Phonetics/Phonology/Phonotactics /Morphophonology/Prosody

Kazakh phonology

Kazakh phonology

Introduction

One of the most immediately striking facts about Kazakh phonology that distinguishes the language from most of its sister languages in the Turkic family is the extent to which the language uses consonant assimilatory processes, along with the role sonority plays in doing so. In fact, various syllabification processes in the language provide some of the best types of evidence for the place of Sonority Hierarchy in phonological theory. Likewise, as with consonants, there are particular properties of Kazakh vowels that make them immediately prominent within the Turkic language family, such as the presence of several diphthongoid vowels, vowels that contain an initial glide element (although this feature is also present in some NW Turkic languages, especially Karakalpak). At the same time, there are numerous properties of Kazakh vowels and consonants, such as vowel harmony and the presence of harmonic variants of consonants, which make its phonology very similar to other Turkic languages.

Segmental phonology

Vowels

The presence of nine vowel phonemes is usually agreed upon (Krueger 1980, Kirchner 1992, 2006, Somfai Kara 2002, Muhammedowa 2016, Washington 2016; cf. Трубецкой 1960) consisting of the back vowels /ɑ, o, ɤ, ʊ / and their front counterparts /e, ø, ɪ, ʏ/, as well as the sometimes contested low front vowel /æ/, which is mostly used in words borrowed from Arabic and Persian and is limited at large to initial syllables, never occurring in suffixes. It emerged in the language through fronting of /ɑ/, and does not contrast with /e/. These vowels respectively correspond to the symbols <a, o, ï, ụ, e, ö, ị, ü, ä> of the Turkological literature. Table 1 below illustrates the vowel inventory of Kazakh:

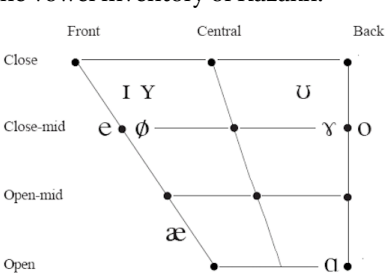

Figure 1: Kazakh vowel inventory

The vowels /e, o, ø/ are often pronounced with an initial glide, a phenomenon that is most obvious in word-initial position, as in the examples of /ʲeski/ 'old, ancient' and /ʷol/ 'he, she, it'. This led some authors to propose that these are phonetically diphthongs although they phonologically behave like single units (e.g. Vajda 1994, Джунисбеков 1980). In fact, Vajda

recommends using the following characters instead: /ʃi̯, wo̯, wʉ̯/, based on Джунисбеков's (1972) x-ray analyses of the oral tract. They are transcribed furthermore as ʸe, ʷo, ʷö.

Consonants

Kazakh has a large inventory of consonant sounds, many of which are allophonic variants of other phonemes and some are recent borrowings from Russian. The following chart summarizes these based on place (each column) and manner of articulation (each row) (see Krueger 1980, Kirchner 1992, Vajda 1994). Some of these appear only as allophones, and /f, v, x, ts, tʃ, h/ are the borrowed phonemes. Some, such as [ts] and [h], rarely appear in normal speech (and thus presented in parentheses). The language is particularly rich in velar and uvular consonants, many of which are in complementary distribution, as will be demonstrated below. Excluding allophonic variants, and those borrowed from other languages, it has 17 native consonant phonemes; these are the stops /p, b, t, d, k, g/, fricatives /s, z, ʃ, ʒ/, nasals /m, n, ŋ/, liquids /r, l/, and two glides /w, j/.

	Bilabial	**Labiodental**	**Alveolar**	**Alveo-palatal**	**Velar**	**Uvular**	**Glottal**
Stop	p b		t d		k g	q ɢ	
Fricative		f v	s z	ʃ ʒ	x ɣ	χ ʁ	(h)
Affricate			(ts)	tʃ			
Nasal	m		n		ŋ	N	
Liquid			r l				
Glide	w			j			

Figure 2: Kazakh consonants

These consonants, given in IPA above, demonstrate certain changes from their Turkic counterparts, changes that are in general principled. Four such patterns are immediately recognizable: (i) Turkic [tʃ] corresponds to Kazakh [ʃ], e.g. /qatʃ/ à /qaʃ/ 'run away'; (ii) Turkic [ʃ] in turn corresponds to Kazakh [s] in final position, e.g. /tyʃ/ à /tys/ 'fall down'; (iii) Turkic [j] corresponds to [ʒ] in initial position, e.g. /jaz/ à /ʒaz/ 'write' (a change that first

led to /j/ à /dʒ/ before resulting in /dʒ/ à /ʒ/, but stayed as /dʒ/ in some dialects, such as in the south and east, e.g. /dʒaz/, Jankowski 2010); and, perhaps most notably, (iv) Turkic [ɣ] corresponds to Kazakh [w] in final position /aɣ / à /aw/ 'net' (see also Krueger 1980, Johanson 2009).

Kazakh shows numerous allophonic alternations, which mostly involve, but are not limited to, its rich inventory of velar and uvular consonants. For example, velar [g] and uvular [ʁ] are in complementary distribution; while the former appears in the environment of front vowels, the latter appears in the environment of back vowels. This is similar to the [g] and [ɣ] alternation observed in many other Turkic languages, although, in Kazakh, the two sounds differ not only in manner (stop vs. fricative), but also in place of articulation (velar vs. uvular), showing that an additional operation (one that targets place) is in place. A similar pattern that shows this operation is observed with [k] and [q], with the former appearing in front and the latter in back vowel environments, although [χ] tends to be used, instead of [q], following the vowel [a] (e.g. /ǰaχsï/, 'good', Kirchner 2006, p. 321). Furthermore, according to some authors (e.g. Vajda 1994), this alternation involving velar and uvular stops holds with nasal stops, too, meaning that while velar [ŋ] appears in front environments, its uvular counterpart [N] appears in back environments, although, in this case, the difference is not marked in the alphabet. These alternations are illustrated in (1) below:

(1) a. [g] ~ [ʁ]		b. [k] ~ [q]		c. [ŋ] ~ [N]	
[gʏl]	~ [ʁalɯm]	[kel]	~ [qal]	[keldɪŋ]	~ [ʷʊN]
'rose'	'researcher'	'come'	'stay'	'you came'	'left (direction)'

Suprasegmental phonology

Kazakh suprasegmental phonology offers a wide array of phonological processes, which are highly informative for formal phonological theory.

Syllabification and consonant assimilatory processes

Kazakh does not permit onset clusters; therefore, no Kazakh words start with a sequence of two consonants. Although there are words in the language which are orthographically represented with a sequence of two consonants in initial position, these are pronounced with an epenthetic vowel, which is inserted either before the cluster or in between the two members of the cluster, as in < класс > [kɯ.las] and <спорт> [ɪs.port], thereby preventing a complex onset. Although this is the pattern in many parts of Kazakhstan (especially the south), some educated speakers in northern and western areas pronounce these words with a cluster, perhaps due to Russian influence.

Kazakh does, however, have coda clusters, although there are significant conditions on their distribution. The most important one of these is that while the first member of the cluster needs to be a sonorant, i.e. /m, n, ŋ, r, l, w, j/, the second member has to be an obstruent, meaning that Kazakh strictly abides by the Sonority Sequencing Principle, which dictates that coda clusters fall in sonority, e.g. [kɪlt], 'key, [ʒʊrt] 'people', [ʒent] '(a type of)

dessert'. There are two additional conditions however, both of which target the second member of the cluster. First, it has to be homorganic, i.e. have the same place of articulation as the preceding consonant. Further, it needs to be voiceless, which is true irrespective of the quality of the initial member. Exceptions to this pattern exist, but are restricted to liquid-initial clusters in onomatopoeia, e.g. [qark qark], 'laughter' (Vajda 1994), or [ʒalp] (falling noise). Other words, which contain an underlying final cluster and appear with a cluster in their dictionary form in other Turkic languages surface with an epenthetic vowel breaking the cluster, as in /χɑlq/ appearing as [χɑlɨq] ~ [qalɨq] and /murn/ as [murun] ~ [murɨn]. That the underlying representation of these forms contains a cluster is clear from their affixed versions, where they appear as a cluster e.g. [χɑlq-ɨm] ~ [qalq-ɨm] 'my people' and [murn-ɨm] 'my nose' (Washington, 2010).

One may conclude that the second requirement above, that final clusters end in a voiceless (or 'strong/fortis', as many would argue for Kazakh) obstruent, is due to a more general constraint in Kazakh phonology, namely that certain obstruents are devoiced (or strengthened) in syllable-final position. Although this is true, and although voiced stops /b, d, g/ are indeed devoiced/strong in final position (and therefore the contrast with /p, t, k/ is lost), this condition does not target fricatives (e.g. [qɯz] 'girl' is fine), but consonant clusters ending in a voiced fricative are still not permitted, e.g. *rz].

In Kazakh, processes involving voicing (or lenition) assimilation do not only occur in word- or syllable-final position. Similar processes target consonant-initial suffixes; thus, the initial consonant of suffixes alternate between a voiced/weak and a voiceless/strong variant, compare e.g. [balaq-tar] 'fish-pl' vs. [qɯz-dar] 'girl-pl'.

In addition to assimilation, changes in the quality of Kazakh consonants occur also because of reasons related directly to markedness constraints targeting coda/onset syllabification. One such constraint is the Syllable Contact Law (SCL), which Kazakh strictly abides by, unlike many other Turkic languages. As per the SCL, an onset consonant in Kazakh has to be less sonorous (or at least cannot be more sonorous) than the preceding coda consonant. The effects of SCL can be observed through a look at the phonologically conditioned allomorphs of the plural morpheme and the yes/no question particle, as illustrated in (2):

(2)	i. Plural /-Lar/	ii. Question /-MA/	Gloss
	a. alma-lar	alma-ma	'apple'
	b. maŋday-lar	maŋday-ma	'forehead'
	c. ḳ̇iyar-lar	ḳ̇iyar-ma	'cucumber'
	d. ḳol-dar	ḳol-ma	'arm'
	e. mụrụn-dar	mụrụn-ba	'nose'
	f. koŋɯz-dar	koŋɯz-ba	'bug'

g. süyek-ter süyek-pe ‘bone’ (modified from Davis 1998:191)

These data indicate that when an /-l/ or /-m/-initial suffix is attached to a root that ends in a coda consonant with equal or lower sonority, the initial consonant of the suffix, which also functions as the onset of the final syllable, desonorizes, leading to the [d]- and [t]-initial variants for the plural suffix and [b]- and [p]-initial variants for the question particle. For the plural suffix, for example, the allomorph *-lar* is used after a vowel-, glide-, or /r/-final root (2a to 2c), whereas *-dar* (or *-tar* depending on voicing assimilation) is used after a /l, n, z, k/-final root, as all of these consonants have equal or lower sonority than /l/, and thus, using the allomorph *-lar* would have violated the SCL: *l.l, *n.l, *z.l, *k.l. Notice also that the change from *-lar* to *-dar* does not occur after -r-final words, providing evidence that /r/ is more sonorous than /l/, helping contribute to a long held debate in phonology as to whether all liquids are equally sonorous (e.g. Gouskova 2004). Notice also that the onset consonant in (2g) has to be equally sonorous as the coda consonant (e.g. k.t and k.p), which is because the onset consonant cannot desonorize further than a stop. So the crucial generalization is that although codas should ideally be more sonorous than the following onset, equally sonorous codas are not banned outright and are resorted to when necessary.

Vowel harmony

As with the great majority of Turkic languages, Kazakh has two types of vowel harmony, (i) backness harmony and (ii) rounding harmony, out of which only the former is represented orthographically. Furthermore, similar to other Turkic languages, backness harmony is quite strong, applying across all vowels and regardless of word length; so a back suffix vowel is used following a back vowel, and a front suffix vowel following a front vowel, e.g. [at-qa] ‘horse-DAT’ and [jet-ke] ‘meat-DAT’. As for rounding harmony, which ensures that a suffix vowel agrees in rounding with the preceding vowel, although it applies more strongly in Kazakh than many other Turkic languages like Turkish, Uyghur and Tuvan, thereby targeting not only underlyingly high vowels ([süj-dü] ‘kiss-ed’), but also – for most speakers – some non-high vowels (e.g. [süj-mö] ‘don’t kiss’), it is also more restricted than in languages like Kirghiz and Altay in that the vowel [o] is blocked, e.g. [quj-ma] ‘don’t spill’, but not *[quj-mo]. Furthermore, the effect of rounding decreases as the distance increases from the source, as can be seen through a comparison of [ʒüzüm-dö] ‘in the grape’ vs. [ʒüzüm-ümüz-de] ‘in our grape’.

Although both backness and rounding harmony target both root and suffix vowels, certain words, especially those borrowed from other languages or formed through compounding, contain vowel combinations not in line with vowel harmony. The suffix vowel in such cases harmonizes with the final vowel of the stem, e.g. [muɣalim-nen] ‘from the teacher’. Many consonants (such as /k, g, l/) are also involved in vowel harmony in Kazakh, a phenomenon especially true for backness harmony (Kirchner, 1992; Johanson, 2009, Vajda 1994), as can be observed in the contrast between [ʁ] and [g] in the words [bal-ʁa] ‘honey-DAT’ and [bel-ge] ‘waist-DAT’.

Stress and intonation

As with most Turkic languages, accent in Kazakh falls on the final syllable of a prosodic word. As such, each time a suffix is added to a word, accent appears to move to the right, as can be seen in the following example: *almá* 'apple', *alma-lár* 'apples', *alma-lar-ím* 'my apples', *alma-lar-ím-dá* 'on my apples'.

A comparison of Kazakh with other Turkic languages demonstrates, however, that word-level prominence may be closer to 'stress' in Kazakh than in most other Turkic languages, such as Turkish and Uyghur, because accented syllables in Kazakh bear greater duration (in addition to higher pitch) than their unaccented counterparts, a situation not observed in languages like Turkish and Uyghur (Özçelik 2015). In fact, as Özçelik demonstrates, many of the apparent exceptions to stress assignment in Kazakh are aligned with the weight of the stressed syllable, i.e. that heavy syllables (whether long or ending in a coda) attract stress, as with the words *kép.ka* 'hat', *ʃe.gírt.ke* 'grasshopper', *dar.ba.zá:* 'gate', and *sa.rum.sák* 'garlic'. As such, finally stressed syllables, if open, end in a long vowel, a pattern that is typical of iambic languages in general (Hayes 1995), as opposed to the pattern displayed by many other Turkic languages in which word level accent is more like intonational prominence than stress (see e.g. Özçelik 2014, 2017) in that final accent is mostly manifested through pitch only (Johanson 1998, Özçelik 2014).

In addition to apparent exceptions observed in roots mentioned above, there are also certain suffixes, which, when attached to a word, lead to exceptional stress, for these cannot bear stress, or are, in other words, pre-stressing. Such suffixes include the negation suffix {-MA}, the question particle {-MA}, the adverbial marker {-ʃA}, and perhaps most notably the copula, thereby affecting all person agreement suffixes, too.

(3) Regular stress: Exceptional stress:

a. [ʲeʃek-síz] 'without a donkey' [ʲeʃék-siz] 'You're a donkey.'

b. [alma] 'apple' [ál-ma] 'Don't take.'

(3a) compares the morpheme that means 'without', which is regular, with the second person agreement marker, which is pre-stressing, as it is followed by the silent/null copula. (3b) compares the root [alma], which is regular with a form where the pre-stressing negation suffix follows the regular root [al], leading to the same segmental profile but differing at the autosegmental level, which results in a contrast in meaning.

Bibliography

Davis, Stuart. 1998. Syllable contact in Optimality Theory. *Korea Journal of Linguistics* 23. 181–211.

Джунисбеков, Алимхан. 1972. *Гласные казахского языка* (*экспериментально-фонетическое исследование*). Алма-Ата: Наука.

Джунисбеков, Алимхан. 1980. Сингармонизм в казахском языке. Алма-Ата: Наука.

Hayes, Bruce. 1995. *Metrical stress theory: principles and case studies*. Chicago.

Johanson, Lars. 1998. The structure of Turkic. In: Johanson, Lars & Csató, Éva Á. (eds.) *The Turkic languages*. London. 30–66.

Johanson Lars. 2009. Kazakh. In: Brown, K. & Ogilvie, S. (eds.) *Concise encyclopedia of languages of the world*. Oxford. 588–592.

Jankowski, Henryk. 2010. The Consonants ž - ğ in Kazakh. In: Mańczak-Wohlfeld, Elżbieta & Podolak, Barbara (eds.) *Studies on the Turkic word: A festschrift for Professor Stanislaw Stachowski on the occasion of his 80th birthday*. Kraków. 23–32.

Kirchner, Mark. 1992. *Phonologie des Kasachischen: Untersuchungen anhand von Sprachaufnahmen aus der kasachischen Exilgruppe in Istanbul, Teil 1: Untersuchung*. Wiesbaden.

Kirchner, Mark. 2006. Kazakh and Karakalpak. In: Johanson, Lars & Csató, Éva Á. (eds.) The Turkic languages. London. 318–332.

Krueger, John R. 1980. *Introduction to Kazakh*. Bloomington.

Muhammedowa, Raihan. 2016. *Kazakh: A comprehensive grammar*. London.

Özçelik, Öner. 2017. The Foot is not an obligatory constituent of the Prosodic Hierarchy: "stress" in Turkish, French and child English. *The Linguistic Review* 34:157–213.

Özçelik, Öner. 2015. 'Stress' or 'Intonational prominence'? Word accent in Kazakh and Uyghur. *Turkic Languages* 19:163–192.

Özçelik, Öner. 2014. Prosodic faithfulness to foot edges: the case of Turkish stress. *Phonology* 31:229-269.

Somfai Kara, Dávid. 2002. *Kazak*. München.

Трубецкой, Н. С. 1960. *Основы фонологии* [Principles of phonology]. Москва.

Vajda, Edward J. 1994. Kazakh Phonology. In: Kaplan, E. et al. (eds.) *Opuscula Altaica: Essays Presented in Honor of Henry Schwarz*. Bellingham, WA. 603–650.

Washington, Jonathan. 2010. *Syllable margin phonology in Kazakh*. Manuscript, Indiana University.

Washington, Jonathan. 2016. An investigation of vowel anteriority in three Turkic languages using ultrasound tongue imaging. PhD Dissertation, Indiana University.