ISSP 2024 - 13th International Seminar on Speech Production
13-17 May 2024, Autrans, France

# Lingual and epilaryngeal articulation of vowels in Mundabli

Matthew Faytak[1], Mariana Quintana Godoy[1], Tianle Yang[1]

[1]*University at Buffalo, USA*
{faytak, marianaq, tianleya} @buffalo.edu

## Abstract

*The vowel system of Mundabli (Yemne-Kimbi, Cameroon) is rich in contrasts involving lower vocal tract activity. In this study, we aim to characterize the acoustics and articulation of the three sets of vowels in the language, which we refer to as pharyngeal, plain, and lax. Acoustic time series data reveal that pharyngealization raises F1 and lowers F2 and F3; it also conditions tense or creaky voice quality relative to the plain and lax vowels. Ultrasound data suggest that these acoustic properties can typically be attributed to a lower pharyngeal or epilaryngeal constriction. The data also suggest that the lax and plain vowels may exhibit an advanced tongue root contrast. Variation in the articulatory implementation of pharyngealization observed in the ultrasound data is discussed.*

**Keywords:** speech production, ultrasound tongue imaging, laryngeal articulator, pharyngealization, advanced tongue root

## 1. Introduction

Mundabli (ISO 639-3 [boe]) is a Yemne-Kimbi language spoken in the Lower Fungom region of northwestern Cameroon by no more than 800 inhabitants of a single village of the same name (Good et al. 2011; Voll 2017). Mundabli's vowel system features ten monophthongs /i ɪ e ɛ ɨ a u ʊ o ɔ/ plus six monophthongs which have been described as pharyngealized /iˤ eˤ ɨˤ aˤ uˤ oˤ/ (Voll 2017). The paired arrangement of the front and back non-pharyngealized vowels suggests two sets contrasting in height or tongue root advancement, here referred to as PLAIN /i e u o/ and LAX /ɪ ɛ ʊ ɔ/.

The PHARYNGEAL vowels are unusual both in Mundabli's local area and more broadly cross-linguistically. They have developed recently in the language's history from vowels formerly followed by coda **k* or **ʔ*. Fieldwork by the first author has revealed that a closely related neighboring language, Mufu, has /k/ or /ʔ/ in cognate lexical items, e.g. Mufu [bàʔ], Mundabli [bàˤ] 'scar'; Mufu [cōk], Mundabli [tsōˤ] 'banana'; Mufu [dàk], Mundabli [dèˤ] 'place'.

Morphophonological alternations between plain and pharyngeal vowels suggest a phonological organization of the vowels into a system summarized in **Table 1**. For instance, pharyngealization marks imperfective aspect on many open-syllable verb stems, e.g. [fī] 'press.PERF' *vs.* [fīˤ] 'press.IPFV'; [bú] 'give birth.PERF' *vs.* [búˤ] 'give birth.IPFV'. This alternation is a trace of an imperfective suffix **-k(ə)* which is still overtly realized in Mufu as *-k* or *-ʔ*; compare Mufu [fjàk] 'press.IPFV', [búk] 'give birth.IPFV'. We refer the reader to Voll (2017) for further reading on the relationships among the Mundabli vowel sets.

Mundabli is notably rich in contrasts based on lower vocal tract activity, in that both the lax and pharyngeal vowels may make use of the epilaryngeal tube as an articulator. Pharyngeal consonants and pharyngealized vowels are known to involve strong constriction of the epilaryngeal tube (Catford 1983; Arkhipov et al. 2019). On the other hand, the lax vowel set may exhibit retracted tongue root (RTR), and the plain vowel set advanced tongue root (ATR), a common arrangement elsewhere in West Africa (Casali 2008). RTR vowels have *also* been identified as making use of epilaryngeal constriction to distinguish themselves from ATR vowels (Esling 2005; Edmondson et al. 2007; Starwalt 2008).

It is unclear how Mundabli speakers would organize lingual and epilaryngeal articulation in its vowels to accommodate a three-way lower vocal tract activity distinction, since such a situation is (at the least) very rare in the world's languages. As such, this exploratory study aims to clarify the acoustic differences among Mundabli's pharyngeal, plain, and lax vowel sets, and the lingual and epilaryngeal articulatory basis of these distinctions. In particular, we hope to clarify the articulatory nature of the pharyngeal and lax vowels, and how they differ from the plain vowels and each other.

Table 1: *A possible phonological organization of the Mundabli monophthongal vowels.*

| Vowel set | Plain | Pharyngeal | Lax |
|---|---|---|---|
| I | i | iˤ | ɪ |
| E | e | eˤ | ɛ |
| II | ɨ | ɨˤ | |
| A | a | aˤ | |
| U | u | uˤ | ʊ |
| O | o | oˤ | ɔ |

## 2. Methods

Time-aligned acoustic and ultrasound data were collected from 15 Mundabli speakers (7M 8F, mean age 31.9, SD 8.01) in 2022 and 2023 in Douala, Cameroon using Articulate Assistant Advanced (v221.2.0). Audio was recorded at a sampling rate of 22.05 kHz using a Røde NTG2 supercardioid condenser microphone mounted on a tabletop tripod; the signal was digitized using a Scarlett Solo 2i2 USB audio interface. Ultrasound video was recorded using a Telemed MicrUs and an MC4 micro-convex probe secured by an Articulate Instruments UltraFit headset (Spreafico, Pucher, and Matosova 2018); recordings were made at a rate of 82.1 Hz in a 101.2° field of view.

Stimuli were a set of 32 open-syllable monosyllabic words containing all 16 Mundabli vowels (two word types per vowel) read in one of three frame sentences, verbally prompted by the first author. The frame sentences were varied to provide a syntactically appropriate frame for each target word. Nouns were presented in the frame ⁿkáˤn _ 'I have (a) (noun)"; verbs were


10.21437/issp.2024-43

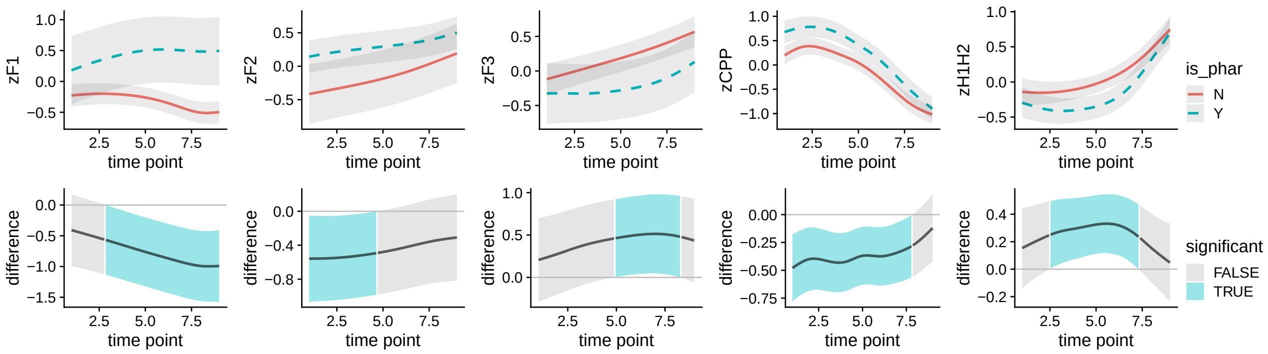


Figure 1: *GAMM smooths and difference smooths for z-scored F1, F2, F3, CPP, and H1*-H2* over vowels' normalized duration.*

presented one of two frames, $^{n}$fã$^{ʕ}$ á$^{ʕ}$ $^{n}$ _ 'I am (verb)-ing' for imperfective forms of verbs, which invariably take a low-toned and prenasalized form in this context, and $^{n}$fá _ 'I have (verb)-ed' The varied frame sentences were needed, in part, to elicit the correct aspect-inflected form of verbs which are pharyngealized only in the imperfective. During recording, speakers repeated the frame with the embedded target between four and six times (typically five).

The final acoustic data set contains 4,726 vowel tokens across all 15 speakers. Acoustic measures were extracted from this data using PraatSauce (Kirby 2019) at nine evenly spaced time points across vowels' durations. Formant frequencies (F1-3) were extracted, and the amplitude difference between the first and second harmonics (H1-H2) and cepstral peak prominence (CPP) were extracted as measures of voice quality. All measures were $z$-scored. After removing outliers more than $\pm$2SD from the mean for each measure, 37,092 samples remained (from roughly 4,121 vowel tokens). AR1 GAMMs were carried out for each acoustic measure using the *mgcv* package in R (Wood 2023), with factor smooths for vowel class (pharyngealized *vs* pooled lax and plain) and random smooths for speaker and word.

The co-collected ultrasound data was analyzed for six of the 15 speakers (3M 3F, mean age 30.0, SD 7.92). Tongue surface contours were segmented from ultrasound video using the DeepLabCut model implemented in Articulate Assistant Advanced (v221.2.0). DeepLabCut contours were converted to fan splines and trimmed of any knots not originally estimated by the DeepLabCut model; contours at vowel midpoint were extracted and converted to polar coordinates. Smoothing-spline ANOVAs (SSANOVAs) by vowel group and category (pharyngeal *vs.* plain *vs.* lax) were carried out using the *gss* package in R (Gu 2023).

## 3. Results

### 3.1. Acoustic measures

GAMM smooths and difference plots for formant data are shown in the first three columns of **Figure 1**; difference smooths are located below each measure's smoothed estimates. As a group, pharyngealized vowels exhibit raised F1, raised F2, and lowered F3 (Figure 1B) relative to the plain and lax vowels. The difference in F1 is largest, and increases in size towards the end of the vowel. This coincides with the development of a smaller, significant difference in F3 during the last half of the vowel. By contrast, the F2 difference reaches significance only during the first half of the vowel and diminishes later in the vowel.

GAMM smooths and difference plots for voice quality data are shown in the two rightmost columns in **Figure 1**. Pharyngealized vowels exhibit elevated CPP and reduced H1*-H2* relative to their plain and lax counterparts, suggesting that pharyngealized vowels are characterized by relatively tense or creaky phonation. Values for both voice quality measures converge towards the end of vowel duration in a direction that suggests that all vowels end in breathy phonation; this is likely a reflection of the prepausal devoicing of the end of all target words. Phrase-final devoicing is a tendency noted independently by Voll (2017, p. 32).

### 3.2. Ultrasound

Data for six of fifteens speakers are analyzed here, with SSANOVAs carried out separately for each speaker. For reasons of space, we present results pooled across all vowels by speaker, then provide by-vowel group breakdowns for two representative speakers.

We first focus on articulatory differences between the pharyngeal and non-pharyngeal vowels (lax and plain vowels as a group). For the pharyngeal vowels, all speakers exhibit substantial lowering of the posterior tongue dorsum and bunching and fronting of the anterior dorsum and blade (**Figure 2**). Some speakers tend toward a double-bunched shape, particularly speaker F1 (**Figure 3**). The pharyngeal-non pharyngeal difference is reduced somewhat in the front vowels, particularly in the I group (**Figures 3, 4**). For four of the six speakers (F1, M1, M2, and M3) the lower portion of the tongue root shows additional retraction relative to the non-pharyngeal vowels. This difference obtains across all vowel groups, though less so for the back vowel groups (U and O), as shown for speaker F1 in **Figure 3**. Speaker F3 exhibits no differences in the position of the lower tongue root, but her pattern of articulation otherwise resembles the aforementioned four speakers.

Speaker F2 diverges from the majority pattern, with the entire tongue root fronted during production of pharyngealized vowels. This pattern also holds in most of the individual vowel groups as shown in **Figure 4**, though the I group exhibits no clear distinction in root frontness and the E group exhibits the majority pattern by which the root is retracted. Speaker F2's articulation of the pharyngeal vowel set could thus be characterized as involving fronting of the entire tongue and a bunched configuration, rather than the simultaneous fronting/bunching and lower root constriction which characterizes four of the five remaining speakers.

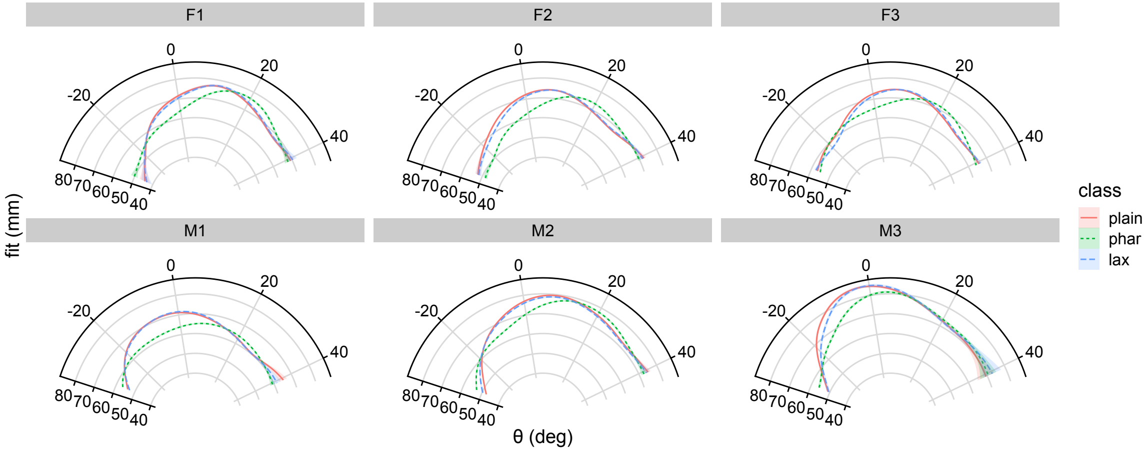


Figure 2: *SSANOVA splines at vowel midpoint for all speakers, pooling across vowel groups.*

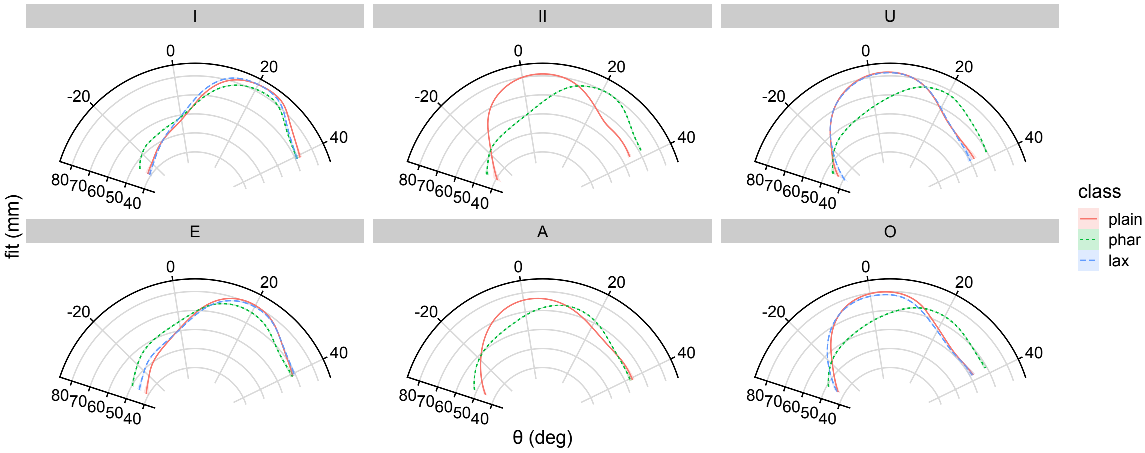


Figure 3: *SSANOVA splines at vowel midpoint by vowel group for speaker F1, as an example of the majority pattern of root retraction in pharyngealized vowels.*

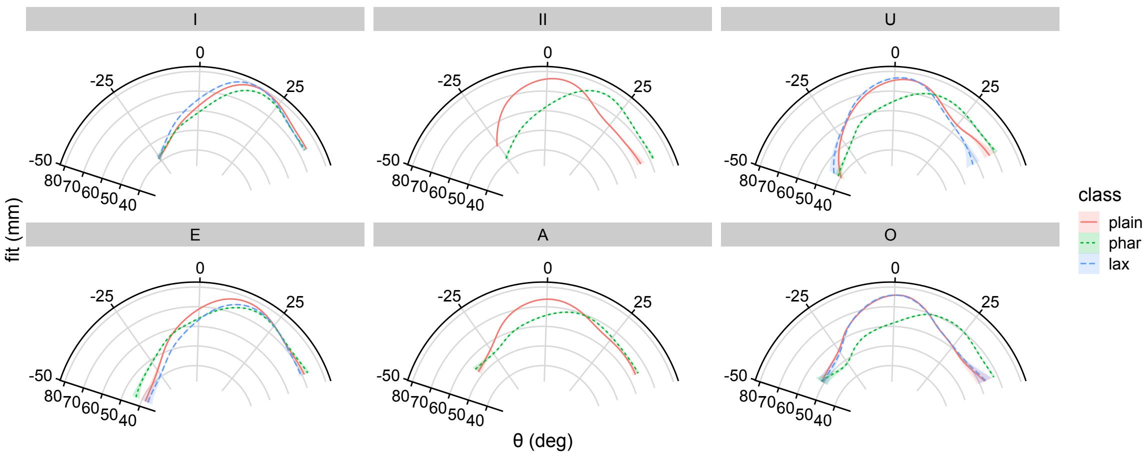


Figure 4: *SSANOVA splines at vowel midpoint by vowel group for speaker F2, who generally exhibits a distinct pattern of tongue fronting in pharyngealized vowels.*

We now turn to differences in articulation between the lax and plain vowel sets. Unpooling the vowels as in **Figures 3, 4**, for both speakers F1 and F2, a majority of the data suggests slight tongue root retraction for lax /ɪ ɛ ʊ ɔ/, relative to plain /i e u o/, with little consistent involvement of dorsum height differences. Both speakers exhibit reversals of this relationship (F1's /u-ʊ/ pair, F2's /e-ɛ/ pair), but these may be due to gaps in the lexicon which required the use of stimuli with lingual onset consonants in the E and U vowel groups (i.e. gbɛ̋ 'wind', gbʊ̄ 'fall').

Finally, we compare the lax and pharyngeal vowels for speaker F2 in **Figure 3**, whose pharyngeal vowels involve root retraction. The root retraction observed for the lax set is less extreme than for the pharyngeal set; the lax set's root retraction also recruits tongue dorsum lowering to a lesser extent.

# 4. Discussion and conclusion

The articulatory and acoustic data suggest that the pharyngeal vowels involve an epilaryngeal constriction for most speakers, a maneuver readily distinguished from lower vocal tract articulations often called "pharyngealization", such as emphasis or uvularization (Evans et al. 2016; al-Tamimi 2017). Lower pharyngeal or epilaryngeal constriction is clearly suggested by both the F2-raising effect seen in the pharyngeal vowels (as opposed to F2-lowering for uvularization) and the characteristic double-bunching also observed in languages with lower pharyngeal constrictions (Catford 1983; Arkhipov et al. 2019). The involvement of tense or creaky phonation in pharyngeal vowels also specifically implicates the laryngeal articulator, as constriction of the epilarynx is known to yield non-modal phonation (Edmondson et al. 2007; Moisik, Czaykowska-Higgins, and Esling 2021).

The "lax" and plain vowels appear to exhibit ±ATR differences. This finding suggests that Mundabli may exhibit two articulatory mechanisms for the lax and pharyngealized vowels which involve different types or degrees of epilaryngeal constriction. The Mundabli pharyngealized vowels often show dorsum lowering and "bunching" around the level of the hyoid bone, a configuration which Esling (2005) describes as characteristic of the most epilaryngeally constricted sounds; they especially resemble pharyngeal approximant consonants such as [ʕ]. The lax vowels are lightly retracted by comparison, without a great deal of tongue dorsum lowering in support of this goal. The lax vowels thus more closely resemble vocalic ATR or registral differences as described in a number of previous works (Edmondson et al. 2007; Esling 2005). Further analysis of more speakers' articulations is needed within each vowel group to confirm the robustness of this finding. More detailed analysis of the acoustic data to examine the voice quality of the lax vowels is also merited, given that ATR distinctions are often associated with nonmodal phonation to some degree (Casali 2008; Starwalt 2008; Akinbo et al. 2023).

Speaker F2 appears to use fronted or singly bunched, rather than double-bunched, tongue shapes for the pharyngeal vowels, with no obvious lower pharyngeal constriction. The lack of epilaryngeal activity for this speaker is also suggested by the absence of CPP and H1*-H2* differences between the pharyngealized and non-pharyngealized vowels for this speaker; on the other hand, her formant data are comparable to that of other speakers. For speaker F2, the contrast between the pharyngeal vowels and the other vowel sets may have shifted to rely more on the associated formant frequency differences, possibly due to a misattribution of the acoustic effects of double-bunching to a fronted, singly bunched tongue position. Three other speakers in the data set whose ultrasound data were not analyzed here may exhibit similar variants to speaker F2's; further analysis may shed light on variation across the larger population of Mundabli speakers.


# 5. Acknowledgements

Yùŋ kɛ̋n mòˤmòˤ to the Mundabli community members for your contributions as language experts. We also thank Alan Wrench, Bowei Shao, and Jalal al-Tamimi for helpful comments, Connor Quimby for annotation of Mufu data, and Ikom Christopher, Jeff Good, and Pierpaolo DiCarlo for logistical support.


# 6. References


Akinbo, Samuel, Avery Ozburn, Gerald Nweya, and Douglas Pulleyblank (2023). "Eleven vowels of Imilike Igbo including ATR and RTR schwa". In: *JIPA*, pp. 1–24.

Arkhipov, A, M Daniel, O Belyaev, G Moroz, and JH Esling (2019). "A reinterpretation of lower-vocal-tract articulations in Caucasian languages". In: *Proc ICPhS 19, Melbourne*, pp. 5–9.

Casali, Roderic F (2008). "ATR harmony in African languages". In: *Lang and Ling Compass* 2.3, pp. 496–549.

Catford, John C (1983). "Pharyngeal and laryngeal sounds in Caucasian languages". In: *Vocal fold physiology: Contemporary research and clinical issues*, pp. 344–350.

Edmondson, Jerold A, Cécile M Padayodi, Zeki Majeed Hassan, and John H Esling (2007). "The laryngeal articulator: Source and resonator". In: *Proc ICPhS 16, Saarbrücken*. Vol. 3, pp. 2065–2068.

Esling, John H (2005). "There are no back vowels: The larygeal articulator model". In: *Canadian Journal of Linguistics/Revue canadienne de linguistique* 50.1-4, pp. 13–44.

Evans, Jonathan P, Jackson T-S Sun, Chenhao Chiu, and Michelle Liou (2016). "Uvular approximation as an articulatory vowel feature". In: *JIPA* 46.1, pp. 1–31.

Good, Jeff, Jesse Lovegren, Jean Patrick Mve, Carine Nganguep Tchiemouo, Rebecca Voll, and Pierpaolo Dicarlo (2011). "The languages of the Lower Fungom region of Cameroon: Grammatical overview". In: *Africana Linguistica* 17.1, pp. 101–164.

Gu, Chong (2023). *Package 'gss'*. R package version 2.2.7.

Kirby, James (2019). *praatsauce*. GitHub repository. URL: https://github.com/kirbyj/praatsauce.

Moisik, Scott R, Ewa Czaykowska-Higgins, and John H Esling (2021). "Phonological potentials and the lower vocal tract". In: *JIPA* 51.1, pp. 1–35.

Spreafico, Lorenzo, Michael Pucher, and Anna Matosova (2018). "UltraFit: A speaker-friendly headset for ultrasound recordings in speech science". In: *Proc Interspeech 2018, Hyderabad*. ISCA, pp. 1517–1520.

Starwalt, Coleen G. A. (2008). "The acoustic correlates of ATR harmony in seven-and nine-vowel African languages: A phonetic inquiry into phonological structure". PhD thesis. U Texas Arlington.

al-Tamimi, Jalal (2017). "Revisiting acoustic correlates of pharyngealization in Jordanian and Moroccan Arabic: Implications for formal representations". In: *LabPhon* 8.

Voll, Rebecca (2017). "A grammar of Mundabli: a Bantoid (Yemne-Kimbi) language of Cameroon". PhD thesis. U Leiden.

Wood, Simon (2023). *Package 'mgcv'*. R package version v1.8.42.