

Phonological conditioning of affricate variability in Emirati Arabic

Marta Szreder 

United Arab Emirates University
marta.szreder@uaeu.ac.ae

Donald Derrick 

University of Canterbury
donald.derrick@canterbury.ac.nz

This study investigates the conditioning effects of neighbouring consonants on the realisation of the phonemes /k/ and /dʒ/ in Emirati Arabic (EA), which are optionally realised as [tʃ] and [j], respectively. Based on previous accounts of EA and other Gulf Arabic (GA) dialects, we set out to test the prediction that proximity of other, phonetically similar coronal (COR) obstruents [COR, –son, –cont] and coronal postalveolar fricatives [COR, –ant] inhibit the surface realisation of the affricate variants of these phonemes. We examine elicitation data from twenty young female native speakers of EA, using stimuli with the target segment in the presence of a similar neighbour, as compared to words with the neighbour at a longer distance or with another coronal consonant. The results point to an asymmetry in the behaviour of the voiced and voiceless targets, such that the predicted inhibitory effect is confirmed for the voiced, but not the voiceless target. We argue that this finding, coupled with a consideration of the intra-participant and lexical trends in the data, is compatible with an approach that treats the two processes as being at different stages of development, where the [k~tʃ] alternation is a completed phonemic change, while the [dʒ~j] alternation is a synchronic phonological process.

1 Introduction

Differences in affricate use and distribution are one of the defining features of cross-dialectal variation in Arabic. While the behaviour of affricates and their allophones is conditioned by a variety of historical, geographical, social, lexical, and phonological factors, detailed analyses of the specifics of that conditioning in different dialects are still needed. Recently, Szreder, Derrick & Ben-Ammar (2021) offered an exploratory analysis of some of the factors in intra- and inter-speaker variability in affricate realisation in Emirati Arabic (EA). The current study presents the results of a follow-up experiment, designed to illuminate one of the phonological factors identified as relevant in the original study, which is the potential role of neighbouring coronal consonants in facilitating the realisation of the affricates. Below, we provide a general overview of the historical origins of affricate variation in Arabic (Section 1.1), and previously reported phonological effects in their distribution in Gulf Arabic (GA) dialects in general

(Section 1.2), and in EA in particular (Section 1.3). Based on this context, we present the objectives of the study in Section 1.4.

1.1 The origins of affricate variation in modern Arabic dialects

In the dialects in which they are present, alveolar and postalveolar affricates are derived from Old Arabic (OA) [k] (voiceless) and [g] (voiced) (Holes 1995). As such, they form a subset of the larger pattern of dorsal stop variation, which also includes the uvular plosive [q]. The three OA dorsal stops [k], [g], and [q], have historically been prone to fronting and/or affrication – a variation reported already by the ancient grammarians Sibawayhi and Ibn Jinni, and later confirmed experimentally (Al-Ani 1970). Over time, this has led to modern dialects exhibiting different combinations of the stops and their variants. Table 1 illustrates their distribution in selected Arabic dialects.

Table 1 Examples of Old Arabic dorsal stop distribution in selected dialects of Arabic.

Old Arabic (Holes 1991, Watson 2002)	k	g~ʒ	q
Cairene Arabic (Monassar 2014)	k	g	ʔ
Gulf Arabic (Feghali 2008, As-Sammer 2010)	k~ts~ʃ~tʃ	g~dz~ʒ~dʒ~j	q~g
Ha'ili Arabic (Saudi Arabia) (Prochazka 1988, Al Rasheedi 2015)	k~ts	g~dz	g
Lebanese Arabic (Watson 2002)	k	ʒ	ʔ
Mesopotamian Qeltu Arabic (Holes 1995, Youssef 2021)	k	dʒ	q
Modern Standard Arabic (Watson 2002, Monassar 2014)	k	dʒ~ʒ	q
Moroccan Arabic (Watson 2002)	k	ʒ	q
Ruralite Levantine (Fischer, Jastrow & Behnstedt 1980, Youssef 2021)	tʃ	dʒ	k
Sudanese Arabic (Dickins 2007)	k	g ^j	g

As can be seen in Table 1, the fronting and assibilation of the stops exhibits significant within-dialect variation, but all dialects maintain a three-way contrast, albeit using different phonemes to realise it. The dialects which use the uvular or the glottal stop for OA [q] may contrast it with either [g] (e.g. Cairene Arabic), [ʒ] (e.g. Lebanese Arabic), or [dʒ] (e.g. Qeltu Arabic) in the place of OA [g]. Furthermore, there is often some degree of variability in the use of the OA [g] variants. For example, in Ha'ili Arabic [g] varies with [dz], while in some GA varieties the voiced affricate further undergoes lenition into the glide [j].

As regards OA [k], it has mostly remained stable across dialects, with the notable exception of the Peninsula, where it is often replaced with a voiceless affricate [ts] or [tʃ], or fricative [ʃ] (a process traditionally termed *kashkasha*), particularly in the varieties of GA (Holes 1995). GA is a group of dialects of Peninsular Arabic spoken in the United Arab Emirates,

Qatar, Kuwait, and parts of Bahrain, Oman, Saudi Arabia, and Iraq (Versteegh 1997, Holes 2001). Indeed, the variable affrication of the OA [k] is one of the characteristic features of all varieties of these dialects.

Gulf Arabic is therefore a group of dialects exhibiting very high variability of realisations of both historical /k/ and /g/. The words containing these sounds have variable pronunciations both within and across different varieties of the dialect, as illustrated by the examples in (1):

- (1) [kalb ~ ʃalb ~ tsalb ~ tʃalb] 'dog'
 [gebel ~ ʒebel ~ dzebel ~ dʒebel ~ jebel] 'mountain'

The variability of the GA affricates is due to two processes: affrication of the two OA stops and lenition of the derived voiced affricate. Lexical and sociolinguistic factors have been observed to affect these two processes. As far as lexical identity is concerned, neither of them are supposed to occur in borrowings (Qafisheh 1977), as illustrated by the examples of loanwords below, where k-affrication and dʒ-lenition are blocked in in (2a) and (2b), respectively (examples after Qafisheh 1977).

- (2) (a) [sikri:m] ~ *[sitʃri:m] 'ice-cream' (English)
 [stika:n] ~ *[stitʃa:n] 'teacup' (Persian)
 [sakar] ~ *[satʃar] 'sugar' (Hindu)
- (b) [gara:dʒ] ~ *[gara:j] 'garage' (English)
 [dʒaza:ʔir] ~ *[jaza:ʔir] 'Algeria' (proper noun)
 [dʒawa:z safar] ~ *[jawa:z safar] 'passport' (literary Arabic)

Outside of this restriction, the variation between affricate and non-affricate variants has often been linked to sociolinguistic factors. Studies investigating the correlation between affrication and extra-linguistic factors, including age, educational level, gender of the speaker and area, report that k-affrication appears more frequently in the speech of older, less educated speakers, and within this group, males are more likely to affricate than females (Al Rojaie 2013, Dashti, Akbar & Taqi 2015, El Salman 2016, Dashti 2018). In contrast, while dʒ-lenition also used to be considered an 'incorrect' pronunciation and as such associated with lower socioeconomic status and rural populations (As-Sammer 2010, Hassan 2017, Leung, Ntelitheos & Al Kaabi 2020), more recent reports suggest that the [j] variant has been adopted by young urban speakers in Bahrain, Kuwait, and the rest of the Gulf (Al Qouz 2009, Hassan 2009, Holes 2011). Nonetheless, both processes exhibit variability in application within speakers, and the probability of their occurrence has also been linked to phonological factors.

1.2 Phonological factors

The above sociolinguistic factors have traditionally been considered the sole determinants in the process of dʒ-lenition, rendering it a purely sociolinguistic phenomenon (Hoffiz 1995, Feghali 2008, As-Sammer 2010, Hassan 2017). In contrast, a phonological factor which has often been associated with the likelihood of a surface [tʃ] is the vowel context, as the process is traditionally assumed to have originated from palatalization in front vowel contexts – a

phenomenon widely attested cross-linguistically (Bhat 1978; Bateman 2007, 2011; Kochetov, 2011). Grammars of GA report that the effect of non-back vowels on the process is still active (Feghali 2008, Leung et al. 2020), and therefore [tʃ] is more likely to surface in words with a non-back vowel than with back vowel.

However, this is complicated by consonantal factors which may affect vowel quality. For example, Cantineau, (1936), Johnstone (1963), and Holes (1995) report that in GA dialects, /k/ undergoes affrication when it is in contiguity with front vowels /i i: e:/ and, conversely, that affrication is often blocked by the presence of back vowels /u: o:/. Nonetheless, they also report variability in the application of the process, attributable mainly to the problematic case of the vowels /a a:/, which are generally front or central, but always surface as back [ɑ a:] in the presence of emphatic consonants, and sometimes also in the presence of labial consonants /b f m w/, and the liquids /r l/. A similar process has been observed in Proto-Tamil, where palatalisation of velars normally conditioned by the following front vowel is blocked if the vowel is followed by a retroflex or a liquid, as those segments lead to the retraction of the vowel (Krishnamurti 2003).

The variability in the backness of /a a:/ therefore accounts for some of the variability in the process of affrication. Nonetheless, it is virtually impossible to find clear instances of non-application of affrication due to any consistent phonological influence, and Johnstone (1963) and Holes (1995) are careful to note that the identified regularities do not categorically determine the presence or absence of the affricate. Overall, k-affrication and d₃-lenition thus appear to be unrelated processes which escape rigid phonological descriptions, and their variants are sometimes simply characterised as occurring in free variation (Leung et al. 2020).

However, Mustafawi (2006, 2011) proposes a unified phonological account of both processes in Qatari Arabic (QA) and other Gulf varieties, rooted in Optimality Theory (Prince & Smolensky 2004). Specifically, one factor that blocks either of the phonemes from surfacing as an affricate in this account is the Obligatory Contour Principle (OCP). OCP can be described as a morpheme-structure constraint (Leben 1973), an autosegmental phonology concept (Goldsmith 1976), or Optimality Theory constraint (Frisch, Pierrehumbert & Broe 2004), which restricts two phonetically similar segments from occurring in close proximity.

Mustafawi (2011) relates the OCP to the previously observed constraint on consonantal roots in Arabic, where highly phonetically similar consonants (Cs) are blocked from occurring in a sequence (i.e. as C1 and C2, or C2 and C3, of a root), and only occasionally allowed in the same root at all (i.e. as C1 and C3) (Greenberg 1950). As regards the affricates, the relevant OCP dimensions are *([cont] . . . [cont])_{COR}, and *([ant] . . . [ant])_{COR}, that is phonetic similarity between two coronal obstruents in continuity or place. In Mustafawi's account, two constraints restrict coronal (COR) obstruents which agree in these features from co-occurring, including across an intervening vowel. The first constraint, related to continuity, can be stated as *[COR, –son, –cont . . . COR, –son, –cont]_{WORD}. As a result, the [COR, –son, –cont] affricates [tʃ d₃] are not permitted to surface in the presence of the [COR, –son, –cont] coronal stops [t d], as the examples in (3) illustrate:

- (3) [katəf] ~ *[tʃatəf] 'shoulder'
 [di:k] ~ *[di:tʃ] 'rooster'
 [jad] ~ *[d₃ad] 'grandpa'

The second constraint, related to Place, can be stated as *[COR, –son, –ant . . . COR, –son, –ant]_{WORD}. As a result, [COR, –ant] affricates [tʃ d₃] are not permitted to surface in the presence of the [COR, –ant] fricative [ʃ], even though they contrast in [cont], as the examples in (4) illustrate:

- (4) [ʃijar] ~ *[ʃidʒar] ‘tree’
 [ʃo:ka] ~ *[ʃo:tʃa] ‘fork’

The constraints are said not to apply to emphatic consonants, due to the fact that they contrast with [tʃ, dʒ] in [RTR] (retracted tongue root). While some observed variability is attributed to the fact that the OCP is not categorical and is stricter for immediately neighbouring consonants than for those separated by a vowel, this account offers specific predictions that are worth investigating in other dialects which exhibit the variation. Indeed, the OCP also appears to be compatible with an account of affrication of /k/ to [tʃ] in a dialect outside of the GA group – Ha’ili Arabic (spoken in Saudi Arabia and part of Najdi Arabic), offered by Al Rasheedi (2015). Specifically, affrication is prohibited in the dialect if there are other coronal segments in the word, or if the velar stop is geminated (although the gemination constraint is not applicable to GA, see Mustafawi 2006). Hence, for example, the words in (5) cannot undergo affrication:

- (5) [ka:s] ~ *[tʃa:s] ‘glass’
 [rukka:b] ~ *[rutʃta:b] ‘passengers’

While the gemination constraint is not relevant to GA (Mustafawi 2011), the general convergence of the findings from QA and Ha’ili Arabic regarding the OCP suggests that the effect may be potentially relevant to affrication in Emirati Arabic. Below, we outline the phonology of Emirati Arabic and recent findings relevant to this question.

1.3 Emirati Arabic

EA is a variety of GA spoken in the United Arab Emirates which uses both voiceless and voiced postalveolar affricates. Tables 2 and 3 present the phonemic inventory of EA.

Both the voiced and the voiceless postalveolar affricates are used in EA. Although grammars attest that the voiced plosive variant is still a possible realisation of OA [g] in the dialect (Leung et al. 2020), it is virtually absent from the speech of modern Emirati natives. For example, a cursory search of the EA corpus EMAC (Halefom et al. 2012) reveals that out of all 1,935 tokens with a target /g/, only 39 (0.2%) surfaced with the plosive. It is likely that [g] has come out of use in this context since it is also used for OA [q] in the dialect, except in proper names (which is why /q/ remains part of the phonemic inventory). Therefore, the change of /g/ to [dʒ] can be considered a completed historical change, and the affricate [dʒ] is treated in this study as the underlying variant of the OA /g/ in EA. However, it does exhibit variability, since it undergoes a process of lenition into [j], as in some other GA dialects discussed in Section 1.2.

On the other hand, k-affrication to [tʃ] is an active process in EA. The voiceless affricate [tʃ] occurs as part of two distinct categories of phonological processes: suffix- and stem-based affrication. Suffix-based affrication of /k/ into [tʃ] is a regular morphological feature of GA in the second person feminine singular suffix (Qafisheh 1975, Hoffiz 1995, Feghali 2008). For example:

- (6) MSA [ixwa:nik] – GA [ixwa:nitʃ] ‘your brothers’ (SG.FEM)
 MSA [ʃafik] – GA [ʃafitʃ] ‘he saw you’ (SG.FEM)

Table 2 Emirati Arabic consonants (following Table 2.1 in Leung et al. 2020: 11). Symbols on the right represent emphatic (pharyngealised) consonants.

		Labial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	Voiceless				t t ^ʕ			k	q		ʔ
	Voiced	b			d			g			
Nasal		m			n						
Fricative	Voiceless		f	θ	s s ^ʕ	ʃ		x		ħ	h
	Voiced			ð ð ^ʕ	z			ɣ		ʕ	
Affricate	Voiceless					tʃ					
	Voiced					dʒ					
Approximant	Trill				r						
	Lateral				l l ^ʕ						
	Glide	w					j	w			

Table 3 Emirati Arabic vowels (following Figure 2.1 in Leung et al. 2020: 16).

	Front	Central	Back
Close	i i:		u u:
Close-mid	e e:		o o:
Central		ə	
Open-mid			
Open	a a:		

As such, suffix-based affrication is an obligatory process with a syntactic function, which applies across the board in GA varieties, including EA. In contrast, stem-based affrication applies to many targets with OA /k/, as discussed in the preceding sections.

As in other dialects of GA, the likelihood of k-affrication is also supposed to increase if the segment is preceded or followed by a non-back vowel. However, this generalisation appears to have little predictive power in modern day EA, and at least one study did not find any stable influence of non-back vowels on the probability of affrication in the dialect (Hassan 2017). Recently, Szreder et al. (2021) examined the variability of affrication and lenition in EA based on corpus data and an elicitation experiment. Contrary to Hassan (2017), the corpus analysis did suggest a modest effect of vowel quality, such that the presence of the back vowels /o: u:/ decreased the probability of both k-affrication and d₃-lenition, although the presence of the front vowels /i: e:/ did not have a strong predictive power (i.e. both affricate and non-affricate variants of both phonemes occurred in the presence of front vowels). Furthermore, the experimental part of the study revealed intra-speaker and intra-word variability (despite laboratory conditions), which precluded the formulation of any strict rules.

Interestingly, Szreder and colleagues also reported an unexpected effect of certain coronal segments in the corpus data. Specifically, across different lexical types and tokens, /d₃/ was statistically more likely in contiguity with /s ʃ t ð/, while /j/ – in contiguity with /t d n s/. However, the experimental part of the EA study did not vary this phonological factor systematically, since it focused on intra-speaker variability and lexical effects. While the small number of relevant tokens available and the fact that the corpus contained data mostly from the 1970s did not allow the authors to state any firm conclusions, it appears to warrant further investigation for two reasons. First, while vowel effects have commonly taken priority in previous studies, the sheer variability of vowels in spoken Arabic means that their effect is at the very least bidirectional – i.e. that it is equally likely that the vowel quality is predicted by the consonant as vice versa (as suggested by Johnstone 1965). This could be one of the reasons why previous investigations did not arrive at firm conclusions regarding their role. Meanwhile, consonants are realised much more consistently, and therefore can be more theoretically interesting as a potential explanation of the observed variability. Second, the possible effect of certain coronal consonants appears to bear some resemblance to the findings in QA by Mustafawi (2006, 2011), who proposed that both affricates [tʃ] and [d₃] were less likely to occur next to phonetically similar coronal segments.

This similarity, as well as its plausible phonological motivation (i.e. OCP) make it a direction worth further study, especially since there is partial overlap between the conditioning consonants in EA (/s ʃ t d n ð/) and those described in Mustafawi's account (/ʃ t d/). Moreover, in the old corpus data, the direction of the effect appears to be opposite for the voiceless segment (i.e. increased odds of affricate variant in the presence of these segments). Since the EA corpus data (see Szreder et al. 2021) did not have enough relevant tokens (and the ones available were from a different point in time), it remains to be shown whether these differences were due to sampling, or whether the process in EA is sensitive to different phonological properties (or, perhaps, not active at all.)

1.4 Objectives

The current study was designed to provide a more in-depth analysis of the effects of coronal consonants on surface affricates in EA. Our objective was to investigate whether a conditioning effect of coronal consonants on affricates can be demonstrated in experimental data from modern speakers, and if so, whether its pattern is compatible with a constraint on the co-occurrence of similar consonants previously found for QA. Given this, we propose two hypotheses outlined below.

Hypothesis 1 holds that there are OCP constraints such that adjacent word-internal [t], [d], and [ʃ] would reduce the likelihood of affrication of /k/ as [tʃ] and increase the likelihood of lenition of [dʒ] to [j]. It is important to note that these groups of stimuli were not meant to elicit categorical differences, and variability is expected. All processes are expected to be optional and sensitive to other lexical, sociolinguistic, and vowel-related factors, in addition to the (postulated) OCP. Nonetheless, if the OCP holds for EA, our first hypothesis makes the following four predictions:

- (i) [k] will be significantly preferred over [tʃ], and
- (ii) [j] will be significantly preferred over [dʒ], in the presence of the phonetically similar segments /t d ʃ/, but not other coronal segments;
- (iii) [k] will be significantly preferred over [tʃ], and
- (iv) [j] will be significantly preferred over [dʒ] if the segments /t d ʃ/ are neighbouring with the target or separated only by a vowel, but they will be less affected in cases where these consonants are at a greater distance – i.e. separated by another consonant or an additional syllable.

Hypothesis 2 holds that affrication of /k/ and lenition of /dʒ/ are associated with lower socioeconomic status, in agreement with previous studies. Therefore, our fifth prediction is as follows:

- (v) there is a relationship within individual speakers such that speakers who affricate /k/ will also be more likely to lenite /dʒ/.

2 Method

2.1 Speakers

Twenty native speakers of EA were recruited, all female, aged 18–23 years (mean = 20 years). All participants were students at the United Arab Emirates University and as such were all bilingual in Arabic and English. Some participants also reported some knowledge of Turkish, Spanish, and Korean as a third language. All participants reported having no history of speech or hearing impairments and were compensated for their time.

2.2 Materials

The experiment was a picture naming task. The stimuli consisted of 20 Arabic words, 10 containing /k/ and 10 containing /dʒ/. Approximately half of the words (depending on availability and imageability) were selected to include the target segments in the context of a phonetically similar consonant /ʃ t d/ separated only by a vowel. The remaining words included the target segments with other coronal consonants, e.g. /s ð r/, or a phonetically similar consonant at a farther distance. The words were also chosen to include the target segments in word initial, word medial and word final positions, wherever possible. Table 4 presents the list of stimuli by experimental group and by relevant segment.

Table 4 Stimuli by experimental category. Forms in square brackets represent hypothetical variants with the one expected to be more common underlined.

Underlying Representation	Variants	Gloss	Condition
Expected [k]			
/katəf/	[katəf ~ tʃatəf]	'shoulder'	/k/ across V from /t/
/di:k/	[di:k ~ di:tʃ]	'rooster'	/k/ across V from /d/
/ʃo:k/	[ʃo:k ~ ʃo:tʃ]	'thorn'	/k/ across V from /ʃ/
/ʃo:ka/	[ʃo:ka ~ ʃo:tʃa]	'fork'	/k/ across V from /ʃ/
Expected [tʃ]			
/kaðb/	[kaðb ~ tʃaðb]	'lie'	/k/ across V from another coronal
/ki:s/	[ki:s ~ tʃi:s]	'bag'	/k/ across V from another coronal
/məswa:k/	[məswa:k ~ məswa:tʃ]	'toothbrush*'	/k/ across V from another coronal
/səkkin/	[səkkin ~ sətʃin]	'knife'	/k/ across V from another coronal
/kabʃ/	[kabʃ ~ tʃabʃ]	'ram'	/k/ across VC from /ʃ/
/kəbrit/	[kəbrit ~ tʃəbrit]	'matches'	/k/ across VCCV from /t/
Expected [j]			
/dʒad/	[dʒad ~ jad]	'grandpa'	/dʒ/ across V from /d/
/dʒidar/	[dʒidar ~ jidar]	'wall'	/dʒ/ across V from /d/
/dʒədi:d/	[dʒədi:d ~ jədi:d]	'new'	/dʒ/ across V from /d/
/məsdʒəd/	[məsdʒəd ~ məsjəd]	'mosque'	/dʒ/ across V from /d/
/ʃadʒara/	[ʃadʒara ~ ʃajara]	'tree'	/dʒ/ across V from /ʃ/
Expected [dʒ]			
/dʒaras/	[dʒaras ~ jaras]	'bell'	/dʒ/ across V from another coronal
/madʒləs/	[madʒləs ~ majləs]	'gathering'	/dʒ/ across V from another coronal
/dʒəld/	[dʒəld ~ jəld]	'skin'	/dʒ/ across VC from /d/
/daradʒ/	[daradʒ ~ daraj]	'stairs'	/dʒ/ across VCV from /d/
/dʒari:da/	[dʒari:da ~ jarida]	'newspaper'	/dʒ/ across VCV from /d/

*A specific type of traditional natural toothbrush

2.3 Procedure

Participants were seated on a chair located in the Phonetics Lab at the United Arab Emirates University and were first given instructions regarding the experiment. The participants then had an initial training run in which they were asked to provide the target words based on pictures, so that the experimenter would not model the pronunciation. They were specifically instructed to use Emirati Arabic rather than MSA. If they struggled with elicitation, they were provided with the English equivalent. If that did not result in elicitation, the subsequent trials for the unrecognized word from that participant were excluded from analysis. Following the training, the participants were presented with the full list of stimuli ten times in a pseudorandomized manner on Articulate Assistant Advanced software (AAA) (Articulate Instruments Ltd. 2012), to allow for estimates of intra-speaker variability.

Acoustic data were collected using a Neumann U87 Ai microphone positioned near the participant's mouth with a USBPre2 microphone preamp and A/D converter. The ultrasound and acoustic data were recorded and automatically synchronized on an ASUS ROG G701 VIK laptop running AAA.

2.4 Analysis

The stimulus list was expected to produce 200 tokens per participant and 4,000 tokens in total (20-word items × 10 trials × 20 participants). Only three tokens had to be excluded due to elicitation errors, and the final number of responses included in the analysis was 3,997. All

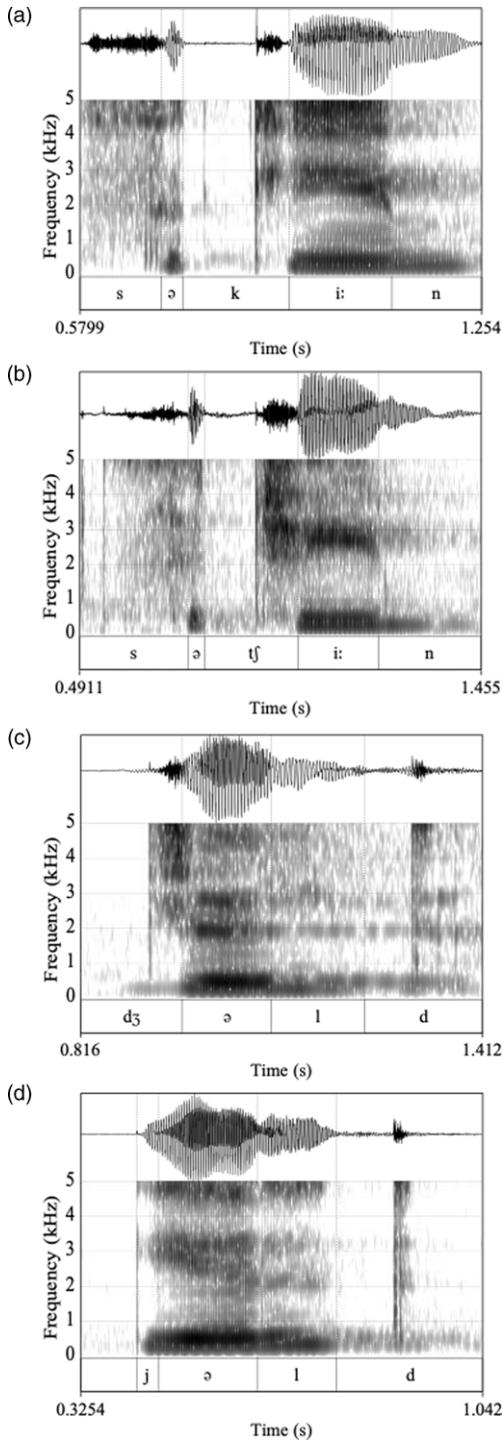


Figure 1 Sample annotated spectrograms of responses for the four segments. (a) Velar stop variant in the word 'knife' [səki:n] from subject O7. (b) Voiceless affricate variant in the word 'knife' [sətʃi:n] from subject O2. (c) Voiced affricate variant in the word 'skin' [dʒəld] from subject 19. (d) Glide variant in the word 'skin' [jəld] from subject O3.

responses were phonetically transcribed in AAA, by a research assistant who was a native speaker of Arabic (bidialectal in Emirati and another dialect), coded for the presence of the target segment, and double-checked by the first author.

The transcription was based on impressionistic native speaker judgment and visual inspection of the spectrograms in AAA. Figure 1 presents sample spectrograms for each of the four segments. Since the target segments are quite perceptually distinct, there were no tokens judged as ambiguous by the transcriber, and the first author, having listened to each token, agreed with the transcriber's judgements.

The data were then analysed using generalized linear mixed-effects modelling using the *lme4* package (Bates et al. 2015) in R (R Core Team 2022). As each token was described as either affricated (assigned a value of 1) or unaffricated (assigned a value of 0), analysis used a generalized linear model 'binomial' distribution. We began with a full random effects model as described in the R statistical formula in (7):

$$(7) \text{ Affricate} \sim \text{OCP} * \text{Phoneme} + (1 + \text{Condition} * \text{Phoneme} | \text{Subject_ID})$$

In this formula, *Affricate* is 0 for an unaffricated segment ([k], [j]), and 1 for an affricated segment [tʃ] and [dʒ]). The variable *OCP* is one of three codes, 'adjacent' for words where (/t d ʃ/) are separated from the relevant phoneme (/k dʒ/) by only a vowel, 'nearby' for where (/t d ʃ/) are more distant, and 'no' for where there are no other examples of (/t d ʃ/) within the word. The variable *Phoneme* is one of the two phonemes, /k/ or /dʒ/. This model converged but had singularity errors as identified using the *ISSINGULAR* function, so the model used for our analysis is the one shown in (8):

$$(8) \text{ Affricate} \sim \text{OCP} * \text{Phoneme} + (1 + \text{OCP} + \text{Phoneme} | \text{Subject_ID})$$

The only difference between (7) and (8) is that model (8) does not have a random-effect interaction between *OCP* and *Phoneme*.

3 Results

We begin by focusing on Predictions (i)–(iv) of Hypothesis 1 – that there are two OCP Constraints, *[COR, –son, –cont . . . COR, –son, –cont]_{word} and *[COR, –ant . . . COR, –ant]_{word}, such that word-internal [t], [d], and [ʃ] would reduce the likelihood of affrication of [k] to [tʃ] and increase the likelihood of lenition of [dʒ] to [j].

Overall, the data from all tokens show many examples of both affricated and non-affricated /k/ and /dʒ/. However, speakers produced /k/ in affricated [tʃ] form more often (53% of tokens) than they produced /dʒ/ in its affricated [dʒ] form (29% of tokens), as seen in Figure 2.

Examination of the descriptive statistics for the /k/ phoneme appears to show a high degree of variability in production, regardless of whether the word with /k/ was produced with a /t d ʃ/ within one vowel of the /k/ (Adjacent), elsewhere in the word (Nearby), or not in the word at all (No), as seen in Figure 3.

Examination of descriptive statistics for the /dʒ/ phoneme appears to show low levels of affrication in any of the five words with adjacent /t d ʃ/. In contrast one of the three words that has a /t d ʃ/ further away from the /dʒ/ phoneme than one vowel, and one of the two words without any /t d ʃ/ were most commonly produced in their affricated form, as seen in Figure 4.

Generalized linear mixed-effects modelling shows a main effect that affrication is significantly higher for tokens where /t d ʃ/ were not adjacent to /k/ or /dʒ/. The results show an interaction such that for the phoneme /k/, this main effect is negated, as seen in Table 5.

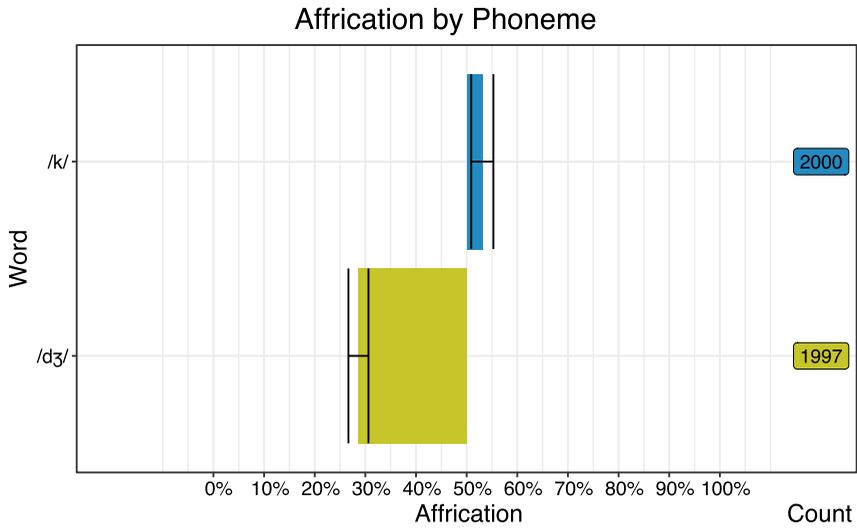


Figure 2 (Colour online) Graph showing the percent affricated ([tʃ], [dʒ]; blue) and unaffricated ([k], [j]; gold) of tokens from the /k/ and /dʒ/ phonemes. The number of recorded tokens is seen on the right.

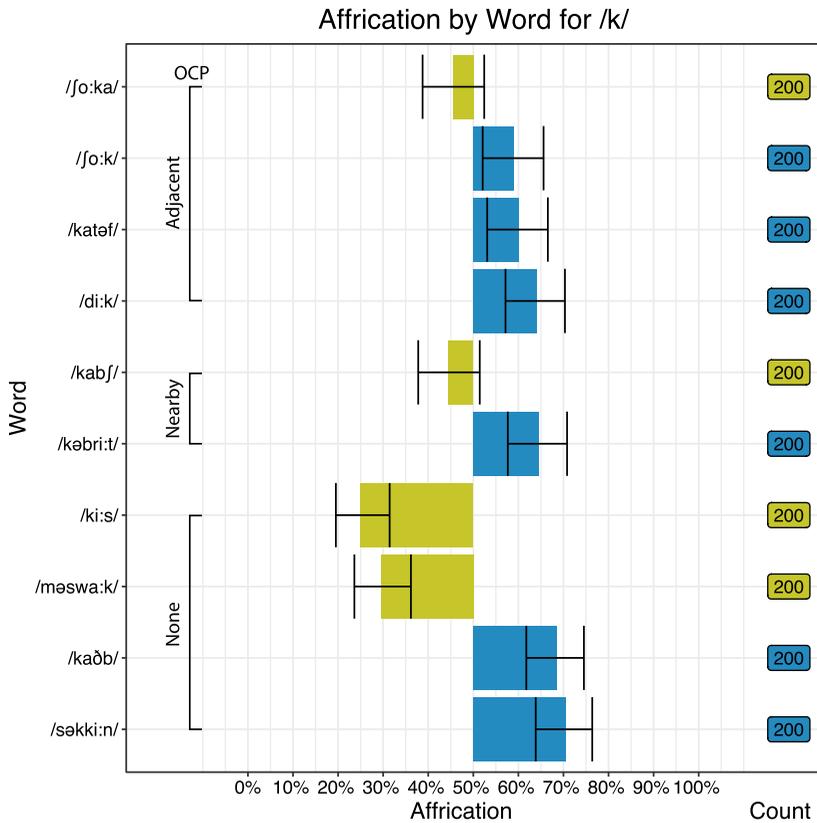


Figure 3 (Colour online) Graph showing the percent affricated ([tʃ]; blue) and unaffricated ([k]; gold) of the /k/ phoneme. The graph also shows which words have /t d ʃ/ across a vowel (adjacent), within the word (nearby), or not present in the word. The number of recorded tokens is seen on the right.

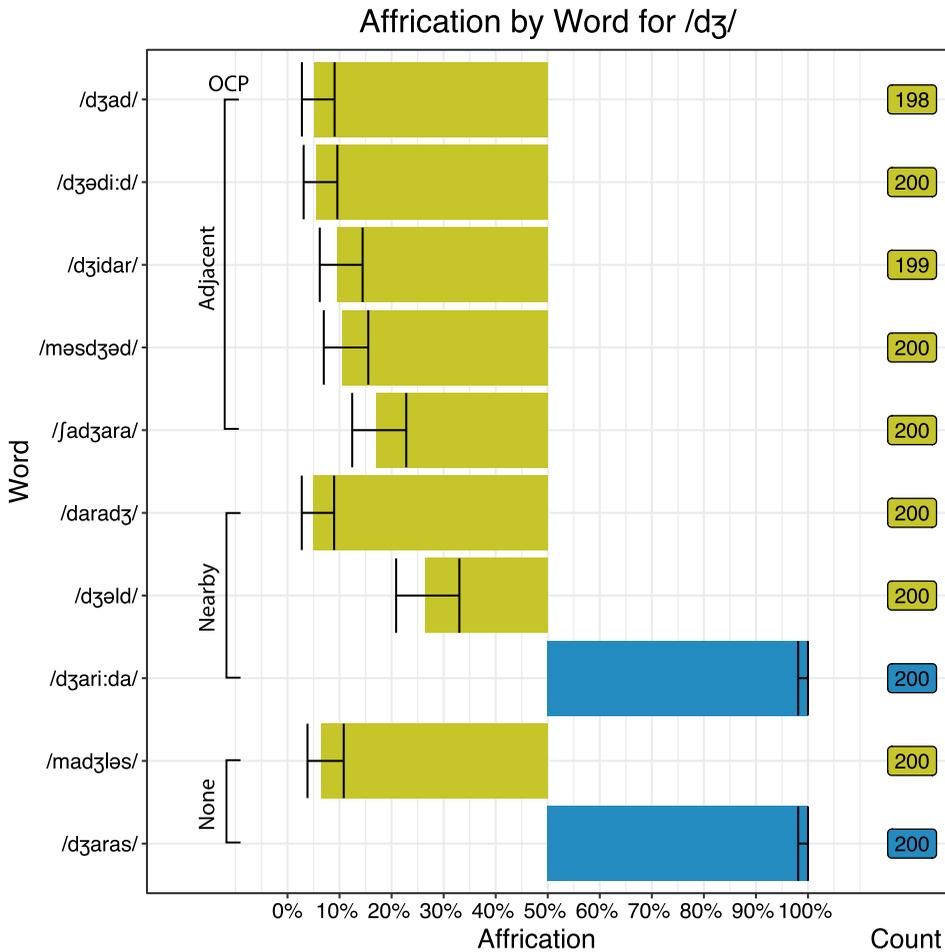


Figure 4 (Colour online) Graph showing the percent affricated ([dʒ]; blue) and lenited (unaffricated) ([j]; gold) of the /dʒ/ phoneme. The graph also shows which words have /t dʒ/ across a vowel (adjacent), within the word (nearby), or not present in the word. The number of recorded tokens is seen on the right.

These results were graphed using an interaction plot, showing that the phoneme /dʒ/ was produced with lower rates of affrication for adjacent /t dʒ/, but there was little to no differences between /t dʒ/ further than a vowel's distance away in the same word, or completely absent. In contrast, /k/ was produced with the most affrication in words with adjacent /t dʒ/, and less in words with /t dʒ/ further than a vowel's distance away in the same word, or completely absent, as seen in Figure 5.

Next, we tested Hypothesis 2, following Prediction (v), since /k/ affrication and /dʒ/ lenition are both considered to be features of the local dialect, and have both been proposed to relate to sociolinguistic factors, we expected that participants who used more affrication would also use more lenition. The participants were not fully consistent in their preference for each variant, although some trends could be observed. Specifically, no participant always produced the more local variant for each type of stimulus (i.e. [j] and [tʃ]), five out of the 20 participants never affricated the /k/, and one more produced /k/ in 99% of tokens. However, there was a slight general tendency for the local variants to pattern together. That is, the participants who used more [j] were also slightly more likely to use more [tʃ]. Figure 6 illustrates the relationship between the two variants. Linear modelling showed the relationship between

Table 5 Fixed-effects results comparing rates of affrication based on comparing /t d ʃ/ adjacency (Adjacent, Nearby) to None, and based on underlying phoneme (/k dʒ/).

Variable	Estimate	Std. Error	z-value	p-value ($\alpha = 0.05$)
OCP = Nearby	-0.428	0.260	-1.646	.099
OCP = Adjacent	-4.661	0.680	-6.852	< .001
Phoneme = /k/	-2.279	1.364	-1.671	.095
Nearby: /k/	1.080	0.247	4.371	< .001
Adjacent: /k/	6.655	0.549	12.113	< .001

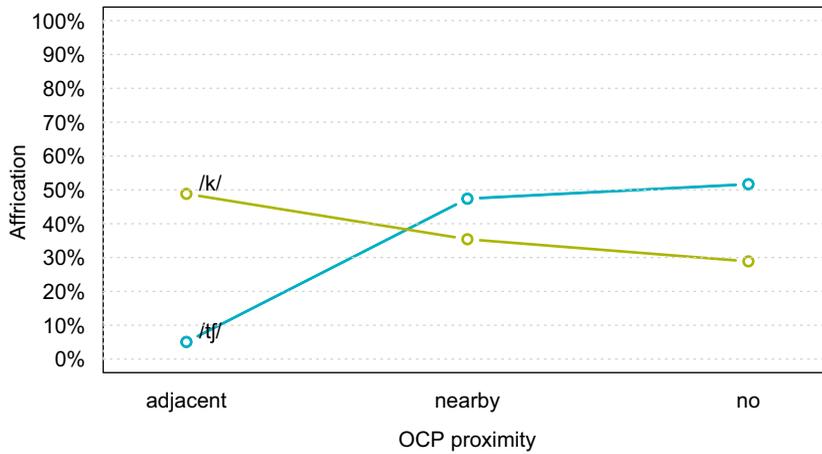


Figure 5 (Colour online) Interaction graph showing relative affrication in /k/ and /dʒ/ phonemes based on whether there was a /t d ʃ/ within one vowel, somewhere further away in the word, or completely absent from the word.

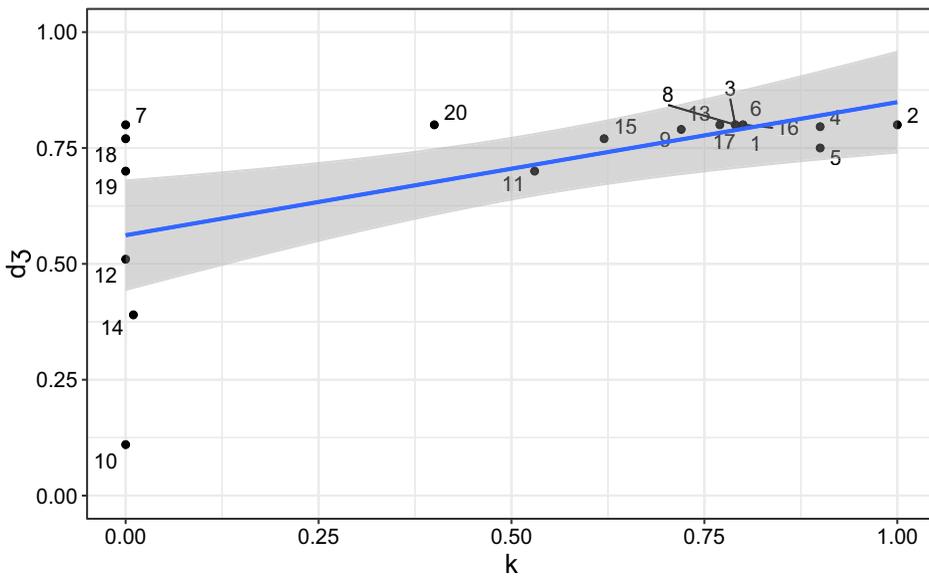


Figure 6 (Colour online) Trendline graph showing the relationship between rates of affrication of /k/ and lenition of /dʒ/ based on participant. Numbers represent participants.

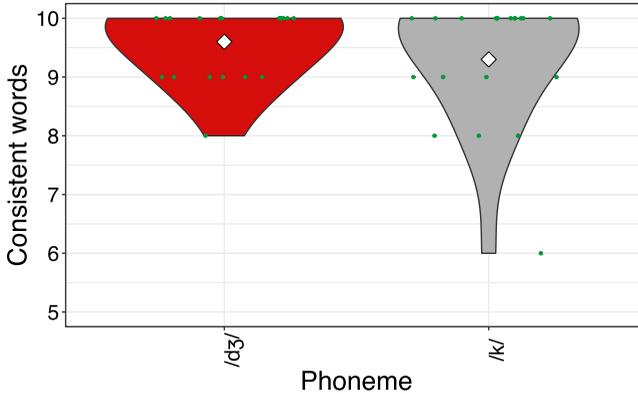


Figure 7 (Colour online) Number of words consistently produced with a specific variant of /k/ and /dʒ/ per participant, shown in violin plots. Individual subjects shown as green jittered dots.

the two was statistically significant, with a t -value of 3.244 (18 degrees of freedom), a p -value of .005, an F -statistics of 10.53, and an adjusted R^2 of 0.334.

However, even though participants were not generally consistent in their preference for each variant, they usually produced each word with the same variant in all tokens. Figure 7 shows that speakers were more likely to produce the /dʒ/ phoneme consistently across repetitions for each word. In contrast, speakers had more variable productions of the /k/ phoneme.

4 Discussion and conclusion

Traditional accounts of the [k~tʃ] and [dʒ~j] alternation in EA attribute most of their variability to sociolinguistic, lexical, and vowel quality-related factors. However, a potential conditioning role of consonants has previously been proposed for equivalent processes in Qatari Arabic (Mustafawi 2006, 2001) and Haʿili Arabic (Al Rasheedi 2015), and recently speculated for EA (Szreder et al. 2021). This study investigated the realisation of the variants in young, female, educated speakers, to examine the relevance of those accounts for modern EA. Based on previous research, we expected to find that the non-affricate variants [j k] would be preferred in the presence of homorganic coronal obstruents, demonstrating a phonological conditioning of the process (Hypothesis 1, Predictions (i)–(iv), and the traditionally ‘local’ variants [j tʃ] would pattern together in the productions of individual participants, illustrating sociolinguistic factors (Hypothesis 2, Prediction (v)).

The results of this study show that Hypothesis 1 was supported for /dʒ/, but not /k/. Specifically, the data confirm the Prediction (ii), that [j] would be significantly preferred over [dʒ] in the presence of homorganic segments /t d ʃ/. Prediction (i), that [k] would be significantly preferred over [tʃ] in the same context, was not supported. However, this is the case only when the homorganic segments are separated from the relevant phoneme by only a vowel. When the distance is greater, the OCP applies to neither phoneme. That is, Predictions (iii) and (iv) were not supported. Overall, our findings support a presence of an active OCP as a conditioning factor for /j/ lenition, compatible with that proposed by Mustafawi (2011) for QA, but only active in immediate neighbourhood. In contrast, the proposed OCP does not appear to be active for /k/-affrication in EA, and it potentially has an inhibitory effect (since /k/ is more likely to be affricated if the phonetically similar segment is farther away), providing an experimental confirmation of the effect found in the previous corpus analysis of EA (Szreder et al. 2021).

In addition, although we did not specifically focus on vowels, our data did not show an inhibitory effect of back vowels on /k/-affrication. The percentages of affricated tokens from the experimental study for words with a back vowel in (9) and with /i/ in (10) illustrate that both back and front vowels surfaced variably with the affricate:

(9)	[ʃo:tʃ]	‘thorn’	59%
	[ʃo:tʃa]	‘fork’	46%
(10)	[di:tʃ]	‘rooster’	64%
	[tʃi:s]	‘bag’	25%

Although not conclusive because of our limited data, this trend suggests that vowels do not currently have a strong effect on k-affrication in EA.

Therefore, k-affrication appears not to be consistently responsive to phonological conditioning of either consonants (OCP) or vowels. This suggests that it is no longer a synchronic phonological process in Emirati Arabic, but rather, that /k/ and /tʃ/ now function as separate phonemes in the dialect. A similar interpretation was proposed by Youssef (2014) for affrication in Baghdadi Arabic, which no longer appears to be a productive process, but rather a completed historical change, which resulted in two separate phonemes: /k/ and /tʃ/. In Baghdadi Arabic, this is evident in the presence of minimal pairs, as well as the variability being mostly restricted to lexical and sociolinguistic factors. Indeed, much of the variability of affrication in our experimental data could also be attributed to across-participant differences. Even though we do not have current data regarding minimal pairs in EA, and our analysis was restricted to a socially homogeneous sample, and even though within-word variability is still observable, it is likely that the two voiceless variants are better understood as phonological doublets.

In contrast, /dʒ/-lenition appears to be more dependent on phonological factors than previously reported, often being triggered by OCP conditions in the experimental data. This suggests that /dʒ/-lenition can still be described as an active phonological process in EA. Nonetheless, the process remains optional, and is often also sensitive to lexical effects and across-word, across-speaker variability.

Our test of Hypothesis 2, that affrication of /k/ and lenition of /dʒ/ are associated with socioeconomic status, supported Prediction (v), of a relationship between individual speakers such that speakers who affricate /k/ are also more likely to lenite /dʒ/. Nonetheless, there was a strong overall preference for lenited items, and an equally strong preference against affricated items. In other words, the asymmetry between the two processes was not restricted to their sensitivity to OCP, but also extended to participant preference. The observed pattern could suggest that there is a difference in the social distribution and perception of the two processes, such that affrication, but not lenition, is perhaps more stigmatised and less preferable, at least for our educated female speakers. A similar observation is made by Holes (2011), who takes it to suggest that the two processes do not pattern together anymore, especially among the educated urban speakers, such as our participants.

However, the interpretation of this finding in our data is complicated by the fact that while speakers were overall more likely to use dʒ-lenition than k-affrication, the speakers who did use the latter process were the ones who also relied more on the former. One possible explanation is that this somewhat paradoxical pattern was due to the fact that the participants were instructed to use the local dialect (as opposed to MSA) during the experiment. However, seeing that 25% of them still used the standard variant [k] 100% of the time, it is unlikely that the instruction strongly affected the participants’ performance. Another possibility is that the social distribution and perceptions of the variants are currently changing and/or are sensitive to sociolinguistic factors, which we did not investigate (e.g. family background,

regional differences). It is also the case that because we only studied the speech of educated female participants, we do not know if the same trend follows amongst the entire EA population, or just the educated elite. Further research on more heterogeneous samples is needed to systematically study the sociolinguistic effects on this variation.

In conclusion, the examination of phonological, lexical, and speaker effects in our elicited data reveals that the use of affricates in contemporary EA differs from the patterns reported for older data, as well as from those predicted by newer Optimality-Theoretic accounts of other Gulf dialects (Mustafawi 2006, 2011). Specifically, /k/-affrication is no longer sensitive to the original phonological conditioning effects of front vowels or the OCP constraint proposed for QA, and is overall less preferred, and sometimes entirely absent from the speech of young, educated EA speakers. Conversely, lenition is highly preferred by those speakers, and strongly facilitated by coronal stops. We propose that these results are best explained by treating affrication as a completed language change, and lenition as a synchronic process that is still phonologically active. At the same time, we note that changes in the social perceptions of this variation in EA may also be changing. Future research should further investigate whether these patterns are observable in the wider UAE population and across social groups.

Acknowledgements

This study was partially supported by the UAEU Research Startup grant 31H132, awarded to Marta Szreder. We are grateful to the participants and native EA informants who helped with the study, as well as to Mark Scott and Tommi Leung, without whom we would not have had a chance to collaborate. We also thank our research assistant, Chahla Ben-Ammar, who collected the data and provided invaluable help with all practical aspects of the project. Finally, we would like to extend our thanks to the anonymous reviewers and *JIPA* editors, who helped improve previous versions of this paper. Special thanks go to Associate Editor Alexei Kochetov, whose thorough and insightful comments were formative in shaping the final version of the manuscript. Any remaining errors are our own.

References

- Al-Ani, Salman H. 1970. *Arabic phonology: An acoustical and physiological investigation*. The Hague: Mouton.
- Al-Qouz, Muna Y. 2009. *Dialect contact, acquisition and change among Manama youth, Bahrain*. Ph.D. dissertation, University of Essex.
- Al Rasheedi, Eisa S. 2015. Affrication in Ha'ili Arabic: A rule-based approach. *International Journal of Linguistics* 7, 27–41.
- Al Rojaie, Yousef. 2013. Regional dialect levelling in Najdi Arabic: The case of the de-affrication of [k] in the Qasimi dialect. *Language Variation and Change* 25(1), 43–63.
- Articulate Instruments Ltd. 2012. *Articulate Assistant Advanced User Guide: Version 2.15*. Edinburgh: Articulate Instruments Ltd.
- As-Sammer, Maas. 2010. Phonological variation in Modern Standard Arabic: The case of the affricate /dʒ/. Oman as a sample. *Journal of Basrah Researches* (Humanities Series) 36(4), 29–57.
- Bateman, Nicoleta. 2007. *A crosslinguistic investigation of palatalization*. Ph.D. dissertation, University of California, San Diego.
- Bateman, Nicoleta. 2011. On the typology of palatalization. *Language and Linguistics Compass* 5(8), 588–602.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using *lme4*. *Journal of Statistical Software* 67(1), doi: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Bhat, Darbhe N. S. 1978. A general study of palatalization. In Joseph H. Greenberg (ed.), *Universals of human language*, vol. 2: *Phonology*, 47–92. Stanford, CA: Stanford University Press.
- Cantineau, Jean. 1936. Tadmorea (suite). *Syria* 17(3), 267–282.

- Dashti, Abdulmohsen A. 2018. A consonant shift in Kuwaitis' speech: Challenging the Bedouin Vs Sedentary Hypothesis. The case of [tʃ]. *International Journal of Applied Linguistics and English Literature* 7(1), 12–18.
- Dashti, Abdulmohsen A., Rahimah S. Akbar & Hanan A. Taqi. 2015. A change in consonant status: The Bedouinisation of the [j] sound in the speech of Kuwaitis. A case study. *International Journal of Applied Linguistics and English Literature* 4(5), 104–112.
- Dickins, James. 2007. *Sudanese Arabic: Phonematics and syllable structure*. Wiesbaden: Otto Harrassowitz.
- El Salman, Mahmoud. 2016. The use of the [ts] variant in the Arabic Bedouin dialects. *Ijel* 6(1), 118–127.
- Feghali, Habaka J. 2008. *Gulf Arabic: The dialects of Kuwait, Bahrain, Qatar, UAE, and Oman*. Hyattsville, MD: Dunwoody Press.
- Fischer, Wolfdietrich, Otto Jastrow & Peter Behnstedt. 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: Harrassowitz.
- Frisch, Stephen A., Janet Pierrehumbert & Michael E. Broe. 2004. Similarity avoidance and the OCP. *Natural Language & Linguistic Theory* 22, 179–228.
- Goldsmith, John A. 1976. *Autosegmental phonology*. Ph.D. dissertation, MIT.
- Greenberg, Joseph H. 1950. The patterning of root morphemes in Semitic. *Word* 6(2), 162–181.
- Halefom, Girma, Dimitrios Ntelitheos & Tommi T-C. Leung. 2012. A corpus of Emirati Arabic https://faculty.uaeu.ac.ae/dimitrios_n/Research/A-Corpus-of-Emirati-Arabic/a-corpus-of-emirati-arabic.html (accessed 4 January 2023).
- Hassan, Batoul. 2009. *Ideology, identity, and idealistic capital: A sociolinguistic investigation of language shift among the Ajam of Kuwait*. Ph.D. dissertation, University of Essex.
- Hassan, H. Mohamoud. 2017. The palatalization of /dʒ/ into /j/ in Emirati Arabic (EA): A rule-governed or random alternation? *International Journal of Linguistics* 9(5), 173–183.
- Hoffiz, Benjamin T. III. 1995. *Morphology of United Arab Emirates Arabic, Dubai dialect*. Ph.D. dissertation, The University of Arizona. <https://repository.arizona.edu/handle/10150/187179>.
- Holes, Clive. 1991. Kashkasha and the fronting and affrication of the velar stops revisited: A contribution to the historical phonology of the peninsular Arabic dialects. In Alan S. Kaye (ed.), *Semitic studies: In honor of Wolf Leslau*, vol. 1, 652–678. Wiesbaden: Harrassowitz.
- Holes, Clive. 1995. *Modern Arabic: Structures, functions and varieties*. London & New York: Longman.
- Holes, Clive. 2001. *Dialect, culture, and society in Eastern Arabia*, vol I: *Glossary*. Leiden: Brill.
- Holes, Clive. 2011. Language and identity in the Arabian Gulf. *Journal of Arabian Studies* 1(2), 129–145.
- Johnstone, Thomas M. 1963. The affrication of “kaʔ” and “qaʔ” in the Arabic dialects of the Arabian Peninsula. *Journal of Semitic Studies* 8(2), 210–226.
- Johnstone, Thomas M. 1965. The sound change j > y in the Arabic dialects of peninsular Arabia. *Bulletin of the School of Oriental and African Studies* 28(2), 233–241.
- Kochetov, Alexei. 2011. Palatalization. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, vol. 3, 1666–1690. Malden, MA: Wiley-Blackwell.
- Krishnamurti, Bhadriraju. 2003. *The Dravidian languages*. Cambridge: Cambridge University Press.
- Leben, William R. 1973. *Suprasegmental phonology*. Ph.D. dissertation, MIT.
- Leung, Tommi Tsz-Cheung, Dimitrios Ntelitheos & Meera Al Kaabi. 2020. *Emirati Arabic: A comprehensive grammar*. New York: Routledge.
- Monassar, Hisham. 2014. Modern Standard Arabic /ʒ/ and /dʒ/. *International Journal of Linguistics, Literature and Culture* 1(3), 56–69.
- Mustafawi, Eiman M. 2006. *An optimality theoretic approach to variable consonant alternations in Qatari Arabic*. Ph.D. dissertation, University of Ottawa.
- Mustafawi, Eiman M. 2011. The OCP as a synchronic constraint in Arabic. *Canadian Journal of Linguistics* 56(2), 229–246.
- Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in Generative Grammar*. Malden, MA: Blackwell.
- Prochazka, Theodore. 1988. *Saudi Arabian dialects*. New York: Kegan Paul.

- Qafisheh, Hamdi A. 1975. The affricates /tʃ/ and /dʒ/ in Gulf Arabic. *Journal of the American Association of Teachers of Arabic* 8(1), 24–40.
- Qafisheh, Hamdi A. 1977. *A short reference grammar of Gulf Arabic*. Tucson, AZ: The University of Arizona Press.
- R Core Team 2022. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Szreder, Marta, Donald Derrick & Chahla Ben-Ammar. 2021. Affricate variation in Emirati Arabic: An exploratory study. In Dimitrios Ntelitheos & Tommi T-C. Leung (eds.), *Experimental Arabic linguistics* (Series on Arabic Linguistics 10), 57–81. Amsterdam: John Benjamins.
- Versteegh, Kees. 1997. *The Arabic language*. New York: Columbia University Press.
- Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Youssef, Islam. 2014. Affrication in Baghdadi Arabic: Synchrony and diachrony. *Arabica* 61, 746–757.
- Youssef, Islam. 2021. Contrastive feature typologies of Arabic consonant reflexes. *Languages* 6(3), 141.