UNIVERSITY OF CALIFORNIA

Los Angeles

Word-initial /p/ in Khalkha Mongolian: Variation in Connected Speech

A thesis submitted in partial satisfaction

of the requirements for the degree Master of Arts in Linguistics

by

Corrina Lee Fuller

2024

ABSTRACT OF THE THESIS

Word-initial /p/ in Khalkha Mongolian: Variation in Connected Speech

by

Corrina Lee Fuller

Master of Arts in Linguistics
University of California, Los Angeles, 2024
Professor Kie Ross Zuraw, Chair

This thesis explores the variable realization of word-initial /p/ in Khalkha Mongolian through the lens of connected speech. While phoneme /p/ has two phonologically conditioned allophones word-internally ([p] and [w]), word-initial /p/ does not alternate in careful speech. In connected speech, however, not only is word-initial /p/ variably realized as [p] or [w], it is at times deleted altogether, yielding a "zero" realization. In this study, I offer a characterization of the nature of the variable realization of word-initial /p/ through careful investigation of a small corpus of naturalistic speech (four Khalkha-language TEDx Talks). Then, I model the variation illuminated within the corpus data using a Maximum Entropy grammar (MaxEnt; Goldwater &

Johnson, 2003), a probabilistic version of Optimality Theory (Prince & Smolensky, 1993) that generates probability distributions over all candidates. I argue that the variation in the realization of word-initial /p/ is patterned, and that processes related to ease of articulation in connected speech motivate lenition, while processes particular to certain morphosyntactic configurations condition reduction.

The corpus data show a strong preference for the [p] realization, followed by a slight preference for the zero realization, with [w] being the least preferred variant of word-initial /p/. The results also reveal a massive split between content and function words, with the less faithful realizations of /p/ ([w] and zero) occurring almost exclusively in function words. However, the function words do not all behave as one. Markedness constraints favoring the [w] and zero realizations reflect the articulatory pressure to lenite in different phonological environments, while lexically-indexed faithfulness constraints represent morpheme-specific resistance to this pressure. Finally, a constraint encoding a structurally-conditioned listed allomorph of auxiliary *pai-* captures a case of auxiliary reduction.

The lens of connected speech affords insights into the interplay of articulatory pressures affecting lenition and the morphosyntactic configurations that select phonetically-reduced allomorphs, both of which affect Khalkha word-initial /p/. The core finding of this study is that the variation concerning word-initial /p/ is systematic, but not categorical, reflecting Weinreich et al.'s (1968) postulation of "ordered heterogeneity" as the norm for all linguistic systems and the heart of the variationist enterprise.

The thesis of Corrina Lee Fuller is approved.

Bruce P. Hayes

Claire Moore-Cantwell

Benjamin J. Eischens

Kie Ross Zuraw, Committee Chair

University of California, Los Angeles

2024

Contents

List of Figures

List of Tables

1 Introduction

This study explores the variable realization of word-initial /p/ in Khalkha [xaɬɮx] Mongolian through the lens of connected speech. While the Khalkha phoneme /p/ has two phonologically-conditioned allophones word-internally ([p] and [w]), in slow, careful speech, word-*initial* /p/ does not alternate. In connected speech, however, not only is word-initial /p/ variably realized as [p] or [w], it is at times deleted altogether, yielding a "zero" realization. This is the case for auxiliary *pai-* (often translated as "to be") below:[1]

(1) **Variable realization of /p/ in *pai-***

Saraa	*khooloo*	*ide-j*	***pai-na***
[sara	xɔːɬɮɔ	itə͡ʧ	**pain~wain~ain**]
PN	food	eat-CNV	AUX-NONPST

"Saraa is eating"

Though all three realizations ([p], [w], "zero") appear *a priori* to be available to any word-initial /p/ in Khalkha, this study identifies /p/-initial function words as the locus of variation, with /p/-initial content words overwhelmingly displaying the maximally-faithful [p] realization. However, not all /p/-initial function word types pattern the same way. Some of these function word types feature all three realizations with considerable frequency, while other function word types show near-categorical preferences for a single realization. Therefore, an investigation into word-initial /p/ requires not only an account of the observed variation in realization but also an account of the differing patterns of variation.

[1] Abbreviations used in this paper: ACC = Accusative; AUX = Auxiliary; COMPL = Complementizer; COND = Conditional; CNV = Converb; FOC = Focus; FUT = Future; HAB = Habitual; IMPF = Imperfective; LBE = Lenition-blocking Environment; LL = Log Likelihood; NEG = Negation; NONPST = Non Past; PART = Particle; PERF = Perfective; PN = Proper Noun; PST = Past; Q = Question Particle; RESTR = Restrictive (Focus); VN = Verbal Noun

Kaisse's (1985) pioneering work into connected speech lends a useful distinction to the discussion of the distribution of word-initial /p/ in Khalkha. Namely, Kaisse separates "fast speech rules" from "external sandhi rules", both of which are processes that appear to affect the realization of certain word-initial /p/s in Khalkha. Despite the adjective in their name, "fast speech rules" need not apply only to the most rapid rates of speech. Instead, they can apply to both normal and faster-paced speech, are informed by adjacent phonetic and phonological environments, and describe processes primarily motivated by ease of articulation. Lenition of word-initial /p/ in Khalkha to either [w] or zero is one such process.

On the other hand, Kaisse's "external sandhi rules" are sensitive to syntactic and/or morpholexical environments. These processes are largely speech rate invariant, with Kaisse going so far as to "deny[] their phonological nature altogether" (p. 3). Kaisse presents English auxiliary reduction as a key example of an external sandhi rule. Aligning with Kaisse, I argue that the realization of /p/ in certain morphosyntactically-defined instances of the Khalkha auxiliary *pai-* is an example of an external sandhi rule, a process that is to be distinguished from general patterns of lenition affecting other function words in the language.

Crucially, I argue that processes of lenition (a type of "fast speech rule") and auxiliary reduction (an "external sandhi rule") are both necessary components of an account of the variable realization of Khalkha word-initial /p/. In articulating this argument, I first present the relevant phonological properties of Khalkha (Section 2) before exploring their connection to the three realizations of word-initial /p/ ([p], [w], and zero; Section 3). While word-internally, there exists a highly-phonologized alternation of [p] and [w] as realizations of /p/, this alternation alone cannot account for the variation observed word-initially.

In Section 4, I introduce the data utilized for this study – four Khalkha-language TEDx Talks

– as well as the methods for their analysis. I then offer a characterization of the nature of the variable realization of word-initial /p/ through careful investigation of this small corpus of naturalistic speech (Section 5), identifying /p/-initial function words (as opposed to content words) as the heart of the observed variation. However, it is not the case that all /p/-initial function word types feature one and the same distribution of realizations of /p/. Some function words strongly prefer the [p] realization while others almost categorically prefer the zero realization, with others still occupying the space between these two poles.

In Section 6, I model this variation within function word types using a Maximum Entropy (MaxEnt) grammar (Goldwater & Johnson, 2003), a probabilistic version of Optimality Theory (Prince & Smolensky, 2004) in which harmony scores for candidate output forms are mapped onto probabilities. This model captures the interplay of markedness constraints dispreferring the [p] realization of /p/ across all functional types and morpheme-specific faithfulness constraints that prefer the [p] realization at degrees specific to each functional type. Section 7 features a discussion of this model, positioning the formal separation of lenition and auxiliary reduction as a key component. Though both processes are encoded as markedness constraints that disprefer the maximally-faithful realization of /p/, [p], the function word type that is affected by auxiliary reduction, *pai-*, is realized as zero to such an extreme degree that the model simply cannot succeed if auxiliary reduction is folded into lenition. Instead, I contend that auxiliary reduction here is a Kaisse-style external sandhi rule that involves auxiliary *pai-* and its listed allomorph, *ai-*. This listed allomorph is morphosyntactically conditioned, which is distinct from the phonetic and phonological conditioning of the lenition of word-initial /p/ elsewhere in Khalkha.

In Section 8, I conclude with a brief discussion of the major implications of this case study

of connected speech, while also highlighting pathways for future research.

A note on transcription and transliteration

In this study, the transliteration of Cyrillic Khalkha Mongolian is given in italics and mostly follows the Tibetan and Himalayan Library system (Atwood, n.d.). Under this system, the Khalkha phoneme /p/ is transliterated from the Cyrillic script using grapheme <b>. Because phoneme /p/ forms the core linguistic unit of analysis in this study, all transliterations of Cyrillic-script Khalkha will utilize <p> in place of <b>, a departure from the Tibetan and Himalayan Library system's schema.

Phonological representations are transcribed using the IPA given in slashes //, while phonetic representations are transcribed using the IPA and are given in brackets [].

2 Khalkha Phonology: An Overview

Khalkha (/xaɮx/) Mongolian is the majority language of the independent nation of Mongolia. It has some 2.6 million speakers (Svantesson, 2020, p. 334) and is often referred to as simply "Mongolian". Khalkha Mongolian is a dialect of a language variously referred to as "Mongolian", "Mongol proper", "Central Mongol", or "Common Mongol". Other mutually-intelligible Mongolian dialects are spoken across Mongolia and the People's Republic of China. The data analyzed in this study come exclusively from Khalkha Mongolian as it is spoken in the Mongolian capital of Ulaanbaatar.

2.1 Khalkha Consonants

The consonants of Khalkha are given as follows:

(2) *Khalkha Consonant Phonemes – adapted from Svantesson et al. (2005, p. 25)*

	labial	*palat. labial*	*dental*	*alveopalatal*	*palatal*	*velar*	*uvular*
VOICELESS ASP. STOPS	p^h	p^{jh}	t^h	t^{jh}			
VOICELESS UNASP. STOPS	p	p^j	t	t^j			
VOICED STOPS					g	g^j	ɢ
VOICELESS ASP. AFFRICATES			$\widehat{ts}^h$	$\widehat{tʃ}^h$			
VOICELESS UNASP. AFFRICATES			$\widehat{ts}$	$\widehat{tʃ}$			
VOICELESS FRICATIVES			s	ʃ	x^j	x	
NASALS	m	m^j	n	n^j		ŋ	
LATERAL FRICATIVES			ɬ	$ɬ^j$			
RHOTICS		r	r^j				
GLIDES	w	w^j			j		

This study focuses solely on the distribution of the voiceless bilabial unaspirated stop /p/ in Khalkha Mongolian. Though palatalized /p^j/ does participate in this alternation, it occurred only once in word-initial position within the present corpus and, as such, will not be investigated further in this study. The aspirated and palatalized aspirated stops (/p^h/ and /p^{jh}/, respectively) do not participate in this alternation, a fact which pushes them, in turn, beyond the scope of the present work.

Because this study concerns the [w] realization of word-initial /p/, some discussion of phoneme /w/ is in order. There are very few native Mongolic words beginning with /w/. Out of over 100,000 Khalkha entries in the online *MongolToli* database, only 117 begin with /w/.[2] Over half of these entries are readily identified as being loans from the Chinese/Buddhist world

[2] *MongolToli* is a web-based Khalkha Mongolian dictionary.

or recent loans from English or Russian. Futhermore, though sound [w] occurs in abundance word-medially and word-finally, this non-initial [w] is overwhelmingly a reflex of historical **p*, with historical Mongolic varieties lacking a phoneme /w/ (see Section 3.1 for further discussion). The scarcity of word-initial /w/ in Khalkha, coupled with the fact that non-initial [w] is not descended from any phoneme /w/, offers a characterization of /w/ as a marginal phoneme.

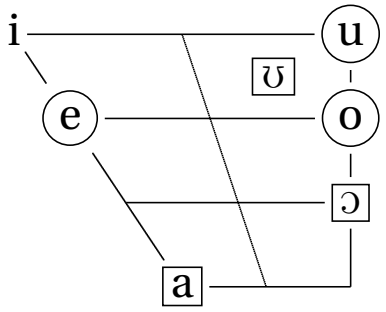
2.2 Khalkha Vowels: Vowel Phonemes, Length, and Harmony

2.2.1 Vowel Phonemes

The vowels of Khalkha are given as follows:[3]

(3) Khalkha Vowel Inventory

a. *Khalkha Monophthongs*



b. *Khalkha Diphthongs*
[ai], [oi], [ʊi], [ui]

[3]Svantesson et al. (2005) and Ueta (2018) argue that /e/ and /i/ have merged to [i] in initial syllables, while the quality of /o/ is [ɵ] when phonetically shorter.

2.2.2 Vowel Length

Though vowel length has traditionally been argued to be a core component of Khalkha phonology (chief among them being Chinggaltai and Lapsangwangtan (1952)), recent advances in the study of Khalkha vowels have questioned the centrality of a vowel length distinction in contemporary Khalkha. Svantesson et al. (2005) find vowel length to be contrastive only in initial syllables as in (4):

(4) *Khalkha vowel length is contrastive in initial syllables*
   a. [uɮ] "not" vs. [u:ɮ] "cloud"
   b. [oɮ] "grey" vs. [o:ɮ] "to trim"

In non-initial syllables, Svantesson et al. (2005) and Ueta (2018) argue that vowels can be either full or reduced.

Because this study looks exclusively at word-initial instances of /p/, we expect the vowels in this initial syllable to demonstrate a vowel length distinction and to be full (non-reduced). While this is largely the case for the vowels following /p/ in lexical words, the vowels following /p/ in content words are typically reduced, be they in the initial syllable or non-initial syllable(s).

2.2.3 Vowel Harmony

The vowels in (3a) are organized with respect to ATR harmony following the works of Ueta (2018) and Svantesson et al. (2005). Vowels /e/, /o/, and /u/ (circles) constitute the +ATR group, while /a/, /ɔ/ and /ʊ/ (squares) form the +RTR group. Vowel /i/ is transparent to ATR harmony when not in an initial syllable, and it can occur in non-initial position of words with either +ATR or

+RTR stems. When in the initial syllable, /i/ (phonetically +ATR) acts as a +ATR vowel and transmits +ATR harmony.

Rounding harmony is also active in Khalkha, and it applies to [-high] vowels. [+high] vowels (/i/, /u/, /ʊ/) are opaque to rounding harmony, as they are neither affected by the rounding feature of the previous vowel, nor do they transmit the [+round] feature to any subsequent vowel.

The domains of both ATR and rounding harmony are the non-compound word, with compound words like the given name "Mönkh(-)bayar" [monx-pajər] (lit. "eternal-joy") violating both ATR and rounding harmony. This study examines only word-initial /p/ and, therefore, we do not expect vowels that follow word-initial /p/ to change based on the vowels of the preceding word.

2.3 The Khalkha Syllable

All possible Mongolian syllable structures are a subset of the following template:

(5) *Khalkha syllable structure (Svantesson et al. (2005, p. 62))*
(C) V (V) (C) (C) (C)

From this syllable schema, a few important generalizations are made apparent. First, Khalkha does not allow for complex onsets. In this way, if a native Khalkha word begins with /p/, the sound following /p/ is, necessarily, a vowel. Loanwords featuring onset clusters are occasionally used in the language, as in *prend* "brand" (as in, "brand of clothing") or *prigad* "brigade". Word-initial /p/ in loanwords of this nature is not known to alternate.

Second, Khalkha *does* allow for complex codas. However, these codas must be strictly decreasing in sonority. The word *elst* [eɮstʰ] "sandy" exemplifies this condition, with the voiced

consonant /ɮ/ being the most sonorous member of the coda, followed by /s/ and then /t^h/. The [p] realization of /p/ may also occur in coda position, though only following [ɮ] or [m]:[4]

(6) a. [aɮ(ə)p] *alba* "service"
b. [sɔjəmp] *soyombo* "soyombo" (name of the national emblem of Mongolia)

The [w] realization of /p/ does not occur in coda clusters.

Finally, though a Khalkha syllable is minimally V, a Khalkha word consists minimally of two moras. Following Karlsson (2014), long vowels are considered bimoraic, while short vowels, nasals, and the lateral fricative /ɮ/ constitute a mora when they are in a syllable nucleus.[5] Thus, the minimal shape of a monosyllabic /p/-initial word in Khalkha is CVV or CVC:

(7) a. [pii] *bi* "I" (1SG.NOM)
b. [pee] *be* (question particle)
c. [paɮ] *bal* "honey"
d. [pur] *bur* "every"

2.4 Stress in Khalkha

The role of stress in Khalkha Mongolian is the subject of much debate. Earlier grammars of Khalkha gesture to the role of stress in the language, especially as it pertains to the distinction between initial syllables and non-initial syllables. In Chinggaltai and Lapsangwangtan's (1952) terms, a syllable is "accented" if it is "distinct and strong" (p. 34-35). They further note that this is typically the case for the first syllable of a word.

[4] In word-internal position, the [p] realization of /p/ is extraordinarily limited, occurring only after [m], [n], [ŋ], [ɮ], or [w]. Coda clusters [np], [ŋp], and [wp] do not occur.

[5] Nasals and the lateral fricative /ɮ/ can be syllable nuclei in casual speech.

Walker (1997) presents quite a different view of Khalkha stress, deeming it a default-to-opposite system of the following type:

1. Primary stress falls on the last syllable if it is the only heavy syllable;
2. Otherwise primary stress falls on the rightmost non-final heavy syllable.
3. If there are no heavy syllables, primary stress falls on the initial syllable.
4. Heavy syllables and (possibly) the initial syllable receive secondary stress when not primary by (1-3). (from Walker, 1997, p.25)

Svantesson et al. (2005) provide their own account of this debate and, indeed, use the fact that there is a debate at all as a key justification for their decision to declare lexical stress phonologically irrelevant in Khalkha. Karlsson (2014) takes this same approach, and I follow suit. In this way, this study will not operationalize stress in analyzing the realization of word-initial /p/.

2.5 Pitch in Khalkha

While stress in Khalkha has been treated variously throughout the literature, there is a remarkable degree of convergence in analyses of pitch patterns in the language. Harnud (2003), Svantesson et al. (2005), Karlsson (2014), and Ueta (2018) all discuss a left-edged Low-High gesture at the level of the word. Only Karlsson (2014) specifies that "word" here is taken to be the prosodic word.

This tonal rise is associated with the first two moras of the word, with the H component of the LH gesture being synchronized with the second mora. Karlsson (2014) further specifies that this gesture applies to content words, though Karlsson et al. (2020) offer insight into some

function words, in that "[p]ost-positions always share the same prosodic phrase (or accentual phrase) as their left-edge host" (p. 212).

Though the nature of the data used in this study differs considerably from the data used in Karlsson (2014) (a corpus of semi-naturalistic data vs. utterances produced in isolation in a laboratory setting), we might still expect to observe a LH gesture associated with the /p/-initial content words in the present corpus. But what of function words? The overwhelming majority of non-[p] realizations of /p/ throughout the present corpus come from instances of /p/-initial function words, though none of them are postpositions. Though Karlsson's (2014) work does not focus on the prosody of function words, many of her examples feature sentence-medial function words, and these are analyzed as not having their own LH gesture. Instead, these function words are grouped with the content word that precedes them into a singular accentual phrase.

The differential treatment of pitch patterns between content and function words resembles the differential distribution of realizations of /p/ between content and function words. Furthermore, the initial /p/ in function words does not occur at any prosodic boundary, while the initial /p/ in content words does. This positioning of initial /p/ in function words is one of prosodic non-prominence, a fact that may offer an explanation as to why initial /p/ is subjected to lenition and reduction in function words but not in content words.

2.6 Section Summary

This section has presented the basics of Khalkha phonology and phonotactics for the purposes of contextualizing the environments in which we find word-initial /p/ in the language. We can now determine the following:

1. Word-initial /p/ will always be followed by a vowel in native Mongolic words.

2. Word-initial /p/ may be preceded by up to three consonants.

3. /p/ is phonotactically legal in syllable codas, but it must occur after /m/ or /ŋ/.

4. The varying analyses of stress in Khalkha make it difficult to associate any stress pattern with realizations of word-initial /p/. Therefore, this study will not consider the stress pattern of a word to be a predictor of the realization of word-initial /p/.

5. Given that a LH gesture characterizes the F0 of each prosodic word, we can predict a tonal rise across the initial two moras of a content word. This puts word-initial /p/ in a lower-pitched environment. The literature has offered a brief glimpse at how function words may pattern with respect to this LH gesture. Namely, this class of words is not expected to have its own LH gesture, but will instead be tacked onto the immediately preceding content word.

3 *p* and *w* in Khalkha

Whereas Section 2 featured overviews of key phonological properties of Khalkha as an introduction to factors potentially contributing to the variable realization of /p/, this section serves to contextualize why [w] would be a realization of /p/ at all. We begin by tracing the diachronic development of the sound *w* in Khalkha. This, in turn, sets the stage for the analysis of /p/ and /w/ as being in an intermediate phonological relationship (Hall, 2013). Though *p* and *w* are contrastive word-initially, they are in an allophonic relationship elsewhere. Nonetheless, with /w/ being so rare word-initially, the contrast between word-initial /p/ and word-initial /w/ is,

itself, marginal. Nowhere in this relationship is a zero realization expected, a fact that puts the zero realization of /p/ at odds with whatever processes may be modulating the [p] and [w] realizations word-internally.

3.1 The Diachronic Development of *w*

Though the consonant inventory in (2) gives /w/ phonemic status following Svantesson et al. (2005), this decision is not universal across all treatments of Khalkha phonology. Battsogt and Önörbayan (2008) present [w] as an allophone of /p/, as does Tömörtogoo (2017).[6] Neither gives *w* phonemic status in Khalkha. In addition to positing a phoneme /w/ in Khalkha, Svantesson et al. (2005) also provide an account of an allophone [w] as having developed from Old Mongolian (c. 12th century) **p*.

The varying treatments of *w* in Khalkha become more understandable when one considers two strands of diachronic development involving the sound *w*. The first is related to the marginal status of /w/ across the Mongolic languages. Old Mongolian had no phoneme /w/, but the language incorporated many /w/-initial loanwords from Chinese such as *wang* "king" and *waar* "tile". Thus, *w* became a marginal phoneme in Old Mongolian, occurring only in loanwords and almost exclusively in word-initial position.

The second strand concerns a process of lenition that affected reflexes of the Old Mongolian **p*. Over time, reflexes of this **p* have come to be lenited to [w] in most non-initial positions within Khalkha words. (8), below, demonstrates the lenition process occurring between attested Classical Mongolian (17th-19th centuries) written forms and Modern Khalkha forms.[7]

[6] Both Battsogt and Önörbayan (2008) and Tömörtogoo (2017) refer to phoneme /p/ as <b>.

[7] The Classical Mongolian written forms are taken from Lessing (1960).

(8) a. Classical Mongolian: <pagatur> “hero” > Modern Khalkha [pa:t^{h}ər] “hero”
b. Classical Mongolian: <arpan> “ten” > Modern Khalkha [arwən] “ten”
c. Classical Mongolian: <eɮtep> “abundance” > Modern Khalkha [eɮtəw] “abundance”

In this way, the *w*-initial loanwords are synchronically analyzed by Svantesson et al. (2005) as involving phoneme /w/.[8] The status of these *w*-initial words as loanwords may be behind the decisions of Tömörtogoo (2017) and Battsogt and Önörbayan (2008) to not consider *w* to be a phoneme in Khalkha. By contrast, all of the above authors recognize the historical relationship between **p* and its modern reflexes of [p] and [w], a point that is illustrated by [w] being recognized as an allophone of /p/ across all three works.

3.2 An Intermediate Phonological Relationship

The different sources of modern Khalkha *w* have created different contexts for contrast. The incorporation of *w*-initial loanwords, coupled with the fact that word-initial position is one of the few environments in which native Mongolic /p/ does *not* lenite to [w], have led to /p/ and /w/ being contrastive in word-initial position, as in (9).

(9) a. [pa:r] “bar” (a place to drink, loan from Indo-European)
b. [wa:r] “tile” (loan from Chinese)

In non-initial position, [p] and [w] alternate, and this alternation is phonologically conditioned. Prescriptive grammars of Khalkha (Kullmann & Tserenpil, 2015; Legden & Luethy, 2009) give the following pair of rules to predict the alternation word-internally:

[8] The more modern influences of Russian and English on Khalkha Mongolian have resulted in additional strata of *v*- and *w*-initial loanwords in the language. Ex: Khalkha *walyut* “currency exchange” < Russian *valyuta* “currency”; Khalkha *web* “web, internet” < English “web, internet”.

(10) a. /p/ -> [p] / {m, n, ŋ, ɬ, w}___

b. /p/ -> [w] elsewhere

According to these rules, word-internally, the distribution of [p] and [w] as allophones of /p/ is highly predictable. /p/ is realized as [p] after nasals ([m], [n], or [ŋ]), the lateral fricative [ɬ], and the approximant [w].[9] This pattern of alternation holds across a morpheme boundary, as the conditional affix [-*p*əɮ]/[-*u*əɮ] demonstrates:

(11) a. [xa**m-p**əɮ] "to rake up-COND"

b. [xeɮ**-p**əɮ] "to say-COND"

c. [a**w-p**əɮ] "to take-COND"

d. [i**t-w**əɮ] "to eat-COND"

e. [**uː-w**əɮ] "to drink-COND"

f. [ɔtʃʰ-wəɮ] "to go-COND"

g. [ɔ**r-w**əɮ] "to enter-COND"

Thus, I argue that /p/ and /w/ are in an intermediate phonological relationship (Hall, 2013) in that they are contrastive word-initially, as in (9), and that [p] and [w] are in an allophonic relationship elsewhere within the domain of at least the non-compound word.[10]

While the domain of this alternation is most transparently viewed as the non-compound word, it also extends to the domain of the clitic group, as in (12):

[9] Together, the phonetic environment conditioning the [p] realization of /p/ can nearly be generalized as being "following a sonorant". However, the exclusion of [r] as a conditioning environment for [p] would make this generalization false.

[10] Compound words will not be treated in this paper. They are often names of people or places such as *Ulaan(-)baatar* (name of the capital city of Mongolia, lit. "red-hero") or *Mönkh(-)bayar* (a given name, lit: "eternal-joy"). Compound words often violate vowel harmony (*Mönkh(-)bayar* violates both rounding and ATR harmony), and do not demonstrate the [p] ~ [w] alternation described in this section.

(12) a. *ta* *yu* *id-sen* ***be***
t^{h}a: ju: it-sən pe:
you what eat-PST.VN **Q**

"What did you eat?"

b. *ta* *yu* *id-deg* ***we***
t^{h}a: ju: it-təg we:
you what eat-HAB.VN **Q**

"What do you eat?"

c. *ta* *yu* *id-ekh* ***we***
t^{h}a: ju: it-əx we:
you what eat-FUT.VN **Q**

"What will you eat?"

In (12a), the question particle follows a verbal noun ending in a nasal, and is realized as [pe]. (12b-c) feature verbal nouns ending in sounds in the "elsewhere" condition in the rules in (10). Predictably, in these cases, the question particle is realized as [we].

The alternation in the initial segment in the conditional affix [-pəɮ ~ -wəɮ] and the question particle [pe:~we:] is reflected orthographically in Cyrillic-script Khalkha (in use since the 1940s), with [p] realizations being written using grapheme <b> and [w] realizations using grapheme <w>. This is a departure from the earlier Vertical Script (used widely from the 13th century up to the 1940s), wherein only word-initial instances of *w* (i.e., those coming from loanwords) were written with the grapheme <w>. While the relationship between phonological representations and orthographic conventions is often not one-to-one, in this particular case, the new (c. 1940s) orthographic conventions could be representative of an official codification of the phonologization of this alternation.

3.3 The Zero Realization of /p/

Given the relationship between sounds [p] and [w], the presence of a zero realization of word-initial /p/ is striking. As was demonstrated in Section 3.2, the language is able to accommodate phoneme /p/ after any segment, yielding [p] when following nasals, [ɲ̟], or [w] and yielding [w] elsewhere. Though these processes occur within the domain of the word or clitic group, this study will show that they can at times occur between words in naturalistic speech, where certain prosodic boundaries may be weaker.

None of this works to explain why initial /p/ at times deletes altogether. This deletion is not part of extremely careful speech, nor is it reflected anywhere in the standard orthography. Indeed, this study argues that the zero realization of /p/ is the result of two different processes inherent to connected speech, a fact that can account for the zero realization's conspicuous absence in analyses of Khalkha speech that is produced in more careful and controlled contexts.

3.4 Section Summary

Though [p] and [w] alternate within the domain of the word and clitic group, the data analyzed in this study concern solely word-initial instances of /p/. Recall that word-initial position is a contrastive context for /p/ and /w/. Furthermore, word-initial /p/, being at the left edge of a word boundary, is not part of the domain over which the [p~w] alternation applies. Thus, in investigating the realizations of word-initial /p/, our null hypothesis is that every single occurrence will feature the [p] realization. Clearly, this is not the case, as the subsequent sections will demonstrate.

That /p/ of any type (word-initial, not word-initial) can be deleted altogether does not follow from the facts of the [p~w] alternation. Section 5 will show that the zero realization of

word-initial /p/ occurs almost exclusively within a specific class of function words (auxiliaries) within a restricted set of syntactic configurations, indicating that the zero realization of /p/ is likely *structurally* conditioned rather than phonologically conditioned. This structural conditioning is an instantiation of a Kaisse-style (1985) "external sandhi rule", operating outside of the realm of phonology and resulting in auxiliary reduction. This is to be distinguished from /p/-lenition (a Kaisse-style "fast speech rule") wherein realizations of word-initial /p/ vary on a continuum from [p] to [w] to zero, but can be predicted based on the phonetic and phonological environment. Both auxiliary reduction and /p/-lenition are needed to capture the distribution of word-initial /p/ in Khalkha.

4 Data and Methods

4.1 Data Collection

This study utilizes data from four Mongolian language TED^{x} Talks, all given at either TED^{x} Ulaanbaatar or TED^{x} Baga Toiruu. Both locations are within the capital city of Mongolia, Ulaanbaatar, with Baga Toiruu referring to the centermost region of the capital, encompassing many government offices and buildings. These talks were uploaded to YouTube between the years of 2014 and 2017 and were given by adults in their 20s and 30s. The table below summarizes information about each talk.[11]

Taken together, these four talks form a small spoken corpus that represents 53 minutes and 25

[11] CODE = Code name given to the talk, and the name that will be used for each talk throughout this paper; LENGTH = Length of the talk in minutes:seconds; GENDER = The assumed gender of each speaker; AGE = The age of the speaker at the time the video was uploaded; UPLOAD YEAR = Year the recording of the talk was uploaded to YouTube; LOCATION = Place where each talk was given; TITLE = Title of the talk. The titles for TED1, TED2, and TED4 were translated into English by the author. The title of TED3 was given in English by the uploader.

CODE	LENGTH	GENDER	AGE	UPLOAD YEAR	LOCATION	GENERAL TOPIC
TED1	16:01	M	23	2015	Baga Toiruu	mindfulness, gratitude
TED2	15:23	F	28	2017	Ulaanbaatar	compassion in the workplace
TED3	8:27	M	26-27	2014	Ulaanbaatar	media censorship
TED4	13:34	F	~30	2015	Baga Toiruu	youth participation in politics

Table 1: Coding of Sample TED^{x} Talks

seconds of recorded speech.

TED^{x} Talks are particularly useful in an investigation into connected speech. The talks are characterized as being somewhat memorized and, therefore, quasi-performed. Still, there is room for spontaneous utterances, and the speech of all four talks is characterized (to varying degrees) by false starts, filler words, and speech errors, to name just a few of the salient trappings of spontaneous speech. Moreover, these talks are given to audiences of Khalkha Mongolian speakers, thereby negating certain tendencies toward hypercorrection and careful pronunciation that are often found in speech directed toward non-native speakers. Most importantly, the nature of the speech in all four talks is such that the realization of word-initial /p/ is varied. Though each token analyzed in this study would feature a [p] realization of /p/ when spoken in isolation or with extreme care, all four speakers display variation in their actual realizations of word-initial /p/. Crucially, I argue that this variation is patterned, robust across speakers, and analyzable in terms of the dual processes of lenition and auxiliary reduction.

The four talks were selected based on their audio quality and the presence of a corresponding human-generated transcript. These time-aligned transcripts are available for download on the YouTube video itself. There are many instances where the text of the transcript diverges from the actual speech of the speaker. This appears to have been done to remove filler words, errors, and circumlocutions on the part of the speaker. In this way, transcripts provide a base from which to begin identifying /p/-words. Given the discrepancies between each transcript and

the actual words spoken by each speaker, this study considers all instances of word-initial/p/, whether present in the transcript or not.

4.2 Data Coding

All talks for this study were analyzed in Praat (Boersma & Weenick, 2023), with each occurrence of a word-initial /p/ receiving a code of [p], [w], or "zero" to represent the speaker's phonetic realization of /p/. In addition to matching what was heard, the presence or absence of a release burst in conjunction with the presence or absence of a rising gesture F2 served as the defining characteristics of the [p], [w], or "zero" codes.

	[p]	[w]	ZERO
Release Burst	+	-	-
F2 Rise	+	+	-

Table 2: Coding of Word-initial /p/

Naturally, there is considerable phonetic variation within each of the three realizations of /p/ that I analyze in this study. This is particularly the case for realizations which I have labeled [w], which is a stand-in for a range of realizations of /p/ that carry the +CONTINUANT feature. While IPA [w] is the most canonical of them, a narrow transcription could reveal [v], [ʊ], or even [β], all of which lack a release burst but feature an F2 rise due to a constriction at the lips.[12]

Thus, the labels of [p], [w] and "zero" can be thought of as degrees on a spectrum of /p/-lenition. The maximally faithful, unlenited realization is coded as [p], while the [w] code indicates lenition to a continuant. The "zero" code represents a fully lenited form. However, I argue that there are two sources of a "zero" realization. The first source comes from this spectrum

[12]To better probe the phonetic contouring of each sound in question, this author recommends studies conducted in a laboratory setting, which are beyond the scope of this paper.

of /p/-lenition and is a "fast speech rules" feature of connected Khalkha speech. The second comes from auxiliary reduction, a separate process whereby certain morphemes in specific morphosyntactic configurations are realized as "zero" regardless of speech rate, phonetic environment, or any other phonological consideration. I argue that these sources of the "zero" realization are distinct, but that both must be encoded in a grammar of the language in order to account for the variable realization of word-initial /p/.

This small corpus has yielded 1334 analyzable tokens of word-initial /p/ spread across all four talks. Only tokens that feature a non-alternating word-initial /p/ are included in this study. This means that question particle *be/we* [pe~we], while its own orthographic word in standard Khalkha orthography, is not included in this study due to its highly phonologized distribution that features strict alternation ([p] after {m. n, ŋ, ɮ, w}; [w] elsewhere).

Tokens are considered "analyzable" if they meet the following criteria:

1. Are uttered by the speaker
2. Are of sufficient audio quality
3. Are not overlapping with audience noise (clapping, laughing)

Tokens failing to meet any one of these criteria were excluded from the study.

5 Main Results

5.1 The Distribution of Word-initial /p/

Figure 1 gives a generalized breakdown of the distribution of word-initial /p/ across the corpus, where the x-axis represents the realization of /p/ produced ([p], [w], or "zero") and the y-axis represents the number of occurrences of each realization.

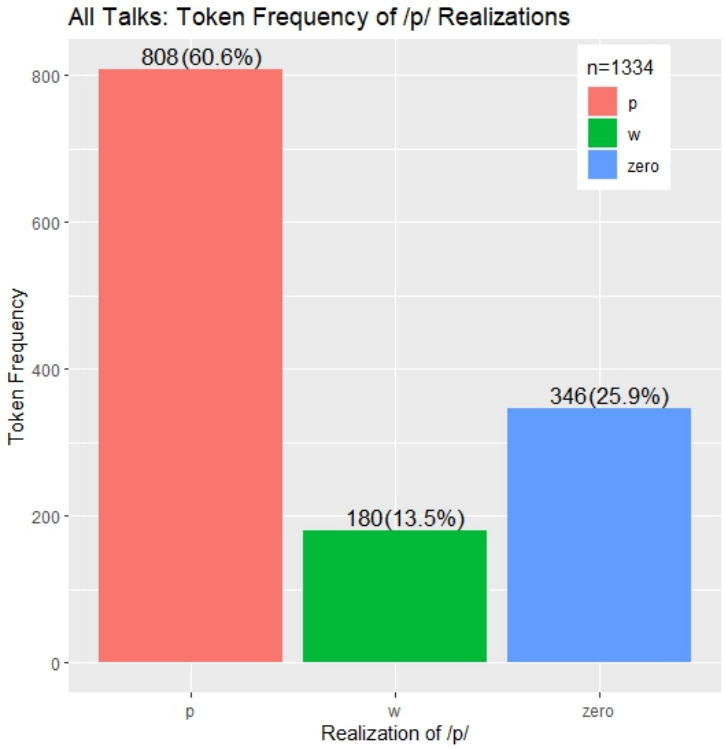


Figure 1: Aggregate Distribution of /p/

The corpus data show a strong preference for the [p] realization of /p/ and a slight preference for the zero realization, with the [w] realization being the least preferred. This general trend holds when looking at the realizations of word-initial /p/ for each individual speaker:

Figure 2: /p/ Realizations in TED1

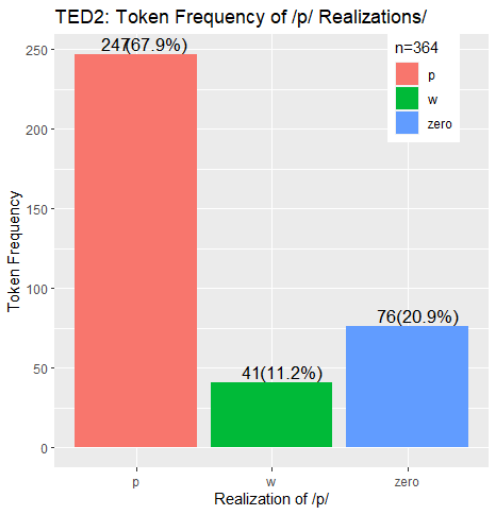


Figure 3: /p/ Realizations in TED2

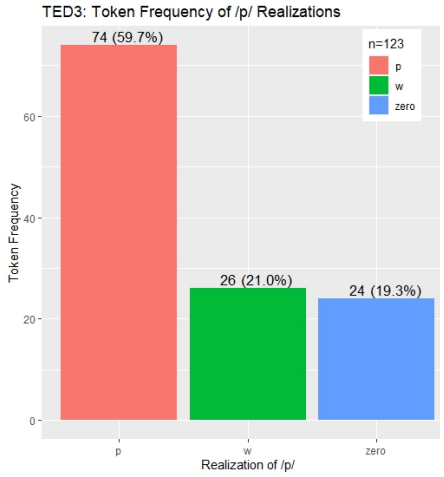


Figure 4: /p/ Realizations in TED3

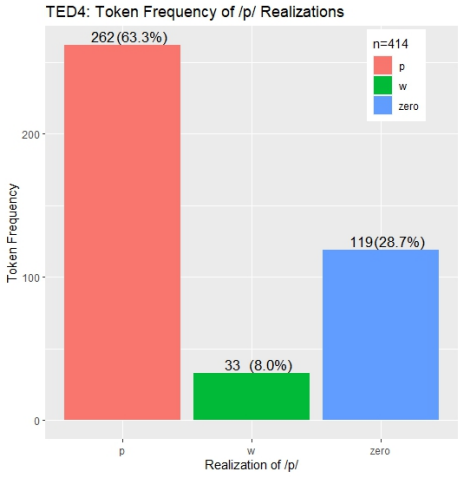


Figure 5: /p/ Realizations in TED4

TED3 is the only talk to feature a higher proportion of [w] realizations than zero realizations. This is due in part to the low overall token count of the talk, but also due to the kinds of tokens in the talk. As Section 5.3 will demonstrate, two function word types (*pol* and *pai-*) constitute the majority of the zero realizations of /p/ throughout this corpus. The speaker of TED3 simply does not use these functional types at the same rate that the other speakers do, causing the relative proportion of zero realizations of word-initial /p/ to appear lower. I argue that this is not a meaningful difference in overall treatment of word-initial /p/. Instead, it is more of a reflection of the types of words in each talk.

5.2 /p/ in Function vs. Content Words

A closer inspection of the data reveals a marked split in the treatment of /p/ when it begins content words versus when it begins function word. For the purposes of this study, "function words" are all closed-class lexical items including pronouns, negation markers, particles, auxiliaries, quantifiers, conjunctions, modals, qualifiers, and postpositions. "Content words" are all open-class lexical items, consisting mainly of verbs, nouns, adjectives, and adverbs.

Whereas content words display a near-categorical preference of the maximally faithful [p] realization of /p/, the function words are responsible for the majority of the variation in the realization of /p/. Figure 11 demonstrates this split in the realization of /p/ in function and content words throughout the corpus:

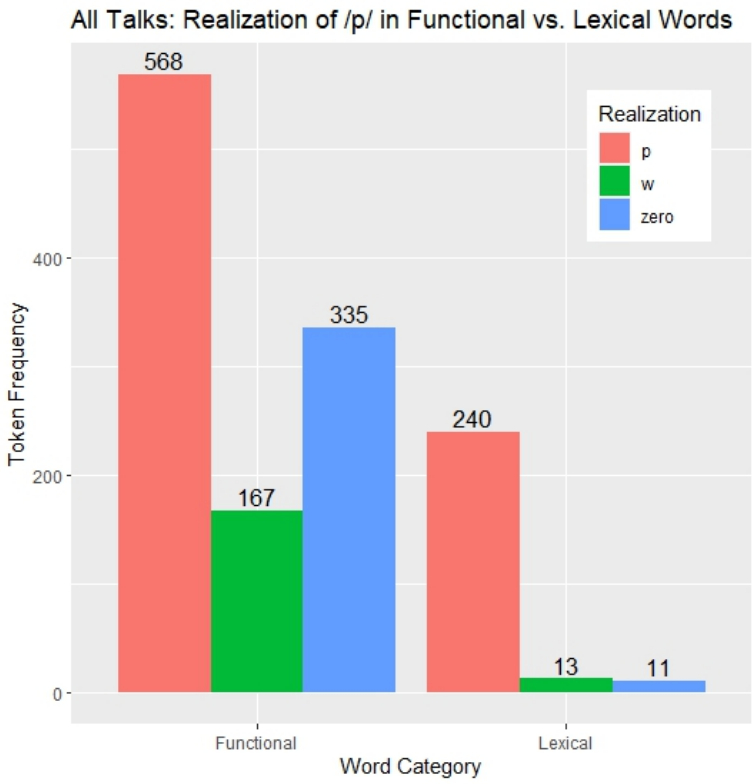


Figure 6: Realization of /p/ in Function vs. Content Words

5.3 Word-initial /p/ Varies across Function Words

The majority of the variation in word-initial /p/ comes from the function words, making them the core object of inquiry in this study. That word-initial /p/ is variably realized in function words to a much greater degree than it is in content words is entirely unsurprising in light of Beckman's (1998) account of positional faithfulness, wherein function words are argued to be psycholinguistically less prominent and thus occupy non-privileged positions within the phonological system.

Function words account for 1,070 out of 1,334 tokens (80%) of word-initial /p/ within my small corpus. These 1,070 tokens are represented by just 26 types (cf. 264 tokens spread across 65 types for content words). 911 of the 1,070 function word tokens are represented by just four types. These are the only function word types to occur with a frequency of $n > 30$:

1. *pi*: 1st-person pronoun, can be inflected to reflect plurality (*pit* 1PL.NOM).
    - Token frequency: 191
2. *pol-*: Copula/auxiliary, often translated as "to become".
    - Token frequency: 112
3. *pol*: Particle that can mark interrogatives, conditional statements, or serve as a neutral focus marker.
    - Token frequency: 162
4. *pai-*: Copula/auxiliary, often translated as "to be".
    - Token frequency: 442

Figure 7 presents a breakdown of word-initial /p/ realizations across these four types.

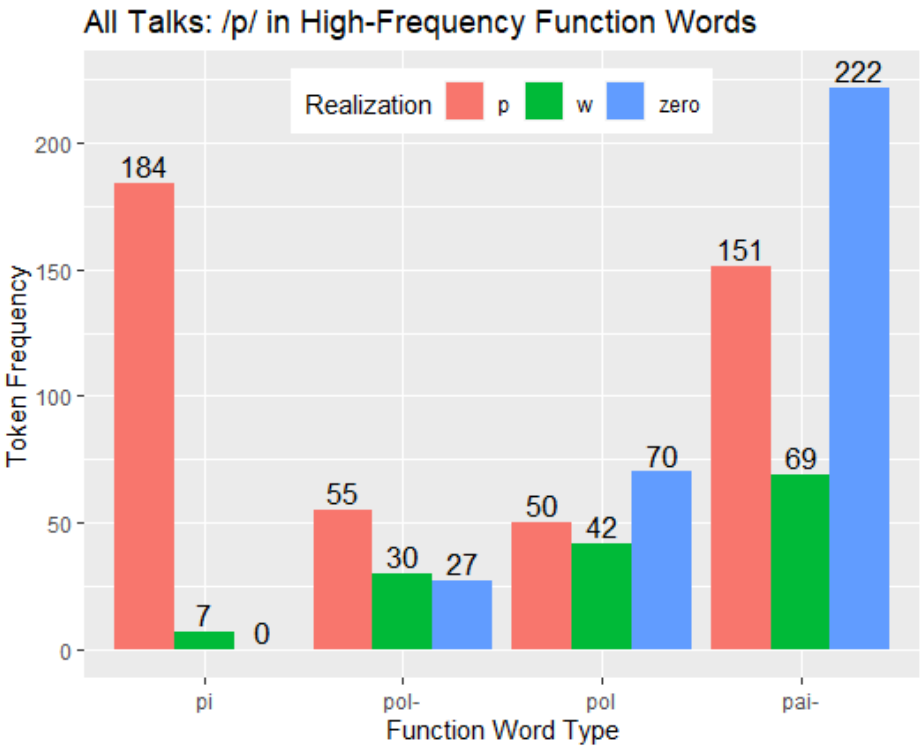


Figure 7: Realization of /p/ in High-Frequency Function Words

Whereas *pai-* displays a marked preference for the zero realization, first-person pronoun *pi* [pii] shows a near-categorical preference for the maximally-faithful [p] realization. The distribution of /p/ across these types demonstrates that the function words are not acting as a unified class.

5.4 /p/ in Fixed Expressions and Collocations

5.4.1 Content words

While it may be tempting to declare variation in word-initial /p/ to be a property of function words alone, the content words displaying variation shed light on potential motivations for variation throughout the entire language. First, the content words with non-[p] realizations represent 9% (24/264) of all content words throughout the corpus. This amount is non-negligible. Second, all but four of these 24 content words can be readily identified as being the second or third element of a fixed expression or a collocation.

Fixed expressions are sequences of words that are used together to convey a specific meaning. The fixed expressions identified in this corpus all have separate dictionary entries in the *MongolToli* database, suggesting lexicalization of these phrases. Seven such expressions featuring a content word beginning with /p/ are given as follows (/p/-initial content words in bold):

WORD	MEANING	COUNT	REALIZATION	EXPRESSION	EXPRESSION MEANING
podol	"thought"	3	zero	*uzel* ***podol***	"opinion"
polgo-	"cause (V)"	2	zero	*pii* ***polgo-***	"form (V)"
paiguullaga	"organization"	1	zero	*töriin pus* ***paiguullaga***	"NGO"
piye	"body"	1	w	*piye* ***piyee***	"each other"

Table 3: Fixed Expressions with Content Words

Table 3 catalogs the only occurrences of these fixed expressions in the corpus. That is, there are no tokens of these expressions that feature a [p] realization of the lexical item.

Unlike fixed expressions, words within a collocation do not necessarily take on a new meaning. Collocations instead represent words that frequently co-occur or are otherwise mutually informative of one another. The present lack of access to a broader, lemmatized corpus of Khalkha texts limits this study's ability to quantify the degree to which members of potential collocations are, in fact, mutually informative of one another. However, the fact that many (13/24) of the non-faithful realizations of /p/ in content words appear to be in collocations indicates a fruitful avenue for further exploration of variation of word-initial /p/ involving content words.

The most frequent collocation in this group is the bigram *gej pod-*, formed from the complementizer *gej* and verb root *pod-* "to think", as in:

(13) *pi eniig engiin zuil* ***gej*** ***pod-okh-gui*** *paina*
I this.ACC normal thing COMPL think-NONPST-NEG AUX
"I do**n't think that** this is normal"

This collocation occurs 28 times throughout the corpus and is realized as [w] or "zero" five of those times. Though this hardly shows a preference for a non-faithful realization of /p/ in this configuration, the relative frequency at which these two morphemes co-occur may suggest that articulatory gestures are more coordinated across these words and that interlocutors are able to predict the second, /p/-initial word on the basis of the first word.

This connection between commonly co-occurring forms and processes of lenition is not a novel one. Bybee (2001, 2006) and Jurafsky et al. (2001) build upon over a century of research into the key roles of frequency and predictability in modeling human language processing and

extend these variables to models of human language production. They argue, chiefly, that commonly co-occurring words are more susceptible to reduction processes, with Bybee (2001, 2006) identifying the routinization of neuromotor procedures (achieved through frequent co-occurrence) as the cause of overlap and reduction in articulatory gestures.

The remaining four content words with non-[p] realizations do not have any particular relationship with the words that immediately precede them:

(14) a. [w] realization
paigaa ***paidal***
be situation

"a situation that..."

b. [w] realization
ene ***painga***
this always

c. [w] realization
podloo ***pich-eed***
thought write-CNV

"write (one's) thoughts, and then..."

d. "zero" realization
ter-iig ***polowsrol***
this-ACC education

These four non-[p] realizations of /p/-initial content words constitute such a small portion of the available data that the variation here will be considered trivial.

Conversely, that 20 out of 24 of the content words with either [w] or "zero" realizations are parts of fixed expressions or collocations creates a conspicuous profile. In particular, this profile suggests that variable realizations of /p/ in content words in Khalkha may be sensitive to frequency and predictability of occurrence, aligning with the findings of Bybee (2001, 2006) and

Jurafsky et al. (2001).[13]

5.4.2 Function words

Function words can be part of fixed expressions, just as content words can. Indeed, the function words in fixed expressions bear striking similarities to their content word counterparts. Table 4 demonstrates the preference for function words in fixed expressions to lenite to [w] or zero.

WORD	MEANING	COUNT + REALIZATION	EXPRESSION	EXPRESSION MEANING
pukhen	"all"	2p, 2w, 7 zero	*ta* ***pükhen***	"you all"
puriin	"every"	1p, 2w, 2 zero	*yanz* ***püriin***	"various"
pus	NEGATION	1w	*töriin* ***pus*** *paiguulaga*	"NGO"
pol-	"to become"	1 zero	*tus* ***polokh***	"to assist, avail"
pol-	"to become"	1 zero	*il* ***polokh***	"to come out" (as in a secret)
pai-	"to be"	4 zero	*yamar ch* ***paikh***	"in any case..."

Table 4: Function Words in Fixed Expressions and Collocations

Because function words and content words pattern so differently elsewhere in the language with respect to word-initial /p/, fixed expressions and collocations offer a way to unite the variable realization of word-initial /p/ across the entire lexicon. That is, both fixed expressions and collocations involve forming a multi-word unit or "chunk" of speech wherein the word forming the latter part of the chunk is highly predictable based on the appearance of the first part of the chunk, a fact that is irrespective of "function" or "content" word status. This chunking can involve overlap and/or reduction of articulatory gestures, leading to non-[p] realizations of /p/. Additionally, the chunks formed from these frequently co-occurring words could result

[13]The findings of Bybee (2001, 2006) and Jurafsky et al. (2001) are not unquestioned. See MacKenzie and Yang (2013) for a case study that finds neither frequency nor predictability to be relevant to *h*-deletion in English.

in a weakening of the prosodic boundaries between each individual word and thus make an expression-internal /p/ (like in *gej **pod-*** “to think that...”) more susceptible to lenition.

5.5 Section Summary

The results of this study yield three key insights into the distribution of word-initial /p/ in Khalkha. First, the [p] realization of /p/ is most preferred across all speakers, with the zero and [w] realizations being increasingly dispreferred. Second, a primary split in the data reveals that content (open-class) and functional (closed-class) words pattern differently with respect to their realizations of word-initial /p/. Whereas content words have a near categorical preference for the [p] realization (91% of all content word tokens), function words display considerable variation in realization. Third, function words do not act as a unified class with respect to the realization of /p/. The four function word types with token frequencies greater than 30 (Figure 7) capture this apparent fact, suggesting that /p/-lenition rates are not constant throughout the class of function words.

A secondary finding of this corpus study concerns the phonology of fixed expressions and collocations. Word-initial /p/s that are present in these environments show a preference to lenite regardless of whether the /p/-initial item is a content word or a function word. While the low token frequency of each fixed expression/collocation in this study precludes a rigorous investigation of this environment, these preliminary findings echo those of Bybee (2001, 2006) and Jurafsky et al. (2001) which view reduction processes as a probable outcome of frequently and predictably co-occurring units of language.

6 Modeling the Distribution of Word-initial /p/

Whereas Section 5 outlined the major trends of the corpus data, the present section will formally define a phonological grammar to model the varying realizations of /p/. Given that the data comprising the corpus is naturalistic and that the phenomenon under investigation is variable in nature, I employ a Maximum Entropy grammar (Goldwater & Johnson, 2003) in modeling /p/-lenition in Khalkha.

This grammar is concerned solely with the realization of word-initial /p/ within the four high token-frequency function word types outlined in Section 5 (*pi, pai-, pol-,* and *pol*). Content words are excluded from the grammar as they feature virtually no variation outside of fixed expressions and collocations. Furthermore , these four functional types are the only types to occur with a token frequency greater than 30.

Each input in this grammar is one of the four high token-frequency function word types, and each of these inputs has three candidate output forms: one featuring a [p] realization, one featuring a [w] realization, and one featuring a zero realization. Constraints are fitted with weights (w) and the harmony (H) of each candidate is computed by taking the weighted sum of that candidate's violations. The probability (p) of each candidate is exactly proportional to the negative exponential of its harmony. This probability can be compared to the observed proportion of realizations for a given candidate output form to assess how well the predictions of the grammar match the observed data.

A single tableau thus looks like the following:

/pai-/	Token Freq.	Obs. Prop.	Constraint$_i$	Constraint$_n$	H	p
			weight	*weight*		
pai-	151	.34				
wai-	69	.16				
ai-	222	.50				

Table 5: Sample Tableau

6.1 Markedness as Preferring Lenition

Recall from Section 5 that the four high-frequency function word types prominently feature lenited realizations of /p/ (see Table 7). This can be encoded in the grammar as a markedness constraint that penalizes non-lenited realizations.

(15) LENITE: Assign a violation for every realization of /p/ that is not lenited. [p] realizations incur 1.7 violations and [w] realizations incur 1 violation.

The number of LENITE violations incurred can be thought of as a hyperparameter of the model. The value of 1 for [w] was chosen arbitrarily, while the value of 1.7 for [p] was chosen by allowing Solver in Microsoft Excel to adjust both this number and the constraint weights when maximizing the objective function.

Each input has three corresponding candidate output forms (one for each realization of /p/), and these realizations represent a cline of lenition. The zero realization of /p/ indicates the strongest degree of lenition and incurs zero violations to LENITE. The [w] realization is lenited, but is not the strongest degree of lenition. It incurs a single violation. The [p] realization features no lenition and incurs 1.7 violations.

Thus, LENITE is a markedness constraint. More importantly, it makes an analytical claim about a preferred output configuration. Namely, this constraint encourages lenition to zero as opposed to lenition to [w]. This type of claim follows naturally from the insights of a Kaisse-

style fast speech rule. This approach takes the zero realization to feature the greatest degree of articulatory ease, with the [p] realization requiring the greatest amount of articulatory effort.

6.2 Morpheme-specific Faithfulness Constraints

The LENITE constraint alone treats each input form as if it undergoes lenition at the same rate. As Figure 7 shows, this is simply not the case, with *pi* featuring almost no lenition whatsoever while the other high-frequency function words display varying rates of lenition.

What emerges, then, is a picture of each functional morpheme patterning differently. This apparent fact can be implemented in a MaxEnt grammar by utilizing lexically-indexed constraints, as are put forth by Pater (2010). These constraints largely work to moderate the degree to which each individual morpheme engages in /p/-lenition.

(16) FAITH-(MORPHEME): Assign a violation to every candidate that features an initial segment that is not faithful to its corresponding input. [p] realizations incur no violations, [w] realizations incur 1 violation, and zero realizations incur 1.6 violations.

As in the case of the LENITE violations, for these faithfulness constraints, the value of 1 for [w] was chosen arbitrarily while the value of 1.6 for zero was fitted by Solver.

In reality, this singular FAITH-(MORPHEME) constraint has behind it two faithfulness constraints: ID(CONTINUANT)-(MORPHEME) and MAX(C)-(MORPHEME). This means that the expanded model contains eight faithfulness constraints (two per morpheme). For the purposes of succinct exposition, these pairs of morpheme-specific faithfulness constraints are collapsed into a single faithfulness constraint, all with no significant effect on the overall log-likelihood of the model.

The morpheme-specific FAITH constraints, like LENITE, operate over a cline. This time, the

cline represents degrees of faithfulness, with [p] being maximally faithful, [w] being less faithful, and zero being the least faithful. A tableau featuring these constraints thus looks like the following, with the numbers in the cells showing the number of violations:

/pai-/	LENITE	FAITH-(PAI-)	FAITH-(POL-)	FAITH-(POL)	FAITH-(PI)
	weight	*weight*	*weight*	*weight*	*weight*
pai-	1.7				
wai-	1	1			
ai-		1.6			

Table 6: Sample Tableau with LENITE and FAITH Constraints

Note first that because the faithfulness constraints are morpheme-specific, the input form /pai-/ in Table 6 interacts only with LENITE and FAITH-(PAI-).

Second, and more importantly, note that the two clines that LENITE and FAITH operate over are the reverse of one another. That is, while the [p] realization incurs 1.7 violations of LENITE and no violations of FAITH, the zero realization incurs no violations of LENITE but 1.6 violations of FAITH. The interaction of these constraints has the benefit of allowing each morpheme-specific FAITH constraint to pivot around the morpheme-general constraint LENITE. In other words, these faithfulness constraints represent a morpheme's resistance to fast speech rules that are motivated by ease of articulation (and, thus, favoring lenited realizations). As such, we might expect the weight of FAITH-(PI) (indexed to the nearly non-leniting type *pi*) to be far greater than the weight of FAITH-(POL-).

A surface-level disadvantage of this approach is that when looking only at these two constraints, [w] realizations never emerge as the preferred realization of /p/. Because it incurs a violation of 1 for both the markedness and faithfulness, the [w] realization in the present grammar resembles something like harmonic boundedness in Classical Optimality Theory. This would

be fatal if the data set showed a function word type that preferred the [w] realization. As the data do not display such a word type, the grammatical encoding of a dispreference for [w] realizations is intentional and should be construed as a beneficial feature of the model rather than a proverbial "bug".

6.3 Predicting [p]: The "Lenition-Blocking Environment"

Aside from general principles such as "ease of articulation" or morpheme-specific resistance to lenition, certain environments appear to predict one realization over all others.

We will take functional type *pai-* as a point of departure into the investigation of these environments, as tokens of one and the same morpheme show a preference for the "zero" realization, but also a slight preference for the /p/ realization. The vowel sound following the initial /p/ is, phonologically, always /ai/.[14] Thus, the immediately succeeding environment offers little insight into the variable realization of /p/ in *pai-*.

The segmental environments immediately *preceding* the function words offer a different perspective. Recall from Section 3 that the [p] realization of word-internal /p/ was conditioned by a distinct subset of sounds, namely, [m], [n], [ŋ], [ɲ], and [w], while the [w] realization worked as the "elsewhere" condition. In terms of word-internal /p/, these specific preceding contexts work to block lenition of /p/ to [w]. These contexts appear to have the same effect on word-initial /p/ in function words. Figure 8 shows a breakdown of the realization across the four key functional types if we applied the same rules to word-initial /p/ as we do to word-internal /p/.

[14] Phonetically, there is considerable variation in the phonetic realization of /ai/, even between tokens featuring the same realization. This variation falls outside the scope of this paper.

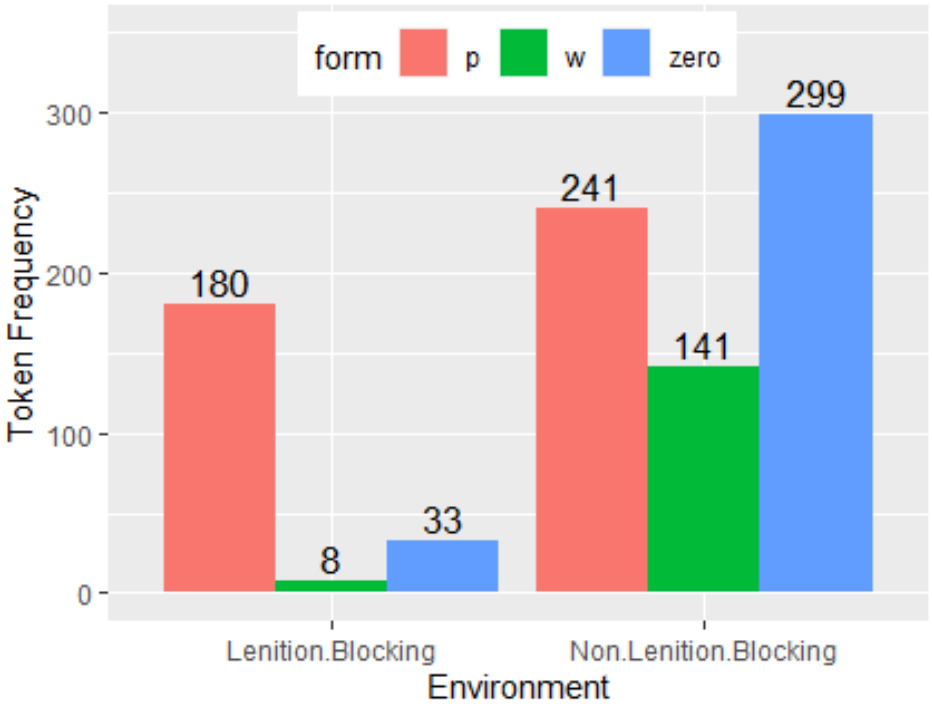


Figure 8: /p/ in Lenition-Blocking Environments

These distributional facts appear to reflect a relevant phonological environment that conditions a [p] realization.

(17) LENITION-BLOCKING ENVIRONMENT (LBE): Assign a violation to every lenited realization of /p/ that follows the segments [m], [n], [ŋ], [ɲ], or [w]. This penalizes the [w] and zero realizations that occur in these environments, with both realizations incurring 1 violation.

6.4 Predicting [w]: More Ease of Articulation in Fast Speech Rules

Figure 8 gets at the heart of the variable realization of /p/ with its depiction of the "Non-Lenition-blocking" environment. Whereas the word-internal environment conditioning the [p] realization appears to condition [p] word-initially for function words, there is no clear mapping of the word-internal "Non-Lenition-blocking" environment onto what we see word-initially. Recall that word-internal /p/ is not realized as zero. This is in stark contrast to the abundance of zero realizations that we see word-initially. Moreover, applied word-initially, "Non-Lenition-blocking" does not predict [w] realizations. All three realizations appear with considerable frequency in this pseudo-environment, though I argue that the realizations of word-initial /p/ are *not* in free variation in this environment.

A different breakdown of the data reveals an environment that conditions the [w] realization of /p/. Following [g], [x], [ɢ], and [ui], we can see a distinct preference for the [w] realization, as seen in Figure 9:

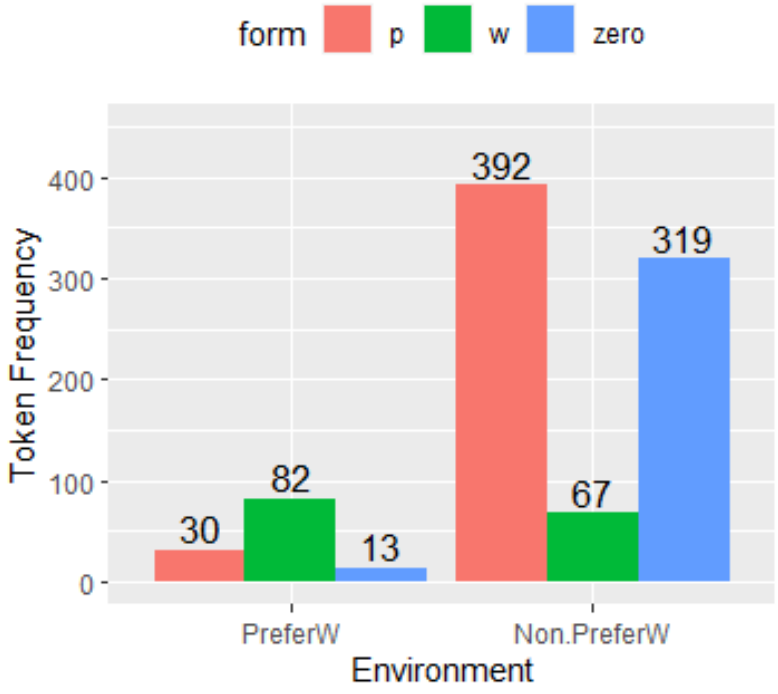


Figure 9: /p/ in Environments Preferring [w]

This conditioning appears to be assimilatory or co-articulatory (or both) in nature. The predictive power of this phonetic environment can, in turn, be implemented as a "fast speech rules" constraint in this grammar:

(18) PREFERW: Assign one violation to every non-[w] realization of /p/ when /p/ follows [g], [x], [ɢ], or [ui]. This penalizes [p] and zero realizations.

6.5 Predicting "zero": The Case for Structurally-Conditioned Allomorphy

With a key [w]-predicting environment accounted for, we are now left to deal with the high token frequency of the zero realization in these functional types (minus seemingly invariant *pi*,

which is never realized as zero within this corpus; see Figure 7). Breaking down the realizations of /p/ by the remaining three functional types, the incredible preference of *pai-* for the zero realization is at odds with the distributions of auxiliary *pol-* and particle *pol*.

In looking for a phonetic environment that conditions the zero realization of *pai-*, we come across [t͡ʃ], [t͡ʃʰ], [t], and [r]. Defining a natural class to accommodate these four sounds at the exclusion of all others in the language proves difficult. The four sounds share neither voicing features nor a manner of articulation, though they are united in being +CORONAL. However, this language has many other +CORONAL sounds that do not appear to trigger the zero realization of /p/ in *pai-* (ex: [n], [s], [t͡s], [ʃ], [tʰ]).

Defining the environment conditioning the zero realization of *pai-* on morphosyntactic grounds not only better accounts for the data, it also explains why [t͡ʃ], [t͡ʃʰ], [t], and [r] appear to constitute a phonetic environment. As a highly-agglutinating, exclusively suffixing, head-final language, the morphosyntactic environments preceding *pai-* often involve the inflectional and derivational morphology of the immediately preceding word. Taken in this light, we can better understand that [t͡ʃ], [t͡ʃʰ], [t], and [r] represent neither a phonetic nor phonological environment conditioning the zero realization of *pai-*, but instead the tail end of a set of morphemes (*-j/-ch, -AAd,* and *-sAAr*) that very frequently precede *pai-*.[15]

This environment is exactly when a converb (specifically, a verb ending in a converbal suffix) precedes auxiliary *pai-*. Converbs are utilized widely throughout the agglutinating SOV languages of Eurasia (Khalkha is one such variety). Haspelmath (1995) defines converbs as "verbal adverbs" or "non-finite verbal forms whose main function is to mark adverbial subordination" (p. 3). The converbal suffixes in Khalkha are a closed set, with the suffixes that can immediately

[15]Note that [t͡ʃ] and [t͡ʃʰ] are allomorphs of the same imperfective converbal suffix

precede *pai-* constituting an even smaller subset. These suffixes are given in (19b-d) and happen to end in [t͡ʃ], [t], and [r]. Auxiliary *pai-* following a converb (that is, a verb with a converbal suffix) yields a periphrastic TAM construction of the following sort:

(19) a. **Simple present tense, no converb**
Saraa unt-na
[sara ʊnt^{h}ə-n]
PN sleep-NONPST

"Saraa sleeps"

b. **Periphrastic progressive, imperfective converb *-j* + *pai-***
Saraa unta-j pai-na
[sara ʊnt^{h}ə-t͡ʃ pain~wain~ain]
PN sleep-IMPF.CNV AUX-NONPST

"Saraa is sleeping"

c. **Periphrastic perfective, perfective converb *-AAd* + *pai-***
Saraa unt-aad pai-na
[sara ʊnt^{h}-at pain~wain~ain]
PN sleep-PERF.CNV AUX-NONPST

"Saraa has been sleeping"

d. **Periphrastic progressive-perfective, progressive converb *-sAAr* + *pai-***
Saraa unt-saar pai-na
[sara ʊnt^{h}-sar pain~wain~ain]
PN sleep-PROG.CNV AUX-NONPST

"Saraa is still sleeping"

Occasionally, focus particles *ch* (additive focus) and *l* (restrictive focus) intervene between the converb and auxiliary *pai-*, as in (20):

(20) **Periphrastic progressive with restrictive focus particle *l***

Pi	*aldaa*	***khii-j***	***l***	***pai-san***
[pi	alda	xit͡ʃ	ɮ	paisən~waisən~aisən]
1SG.NOM	mistake	**do-IMPF.CNV**	**RESTR.PART**	**AUX-PST**

"I **just kept on making** mistakes"

There are three instances of the restrictive focus particle occurring between a converb and auxiliary *pai-* in the present corpus, two in TED1 and one in TED2.[16] In all three instances, the /p/ of auxiliary *pai-* is realized as zero. This is especially noteworthy as [ɮ] is noted as being part of the set of sounds that constitute the lenition-blocking environment outlined in Section 6.3. Though limited in number, the *pai-* of these three examples is patterning more in concert with what we expect of the overall [CONVERB (focus particle) *pai-*] construction (conditioning the zero realization) than what we would expect given the phonetic environment (the lenition-blocking environment, conditioning the [p] realization).

Breaking up *pai-* into two groups, then, allows us to see the degree to which *pai-* prefers the zero realization when in a periphrastic tense construction as in (19b-d) and (20). *Pai-2* represents this morphosyntactically-conditioned instance of *pai-*, while *pai-1* represents *pai-* as it is used elsewhere throughout the language:

(21) a. ***pai-2*, following converbs**

Saraa	*unt-aj*	*pai-san*
[sara	ʊntʰə-t͡ʃ	paisan~waisan~aisan]
PN	sleep-IMPF.CNV	AUX-PST

"Saraa was sleeping"

b. ***pai-1*, elsewhere in the language; here, as a copula**

[16] There are no instances of additive particle *ch* in this construction within the corpus.

Saraa pagsh pai-san
[sara pagʃ paisan~waisan~aisan]
PN teacher COP-PST

"Saraa was a teacher"

Figure 10 shows the breakdown of the high-frequency functional types with *pai-2* separated out from *pai-1*:

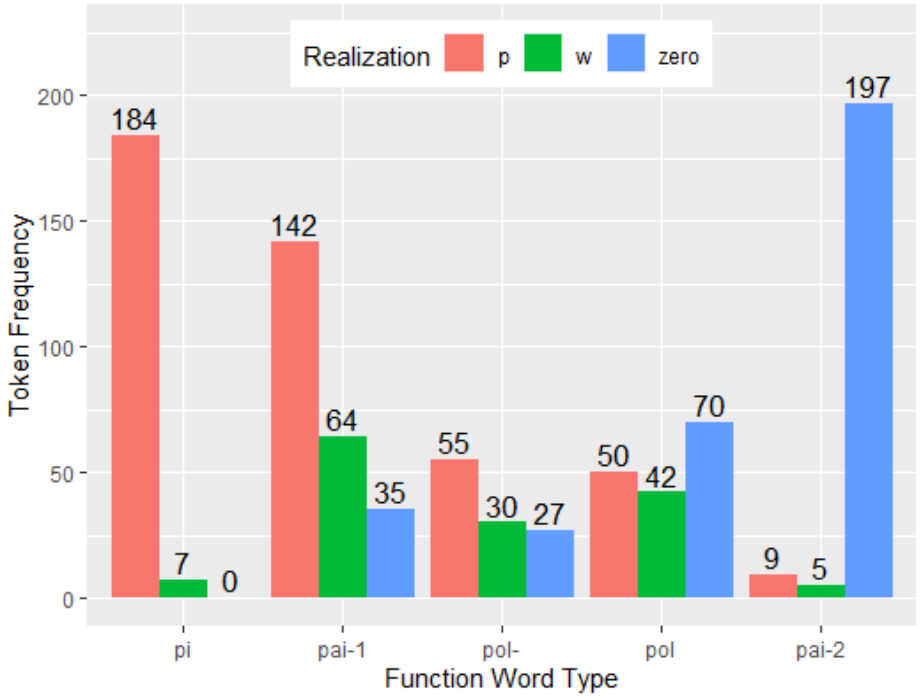


Figure 10: Breakdown of high-frequency function words; *pai-1* and *pai-2* separated

Outside of the periphrastic morphosyntactic context, [t͡ʃ], [t͡ʃʰ], [t], and [r] hold no special relation to the zero realization of any word-initial /p/.

However, it is not the case that this morphosyntactic configuration alone triggers a zero realization of /p/-initial function words. *pol-* (often translated as "to become") shares many distributional facts with *pai-* (often translated as "to be"), including its ability to function both as a copula and as an auxiliary. Like *pai-*, *pol-* can also follow the imperfective converb (ending in [t͡ʃ] or [t͡ʃʰ]), as in the following:

(22) a. **Periphrastic progressive, imperfective converb *-j* + *pai-***

Saraa	*unta-j*	*pai-na*
[sara	ʊntʰə-t͡ʃ	pain~wain~ain]
PN	sleep-IMPF.CNV	PAI-NONPST

"Saraa is sleeping"

b. **Modal construction, imperfective converb *-j* + *pol-***

Saraa	*unta-j*	*pol-no*
[sara	ʊntʰə-t͡ʃ	pɔɮən~wɔɮən~ɔɮən]
PN	sleep-IMPF.CNV	POL-NONPST

"Saraa may/is allowed to sleep"

While *pol-* occurs following the converbal suffix *-j* far less frequently than *pai-2* does within this corpus, what tokens we do have of *pol-* within this configuration demonstrate that word-initial /p/ is not being realized as "zero" to the same extent across both function word types.

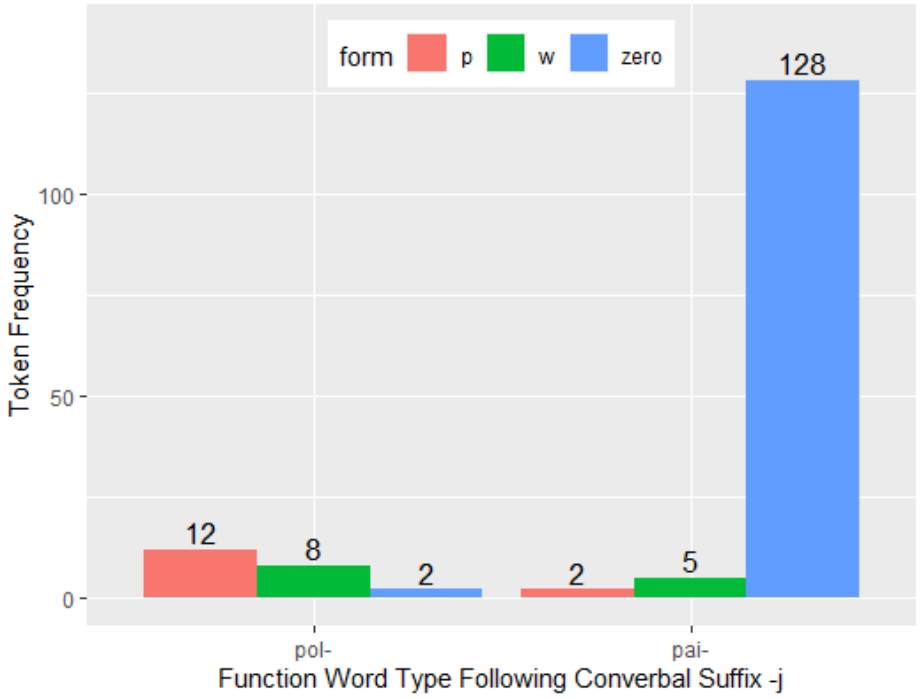


Figure 11: *pol-* vs. *pai-2* Following a Converb

Thus, this structural trigger for the zero realization of /p/ in *pai-2* is specific to morpheme *pai-*.

Because the zero realization of /p/ in *pai-2* occurs with near-categorical frequency across all speakers and all kinds of speech rates, and because no phonological environment works to distinguish *pai-2* from *pai-1*, I argue that *pai-2* is a structurally-conditioned allomorph of *pai-*, and that this fact should be grammatically encoded as the following:

(23) NULLPAI-2: Assign 1 violation to each [p] and [w] realization of *pai-* when *pai-* follows a converbal suffix to form a periphrastic tense. This is the morphosyntactic environment that conditions *pai-2*.

In *pai-2*, we have a Kaisse-style "external sandhi rule". Interestingly, one of Kaisse's (1985) case studies of external sandhi rules is auxiliary reduction in English. Kaisse takes auxiliary reduction to be not a phonological rule, but rather a series of structural rules "stating under what circumstances the reduced allomorph may be inserted" (p. 42). What *pai-2* presents is a phenomenon that exactly parallels English auxiliary reduction, for while *pai-2* is a reduced auxiliary in the phonetic sense, the process of deriving this reduced allomorph proceeds from the morphosyntax and not any phonetic environment.

6.6 Implementing the Grammar

The full grammar outlined in this section has eight constraints: LENITE and PREFERW are markedness constraints that capture the pressure for ease of articulation in naturalistic speech. The four morpheme-specific faithfulness constraints (FAITH-(MORPHEME)), one per functional morpheme) and the LENITION-BLOCKING ENVIRONMENT work to resist these fast speech rules, with the morpheme-specific faithfulness constraint appealing to the underlying representations of their respective morphemes and the LBE specifying a lenition-blocking environment that operates both within and across word boundaries. NULLPAI-2 is the final constraint, a Kaisse-style "external sandhi rule", and it encodes a structurally-conditioned allomorph of *pai-* which strongly favors the zero realization of /p/.

To combat over-fitting of the weights to the data, the model implements a Gaussian prior with parameters $\mu = 0$ and $\sigma = 2$.

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	FAITH-POL	LBE	PREFERW	NULLPAI-2
1.09	4.50	1.83	1.51	0.000	2.69	2.20	3.97

Table 7: Preliminary Constraint Weights

Before settling on this model as the best-fitting analysis of the variable realization of word-initial /p/ in Khalkha, we must first consider the necessity of each constraint. The value of the weight of the faithfulness constraint indexed to particle *pol* (zero to at least 10 decimal places), FAITH-(POL), strongly suggests that the presence of the constraint in the grammar is not significant. To assess the significance of each individual constraint (regardless of its weight in the full model), we can take out each individual constraint, one by one, and compare the resulting smaller models to the full model. In general, a model with more constraints will have a larger log-likelihood, as it features a greater degree of specification. The log-likelihood ratio test can be used to assess whether the increase in log-likelihood is worth the added complexity of extra constraints (which represent an increase in degrees of freedom).

Table 8 presents the results of the log-likelihood ratio tests applied to eight different versions of the present model as compared to the full model. The leftmost column designates the model in question and describes which constraint has been removed. For example, the NO LENITE model is exactly the same as the Full Model, only it lacks the LENITE constraint. The log-likelihood (LL) of each model is presented in the third column. The fourth column calculates the difference in the log-likelihood of each smaller model as compared to the Full Model. Finally, the *p*-value (proportional to the difference in log-likelihood) allows us to assess whether the removed constraint is significant in the grammar. A *p*-value of less than .05 indicates that the removed constraint is significant.

Not only is the FAITH-(POL) constraint not significant, but removing the constraint actually works to improve the log-likelihood of the model.[17] For this reason, I adopt the "No FAITH-

[17]The increase in log-likelihood of the smaller model is at first perplexing, as we expect the larger, more-specified Full Model to feature a strictly better fit. When we strip both the Full Model and the NO FAITH-(POL) Model of their Gaussian priors, this expectation is met:

LL of Full Model, no Gaussian Prior: -468.99

	# of constraints	LL	Δ in LL	*p*-value
Full Model	8	-482.71		
No LENITE	7	-510.07	-27.36	$p < .001$
No FAITH-(PI)	7	-663.31	-180.60	$p < .001$
No FAITH-(PAI-)	7	-530.12	-47.41	$p < .001$
No FAITH-(POL-)	7	-526.43	-43.72	$p < .001$
No **FAITH-(POL)**	**7**	**-481.797**	**0.913**	$\boldsymbol{p = .1766}$
No LBE	7	-544.47	-61.76	$p < .001$
No PREFERW	7	-537.27	-54.56	$p < .001$
No NULLINPERIPHPAI	7	-603.95	-121.24	$p < .001$

Table 8: LL Ratio Tests for Each Constraint

(POL)" model as the best-fitting grammar for the present analysis. The constraint weights of this model are exactly the same as the constraint weights in the previous "Full Model" (given in Table 7), except for the fact that FAITH-(POL) is removed.

This model makes predictions that very closely mirror the observed realizations, as Figure 12 shows:

LL of NO FAITH-(POL) Model, no Gaussian Prior: -469.826
The slight increase in log-likelihood of the smaller model with the Gaussian prior is, therefore, a mathematical consequence of implementing the bias terms.

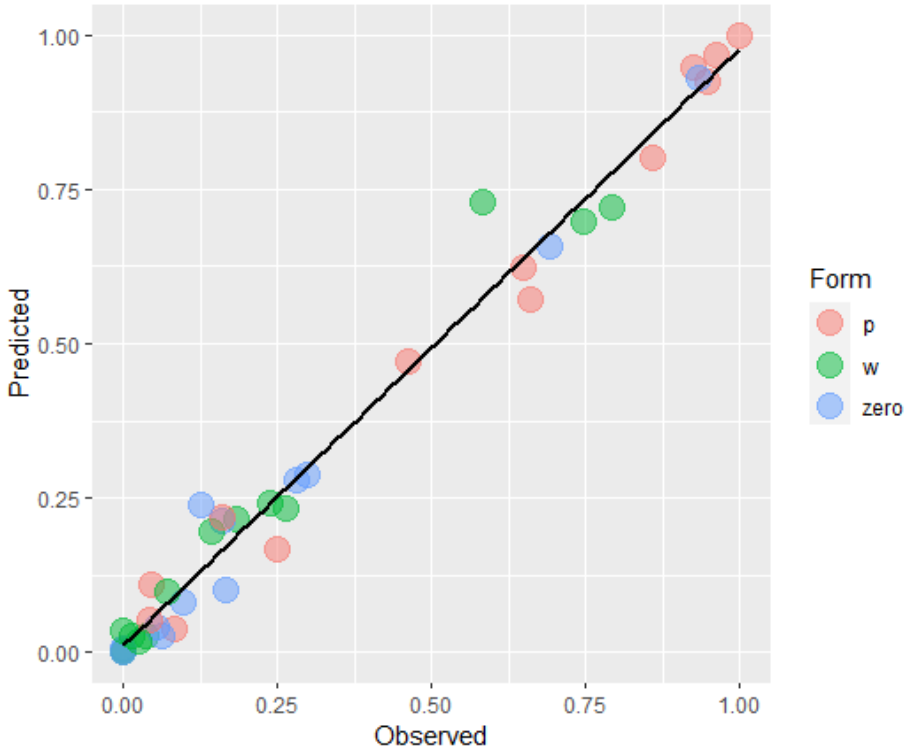


Figure 12: Empirical Observations vs. Predictions of the Grammar
LL = -481.797; y=0.9541x + 0.0153; R^2 = 0.9802

The obvious outlier in Figure 12 is the [w] realization of copula/auxiliary *pol-* in the PREFERW environment, which the grammar over-predicts. While it is observed as [w] 14/24 times (58%) in this environment, it is predicted to occur as [w] 72% of the time. This case is interesting because a potential problem in the design of this model was the capacity for the *underprediction* of [w] realizations resulting from the interaction of LENITE and FAITH-(MORPHEME), neither of which are [w] preferring. To over-predict [w], then, the weight of the PREFERW must be too heavy, but this is at odds with the fact that the other function word types

in this environment are predicted accurately.

Relative speech rate within this environment does not give a clear solution to the over-prediction problem: some of the slowest speech rates (in syllables per second) in this environment yield [w] realizations rather than [p] realizations. In this way, we cannot say that this over-prediction of [w] in the grammar results from failing to account for the effects of a slower speech rate. Instead, the low count of [w] realizations of *pol-* in this environment may come from the fact that many of the data points from this environment (10/24) come from TED4. TED4's speaker had the lowest overall proportion of [w] realizations and appears to be, in general, the most conservative speaker.[18] This trend carries over into the PREFERW environment, as only five of their 10 /p/s in this environment are realized as [w]. Token counts this low preclude any firm conclusions, but this outlier could offer a glimpse into how inter-speaker variation can affect the global picture of the application of fast speech rules.

7 Discussion of the Grammar

7.1 A Formal Separation of Lenition and Auxiliary Reduction

Two constraints in the model encode lenition: LENITE represents the overall pressure to lenite motivated by ease of articulation in connected speech, while PREFERW represents a specific phonetic environment that conditions lenition to [w]. This environment is, in turn, likely motivated by ease of articulation. Both are quintessential Kaisse-style fast speech rules. The model succeeds only when these are separated out from auxiliary reduction, a process that is repre-

[18] See the Appendix (esp. Fig. 13 and Fig. 17) for a by-speaker breakdown of the present grammar. In general, the grammar for TED4's speaker has relatively higher constraint weights for the faithfulness constraints. Furthermore, the grammar as applied to TED4's speaker overpredicts many [w] realizations and underpredicts many [p] realizations.

sented in the NULLINPERIPHPAI constraint. Removing NULLINPERIPHPAI has the numerical effect of an enormous decrease in the log-likelihood of the model, as it receives the second-highest constraint weight in the model. Conceptually, however, the removal of this constraint forces the general tendency toward articulatory ease (the constraint LENITE, specifically) to be the source of all the zero realizations of /p/ in *pai-*. If this were the case, we would expect lenition to zero on a much grander scale throughout the data.

The constraint NULLINPERIPHPAI is, therefore, a grammatical encoding of a non-phonological environment that conditions an allomorph of *pai-*, *pai-2*. Recall from (20) that the conditioning environment in question (the construction [CONVERB + (PARTICLE) + *pai-*]) can still yield a zero realization of /p/ even when it features an otherwise lenition-blocking sound like [ʧ]. This is powerful corroborating evidence for the argument that it is the morphosyntax and not any pseudo-phonetic environment (following [t͡ʃ], [t], or [r], for example) that is conditioning the overwhelming majority of zero realizations of *pai-*.

7.2 A Grammar with no FAITH-(POL)

The removal of the FAITH-(POL) constraint from the model had the numerical result of slightly improving the log-likelihood of the model. The theoretical implications of this removal, however, are more interesting. The set of FAITH-(MORPHEME) constraints was implemented in this grammar to capture the apparently morpheme-specific resistance to the pressure to lenite in connected speech, as not all morphemes showed the same rate of lenition. What does it mean, then, for a FAITH-(MORPHEME) constraint to be insignificant in this grammar?

The answer is fairly straightforward: the distribution of word-initial /p/ in particle *pol* is best modeled as *not* resisting the pressure to lenite. Instead, the interaction of the LENITE,

LENITION-BLOCKING ENVIRONMENT, and PREFERW constraints is sufficient to model the distribution of /p/ in *pol*. Putting *pol* into the broader picture of high token-frequency function words in Khalkha, we see that it is a particle while the others are copulas/auxiliaries (*pai-* and *pol-*) or pronouns (*pi*). An exact definition for "particle" is difficult to pin down across the literature, but such a definition is not required for the present discussion. Instead, what emerges is the perhaps intuitive idea that particle *pol* represents a linguistic unit that is "smaller" than *pai-*, *pol-*, or *pi*.

At this juncture, it is tempting to assign particle *pol* to a category like "clitic". However, recall question particle *pe/we*, described briefly in Section 3. This particle was introduced as a clitic, largely because the distribution of its initial /p/ is phonologized as if it were word-internal. In other words, question particle *pe/we* follows the word-internal [p ~ w] alternation, whereas all of the function words investigated in this study – including *pol* – do not.

In this way, this study offers a diagnostic for distinguishing classes of items variously termed "particle" in Khalkha. The example of *pe/we* shows that there exists a class of items that follow the word-internal [p ~ w] alternation, while there is another class that participates in the word-initial /p/-lenition process defined by the grammar in Section 6, of which *pol* is a prime example.

7.3 Interpreting Constraint Weights: An Appeal to Prosodic Structure

The idea that *pol* could represent some "smaller" linguistic unit may be justified with an appeal to prosodic structure. Karlsson et al. (2020) describe certain function words (namely, postpositions, none of which begin with /p/) as being part of the same prosodic word as the content word that precedes them. Though not the focus of their work, Karlsson et al. even go so far as to

specifically analyze particle *pol* as part of the same prosodic word as the content word that precedes it (p. 215). The data utilized for the present study is naturalistic and thus not optimized for an in-depth prosodic analysis. However, the work of Karlsson et al. (2020) offers an initial step into the interactions of function words and prosodic boundaries. In particular, the weights of the FAITH-(MORPHEME) constraints do present an indication that fast speech rules (LENITE and PREFERW) are operating at varying strengths across different types of prosodic boundaries.

Take, for example, the weight of FAITH-(PI), the constraint receiving the largest weight in the grammar defined in Section 6. Elsewhere in Karlsson et al. (2020), the first-person pronoun is analyzed as being its own prosodic word. Furthermore, being an SOV language, this pronoun is often times at the left edge of an even larger prosodic unit, that of the intonational phrase. *Pai-1* and *pol-* are perhaps somewhere between pronoun *pi* and particle *pol* in terms of boundaries.

The FAITH-(MORPHEME) constraints as implemented in this grammar appear to define three types of boundaries, with *pi* often at the left edge of a larger constituent (perhaps the intonational phrase), and *pai-1* and *pol-* at the left edge of a smaller constituent (perhaps the accentual phrase). Particle *pol*, being incorporated into the same accentual phrase as the content word immediately preceding it, is not at the left edge of a boundary that resists lenition. In this way, the FAITH-(MORPHEME) constraints still represent resistance to lenition, but perhaps this resistance is prosodically defined. That the weights of FAITH-(PAI-) and FAITH-(POL) are similar but not identical suggests that prosodically-defined resistance to lenition does not have to come at the exclusion of morpheme-specific faithfulness.

7.4 Listed Allomorphy in Diachronic Perspective

Whereas Section 7.2 entertained the idea of prosodic units and prosodic boundaries as relevant domains for the application of fast speech rules, the mapping of prosodic structure onto the application of auxiliary reduction is less straightforward. This grammar has proposed that the morphosyntactic configuration that characterizes *pai-2* results in the nearly categorical reduction of [pai-] to [ai-]. Synchronically, this positions *pai-2* as a structurally-conditioned listed allomorph.

A larger, diachronic picture could feature grammaticalization as an analysis of the separation of *pai-1* and *pai-2*, with auxiliary reduction serving as the phonological signature of a structural change. In a very literal sense, this study addresses a type of "grammaticalization", as a constraint encoding *pai-2* as a listed allomorph is directly implemented in the grammatical model defined in Section 6.

However, *pai-2* does interact with more conventional interpretations of grammaticalization. *Pai-2* is involved in forming periphrastic tenses with a lexical verb ending in a converbal suffix (see (19), above). Periphrastic constructions have a strong tendency to coalesce into morphological constructions (typically, affixes) over time. In Khalkha, this could look like the following, wherein a converbal suffix followed by an auxiliary (like *pai-2*) is reanalyzed as a singular TAM affix:[19]

(24) VERB-CONVERB + AUXILIARY > VERB-TAM

Indeed, Brosig (2014) offers an analysis of the progressive tense in Khalkha Mongolian that displays this exact structural reanalysis. Instead of the sequence VERB-*j pai-* forming the pe-

[19] See Hopper and Traugott (2003) for a discussion of this same process in the Romance languages.

riphrastic progressive, Brosig offers the monomorpheme *-jai-* as the synchronic status of the progressive in Khalkha, but this is a break from established convention. Still, Khalkha is not the only Mongolic variety to be analyzed in this way; Khamnigan and Kalmyk also have periphrastic constructions that are written as VERB*-j pai-* and VERB*-j pää-*, respectively, but are analyzed synchronically as progressive monomorphemes *-jai-* and *-jää-*.[20]

Pai-2 does not only follow the imperfective converb *-j* in Khalkha, however. As was discussed in Section 6.5, other converbal endings (*-AAd* and *-sAAr*) can also precede *pai-2* and still trigger auxiliary reduction in *pai-2*. To my knowledge, no treatment of Khalkha – or any other Mongolic language, for that matter – has proposed grammaticalization involving *-AAd* and *-sAAr* and *pai-2*. The difference between *-AAd*/*-sAAr* and *-j* is frequency. Out of the 211 tokens of *pai-2* in the present corpus, 135 of them occurred after *-j*. In this way, structurally-conditioned allomorphy alone is perhaps not indicative of grammaticalization; instead, a structurally-conditioned allomorph when it very frequently co-occurs with another linguistic element may be a better candidate for undergoing grammaticalization.

In any case, the fact remains that the initial /p/ in *pai-2* can still be realized as [p] or [w]. This means that an underlying /p/ is still present in *pai-2*. In this way, any analysis involving the grammatical listing of *pai-2* and/or the potential grammaticalization of certain instances of *pai-2* represents changes-in-progress rather than a categorical or "completed" shift.

8 Conclusions

This study has presented a systematic look into the variable realization of word-initial /p/ in Khalkha Mongolian. Though word-initial /p/ is always realized as [p] in careful speech, ap-

[20] See Janhunen (2003) and Blasing (2003) for discussion of Khamnigan and Kalmyk, respectively.

proaching word-initial /p/ from the lens of connected speech reveals a system of patterned variation, with all four speakers represented in the corpus data engaging with word-initial /p/ in roughly the same, systematized way. A key finding of this study is that variation in the realization of /p/ is largely a property of function words, with content words displaying minimal variation outside of fixed expressions or collocations.

This study positions function words as being especially susceptible to the articulatory pressures that rise to prominence in connected speech. Markedness constraints favoring lenition serve as a way to capture these pressures and serve as examples of the "fast speech rules" put forward by Kaisse (1985). Lenition is to be separated from auxiliary reduction, the latter being a process that applies only in certain morphosyntactic configurations. This auxiliary reduction is an example of a Kaisse-style external sandhi rule. Initial /p/ in this study is affected by both lenition and auxiliary reduction, with lenition affecting all function words (but capable of being blocked at rates particular to each individual morpheme) and auxiliary reduction affecting only *pai-2*. A grammar that formally separates these two processes is necessary to make accurate predictions.

Though the present analysis suffices to make reasonable predictions, a few limitations warrant direct attention. First, /p/ was analyzed as having three realizations, [p], [w], and zero. This method of categorizing, while practical given the scope of this investigation, obscures many of the phonetic details of realizations of /p/. Second, the possible prosodic interpretation of the weights assigned to the lexically-indexed FAITH-(MORPHEME) constraints begs further investigation into the intonational phonology of Khalkha, a task that is not easily made actionable given the naturalistic character of the present data. Next, positing the morphosyntactic conditioning of a particular allomorph (i.e., *pai-2*) was an essential component of the model here

presented. Having allomorphs listed separately in the lexicon gives rise to the expectation that, over time, the meanings and distributions of the full and reduced forms will diverge. To give a full Kaisse-style treatment to the phenomenon of auxiliary reduction, a more thorough investigation into the distribution of occurrences of the reduced form versus the full form is required. Finally, in addition to constituting an incredibly small corpus, the speech analyzed in this study belongs to a very specific, presentationally-oriented genre – that of the TEDx Talk. Though positioned as "naturalistic speech", these data do not necessarily reflect language as it is utilized for everyday communication. Still, the fact that so much variation is present within this genre indicates that this variation is alive and well in everyday speech. The present study is offered as a point of comparison for investigations into different genres of Khalkha connected speech.

The limitations of this study thus serve as sources of inspiration for future work on Khalkha word-initial /p/ along at least four lines of linguistic inquiry: phonetics, prosody, morphosyntax, and sociolinguistics. In general, the spirit of investigations into connected speech involves operating both at and across the linguistic interfaces as they are traditionally defined. In this way, the interaction of the phonologically-oriented "fast speech rules" and the morphosyntactically-defined auxiliary reduction process demonstrate that both components are essential to unpacking to the patterned distribution of word-initial /p/ in Khalkha.

Appendix

Individual Grammars

This appendix features a grammar for each individual speaker. The set of constraints for each individual grammar is the NO FAITH-POL model discussed in Section 6. While the set of constraints, hyperparameters, and bias terms are exactly the same as in the aggregate grammar defined in Section 6, the weights of the constraints are fitted for each individual speaker. The model performs moderately well for the speakers of TED1, TED2, and TED4, but performs poorly for the speaker of TED3. One reason for the poor performance for the speaker of TED3 is likely overall token count: TED3 features only 81 tokens across all four high-frequency function word types. This is as compared to the 296, 234, and 296 tokens for the speakers of TED1, TED2, and TED4, respectively.

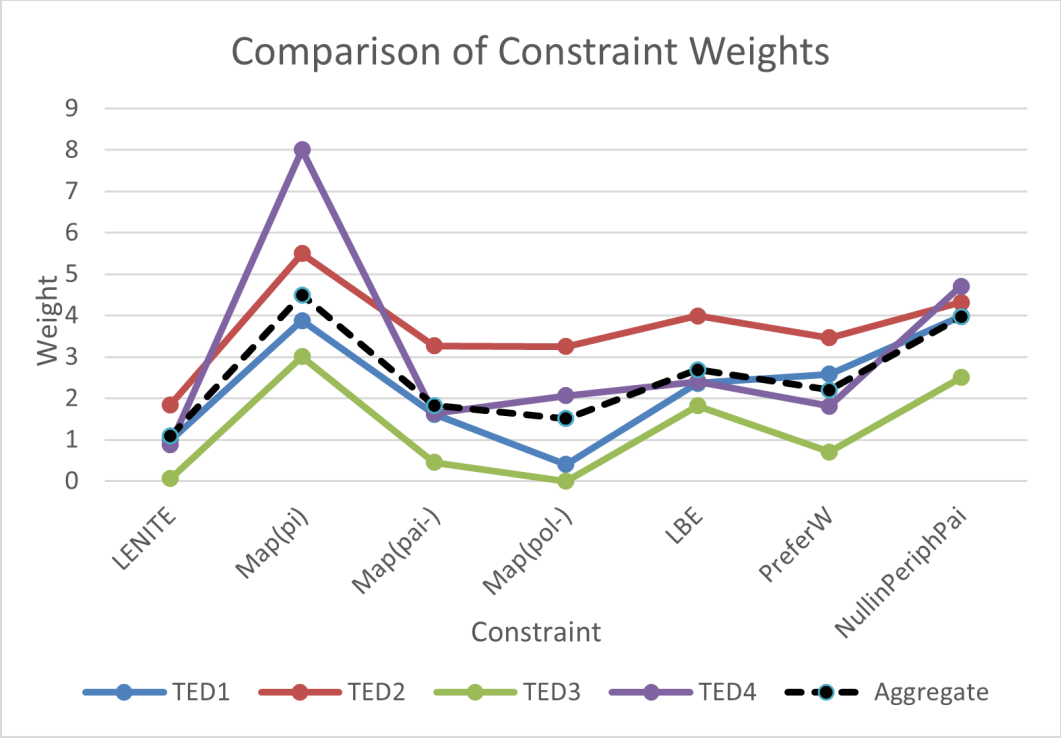


Figure 13: Comparison of Constraint Weights across Talks

Aggregate

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	LBE	PREFERW	NULLPAI-2
1.09	4.50	1.83	1.51	2.69	2.20	3.97

Table 9: Constraint Weights, All Talks

TED1

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	LBE	PREFERW	NULLPAI-2
.99	3.88	1.62	.40	2.36	2.58	3.99

Table 10: TED1 Constraint Weights

TED2

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	LBE	PREFERW	NULLPAI-2
1.85	5.50	3.23	3.25	3.99	3.47	4.32

Table 11: TED2 Constraint Weights

TED3

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	LBE	PREFERW	NULLPAI-2
.06	3.01	.45	0.00	1.82	.71	2.51

Table 12: TED3 Constraint Weights

TED4

LENITE	FAITH-(PI)	FAITH-(PAI-)	FAITH-(POL-)	LBE	PREFERW	NULLPAI-2
.89	8	1.64	0.53	2.41	1.81	4.71

Table 13: TED3 Constraint Weights

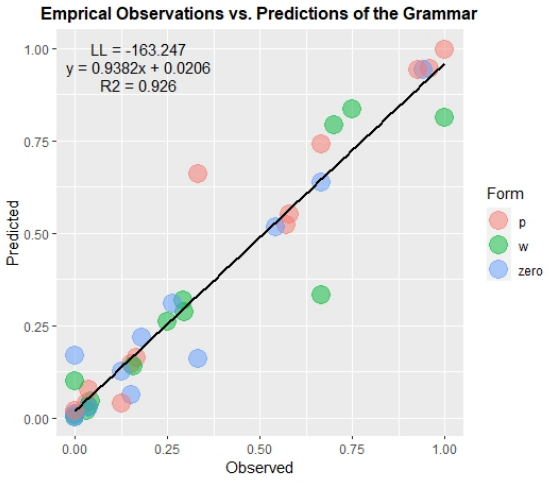


Figure 14: TED1 Grammar

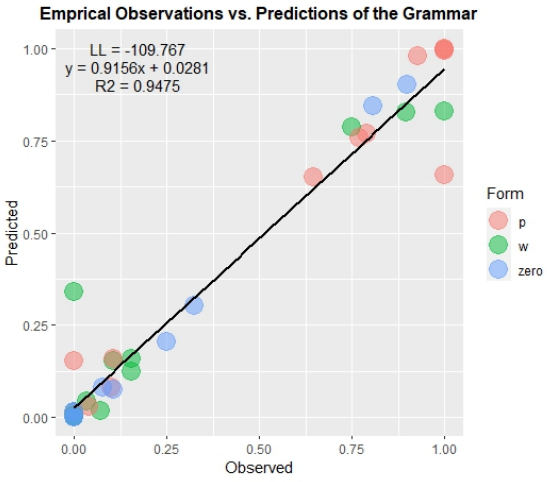


Figure 15: TED2 Grammar

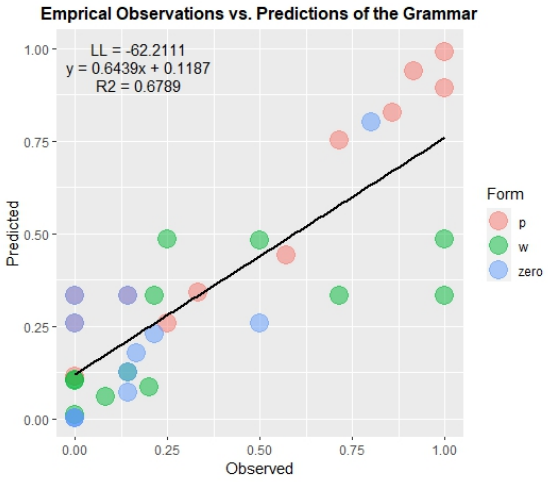


Figure 16: TED3 Grammar

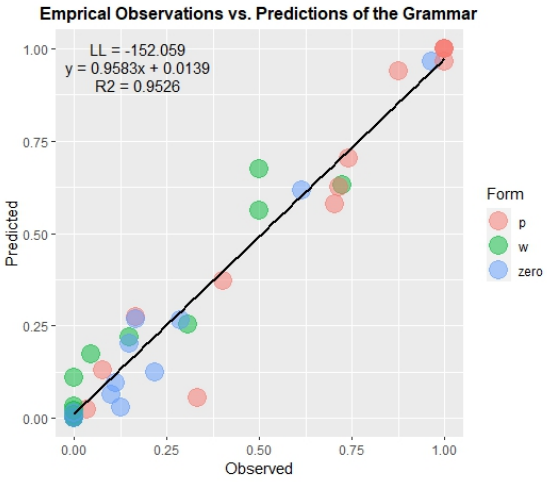


Figure 17: TED4 Grammar

References

Atwood, C. P. (n.d.). Transliteration of Cyrillic Khalkha. https://www.thlib.org/tools/scripts/wiki/Transliteration%20Schemes%20Mongolian%20Cyrillic%20Script.html

Battsogt, D., & Önörbayan, T. (2008). Awiazüi [Phonology]. In *Orchin tsagiin mongol khel modern [Modern Mongolian]*. Interpress.

Beckman, J. N. (1998). *Positional faithfulness* (Dissertation), University of Massachusetts Amherst. https://scholarworks.umass.edu/dissertations/AAI9823717

Blasing, U. (2003). Kalmuck. In *The Mongolic Languages*. Routledge.

Boersma, P., & Weenick, D. (2023). Praat: Doing phonetics by computer [Computer program]. http://www.praat.org/

Brosig, B. (2014). The aspect-evidentiality system of Middle Mongol. *Ural-Altaic studies*. Retrieved July 10, 2022, from https://www.academia.edu/77213974/The_aspect_evidentiality_system_of_Middle_Mongol

Bybee, J. (2001). *Phonology and language use*. Cambridge University Press.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition [Publisher: Linguistic Society of America]. *Language*, *82*(4), 711–733. Retrieved May 30, 2023, from https://www.jstor.org/stable/4490266

Chinggaltai & Lapsangwangtan. (1952). *A grammar of the Mongol language* (Rev). Evangelical Alliance Mission.

Goldwater, S., & Johnson, M. (2003). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Workshop on Variation within Optimality Theory* (pp. 111–120). Retrieved June 3, 2023, from https://www.research.ed.ac.uk/en/publications/learning-ot-constraint-rankings-using-a-maximum-entropy-model

Hall, K. (2013). A typology of intermediate phonological relationships. *The Linguistic Review*, *30*. https://doi.org/10.1515/tlr-2013-0008

Harnud, H. (2003). A Basic Study of Mongolian Prosody. *Helsinki: Publications of the Department of Phonetics, University of Helsinki*, *A*(45).

Haspelmath, M. (1995). The converb as a cross-linguistically valid category. In M. Haspelmath & E. König (Eds.), *Converbs in Cross-Linguistic Perspective: Structure and Meaning of Adverbial Forms - Adverbial Participles, Gerunds* -. Mouton de Gruyter.

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139165525

Janhunen, J. (2003). Khamnigan Mongol. In J. Janhunen (Ed.), *The Mongolic Languages*.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure* (pp. 229–254). John Benjamins Publishing Company. https://doi.org/10.1075/tsl.45.13jur

Kaisse, E. M. (1985). *Connected Speech: The Interaction of Syntax and Phonology* (First Edition). Academic Pr.

Karlsson, A. M. (2014). The intonational phonology of Mongolian. In S.-A. Jun (Ed.), *Prosodic Typology II: The Phonology of Intonation and Phrasing* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567300.003.0007

Karlsson, A. M., Güneş, G., Rahmani, H., & Jun, S.-A. (2020). South West and Central Asia. In C. Gussenhhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198832232.013.13

Kullmann, R., & Tserenpil, D. (2015). *Mongolian Grammar* (2015 Revised Edition). Kullnom Verlag.

Legden, T., & Luethy, S. (2009). *Sain Baina uu? Mongolian Language Textbook One*.

Lessing, F. (1960). *Mongolian-English Dictionary*. Unviersity of California Press.

MacKenzie, L., & Yang, C. (2013). English auxiliary realization and the independence of morphology and phonetics. In A. Freeman (Ed.), *University of pennsylvania working papers in linguistics 19.2: Selected papers from nwav 41* (pp. 121–129).

Pater, J. (2010). Morpheme-specific Phonology: Constraint Indexation and Inconsistency Resolution. In S. Parker (Ed.), *Phonological Argumentation: Essays on Evidence and Motivation.* (pp. 123–154). Equinox.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint INteraction in Generative Grammar*. Blackwell.

Svantesson, J.-O. (2020). Khalkha Mongolian. In M. Robbeets & A. Savelyev (Eds.), *The Oxford guide to the Transeurasian languages* (pp. 334–389). Oxford University Press. https://doi.org/10.1093/oso/9780198804628.003.0022

Svantesson, J.-O., Tsendina, A., Karlsson, A., & Franzén, a. V. (2005). *The Phonology of Mongolian*. Oxford University Press.

Tömörtogoo, D. (2017). *Mongol khelnii tüükhen awia züin ündes [principles of Mongolian Historical Phonology [Mongol]* (2nd ed.). Soyombo Printing.

Ueta, N. (2018). *Mongorugo no boin ni kansuru sougouteki kenkyuu [Comprehensive Research into Mongolian Vowels]* (Doctoral dissertation). Kyoto University.

Walker, W. (1997). Mongolian Stress, Licensing, and Factorial Typology [Publisher: University of California, Santa Cruz]. https://doi.org/10.7282/T3VD6WHR

Weinreich, U., Labov, W., & Herzog, M. (1968). *Emprical Foundations for a Theory of Language Change*. Unviersity of Texas Press.