THE ACOUSTIC CORRELATES OF VOWEL PHARYNGEALISATION IN ARCHI (EAST CAUCASIAN)

Alexandre Arkhipov

Moscow State University
sarkipo@mail.ru

ABSTRACT

The paper offers an exploratory analysis of acoustic properties of pharyngealised vowels in Archi (East Caucasian). Three speakers considered show variation in pharyngealisation correlates. Only one speaker is consistent in F3 lowering and duration increase across most vowels.

Keywords: Archi, pharyngealisation, F3, duration.

1. INTRODUCTION

Archi has a basic set of 5 vowel qualities plus schwa. The 5 full vowels show a length contrast (although long *u:* and *i:* are rare); moreover, all vowels including schwa can bear pharyngealisation which is phonologically distinctive. S. Kodzasov [2] treats pharyngealisation as a prosodic feature that applies to a syllable or to a group of syllables; its phonetic exponence is governed by certain rules. All segments except dentals can show different degrees of pharyngealisation. In particular, a uvular consonant, if present, becomes the primary locus of pharyngealisation, with neighbouring segments having weaker pharyngealisation. In the absence of a uvular consonant, it is generally the stressed vowel that becomes the main pharyngealisation locus, also spreading the feature over to its neighbours. In this paper, I focus on the pharyngealisation effects on stressed vowels.

2. GOALS

The main objective of this study was to reassess the acoustic part of the description of the Archi pharyngealisation as appears in [2]. The main findings so far can be summarized as follows. The articulatory mechanism of Archi pharyngealisation is the narrowing of the lower pharynx along with the backwards tongue displacement. The body of the tongue is retracted and lowered, while its front part is retracted and raised. The acoustic effect on vowels is different depending on vowel backness: for the front vowels *e*, *i* pharyngealisation disfavours higher frequencies, either by lowering F2 and F3, or by lowering the overall intensity in the corresponding part of the spectrum. For central and back vowels *a*, *o*, *u*, however, F2 appears to be raised, which is attributed to the raised tongue tip position. In addition, it was observed that pharyngealised vowels show somewhat longer duration than the corresponding plain vowels (both short and long).

Generally for Caucasian pharyngealisation, F3 lowering is seen as the main acoustic effect, similarly to the rhotacized vowels in American English [3].

3. MATERIAL

For the present study, a subset of data recorded under the "Five Languages of Eurasia" documentation project (2006-2010) was used, namely a part of the "phonetic database". It is a selected list of words (mostly citation forms, but occasionally specific grammatical forms) designed by S. V. Kodzasov to cover all the Archi phonemes in diverse phonological contexts. It is however not balanced with respect to different phonemes or kinds of contexts. The list was recorded with 6 adult speakers (3 men and 3 women) in 2006; in this paper, I consider data from 1 man and 2 women. Each word was pronounced twice in isolation, without a carrier phrase. The equipment used was a digital MicroTrack 24/96 recorder with an AKG C-1000 condenser microphone, recording at 44 kHz/16 bit.

The number of vowel tokens was not uniform across vowel features and differed slightly between speakers. A summary is given in Table 1.

Table 1: Stressed vowel token counts for the three speakers (JSP, HDK, PSX). Short and long vowels are collapsed.

	JSP (m)		HDK (f)		PSX (f)	
vowel	–ph	+ph	–ph	+ph	–ph	+ph
a	76	22	59	21	65	21
e	32	26	28	20	25	27
i	49	4	37	6	50	7
o	42	21	24	22	34	24
u	22	9	20	10	36	5

4. PROCEDURE

4.1. Transcription, segmentation and measurements

The recordings were transcribed and manually segmented in Praat. The relevant part of annotation included, for each pronunciation, a phonological transcription and a (broad) phonetic transcription, which was subjected to phone-level segmentation.

The following measures were taken: for each vowel segment, total duration and the mean frequencies of F1, F2, F3 within the 40% central part of the segment's duration (measurements automated and written to a table).

4.2. Segmentation issues and duration

Concerning vowels, the most problematic cases for segmentation were word-final vowels. Given the isolated words pronunciation style, these vowels systematically present very long "tails" of decreasing intensity and waning formant structure. In order to improve comparability with word-medial positions and get more reliable formant frequencies, it was decided to cutoff at places where the higher formants cease to be visible. It should thus be kept in mind that these borders, and accordingly the duration figures, are to some extent arbitrary.

Although the durations of all vowels in the data are higher than in normal speech, the overall duration correlations seem to be preserved quite well. The phonemically long vowels are in general longer (~300-450 ms) than the short ones (~150-250 ms) and the schwa (~50-100 ms).

4.3. Formant frequencies

It is well known that the formant frequencies measured with the BURG algorithm in Praat depends considerably on the initial settings, mainly on the frequecny ceiling below which the preestimated number of formants is calculated. As was shown in [1] for Portuguese vowels, the optimal ceiling for the F1 and F2 calculation varies for different vowel qualities increases from back (rounded) to front (spread) vowels.

The optimal settings for the formant analysis were chosen with a script[1] inspired by the algorithm described in [1]: for each vowel category (e.g. "i"), all formants measures for all segments were taken with the default BURG algorithm in 21 iterations, changing the ceiling of the presumed F1-F5 range in 50 Hz increments (e.g. 4000 Hz, 4050 Hz,... 5000 Hz). From these measures, the setting yielding the minimal variation of the F1 and F2 was retained. As expected, the settings retained by the algorithm were generally increasing along the back-front axis.

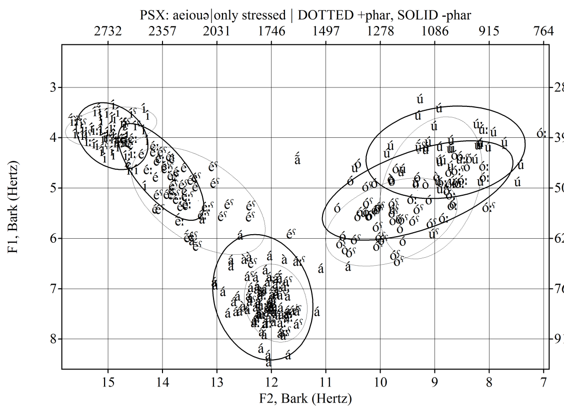
However, the ceilings optimized for F1 and F2 did not always perform equally well for F3 measurements. The algorithm was then changed to optimize for F1, F2 and F3 variation, and the ranges of the ceiling iterations were separated for front, mid and back vowels, starting 1 kHz higher for *i*, *e* and 0.5 kHz higher for *a* than for *o*, *u*. However, upon inspection of selected tokens it was noticed that F3 was still incorrectly estimated in some cases.

More consistent results were obtained with separate calculations for pharyngealised and non-pharyngealised vowels. It should be noted, however, that a more thorough manual checking of the formant values might be required, since the technique used does not guarantee from errors, especially in the extreme points of the vowel space.

4. RESULTS AND DISCUSSION

The data from the three speakers turned out to be dissimilar in some ways. Only one speaker (PSX, f, 60 y.o.) showed consistent F3 lowering for almost all [+phar(yngealised)] vowels (except *i*). She also presented duration increase in more vowels than other speakers. Several of her (short) vowels, namely *e*, *o*, *u*, are slightly more open (higher F1) and centralised (higher/lower F2) when pharyngealised.

Figure 1: PSX: pharyngealised (dotted) vs. plain vowels (solid) in F2-F1 space (Bark scale). Ellipses at 2 sigmas.



The second speaker, JSP (m, 80 y.o.), only showed F3 lowering in back vowels, while in front vowels F3 was, suprisingly, higher; it remains to be verified if this could be due to an algorithm error. The increase in duration is seen in $e^{ʕ}$, $eː^{ʕ}$, and $o^{ʕ}$, with a slight decrease for $u^{ʕ}$. Long pharyngealised $eː^{ʕ}$ and $oː^{ʕ}$ are more open. Finally, both $a^{ʕ}$ and $aː^{ʕ}$ are only distinguished by fronting (higher F2).

Figure 2: PSX: pharyngealised (dotted) vs. plain vowels (solid) in F2-F3 space (Bark scale).

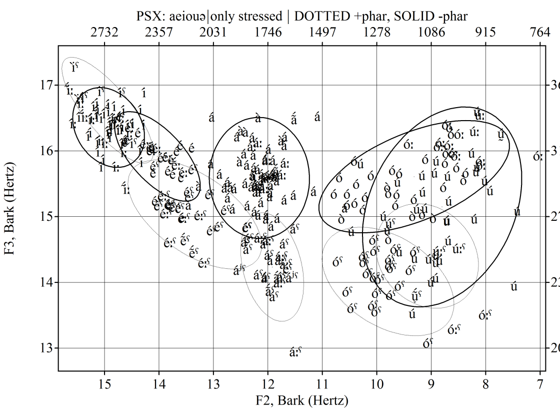


Figure 3: JSP: pharyngealised (dotted) vs. plain vowels (solid) in F2-F1 space (Bark scale).

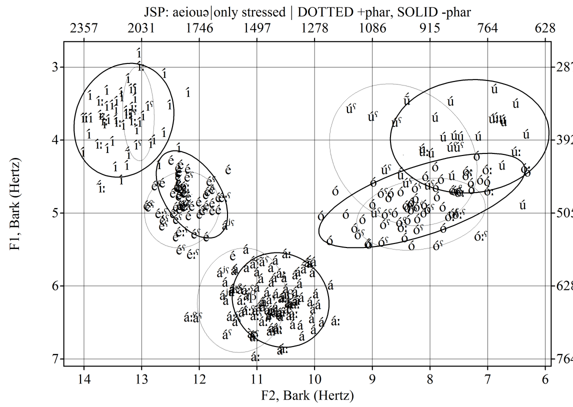
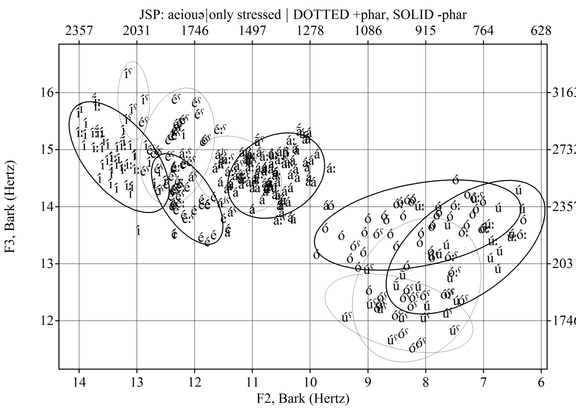


Figure 4: JSP: pharyngealised (dotted) vs. plain vowels (solid) in F2-F3 space (Bark scale).



The third speaker, HDK (f, 35 y.o.), also presented F3 lowering mostly in back vowels (both short and long), with a slight lowering for *e*. The short $i^{ʕ}$, $e^{ʕ}$ and $u^{ʕ}$ tend to be slightly more open than their non-pharyngealised counterparts. Increase in duration is only observed for $a^{ʕ}$ (both short and long) and short $e^{ʕ}$, while the duration of $u^{ʕ}$ decreases slightly, as for JSP.

Figure 5: HDK: pharyngealised (dotted) vs. plain vowels (solid) in F2-F1 space (Bark scale).

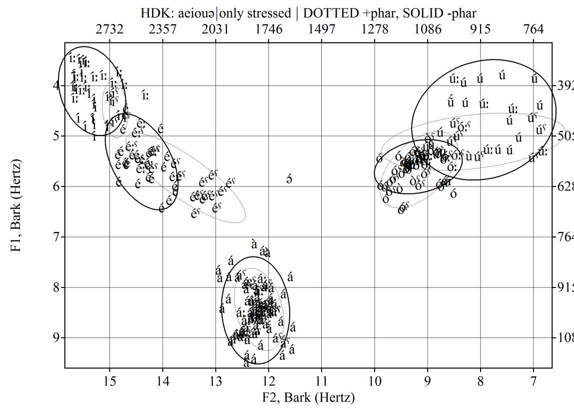


Figure 6: HDK: pharyngealised (dotted) vs. plain vowels (solid) in F2-F3 space (Bark scale).

As a general conclusion, on the data and parameters studied, the speakers were shown to differ in their manifestation of vowel pharyngealisation. Several enhancements can be proposed to get a more consistent view of the phenomenon. Most obviously, including more tokens and data from more speakers are likely to improve the correlations observed. Another step would be distinguishing between various phonetic contexts (e.g. word-initial, interconsonantal and word-final).

Furthermore, a number of other parameters may be suggested for further study: F0 and phonation, spectral slope, and the dynamics of the formant frequencies.

7. ACKNOWLDEGEMENTS

This paper is dedicated to the memory of Sandro V. Kodzasov (1938-2014).

The study was partially supported by RFH grant № 14-04-12038 and RFBR grant № 15-06-06103.

8. REFERENCES

[1] Escudero, P., Boersma, P., Schurt Rauber, A., Bion, R. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *J. Acoust. Soc. Am*. 126, 1379–1393.

[2] Kodzasov, S. V. 1977. Phonetics [Fonetika]. In: Kibrik, A. E., Kodzasov, S. V., Oloviannikova, I. P., Samedov, D. S. *Opyt strukturnogo opisanija archinskogo jazyka*, v. 1: *Leksika. Fonetika*. Moscow: MGU.

[3] Ladefoged, P., Maddieson, I. 1996. The Sounds of the World's Languages. Oxford: Blackwell.

[1] The core part of this script for 2 formants was written by Ruslan Idrisov. All the other scripting used for this paper was mine.