Horne, M., J. Frid, B. Lastow, G. Bruce and A. Svensson. 2003. 'Hesitation disfluencies in Swedish: prosodic and segmental correlates'. *Proceedings ICPhS* 15, 2429-32. Barcelona.

Levelt, W. 1989. *Speaking: from attention to articulation*. Cambridge, Mass.: MIT Press.

Miller, G. 1956. 'The magical number seven, plus or minus two: some limits on our capacity for processing information'. *Psychological review* 63, 81-97.

Nakajima, S. & J. Allen. 1993. 'A study on prosody and discourse structure in cooperative dialogues'. *Phonetica* 50, 197-210.

Nordling, J. 1998. 'Reparationer i spontant tal'. *Kandidat* ('B.A.') term paper, Dept. of Linguistics, Lund University.

Shriberg, E. 1994. *Preliminaries to a theory of speech disfluencies*. Ph.D. diss., Univ. of Berkeley.

Swerts, M., A. Wichmann & R.-J. Beun. 1996. 'Filled pauses as markers of discourse structure'. *Proceedings of the International congress on spoken language processing (ICSLP)*, 1033-36. Philadelphia.

Merle Horne <Merle.Horne@ling.lu.se>
Johan Frid <Johan.Frid@ling.lu.se>
Mikael Roll <Mikael.Roll@ling.lu.se>

Lund University, Dept. of Linguistics
Working Papers 51 (2005), 105–123

Vowels in Mongolian speech: deletions and epenthesis

Anastasia Mukhanova Karlsson

Mongolian has phonologically long and short vowels. Besides them there are epenthetic (non-phonemic) vowels which are inserted on the phonetic level in order to satisfy syllable well-formedness conditions, governed by the sonority law. However, in Mongolian non-formal speech phonemic vowels are often missing, and epenthesis often fails to take place where it is predicted by syllabification rules. The description of rhythmic features of Mongolian would be far from complete without an investigation of these processes.
The following is an attempt to give a systematic picture of vowel deletion in Mongolian. Two different events are identified: devoicing of phonemic vowels and non-epenthesis of schwa. We argue that vowel deletions and epenthesis depend on the consonantal composition of the words: both voicing and some articulatory features are important for these processes.

1 Reductions of phonemic vowels

1.1 Vowels in Mongolian

Several phonemic representations of vowels in Mongolian have been proposed. We follow the latest description based on acoustic analysis (for more details, see Svantesson et al. 2005) which distinguishes between full, or phonemic, vowels and reduced, or non-phonemic, vowels. There are also diphthongs in Mongolian, but they are not treated here.

Full vowels contrast as long and short, but only in initial syllables, and the duration of a short vowel in this position is on the average 48% of the duration of a long vowel (Svantesson et al. 2005). The length opposition does not occur in non-initial syllables, and we will describe phonemic vowels in this position as non-initial full vowels, without specifying their length. In addition to full vowels, there are reduced non-phonemic vowels in non-initial syllables. They are inserted at the surface level to form syllables, and are purely epenthetic. Reduced vowels are articulatorily underspecified and their quality depends on the preceding full vowel. In the following they will be transcribed as [ə] independently of their quality. In most traditional descriptions the reduced vowels are analysed as phonemic, so that non-initial full vowels are regarded as allophones of long vowels, and reduced vowels are treated as allophones of short vowels (Poppe 1951). This analysis is reflected in the Cyrillic Mongolian script, but not in Kalmuck (another Mongolic language), in

Table 1. Vowel phonemes

vowels in initial syllables	*long phonemic vowels*	*short phonemic vowels*
	i: u: ʊ: e: o: a: ɔ:	i u ʊ o a ɔ
vowels in non-initial syllables	*phonemic vowels*	*non-phonemic (reduced) vowel*
	i u ʊ e o a ɔ	ə

which reduced vowels are not represented orthographically. The vowels are given in Table 1.

1.2 Consonant phonemes in Mongolian

In Table 2, the Mongolian consonant phonemes are given, excluding some rare consonants that occur only in loanwords.

Aspirated stops and affricates are postaspirated word-initially and preaspirated word-internally and word-finally (see Karlsson & Svantesson 2002). Even /s/ is postaspirated. The nasal /ŋ/ occurs only before consonants (except dentals) or word-finally. The lateral /ɮ/ is often realised as a voiceless fricative [ɬ] in our material.

1.3 Method

To investigate the behaviour of phonemic vowels, two texts read by four female speakers of the Halh dialect were recorded. All words in the texts were analysed, except some words in utterance final positions, because of weak acoustic features. Deletion of a vowel is defined here as the acoustic absence of any vocalic features.

1.4 Results

Totally 432 word occurrences (312 polysyllabic and 120 monosyllabic) were analysed. In monosyllables, no vowel deletions were attested, but only strong devoicing. The number of deleted vowels in different positions in polysyllabic words is shown in Table 3. Long vowels are never deleted in our material. Non-long vowels (i.e. short vowels in initial syllables and phonemic vowels in non-initial syllables) are only deleted in initial syllables (in 14% in the material).

Table 2. Mongolian consonant phonemes

	labial	*palatalised labial*	*dental*	*alveo-palatal*	*pal-atal*	*velar*	*uvular*
voiceless aspirated stops			t^{h}	t^{jh}			
voiceless unaspirated stops	p	p^{j}	t	t^{j}			
voiced stops					g^{j}	g	ɢ
voiceless aspirated affricates			c^{h}	$č^{h}$			
voiceless unaspir. affricates			c	č			
voiceless fricatives			s	š	x^{j}	x	
nasals	m	m^{j}	n	n^{j}		ŋ	
voiced lateral fricatives			ɮ	$ɮ^{j}$			
rhotics			r	r^{j}			
glides	w	w^{j}			j		

Features that were chosen as possible triggers of reductions were: vowel quality, place and manner of articulation of surrounding consonants, and voicing. Besides the position in the word, only voicing influenced the deletion of non-long phonological vowels, see Table 4. (Henceforth voiced consonants are denoted C^{+}, voiceless C^{-}, and C stands for either when the difference is not relevant.)

Clearly, the deletion of initial vowels in our material depends on the phonetic environment: it is triggered by voiceless segments. Deletion never occurs between, and is very rare after or before voiced consonants, while 61% of the vowels are deleted between voiceless consonants. Voiceless environment causes devoicing as well as complete deletion of vowels (for an example, see Figure 1).

We regard this type of deletion as vowel devoicing, and the representation of reduction stages (Kohler 1991) in a voiceless context will be e.g. [ciws^{h}əgɮəɮ] → [ci̥fs^{h}əxɬəɮ] →[cfs^{h}əxɬɮ] ‘armament’, where devoicing spreads from the voiceless segments [c] and [s^{h}] to the surrounding segments and causes complete deletion of the vowel.

Table 3. Vowel deletion in 312 occurrences of polysyllabic words for all speakers. The symbol V: stands for long phonemic vowel, V for non-long phonemic vowel.

vowel	*total*	*deleted*	*%*
V:	42	0	0
V in initial syllable	270	38	14
V in internal syllable	24	0	0
V in final syllable	162	1	1

Table 4. Deletion of non-long phonemic vowels with reference to their segmental environment. C^+ stands for a (phonetically) voiced and C^- for a (phonetically) voiceless consonant allophone, # stands for word boundary and _ shows the place of the vowel.

	initial syllable		*internal syllable*		*final syllable*	
	total	*deleted*	*total*	*deleted*	*total*	*deleted*
[C^+_ C^+]	24	0	10	0	66	0
[C^+_ C^-]	18	1	—		8	0
[C^-_ C^+]	120	4	10	0	36	0
[C^-_ C^-]	46	28	4	0	4	0
[#_ C^+]	50	1				
[#_ C^-]	12	3				
[C^+_#]					20	0
[C^-_#]					28	1
total	270	37	24	0	162	1

2 Epenthesis

2.1 Syllable and syllabification generally

We accept the view of the syllable as a surface phenomenon, so that the syllabification of underlying phonemic strings in speech is governed by rules, which are language-specific. According to this view, there are no extrasyllabic elements. Insertion of epenthetic vowels in Mongolian is regarded as functioning to build up syllables.

We represent Mongolian syllable structure as in (1), with the onset and the rhyme, divided into nucleus and coda, recognised as structures relevant for syllabification. We also propose mora counting for Mongolian.

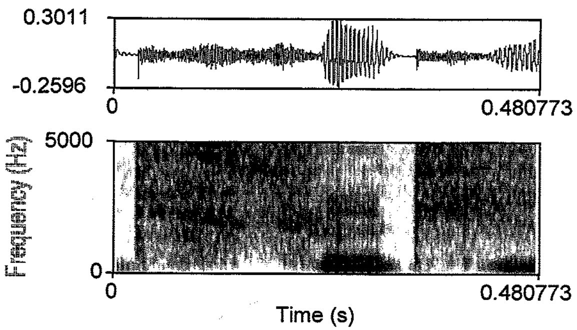
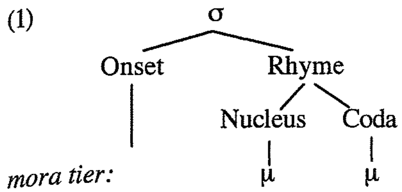


Figure 1. The word *ciwsəgʒəʒ* 'armament', realised with complete reduction of the first phonemic /i/: [cfsʰəxɬʒ]

(1)



The syllable is traditionally seen as containing a sonority peak with less sonorous segments attached to it. According to Jespersen 1897-99:525, an utterance contains as many syllables as there are sonority peaks, and the relation between the members of a syllable is stated by the Sonority sequencing generalisation (or Sonority law), which is seen as a language universal principle: syllables tend to have sonority that increases from the onset to the nucleus, the sonority peak of the syllable, and then decreases to the coda. The sonority level is stated in the Sonority hierarchy, or Sonority scale, a general version of which is (from most sonorous to least sonorous segments): vowels – liquids – nasals – obstruents. The sonority relation between segments is, however, language specific, and for Mongolian the scale vowels – voiced consonants – voiceless consonants applies.

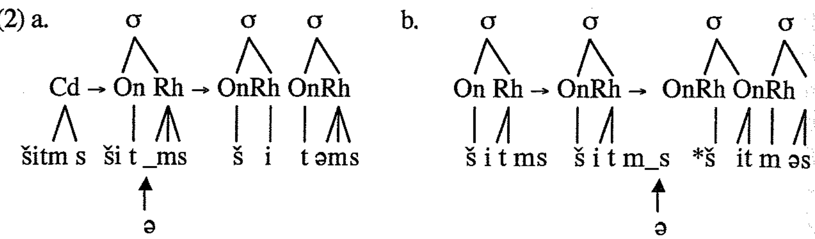
2.2 Syllabification in Mongolian

Reduced vowels are epenthetic in Mongolian and serve to build syllables at the surface level. Syllabification is governed by three principles, the Sonority law, Right-to-left syllabification (Directionality principle, Itô 1989) and Cyclicity. For details we refer to Svantesson et al. 2005. Mongolian surface syllable structure is (C){V:/V/ə}(C)(C)(C), with V:, V or *ə* as possible nuclei. Onsetless syllables are only allowed word-initially, and at vowel hiatus within a word, *g*/*ɢ*/*j* is inserted to build an onset: Ø → *g* (*ɢ*, *j*) / V __ V. Obeying Onset satisfaction, a consonant followed by a vowel becomes an onset, so that the syllable division of, e.g., *xʊʊšʊr* 'dumpling' is *xʊʊ.šʊr*. Only one-consonant onsets are allowed (Complex onset constraint).

Medial consonant clusters are divided between coda and onset, and if the consonant combination cannot form a coda, an epenthetic vowel is inserted. Possible codas are those that have decreasing sonority, following the Sonority law, so that voiced+voiceless combinations are allowed, while the reverse as well as two voiceless or two voiced consonants trigger epenthesis: /pugt/ 'all' is syllabified as [pugt], but /atg/ 'end' as [a.təg]. The only allowed clusters of voiceless consonants are a fricative followed by an aspirated stop or affricate (except c^h): *št, xt^h*, etc., and there are three-consonant codas consisting of a voiced consonant followed by one of these allowed clusters.

The direction of syllabification is important for predicting the correct place of schwa insertion. In Mongolian, epenthesis is processed from right to left, so that a maximal coda is found at the right edge of the word and combined with the preceding vowel to form a syllable rhyme. When a consonant cluster cannot form a coda, a schwa is inserted. For example, /šitms/ 'fillet' is syllabified as [ši.təms] (2a), and not as [*šit.məs] as were the case if the syllabification had the direction left-to-right (2b):

(2) a.



Due to the agglutinative nature of Mongolian, derivations and inflections are formed by adding suffixes to the stem. Syllabification of affixed words is cyclic so that words are syllabified at each addition of a new affix, i.e. at each morphological cycle. Words are syllabified as monomorphemic words until a new affix requires insertion of a schwa into the already syllabified part of a word. This is blocked by the Resyllabification constraint (Svantesson et al. 2005):

Resyllabification constraint: On each morphological cycle, an epenthetic vowel cannot be inserted into the already syllabified part of a word.

Example: t^h*ogrg* → t^h*ogr_g* → t^h*og.rəg* → t^h*og.rəg* + *ʤ* → t^h*og.rə.g_ʤ* → t^h*og.rə.gəʤ* 'to circulate', while with non-cyclic syllabification we would get: t^h*ogrg* + *ʤ* → t^h*og_rg_ʤ* → *t^h*o.gər.gəʤ*

These rules were worked out based on careful pronunciation of isolated words. Our material consists of longer utterances and texts read by several informants. We will discriminate between citation forms and our type of material by defining the first speech style as formal speech, in contrast to casual. It will be shown that syllabification in casual speech differs from formal speech in systematic ways. The Sonority law is active in casual speech though it needs to be redefined. It will be shown that it is only voiced consonants that

are marked for the sonority feature in speech, while voiceless segments are unmarked for this feature and can combine freely to form a coda until articulation disallows this. Articulation assimilation is an important factor in speech and some systematic violations of the Sonority law as well as of the Simple onset constraint are found. Open syllables with an epenthetic vowel as nucleus [Cə] are avoided in casual speech, and we analyse this as a change towards non-cyclic syllabification. This is the reason why we actually find such forms as (*)t^h*o.gər.gəʤ* instead of t^h*og.rə.gəʤ* (see the example above) in casual speech. It will also be demonstrated that nasals become syllabic in some positions, so that a nasal N should be added as a potential nucleus to the Mongolian syllable structure : (C){V/V:/ə/N}(C)(C)(C).

2.2.1 Voicing assimilation. Prediction of the syllabification of a phonemic string in speech is obscured by several assimilation processes that lead to variation in the surface realisation of the same phoneme. One of the most frequent processes found in the material is consonant devoicing, causing voiced phonemes to have a voiceless realisation.

To establish rules for voice assimilation, we analysed words from texts and isolated utterances. The material was read by totally seven speakers, four women and three men. Words where insertion of one or more schwas was expected (following regular syllabification), but did not take place were investigated. It is clusters that resulted from the non-insertion of schwa that are the object of this investigation of voice assimilation. Clusters with nasals and rhotics are not included. Totally 183 cases with different consonant clusters were investigated. The results are summarised in Table 5.

Only three cases with the [Casp _] context were found, and they are not included in the table. In two of them, the following voiced segment became devoiced. Progressive voicing of voiceless unaspirated stops was found in four cases and regressive in six cases.

The results shown in Table 5 suggest the following rules:

1. Voiced consonant phonemes become voiceless before (pre)aspirated stops. The assimilation is almost obligatory in this context (94%).
2. Voiced consonant phonemes are realised as voiceless before voiceless fricatives and affricates. The assimilation is optional but frequent in this context (79%).
3. Assimilatory devoicing is regressive in Mongolian.
4. Assimilatory voicing is rare and occasional in Mongolian.

Table 5. Voicing assimilation of voiced stop and fricative phonemes before or after voiceless consonants. Casp denotes aspirated stops and affricates, Cfric voiceless fricatives and (unaspirated) affricates, and Cstop voiceless unaspirated stops.

context	*direction*	*total*	*devoicing*	*%*	*getting voice*
[_ Casp]	regressive	31	29	94	
[_Cfric]	regressive	14	11	79	
[_ Cstop]	regressive	77	19	25	4 (progressive)
[Cfric _]	progressive	10	2	20	
[Cstop _]	progressive	48	3	6	6 (regressive)

Based on these data, we can conclude that devoicing is non-obligatory, except before aspirated stops. The optional nature of assimilation obscures the prediction of surface forms. It will be shown that codas with the combination voiceless + voiced or voiced + voiced consonants are split by epenthesis, obeying the Sonority law, while voiceless + voiceless clusters are allowed in speech. Since the same phonemic string can be realised differently due to the optional nature of voice assimilation, syllabification will also be different. Thus we find different syllabifications by the same speaker, for example /xiiɮgsŋ/ can be syllabified as [xii.ɮəɣ.səŋ] with the voiced cluster [ɮɣ] split by *ə*, or as [xiiɬx.səŋ] with a voiceless coda [ɬx].

The fricative /ɮ/ often varies in its voice feature independently of the segmental environment both within and between speakers. It is most often realised as voiceless [ɬ]. The voiced stops /g ɢ/, are almost always realised as voiceless fricatives before voiceless fricatives and stops.

To sum up, there is surprisingly little voiced material in Mongolian speech. Voiced consonant phonemes are few, and they are frequently realised without voice. The other feature that plays a role is aspiration. Voiceless stops and affricates contrast by aspiration, also realised as preaspiration. When an aspirated stop is preceded by a vowel or a sonorant, preaspiration causes partial devoicing of the sonorant or vowel, and complete devoicing of voiced obstruents (see Table 5) and sometimes vowels, as shown in 1.4. In addition to this, the fricative /s/ is postaspirated. The low number of voiced consonants as well as the occurrence of aspiration noise at voiceless obstruents leads to devoicing of both consonants and vowels.

Informal listening to Mongolian speech gives us the impression that the younger generation becomes more and more 'voiceless'.

Table 6. Schwa epenthesis with reference to the voice feature of adjacent consonants. The material consisted of 236 word occurrences with 263 cases of expected schwa epenthesis.

context	*total*	*epenthesis*	*%*
$[C^+\_C^+]$	54	41	76
$[C^-\_C^+]$	122	68	56
$[C^+\_C^-]$	23	12	52
$[C^-\_C^-]$	64	4	6
total	263	125	47

2.3 Epenthesis word-internally

Epenthesis in word-internal syllables was investigated using a material of 236 word occurrences where insertion of one or more schwas were expected, following regular syllabification. The material was taken both from isolated utterances and from texts. Words with rhotics were not included, since it is difficult to decide whether the voiced fragment belongs to the consonant or to the inserted schwa. Different parts of the material were read by totally eight speakers (five women and three men). The results are presented in Table 6.

Most of the words were produced by two or more speakers. No deviations that could be regarded as individual were found, and all cases fitted into a systematic picture. This is true for the whole material used to investigate epenthesis.

Epenthesis occurs most often before voiced consonants, so that a combination CC^+ (i.e. C^+C^+ or C^-C^+) is split to form a new syllable with a schwa nucleus. For example, in /tʰimtgɮɮ/ → [tʰim.təg.ɮəɮ] 'note', the /tg/ cluster is split by [ə]. This agrees with the earlier established rules: a CC^+ coda is avoided due to its non-decreasing sonority, and epenthesis is triggered. The sonority law is thus active in speech. There are, however, many CC^+ codas which are not split by epenthesis. Such words can be divided into two homogeneous groups.

In the first group we find syllabification of the type /arɢɮ+ig/ → [arɢ.ɮig] (+ denotes morpheme boundary), with the coda /rɢ/. Cyclic syllabification would result in [ar.ɢə.ɮig] with the formation of an open [Cə] syllable with schwa nucleus. We found that [Cə] syllables are systematically avoided in our material (see below for other contexts) by all informants except one, Speaker ST, who syllabifies according to the rules established for formal speech. This speaker actually has a formal speaking style (she is a teacher of Mongolian for foreigners). If the other speakers do insert a schwa in this type of words, they avoid building [Cə] syllables, while Speaker ST syllabifies following the established rules. For example, she syllabifies /xʊwʲsɢɮ+iŋ/ as [xʊwʲs.ɢə.ɮiŋ], obeying the Resyllabification constraint and with an open syllable [ɢə]: *xʊwʲsɢ_ɮ* → *xʊwʲs.ɢəɮ* → *xʊwʲs.ɢəɮ* + *iŋ* → [xʊwʲs.ɢə.ɮiŋ], while speaker S pronounces [xʊwʲ.səɢ.ɮiŋ], splitting the *sɢ* cluster and forming only closed syllables.

How can we then interpret the 'irregular' syllabification of this type? If we assume that syllabification in casual speech is cyclic, as in formal speech, the irregular forms should be considered as violations of the Resyllabification constraint to avoid open syllables. An argument for this would be words formed cyclically, when no [Cə] syllable can be built. However, almost all words in our material would surface with the same syllable division irrespective of cyclicity. There is only one example: /paɮmt+ɮ+x/ is realised as [paɮ.mət.ɮəx] by all speakers in cyclic order (right-to-left direction), non-cyclic syllabification would result in [pa.ɮəm.təɮx]. The Resyllabification constraint is not violated here because no [Cə] syllables can be build. Notice that the word is realised in a frame sentence *pii paɮmətɮəx gisəŋ* 'I TARGET WORD said' and is pronounced in a careful way.

Alternatively we can interpret irregular forms with avoidance of open syllables as a result of non-cyclic syllabification. Again, only one example is found to confirm it. The word /tʰɔɢɮ+x+tɔ/ is realised by two speakers as non-cyclic [tʰɔ.ɢəɮx.tɔ] instead of cyclic [tʰɔɢ.ɮəx.tɔ].

An interesting detail is that consonant clusters, even with adjacent identical segments, are seldom simplified in our material, and all segments are fully realised. Such non-simplification of clusters is an important feature for syllabification since underlying phonemic consonant strings are fully realised. An example is given in Figure 2: the word /caaɮttgar/ is realised as [caaɬtʰtxar], not as formal [caaɮt.tə.gar] with an open [Cə] syllable. Simplifications of consonant clusters are very rare in the material.

In the second group with no split of CC^+ codas we found examples like [xot_ɮgoŋ], [xʊt_ɮtəgčʰ], [paac_ɮtgar], etc. (_ denotes non-insertion of a schwa which is predicted by rules for formal speech). The common features for these words is that the non-split coda is a combination of *t/c+ɮ*. From an articulatory point of view, the combinations *tɮ* and *cɮ* are homorganic gesture processes (both consonants are dentals). Stop + fricative combinations are similar to affricates, and affricate + fricative (*cɮ*) is articulatorily a combination of stop + fricative + fricative. It can be assumed that the clsters *tɮ* and *cɮ* are interpreted as articulatorily homogeneous units by Mongolian speakers, like

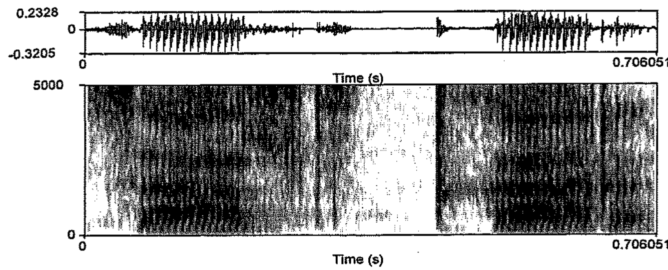


Figure 2. Oscillogram and spectrogram of [caaɬtʰtxar]. Both adjacent occurrences of the stop [t] are fully realised: two occlusions are observed.

the affricates, so that no schwa insertion is needed. It will be shown, that the interpretation of stop/affricate + fricative can help to explain differences in forming final syllables.

2.4 Word-internal codas

A maximal coda is found to contain three consonants, so that the first three consonants of a four-consonant cluster join into a coda, and the last one becomes an onset. This is true when all consonants surface as voiceless or when the first consonant of the cluster is voiced, e.g. /ɔŋɢcʰni/ → [ɔɴχcʰ.ni]. Consonants in such clusters combine independently of their place of articulation. It is however more difficult to make any statement about the role of the manner of articulation. None of the surface three-consonant codas in our material has only stops, and cases with adjacent stops are extremely rare. This partially depends on our material, and partially on the realisation of the frequent stops /g ɢ/ as fricatives before voiceless consonants, and, for example, /tgt/ is realised as [txt]. Codas with only fricatives are frequent.

The voice feature is assumed by us to have a crucial role in the syllabification of four-consonant clusters. For example, /paacɮtgar/ is syllabified either as [paa.cəɮt.gar] or as [paacɬt.gar] by the same speaker. In the first case, /ɮ/ is realised as voiced, causing the ill-formed cluster [$C^-C^+C^-.C^-$], and epenthesis is forced between C^- and C^+. In the second case, the three first consonants in the cluster are realised without voice, and no epenthesis is triggered.

The analysis of intervocalic CCCC as grouped into a complex coda and simple onset results from the Complex onset constraint in Mongolian. Some

Table 7. Epenthesis in word-final consonant clusters

context	*total*	*epenthesis*	*%*
[C$^+$_ C$^+$]	14	10	71
[C$^+$_ C$^-$]	24	22	92
[C$^-$_ C$^-$]	102	62	61
[C$^-$_ C$^+$]	32	32	100
all words	172	126	73

words are similar in that they have a complex consonant cluster ending with a combination *sn: CCsn*, e.g. [t^{h}an^{j}əɮc^{h}s_nar] with failed insertion of schwa within *sn*. Examples with more than four consonants in a cluster are also found. For example, /carc^{h}ʊɮgtsnig/ is realised as [carc^{h}ʊɬxtsnig] instead of regular [car.c^{h}ʊ.ɮəgt.sə.nig]. Such examples can be syllabified in two ways. Since *ts* is a combination of a dental stop and a dental fricative it is similar to the previously identified *tɮ*, analysed as one articulatorily homogeneous unit. With this assumption, the syllabification of /carc^{h}ʊɮgtsnig/ would be [car.c^{h}ʊɬxts.nig], with a four-consonant coda. Alternatively, we can analyse the combination *sn* as an onset: [car.c^{h}ʊɬxt.snig]. As the two dentals *s* and *n* are articulatorily homogeneous, analysing them as combining to an onset is sensible. This would force us to allow exceptions to the Complex onset constraint.

2.5 Formation of word-final syllables

The formation of word-final syllables demonstrates a more complicated picture than word-internal syllabification. We are conscious about the potential interplay with the following context, as well as the role of word frequency in the language. These aspects are, however, outside the scope of the present study, and, having them in mind, we describe only those processes that occur more or less systematically in the material, avoiding the unclear cases because of lack of material.

The same texts and utterances as for the investigation of word-internal epenthesis were used to analyse word-final clusters. Words with final consonant combinations that should be split by epenthesis, if syllabified regularly, were chosen. Forms with final consonant + nasal clusters are not included here, since they are the subject of a separate investigation (see below).

Word-final consonant clusters are split by epenthesis more frequently than non-final ones (Table 7). This is especially remarkable for combinations of voiceless consonants: as many as 61% of the voiceless clusters have schwa

Table 8. Word-final consonant clusters and occurrences of epenthesis within them. Only clusters that occur more than three times in the material are included. 166 cases of totally 172 are presented.

cluster	*realisation*	*total*	*epenthesis*
tx	[tx]	10	0
t^{h}ɮ/tɮ	[t^{h}ɮ,t^{h}ɬ, tɮ, tɬ]	8	0
nɮ	[nɮ, nɬ]	8	0
sx	[sx]	4	0
tg	[tg, tk]	48	48
ɢɮ	[ɢɮ, ɢɬ]	24	24
tt	[tt]	12	12
ɮg	[ɮg, ɬg, ɮk, ɬk]	8	8
tč	[tč]	4	4
xt	[xt]	4	4
xɮ	[xɮ, xɬ]	4	4
cg	[cg]	4	4
gč/ɢč	[xč/xč]	12	9
ɮx	[ɮx]	8	3
cx	[cx]	4	2
šx	[šx]	4	3

insertion, compared to only 6% word-internally (see Table 6). Taking the manner of articulation into account (see Table 8), some tendencies can be found.

The combinations *t^{h}ɮ/tɮ/nɮ* do not force epenthesis. All three are clusters of dental stops + *ɮ* similar to the earlier established *tɮ/cɮ* coda. Non-split *tx, sx* clusters, besides containing only voiceless segments, are also easily articulated, being similar to aspirated *t^h* and *s^h*. All these clusters can be interpreted as articulatorily homogeneous gestures and no insertion is needed.

The clusters *tg, cg, ɮg, xt, ɢɮ, xɮ*, which are split by epenthesis, contain heterorganic consonants. The clusters *tg, cg, ɮg, xt, tt* and *tč* have a stop (or affricate) as their second segment. Such C + stop/affricate clusters are more difficult to produce than stop + fricative, and they are more likely to force epenthesis. The systematic epenthesis in clusters ending with a stop is independent of voicing: insertion is forced even between voiceless segments. It was mentioned above that no simplification of consonant clusters was found word-internally and that all the segments in a cluster are fully realised by the speakers. The absence of consonant deletion is assumed here to trigger epenthesis in clusters like C + stop/affricate (*tt tč*): schwa insertion helps to make the second segment beginning with an occlusion to be heard.

Optional insertion is found in clusters fricative/affricate + *x*/*č* (in the combinations *gč*/*ɢč*, the stops are realised as a voiceless fricative [x]. The examples are too few to permit any conclusions, but it seems that voicing has no primary role in deciding epenthesis in these clusters: the same consonant combinations do not trigger epenthesis when they occur word-internally.

Summing up, we can indicate two features of epenthesis in word-final consonant clusters: word-final clusters are more frequently split by schwa than word-internal ones; both voicing and articulation have crucial roles in triggering epenthesis, but articulation has the primary role.

Although all aspects cannot be clarified based on our material, the differences found between occurrences of epenthesis word-internally and word-finally can be explained by two facts. Firstly, phonetic clusters similar to those which obligatorily trigger epenthesis word-finally do not occur in word-internal positions, and adjacent stops are rare. The low rate of [C+stop] clusters is a result of the realisation of the frequent stops /g, ɢ/ as a fricative [x] before voiceless segments. This kind of assimilation does not occur word-finally in our material, where [Cg] clusters are frequent and trigger epenthesis (see Table 8).

Secondly, the number of possible consonant combinations is greater within words, than word-finally. This can be exemplified with the word /tʰɔɢɮ+x+tɔ/, when realised as [tʰɔxɬxtɔ]. The division into syllables, following our analysis, is *tʰɔxɬx.tɔ*, combining *xɬx* into a three-consonant coda. This kind of syllabification is done on phonological grounds. The divisions into phonological and phonetic syllables need not overlap, and joining of *x* and *t* to an onset is sensible from a phonetic point of view: *xɬx.t* is articulatorily more likely chunked as *xɬ.xt*, with split of the uncomfortable combination of the three fricatives *xɬx*. We can assume that word-final *xɬx* would be split by epenthesis for articulatory reasons, so that we would find *tʰɔxɬəx*. The lower rate of epenthesis word-internally can thus be the result of greater choice of articulatory chunking: uncomfortable consonant clusters are split by schwa word-finally, while the same segments can be joined with a following consonant to form an articulatorily homogeneous onset word-internally.

2.6 Epenthesis in final CN(C) clusters

Mongolian has five nasal phonemes /m mʲ n nʲ ŋ/, and they occur both in onsets and codas, except that velar /ŋ/ occurs only in syllable codas. /ŋ/ has two allophones: velar [ŋ] in words with non-pharyngeal vowels, and uvular [ɴ] in words with pharyngeal vowels. Even in formal speech, /ŋ/ is realised as

strong nasalisation of a preceding phonemic vowel, with frequent loss of the nasal consonant. The rules for insertion of non-phonemic schwa before nasals seem to be different from the ones established here for other segmental contexts, and we will examine this in more detail. Henceforth, nasals will be denoted N.

A final CN(C) cluster violates the sonority law: since N is a voiced sonorant, a CN cluster does not have decreasing sonority, and schwa epenthesis is needed, so that the expected phonetic representation is [CəN]. When epenthesis occurs, the schwa becomes strongly nasalised before the following nasal. Schwa insertion is, however, often missing in our material: in 190 words with a final CN(C) cluster, occurring in isolated utterances and texts read by eight speakers (five women and three men) schwa was inserted in 86 cases (45%). For example, /saixŋ/ 'beautiful' and /ʊʊɮnt/ 'mountain-DAT' are realised without schwa insertion [saixŋ], [ʊʊɮnt] instead of regular [saixəŋ], [ʊʊɮənt].

The rate of schwa epenthesis in final CN(C) clusters is markedly lower than in final clusters ending with a voiced non-nasal consonant CC$^+$: 45% in CN(C)# clusters compared to 91% (42/46) in C _C$^+$# clusters with a non-nasal C$^+$.

Moreover, schwa is inserted more often in word-internal CN clusters than in word-final ones. 32 of the 236 words in the material contain word-internal CNC clusters between vowels, in which N cannot form the onset, and schwa occurs in 22 of these 32 cases (69%), while schwa is inserted only in 86 of 190 word-final clusters (45%)

This indicates that not only the nasality, but also the position of the CN cluster within the word affects schwa insertion. More material is needed to provide a certain explanation of this positional dependency of schwa epenthesis in CN clusters, however.

2.7 Epenthesis in CN clusters and phrasal prominence

In a previous investigation (Karlsson & Svantesson 2004) a dependency of schwa insertion on the degree of phrasal prominence is indicated. Focal gesture and word accent gesture are signalled in Mongolian by a tonal rise. Somewhat unexpectedly, schwa insertion occurs more frequently in words without any rising pitch gesture, in 72% (16 cases out of 22) of all non-accented words, but only in 29% of the accented words (19/66).

The dependency of schwa insertion on phrasal accentuation gets more evidence from our new material, spoken by one female informant. This material consists of pairs of frame sentences, with and without copula (*t^{h}iim in TARGET WORD pain* 'yes, this TARGET WORD is' and *t^{h}iim in TARGET WORD* 'yes, this TARGET WORD'), both said as answers to the question *in TARGET WORD uu?* 'is this TARGET WORD?'. Words with a final CN(C) cluster occur 102 times in the frame sentences. The target word is narrowly focused in both frame sentences, the difference is that in the first case the target word is placed before the final copula verb and gets high fundamental frequency (Figure 3, top), while in the second frame sentence the target word is placed finally, and gets lower F0 due to the sentence final position (see Figure 3, bottom). The mean F0 value of the word-final CN(C) clusters (with or without schwa insertion before N) for this speaker is 194 Hz for words in the position before the copula verb, and 177 Hz for words in final position (154 Hz if we choose the absolute final measurement point within the CN(C) cluster). Schwa is inserted in 10 cases of 51 (20%) in the position before the copula verb, but in 42 cases of 51 (82%) in phrase-final position.

The data indicates that non-insertion of schwa before a nasal occurs in words marked with rising gesture, which signals focus and word accent. When there are no such gestures, epenthesis is frequent. For instance, in phrase final position, even when focused, words get a strong falling tonal gesture to low frequencies, signalling phrase boundary. In such cases, schwa is frequently inserted before the nasal, see Figure 3 (bottom) for exemplification.

2.8 Analysis

The frequent non-insertion of schwa in CN clusters violates the sonority law. The high sonority of nasals often leads to their cooccurrence with the feature [+syllabic] (Cohn 1993). Our data suggests that the analysis of Mongolian nasals as having the feature [+syllabic] is the most appropriate one. Consider for example surface forms like [ɔrtnt, icnt^{h}, ontgnt] (regular syllabification should be [ɔr.tənt, i.cənt^{h}, ont.gənt]). Assuming that the nasals are non-syllabic would result in the syllable divisions [ɔrtnt, icnt^{h}, ontgnt], i.e. as monosyllables with complex codas. No codas with the combinations $(C^{+})C^{-}C^{+}(C^{+})C^{-}$ were attested in our previous data, and they seem to be highly ill-formed. Assuming that the nasals are syllabic, we get a much more natural syllabification: [ɔr.tnt, i.cnt^{h}, ont.gnt] with the nasal as the nucleus of the second syllable. A rule can be established: a nasal becomes a nucleus between consonants or between a consonant and a word boundary. Following the earlier analysis of nuclei as moraic in Mongolian (Karlsson 2003), we have to consider these nasals as moraic segments, like vowels.

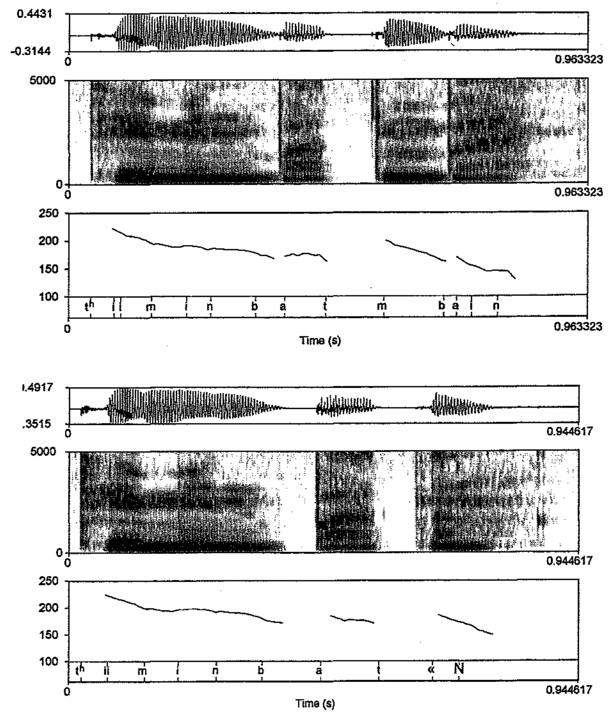


Figure 3. The utterances [t^{h}iim in batm bain] 'yes, this Badan (proper name) is' (top), and [t^{h}iim in batəŋ] 'yes, this Badan' without copula (bottom). /ŋ/ is assimilated to the following bilabial [b] in the first utterance, gets focal H and no schwa is inserted. In the second utterance, [batəŋ] in final position gets a low tone (F0 falls to under 150 Hz) and schwa is inserted.

Our data suggests a dependency of non-insertion of schwa before nasals on rising tonal gestures. Schourup 1973 shows a correlation between stress and nasalisation. Krakow 1987, investigating influences of suprasegmental events on velic movements, found that her subjects kept the velum lowered for a longer time in stressed than in unstressed syllables beginning or ending with a nasal consonant. By equating stress and prominence, we can assume that we find a similar process in Mongolian: lowering of the velum in nasals is more active in words marked by rising tonal gestures. This hinders schwa insertion. Furthermore, non-insertion is favoured by the syllabic function of nasals.

3 Summary

We have investigated syllabification using a material consisting of isolated utterances and texts. The point of departure was syllabification rules established on the basis of formal speech. Two basic rules, the Sonority law and Right-to-Left syllabification are active in both formal and non-formal speech. Whether syllabification in casual speech is cyclic or non-cyclic cannot be decided from our material. However, the frequent violation of the Resyllabification constraint to avoid formation of Cə suggests that syllabification in casual speech is non-cyclic.

The Sonority law is active but is often overridden in casual speech by articulatory effects, and the codas *tʒ/cʒ* with increasing sonority are allowed. Voiceless segments have unmarked sonority value, and three-consonant codas with only voiceless segments are allowed in casual, but not in formal speech. Only simple onsets are found in formal speech, while the two-consonant onset *sn*, supposedly the result of articulatory accommodation occurs in our material

Epenthesis before nasals is governed by other rules than before non-nasal segments, due to the occurrence of syllabic nasals in casual speech.

Acknowledgements

I would like to thank Jan-Olof Svantesson and Klaus Kohler for their constructive comments on earlier versions of this paper. Any errors remain my own.

References

Cohn, Abigail C. 1993. 'The status of nasalized continuants'. In M. Huffman & R. Krakow (eds), *Phonetics and phonology* 5: *nasals, nasalization, and the velum,* 329-67. San Diego: Academic Press.

Itô, Junko. 1989. 'A prosodic theory of epenthesis'. *Natural language and linguistic theory* 7, 217-59.

Jespersen, Otto. 1897-99. *Fonetik: en systematisk fremstilling af læren om sproglyd.* København: Det Schubotheske forlag.

Karlsson, Anastasia M. 2003. 'Rhythmical and accentual structure of Mongolian'. In M.-J. Solé, D. Recasens & J. Romero (eds.), *Proceedings of the 15th International congress of phonetic sciences,* 2465-67. Barcelona: Universitat Autonòma de Barcelona.

Karlsson, Anastasia M. & Jan-Olof Svantesson. 2002. *Preaspiration in Halh Mongolian.* Paper presented at the Workshop on Mongolic languages, Taibei, September 2002.

Karlsson, Anastasia M. & Jan-Olof Svantesson. 2004. 'Prominence and Mora in Mongolian'. *Proceedings of International Conference on Speech Prosody, March 23-26, Nara, Japan*, 65-68.

Kohler, Klaus J. 1991. 'The phonetics/phonology issue in the study of articulatory reduction'. *Phonetica* 48, 180-92.

Krakow, Rena A. 1987. *Stress effects on the articulation and coarticulation of labial and velic gestures.* Paper presented at the meeting of the American Speech-Language-Hearing Association, New Orleans, LA.

Poppe, Nikolaus. 1951. *Khalkha-mongolische Grammatik.* Wiesbaden: Steiner.

Schourup, A. 1973. 'A cross-language study of vowel nasalization'. *Ohio State University Working papers in linguistics* 15, 190-221.

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

Anastasia Mukhanova Karlsson <Anastasia.Karlsson@ling.lu.se>