

Influence de la forme du palais sur la variabilité articulatoire

Jana Brunner¹, Pascal Perrier² et Susanne Fuchs³

¹Humboldt-Universität zu Berlin, INP Grenoble et Zentrum für Allgemeine Sprachwissenschaft, Berlin

²Institut de la Communication Parlée, INPG & Univ. Stendhal, Grenoble

³Zentrum für Allgemeine Sprachwissenschaft, Berlin

ABSTRACT

As has been noted previously, speakers with coronally low "flat" palates exhibit less articulatory variability than speakers with coronally high "domeshaped" palates. This phenomenon is investigated by means of a tongue model and an EPG experiment. The results show that acoustic variability depends on the shape of the vocal tract. The same articulatory variability leads to more acoustic variability if the palate is flat than if it is domeshaped. Furthermore, speakers with domeshaped palates show more articulatory variability than speakers with flat palates. The results are explained by different control strategies by the speakers. Speakers with flat palates reduce their articulatory variability in order to keep their acoustic variability low.

1. INTRODUCTION

Ainsi que l'ont déjà noté un certain nombre de publications, il semble qu'il y ait une relation entre la forme du palais d'un locuteur et sa variabilité articulatoire. Par exemple, Perkell [9] a comparé les productions de six locuteurs ayant des palais différents pour les voyelles /i/, /ɪ/ et /ɛ/. Cinq de ces locuteurs avaient des palais en forme d'arche dans le plan coronal, alors que le palais du sixième était plutôt plat. Perkell a observé que ce dernier locuteur exploitait des ajustements articulatoires de très faible amplitude, tandis que des mouvements plus amples étaient trouvés pour les autres. De la même manière, Mooshammer et al. [7] ont décrit un locuteur qui présentait une variabilité articulatoire moins grande que celles d'autres locuteurs. Ce locuteur aussi avait un palais plus plat que les autres locuteurs étudiés.

Une explication possible pour ce phénomène pourrait être associée au contrôle des gestes de la parole par le locuteur en relation avec la perception de la parole produite. Il est en effet possible que les locuteurs ayant des palais plats soient obligés de réduire leur variabilité articulatoire parce que leur variabilité acoustique serait sinon trop élevée et l'intelligibilité ne serait alors plus assurée. Cette explication se justifie de la manière suivante : un locuteur avec un palais plat a un conduit vocal dont la coupe dans le plan coronal peut grossièrement être assimilée à un quadrilatère, dont le palais est la frontière supérieure, les dents ou les joues les cotés, et la langue la frontière inférieure. Au contraire, un locuteur avec un palais en forme d'arche, a un conduit vocal que l'on peut assimiler à un triangle où la langue est la base, et le palais les deux cotés. Si la position verticale de la langue change un peu, la surface du quadrilatère change plus que la surface du triangle. En d'autres termes, pour le même changement articulatoire, la fonction d'aire change plus pour un conduit vocal dont la section est proche d'un quadrilatère que pour celui dont la section est plutôt trian-

gulaire. En conséquence, le son produit est susceptible de changer plus pour un locuteur avec un palais plat que pour un locuteur avec un palais en forme d'arche. Ceci signifie qu'un locuteur avec un palais en forme d'arche peut se permettre plus de variabilité articulatoire sans changer le son considérablement, tandis que le locuteur avec un palais plat doit articuler plus précisément pour qu'il puisse maintenir la qualité perceptive des sons produits.

Une autre question intéressante concerne la stratégie utilisée par les locuteurs pour réduire la variabilité. Pour le locuteur possédant un palais plat dans le plan coronal, il est nécessaire de trouver une position assez stable de la langue où elle ne peut pas remuer beaucoup. Une telle position est atteinte en particulier quand la langue a beaucoup de contacts palataux. Ainsi, on peut s'attendre à ce qu'un locuteur qui produit une grande quantité de contacts entre le palais et la langue ait moins de variabilité articulatoire qu'un locuteur qui a moins de contacts.

L'objectif de la présente étude est d'étudier ces relations entre la forme du palais, la variabilité articulatoire, la variabilité acoustique et les contacts linguo-palataux. Premièrement, nous étudierons la relation entre la variabilité articulatoire et la variabilité acoustique. Nous observerons en particulier si la variabilité acoustique diffère en fonction de la forme du conduit vocal pour une même variabilité articulatoire. Pour cela, nous avons effectué des simulations avec un modèle biomécanique de la langue couplé à un modèle acoustique harmonique du conduit vocal.

Ensuite, les relations entre la variabilité articulatoire, les contacts linguo-palataux et la forme du palais seront étudiées expérimentalement par électropalatographie pour 16 locuteurs allemands et norvégiens. Nous mesurerons la variabilité articulatoire, les contacts linguo-palataux et la forme du palais. Selon nos prédictions, il devrait y avoir une corrélation entre la variabilité articulatoire et la forme du palais, la variabilité étant plus grande si le palais est en forme d'arche dans le plan coronal que s'il est plat. De plus si le recours aux contacts linguo-palataux est une stratégie pour réduire la variabilité articulatoire, on peut attendre une corrélation négative entre la variabilité articulatoire et le nombre de contacts linguo-palataux.

2. MÉTHODE

La première partie de cette section traite des simulations, tandis que la deuxième partie concerne l'étude expérimentale sur les locuteurs.

2.1. Simulations

Nous avons utilisé la version la plus récente du modèle biomécanique de la langue (Payan & Perrier [8] et [12]).

La configuration de la langue pour les trois voyelles (/a/, /i/ et /u/) a été ainsi déterminée pour ce modèle.

Ces configurations de la langue ont été associées à cinq palais qui différaient dans le plan coronal. Cette courbure a été spécifiée par le coefficient α (du modèle de Heinz & Stevens [4], adapté par Perrier et al. [11]). Si α est grand, le palais est plat, s'il est petit, le palais est en forme d'arche. Pour la présente étude les valeurs suivantes ont été choisies : 3.0, 2.5, 2.0, 1.5 et 1.3. Avec cette association de la langue avec une forme spécifique du palais, les fonctions d'aire des conduits vocaux obtenus pour une configuration donnée de la langue étaient différentes selon la forme du palais. Pour cette raison, il a fallu jouer sur la position verticale de chacun de ces palais de façon à ce que les fonctions d'aire d'une même voyelle soient les mêmes quel que soit le palais considéré.

Ensuite, d'autres simulations ont été réalisées dans l'objectif de faire varier légèrement la position linguale et de simuler ainsi de la variabilité articulaire. Ici encore, ces positions de la langue étaient combinées avec les cinq palais et les fonctions d'aire étaient calculées (pour des détails sur ces simulations cf. Brunner et al. [2]).

Finalement, à partir des fonctions d'aire, les formants des sons associés à ces différentes configurations articulaires ont été calculés avec un modèle harmonique du conduit vocal (Badin & Fant [1]).

2.2. Expérience

16 locuteurs (10 allemands et 6 norvégiens) ont été enregistrés par électropalatographie (EPG 3.0, Reading system). Parallèlement, le signal acoustique de parole a été enregistré. Les locuteurs ont produit des phrases dans leur langue maternelle avec des logatomes qui contenaient les sons suivants : /s, f, ɛ, j, i, e, r, ε/. Chaque phrase était répétée trente fois.

Pour chacun des segments cibles le pourcentage des contacts linguo-palataux a été calculé à chaque instant d'échantillonnage des données électropalatographiques. Comme les sons avaient des durées différentes, ces pourcentages ont été interpolés puis rééchantillonnés sur 20 points quelle que soit la durée de la voyelle étudiée. Puis, la valeur moyenne et l'écart-type des trente répétitions ont été calculés pour chaque locuteur et pour chaque son à chacun des 20 points d'interpolation. Ensuite, les moyennes des 20 valeurs moyennes et des 20 écart-types ont été calculées. Si la langue est proche du palais, par exemple pour des voyelles hautes, il y a plus de contacts que s'elle est plutôt basse. En conséquence, dans le premier cas, un mouvement de la langue pourrait provoquer une variabilité du nombre de contact plus importante que dans le second. Pour limiter cette influence du nombre moyen des contacts sur la variabilité, la valeur moyenne de l'écart-type a été normalisée par la valeur moyenne du nombre de contacts. Ce coefficient a été considéré comme la mesure de la variabilité articulaire.

L'importance du contact linguo-palatal latéral a été calculée comme le pourcentage des contacts dans la région latérale. Comme pour le pourcentage des contacts globaux, une valeur moyenne et un écart-type moyen pour les 30 répétitions et les 20 points d'interpolation ont été calculés.

La forme des palais a été caractérisée d'une part, par la mesure des coordonnées x, y et z des électrodes du pa-

lais d'électropalatographie. Puis la courbe décrite par la sixième rangée d'électrodes de ce palais a été approchée au sens des moindres carrés par une fonction exponentielle caractérisée par le coefficient alpha, conformément aux propositions de Perrier et al. [11] (cf. Brunner et al. [2] pour des détails sur les calculs mis en oeuvre).

3. RÉSULTATS

Dans la première partie les résultats des simulations seront présentés. Ils suggèrent effectivement que, pour la même variabilité articulaire, il y a plus de variabilité acoustique si le palais est plat que s'il est en forme d'arche. La deuxième partie est consacrée aux résultats expérimentaux. Ils indiquent que les locuteurs avec des palais plats ont moins de variabilité articulaire que les locuteurs avec des palais en forme d'arche. De plus, on observe une corrélation négative entre le nombre moyen de contacts linguo-palataux et la variabilité articulaire. Cependant, ce résultat doit être traité avec précaution dans la mesure où la formulation même de la mesure de la variabilité, qui est normalisée par rapport au nombre moyen de contacts, introduit de facto un biais statistique.

3.1. Simulations

Sur la figure 1 on a représenté sous forme d'intervalles de confiance la variabilité des formants F1 et F2 de la voyelle /a/ générée pour les différents palais par les changements de positionnement de la langue.

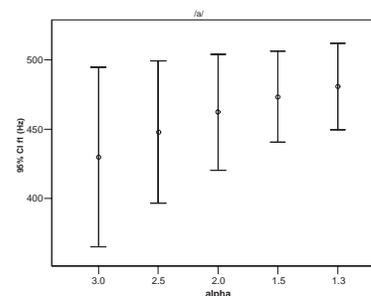


FIG. 1: Intervalles de confiance à 95% pour F1 pour les cinq palais différents. La dispersion est la plus grande pour le palais le plus plat (palais 1) et diminue pour les palais plus arqués.

C'est pour le premier palais, le plus plat avec $\alpha=3.0$, que cet intervalle est le plus grand (de 365 à 494 Hz, soit 129 Hz). L'intervalle de confiance obtenu pour le palais caractérisé par $\alpha=1.3$, c'est-à-dire pour le palais ayant la forme d'arche la plus marquée, est le plus petit (de 450 à 512, 62 Hz). Les autres intervalles diminuent avec α , c'est-à-dire au fur et à mesure que les palais prennent une forme d'arche plus marquée.

Les valeurs pour les autres voyelles et pour F2 sont données dans le tableaux 1. Ils vont dans le même sens que ceux de la figure 1. Les intervalles de confiance diminuent pour les palais plus arqués, c'est-à-dire ceux dont les valeurs α sont les plus petites. Il y a une seule exception, c'est le premier formant de /i/ pour lequel l'intervalle de confiance obtenu pour le palais défini par $\alpha=1.5$ est plus petit que celui du palais caractérisé par $\alpha=1.3$.

TAB. 1: Variabilité pour F1 et F2 (sous forme d'intervalle de confiance, en Hz) pour les cinq palais différents. La dispersion est la plus grande pour le palais le plus plat ($\alpha=3.0$) et diminue pour les palais plus arqués.

	α	95% CI F1	$\lambda F1$	95% CI F2	$\lambda F2$
a	3.0	365-494	129	1561-1770	209
	2.5	397-499	102	1569-1757	188
	2.0	420-504	84	1583-1741	158
	1.5	440-506	66	1566-1723	157
	1.3	450-512	62	1540-1693	153
i	3.0	329-501	172	1699-2263	564
	2.5	323-492	169	1765-2283	518
	2.0	313-478	165	1842-2276	434
	1.5	304-458	145	1923-2258	335
	1.3	297-449	152	1949-2235	286
u	3.0	309-397	88	718-1553	835
	2.5	313-395	82	705-1508	803
	2.0	314-392	78	684-1453	769
	1.5	313-386	73	655-1326	671
	1.3	315-379	64	620-1272	652

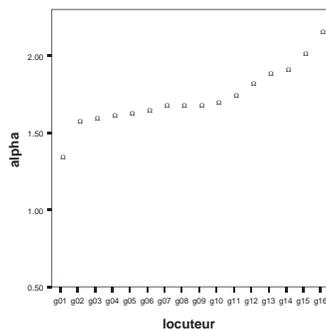


FIG. 2: Valeurs α pour tous les locuteurs. Si α est petit (à gauche) le palais est arqué dans le plan coronal, si α est grand (à droite) le palais est plat.

3.2. Résultats expérimentaux

Sur la figure 2 on peut voir les valeurs α des 16 locuteurs. Si α est petit (à gauche dans la figure) le palais du locuteur est plutôt arqué. Si, à l'opposé, α est grand, le palais est plutôt plat. Sur la figure 3 les locuteurs sont représentés dans le même ordre, mais les barres indiquent cette fois-ci la variabilité articuloire moyenne. Même si, en général, les résultats de l'expérience sont moins clairs que ceux des simulations, on peut voir qu'il y a une tendance pour les locuteurs avec des palais en forme d'arche (à gauche) à avoir plus de variabilité que les locuteurs avec des palais plats. Les barres noires indiquent les locuteurs qui ne respectent pas cette tendance générale. Les locuteurs g09 et g15 ont en effet une variabilité assez grande même si leurs palais sont plutôt plats. Les locuteurs g03 et g11, de l'autre côté, ont moins de variabilité que ne le laissait prévoir la forme de leur palais.

Une explication possible pour ces exceptions pourrait être trouvée dans leur vitesse. En effet, les locuteurs g09 et g15 ont parlé avec un débit d'élocution plutôt rapide ca-

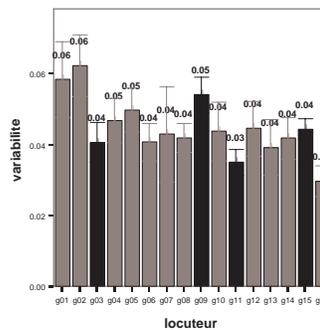


FIG. 3: Variabilité articuloire des locuteurs présentés dans le même ordre que sur la figure 2. On observe une tendance pour les locuteurs avec des palais en forme d'arche dans le plan coronal (à gauche) à avoir plus de variabilité articuloire que les locuteurs avec une palais plats (à droite). Les intervalles de variation correspondent à l'erreur standard.

ractérisé par une durée moyenne des sons cibles étudiés respectivement égale à 77 et 82 ms, tandis que les locuteurs g03 et g11 étaient moins rapides avec une durée respective des sons cibles de 177 et 145 ms. Or d'après la loi de Fitt, bien connue dans le domaine du contrôle moteur, des gestes plus rapides sont intrinsèquement associés à une moins grande précision gestuelle, et vice-versa. On peut donc faire l'hypothèse que les contraintes induites par le débit d'élocution à chaque locuteur se superposent aux contraintes liées à la morphologie du palais jusqu'à même les contrecarrer pour ce qui concerne leurs conséquences sur la variabilité articuloire. La comparaison de ces 4 locuteurs dont la mesure de la variabilité articuloire va à l'opposé de nos prédictions, avec ceux parmi les autres locuteurs qui ont des palais similaires, soutient cette hypothèse. Ainsi, par exemple, les locuteurs g08 et g10, qui ont un palais aussi arqué que celui du locuteur g09, ont moins de variabilité articuloire, et ils parlent effectivement moins vite (durée respective des sons cibles : 83 et 113 ms). Il en va de même pour le locuteur g13 comparé au locuteur g15. Ils ont des palais similaires, et présentent des durées de segments cibles respectivement égales à 118 et 74 ms. Le locuteur g13, à l'élocution plus lente, est celui des deux qui a le moins de variabilité articuloire. Dans l'autre sens la comparaison reste aussi valable, puisque si on compare les locuteurs g03 et g11, dont la variabilité articuloire est plus faible que ne laisse prévoir la forme de leur palais, aux locuteurs g02 et g04 (durées respectives des sons cibles : 79 et 11 ms), et g10 et g12 (durées respectives des sons cibles : 77 et 70 ms) respectivement, on constate bien que le débit d'élocution est particulièrement lent.

En ce qui concerne la relation entre la variabilité articuloire et le nombre de contacts linguo-palataux une corrélation négative a été trouvée ($r=-.691$, $p=.000$). Cependant comme nous l'avons signalé plus haut cette corrélation possède un biais statistique intrinsèque qui incite à la considérer avec précaution. En effet, pour que la mesure de la variabilité articuloire prenne en compte le fait que naturellement la variation absolue du nombre de contacts croît avec le nombre moyen de contacts, nous avons été

amenés à normaliser cette variabilité par le nombre moyen de contacts. On corrèle donc une variable x (le nombre moyen de contacts) avec une variable k/x (la variabilité normalisée) ce qui induit de facto un résultat négatif. Effectivement, un ensemble de tests effectués sous MATLAB sur des distributions purement aléatoires du nombre de contacts et de la variabilité normalisée, nous ont montré qu'un coefficient de -0.691 est susceptible d'être expliqué par ce seul biais statistique.

4. DISCUSSION

L'objectif de la présente étude était de vérifier le phénomène selon lequel des locuteurs avec des palais plats dans le plan coronal ont moins de variabilité articulatoire que d'autres locuteurs avec des palais plus arqués, et de trouver une explication pour ce phénomène.

Les résultats des simulations indiquent que la variabilité acoustique dépend de la forme du conduit vocal telle qu'elle est déterminée par la forme du palais dans le plan coronal. Si la forme du conduit vocal dans le plan coronal est triangulaire, un changement de la position de la langue n'a pas des conséquences aussi importantes pour l'acoustique que si le conduit vocal est en forme de quadrilatère.

Les résultats de l'étude expérimentale basée sur 16 locuteurs ont démontré que les locuteurs avec des palais plats ont, en général, moins de variabilité articulatoire que les locuteurs avec des palais en forme d'arche.

En associant ces deux résultats, celui des simulations et celui de l'étude expérimentale, on pourrait suggérer que, même si la variabilité articulatoire est différente selon que les locuteurs ont des palais plats dans le plan coronal ou des palais arqués, la variabilité acoustique pourrait rester la même. En effet, nous avons montré par nos simulations avec le modèle de langue que la variabilité articulatoire des locuteurs avec des palais arqués n'a pas des conséquences aussi importantes pour la variabilité acoustique que celle des autres locuteurs. Pour les deux catégories de locuteurs, il se pourrait donc que la variabilité articulatoire soit adaptée par rapport à une contrainte de variabilité maximale dans le domaine acoustique et donc dans le domaine perceptif.

Ainsi, nos résultats, tant expérimentaux que de modélisation, soutiennent l'hypothèse que le contrôle de la précision du geste articulatoire pourrait être spécifiquement adapté à la morphologie intrinsèque du conduit vocal de chaque locuteur, de façon à préserver l'objectif ultime de la production de la parole, c'est-à-dire la qualité perceptive du son produit. Dans le célèbre débat sur la caractérisation plutôt acoustique ou plutôt gestuelle de l'espace de la tâche de la production de la parole (cf. par exemple Liberman et al. [5], Stevens [13], Lindblom [6], Fowler [3]), nos résultats plaident donc en faveur de la primauté de la composante acoustique sur la composante gestuelle (cf. Perrier [10] pour un tutorial sur cette question).

Remerciements : Cette étude est menée dans le cadre du projet POPAART financé par le CNRS et le Ministère des Affaires Étrangères français, et par la Deutsche Forschungsgemeinschaft (projet PO 334/4-1). Merci à O.Panzyga de l'université Humboldt et à A.Busler du ZAS Berlin pour la segmentation acoustique et les mesures des palais, aux locuteurs allemands du ZAS à Berlin, de l'Institut für Phonetik und Sprachliche Kommunikation de la LMU à Munich, et de l'IPDS de la CAU Kiel, et aux locuteurs norvégiens

de l'Institut for lingvistiske og nordiske studier à l'université d'Oslo, et tout spécialement à I.Moen. Pour l'aide technique merci à J.Dreyer et D.Pape.

RÉFÉRENCES

- [1] Pierre Badin and Gunnar Fant. Notes on vocal tract computation. *STL-QPSR*, 2-3 :53-108, 1984.
- [2] Jana Brunner, Susanne Fuchs, and Pascal Perrier. The influence of the palate shape on articulatory token-to-token variability. *ZAS Papers in Linguistics*, 42 :43-66, 2005.
- [3] Carol A. Fowler. Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, 99(3) :1730-1741, 1996.
- [4] John M. Heinz and Kenneth N. Stevens. On the relations between lateral cineradiographs, area functions, and acoustic spectra of speech. In *Proceedings of the Fifth International Congress of Acoustic*, page A44. Liège, 1965.
- [5] Alvin M. Liberman, Franklin S. Cooper, Donald P. Shankweiler, and Michael Studdert-Kennedy. Perception of the speech code. *Psychological Review*, 74 :431-461, 1967.
- [6] Björn Lindblom. Phonetic invariance and the adaptive nature of speech. In Ben A.G. Elsendoom and Herman Bouma, editors, *Working Models of Human Perception*, pages 139-173. Academic Press, London, 1988.
- [7] Christine Mooshammer, Pascal Perrier, Christian Geng, and Daniel Pape. An emma and epg study on token-to-token variability. *AIPUK*, 36 :47-63, 2004.
- [8] Yohan Payan and Pascal Perrier. Synthesis of v-v sequences with a 2d biomechanical tongue model controlled by the equilibrium point hypothesis. *Speech Communication*, 22 :185-205, 1997.
- [9] Joseph S. Perkell. Articulatory processes. In William J. Hardcastle and John Laver, editors, *Handbook of Phonetic Sciences*, pages 333-370. Blackwell Handbooks in Linguistics, Oxford and Cambridge, Massachusetts, 1997.
- [10] Pascal Perrier. Control and representations in speech production. *ZAS Papers in Linguistics*, 40 :190-132, 2005.
- [11] Pascal Perrier, Louis J. Boë, and Rudolph Sock. Vocal tract area function estimation from midsagittal dimensions with ct scans and a vocal tract cast : modelling the transition with two sets of coefficients. *Journal of Speech and Hearing Research*, 35 :53-67, 1992.
- [12] Pascal Perrier, Yohan Payan, Majid Zandipour, and Joseph Perkell. Influences that shape tongue biomechanics on speech movements during the production of velar stop consonants : A modeling study. *Journal of the Acoustical Society of America*, 114(3) :1582-1599, 2003.
- [13] Kenneth N. Stevens. The quantal nature of speech : evidence from articulatory-acoustic data. In Edward E. David and Peter B. Denes, editors, *Human Communication : A Unified View*, pages 51-66. McGraw-Hill, New York, 1972.