# Coarticulatory propensity in Khalkha Mongolian

Ushasi Banerjee[*], Indranil Dutta[*] and Irfan S.[**]

[*]The EFL University
[**]University of Illinois at Urbana-Champaign
banerjeeushashi@gmail.com, indranil@efluniversity.ac.in, irfans2@illinois.edu

## Abstract

Acoustic variation brought about by V-to V coarticulation needs to be perceptually compensated by listeners, the lack of which results in a language developing vowel harmony. This could either be mechanico-inertial where the articulatory gestures of $V_1$ perturb the $V_2$ vowel space, or anticipatory where the planning of the following vowel perturbs the space of the former. Khalkha Mongolian exhibits vowel-feature sharing of two kinds: [ATR] and Round. In this paper we examine non-harmonic sequences in Mongolian to verify coarticulatory directionality and propensity. The results show that these non-harmonic sequences exhibit a directionality and propensity which is quite the opposite of those in harmonic sequences. This finding suggests that coarticulation works as an antithetical force in non-harmonic patterns, the primary motive of which would be to maintain contrast.

## 1 Introduction

Continuous acoustic variation in speech is understood to be brought about by overlapping articulatory gestures (Öhman, 1966). Vowel-to-Vowel (V-to-V) coarticulation is a special case of non-contiguous coarticulation which helps model acoustic patterns resulting from anticipatory articulatory planning on the one hand, and carryover coarticulation on the other. Acoustic variation in speech could also be attributed to vocal tract shape differences, size, shape, and density of the segmental inventory (Manuel, 1990; Manuel, 1999). A uniform way of understanding the variation has been to locate it in the dynamic and kinematic differences in articulatory overlap. These differences have been modelled variously in space and time (Kirchhoff and Bilmes, 1999; Deng et al., 2006). In $V_1CV_2$ sequences, the standard assumption has been that the acoustics of $V_1$ is perturbed by the anticipatory production planning of $V_2$ gestures, while the mechanico-inertial properties of articulatory gestures of $V_1$ are known to be responsible for the acoustic perturbation of $V_2$. Coarticulatory acoustic variation in V-to-V sequences once phonologized are known to provide the conditioning for the development of vowel harmony patterns (Przezdziecki, 2000; Ohala, 1994). Typically, in non-harmonic languages, acoustic variation resulting from coarticulation is perceptually compensated by listeners. Lack of such perceptual compensation has been shown to be responsible for the development of vowel harmony patterns (Beddor et al., 2002).

Coarticulatory propensity is the spatial and temporal extent of coarticulation between or across segments. Broadly speaking, coarticulatory resistance or the intrinsic resistance to coarticulation that segments with variable articulatory gestures possess has also been shown to affect acoustic variation (Martin and Bunnell, 1981; Recasens, 1984). Nature and extent of coarticulation is also governed by language specific properties such as syllable structure (Manuel, 1999), coarticulatory resistance of the intervening consonant in V-to-V segment sequences (Recasens et al., 1995). In addition, the nature of contrast between segments has also been shown to affect the magnitude of coarticulation in V-to-V sequences, especially in languages that exhibit vowel harmony of the type where distinctive features are shared between vowels (Dutta et al., 2017) and in those where vowel harmony results in vowel copying (Dutta et al., 2016).

While there has been a substantial amount of work on V-to-V coarticulation in both harmony and non-harmony languages, in this paper we examine the magnitude and extent of V-to-V coarticulation in Khalkha Mongolian. Khalkha Mongolian exhibits both Advanced Tongue Root [ATR]



Proc. of ICON-2017, Kolkata, India. December 2017 ©2017 NLPAI, pages 356–361

harmony and rounding harmony. Unlike, suffixal harmony systems like Assamese (Mahanta, 2008), Bengali (Shamim, 2011), and Telugu (Wilkinson, 1974; Kissock, 2009) where the direction of the harmony is right-to-left, Khalkha, primarily, but not exclusively is a stem internal harmony system where the direction of the harmony proceeds from left-to-right. Specifically, we examine the effect of formants from $V_1$ on $V_2$ and vice-versa using a set of linear fixed effect model. We observe that the direction of coarticulation in non-harmonic sequences is greater in the anticipatory direction which is opposite to the direction of suffixal vowel harmony in Khalkha Mongolian. We also find this effect only in the $F_1$ values and not in the $F_2$, which suggests that the coarticulatory formant perturbation runs in the tongue height dimension and not in the tongue front-back dimension. In section 2, we present a detailed phonological account of the vowel harmony system in Khalkha Mongolian. Following that in section 3, we provide details of our experimental methodology and speech materials. In section 4, we present the primary findings of our study, and we conclude in section 5 by motivating the need to understand the nature of vowel contrasts that mitigate the coarticulatory propensity and directionality, especially in vowel harmony languages.

## 2 Vowel Harmony in Khalkha Mongolian

Vowel harmony is a phonological process that restricts the co-occurrence of vowels (usually) within a non-compound word. The Mongolian vowel inventory is unevenly distributed in features, consisting of seven vowels *i, e, a, o, ɔ, u and ʊ*. These can be classified as pharyngeal: *i, e, u, o* and non-pharyngeal:*a,ʊ, ɔ*. As can be noticed, *i* does not have a pharyngeal counterpart. We find two types of vowel harmony in Khalkha Mongolian : [ATR] and Rounding. It affects roots as well as derivational/inflectional suffixes. While the initial vowel is assumed to be specified, the non-initial vowels can be underlyingly /i/, /E/, or /U/. The archiphonemes /U/ and /E/ have missing features that are filled in through vowel harmony. /U/ undergoes just [ATR] harmony, becoming either /u/ or /ʊ/, while, /E/ undergoes both [ATR] and [round] harmony, becoming /e/, /a/, /o/, or /ɔ/ (Godfrey, 2012).

### 2.1 [ATR] Harmony

ATR Harmony distinguishes tense and lax vowels in Mongolian, which Svantesson (2005) refers to as ‘pharyngeal’ and ‘non-pharyngeal’. Vowels in non-compound words must share values for [+/- ATR], depending on the root/stem vowel (at the morpho-phonemic level).

- Trigger vowels: *o, ɔ, u, ʊ, a, e* (in initial positions)
- Target vowels: Archiphonemes /E/ and /U/

**Alternations:**

- /E/: /e/, /a/, /o/, or /ɔ/
- /U/: /u/, or /ʊ/

The vowel *i* in the initial syllable also forces the following vowels in the non-compound word to be [+ATR] (Svantesson, 1985).

In Nevins (2010) and Godfrey (2012), the harmony is conceptualized as a search-copy mechanism by ‘needy’ vowels instead of there being harmony ‘trigger’s. In [ATR] harmony, the search proceeds leftward and looks for the nearest contrastive instance of [ATR]. Once found, the value is copied. If none is found, default [+ATR] is inserted.

### 2.2 Rounding Harmony

This phonological process influences vowels to surface as rounded when the neighbouring vowel (the root/stem vowel for Khalkha Mongolian) is rounded. However, in most cases, conditions referring to tongue body position (height and/or backness) are imposed on either the triggering element, the target, or both (Kaun, 1995).

In Khalkha Mongolian, we observe two conditions for rounding harmony:

- *The trigger must be nonhigh.*
- *The trigger and target must agree in height.*

This kind of a system is similar to one seen in Sibe, a Tungusic language of China (Li., 1996).

- Trigger vowels: /o/, and /ɔ/
- Target vowels: Archiphoneme /E/

The archiphoneme /E/ surfaces as open rounded vowels *o, or ɔ* in the non-initial syllable when preceded by the same vowel. An open vowel that follows a non-open rounded vowel (*u, u*) must be unrounded (*e or a*) (Svantesson, 1985)

### 2.3 Transparent *i*

Transparent vowels are those vowels that may intervene between the trigger and the target of harmony even when they bear the opposite value for the harmonizing feature (Benus, 2010). Non-initial *i* in Mongolian is transparent, i.e., it is completely ignored by vowel harmony; neither does it participate in vowel harmony, nor does it block the process. *i* is the only vowel phoneme that is fully specified in non-initial vowels.

**Example:**
/po:r-ig-**E**/ po:r-ig-o *po:r-ig-e
(gloss) 'kidney-**ACC-RFL**'

In Benus (2010), phonetic and phonological investigation of transparent vowels under a dynamic model show that transparent vowels are in fact integral parts of harmonic domains. The phonetic properties of transparent vowels get integrated over phonological selection of suffixal vowels.

### 2.4 Opaque *u and* ʊ

Opaque vowels, in contrast, require a local agreement relationship between the trigger and the target, i.e. there can be no intervening vowel (Benus, 2010). In Khalkha, intervening non-open velar vowels block rounding harmony. Not only are they not affected by vowel harmony, they also prevent rounding harmony to spread across them.

**Example:**

/ɔr-ʊʒ-ʒ**E**/ ɔr-ʊʒ-ʒa *ɔr-ʊʒ-ʒɔ
(gloss) 'enter-**CAUS-DPST**'

This opacity, however, is restricted to rounding harmony; these segments don't behave so in the process of [ATR] harmony (Godfrey, 2012). Phonetic factors have also been implicated in grounding the phenomenon of opacity in 'tongue root harmony systems (Archangeli and Pulleyblank, 1994).

## 3 Materials and methods

Four female native Khalkha Mongolian speakers were recorded at The EFL University. The material block consisted of fifty-nine target words and six distractors in carrier phrases of the type "pi <target word/distractor> gesen". Four repetitions of each block were recorded for all speakers with a five minute break within each block. Recordings were conducted in a quiet environment. The total number of critical items was 59*4*4=944. The data was presented to the speakers in Cyrillic script. The recorded speech was segmented and annotated manually in Praat (Boersma and Weenink, 2009).

Ten vowel formants from each $V_1CV_2$ sequence were extracted using a Praat script (Boersma and Weenink, 2009) FormantPro (Xu, 2007 2015). Extracted formants where normalized by using the Lobanov method to eliminate specific speaker effects (Lobanov, 1971).

$$F^N_{n[V]} = \frac{(F_{n[V]} - MEAN_n)}{S_n} \tag{1}$$

Where $F^N_{n[V]}$ is the normalized value for formant n of vowel V. $MEAN_n$ is the mean value for formant n for the speaker in question and $S_n$ is the standard deviation for the speaker's formant n.

Linear Mixed Effects model from the lme4 package in R (Bates et al., 2015) was implemented on the formant data to ascertain the effects of the distal and proximal formants on coarticulation. Distal formant data consisted of $F_1$ and $F_2$ data from vowel midpoints representing the steady-state formants where the V-to-V coarticulation effects are most distal from the formant perturbations due to the intervening consonant. Proximal formant data consisted of $F_1$ and $F_2$ data from vowel offsets of $V_1$ and vowel onsets of $V_2$.

## 4 Statistical analyses and results

Vowel $F_1$ and $F_2$ plots on an X-Y euclidean plane where plotted using the PhonR package (McCloy, 2016). The formant data visualizations provided information about the relative positions of the vowels and the areas of the vowel polygons for $V_1$ and $V_2$. In Fig.1 below the left panel represents the steady-state positions and the vowel area polygons for $V_1$ and the right panel represents the same for $V_2$. We make two observations with respect to the relative positions and the polygonal areas. The mid-front vowel *e* is raised and is overlapped significantly with the high front vowel *i*, while the *e* is lower than the *i*. This pattern suggests that the $F_1$ and $F_2$ values of *e* are perturbed by the presence of a high vowel in the $V_2$ position, which in turn implies a stronger anticipatory effect, at least for the vowel *e*. ?.

all the V1s at 50%

all the V2s at 50%

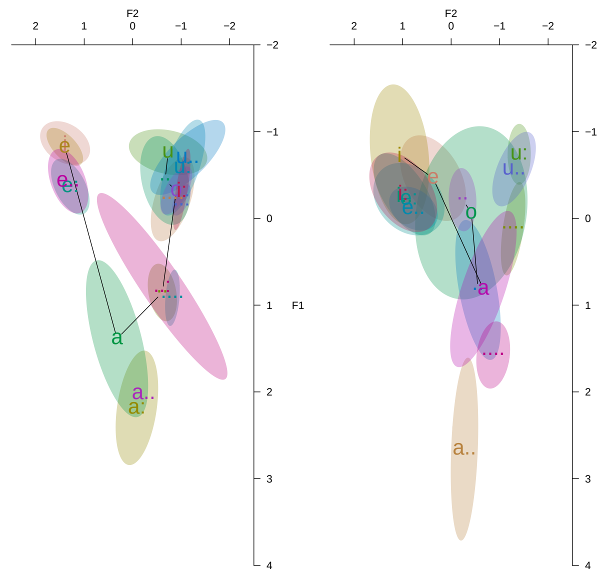


Figure 1: Steady state distal formants: Vowel ellipses and polygons from $V_1T6$, i.e., $V_1$ mid-point (Left panel) and $V_2T_5$; i.e.,$V_2$ mid-point (Right panel)

In Fig.2, the left panel represents proximal positions and vowel area polygons for $V_1$ and the right panel represents those for $V_2$. It shows that in spite of more overlapping effect due to the presence of a consonant, the polygon area of $V_1$ is still larger than that of $V_2$. From this, we make the observation that the effect of $V_2$ on $V_1$ is robust across their different positions in time (vowel-mid position, onset and offset). In Table 1., we present the vowel polygon areas from these models (distal and proximal). The values show that for all subjects, the vowel polygon area for $V_1$ is larger in comparison to that of $V_2$.

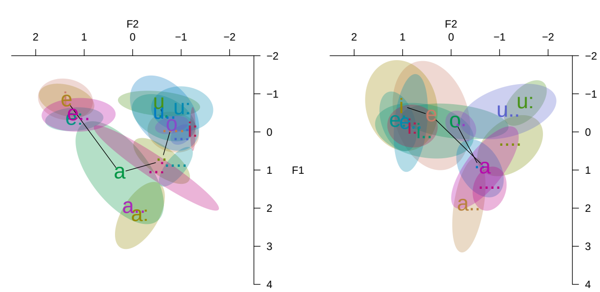


Figure 2: Proximal formants: Vowel ellipses and polygons from $V_1T6$, i.e., $V_1$ mid-point (Left panel) and $V_2T_5$; i.e.,$V_2$ mid-point (Right panel)

We present results from a linear mixed effects model which tested the variable and mutual effect of $F_1$ and $F_2$ values from distal and proximal V positions in order to quantify the coarticulatory propensity in non-harmonic $V_1CV_2$ sequences.

| Polygon Area | Subject f1 | Subject f2 | Subject f3 | Subject f4 |
|---|---|---|---|---|
| *Distal* | | | | |
| V1 at T6 | **3.48591494** | **4.1172744** | **2.5307905** | **2.8948019** |
| V2 at T5 | 0.09251035 | 0.6124816 | 0.1368078 | 0.3408551 |
| *Proximal* | | | | |
| V1 at T10 | **2.725822** | **3.827627** | **1.998709788** | **2.1165014** |
| V2 at T1 | 0.3790905 | 0.536976 | 0.008869308 | 0.1819477 |

Table 1: Vowel polygon areas of $V_1$ and $V_2$ from distal and proximal formant measures

effect of the F1 value

effect of the vowel ID

In the first model, we test the effect of $F_1$ from T5 of $V_2$ and a fixed effect of $V_2$ on the $F_1$ from $V_1$ at T6. This model returns a t=3.543 with no significant effect of $V_2$. In the second model, we test the effect of $F_1$ from T6 of $V_1$ and a fixed effect of $V_1$ on the $F_1$ from $V_2$ at T5. This model returns a t=3.233 and a t=2.185 for vowel $V_1$, /u:/. Here Subject and Item function as random effects. Our results indicate that there is significant covariation of $F_1$ values with slightly greater effect of $V_2$ on $V_1$ suggesting that the directionality of coarticulation as seen in the $F_1$ values is in the anticipatory direction, which is opposite to the direction of the suffixal harmony system in Khalkha Mongolian. Similar models on $F_2$ do not show significant effects, with t=-0.044 and t=-0.020 for $F_2$ of $V_1$ at T6 and $F_2$ of $V_2$ at T5, respectively.

## 5 Contrast and coarticulatory propensity in vowel harmony systems

Stem-internal and stem + suffix harmonic sequences in Mongolian exhibit a left-right directionality. Lack of perceptual compensation (Beddor et al., 2002) is known to contribute to the development of vowel harmony systems, to the extent that it might regress to a radical case of vowel-copying, such as in Telugu (Dutta et al., 2016; Kissock, 2009; Sailaja, 1999). In this paper, we argue that in non-harmonic Khalkha Mongolian sequences the coarticulatory propensity is greater in the anticipatory direction, opposite to the direction of both ATR and rounding harmony, in an effort to maintain contrast, where harmonic sequences may lead to contrast obliteration in terms of featural neutralization. The directionality of coarticulatory propensity, also suggests that articulatory planing seeks to preserve contrast. Unlike the intuitive notion of coarticulation as a force that contrives against contrast, Mongolian shows that it might also be one that does not, and indeed it augments the force of contrast. Coarticulation could be viewed as a force that is not merely functioning at the production end of language but also at the planning end. In vowel harmony languages, where lack of perceptual compensation for coarticulatory acoustic variation may have lead to the development of complex feature sharing in V-to-V contexts, the V-to-V coarticulation patterns in non-harmonic sequences seek to enhance contrast by providing perceptual advantage through variable coarticulatory propensity.

## References


D. Archangeli and D. Pulleyblank. 1994. *Grounded Phonology*. Current Studies in Linguistics. MIT Press.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Patrice Speeter Beddor, James D. Harnsberger, and Stephanie Lindemann. 2002. Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics*, 30(4):591 – 627.

S. Benus. 2010. *Dynamics and Transparency in Vowel Harmony*. Universal Publishers.

Paul Boersma and David Weenink. 2009. Praat: doing phonetics by computer (version 5.1.13).

Li Deng, Dong Yu, and Alex Acero. 2006. A bidirectional target-filtering model of speech coarticulation and reduction: Two-stage implementation for phonetic recognition. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(1):256–265.

Indranil Dutta, S Irfan, and KRS Harsha. 2016. Using ANNs for vowel identification from V-to-V coarticulation in non-harmonic VCV sequences. Presented at the 15th Conference on Laboratory Phonology, Ithaca, NY.

Indranil Dutta, Irfan S., Pamir Gogoi, and Priyankoo Sarmah. 2017. Nature of contrast and coarticulation: Evidence from Mizo tones and Assamese vowel harmony. In *Proceedings of Interspeech 2017*, pages 224–228.

Ross Godfrey. 2012. Opaque intervention in Khalkha Mongolian vowel harmony: A contrastive account. Technical report, McGill Working Papers in Linguistics, Volume 22.1,Winter 2012.

Abigail R. Kaun. 1995. *The Typology of Rounding Harmony: An Optimality Theoretic Approach*. Cambridge University Press.

Katrin Kirchhoff and Jeff A Bilmes. 1999. Statistical acoustic indications of coarticulation. In *Proc. ICPhS*, volume 99. Citeseer.

Dworak Kissock. 2009. Telugu vowel assimilation: Harmony, umlaut, or neither? In *Proceedings of the Seventeenth Manchester Phonology Meeting*.

Bing Li. 1996. *Tungusic vowel harmony: description and analysis*. Holland, Institute of Generative Linguistics, Amsterdam.

B. M. Lobanov. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, 49(2B):606–608.

Shakuntala Mahanta. 2008. *Directionality and locality in vowel harmony: With special reference to vowel harmony in Assamese*. Netherlands Graduate School of Linguistics.

S. Manuel. 1990. The role of contrast in limiting vowel-to-vowel coarticulation indifferent languages. *Journal of the Acoustical Society of America*, 88:1286–1298.

S. Manuel. 1999. Cross-language studies: Relating language-particular coarticulation patterns to other language-particular facts. In W. Hardcastle and N. Hewlett, editors, *Coarticulation: Theory, data and techniques*. Cambridge University Press.

James G. Martin and H. Timothy Bunnell. 1981. Perception of anticipatory coarticulation effects. *The Journal of the Acoustical Society of America*, 69(2):559–567.

Daniel R. McCloy. 2016. Tools for phoneticians and phonologists phonR. R package version 1.14.4.

Andrew Nevins. 2010. *Locality in vowel harmony*, volume 55. MIT Press.

John J Ohala. 1994. Towards a universal, phonetically-based, theory of vowel harmony. In *Third International Conference on Spoken Language Processing*.

Sven EG Öhman. 1966. Coarticulation in vcv utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America*, 39(1):151–168.

Marek Przezdziecki. 2000. Vowel harmony and vowel-to-vowel coarticulation in three dialects of Yoruba. *Working Papers of the Cornell Phonetics Laboratory*, 13:105–124.

Daniel Recasens, Jordi Fontdevila, and Maria Dolors Pallarès. 1995. Velarization degree and coarticulatory resistance for /i/ in Catalan and German. *Journal of Phonetics*, 23(1–2):37 – 52.

Daniel Recasens. 1984. Voweltovowel coarticulation in catalan vcv sequences. *The Journal of the Acoustical Society of America*, 76(6):1624–1635.

P Sailaja. 1999. Syllable structure of Telugu. 4:743–746.

Ahmed Shamim. 2011. A reanalysis of Bengali vowel assimilation with special attention to metaphony. Master's thesis, Graduate Center, City University of New York.

Jan-Olof Svantesson. 1985. Vowel harmony shift in Mongolian. *Lingua*, 67(1985):283–327.

Jan-Olof Svantesson. 2005. *The Phonology of Mongolian*. Oxford University Press Inc, New York.

Robert W Wilkinson. 1974. Tense/lax vowel harmony in Telugu: the influence of derived contrast on rule application. *Linguistic Inquiry*, pages 251–270.

Yi Xu. 2007–2015. Formantpro.praat.