



**HAL**  
open science

## Rôle de l'activité laryngale dans la production des consonnes d'arrière en arabe levantin

Jalal Al-Tamimi

► **To cite this version:**

Jalal Al-Tamimi. Rôle de l'activité laryngale dans la production des consonnes d'arrière en arabe levantin. 35èmes Journées d'Études sur la Parole (JEP 2024) 31ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2024) 26ème Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2024), Jul 2024, Toulouse, France. pp.521-531. hal-04623113

**HAL Id: hal-04623113**

**<https://inria.hal.science/hal-04623113>**

Submitted on 1 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Rôle de l'activité laryngale dans la production des consonnes d'arrière en arabe levantin

Jalal Al-Tamimi<sup>1</sup>

(1) Université Paris Cité, CNRS, Laboratoire de linguistique formelle, UMR 7110, 5 Rue Thomas Mann, 75013, Paris, France

jalal.al-tamimi@u-paris.fr

## RÉSUMÉ

---

Cette étude examine le rôle de l'activité laryngale dans la production des consonnes d'arrière en arabe levantin. 26 mesures incluant la hauteur du larynx (HL), le contact de glotte (quotient fermé ; QF) et la pente spectrale (PS) ont été obtenues de données d'électroglottographies et d'acoustiques synchronisées. À partir des classifications via des forêts aléatoires (Random Forests), sept mesures ont été identifiées comme les plus importantes pour discriminer entre les six classes. Ensuite, une modélisation via des Régressions Additives à Effets-Mixtes montre que les consonnes pharyngales sont associées à  $\uparrow$ HL,  $\uparrow$ QF et  $\uparrow$ PS, résultant d'une différence de la saillance spectrale causée par une constriction épilaryngale. Les consonnes pharyngalisées induisent des traits  $\downarrow$ HL,  $\downarrow$ QF et  $\downarrow$ PS causés par une fermeture abrupte de la glotte ; les consonnes uvulaires induisent des traits  $\uparrow$ HL,  $\downarrow$ QF et  $\downarrow$ PS. Ces changements sont corrélés avec le trait [+Constricted Glottis] et suivent les prédictions du Laryngeal Articulator Model.

## ABSTRACT

---

### Role of the laryngeal activity during the production of back consonants in Levantine Arabic

This study examines the role of the laryngeal activity during the production of back consonants in Levantine Arabic. 26 measures including Larynx Height (LH), glottal contact (Closed Quotient; CQ) and Spectral Tilt (ST) were obtained from synchronised electroglottographic and acoustic data. Using classifications via Random Forests, seven measures were identified as the most important to discriminate between the six classes. Next, modelling the data via Generalised Additive Mixed-Models revealed that pharyngeal consonants to be associated with an  $\uparrow$ LH, a  $\uparrow$ CQ and an  $\uparrow$ ST, resulting from a difference in spectral saliency due to an epilaryngeal constriction. Pharyngealised consonants induced the features  $\downarrow$ LH,  $\downarrow$ CQ and  $\downarrow$ ST caused by an abrupt closure of the glottis. Uvular consonants show the features  $\uparrow$ LH,  $\downarrow$ CQ and  $\downarrow$ ST. These changes are correlated with the feature [+Constricted Glottis] and follow predictions of the Laryngeal Articulator Model.

---

**MOTS-CLÉS** : Epilarynx ; Electroglottographie ; Acoustique ; Qualité de voix ; Consonnes d'arrière ; Arabe levantin.

**KEYWORDS** : Epilarynx; Electroglottography; Acoustics; Voice quality; Back consonants; Levantine Arabic.

---

# 1 Introduction

Les consonnes d'arrière, ou « gutturales », en arabe sont généralement considérées comme faisant partie d'une classe naturelle à cause d'alternance phonologique particulière et/ou une réalisation dans une zone oro-sensorielle commune tout au long de la cavité pharyngale (McCarthy, 1994; Sylak-Glassman, 2014). Cette classe de consonnes d'arrière est traditionnellement composée de consonnes pharyngales /ħ ʕ/ (ou épilaryngales /ʔ ʕ ɦ/) et uvulaires /χ ʁ q/ (McCarthy, 1994). Les consonnes pharyngalisées /tʕ dʕ ðʕ sʕ/ (Sylak-Glassman, 2014) et/ou glottales /h ʔ/ (Zawaydeh, 1999) sont également considérées comme en faisant partie, à cause d'une similarité du lieu (et non pas de degré) de constriction entre les consonnes pharyngalisées et pharyngales (Laufer & Baer, 1988) ou à cause de la similarité d'augmentation du premier formant (F1) associée aux consonnes d'arrière en comparaison avec les consonnes alvéolaires (Zawaydeh, 1999). En suivant les prédictions du Laryngeal Articulator Model (LAM, Esling *et al.*, 2019), les consonnes épilaryngales sont produites dans la partie inférieure de l'épilarynx, les consonnes uvulaires dans la partie supérieure et les consonnes pharyngales/pharyngalisées entre les deux. L'élévation du larynx a pour rôle de faciliter la constriction de l'épilarynx et est associée à une réalisation épilaryngale (Esling *et al.*, 2019).

Utilisant l'échographie de la langue, Al-Tamimi & Palo (2023) ont montré qu'il existe une similarité articulatoire entre les consonnes d'arrière concernant les parties de la langue impactées : une dépression de la partie antérieure de la langue est observée avec une élévation et rétraction variable du dos et de la racine de la langue, respectivement. De plus, en utilisant les prédictions des Modèles GAMMs (Régressions Généralisées Additives à Effets-Mixtes) qui s'étendent au-delà des zones observables à partir de l'échographie de la langue, Al-Tamimi & Palo (2023) ont montré que les consonnes d'arrière induisent une élévation variable du larynx. En effet, l'activité laryngale est souvent négligée dans la littérature décrivant la réalisation des consonnes d'arrière. Néanmoins, quelques études ont examiné le rôle du larynx en arabe. Heselwood (2007) a montré l'existence d'une voix craquée variable quantifiée via la mesure  $H1-H2$ <sup>1</sup> dans les réalisations variables de /ʕ/ dans les dialectes arabes. Al-Tamimi & Heselwood (2011) ont montré une position plus élevée du larynx lors de la production des consonnes pharyngalisées en utilisant la vidéofluoroscopie. Heselwood & Al-Tamimi (2011) ont également montré une élévation du larynx et une voix craquée dans les consonnes pharyngales en utilisant la vidéofluoroscopie et les analyses acoustiques de  $H1-H2$ . Enfin et plus récemment, Al-Tamimi (2015, 2017) a montré une baisse générale de la pente spectrale comme corrélat secondaire de la voix tendue associée aux consonnes pharyngalisées en arabe jordanien et en arabe marocain. Les résultats de ces études montrent que l'activité laryngale peut être utilisée afin de valider les prédictions du LAM. Comme indiqué ci-dessus, le LAM prédit qu'une constriction extrême de l'épilarynx entraîne une élévation du larynx, induisant une voix tendue, craquée et/ou laryngalisée, associée à une rétraction importante du dos et de la racine de la langue (Esling *et al.*, 2019). Ces changements doivent être vue comme combinés et non distincts l'un de l'autre qui d'après Sylak-Glassman (2014) contribuent à la caractérisation de la classe des consonnes d'arrière dans la grande majorité des langues.

Ces observations ont motivé l'étude que nous avons menée. Notre but ici est d'examiner le rôle de l'activité laryngale dans la production des consonnes d'arrière en utilisant une combinaison de mesures articulatoires obtenues à partir de l'électroglottographie (EGG) et de mesures acoustiques de la qualité de voix. Cette combinaison nous permettra d'évaluer la complémentarité entre corrélats articulatoires et acoustiques pour quantifier les similarités et différences entre ces consonnes d'arrière.

---

1.  $H1$  = premier harmonique ;  $H2$  = deuxième harmonique

## 2 Corrélats articulatoires et acoustiques de qualité de voix

### 2.1 Corrélats articulatoires

Kuang & Keating (2014) ont décrit les corrélats articulatoires obtenus de l'EGG de la voix tendue (en comparaison avec une voix relâchée).  $\uparrow$ QF (Quotient Fermé, % contact) causée par une augmentation du contact dans la glotte.  $\downarrow$ PIC (amplitude du pic positif de la dérivée du signal EGG) et  $\downarrow$ PDC (amplitude du pic négatif de la dérivée d'EGG) sont causées par une fermeture abrupte ou rapide de la glotte induisant une augmentation d'énergie dans les hautes fréquences (Michaud, 2004; Kuang & Keating, 2014). Finalement,  $\downarrow$ SQ (quotient de vitesse d'ouverture/fermeture calculé comme le ratio de la phase de fermeture et la phase d'ouverture) est à observer dans la voix tendue causée par une phase de fermeture plus courte et une phase d'ouverture plus longue (Holmberg *et al.*, 1988; Kuang & Keating, 2014). Enfin, une HL (Hauteur du Larynx) variable est observée en fonction de la catégorie examinée. En suivant les prédictions du LAM, les consonnes pharyngales auront un  $\uparrow$ HL, avec une position variable pour les consonnes pharyngalisées et uvulaires.

### 2.2 Corrélats acoustiques

Pour ce qui est des mesures acoustiques, Kuang & Keating (2014) et Al-Tamimi (2017) ont montré qu'une voix tendue, craquée et/ou laryngalisée conduit à un abaissement global de la pente spectrale, en suivant le modèle psychoacoustique de la qualité de voix (Garellek *et al.*, 2016; Kreiman *et al.*, 2021). Ces changements induisent un abaissement des mesures suivantes<sup>2</sup> :  $\downarrow$ H1-H2,  $\downarrow$ H1-A1,  $\downarrow$ H1-A2,  $\downarrow$ H2-H4,  $\downarrow$ H4-H2kHz,  $\downarrow$ H2kHz-H5kHz,  $\downarrow$ HNR et  $\downarrow$ SHR (Fulop *et al.*, 1998; Guion *et al.*, 2004; Aralova *et al.*, 2011; Kuang & Keating, 2014; Al-Tamimi, 2015; Garellek *et al.*, 2016; Al-Tamimi, 2017). De plus, un renforcement de la saillance spectrale entre F1 et F2, quantifié via  $\downarrow$ A1-A2, est observé. Enfin, et comme rapporté précédemment dans Al-Tamimi (2015, 2017) pour les consonnes pharyngalisées en arabe jordanien et marocain, nous prédisons que ce type de voix induit une augmentation de l'énergie spectrale autour de F3 et au-delà, qui peut être causée par deux facteurs. Le premier peut résulter d'une fermeture abrupte de la glotte avec augmentation de l'énergie spectrale autour de F3 avec comme conséquence  $\downarrow$ H1-A3,  $\downarrow$ A1-A3 et  $\downarrow$ A2-A3 (Hanson & Chuang, 1999; Hanson *et al.*, 2001; Al-Tamimi, 2015, 2017). Le second est dû à une qualité vocale renforcée causée par une constriction épilaryngale extrême, comme dans le chant (ex. opéra), avec un renforcement bien marqué de l'énergie spectrale autour de F3, F4 et F5 comme conséquences du « formant du chanteur ». Ce qui induit une saillance spectrale autour de F3, F4 et F5 associée à un changement de la saillance spectrale autour de F1 et F2, induisant  $\uparrow$ H1-A3,  $\downarrow$ A1-A3,  $\downarrow$ A2-A3 (Titze & Story, 1997; Moisik & Esling, 2010; Story, 2019). Tous ces changements doivent être évalués en comparaison avec une voix modale, où les changements de pentes spectrales sont minimaux.

Les prédictions sont donc qu'un abaissement global de la pente spectrale dans les consonnes d'arrière est à observer bien qu'on s'attendrait à des différences liées au degré de constriction épilaryngale. On propose donc que les consonnes pharyngales soient associées à une baisse globale de la pente spectrale avec une augmentation de l'énergie dans les hautes fréquences causée par une constriction épilaryngale extrême. Les consonnes pharyngalisées causeront quant à elles une baisse globale de la pente spectrale avec une augmentation de l'énergie dans les hautes fréquences moins marquée que dans les consonnes

2. Avec H1, 2 ou 4 = Harmonique 1, 2 ou 4; H2kHz ou 5kHz = Harmonique proche des 2 ou 5 kHz; A1-3 = Harmonique proche de F1, F2 ou F3; HNR = Ratio d'énergie dans les harmoniques et le bruit; SHR = ratio d'énergie subharmonique

pharyngales causé par une constriction laryngale. Enfin, les consonnes uvulaires montreront des similarités avec les deux autres classes avec des conséquences intermédiaires. Cette étude examine la complémentarité entre corrélats articulatoires et acoustiques permettant la discrimination entre les consonnes d'arrière en comparaison avec les consonnes alvéolaires vélares et glottales. Nous montrerons s'il existe une similarité entre les consonnes d'arrière en quantifiant la gradience de leurs impacts sur les mesures articulatoires et acoustiques.

### 3 Corpus et méthode

#### 3.1 Participants et enregistrements

Dix locuteurs urbains parlant l'arabe levantin<sup>3</sup> (5 hommes et 5 femmes), âgés entre 25 et 45 ans (moyenne 31,8; écart type = 6,9) ont été recrutés. Trois types de données ont été enregistrés simultanément : échographie de la langue (voir [Al-Tamimi & Palo, 2023](#)), électroglottographie (système à deux-canaux EG2-PCX2 de Glottal Entreprise avec hauteur et contact) et audio (avec un microphone Roland Pro connecté à une carte son Roland Quad-Capture, taux d'échantillonnage à 44.1 kHz, et quantification 16-bits avec distance d'environ 15 cm de la bouche du participant). Les trois systèmes étaient reliés à un système de synchronisation et d'alignement automatique des signaux (pour plus de détails, voir [Wrench & Scobbie, 2008](#)). Après l'installation des électrodes et du casque de stabilisation de la sonde de l'échographie de la langue, nous avons vérifié que les capteurs d'EGG étaient bien positionnés sur le larynx de chaque participant et que le traçage automatique du second canal de l'EGG, qui capte les mouvements verticaux du larynx, était au centre des 15 témoins lumineux.

#### 3.2 Corpus et segmentation

Les participants ont produit une liste de mots réels (près de 75% du corpus) et des pseudo-mots (25% du corpus) répétés trois fois, dans la séquence /ʔV:CV:/, où V: = les voyelles /i: a: u:/ symétriques; C = toutes les consonnes possibles en arabe levantin et d'autres variétés arabes = /b t d m n r f θ ð s z ʃ ʒ l w j k g x ɣ q tʃ dʃ ðʃ sʃ zʃ lʃ h ʕ ʔ h/ (avec un nombre maximal théorique de 2790 items : 31 C \* 3 V \* 3 répétitions \* 10 locuteurs). Ensuite, 21 consonnes ont été retenues pour la suite d'analyse et ont été divisées en six classes, avec un total de 1940 items (21 C \* 3 V \* 3 répétitions \* 10 locuteurs + répétitions additionnelles) : **Alvéolaires** ⇒ /t d ð s z l/, **Vélares** ⇒ /k g x ɣ/, **Uvulaire** ⇒ /q/, **Pharyngalisées** ⇒ /tʃ dʃ ðʃ sʃ zʃ lʃ/, **Pharyngales** ⇒ /h ʕ/ et **Glottales** ⇒ /h ʔ/.

Les données ont été translittérées manuellement utilisant la convention ATR ([Al-Tamimi et al., 2022](#)) et alignées avec le système MAUS ([Schiel, 2015](#)) implémenté dans PraatAlign ([Lubbers & Torreira, 2013](#)). La segmentation automatique a été corrigée manuellement afin de prévenir les erreurs potentielles (suivant les critères définis dans [Al-Tamimi, 2017](#) et [Al-Tamimi & Khattab, 2018](#)).

#### 3.3 Extraction des mesures

Les mesures articulatoires et acoustiques ont été obtenues en utilisant EGGworks ([Tehrani, 2020](#)) pour l'EGG et VoiceSauce ([Shue et al., 2011](#)) pour les mesures de la pente spectrale avec les

---

3. L'arabe levantin regroupe les parlés urbains de la Palestine, Jordanie, Syrie et du Liban. Même si les participants viennent de différentes régions, ils partagent tous des traits spécifiques aux parlars syro-libanais, à savoir la réalisation de /q/ en Arabe Standard Moderne (ASM) comme /ʔ/, de /θ ð ðʃ/ en ASM comme /t d dʃ/ ou /s z zʃ/, de /tʃ/ en ASM comme /ʒ/ et de /ɣ u/ en ASM comme /x ɣ/ ([Embarki, 2008](#))

paramètres spécifiques à Praat (Boersma & Weenink, 2020). Les seuils de  $f_0$  et de formants ont été adaptés à chaque locuteur (suivant Al-Tamimi, 2017, Al-Tamimi & Khattab, 2018 et Al-Tamimi, 2022) et les mesures de  $f_0$  et des quatre premiers formants ont été manuellement vérifiées. Toutes les mesures décrites ci-dessous ont été initialement calculées avec un pas de déplacement d'1 ms. Ensuite, 11 intervalles moyennés sur la durée de chaque segment de la séquence VCV ont été extraits.

Pour les mesures articulatoires, un total de sept mesures a été calculé. La HL a été obtenue du second canal de l'EGG sans normalisation. Ensuite, nous avons calculé les mesures de contact telles que le QF en suivant la méthode hybride (Howard, 1995), le PIC (Michaud, 2004), le PDC, la vitesse d'ouverture et de fermeture, ainsi que leur ratio (voir Kuang & Keating, 2014).

Pour les mesures acoustiques, un total de 19 mesures a été obtenu en suivant deux approches. Pour la première, nous avons calculé toutes les mesures décrites dans le modèle psychoacoustique de la qualité de voix (Garellek *et al.*, 2016; Kreiman *et al.*, 2021) comme : la pente spectrale normalisée suivant Iseli *et al.* (2007) sur diverses bandes d'harmoniques (ex.  $H1^*-H2^*$ ,  $H2^*-H4^*$ ,  $H4^*-H2\text{kHz}^*$ ,  $H2\text{kHz}^*-H5\text{kHz}^*$ ,  $H1^*-A1^*$ ,  $H1^*-A2^*$  et  $H1^*-A3^*$ ); les HNR sur plusieurs bandes; SHR; CPP; soe<sup>4</sup>. Pour la seconde, nous avons calculé les trois mesures d'amplitude dans les hautes fréquences (ex.  $A1^*-A2^*$ ,  $A1^*-A3^*$  et  $A2^*-A3^*$ ) en suivant les prédictions avancées dans Al-Tamimi (2015, 2017).

### 3.4 Approche statistique

Afin d'évaluer la contribution des mesures articulatoires et acoustiques pour montrer les similarités entre les consonnes d'arrière et les différences avec les autres contextes, nous avons procédé à deux types d'analyse. Premièrement, une classification avec les forêts aléatoires (Random Forests) a permis d'évaluer le taux de succès de discrimination suivi de l'évaluation du poids relatif pour chaque variable. Pour ce faire, nous avons entraîné plusieurs forêts aléatoires sur 66.7% des données, avec les mesures articulatoires, acoustiques et leur combinaison afin d'évaluer la robustesse de l'approche. Ces entraînements ont été effectués sur la partie V1, C ou V2 (voyelle précédente, consonne médiane et voyelle suivante, respectivement). Nous avons utilisé deux types de réponses : les six classes (alvéolaires, vélaires, uvulaire, pharyngalisées, pharyngales et glottales), ou deux classes (gutturale = uvulaire, pharyngalisées et pharyngales vs non-gutturale = alvéolaires, vélaires et glottales). Nous avons utilisé l'approche de TidyModels (Kuhn & Wickham, 2020) et la librairie ranger (Wright & Ziegler, 2017). Chaque forêt a été entraînée avec 2000 arbres; le nombre de variables que chaque forêt pouvait utiliser dans l'apprentissage représentait la racine carrée arrondie du nombre total de prédicteurs (3 pour l'EGG; 4 pour l'acoustique; 5 pour l'EGG+acoustique). Une validation croisée avec 10 plis a été effectuée afin d'obtenir un apprentissage robuste sur les données d'entraînement. A la fin, des prédictions sur les données de test (33.3%) ont été effectuées afin d'obtenir le taux de classification et le poids relatif de chaque variable via des tests de permutation (Strobl *et al.*, 2009).

Ensuite, à partir des poids relatifs identifiés dans la classification combinant les mesures articulatoires et acoustiques, nous avons choisi les trois prédicteurs les plus importants pour chacun des V1, C et V2. Au total, nous avons identifié six mesures partagées en plus de la HL. Ces sept mesures ont ensuite été modélisées avec des GAMMs via la librairie mgcv (Wood, 2017). Le but de ces modélisations est de quantifier la dynamique des trajectoires tout au long de la séquence VCV. Pour ce faire, nous avons utilisé le prédicteur comme variable réponse, et l'interaction entre contexte et voyelle (variables ordonnées) et le sexe du locuteur comme variables fixes. Deux trajectoires représentant les 11 intervalles et la position dans la syllabe (1, 2 et 3 pour V1, C, et V2) ont été lissées

4. HNR = Ratio d'énergie dans les harmoniques et dans le bruit; CPP = amplitude du pic cepstral; soe = Force d'excitation

avec une pénalisation en régression cubique, utilisant une interaction entre les deux trajectoires et avec ajustements par les variables fixes. Comme variables aléatoires, nous avons utilisé le locuteur et le mot ajustés par les deux trajectoires, ainsi que par l'interaction entre contextes et voyelles (pour le locuteur), et le sexe du participant (pour le mot). Nos modèles ont suivi l'approche maximale en régression à effets mixtes (Barr *et al.*, 2013); permettant une normalisation entre locuteurs, avec une approche similaire à celle utilisée dans Al-Tamimi & Palo (2023). Nous avons vérifié la structure de nos modèles avec la fonction `gam.check`. Ensuite, les prédictions de chaque modèle ont été obtenues via la fonction `predict.bam`, qui ont été utilisées pour générer les visualisations de la Figure 2 avec `ggplot2` (Wickham, 2009) de la suite `tidyverse` (Wickham *et al.*, 2019).

## 4 Résultats

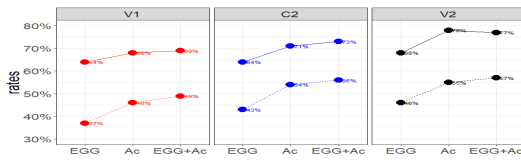
### 4.1 Classification avec les forêts aléatoires, Random Forests

Lorsque les deux classes étaient utilisées comme réponses (Figure 1a), les taux de classification variaient entre 64-78% avec les taux les plus élevés se situant dans la V2 en utilisant soit les mesures acoustiques soit EGG+acoustique. Ceci indique clairement une forme de coarticulation progressive vers la V2. Ces taux sont assez élevés mais ne permettent pas une discrimination parfaite des deux classes; la qualité de voix quantifiée via des méthodes articulatoires et acoustiques joue un rôle secondaire dans la discrimination entre classes. Ce résultat est attendu étant donné que l'arabe n'a pas de contraste phonologique marqué par une différence de phonation. En évaluant la classification par classe, les taux baissent d'une façon spectaculaire avec des taux proches des 37-67%. Il est à noter que lorsque les mesures articulatoires étaient utilisées, les taux étaient les plus bas en comparaison avec les mesures acoustiques, indiquant plus de facilité pour les forêts aléatoires à identifier les schémas spécifiques à nos classes en s'appuyant sur ces dernières. Les Figures 1b, 1c et 1d montrent le poids relatif pour les mesures articulatoires<sup>5</sup>. La HL était le prédicteur le plus important dans la classification dans V1 et V2, mais le QF était le prédicteur le plus important pour C. Les Figures 1e, 1f et 1g montrent le poids relatif des dix mesures les plus importantes en combinant toutes les mesures articulatoires et acoustiques. Le QF reste le prédicteur le plus important pour C, tandis que les mesures acoustiques  $A1^*-A3^*$ ,  $A2^*-A3^*$ ,  $H1^*-A3^*$ ,  $H4^*-H2kHz^*$ ,  $H2kHz^*-H5kHz^*$  étaient parmi les prédicteurs les plus importants dans toutes les séquences pour discriminer les six classes. Toutes ces mesures sont indicatives de variations liées à la pente spectrale, avec des pentes plus raides indicatives d'une augmentation d'énergie dans les hautes fréquences causée soit par une fermeture abrupte de la glotte (Hanson *et al.*, 2001; Al-Tamimi, 2017) soit par une constriction épilaryngale induisant une qualité de voix renforcée (Titze & Story, 1997; Moisik & Esling, 2010; Story, 2019).

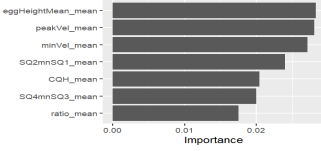
### 4.2 Modélisation avec les GAMMs

Figure 2 montrent les courbes moyennes prédites par nos modèles GAMMs par prédicteur en fonction des six classes par V1, C ou V2. Comme les résultats de  $A1^*-A3^*$  montrent des patterns comparables à ceux de  $A2^*-A3^*$ , nous présentons uniquement ceux de cette dernière. Les changements dynamiques de la qualité de voix au niveau articulatoire et acoustique suivent un pattern générique de  $\downarrow$  dans V1,  $\uparrow/\downarrow$  dans C et  $\uparrow$  dans V2. Plus spécifiquement, une HL variable est observée (Figure 2a) avec les

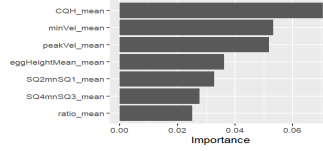
5. CQH = Quotient Fermé hybride; minVel = PDC; peakVel = PIC; SQ2mnSQ1 = phase fermante; SQ4mnSQ3 = phase ouvrante; ratio = Quotient de vitesse; eggHeightMean = Hauteur du Larynx



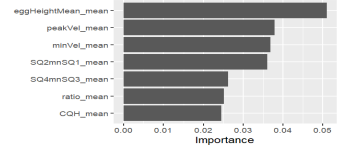
(a) Taux de Classification



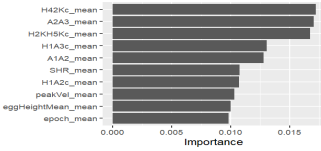
(b) PR - EGG - V1



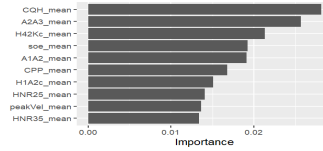
(c) PR - EGG - C



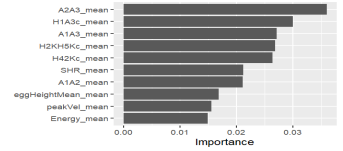
(d) PR - EGG - V2



(e) PR - EGG+acoust - V1



(f) PR - EGG+acoust - C



(g) PR - EGG+acoust - V2

FIGURE 1 – Taux de classification par forêt aléatoire sur les données de tests (a, ac = acoustique), le poids relatif (PR) des variables dans la tâche de classification des six classes combinant les mesures articulatoires (EGG) dans V1 (b), C (c) et V2 (d) et les dix mesures les plus importantes en combinant toutes les mesures articulatoires et acoustiques (EGG+acoust) dans V1 (e), C (f) et V2 (g).

consonnes pharyngales montrant un pattern  $\uparrow$ HL tout au long de la séquence VCV, les consonnes pharyngalisées avec un pattern  $\downarrow$  dans V1,  $\uparrow$  dans C et  $\downarrow$  dans V2; la consonne uvulaire suit un pattern similaire mais avec  $\uparrow$ HL dans la phase de relâchement (intervalle 9). Pour le QF (Figure 2b), les courbes semblent plus variées dans C, avec des changements mineurs dans V1 et V2. La consonne uvulaire montre le QF le plus bas, suivi des consonnes glottales et vélaires, puis pharyngales avec les consonnes alvéolaires et pharyngalisées ayant le QF le plus haut (autour de 0,35). Ces résultats corréleront bien avec un contact moins élevé pour les consonnes uvulaires et pharyngales en comparaison avec les consonnes pharyngalisées. Figures 2c, 2d et 2e montrent des résultats très similaires pour les consonnes uvulaires et pharyngalisées : une réduction globale de  $A2^*-A3^*$ ,  $H1^*-A3^*$  et de  $H4^*-H2kHz^*$  est observée, plus marquée dans les consonnes pharyngalisées. Ceci est indicatif d'un changement abrupt de la pente spectrale autour des 2kHz et 3kHz qui peut être corrélé avec une glotte plus serrée; pour les consonnes pharyngales, l'augmentation dans  $H1^*-A3^*$  et la baisse dans les autres mesures indiquent un épilarynx plus serré, suivant les prédictions avancées ci-dessus. Les résultats de  $H2kHz^*-H5kHz^*$  (Figure 2f) montrent un effet inverse avec un changement abrupt de la pente spectrale dans les consonnes pharyngales causé par une augmentation d'énergie dans les hautes fréquences causée par une constriction épilaryngale (avec  $\downarrow H2kHz^*-H5kHz^*$ ) avec une baisse d'énergie dans les consonnes uvulaires et pharyngalisées causée par une constriction glottale.

## 5 Discussion et conclusions

Cette étude a démontré que les consonnes d'arrière en arabe levantin induisent une élévation et une constriction du larynx variables. Les consonnes pharyngales sont produites avec une élévation



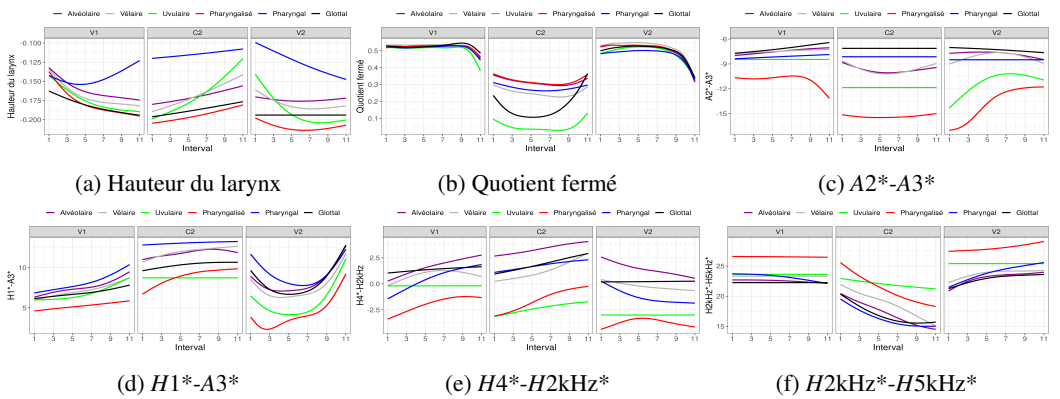


FIGURE 2 – Résultats des modélisations avec les GAMMs, avec les courbes prédites par V1, C ou V2 pour la HL (a), le QF (b),  $A2^*-A3^*$  (c),  $H1^*-A3^*$  (d),  $H4^*-H2kHz^*$  (e) et  $H2kHz^*-H5kHz^*$  (f).

maximale du larynx tout au long de la séquence VCV causée par une constriction épilaryngale, suivant les prédictions du LAM. Les consonnes pharyngalisées sont produites avec un abaissement du larynx, qui augmente vers la période du relâchement de la consonne, mais avec une fermeture abrupte de la glotte. La consonne uvulaire partage un larynx plus élevé avec les consonnes pharyngales mais une fermeture abrupte de la glotte avec les consonnes pharyngalisées. Les résultats articulatoires de HL et de QF et acoustiques liés à l'augmentation de l'énergie spectrale dans les hautes fréquences dans les consonnes pharyngalisées va dans le sens d'un contraste marqué par une différence de voix tendue dans ces dernières en comparaison avec une voix relâchée/modale dans les consonnes alvéolaires comparable aux résultats rapportés dans (Kuang & Keating, 2014) de contraste phonologique entre voix tendues et relâchées. La complémentarité entre corrélats articulatoires et acoustiques dans cette étude confirme que les mesures de la pente spectrale, qui sont validées comme permettant de classifier les qualités de voix (Garellek *et al.*, 2016; Kreiman *et al.*, 2021) jouent un rôle important dans la caractérisation des différences phonatoires entre les consonnes d'arrière de l'arabe. Nos résultats articulatoires et acoustiques fournissent une évidence empirique que les consonnes d'arrière en arabe partagent le trait phonologique [+Constricted Glottis] comme trait secondaire actif avec une variation graduelle bien marquée permettant de les différencier. Ces résultats suivent les prédictions du LAM : les consonnes d'arrière en arabe sont produites avec une constriction graduelle de l'épilarynx qui induit une élévation et une constriction variables de la glotte avec une rétraction variable du dos et de la racine de la langue. Cette complémentarité permet ainsi de conclure que les trois membres de cette classe de consonnes d'arrière partagent des traits phonologiques similaires permettant de les unir.

## Remerciements

Ce travail a bénéficié partiellement d'une aide de l'IdEx Université Paris Cité (ANR-18-IDEX-0001) au titre du Labex Empirical Foundations of Linguistics - EFL. Il a également bénéficié d'un financement de la British Academy/Leverhulme small research grant, Royaume Uni (SG160181; 2017-2019) et d'un financement Leverhulme International Academic Fellowship, Royaume Uni (IAF-2018-016). Ce travail a bénéficié du support pour l'utilisation du High Power Computing (HPC) de l'université de Newcastle, au Royaume Uni, du CNRS/TGIR HUMA-NUM, IN2P3 et du GENCI-IDRIS, France (2022-AD010613733).

# Références

- AL-TAMIMI F. & HESELWOOD B. (2011). Nasoendoscopic, videofluoroscopic and acoustic study of plain and emphatic coronals in Jordanian Arabic. In B. HESELWOOD & Z. HASSAN, Éd.s., *Instrumental Studies in Arabic Phonetics*, p. 165–191. John Benjamins. DOI : [10.1075/cilt.319](https://doi.org/10.1075/cilt.319).
- AL-TAMIMI J. (2015). Spectral tilt as an acoustic correlate to pharyngealisation in Jordanian and Moroccan Arabic (Article : 0436). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*.
- AL-TAMIMI J. (2017). Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic : Implications for formal representations. *Laboratory Phonology : Journal of the Association for Laboratory Phonology*, **8**(1), 1–40. DOI : [10.5334/labphon.19](https://doi.org/10.5334/labphon.19).
- AL-TAMIMI J. (2022). Praat-f0-Accurate-Estimation. <https://jalalal-tamimi.github.io/Praat-f0-Accurate-Estimation/>.
- AL-TAMIMI J. & KHATTAB G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, **71**, 306–325. DOI : [10.1016/j.wocn.2018.09.010](https://doi.org/10.1016/j.wocn.2018.09.010).
- AL-TAMIMI J. & PALO P. (2023). Dynamics of the tongue contour in the production of guttural consonants in Levantine Arabic. In *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS)*, p. 2095–2099.
- AL-TAMIMI J., SCHIEL F., KHATTAB G., SOKHEY N., AMAZOUZ D., DALLAK A. & MOUSSA H. (2022). A Romanization System and WebMAUS Aligner for Arabic Varieties. In *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, p. 7269–7276.
- ARALOVA N., GRAWUNDER S. & WINTER B. (2011). The Acoustic Correlates of Tongue Root Vowel Harmony in Even (Tungusic). In *Proceedings of the 17th International Congress of Phonetic Sciences, ICPhS*, p. 240–243.
- BARR D. J., LEVY R., SCHEEPERS C. & TILY H. J. (2013). Random effects structure for confirmatory hypothesis testing : Keep it maximal. *Journal of Memory and Language*, **68**(3), 255–278. DOI : [10.1016/j.jml.2012.11.001](https://doi.org/10.1016/j.jml.2012.11.001).
- BOERSMA P. & WEENINK D. (2020). *Praat Software*. University of Amsterdam.
- EMBARKI M. (2008). Les dialectes arabes modernes : État et nouvelles perspectives pour la classification géo-sociologique. *Arabica*, **55**(5/6), 583–604. DOI : [10.1163/157005808X364616](https://doi.org/10.1163/157005808X364616).
- ESLING J., MOISIK S., BENNER A. & CREVIER-BUCHMAN L. (2019). *Voice Quality : The Laryngeal Articulator Model*. Cambridge University Press. DOI : [10.1017/9781108696555](https://doi.org/10.1017/9781108696555).
- FULOP S. A., KARI E. & LADEFOGED P. (1998). An Acoustic Study of the Tongue Root Contrast in Degema Vowels. *Phonetica*, **55**(1-2), 80–98. DOI : [10.1159/000028425](https://doi.org/10.1159/000028425).
- GARELLEK M., SAMLAN R., GERRATT B. R. & KREIMAN J. (2016). Modeling the voice source in terms of spectral slopes. *The Journal of the Acoustical Society of America*, **139**(3), 1404–1410. DOI : [10.1121/1.4944474](https://doi.org/10.1121/1.4944474).
- GUION S. G., POST M. W. & PAYNE D. L. (2004). Phonetic correlates of tongue root vowel contrasts in Maa. *Journal of Phonetics*, **32**(4), 517–542. DOI : [10.1016/j.wocn.2004.04.002](https://doi.org/10.1016/j.wocn.2004.04.002).
- HANSON H. M. & CHUANG E. S. (1999). Glottal characteristics of male speakers : Acoustic correlates and comparison with female data. *The Journal of the Acoustical Society of America*, **106**(2), 1064–1077. DOI : [10.1121/1.427116](https://doi.org/10.1121/1.427116).
- HANSON H. M., STEVENS K. N., KUO H.-K. J., CHEN M. Y. & SLIFKA J. (2001). Towards models of phonation. *Journal of Phonetics*, **29**(4), 451–480. DOI : [10.1006/jpho.2001.0146](https://doi.org/10.1006/jpho.2001.0146).

- HESELWOOD B. (2007). The ‘tight approximant’ variant of the Arabic ‘ayn. *Journal of the International Phonetic Association*, **37**(1), 1. DOI : [10.1017/S0025100306002787](https://doi.org/10.1017/S0025100306002787).
- HESELWOOD B. & AL-TAMIMI F. (2011). A study of the laryngeal and pharyngeal consonants in Jordanian Arabic using nasoendoscopy, videofluoroscopy and spectrography. In B. HESELWOOD & Z. HASSAN, Éds., *Instrumental Studies in Arabic Phonetics*, p. 101–128. John Benjamins. DOI : [10.1075/cilt.319](https://doi.org/10.1075/cilt.319).
- HOLMBERG E. B., HILLMAN R. E. & PERKELL J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, **84**(2), 511–529. DOI : [10.1121/1.396829](https://doi.org/10.1121/1.396829).
- HOWARD D. M. (1995). Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice*, **9**(2), 163–172. DOI : [10.1016/S0892-1997\(05\)80250-4](https://doi.org/10.1016/S0892-1997(05)80250-4).
- ISELI M., SHUE Y.-L. & ALWAN A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America*, **121**(4), 2283–2295. DOI : [10.1121/1.2697522](https://doi.org/10.1121/1.2697522).
- KREIMAN J., LEE Y., GARELLEK M., SAMLAN R. & GERRATT B. R. (2021). Validating a psychoacoustic model of voice quality. *The Journal of the Acoustical Society of America*, **149**(1), 457–465. DOI : [10.1121/10.0003331](https://doi.org/10.1121/10.0003331).
- KUANG J. & KEATING P. (2014). Vocal fold vibratory patterns in tense versus lax phonation contrasts. *The Journal of the Acoustical Society of America*, **136**(5), 2784–2797. DOI : [10.1121/1.4896462](https://doi.org/10.1121/1.4896462).
- KUHN M. & WICKHAM H. (2020). *tidymodels : Easily Install and Load the 'Tidymodels' Packages*.
- LAUFER A. & BAER T. (1988). The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and Speech*, **31**, 181–205. DOI : [10.1177/002383098803100205](https://doi.org/10.1177/002383098803100205).
- LUBBERS M. & TORREIRA F. (2013). Praatalign : An interactive Praat plug-in for performing phonetic forced alignment : <https://github.com/dopefishh/praatalign>. Version 2.0a.
- MCCARTHY J. (1994). The phonetics and phonology of Semitic pharyngeals. In P. KEATING, Éd., *Phonological Structure and Phonetic Form*, p. 191–233. Cambridge University Press. DOI : [10.1017/CBO9780511659461.012](https://doi.org/10.1017/CBO9780511659461.012).
- MICHAUD A. (2004). A measurement from electroglottography : DECPA, and its application in prosody. In *Speech Prosody 2004*, p. 633–636.
- MOISIK S. & ESLING J. (2010). Examining the Acoustic Contributions of the Epilaryngeal Tube to the Voice Source and Vocal Tract Resonance. *Canadian Acoustics*, **38**(3), 138–139.
- SCHIEL F. (2015). A statistical model for predicting pronunciation (Article : 0195). In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*.
- SHUE Y.-L., KEATING P., VICENIK C., YU K. & YEN-LIANG K. P. V. C. Y. K. S. (2011). VoiceSauce : A program for voice analysis. In *Proceedings of the 17th International Congress of Phonetic Sciences, ICPhS*, p. 1846–1849.
- STORY B. (2019). The Vocal Tract in Singing. In *The Oxford Handbook of Singing*, p. 144–166. Oxford University Press. DOI : [10.1093/oxfordhb/9780199660773.013.012](https://doi.org/10.1093/oxfordhb/9780199660773.013.012).
- STROBL C., MALLEY J. & TUTZ G. (2009). An introduction to recursive partitioning : Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, **14**(4), 323–348. DOI : [10.1037/a0016973](https://doi.org/10.1037/a0016973).
- SYLAK-GLASSMAN J. (2014). *Deriving Natural Classes : The Phonology and Typology of Post-Velar Consonants*. Thèse de doctorat, University of California, Berkeley.

- TEHRANI H. (2020). <http://www.appsobabble.com/functions/eggworks.aspx> (accessible 18/02/2020).
- TITZE I. R. & STORY B. H. (1997). Acoustic interactions of the voice source with the lower vocal tract. *The Journal of the Acoustical Society of America*, **101**(4), 2234–2243. DOI : [10.1121/1.418246](https://doi.org/10.1121/1.418246).
- WICKHAM H. (2009). *Ggplot2 : Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- WICKHAM H., AVERICK M., BRYAN J., CHANG W., MCGOWAN L. D., FRANÇOIS R., GROLEMUND G., HAYES A., HENRY L., HESTER J., KUHN M., PEDERSEN T. L., MILLER E., BACHE S. M., MÜLLER K., OOMS J., ROBINSON D., SEIDEL D. P., SPINU V., TAKAHASHI K., VAUGHAN D., WILKE C., WOO K. & YUTANI H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, **4**(43), 1686. DOI : [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- WOOD S. N. (2017). *Generalized Additive Models : An Introduction with R*. CRC Press/Taylor & Francis Group. DOI : [10.1201/9781315370279](https://doi.org/10.1201/9781315370279).
- WRENCH A. A. & SCOBIE J. M. (2008). High-speed Cineloop Ultrasound vs. Video Ultrasound Tongue Imaging : Comparison of Front and Back Lingual Gesture Location and Relative Timing. In *Proceedings of the 8th International Seminar on Speech Production (ISSP)*, p. 57–60.
- WRIGHT M. N. & ZIEGLER A. (2017). Ranger : A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, **77**(1). DOI : [10.18637/jss.v077.i01](https://doi.org/10.18637/jss.v077.i01).
- ZAWAYDEH B. (1999). *The Phonetics and Phonology of Gutturals in Arabic*. Thèse de doctorat, Bloomington, IN : Indiana University.