse d'élocution normale et de 1297 Hz (33 Hz) en vitesse d'élocution rapide. De même, aucune différence pertinente n'a été constatée lorsque la vitesse d'élocution du locuteur bègue augmente : F1 passe de 432 Hz à 438 Hz (35 Hz) et F2 de 1153 Hz à 1116 Hz (20 Hz). Il est à noter que ces résultats sont conformes à la littérature, étant donné que Lindblom [4] et Ferbach-Hecker [2] n'ont pas observé de centralisation pour le [a].

Conséquences de la vitesse d'élocution sur le [u]

L'accélération de la vitesse d'élocution entraîne chez NB une légère modification de la structure formantique du [u]. En effet, on peut constater que F2 passe de 805 Hz à 886 Hz (97 Hz). Cependant, cette différence est non significative (à cause d'un écart-type relativement trop élevé) et doit, en conséquence, être prise avec précaution. Quant à la première zone de résonance, elle est située à 268 Hz, lorsqu'aucune contrainte de vitesse n'est exigée et elle est de 292 Hz (11 Hz) lorsque la vitesse d'élocution est accélérée. Des remarques similaires peuvent être effectuées pour AB, étant donné que F2 était de 1004 Hz en vitesse d'élocution normale et qu'il est évalué à 1063 Hz (75 Hz) en vitesse d'élocution rapide.

Enfin, il est intéressant de noter que les valeurs F1 et F2 du [u] sont quasiment identiques pour le locuteur bègue dans les deux vitesses d'élocution, puisque F1 est de 269 Hz en vitesse d'élocution normale et de 258 Hz (29 Hz) en parole rapide et que F2 est à 741 Hz lorsque le locuteur parle sans consignes temporelles et de 724 Hz (66 Hz) lorsqu'il lui est demandé de parler vite.

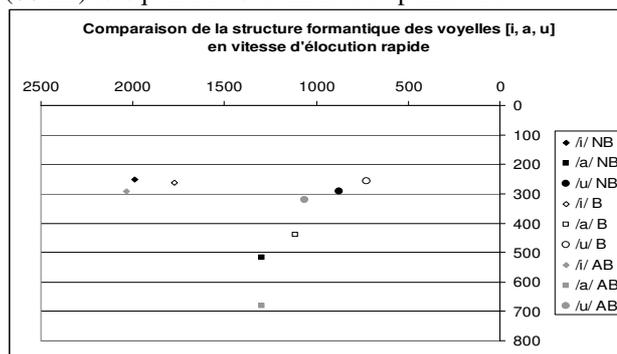


Figure 2 : comparaison des productions des voyelles [i, a, u] produites par un locuteur de contrôle (signes noirs), un locuteur bègue (signes vides) et un ancien bègue (signes gris) en vitesse d'élocution rapide.

3.3. Comparaison de l'aire des triangles vocaliques

Il est important de signaler que le calcul de l'aire et la comparaison des triangles vocaliques n'a pas de signification fonctionnelle à proprement parler. Elle

permet cependant d'obtenir une vue globale sur l'espace vocalique maximal utilisé pour obtenir une opposition qualitative entre les voyelles et sur l'amplitude des mouvements réalisés à cette fin. Les valeurs obtenues pour les trois locuteurs qui font l'objet de cette étude sont données dans le tableau 1.

Tableau 1 : Comparaison de l'aire des triangles vocaliques entre le locuteur de contrôle (NB), le sujet bègue (B) et l'ancien bègue (AB).

Aire (Hz ²)	NB	B	AB
Vitesse d'élocution normale	208.072	81.722	212.856
Vitesse d'élocution rapide	164.620	93.112	173.343

L'étude de l'aire des triangles vocaliques révèle que la différenciation des voyelles s'effectue, chez les bègues, sur une surface deux fois moins élevée par rapport au locuteur de contrôle, puisque l'espace vocalique a été calculé à 81.722 Hz² pour le premier, alors qu'il est de 208.072 Hz² pour le second. Par ailleurs, il est également possible de constater que l'aire obtenue pour les voyelles réalisées par AB est comparable au sujet non-pathologique, étant donné qu'elle est de 212.856 Hz² (vs. 208.072 Hz² pour NB).

En vitesse d'élocution rapide, il est possible de constater une réduction significative de l'aire du triangle vocalique pour NB, puisque cette valeur passe de 208.072 Hz² à 164.620 Hz². Ce constat vaut également pour AB, la surface du triangle étant de 212.856 Hz² en vitesse d'élocution normale et de 173.343 Hz² lorsque le locuteur parle plus rapidement.

Pour le locuteur bègue cependant, aucune réduction de l'aire du triangle vocalique n'est constatée. Au contraire, celle-ci augmente très légèrement, en passant de 81.722 Hz² à 93.112 Hz².

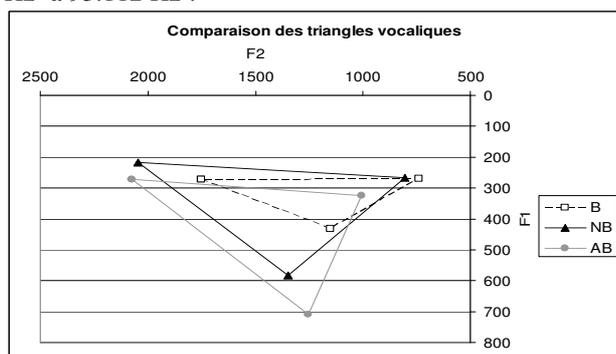


Figure 3 : Comparaison de l'aire des triangles vocaliques entre le locuteur de contrôle (NB), le sujet bègue (B) et l'ancien bègue (AB) en vitesse d'élocution normale.

4. Synthèse et conclusion

La structure formantique des voyelles [i, a, u] produites par l'ancien bègue est comparable à celle obtenue pour le locuteur de contrôle. Il semble alors cohérent de constater que l'aire des triangles vocalique de ces deux locuteurs soit similaire.

Ce n'est cependant pas le cas pour le locuteur bègue, pour qui l'aire du triangle est fortement restreinte, montrant

ainsi une différence de stratégie par rapport aux locuteurs de contrôle et anciennement bègues ; c'est le F2 qui est le grand responsable des différences observées dans les espaces vocaliques. Ce résultat suggère une particularité dans la gestion du lieu d'articulation chez les bègues [1]. Il est également intéressant de noter que l'augmentation de la vitesse d'élocution entraîne, pour les contrôles et les anciens bègues, une diminution de l'aire du triangle. Autrement dit, le locuteur de contrôle et l'ancien bègue utilisent un phénomène d'« undershoot » en accélérant leur vitesse d'élocution, ce qui n'est pas le cas pour le locuteur bègue, étant donné que la structure formantique des voyelles reste stable malgré le changement de rythme. Par conséquent, on peut dire que le locuteur bègue « sait » moins utiliser les possibilités de variation de l'espace vocalique, suivant le contexte de vitesse d'élocution.

Remerciements

Nous remercions Rudolph Sock pour ses remarques et ses suggestions. Cette recherche a été financée en partie par le Programme de Recherche ACI TTT 2003-2006 du Ministère de la Recherche et des Nouvelles Technologies, ainsi que par le Programme MISHA attribué à la Composante Parole et Cognition de l'E.A. 1339 – LiLPa.

Bibliographie

- [1] M. Blomgren M. Robb and Y. Chen. A note on Vowel Centralization in Stuttering and Nonstuttering Individuals. In *Journal of Speech, Language, and Hearing Research*, volume 41, pages 1042-1051, 1998.
- [2] V. Ferbach-Hecker. La résistivité de la qualité des voyelles orales du français. In *SCOLIA*, volume 20, pages 115-134, 2005.
- [3] R. Klich and G. May Spectrographic study of vowels in stutterers' fluent speech. In *Journal of Speech and Hearing Research*, volume 25, pages 364-370, 1982.
- [4] B. Lindblom. On vowel reduction. In *The Royal Institute of Technology, Speech Transmission Laboratory*, volume 29, 1963.
- [5] R. Prosek A. Montgomery B. Walden and D. Hawkins. Formant frequencies of stuttered and fluent vowels. *Journal of Speech and Hearing Research*, volume 30, pages 301-305, 1987.
- [6] K. Stevens and A. House. Development of a description of vowel articulation. In *Journal of Acoustical Society of America*, volume 27, pages 484-493, 1955.
- [7] B. Vaxelaire. Etude comparée des effets des variations de débit –lent, rapide– sur les paramètres articulatoires, à partir de la cinéradiographie (sujets français). Thèse de Doctorat Nouveau Régime, Université des Sciences Humaines de Strasbourg, 1993.