THE VOICELESS VELAR STOP /k/ IN RIJAL ALMA ARABIC: REVISITED

Leila Yaqoub[1], Sam Hellmuth[2], George Bailey[3]

[1]King Khalid University, [1, 2, 3]University of York
lyacob@kku.edu.sa, sam.hellmuth@york.ac.uk, george.bailey@york.ac.uk

ABSTRACT

An earlier documentation study showed that the voiceless velar stop /k/ is articulated in Rijal Alma Arabic (spoken in southwest Saudi Arabia) as a post-palatal fricative /x̟/, a phoneme absent from most other Arabic varieties. However, a recent sociolinguistic study in Rijal Alma revealed patterns of variation and change in use of post-palatal /x̟/ in favour of the supra-local voiceless velar stop /k/.

To better understand variation and change in the speech patterns of speakers of Rijal Alma, this paper presents an auditory and acoustic analysis of the realisation of /k/ in three positions (word-initial, intervocalic, and word-final). The data are from wordlist fieldwork recordings of 33 target items produced by 43 Rijal Alma speakers (24M/19F aged 18-60). Results indicate that target /k/ displays more complex phonetic properties than previously reported, pointing to a range of different realisations, which appear to be both auditorily and acoustically distinct: [k x kx xk].

Keywords: Rijal Alma dialect, velar stop /k/, auditory realisation, acoustics measurements.

1. INTRODUCTION

Rijal Alma (RA) is a variety of Tihamat Asir Arabic spoken in the RA province of Saudi Arabia in the south-west with a population of approximately 65406 [1].

The phonology of RA has been reported to differ from Modern Standard Arabic (MSA) and most Arabic dialects in a number of respects. A notable difference is the absence in RA of the voiceless velar stop /k/. In most Arabic varieties, the voiceless velar stop /k/ is a common sound. However, in some Arabic varieties, the voiceless velar stop /k/ can also be realised as either affricate [t͡ʃ], as in some rural dialects of the Levant, parts of Jordan and Iraq dialects and Gulf dialects, or as affricate [t͡s], as in Najdi Arabic [2], where the occurrence of [t͡ʃ] or [t͡s] is conditioned mainly in front vowel environments [3].

In contrast, in RA, the voiceless velar stop /k/ has been proposed to manifest instead as a fricative sound [4, 5, 6, 7, 8]. Prochazka [4] described the realisation of voiceless velar stop /k/ in RA as a voiceless uvular /χ/. A another description by Hamzah cited in [6] suggested that target /k/ in RA is perceived as a mixture of two fricatives /ʃ/ and /χ/.

An acoustic analysis conducted by Asiri [5, 6] describes the production of target /k/ by native RA speakers as a post-palatal fricative [x̟], occurring in all phonological environments. Use of the local [x̟] was, however, observed to be reducing in favour of the voiceless velar stop /k/ particularly among educated speakers [7]. There is, however, a lack of evidence regarding how the observed variants of local [x̟] are phonetically characterised. In the previous studies [5, 6], the focus was only on the phonological behaviour of the local [x̟], with only limited acoustic investigation.

This paper presents the first preliminary results of a wider study into linguistic variation in Rijal Alma. Our focus here is the evidence for variation in the phonetic realisation of the local variant [x̟] in the speech of native speakers of Rijal Alma, by examining the auditory and acoustic characteristics of their speech.

2. METHODS

2.1. Participants, Materials and Procedure

The data for this study stems from recordings of 43 RA speakers, comprising two genders and are balanced by age (24 men and 19 women aged 18 to 60). The participants were asked to read 33 words as naturally as possible. The target words are placed in a carrier phrase, "ʔana ʔagool [target word] marratti:n" ~'I say [target word] twice'.

Each participant was recorded on location in RA using a Marantz model PMD660 with an external microphone (Shure SM10 professional unidirectional head-worn dynamic microphone). The recording was saved in WAV format, with a sampling rate of 44.1 KHz 16 bit.

2.2. Data Preparation and Auditory Analysis

The data was phonetically transcribed using Al-Tamimi Romanization system (ATR) [9] in ELAN [10] then forced aligned at the phone level using WebMaus [11] for further acoustic analysis. Manual auditory coding was performed by the first author. Auditory judgements were made using Praat [12] with close reference to the waveform and spectrogram. The local variant [x̣] was found to display a number of complex phonetic structures that can be grouped into three variants.

A token was coded as a plosive [k] if it has three clear phases of articulation; closure, burst, aspiration (as in Figure 1(a)). Tokens with aperiodic energy typical of a fricative, and perceived as a fricative, were coded as a fricative [x] [i](Figure 1(b)). In some cases, however, potential tokens of [k] had heavy aspiration or frication noise after the closure phase and were perceived auditorily as a combination of stop and fricative [kx] (Figure 1(c)). In addition, there was a further variant classified as [xk], in which there is a period of aperiodic frication before the transient burst begins, and in some cases, an incomplete closure phase is present (Figure 1(d)). The two variants [kx] and [xk] are labels used during coding for ambiguous or intermediate variants which exhibit some characteristics of both plosive and fricative sounds; they comprise 11% of the sample.

Most variants often involved a glottal presence in which it appeared that two constrictions were occurring, one at the glottis and the other at one or more phases of articulation (e.g., before or after closure, before or after frication, or after aspiration), though this rarely developed into a non-pulmonary stop such as clicks.

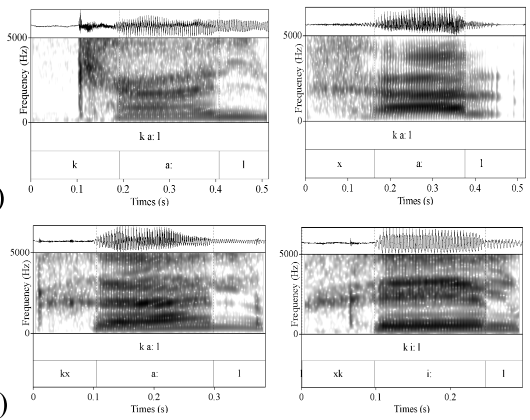

Figure 1: Waveforms and spectrograms of [ka:l] 'weight': a) a plosive variant [k]; b) a fricative variant [x], and two with intermediate variants c) [kx] and d) [xk].

2.3. Acoustic Measurements

Acoustic measurements were also taken to corroborate the auditory coding and provide a better basis for distinguishing tokens heard as plosive [k], fricative [x] or intermediate [kx]/[xk]. Following prior literature on measuring the duration of obstruent consonants [13, 14, 15, 16, 17, 18] each variant [k], [x], [kx], or [xk] was measured according to the various acoustic events associated with each variant (closure, burst, aspiration/frication). This allows comparison of the duration of a certain phase between variants and provides more comparable results to prior literature regarding the distinction between stops, affricates and fricatives. In this report, non-pulmonic variants were excluded due to their small number (0.96% of the sample).

The whole segment duration of variants coded as plosive [k] was defined based on its acoustic events (closure, burst, aspiration), in which the onset and offset of each phase was marked (see Figure 2). In cases where multiple release bursts were observed, the whole segment duration was taken from the first to final transient at the release.

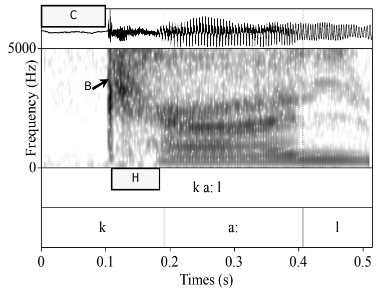

Figure 2: Example of segmentation for [k], showing hold phase (C), burst release (B) and aspiration (H).

As for the variant coded as [x], there was no closure or burst, and the whole segment duration was defined as the interval between the onset and offset of the frication noise. The whole segment duration for variants coded as [kx] or [xk] were defined in a similar manner to the durations defined for the variant coded as [k], where the time between onset and offset of each phase of articulation was defined. However, in cases where no closure phase occurred, only burst and aspiration duration measurements were recorded.

The second acoustic measurements used in this study is rise time to gain a further insight in distinguishing different possibilities among the sound realised as a stop and other obstruent that were identified auditorily. It has previously been suggested that the rise time of the noise energy of obstruent is an important cue to the manner of articulation [17]. In this document, a rise time is measured by obtaining the duration from the time of onset of aspiration/frication noise to the time at which peak amplitude reaches its maximum.

The acoustic measurements (durations and rise time) were analysed statistically using mixed-effects linear regression model in the lme4 package [19] in R to assess whether there were any overall differences among the variant types. The data was fitted with the following model (DV (acoustic measurements) ~ auditory label * word_position + (1 | word) + (1 | speakerID). Acoustic measurements are the dependent variable, which is expressed as a continuous variable. The fixed effects are auditory labels, and word position. Words and speakers were fitted as random intercepts in order to account for individual differences.

3. RESULTS

The data examined in this document were derived from the wordlist task outlined above. The sample included 24 males and 19 females, balanced by age. Specifically, the younger participants were aged between 18 and 35 years, while the older participants were 40 years or older. The total number of auditorily labelled tokens from the wordlist data included 2888 tokens. Among the 2888 tokens, there were 2152 (74%) tokens were realised as [k], of which 48% were produced by younger participants and 26% by older participants, 447 (15%) tokens were realised as [x] , 7 % are produced by younger participants and 8% by older participants, 157 (6%) tokens were realised as [kx], 2% of them are produced by younger participants and 4% by older participants, and 132 (5%) tokens were realised as [xk], 2% of which are produced by younger participants and 3% older participants.

3.1 Duration

The distribution of durations for the whole segment of all variant types are presented in Figure 3. The mean duration of tokens labelled [k] (annotated on the plot) appears to be longer than that of other variants.

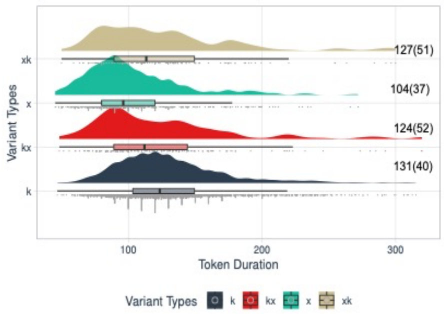

Figure 3: Mean and standard deviation (ms) given in brackets of Total segment duration for all variants.

A closer look at the durations of each component phase is provided in Table 1. As predicted, tokens with intermediate variants [kx, xk] appear to have a shorter mean duration of closure and burst than tokens coded as a stop [k]. Further, the mean duration of aspiration for the intermediate variants [kx, xk], is longer than the mean duration of aspiration for the variant coded as a stop [k]. In a comparison of tokens realised auditorily as fricatives [x] and intermediate variants [kx, xk], it was found that the mean duration of aspiration for intermediate variants was shorter than that for tokens realised auditorily as fricatives [x].

Variant:	k	x	kx	xk
Closure	69.67 (27.8)	- -	28.9 (31)	22 (28.9)
Burst	12.1 (9.09)	- -	12.7 (12)	9.57 (9.56)
Aspiration	47.7 (23)	- -	82.3 (44.1)	94.9 (50.6)
Frication	- -	104 (35.9)	- -	- -

Table 1: Mean and standard deviation (ms) (in brackets) of closure, burst, aspiration and frication durations by variant.

A statistical analysis performed by using linear mixed effects models confirms the observations from the raw data; closure duration was significantly shorter for intermediate variants [kx] (β= -7.005, SE = 3.418, t = -2.049, $p < 0.04$), and [xk] (β= -32.424, SE = 3.009, t = -10.777, $p < 0.001$) than for the stop [k], across all different word positions (see Figure 4). There was no statistically significant difference between the burst durations of the stop and intermediate variants.

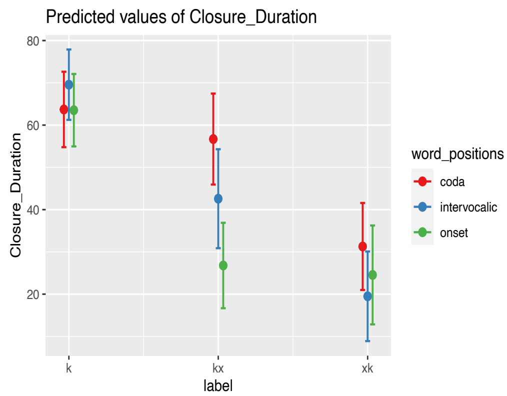

Figure 4: Predicted mean and 95% CI of closure duration for variants [k, kx, xk] by word position.

As for aspiration duration, the statistical results show a significant effect with a longer duration for the intermediate variants [kx] (β= 28.325, SE = 3.519, t = 8.048, $p < 0.001$), and [xk] (41.583, SE = 3.097, t = 13.428, $p < 0.001$) compared to the variant [k], across all different word positions. It appears that these results are in agreement with the literature concerning durational differences between stops and affricates. According to [14] when affricates are compared with stops of the same voicing category and manner of articulation, the noise stage appears to be lengthened while the silence stage is shortened.

Furthermore, the statistical results confirm the longer duration for the fricative form [x] in comparisons with the intermediate variants (β= 27.913, SE = 5.226, t = 5.186, p < 0. 001) compared to [kx], and [xk] (-13.252, SE = 4.897, t = -2.706, p < 0.007) is significantly shorter than [x]. In light of this, the duration means of the fricative variant [x] is longer than that of intermediate variants, which further indicates that intermediate variants might be classified as affricates. Considering the literature that examines the durational differences between affricates and fricatives, this is an expected result in which it suggests that the affricate sounds are found to be distinguished by a shorter duration for the fricative portion than for the corresponding fricative part [14, 15].

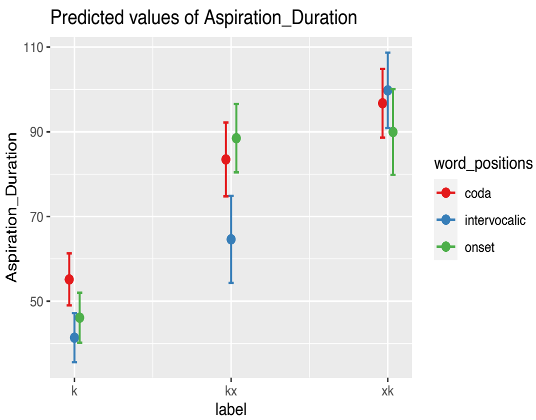

Figure 5: Predicted mean and 95% CI of aspiration duration for variants [k, kx, xk] by word position.

3.2. Rise Time of Stop - Intermediate - Fricative

The values of rise time by variant are presented in Figure 6 and Table 2. The table contains the mean values and (SD) of rise time in aspiration/frication phases across 2152 [k], 447 [x], 157 [kx], and 132 [xk] tokens from 43speakers. In general, the rise time variation between intermediate and fricative variants does not appear as large as the variation between intermediate, stop variants. However, there are still noticeable differences that may be better accounted for by a linear regression model.

Variant:	k	x	kx	xk
Rise time	21.2 (21.3)	48.6 (55.5)	52.7 (45.5)	42.1 (34. 4)

Table 2: Mean and standard deviation (in brackets) of rise time (ms) by variant.

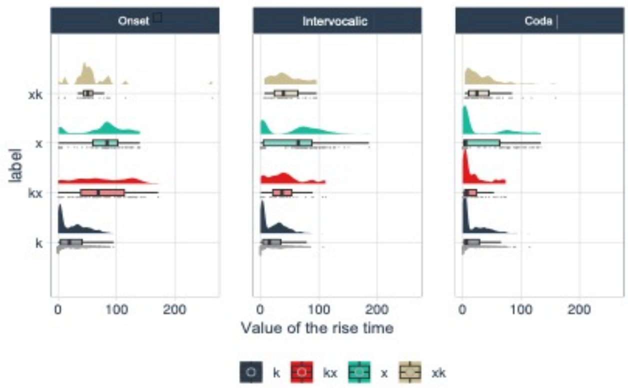

Figure 6: Distribution of rise time values (ms) for each variant type.

A mixed-effects linear regression revealed a significant difference with a longer rise time for the [x] (β= 9.432., SE = 4.346, t = 2.170, p < 0.03) relative to intermediate variant [kx], but no statistically significant difference between [x] and [xk]. Additionally, a significant difference is found between the two intermediate variants with a longer rise time for the variant [kx] (β= 22.769, SE = 2.229, t = 10.21, p < 0.001) compared to the plosive variant [k], and a longer rise time for the intermediate variant [xk] (β= 19.554, SE = 2.130, t = 9.18, P < 0.001) relative to the plosive variant. Accordingly, the plosive variant [k] exhibits the most rapid build-up of acoustic energy, followed by intermediate and then fricative variants. According to the wider literature examining how the obstruent can be phonetically distinguished, this seems to be an expected result, which indicates that noise energy is gradually built up in fricative consonants; it increases faster in affricates than in fricatives, but not as rapidly as in stop consonants [14, 15].

4. CONCLUSIONS

In this document, we have attempted to reconsider the evidence for variations in the phonetic realisation of the local variant [x̹] in the speech of native speakers of Rijal Alma, by examining the auditory and acoustic characteristics of their speech. Although these results are in their preliminary status, they indicate quite clearly that the local variant previously described as (x̹) in Rijal Alma possesses complex phonetic patterns. Based on the acoustic measurements presented in this document, it appears that there are four variants of the local variable (x̹) which are acoustically distinct from each other in Rijal Alma. However, a future sociolinguistic study is still needed to investigate possible associations with specific social categories.

5. ACKNOWLEDGEMENTS

Financial support is gratefully acknowledged from King Khalid University, Saudi Arabia.

6. REFERENCES

[1] General Authority for Statistics Kingdom of Saudi Arabia. (2022). *Population and Housing census.* Available: https://www.stats.gov.sa/en/13

[2] Ingham, B., & Henry Sweet Society for the History of Linguistic Ideas, provenance. 1994. *Najdi Arabic: central Arabian / Bruce Ingham.* Amsterdam; Philadelphia: J. Benjamins Pub. Co.

[3] Johnstone, T. M. 1967. *Eastern Arabian dialect studies / by T. M. Johnstone.* London; New York [etc.]: Oxford U.P.

[4] Prochazka, T. 1988. *Saudi Arabian dialects / Theodore Procházka Jr.* London; New York: Kegan Paul International: Distributed by Routledge, Chapman, and Hall.

[5] Asiri, Y. 2009. "Aspects of the Phonology and Morphology of Rijal Alma'Dialect (south-west Saudi Arabia)", Doctoral dissertation, University of Salford. Manchester

[6] Asiri, Y. 2009. "Remarks on the dialect of Rijal Alma'(South-west Saudi Arabia)". *Wiener Zeitschrift für die Kunde des Morgenlandes*, 9-21.

[7] Assiri, A. 2014. "Sociolinguistic variation in Rijaal Almaʕ, Saudi Arabia: a dialectological study". *King Khalid University Journal for Humanities*, *23*(2), 73–119.

[8] Watson, Janet. 2018. "South Arabian and Arabic dialects". In C. D. Holes (Ed.), *The Historical Dialectology of Arabic.* Oxford University Press. 316–334.

[9] 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". *Proc.13th Conference on Language Resources and Evaluation (LREC 2022)* Marseille, 7269–7276.

[10] Sloetjes, & Wittenburg, L. 2008. "Annotation by Category: ELAN and ISO DCR". *Proc.6th International Conference on Language Resources and Evaluation*. Retrieved from https://archive.mpi.nl/tla/elan".

[11] Kisler, T., Reichel, U., & Schiel, F. 2017. "Multilingual processing of speech via web services". *Computer Speech & Language*, *45*, 326-347.

[12] Boersma, & Weenink. 2020. *Praat: doing phonetics by computer*.

[13] Johnson, K. 1997. *Acoustic and auditory phonetics / Keith Johnson.* Cambridge, Mass.: Blackwell Publishers.

[14] Hayward, K. 2000. *Experimental phonetics/ Katrina Hayward.* Harlow, Eng.: Longman.

[15] Kent, R. D., & Read, Charles. 2002. *The acoustic analysis of speech / Ray D. Kent and Charles Read.* (2nd ed.). Albany, NY: Singular Press.

[16] Ladefoged, P. 2003. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques / Peter Ladefoged.* Oxford: Blackwell Pub.

[17] Mitani, S., Kitama, T., & Sato, Y. 2006. "Voiceless affricate/fricative distinction by frication duration and amplitude rise slope. *The Journal of the Acoustical Society of America*, *120*(3),1600–1607.

[18] Foulkes, P., Docherty, G., & Jones, M. 2011. Analyzing stops. In D. P. Marianna & M. Yaeger-Dror (Eds.), *Sociophonetics: A Student's Guide*. Routledge, 58–71.

[19] Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. 2015. "Fitting linear mixed-effects models using lme4". *Journal of Statistical Software*, *67*(1), 1-48.

[i] The transcription [x] is used here to indicate a fricative variant of the voiceless velar stop /k/, but this does not imply that it is identical to the target realisations of the phoneme /x/ used in other Arabic dialects, and a detailed analysis of this area is beyond the scope of this paper. A further work focused on the fricative variant in RA is required so that the exact place of articulation can be classified.