Northern Language Studies 14: 51-63 (Japan Association of NoLS, 2024)

Vowel Devoicing in Khalkha Mongolian

Effects of Vowel Type and Phonological Environment

Naoki UETA
(Hokuyo University)

Keywords: vowel devoicing, Khalkha Mongolian, aspiration, vowel height, phonological condition

1. Introduction

Vowel devoicing is a phonetic phenomenon whereby a vowel that is normally voiced is pronounced as voiceless. Vowel devoicing is frequently observed cross-linguistically; for example, in Standard Japanese, high vowels /i/ and /u/ are normally devoiced between voiceless consonants, as shown in (1).

(1) a. /kita/ [kʲi̥ta] “north”
b. /kusa/ [ku̥sa] “grass”

High vowel devoicing is a rather common phenomenon across languages due to the fact that high vowels are less sonorous and generally shorter than other vowels (Labrune 2012: 38).

In Khalkha Mongolian (henceforth Mongolian), which is widely spoken in Mongolia and considered Standard Mongolian, vowels are occasionally devoiced. Vowel devoicing in Mongolian is caused by the following aspirated consonant; the aspiration of an aspirated consonant is realized as preaspiration, making the preceding vowel breathy or devoiced (Svantesson and Karlsson 2012, Ueta 2020a).

However, it is still unclear under what conditions and how frequently vowel devoicing occurs in Mongolian because limited basic observation and acoustic analysis have been conducted. The present study observes the frequency of vowel devoicing and phonological conditions in Mongolian through a production experiment and an acoustic analysis.

2. Aspiration Contrast and Vowel Devoicing in Mongolian

2.1 Aspiration Contrast

Mongolian has a laryngeal contrast in stops and affricates: strong (tense, fortis) vs. weak (lax, lenis). In Cyrillic letters and their transliterations, the strong and weak consonants are represented as п, т, к, ц, ч (*p, t, k, ʦ, ʧ*) and б, д, г, з, ж (*b, d, g, ʣ, ʤ*), respectively. The two series are basically differentiated by the presence or absence of aspiration

(Svantesson et al. 2005, Karlsson and Svantesson 2011, 2012, Svantesson and Karlsson 2012). The strong consonants are phonologically regarded as aspirated /pʰ, tʰ, kʰ, ʦʰ, ʧʰ/, while *d, ʣ, ʤ* are interpreted as voiceless unaspirated /t, ʦ, ʧ/. However, the weak velar (and uvular) stop *g* is regarded as voiced /g/ (and /ɢ/), based on the phonological behavior (Svantesson et al. 2005). The weak bilabial stop *b* can also be regarded as voiced /b/ because it is frequently realized as voiced, although Svantesson et al. (2005) regard it as /p/ (see Ueta 2020b for details).

The strong and weak consonants at the word-initial position are distinguished by voice onset time (VOT); shorter VOT for unaspirated consonants and longer VOT for aspirated consonants (Ueta 2018). In word-medial and word-final positions, the strong and weak consonants are distinguished by the presence or absence of preaspiration; the strong consonants are accompanied by aspiration during the preceding segment, resulting in the preceding segment becoming breathy or devoiced (Svantesson and Karlsson 2012, Ueta 2020a).

In contrast, there is no laryngeal contrast in the fricatives /s, ʃ, x/. Although little acoustic investigation has been done, and their phonetic properties, particularly VOT values, are still unknown, these are typically regarded as strong consonants.

2.2 Vowel Devoicing

As mentioned above, the word-medial and word-final aspirated consonants in Mongolian have preaspiration, which can, though not necessarily, cause devoicing of the preceding vowel.

(2) a. *butex* /butʰex/ [bu̥tĕx] ~ [b̥u̥tĕx] “succeed”
b. *gaʦax* /gaʦʰax/ [ɢɑ̯ʦɑ̆χ] ~ [ɢ̥ɑ̯ʦɑ̆χ] “interfere”

Ueta (2020a) points out that the phonetic realization of short vowels followed by an aspirated consonant has several variations and differs among speakers, as shown in Figure 1.

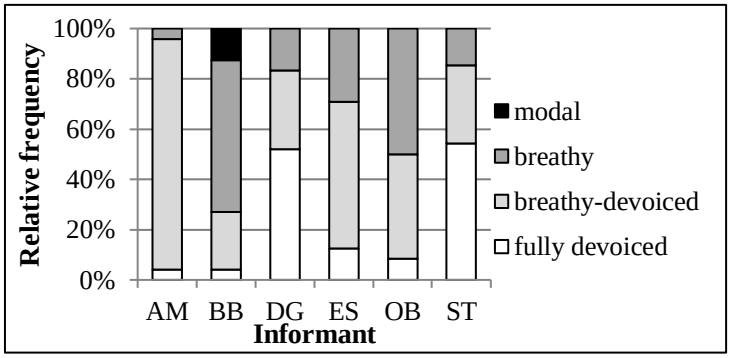

Figure 1. Phonetic variations of short vowels followed by an aspirated consonant (Ueta 2020a: 106, Figure 6)

In conclusion, aspirated consonants can devoice Mongolian short vowels. It is unclear, however, under what conditions and how frequently vowel devoicing occurs and how vowel devoicing is phonologically interpreted in Mongolian because little attention has been paid to the phenomenon and even basic analysis has not been conducted. More specifically, the following questions remain unsolved:

(3) a. Does vowel devoicing consistently occur in most speakers when conditions are fulfilled? In relation to this, is vowel devoicing an important acoustic feature to differentiate between aspirated and unaspirated consonants, or just a redundant phonetic phenomenon?
b. Does the consonant preceding a vowel affect vowel devoicing? In other words, is a vowel more frequently devoiced between two aspirated consonants than between an aspirated and unaspirated consonant?
c. Do fricatives, which are regarded as strong consonants, cause vowel devoicing in the same way as aspirated stops and affricates? If so, does the frequency of vowel devoicing differ depending on whether the following consonant is a stop or affricate, which has a laryngeal contrast between aspirated and unaspirated, or a fricative, which does not have the contrast?
d. Does the frequency of vowel devoicing vary depending on the type of vowel? More specifically, are high vowels more frequently devoiced than other types of vowels?

The purpose of this study is to answer these questions through a production experiment and an acoustic analysis.

3. Survey

3.1 Production Experiment

A production experiment was conducted to clarify the conditions and frequency of vowel devoicing. Hereafter, /p^h, t^h, k^h, $ʦ^h$, $ʧ^h$/ and /b, t, g, ʦ, ʧ/ are represented as *p, t, k, ʦ, ʧ* and *b, d, g, ʣ, ʤ*, respectively.

The target words are 25 two-syllable words (#C_1VC_2-); the first vowel (V) is one of the seven Mongolian short vowels /i, u, e, ʊ, ɵ, ɔ, a/; the consonant following the vowel (C_2) is either an aspirated consonant (*t, ʧ*) or a fricative (*s, ʃ, x*). C_1 is either *t, ʧ, s, ʃ, x, d,* or *ʣ*. The target words in which C_1 is an unaspirated consonant (*d, ʣ*) constitute minimal pairs with those in which C_1 is an aspirated one (*t, ʦ*); for example, *teslegʧ* "detonator" vs. *deslegʧ* "lieutenant." The target words are summarized in Table 1.

Table 1. Target words

Vowel	C_2 = aspirated	C_2 = fricative	
	C_1 = aspirated or fricative		C_1 = unaspirated
i	*titan* “titanium” *xitʃeel* “class”	*ʧixer* “candy”	
u	*ʃuteen* “Buddhist statue”	*tuʃig* “support” *xusel* “hope”	
e	*setgel* “heart”	*teslegʧ* “detonator” *xeseg* “parts”	*deslegʧ* “lieutenant”
ʊ	*xʊtaɢ* “knife”	*tʊslax* “to help” *xʊsam* “scale”	*dʊslax* “to drip”
ɵ	*xɵtlɵx* “to lead”	*xɵʃig* “curtain” *xɵxuur* “leather sack for making kumis”	
ɔ	*xɔtɔix* “to sag”	*ʦɔxix* “to hit” *xɔʃʊʊ* “muzzle”	*ʤɔxix* “to suit”
a	*xatʊʊ* “hard”	*taxix* “to worship” *xasax* “to subtract”	*daxix* “to repeat”

The target words were inserted in the carrier sentences (4), and the whole sentences were read out by native Mongolian speakers. Each target word and two carrier sentences were displayed on a computer screen in Cyrillic orthography.

(4) a. ..... *gedeg nʲ juu we?* “What is .....?”
b. *bi* ..... *geʤ xelsen* “I said .....”

The informants were 14 (8 male and 6 female) native Mongolian speakers, whose ages ranged from 17 to 23. The reading task was repeated three times per informant. Therefore, 2100 tokens (25 words*2 sentences*14 informants*3 times) were obtained. The recording was conducted with a digital recorder (ZOOM H4nPro) and a head-mounted microphone (AKG C520).

3.2 Acoustic Analysis

All recorded material was analyzed with Praat (Boersma and Weenink 2021). Whether the vowels were devoiced was judged by observing wave forms, spectrograms, and pitch curves and by listening to each sound. Although vowels can be partially devoiced, only full devoicing cases were regarded as “devoiced” in this study.

4. Results and Discussion

4.1 Devoicing Rate in Each Informant

First, the rates of vowel devoicing for each informant are shown in Figure 2. This figure demonstrates that vowel devoicing rates differ significantly among informants; informant BB devoiced only 1.3% of the vowels, while informant BO devoiced 76.0% of them. It is impossible to explain the difference among informants from the perspective of gender or regional dialect; most informants were born in Ulan Bator, the capital of Mongolia. The difference in devoicing rates may be attributed to speech rate and voice quality, both of which are regarded as individual differences. This fact suggests that vowel devoicing in Mongolian is not indispensable for realizing the feature of aspiration but just a phonetic phenomenon that can be optionally observed.

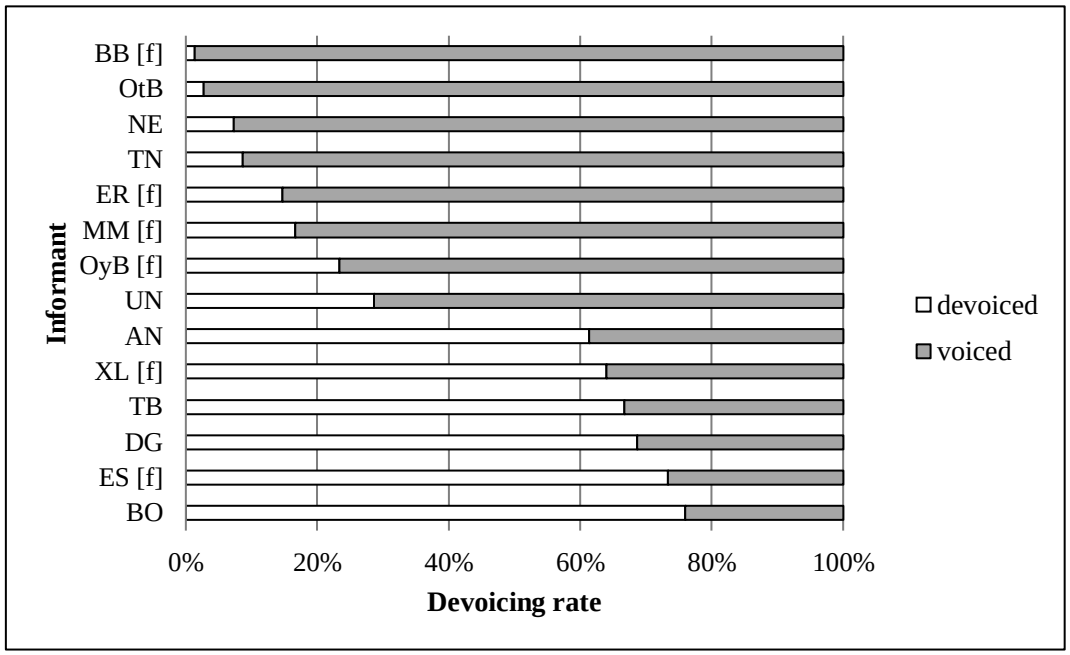

Figure 2. Vowel devoicing rate for each informant ([f] represents female)

4.2 Devoicing Rate for Each Target Word

Figure 3 shows the rates of vowel devoicing for each target word. It is clear from Figure 3 that devoicing rates significantly differ depending on words, although all target words can be devoiced. In particular, devoicing rates are considerably low when the word-initial consonant is unaspirated. This result suggests that not only the following but also preceding consonant of a vowel affects vowel devoicing, as elaborated in Section 4.3. Figure 3 also suggests that the vowels /i/ and /u/, namely high vowels, are frequently devoiced. The relationship between vowel type and devoicing rate is more precisely analyzed in Section 4.5.

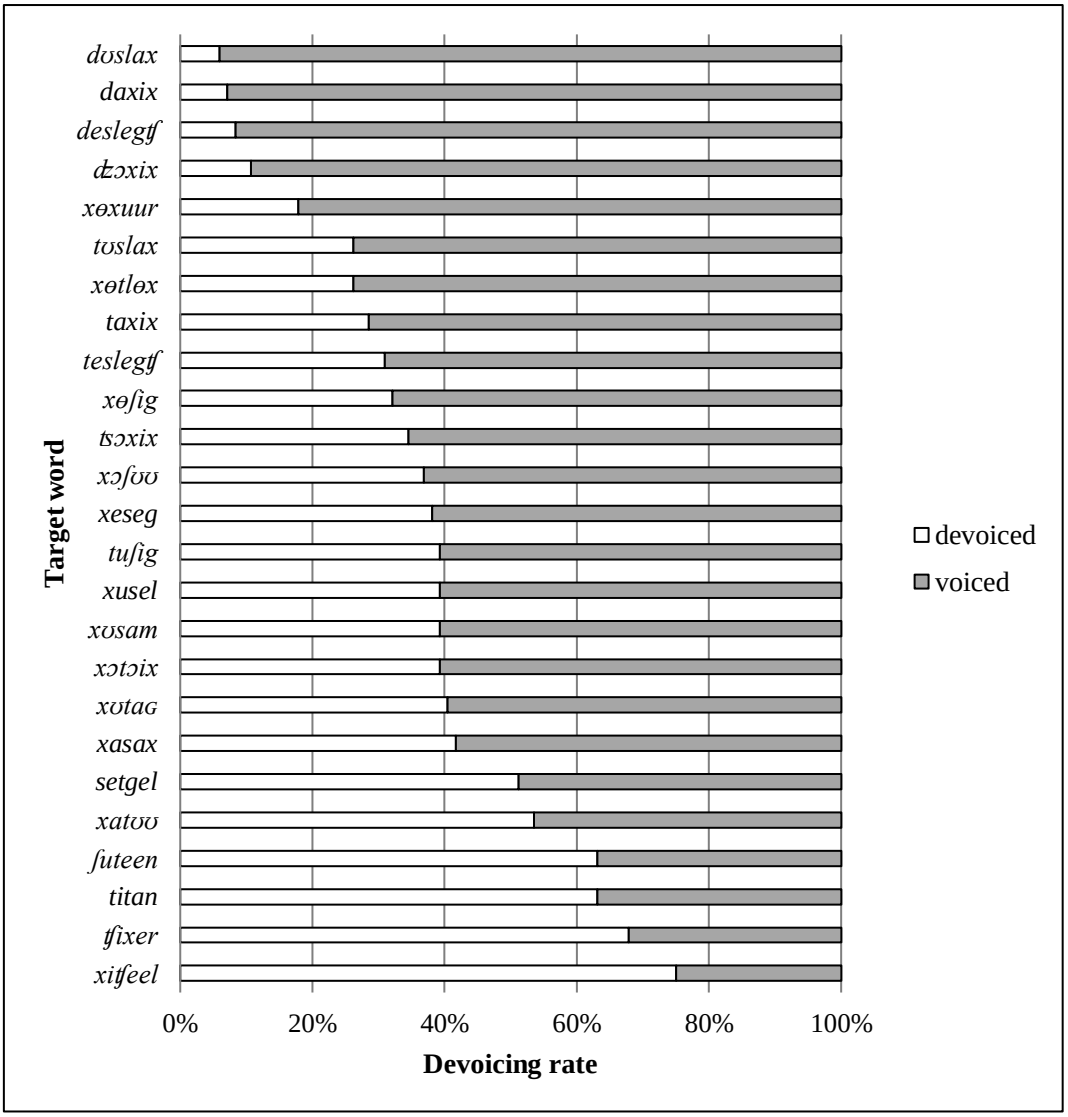

Figure 3. Vowel devoicing rate for each target word

As is clear from Figure 3, vowel devoicing does not necessarily occur even in words where the vowel is frequently devoiced; the devoicing rate is 75.0% in *xitʃeel*, whose vowel is most frequently devoiced. In addition, a speaker can pronounce the same word with and without vowel devoicing; Figure 4 shows wave forms, spectrograms, and pitch curves of the target word *xeseg* by an informant. The left and right figures are the sounds in the carrier sentences (4a) and (4b), respectively. Although these are successively read, the vowel is not devoiced in (4a) and is devoiced in (4b). The difference in devoicing rates between the two carrier sentences is addressed in Section 4.6.

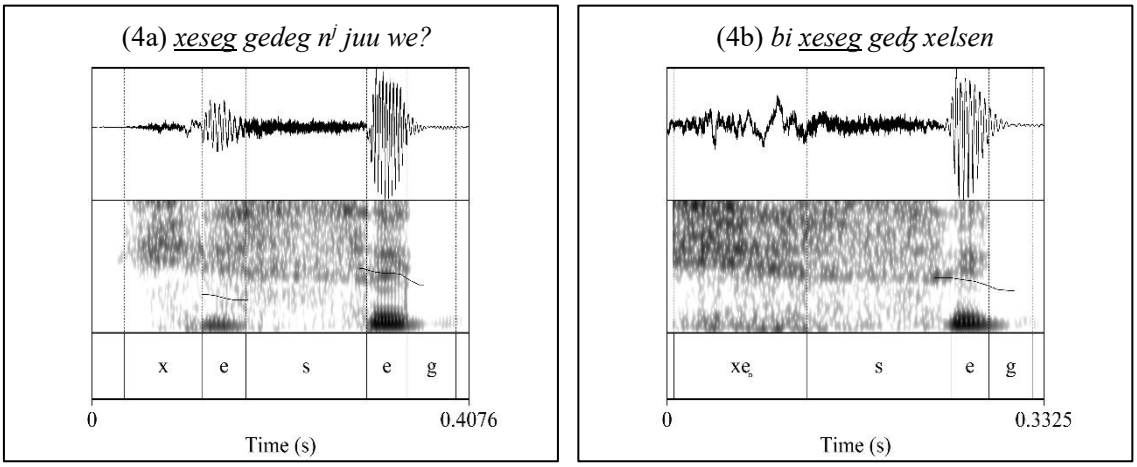

Figure 4. Wave forms, spectrograms, and pitch curves of *xeseg* pronounced by an informant (left: in (4a), right: in (4b))

4.3 Effect of the Preceding Consonant on Vowel Devoicing

As seen in the previous section, vowels do not tend to be devoiced when the preceding consonant is an unaspirated consonant (*d* or *dʑ*). Three different types of preceding consonants—unaspirated stop and affricate, aspirated ones, and fricatives—show the vowel devoicing rates in Figure 5, which reveals that only 8.0% of vowels are devoiced when the preceding consonant is *d* or *dʑ*; the rate is much lower compared to that for other consonants. In minimal pairs based on word-initial aspirated and unaspirated contrast, the devoicing rates significantly differ, as shown in Table 2. These facts suggest that not only the following but also the preceding consonant affects vowel devoicing.

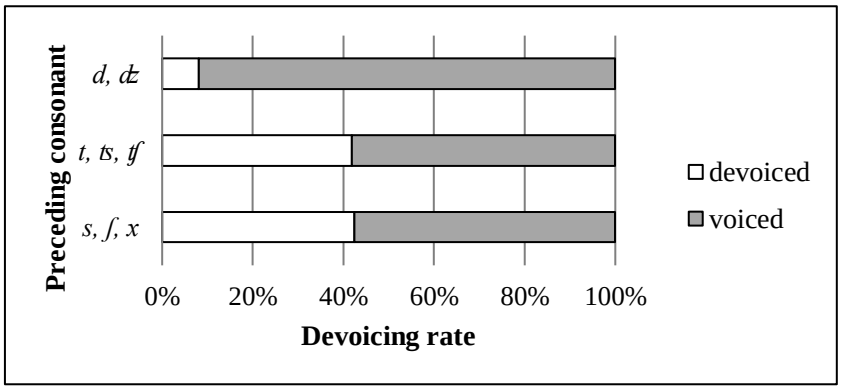

Figure 5. Vowel devoicing rate for each preceding consonant

Table 2. Devoicing rates in minimal pairs

Unaspirated	Devoicing rate	Aspirated	Devoicing rate
deslegʧ	8.3% (7/84)	*teslegʧ*	31.0% (26/84)
dʊslax	6.0% (5/84)	*tʊslax*	26.2% (22/84)
ʤɔxix	10.7% (9/84)	*ʦɔxix*	34.5% (29/84)
daxix	7.1% (6/84)	*taxix*	28.6% (24/84)

4.4 Effect of the Following Consonant on Vowel Devoicing

This section focuses on the type of consonant following a vowel and investigates whether devoicing rates differ between when the following consonant is a stop or affricate and when it is a fricative. Figure 6 shows vowel devoicing rates in terms of the two groups of consonants following a vowel. The results for the target words *deslegʧ*, *dʊslax*, *ʤɔxix*, and *daxix* are excluded because it has been clarified in the previous section that vowels in these words are not normally devoiced due to the preceding unaspirated consonant.

First of all, it is worth noting that a vowel can be devoiced when the following consonant is a fricative. This means that fricatives are interpreted as aspirated consonants in Mongolian in the sense that they can make the preceding vowel devoiced in the same fashion as aspirated stops and affricates, although there is no laryngeal contrast in fricatives.

Figure 6 demonstrates that the devoicing rate is lower when the consonant is a fricative (36.3%) than when it is a stop or affricate (51.5%). This difference is statistically significant (chi-squared test, $p < .001$). The difference may be attributed to the strength of aspiration; the aspiration of aspirated stops and affricates is supposed to be stronger than that of fricatives because stops and affricates have unaspirated counterparts in the Mongolian consonant system, and thus aspiration is an important cue to differentiate aspirated from unaspirated sounds, whereas fricatives do not. In other words, even though the phonological feature for aspirated stops, affricates, and (aspirated) fricatives is the same, the phonetic realization of the feature may be different between them[1].

[1] To clarify whether this hypothesis is true, not only the frequency of full devoicing but also the extent of vowel devoicing should be examined, which is a future challenge.

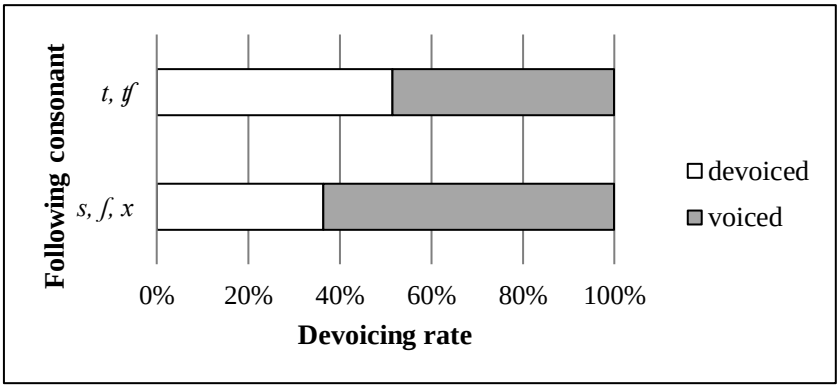

Figure 6. Vowel devoicing rate for each following consonant

4.5 Relationship between Vowel Type and Devoicing

In this section, the difference in devoicing rates depending on the sort of vowel is explored. Figure 7 shows the devoicing rate for each vowel. Here, again, the results for the target words *deslegʧ*, *dʊslax*, *ʤɔxix*, and *daxix* are excluded.

It is clear from Figure 7 that the vowels *i* and *u* are the first and second most frequently devoiced vowels, respectively. This result is natural and corresponds to the universal tendency; high vowels are frequently devoiced. The high vowels *i* and *u* in Mongolian are frequently devoiced, probably because they have a short duration and are less sonorous than other vowels.

However, *i* has been even more frequently devoiced than *u* in this survey. This might be characteristic of Mongolian and be related to the fact that the vowel *i* in Mongolian has no place feature. According to Svantesson et al. (2005) and Ueta (2019), the phonological structure of Mongolian vowels is represented as in (5)[2].

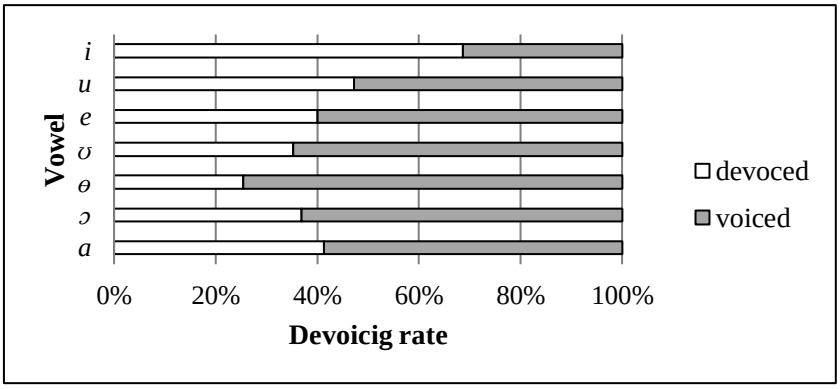

Figure 7. Devoicing rate for each vowel

[2] The phonological representation is justified by vowel harmony. See Svantesson et al. (2005) and Ueta (2019) for details.

(5)

non-pharyngeal vowels	pharyngeal vowels
i [ ]	
u [R]	ʊ [FR]
e [O]	a [FO]
ɵ [OR][3]	ɔ [FOR]

([R]: round, [O]: open, [F]: pharyngeal)

The vowel *i* has neither [R], [O], nor [F], and thus it is regarded as unmarked. This phonological characteristic of *i* can lead to a high rate of vowel devoicing.

Figure 7 also shows that all vowels, including non-high vowels, can be devoiced at any rate. It is noteworthy that the vowel *a* is the third most frequently devoiced despite being a low vowel, demonstrating that vowel devoicing frequency is not solely dependent on vowel height. This implies that high and low vowels in Mongolian are devoiced based on different mechanisms; high vowels are devoiced due to their short duration and less sonority, whereas low vowels are devoiced because the bulk of the oral cavity and expiratory flow rate become large in pronouncing a low vowel, and thus the glottis tends to open widely, resulting in devoicing as an extension of aspiration and/or breathiness. In view of the fact that the realization of preaspiration in Mongolian is consecutive, namely, breathy > partially devoiced > full devoiced (Ueta 2020a), this interpretation seems plausible.

Another remarkable result is that the devoicing rates for *ɵ*, *ʊ*, and *ɔ*, all of which are rounded vowels, are relatively low. This suggests that the feature [R] is likely to function as an inhibitor of vowel devoicing in Mongolian. This explanation holds for the rounded high vowel *u* being less frequently devoiced compared to the unrounded high vowel *i*. Among the rounded vowels, *ɵ* is the least frequently devoiced. This is probably because *ɵ* is phonetically a rounded central vowel, being regarded as marked[4]; marked sounds are assumed to be carefully pronounced and to retain the phonetic characteristics.

4.6 Relationship between Carrier Sentence and Devoicing

In the production experiment, two carrier sentences were utilized, as shown in (4). They differ in where target words are inserted: the utterance initial position in (4a) and the non-initial position in (4b). In addition, the informants read sentences in the order of (4a) > (4b). This difference could affect the vowel devoicing rate.

[3] Although Svantesson et al. (2005) regard *ɵ* as /o/, Ueta (2019) claims that *ɵ* is phonetically a central vowel and should be regarded as /ɵ/.

[4] As shown in (5), /ɵ/ has only the feature [OR], and its markedness is not expressed in the phonological representation. It may need to be rewritten based on the low rate of devoicing and its phonetic characteristics as a central vowel.

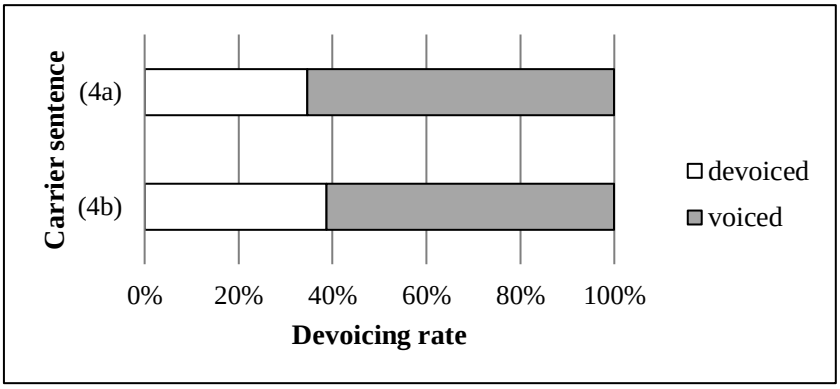

Figure 8. Devoicing rate for each carrier sentence

Figure 8 shows the devoicing rates in each sentence, revealing that the devoicing rate in (4b) is higher than in (4a), and the difference is statistically significant (chi-squared test, $p < .05$). However, the difference is small, and the tendency varies depending on informants; vowels are more frequently devoiced in (4b) than in (4a) for seven informants, whereas the results are opposite for six informants, and the rates are tied for an informant. This indicates, again, that vowel devoicing in Mongolian is just a phonetic phenomenon occurring not necessarily but optionally, and the phonological environment where vowel devoicing frequently occurs differs among speakers.

5. Conclusion

This study observed the relationship between the frequency of vowel devoicing and phonological conditions in Mongolian through a production experiment and an acoustic analysis. This section summarizes the results of the survey and answers the questions (3) provided in Section 2.2.

Firstly, vowel devoicing in Mongolian does not consistently occur even when conditions are fulfilled; the vowel devoicing frequency significantly differs among speakers, and devoicing does not necessarily occur consistently even for the same word pronounced by the same speaker. In this sense, vowel devoicing in Mongolian is considered to be a redundant phenomenon motivated phonetically.

Secondly, the consonant preceding a vowel affects vowel devoicing; vowels are rarely devoiced when preceded by an unaspirated consonant. In other words, vowels are even more frequently devoiced between two aspirated consonants than between an unaspirated and aspirated consonant.

Thirdly, fricatives also cause vowel devoicing in the same way as aspirated stops and affricates. In this sense, fricatives in Mongolian are regarded as aspirated. However, vowel devoicing rates differ between stops, affricates, and fricatives; vowels followed by an aspirated stop or affricate are more frequently devoiced than ones followed by a fricative. The difference probably comes from the presence or absence of phonological

contrast. Stops and affricates have a laryngeal contrast between aspirated and unaspirated, and thus aspiration is an important feature to distinguish them, resulting in that aspiration becomes relatively strong and causes vowel devoicing frequently, whereas fricatives do not have the contrast. Furthermore, aspiration is redundant, leading to the relatively low rate of vowel devoicing.

Lastly, the frequency of devoicing varies depending on the type of vowel; *i* is most frequently devoiced, which corresponds to the cross-linguistic tendency. However, in Mongolian, the low vowel *a* is also frequently devoiced, probably based on a different mechanism than in high vowels. In contrast, rounded vowels, especially *ɵ*, tend to resist devoicing because the feature [round] blocks devoicing, and *ɵ* is a marked vowel.

This study revealed the basic conditions and frequency of vowel devoicing. However, several factors remain unanalyzed; for example, it is possible that vowel devoicing rates differ depending on the speaker's generation. In addition, physiological studies are needed to prove the hypothesis that the devoicing of high and low vowels is based on different mechanisms.

Acknowledgments

This paper is a revised version of the oral presentation "Vowel Devoicing in Khalkha Mongolian" presented at "The 6th Annual Meeting of Japan Association of Northern Language Studies" (in Japanese). I want to thank the participants as well as the two anonymous reviewers who gave helpful and constructive comments. I also thank the students and teachers at the Mongolian University of Science and Technology for their cooperation during the experiment. This research was supported by JSPS KAKENHI Grant Numbers 21K20015 and 23K12168.

References

Boersma, Paul and David Weenink (2021) Praat: Doing phonetics by computer (ver. 6.1.50). https://www.fon.hum.uva.nl/praat/

Karlsson, Anastasia and Jan-Olof Svantesson (2011) Preaspiration in Mongolian dialects: Acoustic properties of contrastive stops. *Paper presented at the 10th Seoul International Altaistic Conference*: 125-140.

Karlsson, Anastasia M. and Jan-Olof Svantesson (2012) Aspiration of stops in Altaic languages: An acoustic study. *Altai Hakpo* 22: 205-222.

Labrune, Laurence (2012) *The phonology of Japanese*. Oxford: Oxford University Press.

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

Svantesson, Jan-Olof and Anastasia Karlsson (2012) Preaspiration in modern and old Mongolian. *Suomalais-Ugrilaisen Seuran Toimituksia* 264: 453-464.

Ueta, Naoki (2018) Voice onset time of word-initial stops and affricates in Khalkha Mongolian. *Journal of the Phonetic Society of Japan* 22 (2): 131-140.

Ueta, Naoki (2019) *Mongorugo no boin: Jikken onseegaku to shakuyoogo on-inron kara no apuroochi* [Monglian vowels: Approaches from experimental phonetics and loanword phonology]. Kyoto: Kyoto University Press.

Ueta, Naoki (2020a) The acoustic realization of preaspiration in Khalkha Mongolian. *Northern Language Studies* 10: 99-115.

Ueta, Naoki (2020b) Mongorugo haruha hoogen no gochuu heesaon no onseeteki barieeshon to on-in kaishaku [Phonetic variations and phonological interpretations of word-medial stops in Khalkha Mongolian]. *Bulletin of the Japan Association for Mongolian Studies* 50: 1-18.

Summary

Vowel devoicing is a common phenomenon across languages. Vowels can be devoiced in Khalkha Mongolian by the following aspirated consonant; aspiration is realized as preaspiration, rendering the preceding vowel breathy or devoiced. However, it is unknown under what conditions and how frequently vowel devoicing occurs in Mongolian. This study examined the relationship between vowel devoicing frequency and phonological conditions and analyzed how vowel devoicing can be interpreted in Mongolian.

The production experiment and acoustic analysis revealed the following facts. Firstly, vowel devoicing in Mongolian does not occur consistently even when conditions are fulfilled, meaning that it is a redundant phenomenon motivated phonetically. Secondly, the consonant preceding a vowel affects vowel devoicing; vowels are even more frequently devoiced between two aspirated consonants compared to when the preceding consonant is unaspirated. Thirdly, fricatives also cause vowel devoicing in Mongolian, indicating that they can be regarded as aspirated, although the vowel devoicing rate is lower when the following consonant is a fricative than when it is an aspirated stop or affricate. Lastly, the frequency of devoicing varies depending on the vowel type; *i* is the most frequently devoiced and *ɵ* is the least frequently devoiced, probably because *i* is a high and unmarked vowel, while *ɵ* is a marked vowel in Mongolian. In addition, the low vowel *a* is also frequently devoiced, suggesting that high and low vowels can be devoiced based on different mechanisms.

(ueta.naoki.82x@gmail.com)