THE ROLE OF LABIOLINGUAL GESTURAL COORDINATION IN SPATIOTEMPORAL FACILITATION OF SPEECH PRODUCTION IN TURKISH, TURKMEN AND HINDI

Ganesh Gupta[1], Indranil Dutta [2]

The English and Foreign Languages University, Hyderabad, India.
gane.eflu@gmail.com[1], indranil@efluniversity.ac.in[2]

ABSTRACT

Sound patterns of vowel harmony reduce articulatory effort which facilitates speech production with increased speech rate and accuracy. We argue that back harmony patterns reduce articulatory effort compared to height harmony and non-harmony patterns and harmony patterns formed from front unrounded (FU) vowels reduce articulatory effort compared to those of front rounded (FR)vowels. We report on a series of priming experiments on Turkish, Turkmen and Hindi. The stimuli consist of items that are primed for back harmony (BH), disharmony (DH) and height harmony (HH) sequences under two conditions; front rounded vowel (FR) and front unrounded vowel (FU). The vowels /i/,/e/,/y/,/ø/,/u/,/o/ and consonants /p/,/t/,/k/ form segments used in the stimuli. Results of the LME test report that harmony patterns of FU condition increase speech rate and accuracy compared to the harmony patterns of FR condition; BH patterns increase speech rate compared to HH and DH patterns of FU condition.

Keywords: Vowel backness harmony, speech planning, labial and lingual gestures, spatiotemporal facilitation, gestural coordination.

1. ROLE OF GESTURAL COORDINATION AND SPEECH FACILITATION

1.1. Motor facilitation and gestural economy of vowel harmony

There exists a complex relationship between phonetics and phonology of vowel harmony. Phonological theories argue that harmony patterns are formed from spreading vowel features across discrete segmental units within the word boundary; [10],[8]. In phonetic accounts, V-to-V coarticulation is the precursor for the emergence of vowel harmony; [25],[20],[23]. Unlike abstract features, speech gestures provide fine-grained phonetic details of sound patterns;[11],[16],[22]. Harmonic patterns are formed from economized speech gestures to save articulatory effort required for speech production. Consequently, harmony patterns yield motor facilitation by way of increasing speaking rate and accuracy; [19],[20],[6],[12].

We argue that motor facilitation can be witnessed only in certain vowel harmony patterns. In theoretical accounts of phonology, labial, lingual and ATR harmony patterns occur independently. On the contrary, we propose that the occurrence of labial harmony and lingual harmony are inter-dependent. Languages which participate in lingual harmony (front/back harmony) also participate in labial harmony and vice verse. As the variations of the lip gesture causes changes in the tongue body position, labial harmony alters the tongue body position;[21],[14],[15].

Lingual harmony is formed from front/back tongue body gestures which is known as back harmony (hereafter BH) and high/low tongue body gestures which is known as height harmony (hereafter HH). The occurrence of lingual harmony is further conditioned within each harmonic language. Languages which participate in BH do not participate in HH. Consequently, languages in which BH patterns yield motor facilitation, HH patterns do not do so and vice verse;[20].

Labial harmony is formed of lip protrusion with larger horizontal/vertical opening of lips and lip compression with narrow vertical and horizontal opening of lips. Due to high amount of muscular effort needed for lip protrusion, discussed in [18],[21],[5], lip protrusion gesture needs longer articulatory duration, larger articulatory gesture and more articulatory effort compared to lip compression gesture. As labial harmony and lingual harmony are inter-dependent, BH patterned with lip compression harmony and lip spreading harmony exhibits a tendency of motor facilitation compared to BH patterned with lip protrusion harmony.

Languages with larger vowel inventories exhibit a tendency of forming front rounded vowels with lip protrusion and lip spreading gestures and back rounded vowels with lip compression gesture. Harmony patterns formed from front unrounded (hereafter, FU) vowels and back rounded (hereafter, BR) vowels yield motor facilitation compared to harmony patterns formed from front rounded (hereafter,FR) vowels.

1.2. Verifying motor planning and gestural economy of vowel harmony in Turkish, Turkmen and Hindi

As discussed in section 1.1, we hypothesize that back harmony languages, Turkish and Turkmen "...as discussed in [17],[26],[7]...", yield higher speaking rate and accuracy for vowel harmony patterns formed of FU vowels compared to those of FR vowels. Further, BH patterns yield higher speaking rate and accuracy compared to HH patterns and disharmony (hereafter DH) patterns within FU condition. As we predict that over all speech rates and accuracy of FU vowel patterns to be higher than those of FR vowel patterns, it is not necessary to verify the facilitatory variations among BH, DH and HH within FR condition. Hindi, a non-harmony language does not yield any facilitation. In this study, we present results from a series of speech production experiments conducted with Turkish, Turkmen and Hindi.

The stimuli are primed for vowel sequences of back harmony (BH), disharmony (DH) and height harmony (HH) under FU and FR vowel conditions. One consonant at a time (/t/, /p/ and /k/) is included in each phrase. The stimuli of BH, DH and HH sequences are formed from FR vowels as well as FU vowels paired with BR vowels. We intend to examine whether FU sequences yield spatiotemporal facilitation with higher speech rate and accuracy compared to FR condition. We also compare the difference of speech rate and accuracy among BH, DH and HH of FU condition. As we predict that stimuli of FR condition does not yield motor facilitation, further, we do not inquire the differences of speech rate and accuracy among harmony and consonantal types.

We examine the phenomenon of motor facilitation through measuring the speech rate and error rate of the stimuli of each harmony type in harmony languages Turkish and Turkmen (speech rate$\rightarrow$ BH$\geq$ DH >HH), (error rate$\rightarrow$ BH$\leq$ DH <HH) and Hindi (speech rate $\rightarrow$ FU>FR)(error rate$\rightarrow$ FU$\leq$ FR). Since the coronal is known to participate in palatalization in Turkish and Turkmen, consonantal influences are measured for speech rates(speech rates$\rightarrow$ t>p>k) but not error rates. Due to the long standing labial gesture in C-V Turkish patterns discussed in [4], we expect the consonantal gestures of /p/ to decrease speech rate. In the following section, we discuss the experimental procedure. Analysis of results of speech rates and error rates in Turkish, Turkmen and Hindi are discussed in the section 3. A detailed discussion on the experimental results is provided in section 4.

2. EXPERIMENT

The stimuli of the experiment are primed under the front rounded vowel condition (FR) with pairs of front rounded vowels /y/, /ø/ and front unrounded vowel condition (FU) with pairs of front unrounded vowels /i/, /e/ and with a pair of common back rounded vowels /u/ and /o/. These vowels are paired to form the back harmony (BH), disharmony (DH) and height harmony (HH) vowel pairs which are represented in Table 1. Disyllabic nonce words of the syllable structure ‘CVCV’ e.g.,‘pipe’,‘popu’ are used for each vowel pair for each harmony type and the consonants/p/, /t/, /k/ being fixed in the onset positions. Words of each harmony type are paired to form the carrier phrase of the kind ‘CVCV la CVCV’ e.g.:‘pipe la popu’ as shown in Table 1. The stimuli comprise of 108 FU and 108 FR phrases.

Five male and five female native Turkish and Turkmen speakers and four male and three female native Hindi speakers (between the age group of 21 to 30 years) participated in the speech production experiment at the EFL University, Hyderabad, India. Each text phrase of the stimuli is presented for a span of 5 seconds on a laptop screen using a MATLAB mathwork tool, [1] that provided a start alert beep before the phrase appeared and a stop alert beep when the phrase disappeared after 5 seconds. Speakers were instructed to utter each phrase as quickly as possible, as accurately as possible, and as many times as possible. The initial five tokens were meant for training and the randomized experimental tokens were presented afterwards. A keyboard interface was provided for transition between tokens with a text command ‘PRESS ENTER’. A Zoom H4n Linear PCM track recorder connected with AKG 420 head worn condenser microphone was used to record the speech on track one and the alert signals were recorded on track 2.

The experiment consists of two categorical variables; ‘harmony type’ (HT) and ‘consonantal type’ (CT) and two continuous variables ‘speech rate’ (SRt) and ‘error rate’ (ERt). ‘Number of syllables uttered’, ‘number of errors’ and ‘speaking duration’ are coded for each speaker and each token, and extracted with an automatic pause detection algorithm; [24]. Speech rates are measured as *number of syllables / speaking duration* and error rates as *number of errors / number of syllables*.

Table 1: Stimuli of FU & FR conditions

Type	FU condition	FR condition
BH	i-e,e-i,u-o,o-u (36)	y-ø,ø- y,u-o,o-u (36)
	pipe la pupo, (36)	pypøla popu,(36)
HH	i-u,u-i,e-o,o-e (36)	y-u,u- y,ø-o,o-ø(36)
	pope la pupi(36)	pypu la pøpo (36)
DD	i-o,o-i,u-e,e-u (36)	y-o,o-y,u-ø,ø-u (36)
	pipo la popi,(36)	pøpu la pypø(36)

In order to normalize the skewness in the distribution of the continuous variables, a log transformation is performed so as to satisfy the assumptions of the linear mixed effects model (LME). Speech rates and error rates are measured against the categorical variables 'harmony type' and 'consonantal type' within each harmony type category. The LME model provides a test of significance of difference between the means of the continuous variables as suggested by [2], [3], [13], [27]. In this study, we adopt the LME model in order to measure the influences of random variables and fixed variables, separately against speech rates and error rates. The LME model measures both the random effects and fixed effects separately by executing the formula provided in [9] using equation 1.

$$(1)\quad \begin{gathered} y_{ij} = \beta_1 x_{1ij} + \ldots + \beta_p x_{pij} + b_{i1} z_{1ij} + \ldots + b_{ip} z_{pij} + \varepsilon_{ij} \\ b_{ik} \sim N(0, \psi_k^2), Cov(b_k, b_{k'}) = \psi_{kk'} \\ \varepsilon_{ik} \sim N(0, \sigma^2 \lambda_{ijj}), Cov(e_{ij}, \varepsilon_{ij'}) = \sigma^2 \lambda_{ijj'} \end{gathered}$$

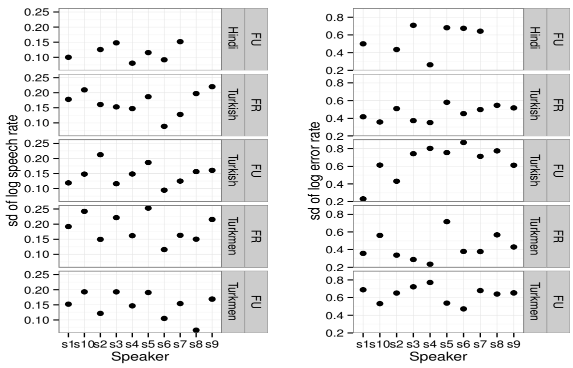
3. ANALYSIS

The LME model reports that speaker variability is less than 0.3 for speech rates and less than 1 for error rates in Turkish, Turkmen and Hindi. Altogether, speech rates and error rates of all the speakers are normally distributed in three languages, shown in Figure 1. In section 3.1, we have provided results of speech rates measured for the influences of FU and FR condition, HT and CT as well. Similarly, we have provided results of error rates measured for the influences of the categorical variables, FR and FU condition, HT and CT in section 3.2.

3.1. Speech rates

Mean speech rates for the two conditions rise in the order of FU (1.57)>FR (1.44) in Turkish, FU (1.55)>FR (1.47) in Turkmen as expected. The LME test reports that there is a significant difference between mean speech rates (p-value<0.05) in which case the FU condition yields higher speech rates

Figure 1: Random influences of speakers



than the speech rates of the FR condition in Turkish and Turkmen. Further, the speech rates of the FU condition are measured for harmony type and consonantal type influences but the speech rates of FR condition are not measured.

The mean values of the speech rates are measured for the three harmony type variables of Turkish, Turkmen and Hindi under the FU condition reported in the Table 2. The mean speech rates of harmony type variables rise in the order of BH>DH >HH in Turkish and Turkmen as expected, but there is no difference between the mean speech rates of harmony types in Hindi.

Table 2: Measurements of FU condition

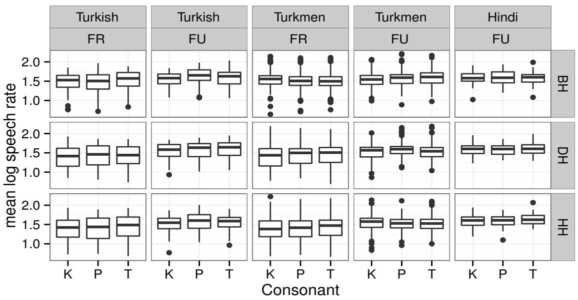
	Turkish		Turkmen		Hindi	
	SRt%	ERt%	SRt%	ERt%	SRt%	ERt%
BH	1.59	2.93	1.57	2.69	1.6	2.4
DH	1.57	2.66	1.55	2.75	1.6%	2.20
HH	1.54	2.74	1.53	2.8	1.6%	2.23

The LME test conducted for Turkish speech rates of FU condition reports that there is a significant difference between the speech rates of BH and HH (p-value<0.05) but the speech rates of DH do not differ from the rest. Speech rates of BH are higher than the speech rates of HH. An LME test conducted for consonantal influences on BH reports that there is a significant difference between the means of consonantal type under BH condition which rise in the order of /t/>/p/>/k/. In BH, the coronal consonant /t/ increases speech rates(p-value<0.05). The LME test conducted for the mean speech rates of consonants under the HH of Turkish reports that there is a significant difference between the mean speech rates of /k/ and /t/ in which the mean speech rate of /t/ is higher than the mean speech rate of /k/(p-value<0.05).

The LME test conducted for Turkmen speech rates under FU condition shows that the difference between mean speech rates of BH and HH is significant (p-value<0.05) and there is no difference between the means of BH and DH (p-value>0.05);

speech rates of BH are higher than the speech rates of HH. The LME test conducted for the consonantal influences on speech rates under BH reports that there is a significant difference between /k/ and /t/ (p-value<0.05) but results of consonant /p/ are not significantly different from /k/ (p-value>0.05). In BH of Turkmen, consonant /t/ increases speech rates. The LME test conducted for consonantal influences on the speech rates of HH report that there is no significant difference between mean speech rates(p-value>0.05). Altogether, BH yields higher speech rates than HH and consonant /t/ increases speech rates in BH only in Turkmen but not in Turkish. The LME test conducted for Hindi speech rates reports that there is no significant difference between mean speech rates of BH, DH and HH (p-value>0.05).

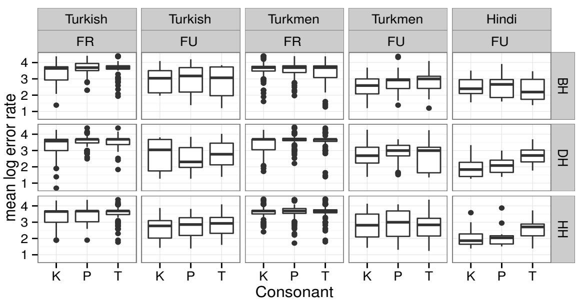
Figure 2: Speech rates-Turkish,Turkmen & Hindi



3.2. Error rates

Mean error rates of the two conditions rise in the order of FU (2.77) <FR (3.5) in Turkish and FU (2.6)<FR (3.6) in Turkmen similar to the expected order as discussed in section 1. An LME test is conducted to ascertain the difference of error rates between FU condition and FR condition. The LME test reports that there is a significant difference between the mean error rates of both the conditions (p-value<0.05) in which case the FU condition yields less error rates than the error rates of the FR condition in Turkish and Turkmen.

Figure 3: error rates-Turkish, Turkmen & Hindi



The mean values of the log transformed error rates are measured for the three harmony type variables of Turkish, Turkmen and Hindi under the FU condition. The mean error rates of the harmony type rise in the expected order of BH <DH <HH in Turkmen but the error rates of Turkish and Hindi rise in the order of BH>DH <HH which is not expected. The LME tests conducted for Turkish, Turkmen and Hindi error rates report that there is no significant difference between the mean error rates of BH, DH and HH (p-value>0.05) which also indicates that the consonantal influences on the error rates of BH, DH and HH are not significant in the three languages.

4. DISCUSSION

According to the results presented in section 3.1, in languages like Turkish and Turkmen, back harmony patterns yield motor facilitation compared to height harmony and non-harmony patterns by way of increasing speech rate but not certainly accuracy. In these languages, harmony patterns articulated with front unrounded vowels reduce the muscular effort compared to harmony patterns articulated with front rounded vowels by way of increasing speech rate and accuracy. Coronal consonants tend to increase speech rate and labial consonants tend to reduce the speech rate in both Turkish and Turkmen. There is no motor facilitation noticed in Hindi as it does not participate in vowel harmony. We argue that vowel harmony patterns yield temporal facilitation but it is conditioned by the coordination between labial and lingual gestures during the formation of V-to-V, C-V and V-C gestures. We argue that it is difficult to represent the factors of reduced articulatory effort which alters speaking rate and the quality of speech sounds to some extent in an acoustic speech signal. Further we intend to develop mathematical models accountable for modelling acoustic speech signals and bio-signals as well with a view to facilitating systems of speech recognition, speech to speech conversion and machine learning.

5. REFERENCES

[1] Matlab:math works. https://in.mathworks.com.

[2] Agresti, A. 2003. *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley.

[3] Baayen, R. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press.

[4] Boyce, S. 1990. Coarticulatory organization for lip rounding in Turkish and English. *The J. of the Acoust. Soc. of Am.* 88(6), 2584–2595.

[5] Catford, J. 2001. *A Practical Introduction to Phonetics*. Oxford linguistics series. Oxford University

Press.
[6] Cole, J. 2009. Emergent feature structures: Harmony systems in exemplar models of phonology. *Language Sciences* 31(2), 144–160.
[7] Crothers, S. M., J. 1980. Issues in the description of turkish vowel harmony. *Issues in Vowel Harmony: Proceedings of the CUNY Linguistics Conference on Vowel Harmony, May 14, 1977* volume 6. John Benjamins publishing 63–68.
[8] Finley, S. 2008. *Formal and cognitive restrictions on vowel harmony*. PhD thesis Johns Hopkins University Baltimore, MD.
[9] Fox, J. 2002. Linear mixed models. *Appendix to An R and S-PLUS Companion to Applied Regression*.
[10] Goldsmith, J. 1979. *Autosegmental phonology*. Outstanding dissertations in linguistics. Garland Pub.
[11] Goldstein, L., Fowler, C. A. 2003. Articulatory phonology: A phonology for public language use. *Phonetics and phonology in language comprehension and production: Differences and similarities* 159–207.
[12] Gupta, G., Indranil, D. 2013. Labial coarticulatory resistance to vowel harmony in telugu. *In proceedings of the Joint meeting of the French Acoustical Society and the Acoustical Society of India, New Delhi*.
[13] Johnson, K. 2011. *Quantitative methods in linguistics*. John Wiley & Sons.
[14] Kaun, A. 1995. *The Typology of Rounding Harmony: An Optimality Theoretic Approach*. UCLA dissertations in linguistics. UCLA.
[15] Kaun, A. 2004. The phonetic foundations of the rounding harmony typology. *Phonetically based phonology* 87–116.
[16] Kent, R. 1983. The segmental organization of speech. In: *The production of speech*. 57.
[17] Kirchner, R. 1993. Turkish vowel harmony and disharmony: An optimality theoretic account. *Rutgers Optimality Workshop I* volume 22.
[18] Lindau, M. 1978. Vowel features. *Journal of the Linguistic Society of America Baltimore, Md* 54(3), 541–563.
[19] Lindblom, B. 1983. Economy of speech gestures. In: MacNeilage, P., (ed), *The Production of Speech*. Springer New York 217–245.
[20] Linebaugh, G. 2007. *Phonetic Grounding and Phonology: Vowel Backness Harmony and Vowel Height Harmony*. PhD thesis University of Illinois at Urbana-Champaign.
[21] Linker, W. 1982. *Articulatory and acoustic correlates of labial activity in vowels: A cross-linguistic study*. Phonetics Laboratory, Department of Linguistics, UCLA.
[22] Mowrey, R., MacKay, I. 1990. Phonological primitives: Electromyographic speech error evidence. *The Journal of the Acoustical Society of America* 88(3), 1299.
[23] Öhman, S. 1966. Coarticulation in vcv utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America* 39(1), 151.
[24] Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/.
[25] Przezdziecki, M. 2000. Vowel harmony and vowel-to-vowel coarticulation in three dialects of yoruba. *Working Papers of the Cornell Phonetics Laboratory* 13, 105–124.
[26] Vago, R. 1980. *Issues in Vowel Harmony: Proceedings of the CUNY Linguistics Conference on Vowel Harmony, May 14, 1977* volume 6. John Benjamins Publishing.
[27] Winter, B. 2013. Linear models and linear mixed effects models in r with linguistic applications. *arXiv preprint arXiv:1308.5499*.