Interspeech 2022
18-22 September 2022, Incheon, Korea

Complex sounds and cross-language influence: The case of ejectives in Omani Mehri

Rachid Ridouane[1], *Philipp Buech*[1]

[1]Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle), Paris
rachid.ridouane@sorbonne-nouvelle.fr, philipp.buech@sorbonne-nouvelle.fr

Abstract

Ejective consonants are known to considerably vary both cross-linguistically and within individual languages. This variability is often considered a consequence of the complex articulatory strategies involved in their production. Because they are complex, they might be particularly prone to sound change, especially under cross-language influence. In this study, we consider the production of ejectives in Mehri, a Semitic endangered language spoken in Oman where considerable influence from Arabic is expected. We provide acoustic data from seven speakers producing a list of items contrasting ejective and pulmonic alveolar and velar stops in word-initial (/#—/), word-medial (V—V), and word-final (V—#) positions.

Different durational and non-durational correlates were examined. The relative importance of these correlates was quantified by the calculation of D-prime values for each. The key empirical finding is that the parameters used to signal ejectivity differ depending mainly on whether the stop is alveolar or velar. Specifically, ejective alveolar stops display characteristics of pharyngealization, similar to Arabic, but velars still maintain attributes of ejectivity in some word positions. We interpret these results as diagnostic of the sound change that is currently in progress, coupled with an ongoing context-dependent neutralization.

Index Terms: Acoustics, ejectivity, pharyngealization, sound change, neutralization, Mehri.

1. Introduction

The complexity of ejective sounds is commonly attributed to articulatory constraints that make them acoustically unstable (e.g., [1,2,3]). Given this instability, ejectives might be particularly prone to sound change, especially under cross-language influence. We consider the production of ejectives in Omani Mehri, a Semitic endangered language particularly threatened by the more prestigious status of Arabic, the only official and national language of Oman.

Mehri is a Modern South Arabian (MSA) language, a subgroup of the Semitic branch of the Afroasiatic family. This subgroup includes 5 other languages: Baṭḥari, Harsūsi, Jibbali, Hobyot, and Soqotri. Arabic and MSA are both Semitic languages. The common classification views MSA languages as part of a South branch of West Semitic and Arabic as part of a Central branch of West Semitic [4]. Spoken by an estimated 50.800 speakers, Mehri is considered by UNESCO as *definitely endangered* [5].

The Mehri variety examined in this study is spoken in Salalah, Dhofar province (Oman). In addition to voiceless and voiced stops and fricatives, the language has ejective stops [t', k'] as well as a series of 4 ejective fricatives [θ', s', ɬ', ʃ']. Previous studies reported a great deal of variability in the realization of these segments, depending on place of articulation, manner of articulation, and prosodic position ([6,7,8,9,10,11]). The transcriptions used to represent these segments have also been variable ([6,7,9,11,12]). This variability raised the question of whether the description of these segments as ejectives was, in fact, correct, and led some researchers to argue for a gradual sound change from an ejective to a dorsopharyngeal contrast, under the influence of Arabic ([7,8,9,13]). For [6], the intensity of the glottalic release in Mehri, and MSA in general, is less than that in the Ethio-Semitic languages. [8], although they also consider these segments as real ejectives, report some conditioned variation, leading to voicing or dorsopharyngealization, but without providing phonetic data on this variation. Dismissing the term 'ejective', [8] argues that only /k'/ is uniformly produced as an ejective, other segments are rather dorsophangealized. [10], based on qualitative analysis of acoustic waveforms and spectrograms from one speaker of Yemeni Mehri, further confirm that only /k'/ is produced on a glottal flow.

Most reports on Mehri ejectives were limited to the variety spoken in Yemen and relied primarily on perception-based segmental transcription. For this study, we recorded data from Omani Mehri and examined various temporal and non-temporal acoustic correlates that potentially differentiate plain alveolar and velar stops [t, k] from their ejective counterparts [t', k'] (see [13] on the acoustics of ejective fricatives). Taking a lead from previous studies, we identify the acoustic attributes used to signal ejectivity and show that they vary depending on place of articulation, word position, and speakers. The findings provide empirical support for the sound change that is currently in progress, probably coupled with an ongoing context-dependent neutralization.

2. Method

2.1. Speakers

Seven male native Mehri speakers, aged 20-27 years (mean 22.8, SD 2.19), were recorded during fieldwork in Salalah, a city in Dhofar province, southern Oman. All the speakers were bilingual in Mehri and dialectal Arabic, as is common for native Mehri speakers. They all learned Standard Arabic at school. None of them had a known history of hearing or speech disorders at the time of recording.

2.2. Procedure

The stimuli used in this study are shown in Table 2. They consisted of 12 items contrasting ejective and pulmonic alveolar and velar stops in word-initial (/#—V/), word-medial (V—V), and word-final (V—#) positions. Eight of these items were triliteral consonantal roots conjugated in two verbal

Copyright © 2022 ISCA

10.21437/Interspeech.2022-10199

paradigms (perf. 3ms and subj. 3ms) to have target segments in word-initial and word-medial positions, respectively. Additional items (shaded columns of Table 1) were used to have target segments in word-final position. Each target consonant is preceded and/or followed by the same vowel for each word position. Some of these vowels were not preceded or followed by the same consonant for each pair due to the lack of actual Mehri words with the relevant structure (e.g., in t/t' contrast in word-medial position, schwa is followed by [r] in [itəre:b] and by [l] in [it'əle:b]). To avoid as much as possible any effect due to the differences between these consonants, relevant measurements were taken at the onset of the following vowels in initial and medial positions and at the offset of the preceding vowels in medial and final positions. Each item was produced within a carrier sentence and repeated 5 times by each speaker, yielding a dataset of 420 tokens (12 items*7 speakers*5 repetitions).

Table 1: *List of Mehri items*

Root	/#—V/	/V—V/	V—#	
√trb 'ablution'	təru:b	itəre:b	kəfu:t	"to catch"
√t'lb 'to ask'	t'əlu:b	it'əle:b	nəfu:t'	"to fight"
√ kbn 'to hide'	kəbu:n	ikəbe:n	səru:k	"behind you"
√ k'br 'to bury'	k'əbu:r	ik'əbe:r	ħəmu:k'	"injuries"

Acoustic data were segmented and labeled using Praat [14]. Different durational and non-durational values were taken from the acoustic records for each item. Non-durational values included:

1. Counts of post-release silent lag (if any) in initial and medial positions
2. Intensity of stop release burst
3. Center of gravity (CoG) of stop release burst
4. ΔF0 at the onset of the following vowel and at the offset of the preceding vowel
5. Formants at the onset of the following vowel and at the offset of the preceding vowel (first 3 Formants)

Temporal measurements included:

1. Closure duration in medial and final positions
2. Release duration
3. Preceding/following vowel duration

In addition, we also looked at phonation differences quantified by H1*-H2* (relative amplitude of the first two harmonics).

Both non-durational and temporal measurements were extracted automatically using Parselmouth in version 0.4.1 [15] with the internal Praat distribution in version 6.1.38, and VoiceSauce version 3.17 [16] was used for the H1*-H2* measurements. The acoustic data were analyzed using the statsmodels package in version 0.11.0 [17]. Student-T tests were performed to compare the groups of plain and ejective plosives for each place and each parameter. The parameters that showed a significant difference between the plain and ejective category were defined as the set of correlates for this distinction. This was done for all word positions. The importance of these correlates was then quantified by the calculation of D-prime values [18]:

$$Dprime = |\mu 1 - \mu 2| \,/\, ((\sigma 1 + \sigma 2) \,/\, 2) \quad (1)$$

3. Results

A total of 30 tokens were removed from the analysis due to mispronunciation. The results are thus based on the analysis of 390 tokens.

3.1. Non-temporal correlates

No significant differences were found for CoG and ΔF0 when comparing Mehri plain and ejective stops. We found differences in the intensity of stop release burst, the formant structure, and the presence/absence of post-release silent lags. The most systematic difference was found in the intensity of the release burst. Ejectives were released with higher burst intensity than their plain counterparts. This difference was significant for alveolars and velars in all word positions (Figure 1 illustrates these differences for medial and final positions).

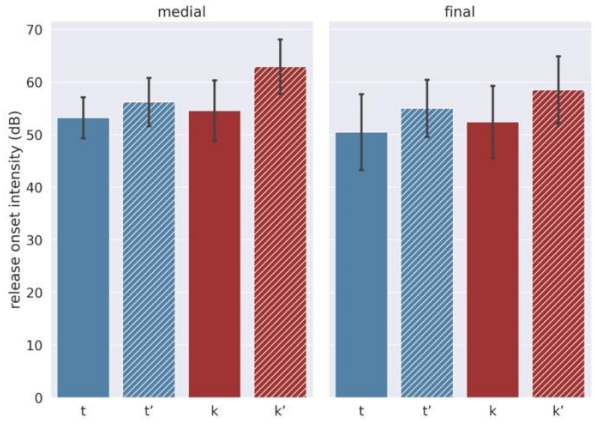

Figure 1: *Stop release intensity (in db) for alveolar and velar ejective/plain stops in medial and final positions.*

A prominent difference between ejective and plain stops was reflected in the formant structure of the following vowels: a substantial lowering of F2 and raising of F1 were observed in the context of ejectives.

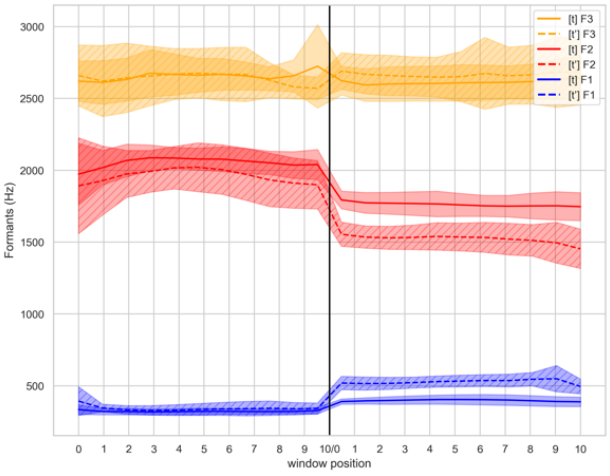

Figure 2: *Formant structure of the preceding (left) and the following vowel (right) for t/t' in medial position. The solid black line marks the stop boundary.*

The drop of F2 is similar to the effect induced by dorsopharyngealized consonants in Arabic [19]. Interestingly, while this effect was systematic for alveolars in all word positions (see Figure 2 for medial position), it noticeably varied for velars depending on word position. Figure 3 illustrates the absence of differences in F2 for plain and ejective velars in medial position.

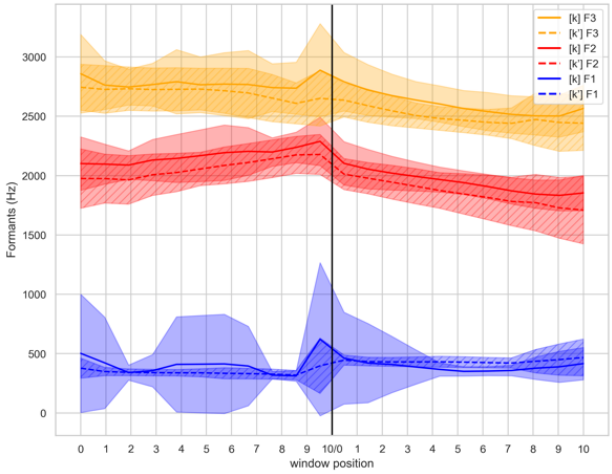

Figure 3: *Formant structure of the preceding (left) and the following vowel (right) for k/k' in medial position. The solid black line marks the stop boundary.*

In addition, the examination of acoustic waveforms and spectrograms showed that Mehri ejectives, unlike their plain counterparts, were sometimes produced with a post-release silent lag, a tell-tale characteristic of ejectivity. An example is shown in Figure 4. The presence of this silent lag varied depending on the speakers, however. While the majority of the productions of S2 (14/20) and S3 (12/20) displayed this feature, the realizations of the other speakers were (almost) devoid of it (4/20 for S4, 1/20 for S1 and 0/20 for S5, S6, and S7). These silent lags were slightly more frequent in velars (18) than in alveolars (13).

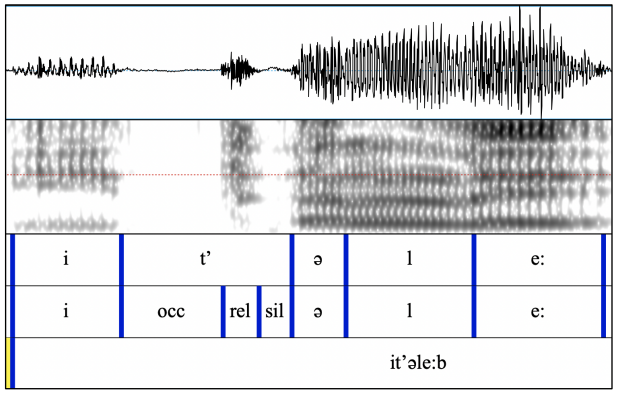

Figure 4: *Acoustic waveform and spectrogram taken from the item [it'əle:b] showing a silent gap following the release of [t'].*

3.2. Temporal correlates

Temporal parameters also displayed variability in the way they implement plain/ejective contrast. Some allowed to distinguish the two series of stops; for example, release duration was significantly shorter for ejectives. But this difference was observed for alveolars only in initial and medial positions (see Figure 5 for medial position).

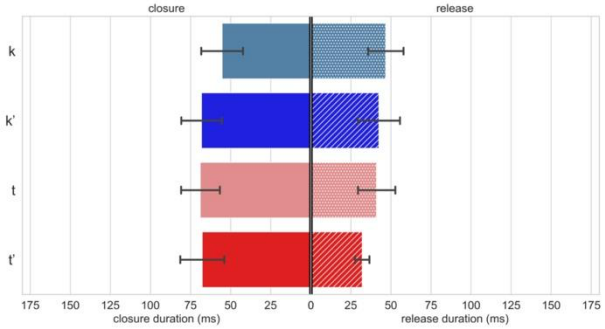

Figure 5: *Closure and release duration differences between velar and alveolar plain and ejective stops in medial position.*

Closure duration was significantly longer for ejective velars compared to their plain counterparts in medial and final positions. Adjacent vowels displayed the same durations for both stops, except in medial position where vowels preceding ejective alveolars were longer.

3.3. Phonation differences

Significant differences in phonation were also found, with vowels following ejectives displaying a lower H1*-H2* value. This difference was significant for alveolars and velars in initial position, and for alveolars in medial position (see Figure 6). No significant differences were observed for H1*-H2* in preceding vowels for both places of articulations.

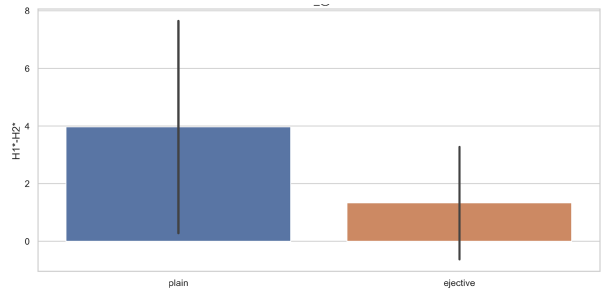

Figure 6: *H1*-H2* differences between plain and ejective velar stops in medial position.*

3.4. Weighting of the acoustic correlates

Figure 7 shows the D-prime values as calculated by formula (1). In word-initial position, the acoustic correlates of ejectivity were similarly distributed across the consonant and the vowel for alveolar stops, with F1 and F2 being the most informative correlates. Consonant-related attributes appeared to be more informative for velar stops with the release burst intensity as the most informative one. Word-medial stops had a strong tendency to signal ejectivity with differences in the formants of the following vowel, targeting again F1 and F2. Alveolar and velar stops in word-final position behaved differently. Consonant-

related attributes were more informative for the ejectivity distinction than correlates in the preceding vowel. Alveolar stops on the other hand relied on both the release onset and F3 similarly, with F1 being negligible.

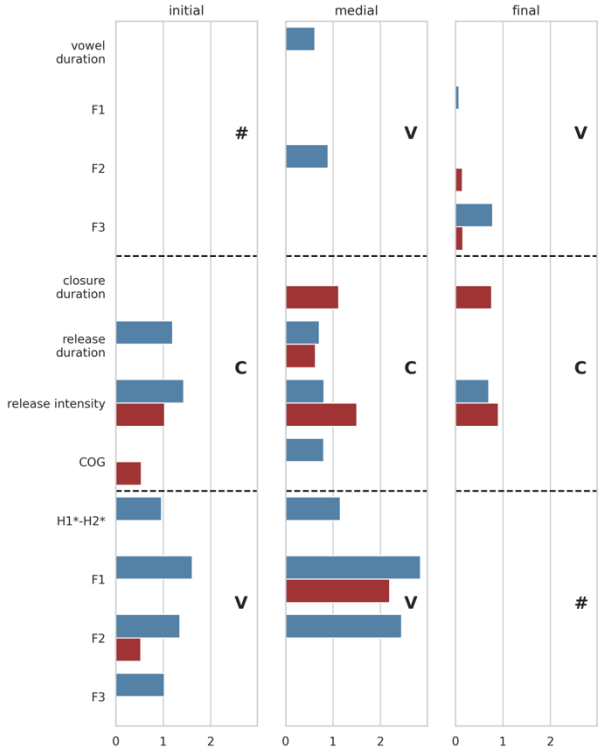

Figure 7: *D-prime values for alveolar (blue) and velar stops (red) in word-initial, word-medial and word-final position. A value of 1 means that the means of the two groups are 1 standard deviation apart from each other; the higher the value, the stronger the distinction.*

Taken together these results show that vowel-related correlates were more informative for alveolars in initial and medial positions, while consonant-related attributes were more informative for velars in the same word positions. This observation is interesting for the alveolar stops, as the high d-prime values and the formant trajectories suggest that these segments are better characterized as dorsopharyngealized in these positions. When looking at the D-prime values in word-final position, the opposite can be observed. Correlates of ejectivity are generally small and all less than 1, meaning the categories' means are less than one standard deviation apart from each other. This may indicate that the ejectivity distinction in this position is at least unstable or that neutralization of the plain/ejectivity contrast is well advanced.

4. Discussion and conclusion

The aim of this study has been to determine the acoustic correlates of ejective stops in Mehri, an endangered language particularly threatened by Arabic. We examined different parameters and showed that they varied depending on place of articulation, word position, and also speakers.

A common acoustic correlate of ejectivity generally observed in the world's languages is the silent lag following stop release. This silent gap, due to glottal closure, has been reported for ejectives, e.g., in Tigrinya [2], Ingush [20], Georgian [21], and Maya Yucatec [22]. Our results show that Mehri ejective stops could also be produced with post-release silent lags. This was not consistent across speakers, however. While two speakers displayed this characteristic for more than half of their productions, the five others never or almost never did.

Ejective stops displayed other characteristics of an ejective articulation. The most prominent one was the intensity of the release burst, which was higher compared to plain stops. This finding, which has also been reported for ejectives in other languages (e.g., in Ingush [20], and in Georgian [21]), suggests that Mehri ejective stops are produced with higher oral air pressure compared to plain stops. It is important to note that for alveolars, ejectives are on average 4dBs greater than their plain counterparts, just above the *Just notable difference* value of 3dB [23]. In addition, significant differences in phonation were also found, with vowels following ejectives displaying lower H1*-H2* values, suggesting a more creaky phonation. However, this difference was not consistent across places of articulation and word positions.

Different places of articulation showed varying patterns in how the formant structure of adjacent vowels was affected by ejectivity. Similar to what has been observed for pharyngealized consonants in Arabic, Mehri ejective alveolar stop /t'/ induced a substantial lowering of F2 combined with a raising of F1. This pattern was not observed for velars. A large decrease of F2 has been shown to be the main perceptual cue to dorsopharyngealization in Arabic [24]. According to [25], even a relatively small F2 lowering may be salient enough to perceive pharyngealization when F1 is high enough. As several scholars have already argued, one can reasonably attribute this pharyngealization to the long-lasting contact with Arabic. Massive borrowing of Arabic words, none of which has an ejective stop, should decrease the frequency of occurrence of these segments in the lexicon, which will naturally decrease the frequency of occurrence of ejective alveolars in Mehri speech, leading to a possible sound change.

The fact that velars keep their ejective articulation and do not pharyngealize could be accounted for on phonetic grounds ([26, 27]; see also [10]). As [28] pointed out, the back articulation for ejectives is based on a preference for a small air chamber, which makes it easier to build up air compression. It is interesting to note that Mehri /k'/ induces the same effect on following vowels as uvulars do in some Arabic varieties, i.e., raising of F1 but no drop of F2 ([29]). A plausible pathway of change is that ejective velar stop /k'/ become a plain uvular /q/, a change that has already been claimed from Proto-Semitic **k'* > to Arabic *q* [30].

The poor acoustic correlates of ejectivity for alveolars in word-final position is a diagnostic that the distinction in this position is at least unstable or that neutralization of the contrast is ongoing. Although this is speculative, the fact that real minimal pairs for the t/t' distinction are quite rare in this position implies that the two segments may be treated as homophones without affecting much comprehension.

5. Acknowledgements

This work was funded by a grant from the French Agence Nationale de la Recherche Scientifique (ANR program ALMAS) and by the Labex EFL (ANR/CGI). We would like to warmly thank the seven Omani participants for their help and enthusiasm during the study visit of the first author in Salalah.

6. References

[1] M. Lindau, "Phonetic differences in glottalic consonants," *Journal of Phonetics,* vol. 12, pp. 147–155, 1984.

[2] J. Kingston, *The phonetics and phonology of the timing of oral and glottal events*, University of California, Berkeley, 1984.

[3] N. Warner, "Acoustic characteristics of ejectives in Ingush," *International Conference on Spoken Language Processing,* pp. 1525–1528, 1996.

[4] R. Hertzon, "La division des langues sémitiques," *Actes du premier congrès international de linguistique sémitique et chamito-sémitque*, pp. 181–94, 1974.

[5] C. Moseley, *Atlas of the World's Languages in Danger*, UNESCO Publishing, Paris, 2010.

[6] T. M. Johnstone, "The Modern South Arabian languages," *Afroasiatic Linguistics* vol. 1, no 5, pp. 93-121, 1975.

[7] A. Lonnet, "Quelques résultats en linguistique sudarabique moderne," *Quaderni di Studi Arabici* vol. 11, pp. 37–82, 1993.

[8] A. Lonnet, and M. C. Simeone-Senelle, "La phonologie des langues sudarabiques modernes," in A. Kaye, (ed.), *Phonologies in Asia and Africa* 1, Winona-Lake: Einsenbrauns, pp. 337–372, 1997.

[9] J. Watson, *The Structure of Mehri,* Wiesbaden: Harrassowitz, 2012.

[10] J. Watson, and A. Bellem, "A detective story: Emphatics in Mehri" *Proceedings of the Seminar for Arabian Studies* vol. 40, 345–356, 2010.

[11] A. Rubin, *The Mehri language of Oman*, Leiden: Brill, 2010.

[12] A. Jahn, *Die Mehri-Sprache in Südarabien: Texte und Wörterbuch*, Wien: Hölder, 1902.

[13] R. Ridouane, and C. Gendrot, "On ejective fricatives in Omani Mehri," *Brill's Journal of Afroasiatic Languages and Linguistics*, vol. 9, no. 1-2, pp. 139–159, 2017.

[14] P. Boersma and D. Weenink, "Praat: doing Phonetics by Com- puter," v. 6.2.09, 2022, https://www.fon.hum.uva.nl/praat/.

[15] Y. Jadoul, B. Thompson, and B. de Boer, "Introducing Parselmouth: A Python interface to Praat," *Journal of Phonetics*, vol. 71, pp. 1–15, 2018.

[16] Y. L. Shue, *The voice source in speech production: Data, analysis and models*, UCLA dissertation, 2010.

[17] S. Seabold, and J. Perktold, "Statsmodels: Econometric and statistical modeling with python," *Proceedings of the 9th Python in Science Conference*, vol. 57, 2010.

[18] M. A. Clayards, "The Ideal Listener: Making Optimal Use of Acoustic-Phonetic Cues for Word Recognition," PhD dissertation, University of Rochester, 2008.

[19] M. Embarki, "Phonetics," in J. Owens, (ed.), *The Oxford handbook of Arabic linguistics*, Oxford University Press, 23–44, 2013.

[20] N. Warner, "Acoustic Characteristics of Ejectives in Ingush," *Proceedings of the International Conference on Spoken Language Processing*, Philadelphia, pp. 1525–1528, 1996.

[21] C. Vicenik, "An acoustic study of Georgian stop consonants," *Journal of The International Phonetic Association,* vol. 40, pp. 59–92, 2010.

[22] E. Bayraktar, and R. Ridouane, "Le vot des éjectives: le cas du maya yucatèque," *Actes des 29e Journées d' Etudes sur la Parole*, Paris, 2016.

[23] F. E. Toole, and S. E. Olive, "The modification of timbre by resonances: Perception and measurement," *Journal of the Audio Engineering Society*, vol. 36(3), 122–142, 1988.

[24] D. Obrecht, "*Effects of the Second Format in the Perception of Verlarization in Lebanese Arabic*," Ph.D. Dissertation, University of Pennsylvania, 1961.

[25] M. Yeou, "Pharyngealization in Arabic: Modelling, acoustic analysis, airflow and perception," *Revue de La Faculté des Lettres El Jadida*, vol. 6, pp. 51–70, 2001.

[26] I. Maddieson, *Patterns of Sounds*. Cambridge: Cambridge University Press, 1984.

[27] H. Javkin. "Towards a phonetic explanation for universal preferences in implosives and ejectives," *Annual Meeting of the Berkeley Linguistics Society*, vol. 3, pp. 559–565, 1977.

[28] J. H. Greenberg, "Some generalizations concerning glottalic consonants, especially implosives," *International journal of American linguistics*, vol. 36, no. 2, pp. 123–145, 1970.

[29] B. A. Zawaydeh, "The interaction of the phonetics and phonology of gutturals," *Papers in Laboratory Phonology* vol. 6, pp. 279–292, 2003.

[30] L. Kogan, "Proto-Semitic Phonetics and Phonology," *The Semitic Languages*, pp. 54–151, 2011.