Variable Word Onset Vowel Deletion in Najdi Arabic

Rana Alsabhan[1], Jane Setter[2]

[1]Princess Nourah Bint Abdulrahman University, [1,2]University of Reading
r.alsabhan@pgr.reading.ac.uk, j.e.setter@reading.ac.uk

ABSTRACT

The literature on Arabic dialects agrees that short vowels are deleted word-initially, resulting in surface consonant clusters. This study supplements the existing literature by acoustically analysing word onset vowel deletion in Najdi Arabic (NA), a dialect spoken in central Saudi Arabia. It aims to examine the phonetic deletion of the vowel within an obstruent context. Given the relatively lower frequency of obstruents cross-linguistically [1], the abundance of such sounds in the Arabic language presents a compelling context for investigating vowel deletion and its conditioning factors.

Forty-five NA native speakers (31 females, 14 males) were recruited to do three production tasks: diapix, reading and shadowing tasks. The results showed that 54% of the data were produced with vowel deletion, the favouring factor being the place of articulation of the first obstruent, where the deletion was more likely to occur following a dorsal obstruent than following coronal, labial and pharyngeal obstruents. The gender of the speaker also influenced vowel deletion, where male listeners were more likely to delete more vowels than female speakers.

Keywords: vowel deletion; Najdi Arabic; obstruents; phonetics; consonant clusters

1. INTRODUCTION

The process of variable short vowel deletion word-initially is attested in several Arabic dialects, such as Algerian Arabic, Jordanian Arabic and Omani Arabic [2, 3]. E.g., the word *kitab* 'book' can be realised with a deleted vowel as /ktab/, or with a retained vowel as /kitab/. The same word can be produced with variable phonetic realisations. Similar processes have been also observed in other languages, such as the variable deletion of the French schwa [4] and the American schwa [5]. This process and its possible causes have received attention in the linguistic literature.

The conditioning factors of short vowel deletion are mostly attributed to speech rate and style [5, 6]. In Arabic dialects, the constraints that account for this phenomenon are attributed to stress and syllable structure, where vowel deletion occurs in unstressed open syllables and does not result in a sequence of more than two consonants [7, 8]. Sociolinguistic research on Arabic attributed vowel deletion more to male speakers than female speakers given that females tend to adhere more closely to Standard Arabic since it has higher status [9].

The literature on phonological contexts that trigger vowel deletion cross-linguistically lists several factors. For example, in Japanese, vowels are deleted/devoiced between voiceless consonants [10], which suggests an effect of the laryngeal feature of the surrounding consonants. In Kozani Greek, back vowels tend to be deleted more frequently than front vowels [11], indicating an effect of the quality of the target vowel. In Montreal French, vowels are deleted when preceded by fricatives [12], implying that the manner of articulation of the surrounding consonants affect the deletion of the vowel.

The phonetic realisation of vowel deletion and the factors that might influence the deletion process have not received much attention in Arabic dialects. We aim to address this issue in our empirical study, which has two objectives: 1) to document the phonetic variability by means of an acoustic study examined over 21,000 occurrences of words; and 2) to examine the effect speaker's gender and phonological contexts on the deletion process.

2. METHODOLOGY

This study employed different speech styles to examine vowel deletion thoroughly. Three production tasks were conducted: a diapix task to elicit spontaneous speech; a reading task to elicit controlled speech based on visual (written) stimuli; and a shadowing task to elicit controlled speech based on audio stimuli. The reading task is a prerequisite for the shadowing task, because this allows participants to access the orthographic forms of stimuli first to help them perceptually recognise the words.

2.1. Participants

45 native NA speakers (31 females, 14 males) were recruited. Their ages ranged from 18 to 50 years (mean= 27, range = 25–34, *SD* = 7.8). They had all given informed consent for their data to be used and did not report any speech or hearing difficulty. A total of 41 participants completed the diapix task and 39 participants completed the shadowing task.

2.2 Recording Procedure

Participants' speech was recorded using a Tascam DR05X stereo handheld digital audio recorder with built-in, omni-directional condenser microphone. Due to the Covid-19 pandamic, an external microphone was not used in order to ensure the safety of the participants. The recording was made at a sampling rate of 44.1 KHz and 16-bits resolution and were saved in waveform audio file (WAV) formats. Data collection rook place in quiet locations within participants' homes and participants were provided with instructions on how to optimise the recording environment. Prior to the recording, participants recived training on how to operate the recorder.

2.3 Diapix Task

A diapix task is an interactive task that elicits spontaneous speech in a controlled setting [13]. It employs two participants, each with a version of the same picture that is different in some ways, and they are asked to spot the differences by conversing with each other without looking at each other's picture.

2.3.1. Stimuli

Six picture pairs were adapted from Baker and Hazan [14]. There were 12 differences between the two versions of each pair. The pictures were printed out in colour in A5 size.

2.3.2. Procedure

Participants were given a folder that included six numbered envelopes (1–6), envelope (1) being for practice. They were instructed to sit at a table facing each other and place the recorder in the middle between them (approximately 30 cm from each participant). Each participant was instructed to take one version and converse with the other participant to find the 12 differences.

2.4. Reading Task

A sentence-reading task was utilised to examine participants' productions at a systematic and controlled level.

2.4.1. Stimuli and Design

We extracted 66 disyllabic and trisyllabic high-frequency NA nouns, in which the first and second consonants were obstruents, from Arabic Web Corpus usinf Sketch Engine. The target words comprised 17 words with a fricative–fricative sequence (FF), 16 words with a fricative–plosive sequence (FP), 19 words with a plosive–fricative sequence (PF) and 14 words with a plosive–plosive sequence (PP). A list of 40 filler words was also included to divert participants' attention.

This task involved two types of stimuli, namely visual and audio, following the protocol used in Hall [15]. The visual stimuli consisted of orthographic forms of the target words, while the audio stimuli were audio recordings of example sentences for each word, produced by a NA native speaker. The target word in each example sentence was excised and replaced with white noise for a duration of 400 ms.

A self-paced PowerPoint presentation was created, where each slide displayed one target word presented within two carrier sentences, accompanied by the matching audio recording of an example sentence which played automatically 300 ms after viewing the slide. Each word was repeated twice in the presentation, which totalled 344 slides.

2.4.2. Procedure

Participants were emailed a link to the task and were instructed to place the recorder on the right side of their computer. They were then instructed to read the sentences after listening to the audio stimulus and press the right arrow on their keyboard to advance to the next slide. Participants were allowed to take as many breaks as needed.

2.5. Shadowing Task

The aim of this task was to elicit the pronunciation of target words with a specific variant, namely, the vowel-absence variant /#CC/.

2.5.1. Stimuli and Design

The target and filler words used in this task were identical to those used in the reading task, except for one PP word. A NA speaker was audio-recorded producing the words with the onset vowel deleted /#CC/. A self-paced PowerPoint presentation was created, where each slide had an audio recording of one word played automatically twice with a 400 ms interstimulus. Each word was repeated twice in the presentation, resulting in a total of 210 slides.

2.5.2. Procedure

Participants were instructed to wear headphones (Beats Solo3 wireless on-ear), place the recorder to the right side of their computer, and sit about 30 cm from the computer. They were directed to listen to the audio stimuli and repeat the word after the second repetition as quickly and clearly as possible. Then, they were asked to press the right arrow on their keyboard to advance to the next slide.

2.6. Acoustic Analysis

Recordings were manually segmented and annotated in Praat [16]. One annotation tier was created to mark word and segment boundaries, following the segmentation guidelines created by Skarnitzl and Machač [17]. Another tier was created to mark the presence or absence of vowels in experimental items.

Vowel presence was identified by visual inspection of the spectrogram and waveform, and was coded as 'CVC' when high amplitude peaks of energy were visible in spectrograms as horizontal dark bands. For example, the onset of the vowel following a fricative was marked at the point where periodicity began. The offset of the vowel preceeding a fricative was marked at the point of onset of turbulent noise, as shown in Figure 1.

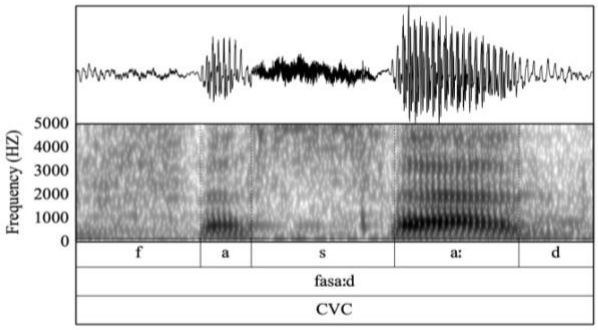

Figure 1: An example of word onset vowel presence for the FF sequence /fasa:d/.

Acoustic evidence for word onset vowel absence was marked when there was a complete absence of voicing and formant structure, and was coded as 'CC'. For example, in an FF sequence, both waveforms and spectrograms showed two distinct aperiodic frequencies for each fricative, with no evidence of formant structure or voicing, as illustrated in Figure 2.

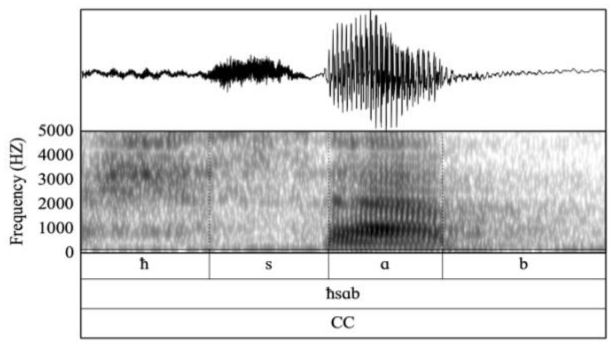

Figure 2: An example of word onset vowel absence for the FF sequence /ħsɑb/.

2.7. Statistical Analysis

The data were analysed by means of mixed-effects logistic regression model in R [18] with the the glmer() function using 'glmmTMB' package [19]. The dependent variable was was dichotomous (0 = vowel was deleted, 1= vowel was retained, 0 was the success level in the analysis). The fixed part of the model consisted of the following independent variables:

- Gender: male vs female speakers. Since past literature shows that gender differences affect deletion rates, this was the only sociolinguistic variable addressed in this study.
- The manner of articulation of the flanking obstruents with four levels: FF, FP, PF, PP.
- The place of articulation of the first obstruent, with four levels: labial–initial, coronal–initial, dorsal–initial and pharyngeal–initial.
- The laryngeal feature of the flanking obstruents (matching laryngeal: both voiced /zb/ or both voiceless /kt/ or different laryngeal features e.g. voiced–voiceless /bs/).
- The order of the place of the articulation of the sequence: front-to-back, e.g. /sk/, and back-to-front e.g. /kt/.
- The front or backness of the target vowel.
- The height of the target vowel.

The random effects considered in this study were participants and items. The significance of each varable was tested against the null hypothesis. If the p-value for a variable was > 0.05, we concluded that there was no effect on the outcome, and that the result was due to chance and therefore not statistically significant. Alternatively, if the p-value was ≤ 0.05, we could reject the null hypothesis and infer that the variable had an effect on the outcome.

For reasons of space, in the next paragraphs we report only the most interesting results in relation to our research aims.

3. RESULTS

The results of the three production tasks were combined, yielding a dataset of 21,205 tokens, with 58% of them exhibiting with vowel deletion. Vowel deletion was attested in all four obstruent sequences: FF, FP, PF, and PP. Word onset vowels were deleted across participants (M = 58%, SD = 4) and across items (M = 54%, SD = 30). Two factors were shown to be significant predictors of vowel deletion, namely gender and place of articulation of the initial obstruents (PoA). The result for each variable is discussed in detail below.

3.1 Gender

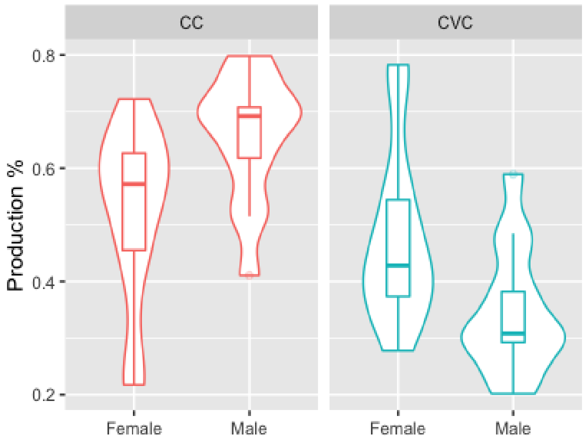

Figure 3: Vowel deletion (CC) and vowel retention (CVC) by gender.

Male speakers (*M* = 65, *SD* = 11) tended to delete more vowels than female speakers (*M* = 52, *SD* = 15). Overall, the rate of word onset vowel deletion among male speakers (66%) was higher than among female speakers (54%). The result of the mixed-effect logistic regression model showed a significant rise in vowel deletion among male speakers ($p < 0.01$), with a positive coefficient estimate of $b = 0.99$, indicating a significant rise in vowel deletion in males compared to females. The likelihood of vowel deletion by male speakers is 2.7 times higher than by female speakers.

3.2 PoA

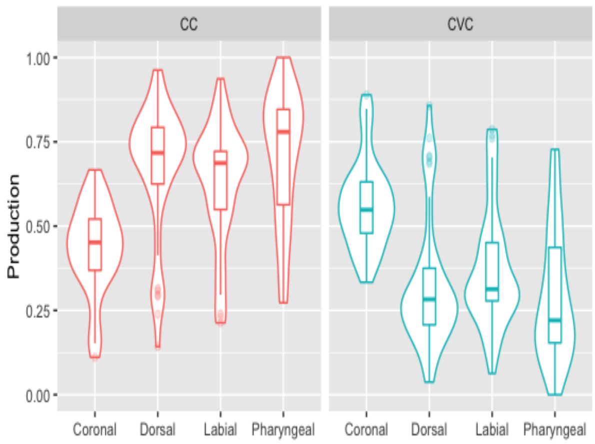

Figure 4: Vowel deletion (CC) and vowel retention (CVC) by PoA.

Participants produced pharyngeal-initial and dorsal initial sequences with the highest percentage of vowel deletion (71% and 69%, respectively), while coronal-initial sequences which were produced with the lowest percentage of vowel deletion (43%).

The findings of the mixed-effects logistic regression model showed a statistically significant effect of PoA for dorsal–initial sequences ($p = 0.03$), with a positive coefficient estimate of $b = 1.28$, indicating that dorsal–initial sequences were more likely to undergo vowel deletion compared to the the baseline reference (coronal-initial sequences).

The probability of vowel deletion in dorsal-initial sequences is 3.63 times higher than in coronal-initial sequences. For the labial-initial sequences, the likelihood of vowel deletion is 2.6 times higher than in coronal-initial sequences. Pharyngeal-initial sequences had the highest likelihood of vowel deletion, being 89% higher than coronal-initial sequences. Overall, coronal-initial sequences exhibited the least percentage of vowel deletion.

4. DISCUSSION AND CONCLUSIONS

In this paper, we have offered an acoustic analysis of vowel deletion in the dialect of NA. We show that vowel deletion is not solely a phenomenon of fast or casual speech, as it occurs at slow rates and with careful pronunciations, which is consistent with findings from previous research on NA [20, 21]. The data also shows that vowel deletion was observed in all four obstruent sequences: FF, FP, PF, and PP

In the present study, we also investigated the factors that contribute to vowel deletion. The findings indicated that two factors affect the rate of the vowel deletion. The first factor is the gender of the speaker, with male speakers tended to delete vowels more frequently than their female counterparts. Vowel-deleted variants are associated with the dialectal form, whereas vowel-present variants are linked with Standard Arabic. This suggests that female speakers tend to adhere to the prestigious variant of the language, as observed in other Arabic dialects [9].

The findings also revealed a statistically significant effect of the place of articulation of the first obstruents, particularly for dorsal-initial sequences. This finding is noteworthy for two reasons: firstly, it is the first time that PoA has been identified to play a role in vowel deletion in NA. Secondly, it contradicts with the findings from other languages, such as Japanese [22], where vowels are more likely to be deleted following coronals than labials or dorsals.

The data, while still preliminary, are enlightening as to the degree of vowel deletion in NA. This study highlights two additional factors that influence vowel deletion, beyond the conventional factors of stress and syllable structure that are widely acknowledged by Arabic linguists.

Acknowledgements: *This work was kindly supported by Princess Nourah Bint Abdulrahman University. We also thank the two anonymous reviewers.*

5. REFERENCES

[1] L. Davidson and K. Roon, "Durational correlates for differentiating consonant sequences in Russian," *Journal of the International Phonetic Association*, vol. 38, no. 2, pp. 137-165, 2008, doi: 10.1017/S0025100308003447.

[2] R. Al-Balushi, "Omani Arabic: More than a dialect," *Macrolinguistics*, vol. 4, no. 4, pp. 80-125, 2016, doi: 10.26478/ja2016.4.4.5.

[3] K. H. AbuAbbas, "Topics in the phonology of Jordanian Arabic: An optimality theory perspective," Dissertation/Thesis Doctoral dissertation, University of Kansas ProQuest Dissertations Publishing, 2003.

[4] A. Bürki, C. Fougeron, C. Gendrot, and U. H. Frauenfelder, "Phonetic reduction versus phonological deletion of French schwa: Some methodological issues," *Journal of Phonetics*, vol. 39, no. 3, pp. 279-288, 2011, doi: https://doi.org/10.1016/j.wocn.2010.07.003.

[5] L. Davidson, "Schwa elision in fast speech: Segmental deletion or gestural overlap?," (in English), *Phonetica*, vol. 63, no. 2-3, pp. 79-112, 2006, doi: 10.1159/000095304.

[6] F. Torreira and M. Ernestus, "Vowel elision in casual French: The case of vowel /e/ in the word c'était," (in English), *Journal of phonetics*, vol. 39, no. 1, pp. 50-58, 2011, doi: 10.1016/j.wocn.2010.11.003.

[7] R. Bani-Yasin and J. Owens, "The phonology of a Northern Jordanian Arabic dialect," *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, vol. 137, no. 2, pp. 297-331, 1987. [Online]. Available: http://www.jstor.org/stable/43379559.

[8] M. Alqahtani, "Syllable structure and related processes in Optimality Theory: An examination of Najdi Arabic," Doctoral dissertation, Newcastle University, Newcastle University, 2014. [Online]. Available: http://hdl.handle.net/10443/2757

[9] A. M. A. Al-Harahsheh, "Language and gender differences in Jordanian spoken Arabic: A sociolinguistics perspective," (in English), *Theory and Practice in Language Studies*, vol. 4, no. 5, p. 872, 2014, doi: 10.4304/tpls.4.5.872-882.

[10] N. Ogasawara, "Lexical representation of Japanese vowel devoicing," *Language and speech*, vol. 56, no. 1, pp. 5-22, 2013, doi: https://doi.org/10.1177/0023830911434118.

[11] N. Topintzi and M. Baltazani, "The acoustics of high-vowel loss in a Northern Greek dialect and typological implications," in *Consonant clusters and structural complexity*, vol. 26, P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, and B. Kühnert Eds. Berlin, Boston: DE GRUYTER, 2012, pp. 373–402.

[12] E. Farnetani and M. G. Busà, "Quantifying the range of vowel reduction in Italian," in *14th International Congress of Phonetic Sciences (ICPhS-14)*, San Francisco, CA, USA, 1999, pp. 491-494., doi: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_0491.html.

[13] K. J. Van Engen, M. Baese-Berk, R. E. Baker, A. Choi, M. Kim, and A. R. Bradlow, "The Wildcat corpus of native-and foreign-accented English: communicative efficiency across conversational dyads with varying language alignment profiles," *Language and Speech*, vol. 53, no. 4, pp. 510-540, 2010.

[14] R. Baker and V. Hazan, "DiapixUK: Task materials for the elicitation of multiple spontaneous speech dialogs," *Behavior Research Methods*, vol. 43, no. 3, pp. 761-770, 2011, doi: https://doi.org/10.3758/s13428-011-0075-y.

[15] N. Hall, "Acoustic differences between lexical and epenthetic vowels in Lebanese Arabic," *Journal of Phonetics*, vol. 41, no. 2, pp. 133-143, 2013, doi: https://doi.org/10.1016/j.wocn.2012.12.001.

[16] *Praat: doing phonetics by computer* (2018). [Online]. Available: http://www.praat.org/

[17] R. Skarnitzl and P. Machač, "Principles of phonetic segmentation," (in English), *Phonetica*, vol. 68, no. 3, pp. 198-199, 2011, doi: 10.1159/000331902.

[18] *R: A language and environment for statistical computing*. (2021). R Foundation for Statistical Computing, Vienna, Austria. [Online]. Available: https://www.R-project.org/

[19] M. E. Brooks *et al.*, "GlmmTMB balances speed and flexibility among packages for Zero-inflated: Generalized Linear Mixed Modeling," *The R Journal*, vol. 9, no. 2, pp. 378-400, 2017, doi: 10.32614/RJ-2017-066.

[20] P. F. Abboud, "The verb in northern Najdi Arabic," *Bulletin of the School of Oriental and African Studies*, vol. 42, no. 3, pp. 467-499, 1979.

[21] B. Ingham, *Najdi Arabic: Central Arabian*. Amsterdam; Philadelphia: John Benjamins Publication Company, 1994.

[22] J. A. Shaw and S. Kawahara, "More on the articulation of devoiced [u] in Tokyo Japanese: Effects of surrounding consonants," (in English), *Phonetica*, vol. 78, no. 5, pp. 467-513, 2021, doi: 10.1515/phon-2021-2011.