What happens to consonant clusters in Mongolian speech?

Karlsson, Anastasia; Svantesson, Jan-Olof

Published in:
[Host publication title missing]

2007

What happens to consonant clusters in Mongolian speech?

Anastasia Mukhanova Karlsson and Jan-Olof Svantesson
Dept. of Linguisics and Phonetics, Lund University

Introduction

As is well known, many languages treat some types of consonant clusters, in particular clusters with two identical consonants, as uncomfortable for speech production, and tend to reduce them, especially in casual speech. For example, the Swedish word *hemskt* 'terrible' is normally pronounced [hemst], although the pronunciation [hemskt] can be heard in more formal speech. Similarly, /ʲ/ in Russian лестница 'staircase' is normally deleted so that this word is pronounced [ˈlʲesʲnʲɪtsɐ].

In Mongolian speech, however, we have observed that consonant clusters, even clusters of two identical stops, are usually not reduced. This is typical not only for careful pronunciation but also for casual speech. An interesting feature of Mongolian speech is that vowels are less stable than consonants. Thus, we found that short phonemic vowels in the initial syllable can be devoiced and even deleted completely in fast speech despite the fact that the vowel of the first syllable determines vowel harmony. For example, *ciwsʰəgʞəʞ* зэвсэглэл 'armament' occurred as [cfsʰəxɬʞ] in casual speech with complete reduction of the first vowel and devoicing of consonants. The epenthetic vowel [ə], which serves to build syllables at the surface level, is often missing in casual speech so that syllable codas with up to four fully realised consonants are found. Some examples are /ačʞʞč/ ажиллаж 'working' realised as [ačɬč] with a three-consonant coda (regular pronunciation would be [a.čəʞ.ʞəč] (where . denotes syllable boundary)) and /carcʰʊʞgtsnig/ зарцуулагдсаныг 'to use-PASS-PSTP-ACC' realised as [car.cʰʊɬxts.nig] with a four-consonant coda (regular pronunciation would be [car.cʰʊ.ʞəgt.sə.nig].

In order to describe the phonotactic features in Halh Mongolian appropriately, it is necessary to distinguish between formal and casual style of speech. Formal speech is characterised by careful pronunciation, while casual speech is everyday speech with less attention to pronunciation. In our previous studies (Svantesson et al. 2005 for formal speech and Karlsson 2005 for casual speech) differences in syllabification between the two styles were described and analyzed.

Syllabification

In our analysis of Mongolian phonology, non-initial "reduced" vowels (here transcribed as [ə] and written with single vowel letters in the Cyrillic alphabet) are epenthetic (non-phonemic) and are inserted in order to build well-formed surface syllables, consisting of a one-consonant onset, a vowel nucleus (which is the only obligatory constituent) and a coda. Consonant

clusters that have strictly decreasing sonority are possible codas. For Mongolian this means (with few exceptions) that the combination voiced + voiceless consonant is allowed as a coda (e.g. /pugt/ [pugt] бүгд ‘all’), while other combinations cannot form codas but trigger schwa epenthesis: voiceless + voiced (/atg/ [a.təg] адаг ‘end’), voiceless + voiceless (/pʊst/ [pʊ.sət] бусад ‘other’) and voiced + voiced (/xamr/ [xa.mər] хамар ‘nose’). Medial consonant clusters are divided between a coda and an onset; if this is not possible, an epenthetic vowel is inserted. These rules hold for formal speech, but for casual speech the syllabification rules are less precise due to several assimilation processes which lead to varying surface realisations of the phonemes.

One such process is voicing assimilation, seldom found in formal speech, but frequent in casual speech. Assimilatory regressive devoicing, causing voiced phonemes to have a voiceless realisation is almost obligatory before aspirated stops and affricates and is frequent, but not obligatory, before voiceless fricatives and unaspirated voiceless stops. Furthermore, the requirement in formal speech that syllable codas have strictly decreasing sonority is relaxed in casual speech to allow codas consisting of consonants with the same sonority value (voiceless + voiceless or voiced + voiced), so that only the combination voiceless + voiced is impossible (at least at the end of a word). As a consequence codas with up to four voiceless consonants may occur in casual speech.

Clusters of identical consonants

In the casual speech material investigated here, many cases of schwa insertion (as expected in formal speech) do not take place and codas with non-decreasing sonority are very frequent, including codas with clusters of adjacent identical segments. Such clusters are seldom simplified in Mongolian, and all segments are fully realised even in casual speech. Thus, while many languages with complex codas, such as Russian and Swedish tend to reduce some consonant clusters, Mongolian prefers not to. An epenthetic vowel can be inserted between two adjacent identical stops to ease the pronunciation, but in many cases the two stops are realised without vowel epenthesis. Examples from our material include the words *caaɮttəg* заалддаг ‘to sue-HAB’, *caaɮttəgar* заалддагаар ‘to sue-HAB-INST’ and *caaɮttəgara* заалддагаараа ‘to sue-HAB-INST-REFL’, spoken in the frame sentence *pii___ gisəŋ* би___гэсэн ‘I said ___.’. Since they are in a focused position, epenthesis should be favoured, but this is not the case in the actual realisations of these words. In 14 of 23 occurrences, both consonants were realised with both occlusion and release, and no epenthesis occurred. The other occurrences of /tt/ were reduced to one [t].

An example is given in Figure 1, where the word /caaɮttgar/ заалддагаар ‘to sue-HAB-INST’ is not realised as formal [caaɮt.tə.gar], but as [caaɬtʰtxar]. The two adjacent occurrences of the stop [t] are fully realised with two occlusions and releases.

[FIGURE 1 ABOUT HERE]

In our material there are clusters with two /g/ and different combinations of aspirated and unaspirated /t/. Our observations on the realisation of these clusters within a coda and across syllable (or word) boundaries can be summarised as follows:

With few exceptions, clusters over word boundaries are reduced to one stop. For example, /cʰait tʰɔstɔi/ цайд тостой is realised as [cʰaitʰɔstɔi]. Clusters with two /g/ (within words and over word boundaries, as in /tɔrɮʲg gitg/ дорлиг гэдэг) are always reduced to one stop [g].

In addition to the 23 occurrences mentioned above, there are 21 words from connected speech (read texts) in our material, which contain clusters with two /t(ʰ)/ (aspirated or unaspirated). These are words where the root ends with /t(ʰ)/ and a following suffix begins with another /t(ʰ)/, and where the second stop is an onset, for example /nat.tʰai/ надтай 'I-COM'.

In 25 of the total 44 occurrences, both stops are realised, each with an occlusion phase and a release. In these cases, the realisation is always [tʰt], independent of the underlying cluster (/tt/, /tʰt/, /ttʰ/ or /tʰtʰ/). Apparently, the release of the first consonant is produced rather strongly, making it aspirated, in order to make it perceptually salient that the cluster consists of two consonants.

In many languages, clusters of identical stops are produced with only one release, but with longer occlusion phase. This is the case in Russian where we find minimal pairs as подтянуть 'to pull up' vs. потянуть 'to pull' or подточить 'to sharpen a little' vs. поточить 'to sharpen', realised as [pətʲːɪˈnutʲ] vs. [pətʲɪˈnutʲ] and [pətːɐˈtʃitʲ] vs. [pətɐˈtʃitʲ]. Here, the pronunciation of the cluster with two stops is simplified to only one occlusion and release, but the different durations of the occlusions in each pair carry the contrastive function.

In other languages, clusters of identical consonants may be simplified to single consonants at least in some positions. Thus, Swedish pairs such as *avvisa* 'reject' vs. *avisa* 'de-ice' or *uttrycka* 'express' vs. *utrycka* 'pull out' are pronounced in the same way (as [ˈɑːˌviːsa] and [ˈʉːˌtʁʏka] in the Southern Swedish pronunciation of the second author), although the morphological and phonological structures of the words in each pair are different (*av+visa* vs. a*v+isa* and *ut+trycka* vs. *ut+rycka*).

ɮ clusters

Mongolian is typologically unusual in having a lateral fricative /ɮ/ (pronounced [ɮ] or [ɬ] in speech independently of the context), and its palatalised counterpart /ɮʲ/, but no plain (non-

fricative) laterals. The fricative pronunciation of /ɮ/ is exceptionless, and independent of the vowel harmony class of the word at least in Ulaanbaatar speech. Even Russian loanwords such as билет ‘ticket’ are typically pronounced with a fricative /ɮ/, except for one of our speakers with a good knowledge in Russian. Since the lateral is rather frequent, its fricativity is a very salient feature of Mongolian pronunciation.

In order to find out whether clusters of two fricatives are simplified or not, we have investigated a material consisting of words with word final /ɮɮ/ clusters. In words ending with /ɮɮ/, such as /tʰimtgɮɮ/ тэмдэглэл ‘note’, the final cluster has a non-decreasing sonority and should therefore be split by epenthesis: [tʰim.təg.ɮəɮ]. Epenthesis often fails in this type of words in casual speech, however. More interestingly, when epenthesis fails, /ɮɮ/ is systematically realised as [ɬɮ], that is, as a voiceless + voiced combination.

In order to find out if the duration of a /ɮɮ/ cluster is longer than that of a single /ɮ/, a small material was recorded from one female speaker. The following target words were recorded: /šaɮ/ шал ‘floor’; /sʊʊtɮ/ суудал ‘seat’; /taatʰɢɮ/ даатгал ‘insurance’; /mantɮ/ мандал ‘rise’; /tʊtʰgtɮ/ дутагдал ‘deficiency’. Each word was put in three carrier sentences, one with the word in utterance-final position, one where it is followed by the copula verb /pain/ байна and one where it is followed by the focus marker /ɮ/ л. This marker is cliticised to the preceding word, resulting in a final /ɮɮ/ cluster. The carrier sentences are illustrated with the word /šaɮ/ below:

1.	/in šaɮ/	Энэ шал.	‘This is a/the floor.’
2.	/in šaɮ pain/	Энэ шал байна.	‘This is a/the floor.’
3.	/in šaɮɮ pain/	Энэ шал л байна.	‘This is the floor.’

After the clitisation of the focal marker *ɮ,* the following realisations of, to give one example, /šaɮ/ occured: [šaɬəɮ], [šaɬəɬ] or [šaɬɮ] (schwa is produced with strong friction). An illustration of a realisation of the word /mantɮɮ/ мандал л ‘rise-FOC’ is given in Figure 2.

[FIGURE 2 ABOUT HERE]

Each sentence was read three or four times, but not all recordings could be used. The duration measurements are given in Table 1.

	(1)	(2)	(3a)			(3b)	
	/ɮ/	/ɮ/	/ɮ/	[ə]	/ɮ/	[ɬ]	[ɮ]
/šaɮ/	136	90	66	36	102	88	128
	124	95	75	51	115		
	119	90					
/sʊʊtɮ/	123	121	75	52	115	82	140
	130	131	84	48	109		
	142	113					
/taatʰɢɮ/	126	102	80	35	119	104	133
	126	98	62	54	111		
	124	112					
/mantɮ/	133	119	101	63	102	122	165
	130	142					
	130	130					
/tʊtʰgtɮ/	117	104	111	57	120	75	195
	134	114	59	40	120		
	205						
mean (*m*)	133	112	79	48	113	94	152

Table 1. Duration of word-final /ɮ/ and /ɮɮ/ in ms. The headings (1)–(3) refer to the three carrier sentences. (3a) are cases where the /ɮɮ/ cluster in sentence 3 is realised with schwa, and (3b) without schwa.

The results show that the single final /ɮ/ is longer (*m* = 133 ms) when it comes in utterance-final position than before the copula /pain/ (*m* = 111.5 ms). A *t*-test shows that the difference is significant ($t(27) = 3.13$, $p = 0.0042$). The longer duration in utterance-final position may be due to final lengthening.

The duration of the focus marker /ɮ/ before /pain/ (3a, 3rd column) differs minimally from the duration of stem-final /ɮ/ in the same position (carrier sentence 1). The means are 113 ms and 112 ms, respectively, and a *t*-test shows no significant difference ($t(21) = 0.185$, $p = 0.855$). This suggests that the focus marker /ɮ/ is not phonetically different from other occurrences of word-final /ɮ/.

The duration of double /ɮɮ/ is much longer than that of single /ɮ/, independent of its realisation with or without a schwa. The mean duration of the whole sequence [ɬɮ] without schwa is 246 ms and that of [ɮəɮ] with schwa is 240 ms (the difference is not significant: $t(12) = 0.410$, $p = 0.689$). This is roughly twice as much as the duration of a single /ɮ/ in this position, and the difference is highly significant ($p < 0.001$).

The duration data suggests that the final cluster [ɬʒ] is not simplified to one segment but includes two fricatives. The resulting cluster has increasing sonority and is ill-formed as a coda, even in casual speech. In casual speech nasals can build syllable nuclei (Karlsson 2005), e.g. regular [ir.sən.tən] ирсэнд нь 'to come-PSTP-DAT-THEME' is realised as [ir.sn.tn] in casual speech with two adjacent syllables both having a nasal as its nucleus. Similarly, [ɬʒ] can be interpreted as a syllable, with the first voiceless [ɬ] as onset and the second voiced [ʒ] as rhyme. In fact, analysing [ɬʒ] as a syllable explains why the second [ʒ] in [ɬʒ] is always voiced. /ʒ/ thus has a potential syllabic function.

Historical implications

As seen above, the consonants of a Mongolian word are not easily deleted even in casual speech, as compared to many other languages. This can be observed not only synchronically but also in the historical development of the Mongolic languages, where deletions affect vowels much more often than consonants. The only consonant that is systematically deleted in Halh and other dialects of Mongolian proper is **h* (as reconstructed by Svantesson et al. 2005, chapter 10): O(ld) M(ongolian) **harpan* 'ten' > Halh *arəw* арав; OM **sehyl* 'tail' > Halh *suuʒ* сүүл. In contrast, all word-final vowels were deleted in Halh e.g. OM **kʰara* 'black' > Halh *xar* хар; **kʰøkʰe* 'blue' > Halh *xox* хөх. Final vowels preceded by **h* are exceptions: OM **uhu* 'to drink' > Halh *ʊʊ* уу.

In Monguor and other Mongolic languages spoken in the Qinghai–Gansu area, initial vowels were deleted, although not as regularly as final vowel deletion in languages as Mongolian, Buriad and Kalmuck. Examples are OM **ønteken* 'egg' > Monguor *ntike* (but Halh *ontəg* өндөг) and OM **emys* 'to wear' > Monguor *mosi*, Shira Yugur *mɤs* (but Halh *oms* өмс).

Thus it seems that the preservation of consonants, sometimes at the cost of changes or deletion of vowels, is a characteristic property of the Mongolic languages both from a synchronic and a diachronic point of view.

References

Karlsson, Anastasia M. 2005. *Rhythm and Intonation in Halh Mongolian*. Travaux de l'Institut de Linguistique de Lund 46. Doctoral dissertation. Lund: Studentlitteratur.

Svantesson Jan-Olof, Tsendina, Anna, Karlsson, Anastasia M. and Franzén, Vivan. 2005. *The Phonology of Mongolian*. Oxford: University Press.

FIGURE CAPTIONS:

Figure 1. Oscillogram and spectrogram of /caaɮttgar/ заалддагаар ‘to sue-HAB-INST’ realised as [caaɬtʰtxar]. Male speaker.

Figure 2. Realisation of /mantɮɮ/ мандал л ‘rise-FOC’ as [mant.ɬɮ] with a [ɬɮ] cluster. Female speaker.