

La mobilisation du tractus vocal est-elle variable selon les langues en parole spontanée ?

Christine Meunier¹, Morgane Peirollo^{1,2}, Brigitte Bigi¹

(1) Aix Marseille Univ, CNRS, LPL, Aix-en-Provence, France

(2) Aix Marseille Univ, Institut Convergence ILCB, Marseille, France

christine.meunier@univ-amu.fr, morgane.peirollo@gmail.com,
brigitte.bigi@univ-amu.fr

RÉSUMÉ

L'objectif de ce travail est de quantifier les positions articulatoires théoriques lors de la production de la parole spontanée dans trois langues. Chaque langue dispose d'un inventaire phonologique spécifique. Mais ces spécificités ne sont pas représentées telles quelles en parole spontanée dans laquelle les phonèmes n'ont pas tous la même fréquence d'apparition. Nous avons comparé trois langues (polonais, français et anglais américain) présentant des différences notables dans leur inventaire phonologique. Des positions articulatoires ont été calculées sur la base des fréquences des phonèmes dans chacune des trois langues dans des corpus de parole spontanée. Étonnamment, les résultats tendent à montrer que les positions articulatoires majoritaires sont très similaires dans les trois langues. Il semble ainsi que l'usage de la parole spontanée, et donc la distribution des phonèmes dans les langues, gomme les disparités des systèmes phonologiques pour tendre vers une mobilisation articulatoire commune. Des investigations plus approfondies devront vérifier cette observation.

ABSTRACT

Does vocal tract use depend on language characteristics in spontaneous speech?

The aim of this work is to quantify the theoretical articulatory positions during the production of spontaneous speech for three languages. Each language has a specific phonological inventory. However, these specificities are not represented as such in spontaneous speech in which phonemes do not have the same relative frequency. We compared three languages (Polish, French and American English) with notable differences in their phonological inventory. Articulatory positions were calculated according to phoneme frequencies in the three languages through spontaneous speech corpora. Surprisingly, the results tend to show that preferred articulatory positions are very similar in the three languages. Thus it seems that spontaneous speech production, and therefore phonemes distribution in languages, erases the disparities of phonological systems in order to provide similar articulation. Further investigation should verify this observation.

MOTS-CLÉS : Position articulatoire, langues, corpus, parole spontanée, fréquences phonétiques.

KEYWORDS : Articulatory position, languages, corpus, spontaneous speech, phoneme frequencies.

1 Introduction

L'ensemble des humains partagent un appareil vocal tout à fait semblable leur permettant de produire le langage oral, mais tous ne parlent pas la même langue. Or, chaque langue dispose d'un inventaire phonologique spécifique puisant dans l'ensemble des sons possibles que l'appareil vocal peut produire. Des tendances universelles ont pu être observées permettant de mettre en évidence la façon dont les systèmes s'organisent et se différencient (Vallée et al., 1999, Ladefoged & Maddieson, 1996). Par exemple, les dialectes arabes sont caractérisés par des positions très arrières (consonnes uvulaires, pharyngales ou glottales) ; le suédois possède de nombreuses voyelles antérieures ; le français n'a pas de consonnes très postérieures (sauf /R/, uvulaire) et un tiers des voyelles sont antérieures tandis qu'un autre tiers est arrondi. Ces particularités permettent de caractériser le système phonologique dans ces langues. Toutefois, il reste à déterminer si l'inventaire phonologique d'une langue est effectivement représentatif de la mobilisation du tractus vocal lors de la production de cette langue dans un usage courant. L'inventaire phonologique d'une langue attribue le même poids à chaque phonème et donc à chacune de ses propriétés articulatoires. Mais les phonèmes sont diversement présents dans l'usage courant de la langue orale et, en conséquence les *préférences* articulatoires d'une langue vont dépendre de la fréquence d'utilisation des phonèmes dans cette langue. Par exemple, le geste d'arrondissement en français pourrait être minoré dans la mesure où les voyelles d'arrière (arrondie) sont peu fréquentes en parole spontanée (Meunier & Espesser, 2011). Par conséquent, dans quelle mesure la mobilisation du tractus vocal dans une langue est-elle pondérée par la fréquence relative des phonèmes dans l'usage courant de la langue?

L'objectif de ce travail est donc de **quantifier les positions articulatoires théoriques** lors de la production de la parole spontanée. Nous comparons les positions articulatoires relatives aux phonèmes dans trois langues : le polonais, le français et l'anglais américain. Pour représenter les positions articulatoires, nous nous sommes inspirés de la terminologie fournie par Articulatory Phonology (AP, Browman & Goldstein, 1992). AP propose une représentation des positions articulatoires au cours de la production de la parole. La terminologie utilisée convient bien à ce travail dans la mesure où il est question de décrire précisément des positions et mouvements articulatoires plutôt que de proposer des traits distinctifs. Il n'est pas besoin en effet, dans ce travail, de distinguer les phonèmes, mais plutôt d'inventorier leurs caractéristiques articulatoires, même si elles sont redondantes. Notons que ce travail ne fournit pas d'informations sur les articulations *réelles* dans la production de la parole et elle ne fournit pas non plus d'informations sur l'articulation *dynamique* des sons. En fait, il ne s'agit pas d'une étude articulatoire. A cette étape, nous souhaitons fournir une indication sur les tendances des positions *attendues* dans les trois langues lors de leur production courante. C'est la raison pour laquelle nos résultats sont basés sur l'annotation phonétique de trois corpus de parole spontanée permettant d'obtenir les fréquences relatives de chaque phonème.

2 Systèmes phonologiques, corpus et affectation des poids articulatoires

Plusieurs critères ont déterminé le choix de ces trois langues, soit le polonais (PO), le français (FR) et l'anglais américain (AE). En premier, nous étions contraints par l'existence de corpus annotés phonétiquement pour lesquels il était possible d'obtenir la fréquence de chaque phonème. D'autre part, nous avons essayé de choisir des corpus assez similaires pour ce qui est du type de parole produite : pour les trois corpus il s'agit de parole spontanée non familière dont les conditions

représentent une forme de parole guidée (conversation téléphonique pour AE et PO, Maptask pour PO et FR). Ensuite, nous avons privilégié ces trois langues car nous disposons de locuteurs natifs phonéticiens susceptibles de pouvoir nous éclairer sur les systèmes phonologiques de leur langue et sur les réalisations effectives. Enfin, ces trois langues présentent des systèmes phonologiques différents : peu de voyelles pour le polonais en comparaison du français et de l'anglais américain; un grand nombre de fricatives et d'affriquées en polonais ; un grand nombre de voyelles arrondies en français. En conséquence, des différences phonologiques pouvant entraîner une mobilisation variée du tractus vocal. Pour chaque langue nous décrivons les caractéristiques du système phonologique, puis le corpus utilisé et enfin quelques informations sur la distribution des phonèmes.

2.1 Polonais (PO)

Le polonais (PO) possède 5 voyelles (/a/ /ε/ /i/ /u/ /ɔ/) et s'apparente ainsi au système « le plus populaire » des langues du monde (Vallée et al., 1999) avec, en plus, une voyelle centrale fermée /ɪ/. Deux voyelles nasales viennent compléter cet inventaire. Le système consonantique est plus riche et plus complexe. Il dispose d'un grand nombre de fricatives et affriquées combinant de nombreux lieux d'articulation (labiales, alvéolaires, rétroflexes palatales et vélaires), modes d'articulation (fricatives et affriquées) et modes de phonation permettant de produire 15 consonnes différentes¹. Avec 6 plosives, 3 nasales et 4 approximantes, la totalité des consonnes représente 31 segments, soit 39 phonèmes au total avec un fort déséquilibre en faveur des consonnes. Notons toutefois que l'inventaire des consonnes (notamment des affriquées) pourrait être surestimé comparé à d'autres sources².

Le corpus polonais utilisé ici est *Paralingua* (Klessa et al., 2013). Il s'agit un recueil d'enregistrements téléphoniques de dialogues dans une tâche de type *Map Task*. Des paires de locuteurs ont été enregistrés dans des séances de 30mn. Le corpus a été transcrit et aligné en utilisant le code SAMPA³. Notre étude porte sur des séquences de 2 à 4mn pour 20 dialogues, soit 145mn d'enregistrement. Nous avons retenu 39 phonèmes-type pour une production de 38.236 segments phonétiques réalisés dans cet extrait. Il s'agit, nous le verrons, d'un corpus de taille très réduite comparé à ceux utilisés pour le FR et l'AE. Notre informatrice polonaise nous a indiqué que certaines consonnes affriquées apparaissent très rarement dans le lexique. Toutefois les phonèmes du corpus sont tout à fait comparables à ceux décrits dans l'inventaire.

2.2 Français (FR)

Le français (FR) possède 11 voyelles orales dont la majorité (7) est arrondie. Trois voyelles nasales dont une arrondie viennent compléter l'inventaire (Fougeron & Smith, 1999). La particularité du français est de posséder des voyelles antérieures arrondies (/y/ par exemple) ce qui est rare dans des langues du monde. Par ailleurs, les voyelles antérieures sont majoritaires (9 sur 14). Le français compte également 20 consonnes dont 6 plosives et 6 fricatives avec alternance de voisement. Les lieux d'articulation sont répartis entre les lèvres et le voile du palais, /R/ étant la consonne la plus

¹ https://en.wikipedia.org/wiki/Polish_phonology

² <http://www.phonetics.ucla.edu/appendix/languages/polish/polish.html>

³ <https://www.phon.ucl.ac.uk/home/sampa/polish.htm>

postérieure. Trois glissantes sont également présentes aux lieux d'articulation correspondant aux 3 voyelles fermées. Le français compte donc en tout 34 phonèmes.

Le corpus français utilisé ici est le *Aix Map Task corpus* (Gorisch et al., 2014). Le protocole suivi et celui établi à Edinburgh (Anderson et al., 1992) dont l'objectif principal est de faire interagir des participants dans une tâche communicative ayant pour but la description d'un itinéraire via une carte. Quatre paires de locuteurs (2 hommes et 6 femmes) ont réalisé cette tâche pour un total de 2h18mn d'enregistrement, soit 26.706 mots. Le corpus a été transcrit et aligné. L'annotation phonétique est en SAMPA⁴. Pour notre étude nous avons utilisé 32 phonèmes-type et 96.090 segments phonétiques produits. La liste des phonèmes du corpus est très similaire à celle de l'inventaire phonologique à l'exception des voyelles /ɔ/ et /o/ qui sont regroupées dans le corpus. De même la nasale palatale /ɲ/ n'est pas représentée.

2.3 Anglais Américain (AE)

L'anglais-américain (AE) est constitué d'un système vocalique riche et complexe. Outre les caractéristiques classiques des voyelles, on note la distinction entre *tense* et *lax* ainsi que de nombreuses diphtongues. La répartition entre voyelles antérieures et postérieures est équilibrée et seules les voyelles postérieures *tense* sont arrondies. Ladefoged (1999) distingue 11 monophthongues et 3 diphtongues. Le système de consonnes comporte 9 fricatives (avec alternance de voisement) et 2 affriquées. Six plosives et 3 nasales viennent compléter l'inventaire. On compte également 4 approximantes (/l/ /r/) dont 2 glissantes (/j/ /w/). L'anglais-américain compte donc 38 phonèmes.

Le corpus anglo-américain utilisé ici est le *Buckeye Corpus of conversational speech* (Pitt et al., 2005) contenant 40 locuteurs enregistrés lors de conversations téléphoniques. Le corpus, contenant environ 300.000 mots, a été transcrit et aligné dans un codage assez spécifique décrit dans le manuel du corpus⁵. La correspondance et l'identification des voyelles n'a pas été simple mais nous avons bénéficié de l'aide de locuteurs natifs. Parmi l'ensemble des symboles décrits, nous avons retenus 39 phonèmes-types représentant 1.536.801 segments phonétiques produits dans le corpus. Là encore, on est assez proche de l'inventaire dressé plus haut. La fréquence des phonèmes présents dans ce corpus est donnée par Yang (2012).

Les limites de la comparaison des productions des trois corpus concernent évidemment la taille très hétérogène des corpus avec un nombre de phonèmes nettement plus réduit pour le polonais. Toutefois, nous avons pu observer que les distributions des phonèmes et donc leur fréquence respective dans un type de corpus sont assez stables et apparaissent très rapidement avec peu de données et peu de locuteurs. Cela est probablement dû au fait que la fréquence des phonèmes est dépendante de la fréquence des mots dont la distribution est très contrainte par le discours.

Dans la mesure où la fréquence des phonèmes pourrait être déterminante dans la représentation des positions articulatoires, nous présentons ici les phonèmes les plus fréquents (50% des occurrences) dans chacune des trois langues (table 1). Notons que les phonèmes fréquents communs aux trois langues sont la plosive alvéolaire /t/ et la voyelle moyenne antérieure /e/ ou /ɛ/.

⁴ <https://www.phon.ucl.ac.uk/home/sampa/french.htm>

⁵ <https://buckeyecorpus.osu.edu/BuckeyeCorpusmanual.pdf> (page 22)

PO	a (11%)	e (13%)	o (9%)	t (6%)	j (5%)	n (4%)	ɪ (4%)		
FR	a (11%)	R (8%)	t (7%)	d (6%)	l (5%)	ɛ (5%)	s (5%)	e (5%)	
AE	ʌ (8%)	ɪ (8%)	n (6%)	s (5%)	ɛ (5%)	t (4%)	i (4%)	r (4%)	k (4%)

TABLE 1 : Phonèmes les plus fréquents dans les trois langues totalisant 50% des occurrences de phonèmes dans les corpus respectifs

On notera également que la répartition des phonèmes dans 50% des occurrences est différentes dans les trois langues : peu de phonèmes (surtout des voyelles) pour PO ; un nombre plus important pour FR et AE avec une proportion de voyelles et consonnes plus équilibré.

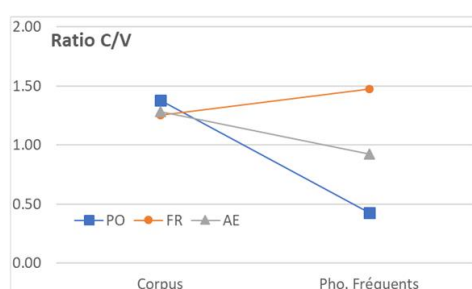


FIGURE 1 : Ratio Consonnes/Voyelles dans totalité de chacun des corpus et au sein des phonèmes les plus fréquents (50% des occurrences, voir table 1).

Ainsi, les trois langues ont un ratio C/V très similaire lorsque l'on considère la totalité des occurrences des corpus. Toutefois, si l'on ne prend en compte que les phonèmes les plus fréquents, les trois langues montrent un ratio C/V très différent (figure 1). Pour PO, 3 voyelles représentent 38% des occurrences du corpus, alors que AE et surtout FR montrent une proportion de consonnes plus importante dans les phonèmes fréquents.

2.4 Attribution des positions articulatoires

Notre objectif est de rendre compte des positions articulatoires les plus fréquentes dans chacune des langues et dans la production courante de la parole. Pour cela, il n'est pas nécessaire que ces critères articulatoires soient distinctifs, mais plutôt qu'ils expriment, le mieux possible, l'état *théorique* des articulateurs au cours de sa production. Nous précisons « théorique » car il ne s'agit pas de la réalisation effective des phonèmes, mais plutôt d'une réalisation attendue. Browman & Goldstein (1992) proposent d'exprimer la représentation phonologique des sons de la parole par un ensemble de *variables* et de *dimensions* du tractus vocal (*Tracts Variables & Dimensions*). Les dimensions précisent l'état de chaque variable. Parmi les différentes propositions des auteurs, nous avons sélectionné quatre variables (table 2) et nous avons effectué un regroupement (les variables *Tongue Tip* et *Tongue Body* ont été regroupées en *Tongue*).

<i>Tract Variables</i>	<i>Dimensions</i>
GLO - Glottis	closed, wide
VEL - Velum	closed, wide
LP - Lip Protrusion	protruded, not protruded
TCL - Tongue Constriction Location	labial, dental, alveolar, postalveolar, palatal, velar, uvular

TABLE 2 : représentation des *Tract Variables* et *Dimensions* selon AP et utilisées dans cette étude.

Pour les variables GLO, VEL et LP, l'attribution des dimensions est binaire et il n'y a pas d'ambiguïté concernant cette attribution. Pour TCL, chaque consonne se voit attribuer le lieu d'articulation qui la caractérise. L'opération est globalement assez simple excepté pour les consonnes rétroflexes de PO et AE qui ont été codées en *postalveolar* (comme il est souvent indiqué dans les inventaires phonologiques). En revanche, les voyelles sont généralement dites antérieures, centrales ou postérieures mais ne sont pas caractérisées par un lieu articulaire précis. Il a été décidé que les voyelles antérieures étaient codées *palatal* et celles d'arrière *velar*, ce qui correspond à leur position effective. Notre méthode consiste donc à assigner à chaque phonème l'ensemble des dimensions qui le caractérise. Ensuite chaque dimension se voit attribuée une valeur correspondant à la fréquence de chaque phonème dans le corpus (table 3). En additionnant l'ensemble de ces valeurs dans chaque colonne nous obtenons la fréquence d'utilisation de chaque dimension. Cette opération est répétée dans chaque langue et pour chacune des variables.

TCL	alveolar	postalveolar	palatal	velar
e			10.75	
s	2.49			
k				3.51
dz'		0.65		
n	4.08			
...				

TABLE 3 : Exemple (pour PO) de l'attribution des fréquences des phonèmes (en %) pour la variable Tongue Constriction Location (TCL). Dans le corpus PO, la fréquence de la consonne /s/ est 2.49%. Cette valeur est reportée dans la colonne *alveolar*, caractéristique de cette consonne.

3 Résultats

3.1 Glottis, Velum et Lip Rounding

Pour ces trois variables l'attribution des dimensions est binaire. GLO et VEL sont en position *closed* ou *wide*. LP est *protruded* ou *not protruded*. Pour GLO et VEL on note une impressionnante similarité pour les trois langues : une forte préférence pour la position *closed* autour de 78% en moyenne pour GLO et 89% pour VEL (figure 2). Il n'y a quasiment pas de différence entre les trois langues. Notons que la dimension *wide* pour VEL représente 15% de l'inventaire de FR, alors que les occurrences *wide* dans le corpus sont à 10.7%. Pour LP, les tendances sont également très similaires pour les trois langues avec la dimension *not protruded* très majoritaire (87% en moyenne). On note une proportion légèrement moins importante de *not protruded* pour FR (81%), probablement due au nombre important de voyelles arrondies dans le système vocalique français. Toutefois, 28% des phonèmes FR sont arrondis et cette proportion se trouve minorée dans le corpus, ce qui suppose une fréquence moins importante des voyelles arrondies en FR.

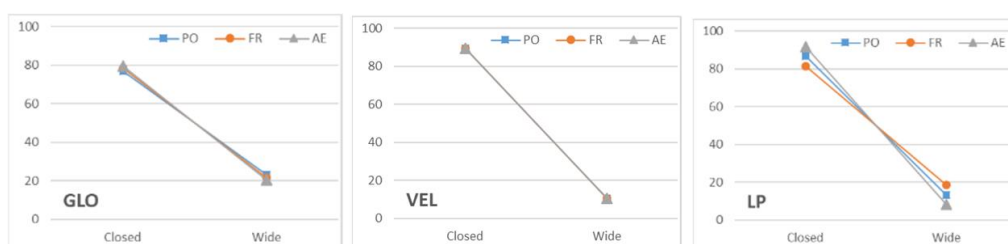


FIGURE 2 : Fréquence d'apparition des dimensions pour les trois variables GLO, VEL et LP

3.2 Tongue Constriction Location

La position de la langue lors de la constriction présente également de nombreuses similitudes dans les trois langues (figure 3). On observe une forte prépondérance des dimensions *alveolar*, *palatal* et *velar* (respectivement en moyenne 25%, 32% et 22%), la dimension *postalveolar* étant nettement moins utilisée (5%). La dimension *dental* n'est présente que pour AE et représente une faible proportion (3.5%). De même, la dimension *uvular* n'est présente qu'en français via le seul phonème /R/ qui, à lui seul, totalise 7.6% des occurrences. Ainsi, au cours de la production de parole spontanée, on observe que dans ces trois langues, la position de la langue est majoritairement positionnée dans trois régions : palais dur, alvéoles puis palais mou.

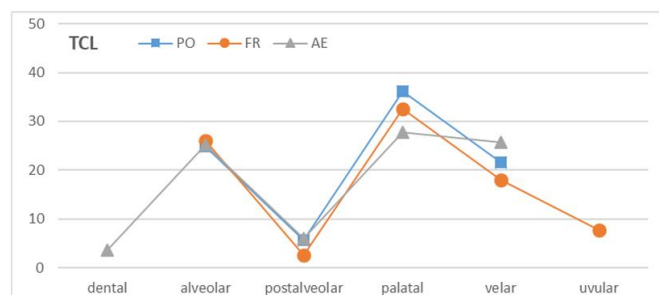


FIGURE 3 : Fréquence d'apparition des dimensions pour la variable TCL

Voyelles et consonnes sont présentées conjointement ici (figure 3) de façon à obtenir la configuration du tractus pendant la parole spontanée. Les voyelles n'étant présentes que sur les dimensions *palatal* et *velar*, il est clair que ces positions sont très sollicitées. Nous avons donc distingué voyelles et consonnes de façon à mieux cerner les parts respectives de ces deux types de phonèmes dans la configuration du tractus. La dimension *alveolar* est nettement plus sollicitée dans la production des consonnes pour les trois langues, ceci au détriment des autres dimensions (figure 4).

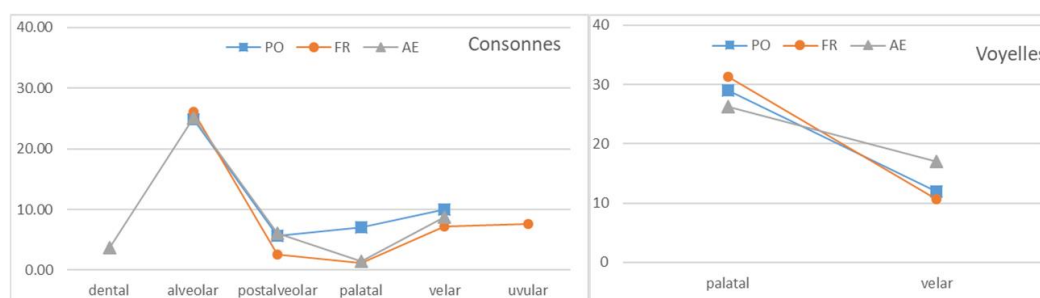


FIGURE 4 : Fréquence d'apparition des dimensions pour la variable TCL. Pour les consonnes (à gauche) et pour les voyelles (à droite)

Certes la dimension *alveolar* est également importante dans les inventaires phonologiques des trois langues, mais de façon variable : 26% pour PO, 19% pour FR et 15% pour AE avec, donc, une forte proportion pour PO. Mais si l'on regarde les phonèmes les plus fréquents (Table 1) dans chacune des trois langues, on trouve 1 alvéolaire en PO (/t/=6%), 4 alvéolaires en FR (/t/ /d/ /l/ /s/=23%) et 3 en AE (/n/ /s/ /t/=16%). La forte proportion des alvéolaires dans l'inventaire phonologique de PO est donc contrebalancée par sa faible représentation dans les phonèmes fréquents. On observe donc que les disparités observées dans les inventaires phonologiques semblent être *gommés* par l'usage des phonèmes dans la production courante. La dimension *alveolar* revêt probablement un caractère

spécifique car il s'agit du lieu d'articulation le plus fréquemment présent dans les inventaires phonologiques des langues du monde (position alvéodentale=15.3%, Vallée et al. 1999).

La dimension *palatal* est majoritaire dans la production des voyelles pour les trois langues (figure 4, droite). S'il est vrai que les voyelles antérieures sont majoritaires dans l'inventaire phonologique FR, ça n'est pas le cas pour PO ni pour AE. La préférence pour la dimension *palatal* dans ces deux langues est donc due à la fréquence des voyelles antérieures dans les corpus (table 1). Là encore on observe que les différences des inventaires phonologiques semblent être compensées par l'usage des sons de chaque langue tendant vers des positions articulatoires communes.

4 Discussion

Dans ce travail nous avons cherché à rendre compte de la configuration *théorique* du tractus vocal au cours de la production de parole spontanée et dans trois langues différentes. Les trois langues utilisées ici PO, FR et AE présentent des spécificités dans leur inventaire phonologique : un système vocalique riche et complexe en AE et réduit pour PO ; des voyelles antérieures plus nombreuses pour FR, idem pour l'arrondissement et la nasalité ; un système consonantique riche en fricatives et affriquées pour AE et surtout PO. Malgré ces spécificités, on remarque une étonnante similitude concernant la fréquence des positions articulatoires pour les trois langues. Au cours de la production de parole spontanée, nous avons pu observer qu'en moyenne 78% des phonèmes sont produits avec la glotte fermée, 89% avec le voile du palais relevé et 13% avec un arrondissement des lèvres. Ces fréquences moyennes sont extrêmement proches pour chacune des langues. De même, la position de la langue lors de la constriction est majoritairement positionnée dans trois régions : palais dur, alvéoles puis palais mou. Là encore la fréquence des positions articulatoires est très proche dans les trois langues avec une prédominance de l'utilisation de la position alvéolaire dans la production des consonnes et de la position palatale pour les voyelles.

On retrouve donc une très forte similarité de la mobilisation du tractus vocal dans les trois langues et ce, malgré des différences notables dans la configuration des trois inventaires phonologiques. Ces tous premiers résultats tendent à montrer que, quelques soient les disparités et spécificités des systèmes phonologiques, la structure lexicale et l'usage du lexique dans le discours spontané tend à gommer ces différences, suggérant des contraintes biomécaniques et/ou phonologiques favorisant des configurations préférentielles du tractus vocal.

Nous devons toutefois indiquer que ce travail préliminaire présente plusieurs limites : 1/ nous avons utilisé des corpus de tailles très hétérogènes (notamment le corpus PO est probablement trop restreint pour représenter fidèlement les distributions de phonèmes en parole spontanée) ; 2/ le lien entre les inventaires phonologiques des trois langues et l'annotation phonétique des corpus a parfois été très difficile à établir et il faudrait probablement utiliser des corpus annotés de façon plus homogène ; 3/ il serait également nécessaire d'obtenir une description précise des réalisations articulatoires des langues de façon à attribuer correctement les dimensions ; 4/ il sera tout à fait nécessaire de poursuivre ce travail avec un nombre plus important de langues présentant des systèmes phonologiques plus diversifiés. Enfin, dans une étape ultérieure, nous envisageons de mettre en correspondance, au travers des corpus, ces positions attendues avec les réalisations effectives des segments phonétiques.

Références

- ANDERSON, A., BADER, M., BARD, E., BOYLE, E., DOHERTY, G., GARROD, S., ISARD, S., KOWTKO, J., MCALLISTER, J., MILLER, J., SOTILLO, C., THOMPSON, H. S., AND WEINERT, R. (1991). The HCRC Map Task corpus. *Language and Speech*, 34:351–366. DOI: [10.1177/002383099103400404](https://doi.org/10.1177/002383099103400404)
- BROWMAN, C. P., & GOLDSTEIN, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49 (3-4), 155–180.
- FOUGERON, C., SMITH, C.L. (1999). French, on *The Handbook of the International Phonetic Association*, Cambridge University Press, 78-81.
- GORISCH, J., ASTÉSANO, C., BARD, E. G., BIGI, B., & PRÉVOT, L. (2014). Aix Map Task corpus: The French multimodal corpus of task-oriented dialogue. *LREC Proceedings*, 5p. HAL: [hal-02158880](https://hal.archives-ouvertes.fr/hal-02158880).
- KLESSA, K., WAGNER, A., OLEŚKOWICZ-POPIEL, M., & KARPIŃSKI, M. (2013). Paralingua – A New Speech Corpus for the Studies of Paralinguistic Features. *Procedia - Social and Behavioral Sciences*, 95, 48–58. DOI: [10.1016/j.sbspro.2013.10.621](https://doi.org/10.1016/j.sbspro.2013.10.621).
- LADEFOGED, P. (1999). American English, on *The Handbook of the International Phonetic Association*, Cambridge University Press, 41-44.
- LADEFOGED, P. & MADDIESON, I. (1996). *The Sounds of the World's Languages*, Wiley, 450 p.
- MEUNIER, C., & ESPESER, R. (2011). *Is Phoneme Inventory a Good Predictor for Vocal Tract Use in Casual Speech?* 1370–1373. HAL: [hal-01514696](https://hal.archives-ouvertes.fr/hal-01514696)
- PITT, M. A., JOHNSON, K., HUME, E., KIESLING, S., & RAYMOND, W. (2005). The Buckeye corpus of conversational speech: Labeling conventions and a test of transcriber reliability. *Speech Communication*, 45(1), 89–95. DOI: [10.1016/j.specom.2004.09.001](https://doi.org/10.1016/j.specom.2004.09.001)
- VALLÉE, N., BOË, L.-J., & STEFANUTO, M. (1999). Typologies phonologiques et tendances universelles. Approche substantialiste. *Linx. Revue des linguistes de l'université Paris X Nanterre*, 11, 31–54. DOI: [10.4000/linx.863](https://doi.org/10.4000/linx.863).
- YANG, B. (2012). [Reduction and Frequency Analyses of Vowels and Consonants in the Buckeye Speech Corpus](#), *Phonetics and Speech Sciences* (말소리와 음성과학), Volume 4, Issue 3, p.75-83.