Erik Anonby*

Emphatic consonants beyond Arabic: The emergence and proliferation of uvular-pharyngeal emphasis in Kumzari

<https://doi.org/10.1515/ling-2019-0039>

Abstract: The complex and cross-linguistically uncommon phonological phenomenon of “emphasis” is best known from Central Semitic languages such as Arabic and Aramaic. It is, however, found to varying degrees in a number of non-Semitic languages in contact with Arabic. This paper describes how in Kumzari, an Indo-European language spoken around the Strait of Hormuz, uvular-pharyngeal emphasis has arisen through language contact and has proliferated through language-internal processes. Beginning with the retention of emphatic consonants in a direct, extensive lexification by Arabic dating back at least 1300 years, emphasis has progressively penetrated the language by means of lexical innovations and two types of sound changes in both borrowed and inherited vocabulary: (i) analogical spread of emphasis onto plain but potentially emphatic consonants; and (ii) a sound change in which *z* has been invariably recast as an emphatic *ẓ* with no plain counterpart. The role of the back consonants *w*, *x*, *q* and *h*, which induce emphasis on potentially emphatic consonants in diachronic processes but not synchronically, highlights the unique way in which this complex phenomenon operates in one non-Semitic language in contact with Arabic.

Keywords: Kumzari, Arabic, Indo-European languages, language contact, emphasis (phonological), uvularization, pharyngealization, sound change

1 Introduction

Although it is typologically uncommon, the phonological phenomenon known as “emphasis” is familiar from the Central Semitic family, where it is found in most varieties of Arabic and Aramaic. Emphasis is often defined simply as pharyngealization or velarization, but in reality it is a bundle of phonetic articulations that varies in its structure and behaviour according to language-specific parameters (Hoberman 1989, Hoberman 1995; Watson 2002; Embarki 2013).

***Corresponding author: Erik Anonby**, Leiden University Centre for Linguistics and School of Linguistics and Language Studies, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada, E-mail: e.j.anonby@hum.leidenuniv.nl; erik.anonby@carleton.ca

Outside of Central Semitic, emphasis is also found in a few languages in contact with Arabic, notably Berber languages and a handful of Indo-European languages including Domari, Kurdish and Kumzari. While mention of emphasis in these contact languages has appeared in the context of other studies (Applegate 1970; Hoberman 1985, Hoberman 1989; Blau 1989b; Matras 2007; Haig 2007), there have been few studies specifically devoted to the topic (but see Kahn's 1976 study of Kurdish, cf. Section 4.1 below). The present study provides, for the first time, an account of how emphasis has appeared and become progressively phonologized in one of the contact languages, Kumzari.

This paper opens with an overview of the Kumzari language, and describes how long-standing contact with Arabic has affected linguistic structures at all levels (Section 2). After an introduction of the emphatic consonant series in Kumzari, I situate this inventory within a general typology of emphasis that makes reference to Arabic, but also takes into account the significant phonetic and phonological variation in the patterning of emphasis in a diverse collection of languages where it has been documented (Section 3). Next, I catalogue Indo-European languages in which emphasis has appeared as a result of contact with Arabic, and describe the patterning of emphasis and the extent to which it has been incorporated in each language (Section 4). Of these languages, it is in Kumzari that emphasis is most profoundly phonologized, permeating the lexicon. How has this situation arisen?

In response to this question, the core of the study is devoted to investigating the appearance and proliferation of emphasis in Kumzari (Section 5). Here, I argue that many of the changes that take place in the patterning of emphasis here are a means of balancing social and linguistic asymmetries that have arisen through language contact. I begin by describing how emphasis initially appeared as a result of retaining emphatic consonants in a direct, extensive lexification by Arabic dating back more than 1300 years. I then move on to an examination of the language-internal innovations through which it has progressively penetrated the phonological system, and as a result, all sections of the lexicon: borrowed Arabic vocabulary, inherited vocabulary, and Kumzari-specific words. I give evidence for three major types of innovation in Kumzari, of which two are diachronic and one is synchronic, and all of which address phonological and sociolinguistic imbalances. First, emphasis has recurrently spread on many items in the lexicon from emphatic or emphasis-inducing consonants to potentially emphatic consonants through the mechanisms of analogical sound change. Second, an across-the-board sound change in which *z* has been invariably recast as an emphatic *ẓ* has resulted in the appearance of emphasis in hundreds of additional words. Third, in contexts where two consonants come together at a word-internal morpheme boundary, there is a co-articulation effect: emphasis

alternates synchronically by spreading onto potentially emphatic consonants. I conclude the article with a summary of the way in which emphatics have been phonologized in Kumzari, and reflect on implications for the typology of emphasis and sound change more generally (Section 6).

The analyses given here are based on a data collection containing 4400 lexical items as well as a number of longer texts collected from various speakers of the Musandam Peninsula dialect of Kumzari (van der Wal Anonby 2015; Anonby and van der Wal Anonby in prep.; Ali Hassan Ali; al-Kumzari 2006).

2 The Kumzari language

Kumzari is an endangered language spoken by about 4000 people in Oman, the United Arab Emirates and Iran. Speakers of the main dialect are found on the Musandam Peninsula of Oman, principally in the village of Kumzar and in part of the town of Khasab, and in small groups in cities along the Gulf coast of the United Arab Emirates. Laraki, a closely related dialect of the language, is spoken across the Strait of Hormuz by a single community on Larak Island in Iran (Anonby 2011a; Anonby and Yousefian 2011).

The Kumzari language was identified by Jayakar (1902), and a brief grammar sketch and lexicon appeared in Thomas (1930). A grammar of the language was produced by van der Wal Anonby (2015), and Anonby (2012) looked at nominal morphophonology. Although Kumzari has been treated as a mixed language (van der Wal Anonby 2014, van der Wal Anonby 2015), its core vocabulary and verbal morphology are in keeping with Skjærvø's (1989) classification of Kumzari within the Southwestern Iranian (SWIr) group of languages, itself within the Indo-European phylum. Still, many of its basic structures, including most of the lexicon, key elements of the phonological system, and a parallel verbal system may be traced to long-standing, acute influence from Arabic, including the neighbouring Shihhi dialects of Arabic (cf. Bayshak 2002; Anonby 2011b; van der Wal Anonby 2015); van der Wal Anonby (2014, 2015, forthcoming) looks at connections with other Semitic languages as well, including a South Arabian substrate.

The effects of language contact are especially evident in the Kumzari consonant inventory (Table 1).¹

In relation to Arabic varieties in the region, it is the array of contrastive stops and affricates (*p b t d ʈ ɖ ʧ j k g q ʔ*) in Kumzari which most noticeably

¹ This chart is based on Anonby (2011b: 375), but additionally includes the peripheral phoneme *l*, which is found in a small number of roots (4.6).

Table 1: Kumzari consonant inventory.

	labial	alveolar	emphatic alveolar	(alveo-) palatal	velar	uvular	pharyngeal	glottal
voiceless stops	<i>p</i>	<i>t</i>	<i>ṭ</i>	<i>č</i>	<i>k</i>	<i>q</i>		<i>ʔ</i>
voiced stops	<i>b</i>	<i>d</i>	<i>ḍ</i>	<i>j</i>	<i>g</i>			
voiceless fricatives	<i>f</i>	<i>s</i>	<i>ṣ</i>	<i>š</i>		<i>x</i>	<i>ħ</i>	<i>h</i>
voiced fricatives			<i>z</i>			<i>ġ</i>		
nasals	<i>m</i>	<i>n</i>						
approximant	<i>w</i>	<i>l / r</i>	<i>ʎ</i>	<i>y</i>		(<i>w</i>)		

distinguishes the consonant inventory. The Kumzari inventory is in most ways typical of a SWIr language, but the existence of an emphatic consonant series is exceptional. Even in New Persian, which has undergone significant influence from Arabic, emphatic consonants in words borrowed from Arabic are consistently reinterpreted as existing members of the phonological inventory (Windfuhr and Perry 2009: 422; see also Paradis and LaCharité 2001).

The initial phonologization of emphasis in Kumzari originates in the retention of emphatic consonants in words borrowed from Arabic (Section 5.1). As a result the robust Kumzari series, which contains the emphatic alveolars *ṭ ḍ ṣ ʎ* as well as the voiceless pharyngeal *ħ*, corresponds closely to the Arabic set from which it is derived (Section 3.1). The local character of language contact is evident in that, as in the neighbouring Shihhi dialects of Arabic, the voiced pharyngeal *ʕ* characteristic of most Arabic varieties is absent from the language.

Still, emphasis in Kumzari is not a simple copy of the Arabic system on which it is based: there are some important aspects of the system that have become accentuated in Kumzari, and a number of Kumzari-internal innovations have arisen. For example, whereas pharyngealization is often the dominant secondary articulation among emphatic consonants in Arabic (Section 2), uvularization is a central quality of the Kumzari system (Section 4.6). In addition,

there is a diachronic process of diffusion of emphasis to non-emphatic consonants which operates in words of Arabic as well as non-Arabic origin (Section 5.3). Most surprisingly, an across-the-board sound change ($z > ẓ$) has resulted in the displacement of a plain consonant from the inventory in favour of its emphatic counterpart (Section 5.4). A number of these phenomena are indirectly and directly driven by the need to account for imbalances in the language and its social context; in Section 5.4.3 below, this idea will be explored in relation to the last of these three innovations.

Before examining the patterning of emphasis in Kumzari in further detail, it is helpful to situate it within a general typology of emphasis (Section 3), and to consider it alongside other systems of emphasis that have arisen among Iranian varieties as a result of contact with Arabic (Section 4).

3 Typology of emphasis

Traditionally, the study of emphasis has been closely associated with Arabic. However, it is pertinent for other languages in the Central Semitic family, especially Aramaic varieties (Dolgopolsky 1977; Hoberman 1985, Hoberman 1989; Khan 1999), and for contact languages in regions where Central Semitic languages are found (Applegate 1970; Matras 2007; see also Section 4 below).² This section first looks at the ways in which emphasis has usually been understood in the literature, but shows that its nature and boundaries are not clearly defined. These variable assessments of emphasis are important in understanding how this rough-edged phenomenon behaves in Kumzari, a non-Semitic language in contact with Arabic.

3.1 Emphatic consonant inventories

Phonologically speaking, the term “emphasis” is prototypically applied to a series of dental consonants, well-known from Classical Arabic, which are distinguished from their “plain” counterparts through a salient co-articulation (Ar. *ʔiṭbāq* ‘spreading and raising’) most commonly defined as pharyngealization (Card 1983: 13–14; Watson 2002; Bakalla 2009: 421) or velarization (Fischer and Jastrow 1980: 56; Holes 2004: 57; Embarki 2013), but also treated as uvularization by some scholars

² An extended category of emphatics, wider than the definition followed here, includes ejective consonants in South Arabian languages that are reflexes of the Central Semitic emphatic series (see, for example, Watson and Bellem 2010).

(Dolgopolsky 1977: 1; McCarthy 1994; Shahin 1997, Shahin 1998; Zawaydeh 1998). There are four contrastive emphatics in Classical Arabic (CA) and, by extension, Modern Standard Arabic (MSA) which fall under this label: *ṭ* [tˤ], *ḍ* (CA [kˤ], MSA [dˤ]; see Section 5.4.2), *ṣ* [sˤ], and *ḍ* [ðˤ] (Bakallah 2009: 421; Lehn 1963: 29).

The idea of emphasis is generalized under the Arabic label *tafxīm* (lit. ‘thickening, enlarging, emphasizing’) to additionally designate emphatic allophonic variants of the consonants *l* and *r* in CA and MSA (Owens 2006: 25; Holes 2004: 57–58; cf. Ferguson 1956) as well as similarly co-articulated consonants in other dialects (Jakobson 1957/1971; Bakallah 2009: 421; Grigore 2011). In some cases, inventories of contrastively emphatic consonants are larger than that of Classical Arabic: for example, Nigerian Arabic has six underlyingly emphatic consonants *ṭ ḍ ṣ ṣ̣ ṣ̣̣ ṣ̣̣̣* and *r*, as well as emphatic allophones of several additional consonants (Owens 2006: 25); and the Northeastern Neo-Aramaic dialect of Amedia in Iraq contains ten contrastive pharyngealized consonants: *p b t d ʕ ʕ̣ ʕ̣̣ ʕ̣̣̣ ʕ̣̣̣̣* and *r* (Hoberman 1985: 224).

In a seminal article on the linguistic nature and behaviour of emphasis, Lehn (1963: 29) claimed (without providing a reference) that writers in the medieval Arabic grammatical tradition included the characteristically Semitic pharyngeal consonants *ḥ* (typically [h] or [ħ]) and *ʕ* ([ʕ], [ʕ̣], [ʕ̣̣])³ with the emphatic dentals under the label of *tafxīm*. This assertion, for which I have not been able to locate an original source, has been taken up and explored by other linguists (Paddock 1970 in Card 1983: 91; see also Marçais 1948). In such a framework, *ḥ* and *ʕ* would be viewed as the emphatic counterparts of the glottal fricative *h* and the glottal stop *ʔ* (cf. Jakobson 1957/1971: 519; Hoberman 1985: 227–228; Kahn 1976: 28–29; Watson 2002: 268). This assessment recognizes a common post-velar articulation for co-articulated emphatics and pharyngeals, and some similar phonetic effects of both groups (see Sections 3.2 and 3.4 below).

In the work of the earliest Arabic grammarians and in much of the literature today, however, *ḥ* and *ʕ* are excluded (Lehn 1963: 29). This arises from the fact that, in Arabic, there is little diachronic connection between co-articulated emphatics and pharyngeals (Hoberman 1985: 223; cf. Section 3.4 below). In most varieties of Arabic, they are also synchronically discrete; and in Cairo Arabic and Palestinian

³ Traditionally, the Semitic consonants *ḥ* and *ʕ* have been considered voiceless and voiced pharyngeal fricatives and transcribed as [h] or [ʕ] respectively in IPA. However, as Esling (1996, 1999) demonstrates, these labels do not provide an accurate characterization of voicing or place of articulation. Such consonants are in many cases actually epilaryngeal fricatives [ħ] and [ʕ̣], some of which are accompanied by aryepiglottic vibration; and *ʕ* in particular can also be realized as an approximant (Laufer 1996) or as an aryepiglottic-epiglottal stop [ʔ̣] (Hessellwood 2007). The phonetic implementation of these phonemes varies greatly within and across Arabic dialects (Hessellwood 2007; Hassan et al. 2011; Moisik 2013).

Arabic, there is even a contrast available between plain and emphatic pharyngeals (Lehn 1963: 31–32; Card 1983: 18, 22). From the perspective of current phonological theory, this may seem unusual, since pharyngeals and pharyngealized consonants are often grouped together as a natural class with a shared distinctive feature of [pharyngeal], [RTR] (retracted tongue root), [guttural] or something similar (Jakobson 1957/1971; Clements and Hume 1995: 273–274; Hayward and Hayward 1989; cf. Davis 1995: 472). However, as Card (1983: 16) and Davis (1995: 483) have pointed out, there are phonetic as well as phonological differences between the two groups of consonants in Arabic (see also Sections 3.2 and 3.3 below).

Despite these differences, and in light of the great variation in patterning across languages, it is worth reconsidering the idea that in some situations, pharyngeal consonants can in fact be classed along with co-articulated emphatics as part of a broader meaning of emphasis. Examples of historical interaction between co-articulated emphatics and pharyngeal consonants have been observed in sporadically distributed varieties of Arabic (Brockelmann 1908 and Blanc 1953 in Hoberman 1985: 223), but no consistent patterns of interaction have emerged. Some synchronic support for a relationship between the two groups of consonants is, however, found in certain varieties of Kurdish: there, the two groups not only demonstrate similar phonetic behaviour (e.g., phonetic effects on vowels; see Kahn 1976: 22), but also exhibit complementary distribution in some contexts (Kahn 1976: 87–89) and free (or stylistic) variation in others (e.g., *taʕzi* ~ *ṭazi* ‘fresh’) (Kahn 1976: 50; Hoberman 1985: 229). Further evidence that the two groups could constitute a single natural class comes from the observation that in some languages, pharyngeal consonants are consistently correlated with the diachronic retention or even appearance of contrastive emphasis on other potentially emphatic consonants; this will be explored in 3.4 below.

3.2 Articulatory properties of emphasis

As mentioned in the preceding section, the articulatory basis of emphasis is usually identified variously as pharyngealization, velarization, or uvularization. In fact, however, it is almost always (or perhaps always) the case that emphasis functions as a complex of secondary articulations which can also include labialization, glottalization and other co-articulations (Jakobson 1957/1971: 511–513; Hoberman 1988, Hoberman 1989: 77, Hoberman 1995: 841; Watson 2002: 269; Embarki 2013: 31–34). This complexity is reflected in an abundance of labels in the linguistic literature, as listed by Lehn (1963: 29): along with pharyngealization, velarization and uvularization, the terms retraction, strong articulation, u-resonance and heaviness have been applied to emphasis.

In articulatory terms, emphatic consonants are typically characterized at least by a narrowing of the upper pharynx (Watson 2002: 269). Acoustically, this narrowing results in the lowering of the second formant (Card 1983: 13–15; cf. Jakobson 1957/1971: 512; Watson 2002: 270). Emphatic consonants also have in common an effect on neighbouring vowels, in particular the low vowels *a* and *ā* (Bakallah 2009: 422; see Section 3.4 below). Hoberman (1989: 77) questions the idea that the production of emphasis necessarily varies from one variety to another, but a wide range of articulatory analyses have been proposed for emphasis in different languages and dialects. In Card's (1983) study of Palestinian Arabic, pharyngealization is equated with emphasis; but Hoberman (1985: 224) observes that pharyngealization is an insignificant component of emphasis in Iranian dialects of Neo-Aramaic. Elsewhere, in varieties such as Shihhi Arabic and Kumzari, stricture at a higher location, namely uvularization, is dominant (cf. Anonby 2011a: 376).

An important point introduced in the previous section, and one that recurs in the literature on emphatics, is the distinction between the secondary articulation of emphatics and the primary articulation of pharyngeal consonants *ħ* and *ʕ*. Although Card (1983) highlights some similarities between the two sets of sounds in Arabic, she concludes that their articulation is fundamentally different: whereas the main secondary constriction of co-articulated emphatics is located in the upper pharynx, the primary constriction of pharyngeals is located consistently further down in the pharynx; and that whereas acoustically, co-articulated emphatics exhibit a strong lowering of the second formant, the effect of pharyngeals on the second formant should be seen as incidental (pp. 13–15, 90–97; this inference will, however, be revisited in the next section). It is precisely this state of affairs which motivates Dolgopolsky (1977: 1), followed by Shahin (1997) and Zawaydeh (1998), to argue that Arabic emphasis should be referred to as uvularization rather than pharyngealization; in Dolgopolsky's words, this would distinguish it from “lower” pharyngealization such as that found in languages of the Caucasus (see also Colarusso 1975).

3.3 Phonological parameters of emphasis

In the simplest scenario, such as that which is often put forward for Classical Arabic (CA) and Modern Standard Arabic (MSA), emphasis may be viewed as an intrinsic phonological property of emphatic consonants. While recognizing that emphatic consonants have predictable phonetic effects on neighbouring segments, most accounts of the phenomenon in CA and MSA, and even some accounts of spoken Arabic varieties, have taken for granted the idea that emphasis can only have as its phonological domain the emphatic consonants themselves (cf. the publications listed in Lehn 1963: 34–35).

In most spoken varieties of Arabic, and in other languages with emphasis, the segment is not in fact the minimal domain; the domain with which this feature is associated tends to vary from one dialect to another (Watson 2002; Holes 2004: 58). In Cairo Arabic, for example, the minimal underlying domain of emphasis is the syllable (Lehn 1963: 32; Hoberman 1989: 83); and in Qatari Arabic (Bukshaisha 1985: 217–219), as well as Neo-Aramaic varieties spoken in Iran, it is typically associated with whole words (Hoberman 1985: 225). In Neo-Aramaic in particular, there are many words in which no intrinsically emphatic underlying consonant can be distinguished.

In most varieties, though, the ultimate source of emphasis is in fact situated in underlyingly emphatic consonants, also known as “primary” emphatics (Fischer and Jastrow 1980: 56–57). Yet often, even here it is not confined to the emphatic consonants themselves; rather, it spreads synchronically within a word according to specific structural constraints (Watson 1999, 2002). In their work on phonological parameters of emphasis, Hoberman (1989: 89–90) initiates and Davis (1995) and Watson (1999, 2002) further develop an overview of three parameters which are particularly helpful in understanding the source and synchronic spread of emphasis in such languages: direction of spread, transparent consonants, and opaque consonants. Regarding the direction of spread in particular, there are variations in different languages and dialects: whereas in Palestinian Arabic, emphasis spreads both progressively (to the right) and regressively (to the left) within a word (cf. also Davis 1995), it spreads only progressively in the Christian Neo-Aramaic dialect of Iranian Azerbaijan; in Kurdish, emphasis does not spread, being systematically limited to a single consonant in any word (Kahn 1976: 49). The extent of spreading in languages where it does occur is affected by the presence of two kinds of segments: “transparent” segments, which although they cannot themselves be associated with the emphatic feature, neither block nor initiate spreading; and “opaque” segments which block the spread of emphasis because they are invariably specified as not associating with emphasis. Examples of transparent segments are *w* and *ū* in Palestinian Arabic; and in several varieties, the high, non-back segments *y*, *ī* and *i* are opaque (Fischer and Jastrow 1980: 57; Card 1983: 15, 118). Hoberman (1985: 224) suggests further that many other dialects of Arabic make use of opaque and transparent segments, as do Neo-Aramaic dialects outside of Iran.

3.4 Variations in the description of the scope of emphasis

The parameters outlined in the preceding section provide a satisfactory account of the way in which emphasis functions synchronically, but the scope of the

phenomenon differs significantly in the various accounts. The limited data available on the topic suggest that diachronic sound changes involving emphasis may operate by means of a more general, or less clearly-defined, mechanism.

There are a number of additional consonants, most of which are post-velar, which in various frameworks have been recurrently identified with co-articulated emphatics: *r w x ġ q ḥ* and *ʕ*. Importantly, along with the core emphatic series *ṭ ḍ ṣ ḏ*, which was characterized by *ʔiṭḃāq* (see 3.1 above), the early Arabic grammarian Sibawaih, and those following his work, considered *x ġ q* (but not *w*) a second subset of consonants exhibiting *ʔistiʕlāʔ* ‘raising (of the tongue back)’ (Sibawaih II.452 [1970], Bakallah 2009: 421; cf. Jakobson 1957/1971: 515–516), although his justification for the broader category was related to phonological effect on neighbouring vowels rather than the mechanics of production.⁴ As mentioned above (Section 3.1), Lehn claims that later Arabic grammarians included the pharyngeals *ḥ* and *ʕ* as part of this larger phonological set of *ʔistiʕlāʔ* (Lehn 1963: 29), but since no reference to his assertion can be traced, this point remains in doubt. In the contemporary literature, Holes (2004: 57–58) observes an articulatory parallel between co-articulated emphatics and the consonants *r x ġ q* in that all of them result in the backing of neighbouring vowels. While Card (1983: 90–97) dismisses a phonological relationship between *r w ḥ ʕ* and co-articulated emphatics, she notes that the acoustic lowering of the second formant characteristic of co-articulated emphatics is to some degree present in these consonants, especially *r* and *w*. Finally, McCarthy (1994: 218–223) and Herzallah (1990: 55) argue for a class of [pharyngeal] “gutturals”; this class, based primarily on co-occurrence patterns, includes the co-articulated emphatics, the uvulars *x ġ q*, and the pharyngeals *ḥ* and *ʕ*; additionally, in contrast to many authors, they include the laryngeals *h* and *ʔ* in this set.

Although the consonants *r w x ġ q* are not themselves contrastively emphasized, and although (as discussed above in 3.2) there are important phonetic and phonological differences between co-articulated emphatics and pharyngeals *ḥ* and *ʕ* in many varieties, there is in certain languages a strong correlation between some of these consonants and the diachronic maintenance of emphasis, or even its appearance on neighbouring, potentially emphatic consonants (i. e., non-emphatic consonants for which an emphatic counterpart exists in the language). Hoberman (1985) carefully demonstrates that in two dialects of Northeastern Neo-Aramaic (the Christian dialect of Urmi and the Jewish dialect of Iranian Azerbaijan), words which in Old Aramaic contained either pharyng-

⁴ This important observation was pointed out by an anonymous reviewer.

ealized consonants or the pharyngeals *ħ* and *ʕ* are reflected as emphatic words (that is, entire words which carry emphasis; see Section 3.3). Similarly, in the Jewish Neo-Aramaic variety of Zakho, the two pharyngeals as well as *q* have been retained in words which in Old Aramaic also had co-articulated emphatics, even though they are dropped in other environments (pp. 224, 228–229). In some Kurdish dialects, where emphasis can be marked on only one point within a word, other emphatic consonants cannot coexist with *q*; this suggests that it is the emphatic counterpart of *k* (Kahn 1976: 23, 26–27, 87–88). In these same dialects, a borrowed *ʕ* has been reinterpreted as pharyngealization on a consonant elsewhere in the word (e. g., Ar. *zaʕfaran* > Kurdish *zafaran* ‘saffron’; p. 50). Historical and comparative examples are limited for *r*, but in the Qabbe dialect of Shihhi Arabic, where the consonant *r* is strongly uvularized, the most likely explanation for the appearance of emphasis on the final consonant of *rōṣ* ‘head’ (cf. Classical/Modern Standard Arabic *ra*’s) is the co-occurrence of *r* (author’s field notes, 2010). Critically, as will be shown for Kumzari, the presence of *w* *x* *q* and *ħ* have repeatedly led to the diachronic appearance of emphasis on potentially emphatic consonants within a word, according to the same mechanism by which emphasis diffuses from co-articulated emphatics (Section 5.3.1). Conversely, for *ħ* in particular, there is evidence that this pharyngeal may arise as a result of emphasis spreading onto *h* from the Kumzari emphasis-inducing consonants mentioned here (Section 5.3.2).

In short, it is clear that the functioning of emphasis, and the boundaries of such a natural class, vary across languages and can be difficult to define in any given language. The manifold possibilities in the evolution and patterning of emphasis are likely traceable to the phonetic and phonological complexity of the phenomenon itself (Section 3.2). As regards diachronic maintenance and spread of emphasis, the consonants *r* *w* *x* and *q* function in some languages like co-articulated emphatics; in this sense, they may be regarded within these systems as somehow intrinsically emphatic. In a discussion of emphasis, then, the idea of opacity (i. e., the obligatory specification of a segment with a feature) need not always be limited to consonants that block the synchronic spread of emphasis (cf. Section 3.3 above); it should be extended to intrinsically emphatic segments. As for *ħ* and *ʕ*, the synchronic evidence that pharyngeals could in some languages be considered emphatic counterparts of *h* and *ʔ* (Section 3.1) is strengthened by the parallels between the two pairs: centrally, the fact that both can function in the same way as co-articulated emphatics with regard to the diachronic maintenance, or even induction, of emphasis on neighbouring consonants; and, as shown in Kumzari, the contrastively emphatic nature of *ħ* as reflected by the recasting of *h* as *ħ* through emphatic diffusion (Sections 5.2.2 and 5.3.2).

3.5 Emphasis in language contact situations

Emphasis in language contact situations is consistently characterized by instability. In areas where Central Semitic and languages from other families are in contact with each other, major changes in its status and distribution have taken place in both directions.

To begin with, languages which historically contained emphatic consonants have responded in two ways. In most varieties of Arabic, emphatic consonants have not only been maintained in contact situations; in many cases, non-emphatic consonants in Arabic words borrowed from contact languages have been reinterpreted as emphatics: for example, *ʔaṭlas* ‘atlas’ (from Greek *átlas*), *baṭāṭā* ‘potato’ (from Portuguese *batata*), *ṣard* ‘extreme cold’ and *ṭāzaj* ‘fresh’ (from Middle Persian *sard* ‘cold’ and *tāzag* ‘flowing, fresh’). (The phonetic basis of this type of reinterpretation is explored in Jakobson 1957/1971: 512–513.) This contrasts with the situation of several geographically peripheral varieties of Arabic, including Maltese, some dialects of Chadian and Sudanese Arabic, and Arabic spoken in Uzbekistan, which as a result of intense contact with other languages have lost the emphatic series from their consonant inventories (Fischer and Jastrow 1980: 56).

Conversely, varieties in contact with languages containing emphatic consonants have responded in a number of ways. In most cases, even when contact varieties have undergone extensive lexification from Arabic, they conform borrowed words to their own phonological inventories, in particular with respect to emphatic consonants (Paradis and LaCharité 2001). Persian and Turkish are two important examples of this. In other cases, the phonological impact of contact is more substantial, and a subset of emphatic consonants has been retained in borrowed words; this is the situation for Keshmi (Section 4.4), Pashto (Section 4.5) and Swahili (Lodhi 2003: 157), although in the case of Pashto and Swahili (and, most likely, many other languages in the broader Islamic sphere), emphatic consonants are further restricted to formal acrolects. Finally, there are a limited number of languages in which emphasis, having arisen through lexical borrowing, has been integrated into the phonological system as a result of intense, long-standing contact.⁵ In these languages, emphasis is not limited to

⁵ Although it is not the focus of this paper, the situation of Aramaic is complex and remarkable, since it has participated in changes in both directions. In earlier centuries, when Aramaic was widespread and dominant, it is possible that its emphatic series was borrowed by neighbouring languages, including Kurdish (see Section 4.1). There are phonological similarities between the behaviour of emphasis in Kurdish and in Neo-Aramaic (Hoberman 1985: 229), and in Amedia, the Iraqi region in which the Neo-Aramaic emphatic inventory is most elaborate

borrowed vocabulary, but over time begins to appear – to varying degrees – in inherited (i. e., non-borrowed) vocabulary. Examples of languages in which emphasis has become integrated to such an extent are Kurdish (Section 4.1), Zazaki (Section 4.2), Dezfuli-Shushtari (Section 4.3), Domari (Matras 2007: 152), some Berber languages (Applegate 1970: 592–593, 604) and, as will be explored in depth below, Kumzari (Sections 4.6–5.5). In each situation, and often even for subdialects of the same language, the mechanisms driving the proliferation of emphasis appear to be unique (Hoberman 1989: 91), but there are no comprehensive accounts of how this happens in any particular variety. The present paper has been written in response to this question.

4 Emphasis in Iranian languages

The meeting of the Arabic and Iranian language blocs follows, roughly speaking, the edge of the mountain ranges to the north and east of the Mesopotamian plain, and the length of the Gulf. While all of the Iranian languages along this boundary have been significantly influenced by Arabic structures, most of the major varieties here – Luri, Bakhtiari, Khuzestani Persian, Bandari, Lari, and Balochi (see the map in Bruk and Apenchenko 1964, plates 70–71) – have not absorbed any emphatic consonants into their phonological inventories. Rather, the appearance of emphatics in this language family has been confined to two main areas: first, the Kurdish-speaking region to the north-west, along with the largely overlapping Zazaki language area and the varieties of Dezful and Shustar a short distance to the south-east of the Kurdish area; and second, the region around the Strait of Hormuz, where Keshmi and Kumzari are spoken. There are no emphatics in any register of present-day Persian, and these consonants have likely never been part of the colloquial language, but alveolar emphatics as well as the pharyngeals *ħ* and *ʕ* were still pronounced in “literary” registers of spoken Persian at least until the mid-1800s (Pisowicz 1985: 103). In contrast, in a similar

(Hoberman 1985: 224), the Kurdish emphatic inventory is also richer than that of other Kurdish varieties (Blau 1989b: 329). But Neo-Aramaic varieties have lost inherited emphatic consonants in many words, likely as a result of long-standing contact with Iranian and Turkic languages. In place of these, they have gained many new emphatics through borrowing from Arabic, Kurdish, liturgical Hebrew and Classical Syriac. Although this issue has yet to be investigated in depth for the various Neo-Aramaic varieties, diverse strategies of borrowing have led to divergent patterning of emphasis in both lexicon and phonological systems (Hoberman 1985: 224; Haig 2007: 165).

scenario, the two pharyngeal consonants are still retained in a prestige variety of Pashto, a language of Afghanistan and Pakistan.

As a point of comparison with Kumzari emphatics, which will later be explored in detail (Sections 4.6–5.5), the following subsections (Sections 4.1–4.5) examine the status of emphatics in the phonological systems of Kurdish, Zazaki, Dezfuli-Shushtari, Keshmi and Pashto. In light of the complex nature of emphasis and its behaviour in some languages (Section 3.4), pharyngeal consonants are included here along with co-articulated alveolar emphatics. It is evident from the following discussion that, although these consonants are contrastive in each of the languages, with the possible exception of Kurdish, they are limited to the periphery of the phonology: for the most part, the inventory of emphatic consonants is small and is limited to Arabic loanwords; and here, in contrast to the situation in Kumzari, emphasis rarely diffuses diachronically, and does not spread synchronically from emphatics to other consonants.

4.1 Kurdish

The best-known case of emphasis in the phonological inventory of an Iranian language is that of the (non-Southwestern) West Iranian language Kurdish; but even here, accounts of the phenomenon are for the most part fragmentary.

MacKenzie (1961) observes the existence of pharyngeal and co-articulated emphatic consonants in a number of Kurdish dialects. The most significant contribution to the topic is that of Kahn (1976), who points out that the behaviour and extent of emphasis in Kurdish varies widely among dialects, and even among speakers of the same dialect (25, 106–108). Kahn examines emphatic consonants as part of a description of borrowing and variation in Northern Kurdish (“Kurmanji”) dialects spoken around Rezaiyeh in the West Azerbaijan Province of Iran. There, emphasis is contrastively (but variably) found with a range of consonants (see the discussion below), yet shows a restricted distribution in the lexicon as well as the phonology (43–44). Most occurrences of emphatic consonants in these dialects are simple retentions in Arabic loanwords; emphasis has, however, arisen spontaneously in a limited number of inherited vocabulary items, and Arabic (but not Persian or Turkish) vocabulary is also susceptible to reinterpretation (see the examples in this section below). While the presence of low vowels *æ* [æ] and *a* [a] or the back vowel *o* [o] does not necessitate the appearance of emphasis, it is a recurring phonological context in which emphasis arises on consonants (90). Due to a language-specific constraint whereby emphasis can be associated with only one consonant in a word, it never spreads, either diachronically or synchronically, from one emphatic consonant to another (42–51) (although in borrowed words it is occasionally displaced from one

consonant to another to provide a more favourable location, generally closer to the beginning of a word; e. g., Ar. *zaʕfarān* > Kurdish *zæfæran* ‘saffron’, p. 50).

Surprisingly, it is in the Northern dialects of Kurdish – which have had less direct contact with Arabic – that the phenomenon appears to be most deeply established (28). The question of whether the initial emergence of emphasis can therefore be at least partially attributed to longstanding contact with Aramaic (cf. Section 3.5) has been raised (cf. Windfuhr’s comment in Hoberman 1985: 229; Haig 2007: 165; cf. Kahn 1976: 5, 12), but not addressed.

In any case, as mentioned above, most Kurdish words with emphatic consonants are borrowed from Arabic. It is widely recognized that Kurdish retains the pharyngeal consonants *ħ* and *ʕ* in Arabic loanwords (Northern Kurdish data from Kahn 1976: 88):

- ħæmlæ* ‘attack (n.)’ (cf. Ar. *ħamla*)
ræħmæ ‘mercy’ (cf. Ar. *rahma*)
ʕæmæɫ ‘task’ (cf. Ar. *ʕamal*)
læʕɪn ‘mercy’ (cf. Ar. *laʕɪn*)

In some dialects, the inventory of borrowed emphatics extends to the pharyngealized alveolars *ʈ* *ʂ* and *ʐ* (28, 81, 88).

- ʈæbæk* ‘layer, level’ (cf. Ar. *ṭabaq*)
ʈayæ ‘tyre’ (cf. English > coll. Ar. *ṭāya*) (Ferhad Shakely, pers. comm. 2011)
ʂabūn ‘soap’ (cf. Ar. *ṣābūn*)
ʂibe ‘morning’ (cf. Ar. *ṣabāḥ*)
ʐilm ‘oppression’ (cf. Ar. *ḍulm*)

A survey of the literature reveals, however, that in various dialects there are a number of cases where these phonemes have diffused into inherited Iranian vocabulary. The Kurdish examples below are taken from Kahn (1976: 25–29, 50 and 108) unless otherwise indicated.

- biħæšt* ‘paradise’ (cf. MP *wahišt*, NP *bihišt*)
ħæft / *ħæwt* ‘seven’ (cf. MP/NP *haft*) (McCarus 2009: 592)
ħæšt ‘eight’ (cf. MP/NP *hašt*)
ʕasman ‘sky’ (cf. MP *asmān*, NP *āsimān*) (Haig and Matras 2002: 5)
mʕær ‘snake’ (cf. MP/NP *mār*) (Haig 2007: 167)
tʕæl ‘bitter’ (cf. MP *taxl*, NP *talx*, Kumzari *ṭaḥl*) (Haig 2007: 167)
ʂæd ‘hundred’ (cf. MP/NP *sad*) (Kahn 1976: 25; McCarus 2009: 592)
ʂæ/ʂæg ‘dog’ (cf. MP/NP *sag*) (Kahn 1976: 28; McCarus 2009: 592)

<i>şal</i>	‘year’ (cf. MP/NP <i>sāl</i>) (McCarus 2009: 592)
<i>tazæ</i> / <i>ṭazi</i>	‘fresh’ (cf. MP <i>tāzag</i> > NP <i>tāza</i> , Ar. <i>ṭāzaj</i> ; Kumzari <i>ṭāzag</i>) (Ferhad Shakely, pers. comm. 2011; Kahn 1976: 50)
<i>ṭirs</i>	‘fear’ (cf. MP/NP <i>tars</i>)
<i>baṣī</i>	‘bird (sp.)’ (cf. MP/NP <i>bāz</i> ‘bird of prey’)
<i>ṣawa</i>	‘bridegroom’ (cf. MP/NP <i>dāmād</i>)

Alongside *ḥ* *ʕ* *ṭ* *ş* and *ṣ*, there are dispersed accounts of additional contrastive emphatics in inherited vocabulary. The best-known of these is the velarized or pharyngealized alveolar lateral approximant *l̤* ([*l̤*] = [*lʕ*]/[*lʕ̥*]), which contrasts with a plain alveolar counterpart *l* in Central and Southern Kurdish varieties (Haig and Matras 2002: 5; Kahn 1976: 32). This opposition is confirmed by minimal pairs such as *gul* ‘rose’ and *gul* ‘person with leprosy’ (McCarus 2009: 592). Additional pairs from Central Kurdish are *kæḷ* ‘water buffalo’ vs. *kæl* ‘hill’, *pæḷæ* ‘spot, stain’ vs. *pælæ* ‘rush (n.)’, and *kuḷ* ‘sorrow’ vs. *kul* ‘dull’ (Jaffer Sheyholislami, pers. comm. 2012).

In her overview of Kurdish, Blau (1989b: 329) mentions that Kurdish varieties spoken in Armenia also contain a contrastive voiceless pharyngealized palato-alveolar affricate *č̤* [*tʃ̤*], and Kahn (1976: 27–28) provides the examples *čæng* ‘fistful’ and *pečan* ‘to wrap’ from Northern Kurdish in Iran.

Kahn cites the additional occurrence of a contrastive *p̤* [*p̤*] in Northern Kurdish (25–26):

<i>pan</i>	‘wide’ (cf. MP/NP <i>pahn</i>)
<i>pænīr</i>	‘cheese’ (cf. MP/NP <i>panīr</i>)
<i>pōpin</i>	‘cowlick’

A further consonant in this set, *r̤*, can be realized as either flapped [*r̤*] or trilled [*r̤̃*] but for which emphatic co-articulation is distinctive. It is found in some varieties of the Mokriyani dialect of Central Kurdish in place of the (reportedly) non-emphatic trilled phoneme *ř* known from other Kurdish varieties (data from Jaffer Sheyholislami, pers. comm. 2012).

<i>bæṛæ</i>	‘gilim (flat-weave carpet)’ (cf. Mokriyani Kurdish [MK] <i>bæṛæ</i> ‘war front’)
<i>jaṛ</i>	‘fallow field’ (cf. MK <i>jar</i> ‘occasion’)
<i>kæṛ</i>	‘deaf’ (cf. MK <i>kær</i> ‘donkey’)

As mentioned in 3.4 above, Kahn (1976: 26–27) includes the voiceless uvular stop *q* as a member of the Kurdish emphatic series, since its distribution mirrors that of other emphatic consonants.

Finally, as stated earlier in this section, it should be pointed out that there are number of Arabic loanwords in spoken Kurdish where the emphatic is not present in the source item, but has arisen in Kurdish subsequent to borrowing (Kurdish items from Kahn 1976: 89–90):

ʕædæb ‘polite’ (cf. Ar./NP *ʔadab*)

ʔælʕan ‘now’ (cf. Ar./NP *ʔalʔān*) (also in Rahimpour and Dovaise 2011: 75)

jaħel ‘young’ (cf. Ar. *jāhil* ‘naïve’, coll. Ar. ‘young’, Gulf Ar. dialects ‘child’)

ʕælal ‘clear (water)’ (cf. Ar. *zulāl* ‘cold water’, coll. Ar. ‘clear (water)’)

Specifically, Kahn sees the introduction of emphasis on non-emphatic consonants in Arabic borrowings as a type of hypercorrection which is at least partly conscious, and labels it “hyper-Arabization” (Kahn 1976: 89–90). This contrasts with the situation in written Kurdish, where Arabic features such as emphasis are deliberately underrepresented. Barry (2019) prefers an analogical, phonetic explanation for the appearance of emphasis in inherited Northern Kurdish structures and, remarkably, shows a strong but imperfect correlation between the emergence of *ħ* and *ʕ* (which he reanalyzes as a pharyngealized vowel feature) and the presence of labials.

Taken together, these data show that emphatic consonants have been phonologized in Kurdish in various ways, and to various degrees. Such consonants are found in most Kurdish varieties, but in each case, they are relatively uncommon and appear to be restricted to the periphery of the grammar. Haig’s observations concerning the Central Anatolian dialects of Northern Kurdish appropriately sum up most scholars’ view of the situation for the language as a whole: “In a sense, the pharyngeals are extraneous to the basic phonology: they are restricted to individual lexical items, they play no part in morphology, their functional load is very limited, and there is considerable cross-speaker and cross-dialect variability in the extent of their presence” (2007: 167). However, Barry’s (forthcoming) findings leave open the possibility of a system in which, as in Kumzari, emphasis is becoming increasingly phonologized and, in fact, may have a connection to the pharyngeal segments traditionally analyzed as phonemes *ħ* and *ʕ*.

4.2 Zazaki

Little has been written on the status of emphatics in Zazaki, a (non-Southwestern) West Iranian language spoken in the north-western portion of the larger Kurdish language area but fundamentally distinct from Kurdish (Paul 2009: 545).

Some of the properties of emphasis in Zazaki appear to be the same as those of Kurdish, although this topic is not addressed in detail in the literature. In their overviews of Zazaki, Blau (1989a: 338) mentions a dental emphatic series, and Paul (1998: 3–6, 11, 2009: 547) presents in the consonant inventory several contrastive but infrequently occurring consonants which could be labelled emphatic: *ʈ*, *ʂ*, velarized *l*, the pharyngeal consonant *ħ* and, in some varieties, the pharyngeal consonant *ʕ*.

While examples of most of the emphatics are not listed along with the inventory, a number of Zazaki words distributed throughout Paul's studies contain the pharyngeal consonant *ħ* in particular. As is the case for this consonant in Kurdish, most of the words in which it is found are borrowings of Arabic items that contain *ħ*:

ħelāl 'fresh' (cf. Ar. *ħalāl*)

ħepis 'prison' (cf. Ar. *ħabs*)

ħeq 'expense' (cf. Ar. *ħaqq* 'right (n.), due (n.)')

In a few instances, however, Paul shows that Zazaki *ħ* is a reflex of Arabic *ʕ* (e. g., *bāħd* / *bāħdo* < Ar. *baʕd* 'after'), and for the numeral 'seven', in parallel with Kurdish, an inherited Iranian root *haft* 'seven' is rendered with an emphatic: *ħewt*.

The emphatic *ʕ*, in turn, is occasionally carried over from Arabic (*ʕerd* < Ar. *ʕard* 'ground'), and in a few cases replaces the glottal stop *ʔ* in words of Arabic origin (e. g., *ʕemel* < Ar. *ʕamal* 'hope'). The emphatic *ʂ* is found in one Iranian root (*ʕī* 'stone'), and *l* is attested in the item *līmīn* 'dirty'; and the emphatic *ʈ* is only found in words of Arabic origin, such as *ʈeyr* (> Ar. *tayr* 'bird'; Paul 1998: 6, 11).

As is the case in Kurdish – and perhaps even more so – emphatics are part of the periphery of Zazaki phonology, appearing in a small set of words and exercising a limited functional load.

4.3 Dezfuli-Shushtari

Dezfuli and Shushtari are closely related Southwestern Iranian varieties spoken by the inhabitants of Dezful and Shushtar, cities in northern Khuzestan Province, Iran. These cities are situated in a linguistically heterogeneous area where Arabic, Khuzi Persian, Southern Kurdish and two Luri languages (Northern Luri and Bakhtiari) are spoken alongside one another. Of these contact languages, the Luri languages are the closest relatives of Dezfuli-Shushtari (Windfuhr 2009: 13).

The consonant inventory of emphatics in Dezfuli-Shushtari differs from that found in Kurdish and Zazaki in that it lacks co-articulated emphatics (ʂ, ʈ, etc.). It does, however, contain both of the pharyngeal consonants found in Arabic, ʕ and ʁ. Most of the words in which pharyngeals occur here are identifiably borrowed from Arabic (Dezfuli-Shushtari data from MacKinnon 1974: 23–34):

<i>ḥamūm</i>	‘bath, shower’ (cf. Ar. <i>ḥammām</i>)
<i>ḥisāb</i>	‘calculation’ (= Ar.)
<i>maḥmūd</i>	‘(proper name)’ (= Ar.)
<i>ʕajīb</i>	‘strange’ (= Ar.)
<i>ʕarūs</i>	‘bride’ (= Ar.)
<i>mōzū(ʕ)</i>	‘topic, subject’ (cf. Ar. <i>mawḍūʕ</i>)

MacKinnon mentions that pharyngeals “occasionally also occur in words not of Arabic origin”, and cites the item ʕas ‘bone’ (1974: 12, 1995; Nirumand [1970]: 181 gives the form ʕass).⁶ Nirumand’s (1970) short lexicon of Shushtari reveals their distribution in a number of other words of non-Arabic origin. Some of these words appear to be lexical innovations, and a small number can be traced to introduction of ʕ into inherited Iranian vocabulary. Both types of words are shown in the following list (Shushtari data from Nirumand [Circa 1970]: 151–153, 181–183):

<i>ḥap</i>	‘place in the mouth’
<i>ḥapant</i>	‘incapable’
<i>ḥifārišt</i>	‘stench’
<i>ḥippa</i>	‘cumbersome’
<i>ʕasp</i>	‘horse’ (cf. MP/NP <i>asp</i>)
<i>ʕars</i>	‘teardrop’ (cf. MP <i>ars</i> , Bakhtiari <i>hars</i>)
<i>ʕūf</i>	‘weary’
<i>ʕuq</i>	‘vomit’

⁶ In fact, this word can be related to Middle Persian *ast* ‘bone’, as can Iranian reflexes such as New Persian *ostoxān*, Luri *has* / *hasaxūn* / *ustuxō* / *suxō* (Anonby 2003: 186–197) and the Kurdish variants *ʕasti* ~ *hasti* ~ [ʔ]asti (Kahn 1976: 45). However, it is also traceable to an ultimate Aramaic source ʔst (*ʔšsett) ‘mortar (tool used with a pestle), bone used as a mortar’ (see Comprehensive Aramaic Lexicon 2012). Importantly, Aramaic exhibits a strong diachronic connection between the glottal stop ʔ and ʕ (3.4), and in at least one historical Aramaic variety, there has been a relation of complementarity between the two sounds (Hoberman 1985: 225–227). These details do not in themselves account for the appearance of ʕ in ‘bone’, but they underscore the complexity of the historical context.

4.4 Keshmi

Keshmi (or Qeshmi) is a Southwestern Iranian variety spoken on Qeshm Island near Bandar Abbas, Iran, at the western opening of the Strait of Hormuz. It shares many traits with Bandari (coastal) Persian, but substantive documentation is lacking. Here, as in Dezfuli-Shushtari, the pharyngeal consonants *ħ* and *ʕ* are found, and co-articulated emphatics are absent from the consonant inventory. However, even the pharyngeal consonants are now being lost in the phonology of younger speakers and most often replaced by their glottal counterparts *h* and *ʔ* (or, in the case of *ʕ* at a word boundary, disappearing completely). In a 240-item wordlist collected from an older speaker of the Dargahān dialect of Keshmi in 2009, most words of Arabic origin retained the two pharyngeals (Anonby 2016). This is evident in the following list, in which Keshmi words are shown along with their Arabic source words and Persian cognates:

Ar. <i>ʕalaf</i> ‘feed (n.), hay’	> Keshmi <i>ʕalaf</i> ‘grass’
> New Persian (NP) <i>alaf</i> ‘grass’
Ar. <i>ʕankabūt</i> ‘spider’	> Keshmi <i>ʕankabut</i>
> NP <i>ankabūt</i>
Ar. <i>ʕadāb</i> ‘suffering’	> Keshmi <i>ʕazab</i>
> NP <i>azab</i>
Ar. <i>ʕayš</i> ‘life, livelihood’	> Keshmi <i>ʕayš</i> ‘marriage’
> Minabi (Bandari) Persian <i>hayš</i> ‘marriage’
Ar. <i>daʕwā</i> ‘dispute’	> Keshmi <i>daʕwā</i>
> NP <i>daʔwā</i>
Ar. <i>ṣaḥrā</i> ‘desert, plain’	> Keshmi <i>saḥrā</i>
> NP <i>sahrā</i>

In two cases, however, the pharyngeal is absent in words of Arabic origin even among older speakers:

Ar. <i>ʕaqrab</i> ‘scorpion’	> Keshmi <i>agrab</i>
> NP <i>aḡrab</i>
Ar. <i>ḥawā</i> ‘wind’	> Keshmi <i>hawā</i>
> NP <i>hawā</i> ‘air’

In addition, the pharyngeal *ʕ* appears in one Keshmi item for which, as for the word *ʕas(s)* ‘bone’ in Dezfuli-Shushtari (4.3), the determination of origin is complex. At first glance, the item *ʕawr* ‘cloud’ appears to be descended from Avestan *awra* or Middle Persian *abr* (note that the *b* ~ *w* correspondence has

been unstable in Persian, and a $b > w$ sound change is widespread in related Southwestern Iranian languages); if this explanation were correct, it would provide an example of a pharyngeal arising spontaneously in Keshmi. However, a similar word for ‘cloud’ *ṣayba* is found in Classical Syriac Aramaic (Nöldeke 1896 in Hoberman 1985: 225), and the Arabic items *ṣabra(t)* ‘teardrop’ and *ṣabar* ‘to cross, to pass over’ cannot be discounted as possible sources.

In sum, pharyngeal consonants are confined to the outer periphery of the phonological system in Keshmi, as is the case in Dezfūli-Shushtari.

4.5 Pashto

Of all the Iranian languages where pharyngeals have been documented, their distribution is most restricted in the East Iranian language Pashto. Parallel to the situation in Dezfūli-Shushtari and Keshmi, the pharyngeal consonants $ħ$ and $ʕ$ are retained in words of Arabic origin. Here, however, these consonants have an acrolectal status: they are confined to the formal speech of educated speakers of the language (Robson and Tegey 2009: 725). Examples of Pashto words in which these fricatives appear are *ḥabīb* ‘dear (n.)’ (= Ar.) and *ṣaqāl* ‘wisdom’ (< Ar. *ṣaql*) (724). As Agnes Korn (pers. comm. 2012) has noted, this situation likely mirrors learned pronunciation of Arabic loanwords within many other languages in the broader Islamic sphere.

4.6 Emphatic consonants in Kumzari

In the five Iranian languages described so far, it is evident that while emphasis has been incorporated in various ways and to various degrees as a result of contact with Arabic, it remains in each case a peripheral aspect of the phonology. In contrast, emphasis has been profoundly phonologized in Kumzari, both in relation to phonological inventory and to the behaviour of emphasis in the language (Section 5).

The Kumzari emphatic series is comprised of the following six consonants:

- $t̤$ voiceless uvularized alveolar stop
- $d̤$ voiced uvularized alveolar stop
- $ʃ̤$ voiceless uvularized alveolar fricative
- $ʒ̤$ voiced uvularized alveolar fricative
- $l̤$ voiced uvularized lateral alveolar approximant
- $ħ$ voiceless pharyngeal consonant (see 3.1)

In addition, three other consonants tend to trigger emphasis in other consonants diachronically, although they are not themselves contrastively emphatic:

- w* voiced labial-uvular approximant (see below)
x voiceless uvular fricative
q voiceless uvular stop

The identification of emphatic consonants in Kumzari is a delicate procedure because there is no single defining structural trait which accounts for all of the elements in the system. There is, however, a set of interrelated phonetic and phonological characteristics that are useful in describing the series. These characteristics are as follows:

1. With one exception, Kumzari emphatics have contrastive non-emphatic counterparts. Contrast between members of the two groups of consonants is evident from the following pairs of words:

<i>ṭ ṭēr</i>	‘bird’	<i>t tēra</i>	‘way, path’
<i>ḍ ḍayf</i>	‘guest’	<i>d dayn</i>	‘loan’
<i>ṣ ṣāl</i>	‘white fish sp.’	<i>s sāl</i>	‘year’
<i>ḷ waḷa</i>	‘or’	<i>l walm</i>	‘dispute’
<i>ḥ ḥējub</i>	‘eyebrow’	<i>h hē</i>	‘yes’

(Sound files with contrasting emphatic/plain consonants pairs, reproduced from Anonby 2011a, are found in the Appendix and audible in the online version of this article.)

To be thorough, the emphatic / non-emphatic contrast is weak (but still defensible) for two of these pairs. In the case of the lateral approximants *ḷ* and *l*, the emphatic member *ḷ* is poorly attested in the Kumzari lexicon, being found in only a small number of stems, including *waḷa* ‘or’, and *ʔaḷḷa* ‘God’ (along with some but not all of the compounds containing this stem; see Section 5.1 below). Concerning *ḥ* and *h*, the inverse is true: *ḥ* is extremely common, but its non-emphatic counterpart *h* is peripheral in the lexicon: it is found in a small set of words, most or all of which are interjections and/or borrowings from Arabic:

<i>bahlul</i>	‘fish sp.’	<i>hišt</i>	‘get! (to large animals)’
<i>dahr</i>	‘very long time’	<i>hud</i>	‘knock-knock!’
<i>ʔēhil</i>	‘relative’	<i>ʔillahī</i>	‘my God!’
<i>fahama</i>	‘understanding’	<i>ʔilmuhum</i>	‘the important thing’
<i>hall</i>	‘what?!; hey!’	<i>nahaba</i>	‘rob, steal from’
<i>hā / hō</i>	‘yes?’	<i>sihl</i>	‘easy thing’
<i>hē</i>	‘yes’	<i>šhōr u dhōr</i>	‘for month and years’
<i>hidī</i>	‘you see?!’	<i>wahaba</i>	‘yawning’

In the case of *z*, however, a plain counterpart is completely absent from the consonant inventory as a result of a surprising sound change (cf. Table 1 in Section 2; see also Section 5.4). Earlier accounts of the alveolar fricative in Kumzari make reference to a plain *z* (Jayakar 1902; Thomas 1930), but Anonby (2011a: 376) describes it as emphatic, and this assertion will be developed throughout the remainder of the article.

2. Uvularization is the main articulatory basis for emphasis in Kumzari. The alveolar emphatics *ɖ ɗ ʒ ʐ* exhibit strong, simultaneous posterior secondary articulation, with uvularization dominating but bounded by a unified stricture all the way from the pharynx up to the velum. Phonetically, this complex articulation can be symbolized as [tʰ dʰ sʰ zʰ] and [ɭʰ].⁷ The remaining member of the Kumzari emphatic series, however, is a pharyngeal consonant *ħ* [ħ]. Although *ħ* is not uvularized, its behaviour suggests that it should be classed as an emphatic (Sections 5.3.1–5.3.2). The status of the retroflex alveolar approximant *r* is problematic for the opposite reason: while it does *not* pattern as an emphatic consonant by inducing emphasis on other consonants (Section 5.3), it is noticeably uvularized ([ɭʰ]) by most speakers in its most frequently attested positions, especially in word-initial and word-final positions:

rū ‘face’

rōr ‘child’ (recorded example #12 in the Appendix)

wīr ‘tuna sp.’

This compares with the non-uvularized realization of *r* in other positions (Anonby 2011a: 377):

- as an alveolar flap [r] in *drāz* ‘length’
- as a retroflex alveolar flap [ɻ] in *brišt* ‘cooked’
- as an alveolar trill [r] in *qarraṣ* ‘mosquito’

⁷ The symbol [ɭʰ] (a superscript inverted capital r) has been proposed by some scholars (e.g., Dolgopolsky 1977) to mark uvularization, but because pharyngealization and uvularization are not known to contrast in any language, this symbol has not been made part of the IPA alphabet (IPA 2015). Since emphasis is most often a bundle of articulatory features (3.2), any of the existing phonetic symbolizations is in fact inadequate. In order to underscore articulatory commonalities among languages where emphasis has been described using the phonetic symbol for pharyngealization [ʕ], even when uvularization is a central or partial feature that makes contributes to emphasis, the phonetic symbol [ɭʰ] is retained here for emphasis in Kumzari.

Still, as these data illustrate, and in opposition to the contrastively uvularized emphatic series outlined in the first point, uvularization of *r* is entirely allophonic.

The uvularized alveolar emphatics, uvular consonants *x q ġ*, the pharyngeal *ħ*, and the uvularized allophone of *r* all cause preceding as well following non-back vowels to be retracted (*ā* [a:] → [ɑ:], *a* [e] → [ʌ]). In the case of non-low vowels, they cause lowering in the transition between the vowel and consonant (*i* [i:] → [i̠ə] before a consonant, [əi̠] after a consonant; *ē* [e:] → [e̠ə] before a consonant, [əe̠] after a consonant).

3. As stated above, *w*, *x* and *q* are three consonants in the Kumzari inventory which, although they are not themselves contrastively emphatic, have diachronically triggered emphasis in the non-emphatic counterparts of the other members of the series (5.3). This behaviour supports their designation, noted for other languages with *x* and *q*, as opaque emphatics (3.4). There are no cases where the voiced uvular fricative *ġ* has been observed to trigger emphasis on neighbouring consonants, but as *ġ* is not abundant in the lexicon, this could reflect an accidental gap.

Notably, the approximant *w* is labial-uvular, i. e., produced by a simultaneous bilabial and uvular double articulation, rather than labial-velar. In addition to inducing emphasis, the approximant *w* – along with *x*, *q* and *ġ* (but not velar consonants *k* and *g*) – causes retraction of non-back vowels (especially *ā*; cf. recording 6 in the Appendix) and centralization of non-low front vowels.

A summary of the emphasis-related properties of consonants in Kumzari, as outlined in the preceding discussion, are summarized in Table 2:

Table 2: Emphasis-related properties of Kumzari consonants.

	<i>ṭ ḍ ṣ ʕ</i>	<i>z</i>	<i>ħ</i>	<i>r</i>	<i>w</i>	<i>x q</i>	<i>ġ</i>	other consonants
primary uvular articulation						✓	✓	
secondary uvular articulation	✓	✓		✓	✓			
contrast with non-emphatic counterpart	✓		✓					(n/a)
cause retraction of back vowels	✓	✓	✓	✓	✓	✓	✓	
trigger emphasis on other consonants	✓	✓	✓	?	✓	✓	?	

This complex situation underlines the reality that place and/or manner of articulation are not in themselves sufficient as predictors of which consonants participate in the language’s system of emphasis.

It is true that, like Kumzari, some Kurdish (and possibly Zazaki) varieties exhibit a highly differentiated inventory of emphatic consonants (Sections 4.1 and 4.2). However, similarities in the degree of phonologization end here.

First of all, the sheer number of emphatics in the Kumzari lexicon (Anonby and van der Wal Anonby in prep.) is remarkable: in a database of 4400 items, there are about 1200 instances of emphatic consonants distributed among almost a quarter (1086) of the words, totalling 5.3% of all phoneme occurrences in the lexicon. While the measures are not identical, this figure compares favourably to the proportion of emphatic consonants in Arabic, where in a random text sample the co-articulated emphatics *ṭ ṣ ḍ ḏ* together show a frequency of only 2.3% of individual phoneme occurrences (al-Xūli 1984 in Bakallah 2009: 423). In reality, such a comparison suggests that the proportion of emphatic consonants may be equivalent or higher in Kumzari than in Arabic.

Second, whereas the overwhelming majority of emphatics in other Iranian languages are limited to Arabic loanwords, there are hundreds of non-borrowed Kumzari words – including both identifiably inherited items as well as apparent Kumzari-specific innovations – in which emphatic consonants have arisen (Section 5.2).

Third, while emphasis shows little inclination toward diffusion in the lexicon and phonology of the related languages examined above, there are numerous examples of diachronic and synchronic spread of emphasis in Kumzari, whether from emphatic consonants themselves or from other emphasis-inducing consonants (Sections 5.3–5.5). In fact, it is this firmly established phonological tendency that has led to the situation described in the previous two points, in which emphasis is not only well-attested in borrowed words, but has permeated the lexicon as a whole.

Now that the stage has been set with a discussion of the typology of emphasis (Section 3) and a review of its status in Iranian languages in contact with Arabic (Section 4), the following section (Section 5) will explore the question of how emphasis has arisen and, over time, become so profoundly phonologized in Kumzari. In short: what are the reasons for the contrasting outcomes of Kumzari versus related languages in similar situations, where emphasis has remained on the phonological periphery?

5 The phonologization of emphasis in Kumzari

A necessarily central theme in a study of language contact is the recognition that the linguistic effects of contact are intertwined with the social context. For Kumzari

in particular, the mechanics of such a relationship are indispensable in accounting for the phonologization of emphasis and its ultimate penetration of the lexicon.

However it became established on the Musandam Peninsula of north-eastern Arabia (see van der Wal Anonby 2015), there is evidence that Kumzari has persisted as a linguistic island in a sea of Arabic for more than 1300 years (see Section 5.1, “Rendering of Arabic consonants and time depth of lexification”). In contrast to the Iranian languages discussed above, which have been influenced by numerous languages, Arabic is the only language with which Kumzari is (and, in the recent history of the language, has been) in significant contact. Seasonal isolation of the main Kumzari-speaking village from Arabic-speaking communities and the historical dominance of Kumzari clans in the region are two of the key factors that have allowed the language to survive until now, but there have long been strong connections between Kumzari speakers and speakers of Arabic: as one of the groups which make up a regional Shihhi Arab ethnic confederation, Kumzari speakers identify as Arabs, and for part of each year, they live in large towns (Khasab and, in smaller numbers, Daba) primarily populated by Arabic speakers (Anonby and Yousefian 2011: 32–33). Being greatly outnumbered by speakers of Arabic, there must have been a long-standing history of bilingualism, and this has made its mark on the Kumzari language (for the significance of each of these sociolinguistic factors in a contact situation, see Aikhenvald 2006: 36–45; Matras 2009: 45–47, 222–225).

The current section (Section 5), which forms the core of the present study, traces the appearance and proliferation of emphasis in Kumzari through five processes. A sequence of phonological developments has been set in motion by the initial and ongoing process of extensive borrowing (cf. Aikhenvald 2006: 21) from Arabic and the resulting implantation of emphasis in the language. In response to the destabilization brought about by the introduction of a new phonological phenomenon, several additional processes have arisen by which the lexicon has regained a measure of homogeneity. New words have appeared in the language, and some of these contain emphatic consonants (Section 5.2). Three types of emphasis spread, of which two are diachronic and one is synchronic, have applied to borrowed words as well as inherited vocabulary. First of all, emphasis has recurrently spread in the lexicon from emphatic or emphasis-inducing consonants to potentially emphatic consonants through analogical sound change (Section 5.3). Secondly, an across-the-board sound change in which *z* has been invariably recast as an emphatic *ẓ* has resulted in the appearance of emphasis in hundreds of additional words (Section 5.4). Finally, in contexts where two consonants come together at a morpheme boundary, emphasis alternates synchronically by spreading onto potentially emphatic consonants (Section 5.5).

Each of these processes shows a clear preference for the amplification of emphasis, and its spread through the language as a whole. In contrast – and despite the fact that even in Arabic, emphasis is diachronically unstable (Sections 3.4 and 3.5; Fischer and Jastrow 1980: 57) – there are no cases in the data in which emphasis is lost from a word where the Arabic cognate contains an emphatic consonant (this generalization does not include ʕ, which is similarly absent in the neighbouring Shihhi dialects of Arabic; see Sections 2 and 5.1).

The role of social context in a systemic bias in favour of emphasis should not be underestimated. In line with their Arab identity, Kumzari attitudes toward Arabic are positive (Anonby and Yousefian 2011: 32–38), and emphasis is recognized as a prototypically Arabic phenomenon. Throughout a year of field research on Kumzari, speakers consistently drew my attention to Arabic features of their language, and on several occasions pointed out that Kumzari, like Arabic, is a *luġat aḍ-ḍād* (Ar., lit. ‘ḍ-language’) – distinct from other languages in the world by virtue of having the typologically rare consonant ḍ as part of its phoneme inventory (Aikhenvald 2006: 41 refers to such features as “emblematic”). So the fact that the phonological processes at work in Kumzari innovations may be in many other respects unconscious does not detract from the significance of the social context. This idea will be explored further in the discussion of the *z > ʒ* sound change (Section 5.4), where the resulting configuration is exceptionally weighted toward emphasis.

Kumzari and its Arabic contact varieties

Up to this point, I have made general reference to Arabic as the language with which Kumzari is in contact, and Shihhi (*šihhi*) Arabic (ShA), spoken by the Shihuh (*šihuh*) Arabs of the Musandam Peninsula region, as the specific contact variety. Several complexities related to a study of emphasis in this contact situation should be mentioned, however.

First: ShA is the primary Arabic variety in contact with Kumzari, but it is not the only one. Increasingly, because Kumzari speakers now travel regularly outside of the Musandam region, and because people from elsewhere are frequent visitors there, Kumzari speakers are exposed to non-local Gulf Arabic dialects – both those spoken in the United Arab Emirates and, to a lesser degree, those spoken elsewhere in Oman. At the same time, due to the penetration of media and formal education in the region, Modern Standard Arabic is also exerting a direct influence on Kumzari.

Second, while ShA is the prevailing Arabic variety throughout the Musandam Peninsula today, this may not have always been the case. According to their oral

traditions, Shihuh Arabs identify Yemen as their original homeland, and record their entrance into the Musandam region as taking place in about the second Century A.D. (Dostal 1972: 2); others have proposed a migration as late as the seventh Century (Zimmermann 1981). This range of dates overlaps with the period in time when ancestors of the Kumzari community likely appeared in the area (see the related discussion in 5.1), and it is not known which group became established there first (Zimmermann 1981; Najmabadi 1988: 67–8). The question therefore arises as to whether ShA was the only early variety in close contact with Kumzari, or even the initial contact variety. I will not attempt to resolve this question here, but the possibility of significant complexity in the history of the contact situation must be acknowledged (see van der Wal Anonby 2014, van der Wal Anonby 2015 for a detailed discussion).

Third, as is often the case with the Arabic “dialects” described in the literature, ShA itself is not a single dialect, but a convenient regional grouping of many distinct varieties. Broadly speaking, ShA can be included under the geographic and linguistic banner of Gulf Arabic, with which it shares many basic structures including the historical fronting of stops, the backing and rounding of Old Arabic *ā*, and an inclination for syllable-initial consonant clusters (cf. Bernabela 2011; Holes 2004: 73–77). However, as is evident in a comparison of studies on several varieties (Jayakar 1902; Bernabela 2011; Anonby in prep. b), ShA is internally heterogeneous; and this observation does not take into account the inland varieties, some of which are reportedly unintelligible to speakers from coastal communities. Thus, the question arises of which ShA sub-dialect (or -dialects) should be treated as the main contact variety. Where there are divergences among ShA varieties, I have based my observations on the dialects of 1) Qabbē, a ShA-speaking village immediately to the north of Kumzar village, and 2) Khasab, a large ShA -speaking town where (as mentioned at the beginning of this section) most Kumzari speakers spend part of each year.

Finally, numerous detailed studies exist which treat emphasis in other Arabic dialects (especially those of Cairo and Palestine; see 3.3 above), but there are none devoted to ShA dialects, or even the dialects of the Gulf in general. The accounts of ShA mentioned in the previous paragraph, while not dealing specifically with emphasis, nonetheless serve as a local comparative framework for the description of emphasis in Kumzari, and these have been supplemented with firsthand research on ShA. In practice, most of the contact-related issues under consideration can in fact be resolved by referring to Modern Standard Arabic forms, but ShA forms are invoked when they are closer to the Kumzari forms under investigation.

5.1 Retention of emphatic consonants in words adopted from Arabic

Like the Arabic system on which it is based, the Kumzari emphatic series is robust. Arabic emphatics *ṭ ɖ ɗ ʂ ʕ* and *ʕ* are all consistently maintained as emphatics in Kumzari words borrowed from Arabic. The glottal stop is, however, found in place of the pharyngeal consonant *ʕ* (*ʕayn*), as in the neighbouring Shihhi dialects of Arabic with which Kumzari is in constant contact. The following paragraphs illustrate the ways in which these consonants have been retained and, in some cases, altered. As mentioned in the previous paragraph, Arabic forms are represented by Modern Standard Arabic unless stated otherwise.

ṭ, ʂ, ʕ and ʕ

In the case of *ṭ ʂ ʕ* and *ʕ*, the primary place of articulation of the Arabic source consonant has been preserved.

<i>ṭ</i> > <i>ṭ</i> (Arabic > Kumzari)	<i>ʂ</i> > <i>ʂ</i>
<i>fuṭām</i> > <i>fuṭman</i> ‘weaning’	<i>ʔaʂl</i> > <i>ʔaʂl</i> ‘origin’
<i>ṭabla(t)</i> > <i>ṭōbil</i> ‘large drum’	<i>muqaʂʂ</i> > <i>mqaʂʂ</i> ‘scissors’
<i>ṭalāq</i> > <i>ṭālaq</i> ‘divorce’	<i>ʂadqa(t)</i> > <i>ʂadqit</i> ‘charity’
<i>ṭayr</i> > <i>ṭēr</i> ‘bird’	<i>ʂuqr</i> > <i>ʂuqr</i> ‘falcon’
<i>xarṭūm</i> > <i>xarṭum</i> ‘peninsula’	<i>ʂūra(t)</i> > <i>ʂūrit</i> ‘picture’
<i>ʕ</i> > <i>ʕ</i>	
<i>ḥadd</i> > <i>ḥadd</i> ‘tip, end’	
<i>ḥadīθ</i> > <i>ḥēdis</i> ‘Hadith (traditions)’	
<i>ḥijāb</i> > <i>ḥējub</i> ‘eyebrow’	
<i>jināḥ</i> > <i>jnāḥ</i> ‘wing’	
<i>maḥār</i> > <i>maḥḥar</i> ‘oyster’	

There are only a small number of Kumzari stems with *ʕ*. Examples in the data include *waʕa* ‘or’ (see Section 5.3.2); *ʂālaṭ* ‘caught hiding’ (cf. Ar. *sallaṭ*); and *ʔalla* ‘God’, where *ʕ* is retained from Arabic (*ʔallāh*). It appears in many expressions that use this root, however:

<i>allāh yihafḏak</i> > <i>ʔalla ḥāfaṭ tu ka</i> ‘may God preserve you’
<i>ʔinšāʔ allāh</i> > <i>ʔinšalla</i> ‘God-willing’
<i>yallāh bilʕafya</i> > <i>yaʕla ba ʔifit</i> ‘hopefully’
<i>waʕlāh</i> > <i>waʕla</i> ‘truly, I swear’
<i>ʕafa llāh</i> > <i>ʔafaʕla</i> ‘bless you!’

ʕ (ʕayn)

The pharyngeal consonant ʕ is the only member of the Arabic emphatic series which is not retained in words borrowed into Kumzari. Instead, it is consistently rendered as a glottal stop ʔ.

ʕ > ʔ	
ʕama(t) > ʔāmit	‘paternal aunt’
ʕīša(t) > ʔīšīʔit	‘life’
jamāʕa(t) > jāmāʔit	‘group’
šaʕb > šaʔb	‘(the) public’
rubʕ > rubʔ	‘quarter’

An important consideration which helps account for the retention of all Arabic emphatics in Kumzari except ʕ is that Shihhi Arabic, a cluster of Arabic dialects which surround Kumzari (see beginning of Section 5), also lacks ʕ in its consonant inventory and has similarly replaced it with the glottal stop (Bernabela 2011: 26).

Rendering of Arabic consonants and time depth of lexification

Although the present article is restricted in scope to the patterning of emphatic consonants in Kumzari, even the limited data here provide a window into the time depth of the language’s initial lexification (large-scale borrowing of vocabulary), from Arabic and the resulting adoption of emphatics in the phonology.

Almost all Iranian languages were lexified by Classical Arabic (as a written language) and Mesopotamian Arabic (as a spoken language) via Persian in the period following the Islamic conquests of south-western Asia beginning in the seventh Century A.D. (Sadeghi 1986; Windfuhr 2009: 419), but Kumzari was not affected by this event. Rather, it has been separately lexified by Peninsular Arabic at some point before this – in other words, more than 1300 years ago. There are almost no New Persian loans into Kumzari, and whereas Persian (and, through it, almost all other Iranian languages) has dispensed with emphatic consonants in Arabic loanwords, Kumzari has consistently maintained them. There are also consistent differences in the primary place of articulation of the emphatic pair *ḍ/ḍ̣*: whereas the Iranian reflex *z* mirrors the fricative articulation *ḍ* in many northern Mesopotamian Arabic dialects (Jastrow 1980: 142; Sadeghi 1986: 29–30), the Kumzari reflex is consistently a stop (*t̪* or *ḍ̣*, depending on the date of the borrowing; see in this section above), in line with the “City” type dialects of the Arabian Peninsula (Section 5.4.2; Holes 2004: 71). This correspondence is shown in Table 3.

Table 3: Kumzari vs. Persian reflexes of Arabic *ḡ/ḡ*.

	Classical Arabic	“City” type Arabic	Kumzari	New Persian	Northern Mesopotamian Arabic
‘agreeing’	<i>rāḡī</i>	<i>rāḡī</i>	<i>rāṭī</i>	<i>rāzī</i>	<i>rāzī</i>
‘Ramadan’	<i>ramaḡān</i>	<i>ramaḡān</i>	<i>rāmaṭan</i>	<i>ramazān</i>	<i>ramazān</i>
‘weak’	<i>ḡaṣīf</i>	<i>ḡaṣīf</i>	<i>ṭṛīf</i>	<i>zaṛīf</i>	<i>ḡaṣīf</i>
‘preserver’	<i>ḡāfiḡ</i>	<i>ḡāfiḡ</i>	<i>ḡāfaṭ</i>	<i>ḡāfiz</i>	<i>ḡāfiz</i>
‘governor’	<i>muḡāfiḡ</i>	<i>muḡāfiḡ</i>	<i>mḡāfiḡ</i>	<i>muḡāfiz^a</i>	<i>muḡāfiz</i>
‘tyrant’	<i>ḡālam</i>	<i>ḡālam</i>	<i>ḡālam</i>	<i>zālim</i>	<i>zālam</i>

^aLiterally, ‘preservative’ (note the semantic shift).

Additional patterns of Kumzari vs. Iranian borrowing and their implications for the linguistic history of Kumzari are developed in greater detail in Anonby (in prep. a).

5.2 Propagation of emphatic consonants through individual innovations

Once emphasis has entered a language through retention of emphatic consonants in loanwords, it may spread into new contexts (Matras 2009: 222–224). In Kumzari, emphasis sometimes appears in individual words, and there is no discernible pattern. In other cases, spread is achieved through systematic changes. Both types of innovations highlight the degree to which emphatic consonants have been incorporated into the core of the phonological system (see analogous examples of loan phonology integration in Aikhenvald 2006: 21, 23; Windfuhr 1990: 543–544). While the following sections will address systematic patterns of spread, the present section will first account for spontaneous appearances of emphasis through lexical innovation (Section 5.2.1) and, in a handful of cases, on potentially emphatic consonants in existing vocabulary (Section 5.2.2).

5.2.1 Lexical innovations

There are numerous Kumzari words with emphatic consonants which cannot be ascribed to inheritance or borrowing: they appear to have arisen from within the language. In Kumzari lexical innovations, *ṭ*, *ṣ* and *ḡ* are not uncommon. However, *ḡ* is rare and *ṭ* is absent. A sample of apparent innovations containing emphatic consonants is shown in Table 4:

Table 4: Apparent lexical innovations in Kumzari containing emphatic consonants.

ṭ		ḥ	
<i>ṛaṭṛaṭa</i>	‘shivering’	<i>čaḥḥa</i>	‘pouring’
<i>ḡāṭaf</i>	‘tree sp. used to make rope’	<i>daḥya</i>	‘sliding’
<i>lāṭī</i>	‘low’	<i>ḥaqm</i>	‘pigeon’
<i>maṇṭa</i>	‘sailfish, marlin’	<i>ḥayl</i>	‘haunches’
<i>n’āṭ</i>	‘strength, resolve’	<i>ḥēlē</i>	‘rust’
<i>nuṭṭēbubō</i>	‘value’	<i>ḥēmqu</i>	‘disputing’
<i>qāpṭ</i>	‘white fish sp.’	<i>ḥwēl</i>	‘chattering’
<i>ṭal’it</i>	‘waiting for fish’	<i>jurbah</i>	‘silver fish sp.’
<i>ṭāma</i>	‘solemn charge, commission’	<i>sīḥak</i>	‘guitarfish’
<i>ṭyāḥ</i>	‘sardine fishing season’	<i>šūḥū</i>	‘sliding’
<i>xatṭar</i>	‘guest’	<i>zīrrah</i>	‘flying insect sp.’
ṣ		ḍ	
<i>ḡaṣṣ</i>	‘peg’	<i>marḍuf</i>	‘type of bread’
<i>ḥampōlaṣ</i>	‘butterflyfish’	<i>qaḍḍuḥ</i>	‘date sp.’
<i>qīṣṣē</i>	‘venomous snake sp.’	<i>qanḍaḥa</i>	‘rainbow’
<i>rsiḡ</i>	‘wrist’		
<i>ṣaḥnē</i>	‘crushed sardines’		
<i>ṣābuṭ</i>	‘jellyfish’		
<i>ṣām</i>	‘handle’		
<i>ṣāmur</i>	‘small fishing weight’		
<i>ṣanduḥ</i>	‘forehead’		
<i>ṣērah</i>	‘animal’s milk’		
<i>šnāṣ</i>	‘jaw’		

5.2.2 Phonological innovations in existing words

In most cases, the appearance of emphasis on non-emphatic consonants in existing Kumzari words is an effect of the spread of the emphatic feature from existing emphatic or emphasis-inducing consonants (Section 5.3), or is the result of a systematic sound change (Section 5.4). However, there is a handful of words in which emphasis seems to have arisen spontaneously on historically non-emphatic consonants; here, its appearance cannot be attributed to the immediate phonological context or to changes in the phonological system as a whole. These words are as follows:

<u>source</u> >	<u>Kumzari</u>	
MP <i>āhan</i> >	<i>ḥan</i>	‘iron’
MP <i>pōst</i> >	(<i>pōst</i> /) ⁹ <i>pōṣṭ</i>	‘skin’
Ar. <i>samar</i> >	<i>ṣumr</i>	‘acacia sp.’
Ar. <i>sarīr</i> >	(<i>sērīr</i> /) <i>ṣērīr</i>	‘bed’

5.3 Diffusion of emphasis through analogical sound change

An additional outcome made possible by the initial phonologization of emphasis in Kumzari is the recurring diffusion of emphasis onto potentially emphatic consonants (Section 3.4) within a word. This diachronic process operates by means of “lexical diffusion” of analogical sound change (Chen and Wang 1975; Kiparsky 1995): there are specific phonological contexts within which emphasis is likely to spread, but it is not automatic; rather, it diffuses on a word-by-word basis. Changes have taken place in a significant and apparently growing number of words in the language, and apply equally to inherited vocabulary and Arabic loanwords.

The first half of this section explores the phonological conditions and mechanisms for the analogical diffusion of emphasis (Section 5.3.1). In the second half, examples of diffusion are catalogued for each of the specific sound changes brought about by this process (Section 5.3.2).

5.3.1 Conditions and mechanisms for diffusion

In almost all cases (cf. Section 5.2.2), there is a discernible phonological motivation (cf. Kiparsky 1995: 642, 653) for the spread of emphasis from one consonant to another within existing Kumzari. Sufficient conditions for diffusion are 1) the presence of an emphatic or emphasis-inducing consonant (see Section 4.6 for an explanation of these terms), and 2) the presence of a potentially emphatic consonant (i. e., a plain consonant which has an emphatic counterpart in the language).

Examples of words in which emphasis has diffused from emphatic consonants to potentially emphatic consonants are as follows:

<u>source</u> >	<u>Kumzari</u>	
Ar. <i>qaṣd</i> >	<i>qaṣḍ</i>	‘intention’
Ar. <i>sāḥir</i> ‘magician’ >	<i>ṣāḥar</i>	‘sorcerer’

⁹ For the Kumzari items with additional variants given in parentheses, only some speakers use the emphatic forms.

MP <i>tazāg</i> ‘flowing’ (cf. NP <i>tāza</i> ‘fresh’)	> (* <i>tāzaḡ</i> >) <i>ṭāzaḡ</i> ‘fresh’ ¹⁰
MP <i>tēz</i> >	(* <i>tēz</i> >) <i>ṭēz</i> ‘sharp’
MP <i>yāzdah</i> >	(* <i>yāzda</i> >) <i>yāzḍa</i> ‘eleven’

Examples of words in which emphasis has diffused from emphasis-inducing consonants *w*, *x* and *q* to potentially emphatic consonants are as follows:

source >	Kumzari	
Ar. <i>wa lā</i> >	<i>waḷa</i>	‘or’ (see 5.3.2)
MP <i>sabz</i> >	(<i>sawz</i> >) <i>ṣawz</i>	‘green’
Ar. <i>suʔāl</i> >	(<i>swāl</i> >) <i>ṣwāl</i>	‘question’
Ar. <i>faxḍ</i> (cf. Shihhi Ar. <i>faxt</i>) >	<i>faxṭ</i> / <i>fōxiṭ</i>	‘thigh’
Ar. <i>wasax</i> >	<i>wāṣax</i>	‘filth’
MP <i>suxr</i> (cf. also NP <i>sorx</i>) >	<i>ṣirx</i>	‘red’
Ar. <i>qahwa(t)</i> >	<i>qaḥwē</i>	‘coffee’
Ar. <i>qāmūs</i> >	<i>qāmuṣ</i>	‘dictionary’
Ar. <i>qyās</i> >	<i>qyāṣ</i>	‘measurement’

The fact that uvular consonants *w*, *x* and *q* have aligned with secondary uvularization (Section 4.6), so that they induce the diachronic spread of emphasis, has resulted in a greater impact of emphasis on the lexicon than would be the case if only the co-articulated alveolar set were involved.

The phonological motivation for the spread of emphasis through analogical sound change fits well within Ohala’s (1993a, 1993b) explanations of sound change in general. Rather than appealing exclusively to the spread of a particular phonological feature such as [guttural] or [RTR], the latter of which is typically limited to emphatics with secondary articulations in synchronic accounts of emphasis (Section 3.1), it is helpful to locate a motivation for sound change in the strong co-articulatory phonetic effects – in the double sense of overlapping gestures as well as secondary articulations (see Ohala 1993a) – shared by all of the emphatic and emphasis-inducing consonants: in particular, there is a common lowering and backing of adjacent vowels, and a corresponding acoustic drop in the frequency of the second formant (Section 3.4). According to Ohala, this type of phonetic variation becomes phonologized through hypo-correction, that is, when listeners fail to correct for allophonic

¹⁰ For an account of possible reinforcement from the Arabic cognate *ṭāzaj* (which is itself borrowed from Middle Persian), see the discussion of words with *ṭ* in the following section. For an account of the prior *z* > *ṣ* shift that has fed the diffusion of emphasis elsewhere in the last three words in this set, see 5.4.

perturbations in the speech signal. He speculates that “the farther away the environment is from the conditioned change ... the more difficult it will be for the listener to be able to establish the causal link between the two and use this link as the basis for correction” (1993b: 246–247). This is relevant for the diffusion of emphasis in Kumzari which, as evident in the data here and taken up below, often spreads to non-adjacent consonants. In sum, the strong incidental phonetic effects that accompany emphatic and emphasis-inducing consonants come to be perceived as phonological emphasis, and the underlying form of neighbouring potentially emphatic consonants is reinterpreted as that of the emphatic counterpart.

The data given above demonstrate that the analogical spread of emphasis is not sensitive to a word’s origin; provided with sufficient phonological conditions, it can apply equally to inherited vocabulary and Arabic loanwords. Any Kumzari lexical innovations not already permeated by emphasis would also be susceptible to further diffusion, although without a record of how Kumzari-specific words were pronounced at an earlier point in history, this cannot be established.

Parameters for diffusion in Kumzari are also apparent from the data. First, as is evident from sound changes in words like *faxt* > *faxt̤* ‘thigh’ and *suxr* > *šixr* ‘red’, the spread of emphasis can be either progressive (left-to-right) or regressive (right-to-left) within a word. In the case of a word like *wasax* > *wāšax* ‘filth’, where emphasis has diffused to a potentially emphatic consonant *s* between two emphasis-inducing consonants *w* and *x*, the ultimate source could be either of the neighbouring consonants. In any case, the presence of emphasis-inducing consonants on both sides of *s* furnishes optimal conditions for diffusion.

Second, as these words also show, emphasis can diffuse to potentially emphatic consonants which are adjacent (for example, *faxt* > *faxt̤* ‘thigh’), but it can also diffuse to potentially emphatic consonants which are separated from the source of emphasis by intervening vowels or transparent consonants (that is, consonants which cannot be contrastively emphasized but which do not block the spread or emphasis, e.g., *qyās* > *qyāš* ‘measurement’; cf. Section 3.3). Although, as shown above, it is induced by a more general set of consonants, this selective, long-distance assimilation of diffusion evokes the autosegmental nature of emphasis as a phonological feature.

Third, emphasis takes place within words (including compound words; see the examples of specific sound changes below) but not across word boundaries. Further data on the synchronic spread of emphasis (Section 5.5) confirm that the domain within which the autosegmental emphatic feature operates is the phonological word.

Finally, the fact that the diffusion of emphasis is not automatic or exceptionless should be underlined. Rather, it is an analogical sound change that is promoted by the specific phonological conditions outlined above, but which takes place among potential candidates one word at a time. This is evident from a comparison of words in which emphatic diffusion has taken place with phonologically similar words in which it has not transpired:

words affected by emphatic diffusion

Ar. *sāḥir* ‘magician’ > K *ṣāḥar* ‘sorcerer’

Ar. *suʔāl* > K (*swāl* >) *ṣwāl* ‘question’

MP *suxr* > K *ṣirx* ‘red’

MP *tēz* > K (**tēz* >) *ṭēz* ‘sharp’

Ar. *wa* ‘and’ + *lā* ‘not’ > K *waḷa* ‘or’

similar words which have not been affected

cf. Ar. *ʔiṣḥabba* ‘love (v.)’ > K *ʔiṣḥabba* ‘loving’

cf. MP *sabuk* > K *swak* ‘light (weight)’

cf. MP *saxt* ‘strong’ > K *saxt* ‘thick’

cf. Ar. *taqwīm* > K *taqwim* ‘calendar’

cf. Ar. *walī* > K *wālī* ‘chief’

5.3.2 Specific sound changes brought about by the diffusion of emphasis

Each of the five potentially emphatic consonants in Kumzari (*t*, *d*, *s*, *l* and *h*) (Section 4.6) has undergone sound changes in which it has acquired emphasis through diffusion within words. In the following discussion, examples of these individual sound changes are brought together.

t > *ṭ*

The diffusion of emphasis onto *t*, resulting in emphatic *ṭ*, has occurred in a small number of Kumzari words. It appears to have taken place in the following inherited items:

SWIr: MP >

taxl >

tazāg ‘flowing’ (cf. NP *tāza* ‘fresh’) >

tēz >

Kumzari

(**taḥl* >) *ṭaḥl* ‘bitter’

ṭāzāg ‘fresh’

(**tēz* >) *ṭēz* ‘sharp’

In each of these cases, it is necessary to posit intermediate stages (such as a prior *z* > *ḡ* sound change; see Section 5.4) to account for the initial appearance of emphasis in the word. In the case of *ṭāzāg* ‘fresh’, it should also be pointed out that emphasis on the initial *ṭ* of the Arabic cognate *ṭāzaj* ‘fresh’, itself a borrowing from Middle Persian, signals the possibility of an alternative or additional historical source in accounting for emphasis on this word in Kumzari; in many Arabic dialects, word-initial *t* followed by *ā* has become emphatic *ṭ* (see Fischer and Jastrow 1980: 56–57).

Arabic loanwords in which this change has transpired are also sparsely represented, but the following examples suggest that it has in fact taken place:

Arabic > Kumzari

ḥūt > (+ Kumzari *šaw* ‘night’) *šawḥaṭ* ‘whale’ (cf. Bayshak 2002)
faxḏ (cf. Arabic: Shihhi *faxt*) > *faxṭ* / *fōxiṭ* ‘thigh’

d > ḏ

As is the case with the diffusion of emphasis onto *t*, there are only a few cases in Kumzari in which *d* has become an emphatic *ḏ*. In inherited vocabulary, emphasis has transformed *d* into *ḏ* in six words, all of which are compound numbers between 10 and 20. In four of these words, a morpheme-final *ṣ* has transmitted its emphatic feature to the adjacent *d* at the beginning of the following morpheme *daʔ* / *-da* ‘ten’.

SWIr: MP > Kumzari

yāzdah > (**yāzda* >) *yāzḏa* ‘eleven’
dwāzdah > (**dwāzda* >) *dwāzḏa* ‘twelve’
sēzdah > (**sēzda* >) *sēzḏa* ‘thirteen’
nōzdah > (**nōzda* >) *nōzḏa* ‘nineteen’

The idea that emphasis in these words has originated with *ṣ* (as a result of a general sound change; see 5.4 below) and not *ḏ* is supported by a comparison of these numbers with the Kumzari numeral ‘ten’ in its independent form *daʔ* as well as its form in the analogous compound numbers *čārda* ‘fourteen’, *ʔafda* ‘seventeen’ and *ʔayda* ‘eighteen’.

Kumzari phonology resists three consecutive consonants within a word; so in the two remaining numbers between 10 and 20, the historical *ṣ* whereby emphasis has been introduced into the word has itself been deleted, but the emphasis that it has brought about remains on the *ḏ*.

SWIr: MP >

Kumzari

pānzdah > (**pānzda* >) *pānḏa* ‘fifteen’
šāzdah (but cf. NP *šānzdah*) > (**šānzda* >) *šānḏa* ‘sixteen’

There are just two Arabic loanwords in the data in which *d* has acquired emphasis under the influence of nearby emphatics:

Arabic >

Kumzari

(Shihhi Ar.) *ḏad* > *ḏaḏ* ‘in a way that hinders’
qaṣd > *qaṣḏ* ‘intention’

The fact that there are few words in which this particular change has taken place – in either inherited or borrowed vocabulary – may be attributable to a relatively recent emergence of *ḍ* in the Kumzari consonant inventory (Section 5.1).

s > ṣ

In a substantial number of Kumzari words, diffusion of emphasis has caused *s* to become emphatic *ṣ*. Inherited items in which this change has taken place are as follows:

<u>SWIr</u> >	<u>Kumzari</u>	
MP <i>sabz</i> >	(<i>sawz</i> >) <i>ṣawz</i>	‘green’
MP <i>sōzan</i> / NP <i>sūzan</i> >	<i>ṣūẓin</i>	‘medicinal needle’
MP <i>suxr</i> >	<i>ṣirx</i>	‘red’
MP <i>*xwāsirg</i> ? (cf. NP: Minabi <i>xwaserg</i>) >	<i>xṣurg</i>	‘sister-in-law’

Arabic loanwords in which *s* has become *ṣ*, as illustrated by the following examples, are even more numerous:

<u>Arabic</u> >	<u>Kumzari</u>	
<i>daqquṣ</i> (Gulf Ar.) >	<i>daqquṣ</i>	‘hot pepper sauce’
<i>ḥaris</i> >	<i>ḥāraṣ</i>	‘guard, watchman’
<i>sāḥir</i> ‘magician’ >	<i>ṣāḥar</i>	‘sorcerer’
<i>suʔāl</i> >	(<i>swāl</i> >) <i>ṣwāl</i>	‘question’
<i>qāmūs</i> >	<i>qāmuṣ</i>	‘dictionary’
<i>qyās</i> >	<i>qyāṣ</i>	‘measurement’

l > ḷ

In one word, Arabic *l* has acquired emphasis as a result of a neighbouring *w*: Arabic *wa* ‘and’ + *lā* ‘not’ > Kumzari *waḷa* ‘or’.

h > ḥ

Finally, in one case, diffusion of emphasis (whether from *q* or *w*) has caused historical *h* to be rendered in Kumzari as a pharyngeal *ḥ* – the possible emphatic correlate of *h* in this system:

<u>source</u> >	<u>Kumzari</u>	
Ar. <i>qahwa</i> (t) >	<i>qaḥwē</i>	‘coffee’

Along with the spontaneous change in MP *āhan* > K *ḥan* ‘iron’ (Section 5.2.2), this change must have taken place some time ago: the *h* in both source items would

have likely become *h* prior to the application of a relatively recent sound change whereby (with a few exceptions; see Section 4.6) non-emphatic *h* came to be pronounced as a glottal stop *ʔ* in Kumzari (see data in Anonby 2011a: 377); otherwise, a glottal stop rather than a pharyngeal consonant *ħ* would be expected in these words.

5.4 The sound change *z* > *ʒ*

The phonologization of emphasis has reached its acme in Kumzari with the implementation of an across-the-board sound change – applying without exception – in which *z* has been recast as an emphatic *ʒ*; non-emphatic *z* has now disappeared from the consonant inventory. The *z* > *ʒ* sound change has applied throughout the Kumzari lexicon, affecting inherited vocabulary as well as words borrowed from Arabic. Examples of inherited vocabulary in which this innovation has taken place in Kumzari are as follows:

MP	Kumzari	cf. NP	
<i>azēr</i>	<i>ʒēr</i>	<i>zīr</i>	‘under’
<i>hazār</i>	<i>ʔāʒar</i>	<i>hizār</i>	‘thousand’
<i>rēz-</i>	<i>-rēʒ-</i>	<i>-rīz-</i>	‘pour (impf.)’
<i>rōz</i>	<i>rōʒ</i>	<i>rūz</i>	‘day’
<i>zād</i>	<i>ʒād</i>	<i>zād</i>	‘give birth (pret.)’
<i>zan</i>	<i>ʒank-</i>	<i>zan</i>	‘woman’
<i>zard</i>	<i>ʒard</i>	<i>zard</i>	‘yellow’
<i>zīndag</i>	<i>ʒīndaġ</i>	<i>zinda</i>	‘alive’

Examples of Kumzari words originally borrowed from Arabic, and in which *z* has similarly undergone this sound change, are as follows:

Arabic	Kumzari	cf. NP	
<i>ʕazā</i>	<i>ʔazyit</i> / <i>ʔēʒē</i>	<i>azā</i>	‘mourning, consolation’
<i>ġazāl</i>	<i>ġāʒalē</i>	<i>ġazāl</i>	‘gazelle’
<i>ʔijāza(t)</i>	<i>ʔijāʒit</i>	<i>ijāza</i>	‘leave (n.), permission’
<i>jazīra(t)</i>	<i>jēʒirit</i>	<i>jazīra</i>	‘island’
<i>lāzim</i>	<i>lāʒum</i>	<i>lāzim</i>	‘(it is) necessary’
<i>wazīr</i>	<i>wēʒir</i>	<i>wazīr</i>	‘government minister’
<i>zamān</i>	<i>ʒāman</i>	<i>zamān</i>	‘era’
<i>ziyāra(t)</i>	<i>ʒāwarit</i> / <i>ʒiyārit</i>	<i>ziyārat</i>	‘visit’

The displacement of *z* by its emphatic counterpart *ẓ* equally applies to Kumzari words recently borrowed from English via Gulf Arabic, such as *flēẓar* ‘freezer’ and *talfaẓūn* ‘television’.

The existence of an emphatic *ẓ* with no plain counterpart is an example of the extraordinary situation in which a marked phoneme is found in the language without an unmarked counterpart. Specifically, this goes against the universal tendency that segments with secondary articulations are only found when there is also a plain counterpart (Gamkrelidze 1978: 11–13).

Since historical *z* was a common phoneme in both inherited and borrowed vocabulary, the effects are far-reaching in present-day Kumzari: the resulting emphatic *ẓ* appears in 327 items (more than 7% of words) in a lexicon of just under 4400 entries. This compares, for example, with about the same number of items (335 words) containing *f*, and 217 items (5%) containing *j* [d͡ʒ], when both of these other sounds are manifestly better represented cross-linguistically than *ẓ*. This pattern runs counter to the cross-linguistic tendency for rare phonemes, and in particular co-articulated emphatics, to be poorly represented in the lexicon (Delattre 1971: 131; Gamkrelidze 1978: 13).

This typological anomaly is addressed here by means of three questions. First, in terms of the Kumzari phonological system, would it be more appropriate to interpret [zʕ] as a plain consonant? (Section 5.4.1) Second, if it really is an underlyingly emphatic consonant with no plain counterpart, can it be attributed to borrowing from Arabic? (Section 5.4.2) And third: if not, how has such a situation arisen? (Section 5.4.3). The discussion closes with reflections on the ultimate motivation of this sound change (Section 5.4.4).

5.4.1 Should [zʕ] be interpreted as a plain consonant?

In light of the improbability that a language would have an emphatic consonant without its plain counterpart, it is worth reviewing the evidence for the phonological status of [zʕ]. If it could be classified as a plain consonant, this would result in a simpler, more symmetrical phonemic inventory, and it would alleviate the need to account for the counter-intuitive, exceptionless sound change *z > ẓ* (cf. the introductory discussion in Section 5.4).

A key indication that [zʕ] is underlyingly emphatic is its distinctly uvularized pronunciation (cf. Anonby 2011a: 376; refer also to recordings of #7 *ẓām* ‘time, occasion’ and #8 *čāẓ* ‘lunch’ in the Appendix). But in itself, this phonetic evidence is not enough. Recall that the liquid *r*, which in some positions is uvularized and others is not, has been analyzed phonologically as a plain consonant (Section 4.6). However, there are additional indicators which suggest

that *z* is underlyingly emphatic. First of all, it is likely that speakers perceive *z* as an emphatic; this is evident in speakers' unprompted choice of the emphatic character ẓ (MSA: ẓ) to write *z*, as observed in Kumzari text messages and as formalized in al-Kumzari's independently designed first version of a standard Kumzari orthography (2006). But crucially, *z* acts like other emphatics in that it initiates the diachronic and synchronic spread of emphasis onto potentially emphatic consonants in the same word. The liquid *r*, in contrast, does not appear to cause emphasis to spread (Sections 5.3 and 5.5).

Taken together, these facts confirm that there is no phonological basis for excluding *z* from the Kumzari inventory of emphatics. It is therefore necessary to account for the appearance of this contrastive emphatic consonant in the language.

5.4.2 The possibility of borrowing from Arabic

Although it is unusual crosslinguistically, the existence of emphatic consonants with no plain counterpart may also be attested in early forms of Arabic, as well as some dialects of modern Arabic, in the form of the Old Arabic (OA) lateral *qʰ* [qʰ], an emphatic voiced dental stop with a lateral release (Holes 2004: 71; al-Wer 2004), and the emphatic voiced lateral fricative ẓ [ẓ] known from Classical Arabic (Lehn 1963: 29; Owens 2006). While voiced dental stops and lateral approximants were both found in early Arabic, a plain voiced dental stop with a lateral release and a plain voiced lateral fricative were not. The phoneme ẓ persists today in a few dialects (Qahtani 2015), but in most varieties descended from OA, it has as its reflex *q* [q], the emphatic synchronic counterpart of an existing consonant *d*. The diachronically unstable phoneme *qʰ* has also been recast as an emphatic counterpart of existing consonants: in most “Bedouin” type dialects, it is now *ḏ* [ḏ], the emphatic counterpart of *d*; in “City” type dialects, it has generally become *ḏ* [ḏ], the emphatic counterpart of *d*; and in many dialects of northern Mesopotamia (and increasingly also in mid-formal registers of urban Arabic dialects in general), it has become *z* [z], the emphatic counterpart of *z*. In Classical Arabic (and its descendent Modern Standard Arabic), a phonemic split was introduced, and OA *qʰ* is now reflected as *ḏ* in some CA words, and *q* in others, but plain phonemic counterparts are found for both of these emphatics (Holes 2004: 71; Fischer and Jastrow 1980: 50). There are still, however, a few modern dialects of Arabic where OA *qʰ* has been retained (Versteegh 2006; Heselwood et al. 2013); and although *qʰ* has no exact plain counterpart, some scholars see it, at least historically, as an emphatic reflex of *š* [ʃ] (Steiner 1977).

Despite the fact that emphatics z and q^l are found in some of the dialects mentioned above, it is not possible to draw a direct link between these consonants and the z found in Kumzari. To begin with, as regards z , Kumzari is not within the historical realm of influence of northern Mesopotamian varieties of Arabic. But more importantly, addressing the contention that there could have been some recent importation of this sound from the mid-level register of urban Arabic dialects into Kumzari, it is important to point out that these dialects use z as a reflex of OA q^l ; whereas in Kumzari, emphatic z is not descended from any Arabic emphatic, but from the plain z which was previously found both in inherited vocabulary and in Arabic borrowings. In contrast, all Kumzari emphatics in borrowed words descended from OA q^l are reflected as t or, in more recent borrowings, q (Section 5.1).

So the central question remains: if emphatic z has not come directly from Arabic, how has it arisen?

5.4.3 Factors which have favoured the appearance of z in Kumzari

There are a number of factors which – although not in themselves sufficient to account for the replacement of z with its emphatic counterpart z – at least point to elements in the linguistic and social context that have been favourable to the appearance of an emphatic z .

First of all, as has emerged in discussions with Pétur Helgason (pers. comm. 2011), the general phonetic distribution of historical z in Kumzari is not unfavourable to the appearance of emphasis: low, non-front vowels would cause a transitional co-articulation with some of the same acoustic characteristics as emphasis (cf. Ohala 1993a). Importantly, low vowels have been associated with the distribution of emphasis in other languages (Owens 2006: 200; see also the comparative Arabic data in Fischer and Jastrow 1980: 56–57, which is especially pertinent for the t ; Hoberman 1989: 90; and the Kurdish data in Section 4.1 above), and these vowels have a high frequency in Kumzari. In a Kumzari lexicon of 4400 entries, historical z occurs beside one of these low vowels in over half (53%) of the items in which it is found.

Secondly, to state the obvious, there is an imbalance in the Kumzari phonological inventory which has been introduced as a result of borrowing from neighbouring varieties of Arabic where a similar imbalance exists: among the alveolar obstruents, it is only historical z which has had no emphatic counterpart. In contrast, other emphatic obstruents have been found in the language for many centuries (dating back to the retention of emphatics in words borrowed from Arabic; see Section 5.1). The system is therefore “primed” to accommodate its inclusion (cf. Kiparsky 1995: 656).

Another prominent linguistic imbalance which has come about as a result of borrowing is situated in the lexicon. While emphatics are frequent in vocabulary borrowed from Arabic, they have been (until the subsequent proliferation of emphasis) absent in inherited vocabulary. The analogical spread of diffusion (Section 5.3) is one way in which the language has counteracted distributional bias, but an exceptionless sound change in which a common, plain *z* becomes an emphatic *ẓ* is far more efficient: this single change in the system immediately introduces emphasis into hundreds of words.

Related to this, for the other plain/emphatic pairs, any merging of consonant pairs resulting only in emphatics would have violated contrast between the members of the pairs; but for *z*, a shift to emphatic status has not affected existing patterns of contrast.

Two remaining factors in the appearance of *ẓ* in Kumzari relate to imbalances brought about by the language's social context. The first of these is specific to Kumzari. As Bo Isaksson (pers. comm. 2011) has pointed out, the fact that Arab identity is of great importance to the Kumzari (see the beginning of Section 5; Anonby 2011b; Anonby and Yousefian 2011: 45) promotes the strengthening of linguistic characteristics such as emphasis that are (unconsciously or consciously) seen as Arabic (compare this also with Kahn's comments, in Section 4.1 above, on "hyper-Arabization" of Arabic loanwords in Kurdish).

Finally, as Nettle (1999) and Wohlgemuth (2010) have pointed out, there is a correlation between the (small) size of a language community, the level of endangerment, and the likelihood that linguistic rarities will emerge. With only 4000 speakers, most of whom are distributed in three communities (Section 2), and with the viability of the language under threat (Anonby and Yousefian 2011: 38–39, 82–87), it is not unreasonable to argue that the unlikely possibility of a highly marked *z > ẓ* innovation in Kumzari has been enhanced by its sociolinguistic situation.

In sum, these five factors provide insight into contextual dynamics which favour the appearance of an emphatic *ẓ* in Kumzari. However, even taken together, they do not actually explain the mechanism by which this sound change has taken place. Each of these factors would be constructive in accounting for the appearance of a *z / ẓ* contrast, but this is not what has happened; the unmarked consonant is replaced by a marked consonant *ẓ* in the inventory. The final factor above is not so much an explanation of the mechanism of change, as it is an acknowledgment of contextual elements with which diachronic anomalies such as the exceptionless sound change found in Kumzari are more likely to take place.

5.4.4 Final reflections on the $z > \text{ʒ}$ sound change in Kumzari

The discussion in this section has shown that the $z > \text{ʒ}$ sound change in Kumzari is a phonological change rather than simply a phonetic change, and that the appearance of ʒ cannot be directly attributed to borrowing from Arabic. It is revealing that no similar sound change has taken place in any variety of Arabic spoken today. Although there are linguistic and social imbalances which have favoured the appearance of an emphatic ʒ in the language, these are not sufficient to account for the highly marked $z > \text{ʒ}$ sound change which results in an unexpected, asymmetrical configuration in the consonant inventory. This stands in opposition to the predictions of theorists ranging from Jakobson (1929/1971) to Kiparsky (1995: 654), who assert that sound changes have an inherent direction, from marked to unmarked, in language systems. The fact that this change has taken place in the context of a small, threatened language community is relevant, but it does not supply us with an explanation for the mechanism of change; it simply confirms the atypical nature of the change. Even a probabilistic explanation which seeks to situate the change in a complex or “accidental” combination of ordinary linguistic and social motivations and processes (Lass 1990; Harris 2008) is therefore problematic. In the end, we are left with an unusual case of sound change which no single explanation is entirely sufficient; rather it appears to be the “accidental result of many different circumstances being lined up in just the right way” (Harris 2008: 55; see also Blust 2005).

5.5 Synchronic alternation

In large part, the spread of emphasis in Kumzari is a diachronic phenomenon (Sections 5.2–5.4). But synchronically, there are two contexts in which emphasis alternates systematically across morpheme boundaries.¹¹

First, when the “count” morpheme *-ta* is added to the bare form of numbers that end in an emphatic consonant, the morpheme-initial *t* becomes emphatic ʈ . There are only a handful of contexts where this process can take place (i.e., number stems ending in emphatic consonants), and other

¹¹ It is notable that this synchronic process of emphatic alternation takes place with only two suffixes. This is, however, due to a limitation in the productive combinatory possibilities of the language rather than any exceptions in the pattern of alternation. Specifically, there are no other stem-affix combinations in the language that meet the conditions of this alternation apparent from the data above: namely that two consonants, of which one is emphatic and the other is potentially emphatic, come into contact with one another at a stem-affix boundary.

morphophonological alternations take place at this boundary in many cases, but a comparison with stems ending in plain consonants makes the scope of its application clear:

<i>šaṣ</i> ‘sixty’	+ -ta	→ <i>šaṣta</i> ‘sixty (count form)’
<i>šaḍ</i> ‘hundred’	+ -ta	→ <i>šaṭṭa</i> ‘one hundred (count form)’
<i>dwēṣṭ</i> ‘two hundred’	+ -ta	→ <i>dwēṣṭa</i> ‘two hundred (count form)’
<i>siṣaḍ</i> / <i>ṣiṣaḍ</i> ¹² ‘three hundred’	+ -ta	→ <i>siṣaṭṭa</i> / <i>ṣiṣaṭṭa</i> ‘three hundred (count form)’

cf. other consonant-final numbers:

<i>čār</i> ‘four’	+ -ta	→ <i>čārta</i> ‘four (count form)’
<i>aft</i> ‘seven’	+ -ta	→ <i>afta</i> ‘seven (count form)’
<i>ašt</i> ‘eight’	+ -ta	→ <i>ašta</i> ‘eight (count form)’
<i>aftad</i> ‘seventy’	+ -ta	→ <i>aftata</i> ‘seventy (count form)’

In a second context, but only with some speakers – perhaps because the co-articulation of emphasis is in the process of phonologization here – when noun stems ending in an alveolar emphatic consonant are inflected with the comparative suffix *-tar*, the initial *t* of the suffix similarly becomes emphatic *ṭ*.

<i>bāruṭ</i> ‘patience’	+ -tar	→ <i>bāruṭṭar</i> (but for some speakers: <i>bāruṭtar</i>) ‘more patience’
<i>ṣawz</i> ‘green thing’	+ -tar	→ <i>ṣawzṭar</i> (/ṣawztar) ‘greener thing’
<i>ṭēz</i> ‘sharp thing’	+ -tar	→ <i>ṭēzṭar</i> (/ṭēztar) ‘sharper thing’
<i>ṭuṣ</i> ‘vile thing’	+ -tar	→ <i>ṭuṣṭar</i> (/ṭuṣtar) ‘viler thing’

As may be expected, the *t* of the suffix does not become emphatic when the stem does not end with an emphatic consonant.

<i>gārad</i> ‘thief’	+ -tar	→ <i>gāradtar</i> ‘more of a thief’
<i>naḥs</i> ‘unlucky person’	+ -tar	→ <i>nḥastar</i> ¹³ ‘unluckier person’
<i>pāk</i> ‘pure thing’	+ -tar	→ <i>pāktar</i> ‘purer thing’
<i>sēr</i> ‘full (with food) person’	+ -tar	→ <i>sērtar</i> ‘fuller (with food) person’

Surprisingly, however, this process does not take place with the pharyngeal consonant *ḥ*, or with the emphasis-inducing consonants *q*, *w* and *x*.

¹² Note *ṣi-*, the emphatic variant of *si-* ‘three’ used by some speakers for this word; this variant has arisen through the analogical diffusion of emphasis (5.3) from the stem-initial *ṣ* of *ṣaḍ*.

¹³ The *aḥ* sequence in this word undergoes metathesis when the *-tar* suffix is attached.

<i>kissah</i> ‘lame person’	+ -tar	→ <i>kissaḥtar</i> ‘lamer person’
<i>ḡilq</i> ‘difficult thing’	+ -tar	→ <i>ḡilqtar</i> ‘more difficult thing’
<i>ṣirx</i> ‘red thing’	+ -tar	→ <i>ṣirxtar</i> ‘redder thing’
<i>taw</i> ‘fever’	+ -tar	→ <i>tawtar</i> ‘more of a fever’

To frame this second process in the terms of the Arabic grammarians introduced above, the synchronic spread of emphasis in Kumzari is limited to consonants which fall under the narrow definition of emphasis as *ṭiḥbāq* (Section 3.1), that is, alveolar consonants with a secondary posterior articulation. This stands in contrast to the diachronic diffusion of emphasis in Kumzari, which operates according to broader conceptualization of emphasis that includes other posterior consonants: the pharyngeal *ḥ* and the uvulars *w*, *x* and *q*, that (Sections 3.4 and 5.3.1).

6 Conclusions

Kumzari is an endangered Indo-European language spoken on both sides of the Strait of Hormuz. Surrounded by Arabic, this language has adopted many characteristics of Arabic contact varieties. One of these features is phonological “emphasis”, a bundle of features most commonly associated with pharyngealization, velarization or uvularization. In contrast to related languages in contact with Arabic, where emphasis remains at the periphery of the language, emphasis has been profoundly phonologized in Kumzari; there, it generally patterns as uvularization, but the phonological parameters for its functioning are complex. The Arab ethnicity of Kumzari speakers and the fact that Arabic is Kumzari’s only significant contact language are two major external conditions that have promoted this degree of contact-induced change.

The Kumzari lexicon, like that of most Arabic dialects, is now permeated with emphatic consonants. The language has acquired many of these consonants as a result of an ongoing process of lexification by Arabic dating back at least 1300 years. Emphatics have been invariably maintained in loanwords, but they are not limited to borrowed vocabulary; rather, they have appeared and spread in all parts of the lexicon: borrowed, inherited and Kumzari-specific vocabulary.

Language-internal innovations, of which two types are diachronic and one is synchronic, have driven the proliferation of emphasis in the Kumzari lexicon and phonological system. To begin with, emphasis has recurrently spread in the lexicon from emphatic or emphasis-inducing consonants to potentially emphatic consonants through the mechanisms of analogical sound change. Additionally, an across-the-board sound change in which *z* has been invariably recast as an

emphatic *ẓ* has resulted in the appearance of emphasis in hundreds of additional words. Finally, in contexts where two consonants come together at a morpheme boundary, emphasis alternates synchronically by spreading onto potentially emphatic consonants.

This paper has recounted for the first time how, in one language, emphasis has arisen through language contact and has become progressively phonologized through language-internal mechanisms. In addition, two parts of this story have implications beyond the linguistic history of Kumzari. The first relates to the typologically improbable sound change from *z* to *ẓ* that has taken place, with system-wide repercussions, in Kumzari. In light of its phonetic characteristics as well as its diachronic and synchronic behaviour, the change must be viewed as phonological rather than simply phonetic. In addition, it cannot be directly attributed to borrowing from Arabic. There are several contextual factors that suggest a general linguistic and social bias favouring the spread of emphasis in the phonology and lexicon, but these are not sufficient to explain the exceptionless replacement of *z* with its emphatic counterpart *ẓ* and the cross-linguistically unique configuration that results. Not only is a co-articulated segment present in the phonological inventory without a plain counterpart; this cross-linguistically marked segment is also frequent in the lexicon. The fact that this change has taken place in the context of a small, threatened language community where structural rarities are more likely to arise is relevant, but it does not explain the mechanism of change; it simply situates the anomalous nature of the change. In the end, we are left with an unusual case of sound change which is not linguistically motivated.

A final observation relates to the typology of emphasis. The phonetic and phonological overview of emphasis makes it clear that this phenomenon is multi-faceted and difficult to delimit, in any language, with a neat and consistent set of features. This inherent complexity is a key factor in the variable extent and patterning of emphasis across languages. Classical Arabic and spoken Modern Standard Arabic exhibit a core set of emphatic alveolar consonants characterized by secondary pharyngealized (or other posterior) co-articulation, defined as *ʔiṭbāq* ‘spreading and raising’ by the early Arabic grammarian Sibawaih. Observing phonological effects on neighbouring vowels, Sibawaih referred to an extended class of consonants, additionally including uvulars *x* & *q*, which is characterized by *ʔistiʕlāl* ‘raising (of the tongue back)’. Emphatic allophonic variants of the consonants *l* and *r* are grouped with the contrastive emphatic consonants under the label of *tafxīm* ‘thickening, enlarging, emphasizing’, and in a few cases, scholars have extended this latter designation to pharyngeals *ħ* and *ʕ*. Analyses of a wide sample of Arabic dialects show many variations in inventory, scope, and function, but the emphatic alveolars are a recurrent component of the system. In Kumzari, an Indo-European language in

close contact with Arabic, further possibilities in the development and configuration of emphasis are evident. Here, a core set of alveolar emphatics is also found, but is characterized by uvularization as a dominant secondary articulation. In keeping with a uvular place of articulation, the consonants *x* and *q*, as well as a uvular *w*, have a clear role in the historical diffusion of emphasis; and evidence for historical spread of emphasis from pharyngeal *ħ* is also found. However, mechanisms of the synchronic spread of emphasis in Kumzari are different from those involved in its diachronic spread. Whereas consonants of the extended emphatic set define its sporadic and analogical diachronic spread, synchronic spread of emphasis is completely regular, and is limited to the core set of co-articulated emphatics.

Acknowledgements: I express my sincere appreciation to all those among the Kumzari people who welcomed us and shared with us insights into the riches of their unique language, and I thank them for their gracious response to what is undoubtedly an incomplete picture of the language, despite my best scientific efforts as a scholar. I also thank all of the scholars who have contributed to the ideas and argumentation in this paper, including Christina van der Wal Anonby, Sami Aydin, Lev Blumenfeld, Bernard Comrie, Michael Cysouw, Pétur Helgason, Geoffrey Haig, Bo Isaksson, Carina Jahani, Agnes Korn, Colin MacKinnon, Hassan Mohebbi Bahmani, Maria Persson, Ferhad Shakely, Jaffer Sheyholislami, Ambjörn Sjörs and Gernot Windfuhr, as well as the anonymous referees of this paper, who provided extensive feedback.

The support of the following institutions in the preparation of this article is gratefully acknowledged: Carleton University, Uppsala University (research position UFV-PA 2010/2580), Sultan Qaboos University, the Oman Studies Centre, Leiden University Centre for Linguistics, and the Alexander von Humboldt Foundation (Fellowship for Experienced Researchers).

References

- Aikhenvald, Alexandra. 2006. Grammars in contact: A cross-linguistic perspective. In Alexandra Aikhenvald & R. M. W. Dixon (eds.), *Grammars in contact: A cross-linguistic typology*, 1–66. Oxford & New York: Oxford University Press.
- al-Kumzari, Ali Hassan Ali. 2006. Kumzari lexicon and grammar notes (ms.).
- al-Wer, Enam. 2004. Variability reproduced: A variationist view of the Dhaa/Daad opposition in modern Arabic dialects. In Martine Haak, Rudolf De Jong & Kees Versteegh (eds.), *Approaches to Arabic dialects*, 21–31. Leiden: Brill.

- Anonby, Erik J. 2003. Update on Luri: How many languages? *Journal of the Royal Asiatic Society* 13(2). 171–197.
- Anonby, Erik J. 2011a. Illustrations of the IPA: Kumzari. *Journal of the International Phonetic Association* 41(3). 375–380.
- Anonby, Erik J. 2011b. The appearance of velaro-pharyngeal emphasis in Kumzari as an effect of Arabic contact. Paper presented for the Semitic and Iranian Seminar series, Department of Linguistics and Philology, Uppsala University, 28 September 2011.
- Anonby, Erik J. 2012. Stress-induced vowel lengthening and harmonization in Kumzari. *Orientalia Suecana* 61. 54–58.
- Anonby, Erik J. 2016. The Keshmi (Qeshmī) dialect of Hormozgan Province, Iran: A first account. *Studia Iranica* 44(2). 165–206.
- Anonby, Erik J. in prep. a. Historical phonology of Kumzari.
- Anonby, Erik J. in prep. b. Dialectology of Musandam Arabic.
- Anonby, Erik J. & Christina van der Wal Anonby. in prep. *Kumzari dictionary/Ktēb majma Kumzārī*.
- Anonby, Erik J. & Pakzad Yousefian. 2011. *Adaptive multilinguals: A survey of language on Larak Island*. Uppsala: Uppsala University Press.
- Applegate, Joseph R. 1970. The Berber languages. In Thomas A. Sebeok (ed.), *Current trends in linguistics, Vol. 6: Linguistics in South West Asia and North Africa*, 596–661. The Hague: Mouton.
- Bakallah, Muhammad Hasan. 2009. Tafxim. In Kees Versteegh, Mushira Eid, Alaa Elgibali, Manfred Woidich & Andrzej Zaborski (eds.), *Encyclopedia of Arabic language and linguistics*, Vol. 4, 421–424. Leiden & Boston: Brill.
- Barry, Daniel. 2019. Pharyngeals in Kurmanji Kurdish: A reanalysis of their source and status. In Songül Gündoğdu, Ergin Öpengin, Geoffrey Haig & Erik Anonby (eds.), forthcoming. *Current issues in Kurdish linguistics*, 39–71. Bamberg: Bamberg University Press.
- Bayshak, Maryam Salam. 2002. Hal aṭṭarat as-Sāsāniyah fī lisān al-Šuḥūḥ wa hal al-Kamzāriyah min awjuh ḍalik? Al-lahjat al-Šihhiyah fī ḍaw' Ṣilm al-luḡāt [Are there traces of Sassanian in the language of the Shihuh, and is Kumzari among the affected varieties? The Shihhi dialect in the light of linguistic science]. *Al-Khaleej* 8541. 12, October 17, 2002.
- Bernabè, Roy S. 2011. *A phonology and morphology sketch of the Šihhi Arabic dialect of alĠēdih, Musandam (Oman)*. Leiden: School of Middle Eastern Studies, Leiden University Master's thesis.
- Blau, Joyce. 1989a. Gurânî et Zâzâ. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 336–340. Wiesbaden: Ludwig Reichart Verlag.
- Blau, Joyce. 1989b. Le Kurde. In Rüdiger Schmitt (ed.), *Compendium Linguarum Iranicarum*, 327–335. Wiesbaden: Ludwig Reichart Verlag.
- Blust, Robert A. 2005. Must sound change be linguistically motivated? *Diachronica* 22(2). 219–269.
- Bruk, Solomon I. & V. S. Apenchenko (eds.). 1964. *Atlas narodov mira [Atlas of the world's peoples]*. Moscow: Glavnoye Upravleniye Geodezii i Kartografii.
- Bukshaisha, F. 1985. *An experimental phonetic study of some aspects of Qatari Arabic*. Edinburgh: University of Edinburgh dissertation.
- Card, Elizabeth Anne. 1983. *A phonetic and phonological study of Arabic emphasis*. Ann Arbor, MI: University Microfilms International.
- Chen, Matthew Y. & William S.-Y. Wang. 1975. Sound change: Actuation and implementation. *Language* 51. 255–281.

- Clements, George N. & Elizabeth Hume. 1995. Internal organization of speech sounds. In John Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Cambridge, MA & Oxford: Blackwell.
- Colarusso, John 1975. *The Northwest Caucasian languages: A phonological survey*. Cambridge, MA: Harvard University dissertation.
- Comprehensive *Aramaic Lexicon*. 2012. <http://cal1.cn.huc.edu> (accessed May 2012).
- Davis, Stuart. 1995. Emphasis spread in Arabic and grounded phonology. *Linguistic Inquiry* 26 (3). 465–498.
- Delattre, Pierre. 1971. Pharyngeal features in the consonants of Arabic, German, Spanish, French and American English. *Phonetica* 23. 129–155.
- Dolgopolsky, Aharon B. 1977. Emphatic consonants in Semitic. *Israel Oriental Studies* 7. 1–13.
- Dostal, Walter. 1972. The Shihuh of northern Oman: A contribution to cultural ecology. *The Geographical Journal* 138. 1–7.
- Embarki, Mohamed. 2013. Phonetics. In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 23–44. Oxford: Oxford University Press.
- Esling, John H. 1996. Pharyngeal consonants and the aryepiglottic sphincter. *Journal of the International Phonetic Association* 26(2). 65–88.
- Esling, John H. 1999. The IPA categories “pharyngeal” and “epiglottal”: Laryngoscopic observations of pharyngeal articulations and larynx height. *Language and Speech* 42(4). 349–372.
- Ferguson, Charles A. 1956. The emphatic *l* in Arabic. *Language* 32. 446–452.
- Fischer, Wolf Dietrich & Otto Jastrow (eds.). 1980. *Handbuch der arabischen Dialekte*. Wiesbaden: Otto Harrassowitz.
- Gamkrelidze, Thomas A. 1978. On the correlation of stops and fricatives in a phonological system. In Joseph H. Greenberg (ed.), *Universals of human language 2: Phonology*, 9–46. Stanford, CA: Stanford University Press.
- Grigore, George. 2011. Les principales caractéristiques de l’arabe parlé à Siirt (Turquie). In Andrei A. Avram, Anca Focșeneanu & George Grigore (eds.), *Festschrift for Nadia Anghelescu*, 255–277. Bucharest: University of Bucharest.
- Haig, Geoffrey. 2007. Grammatical borrowing in Kurdish (Northern Group). In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective* (Empirical Approaches to Language Typology 38), 165–183. Berlin & New York: Mouton de Gruyter.
- Haig, Geoffrey & Yaron Matras. 2002. Kurdish linguistics: A brief overview. *Sprachtypologie Und Universalienforschung/Language Typology and Universals* 55(1). 3–14.
- Harris, Alice C. 2008. On the explanation of typologically unusual structures. In Jeff Good (ed.), *Linguistic universals and language change*, 54–76. Oxford & New York: Oxford University Press.
- Hassan, Zeki M., John H. Esling, Scott R. Moisik & Lise Crevier-Buchman. 2011. Aryepiglottic trilled variants of /ʕ, ħ/ in Iraqi Arabic. In Wai Sum Lee & Eric Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 831–834. Hong Kong: City University of Hong Kong.
- Hayward, Katrina M. & Richard J. Hayward. 1989. “Guttural”: Arguments for a new distinctive feature. *Transactions of the Philological Society* 87(2). 179–193.
- Herzallah, Rukayyah S. 1990. *Aspects of Palestinian Arabic phonology: A non-linear approach* (doctoral dissertation). Ithaca, NY: Cornell University.
- Heselwood, Barry. 2007. The “tight approximant” variant of the Arabic ‘ayn. *Journal of the International Phonetic Association* 37(1). 1–32.
- Heselwood, Barry, Janet C. E. Watson, Munira al-Azraqi & Samia Naïm. 2013. Lateral reflexes of Proto-Semitic *ḏ and *ḏ in Al-Rubū’ah dialect, south-west Saudi Arabia: Electropalatographic and acoustic evidence. In Renaud Kutý, Ulrich Seeger & Shabo Talay

- (eds.), *Nicht nur mit Engelszungen. Beiträge zur semitischen Dialektologie: Festschrift für Werner Arnold*, 135–144. Harrassowitz: Wiesbaden.
- Hoberman, Robert D. 1985. The phonology of pharyngeals and pharyngealization in Pre-Modern Aramaic. *Journal of the American Oriental Society* 105(2). 221–231.
- Hoberman, Robert D. 1988. Emphasis harmony in a Modern Aramaic dialect. *Language* 64(1). 1–26.
- Hoberman, Robert D. 1989. Parameters of emphasis: Autosegmental analyses of pharyngealization in four languages. *Journal of Afro-Asiatic Languages* 2(1). 73–97.
- Hoberman, Robert D. 1995. Current issues in Semitic phonology. In John Goldsmith (ed.), *The handbook of phonological theory*, 839–847. Cambridge, MA & Oxford: Blackwell.
- Holes, Clive D. 2004. *Modern Arabic: Structures, functions and varieties*. rev. edn Washington, DC: Georgetown University Press.
- IPA (International Phonetic Association). 2015. The international phonetic alphabet. <https://www.internationalphoneticassociation.org/content/ipa-chart> (accessed February 2017).
- Jakobson, Roman 1929/1971. Remarques sur l'évolution phonologique du russe comparée à celle des autres langues slaves. *Selected writings 1: Phonological studies*, 7–116. The Hague: Mouton.
- Jakobson, Roman 1957/1971. Mufaxxama: The 'emphatic' phonemes in Arabic. *Selected writings 1: Phonological studies*, 510–522. The Hague: Mouton.
- Jastrow, Otto. 1980. Das mesopotamische Arabisch. In Wolf Dietrich Fischer & Otto Jastrow (eds.), *Handbuch der arabischen Dialekte*, 140–158. Wiesbaden: Harrassowitz.
- Jayakar, Atmaram S. G. 1902. The Shahee dialect of Arabic. *Journal of the Bombay Branch of the Royal Asiatic Society* 21. 246–277.
- Kahn, Elizabeth. 1976. *Borrowing and variation in a phonological description of Kurdish*. Ann Arbor: Phonetics Laboratory, University of Michigan.
- Khan, Geoffrey. 1999. *A grammar of Neo-Aramaic: The dialect of the Jews of Arbel*. Leiden: Brill.
- Kiparsky, Paul. 1995. The phonological basis of sound change. In John Goldsmith (ed.), *The handbook of phonological theory*, 640–670. Cambridge, MA & Oxford: Blackwell.
- Lass, Roger. 1990. How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26(1). 79–102.
- Laufer, Asher. 1996. The common [ʕ] is an approximant and not a fricative. *Journal of the International Phonetic Association* 26(2). 113–117.
- Lehn, Walter. 1963. Emphasis in Cairo Arabic. *Language* 39. 29–39.
- Lodhi, Abdulaziz. 2003. Aspiration in Swahili adjectives and verbs. *Africa and Asia 3: Essays on African and Asian languages*, 155–160. Göteborg: Department of Oriental and African languages, Göteborg University.
- MacKenzie, David Neil. 1961. *Kurdish dialect studies I*. London: Oxford University Press.
- MacKinnon, Colin. 1974. *The phonology and morphology of Dezfūlī-Shushtari: A study in West Persian dialectology*. Los Angeles, CA: University of California Los Angeles.
- MacKinnon, Colin 1995. Dezfūlī and Šūštārī. *Encyclopædia Iranica Online*. <http://www.iranicaonline.org/articles/dezful#pt2> (accessed May 2012).
- Marçais, Philippe. 1948. L'articulation de l'emphase dans un parler arabe maghrébin. *Annales de l'Institut d'Etudes Orientales, Faculté des Lettres de l'Université d'Alger* 7. 5–28.
- Matras, Yaron. 2007. Grammatical borrowing in Domari. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 151–164. Berlin & New York: Mouton de Gruyter.
- Matras, Yaron. 2009. *Language contact*. Cambridge: Cambridge University Press.

- McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In Patricia Keating (ed.), *Papers in laboratory phonology III: Phonological structure and phonetic form*, 191–233. Cambridge: Cambridge University Press.
- McCarus, Ernest M. 2009. Kurdish. In Gernot Windfuhr (ed.), *The Iranian languages*, 587–633. London & New York: Routledge.
- Moisik, Scott R. (2013). *The epilarynx in speech*. Victoria, BC: University of Victoria dissertation.
- Najmabadi, Shahnaz. 1988. Identité ethnique contre nationalité: Le cas de l'Île de Larak (Golfe Persique). In Jean-Pierre Digard (ed.), *Le fait ethnique en Iran et en Afghanistan*, 65–74. Paris: Éditions du CNRS.
- Nettle, Daniel. 1999. *Linguistic diversity*. Oxford & New York: Oxford University Press.
- Nirumand, Mohammad Bāqer. [Circa 1970]. *Nasāb-e Šūštar [Record of Shushtar]*. Tehran: [Publisher unspecified].
- Ohala, John J. 1993a. Coarticulation and phonology. *Language and Speech* 36(2/3). 155–170.
- Ohala, John J. 1993b. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237–278. London & New York: Longman.
- Owens, Jonathan. 2006. *A linguistic history of Arabic*. Oxford: Oxford University Press.
- Paradis, Carole & Darlene LaCharité. 2001. Guttural deletion in loanwords. *Phonology* 18(2). 255–300.
- Paul, Ludwig. 1998. *Zazaki: Grammatik und Versuch einer Dialektologie*. Wiesbaden: Ludwig Reichert.
- Paul, Ludwig. 2009. Zazaki. In Gernot Windfuhr (ed.), *The Iranian languages*, 545–586. London & New York: Routledge.
- Pisowicz, Andrzej. 1985. *Origins of the New and Middle Persian phonological systems*. Krakow: Uniwersytet Jagielloński.
- Qahtani, Khairiah. 2015. *A sociolinguistic study of the Tihami Qahtani dialect in Asir, Southern Arabia*. Colchester: University of Essex dissertation.
- Rahimpour, Massoud & Majid Saedi Doveise. 2011. A phonological contrastive analysis of Kurdish and English. *International Journal of English Linguistics* 1(2). 73–82.
- Robson, Barbara & Habibullah Tegey. 2009. Pashto. In Gernot Windfuhr (ed.), *The Iranian languages*, 721–772. Oxford & New York: Routledge.
- Sadeghi, Ali Ashraf. 1986. Arabic Language, part 1: Arabic elements in Persian. *Encyclopædia Iranica* 2(3). 229–231.
- Shahin, Kimary N. 1997. *Postvelar harmony: An examination of its bases and crosslinguistic variation*. Vancouver, BC: University of British Columbia dissertation.
- Shahin, Kimary N. 1998. Optimized postvelar harmony in Palestinian Arabic. In Elabbas Benmamoun, Mushira Eid & Niloofar Haeri (eds.), *Perspectives on Arabic linguistics*, Vol. 11, 143–164. Amsterdam & Philadelphia: John Benjamins.
- Sibawaih, 'Amr ibn 'Uthmān. 1970. In Hartwig Derenbourg (ed.), *Al-Kitāb Sībawaih II*. Hildesheim: Georg Olms.
- Skjærvø, Prods Oktor. 1989. Languages of southeast Iran: Lārestānī, Kumzārī, Bāškardī. In Rüdiger Schmitt (ed.), *Compendium linguarum iranicarum*, 363–369. Wiesbaden: Ludwig Reichert.
- Steiner, Richard. 1977. *The case for fricative-laterals in Proto-Semitic*. New Haven, CT: American Oriental Society.
- Thomas, Bertram. 1930. The Kumzari dialect of the Shihuh tribe (Musandam), Arabia, and a vocabulary. *Journal of the Royal Asiatic Society* 62. 785–854.

- van der Wal Anonby, Christina. 2014. Traces of Arabian in Kumzari. *Proceedings of the Seminar for Arabian Studies* 44 (supplement): *Language of Southern Arabia*, 137–146.
- van der Wal Anonby, Christina. 2015. *A grammar of Kumzari*. Leiden: Leiden University dissertation.
- van der Wal Anonby, Christina. forthcoming. Arabian elements in Kumzari. In Christopher Lucas (ed.), *Arabic and contact-induced change*. London: Language Science Press.
- Versteegh, Kees. 2006. Dād. In Kees Versteegh, Mushira Eid, Alaa Elgibali, Manfred Woidich & Andrzej Zaborski (eds.), *Encyclopedia of Arabic language and linguistics*, Vol. 1, 544–545. Leiden & Boston: Brill.
- Watson, Janet C. E. 1999. The directionality of emphasis spread in Arabic. *Linguistic Inquiry* 30. 289–300.
- Watson, Janet C. E. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Watson, Janet C. E. & Alex Bellem. 2010. A detective story: Emphatics in Mehri. In Janet Starkey (ed.) *Proceedings of the Seminar for Arabian Studies* 40: Papers from the forty-third meeting, London, 23–25 July 2009, 345–356. Oxford: Archaeopress.
- Windfuhr, Gernot. 1990. Persian. In Bernard Comrie (ed.), *The world's major languages*, 523–546. Oxford & New York: Oxford University Press.
- Windfuhr, Gernot. 2009. Dialectology and topics. In Gernot Windfuhr (ed.), *The Iranian languages*, 5–42. London/New York: Routledge.
- Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian languages*, 416–544. London & New York: Routledge.
- Wohlgemuth, Jan. 2010. Language endangerment, community size, and typological rarity. In Jan Wohlgemuth & Michael Cysouw (eds.), *Rethinking universals: How rarities affect linguistic theory*, 255–277. Berlin & New York: De Gruyter Mouton.
- Zawaydeh, Bushra A. 1998. Gradient uvularization spread in Ammani Jordanian Arabic. In Elabbas Benmamoun, Mushira Eid & Niloofar Haeri (eds.), *Perspectives on Arabic linguistics*, Vol. 11, 117–141. Amsterdam & Philadelphia: John Benjamins.
- Zimmermann, Wolfgang 1981. *Tradition und Integration mobiler Lebensformgruppen: Eine empirische Studie über Beduinen und Fischer in Musandam, Sultanat Oman*. Göttingen: Georg-August-Universität dissertation.

Appendix: Sound files

Seven accompanying sound files can be accessed in the online version of the article. They are provided to illustrate the data and observations in the discussion, as indicated in the relevant sections of the article.

1. *tak* ‘date syrup basket’, *ṭāf* ‘gale’
2. *dār* ‘stick’, *ḍayf* ‘guest’
3. *sā* ‘now’, *ṣām* ‘handle’
4. *zām* ‘time, occasion, turn’, *čāz* ‘lunch’
5. *hē* ‘yes’, *ḥāl* ‘situation’
6. *wāl* ‘slitting (a shark)’
7. *rōr* ‘child’