

PHONETIC ACCOMMODATION IN THE FUNDAMENTAL FREQUENCY OF KOREAN-ENGLISH BILINGUALS AND ENGLISH MONOLINGUALS

Nina Hosseini-Kivanani¹, Stephen J. Tobin², Adamantios I. Gafos¹

¹University of Potsdam, ²University of Michigan
hosseinikiva@uni-potsdam.de, gafos@uni-potsdam.de, sjtobin@umich.edu

ABSTRACT

Phonetic accommodation in fundamental frequency (f0) is of particular interest because it is not contrastive in English yet it covaries with contrastive parameters such as voice onset time (VOT). Accommodation in f0 additionally varies by speaker gender. In this investigation, we tested for f0 accommodation in 10 Korean-English bilinguals and 10 English monolinguals. Acoustic data were drawn from recordings of [k]-initial word productions from an experiment on VOT. Baseline f0 was compared to f0 from a test block, completed after participants shadowed monolingual native English. Our analysis shows convergence towards the model speaker across L1 and gender. All but one group showed convergence towards the model: the female monolingual English group. Their f0 was most similar to the model f0 at baseline, a similarity to which their divergence may be attributed. The results confirm that accommodation in f0 is modulated by social and linguistic factors.

Keywords: phonetic accommodation, bilinguals, Korean, English, f0.

1. INTRODUCTION

Hearing other people's speech exposes us not only to their phonological categories but also to the dialectal, sociolectal and idiosyncractic details of their implementation at the phonetic level. These phonetic details vary both within and between talkers, and it is sensitivity to this variation between talkers that this leads to accommodative change in any given talker's speech. However, many questions remain about accommodation—the necessary and sufficient conditions for accommodation to take place are not fully known.

The question we focus on here is whether accommodation in voice onset time (VOT) observed among a group of Korean-English bilinguals is also observed in the non-contrastive dimension of fundamental frequency (f0). A large number of studies shows that phonetic, lexical, and syntactic structures

that first- and second-language speakers produce are influenced by the speech they hear [9, 14]. An exact match between a pair of similar phonetic categories of an L1 and L2 is surely rare [5]. Some of the L1 phonetic categories of bilinguals, however, are likely to overlap with those in their L2. Researchers have investigated adaptation in various phonetic dimensions such as vowel quality (e.g., [2]) and VOT (e.g., [14]), in conversational settings and in more controlled, laboratory settings in a multitude of studies. Accommodation of f0 is of special interest because it has been reported to be a highly imitable feature and even the primary target of accommodation ([6, 7, 8]).

2. REVIEW OF PHONETIC ACCOMMODATION

Phonetic accommodation is the the process whereby speakers' speech production patterns change as a function of the interlocutor or speech community with whom they are interacting. This phenomenon has been widely studied in various contexts that small but significant effects of accommodation are found across a range of different tasks (e.g., ambient exposure, listening tasks, word-shadowing tasks). In the present investigation results of a shadowing task are reported. In shadowing, listeners repeat words after hearing them, and spontaneous imitation or convergence can be observed in the controlled setting of a phonetics laboratory. Speakers productions have been shown to adapt towards those of the model speech which they have heard [2].

Previous studies have found that phonetic accommodation, which Pardo [15] referred to as a factor intervening between perception and production, is influenced by a range of linguistic and social factors (e.g., [6, 15]). Among these modulating factors, Kim *et al.* [10] found evidence supporting their claim that the linguistic distance between talkers would affect accommodation (i.e., whether talkers shared the same dialect, the same L1 (but different dialects), or did not share the same L1). Kim *et al.* [10] found that talkers shifted more toward their

interlocutor if the model is linguistically closer (i.e., shared the same dialect of the same L1). According to this investigation, sociophonetic similarity between interlocutors facilitates shifting and spontaneous imitation. The pitch of a speaker's voice is a characteristic that pervades their speech, and all voiced segments depend upon vibration of the vocal folds to be distinctly produced and perceived. Moreover, previous studies have argued that f_0 is a critical component of phonetic accommodation (e.g., [7, 8]). F_0 has been reported to be a highly imitable feature and even to be the primary acoustic target in accommodation. In investigations of Gregory and colleagues, both f_0 and amplitude contours of sections of speech were examined in order to evaluate aspects of phonetic accommodation ([7]).

Both f_0 and VOT are important in distinguishing the three-way laryngeal contrast among the stops of Korean (lenis, fortis, and aspirated). With regards to examining adaptation using perceptual judgments in Korean and English speakers, Kim *et al.* [10] found that at the phonetic level, speakers converged toward a partner who had the same language or dialect. That is to say, social factors affect the likelihood and extent of convergence. Additionally, the female participants converged more than their male counterparts when the model talker was female.

Regarding the VOT values of Korean stops, Silva [16] has reported that a VOT merger between the lenis and aspirated stops is underway, and that the distinction between these categories is maintained primarily by an f_0 onset difference between them, that is, it is becoming a tonal contrast. In a related study, Chang [4] examined the phonological system of L2 participants (native English speakers learning Korean). Chang found an influence of L2 Korean on VOT in participants' L1 English. Another related investigation is that of Tobin [18], who compared accommodation in two groups of bilinguals (Spanish-English bilinguals and Korean-English bilinguals). The VOT values of the Korean-English group reduced towards typical values for monolingual English speakers, while the Spanish-English group's VOTs showed no significant change. The difference between these groups is attributed to the stability of oral-laryngeal coordination among the short positive VOT stops of Spanish in comparison with the less stable oral-laryngeal coordination among the long aspirated stops of Korean.

In this study, we test for f_0 accommodation in words produced by Korean-English bilinguals [19]. The additional phonetic variable of f_0 onset was measured in this existing data set. F_0 measurements were made at the onset of each target word's vowel.

Following Kim *et al.* [10], we expect that the likelihood of accommodation will be greater for the English group (henceforth En) because of greater social and linguistic similarity between model speaker and participants. Talkers who share the same L1 with their model talker should process the speech more automatically and are more likely to display phonetic convergence. Thus, if a talker has different L1 from the model speaker (i.e., Korean-English bilinguals (henceforth KoEn)), convergence is less likely in comparison with the En group. Pardo has reported greater convergence among male interlocutors than among female interlocutors [15] while Namy [13] reports that female participants are more likely to converge. Given these varied results on gender we take this opportunity to assess the effect of gender on accommodation in another experimental context.

3. METHODS

3.1. Participants & stimuli

Ten Korean-English bilinguals (7 female) and ten monolingual English controls (6 female) were drawn from an existing data set. Participants in the En group were native monolingual speakers of American English, and those in the KoEn group were native speakers of Korean with English L2. None of the participants had any known speech, language, or hearing impairment. The model speaker, who produced the 40 monosyllabic stimuli for the shadowing task, was a female native speaker of American English.

Data consisted of 40 monosyllabic English words that begin with the voiceless velar stop ([k]). They were recorded by a female native speaker of American English. The participants in this investigation heard audio recordings of these words, repeated them and read them aloud from a screen.

3.2. Procedure

Baseline task: Participants were instructed to read the same set of words aloud. The recording took place in a sound-proofed booth and participant read 10 repetitions of the /k/-initial monosyllabic words of English presented on a computer screen ($n = 400$).

Test task: This task consisted of word shadowing task to induce accommodation and a word reading task, measurements of which could be transparently compared to those from the baseline reading task. Participants repeated words presented to them over headphones as quickly as possible. First, participants completed a short practice block then shadowed two full randomizations of the set of 40 En

glish words. Each block contained two repetitions of each of the stimulus words. In total, five test reading blocks were completed ($n = 400$).

3.3. Measurements

Acoustic measurements of the f_0 onset of the vowel for each of the ten repetitions of forty words per participant were made using semi-automated VOT measurement software [11]. 15841 of the 16225 tokens were measured. Due to mispronunciation and hesitation, 2.4% of tokens were unusable.

3.4. Statistical Analysis

The goal of the analysis is to test for the effects of L1 and gender on accommodation. A linear mixed effects model was run in R [17] to assess changes in f_0 (dependent variable). The lme4 was used to fit the models [3] and the lmerTest [12] package was used to estimate p -values in the R statistical environment. The fixed effects of language, task and gender were included, with random intercepts for participants and random by-participant slopes for task.

4. RESULTS

Our hypothesis was that speakers who share the same language/dialect background as the model speaker would display a larger degree of phonetic accommodation [9]. Thus, if a speaker has different L1 (here Korean) from the model speaker, less convergence in comparison with the En group would be expected. The model's predictions for male participants' f_0 values fell in the range of 100-125 Hz, while those for female participants were between 200 and 260 Hz. These values are typical for male and female speakers and confirm the validity of our measurements. The model revealed significant effects of language, gender and task.

Table 1: Summary of results of linear mixed effect model for f_0 (language, task, gender)

Fixed Factor	Estimate	SE	t-value
(Intercept)	253.77	3.47	73.18 ***
Language(ref = KoEn)	-45.26	5.1	-8.87 ***
Gender(ref= Female)	-144.02	6.33	-22.75 ***
Task (ref = Baseline Task)	-6.81	1.87	-3.65 **
Language:Gender	47.51	8.67	5.48 ***
Language:Task	23.35	2.75	8.5 ***
Gender:Task	17.74	3.41	5.2 ***
Language:Gender:Task	-26.32	4.67	-5.64 ***

***, $p < .001$, **, $p < .01$

Focusing on the comparisons that are relevant to our hypothesis, changes in f_0 from baseline to test reading task were significant for all groups.

Both the female (see Table 1) and male subgroups of the KoEn group converged towards the model speaker's mean f_0 (195 Hz [SD = 8.71]). In the En group, the male subgroup also converged towards the model speaker. However the female English group diverged from the model speaker. Group means along with their standard errors are plotted in Fig. 1. Among two of our subsets of participants then, we observe clear convergence towards the model participant's f_0 : the KoEn group and the English male participants.

Figure 1: Mean f_0 of baseline and test tasks, separated by language and gender groups. The red solid line represents the female model talker.

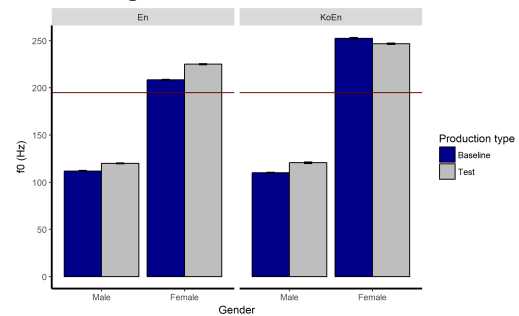
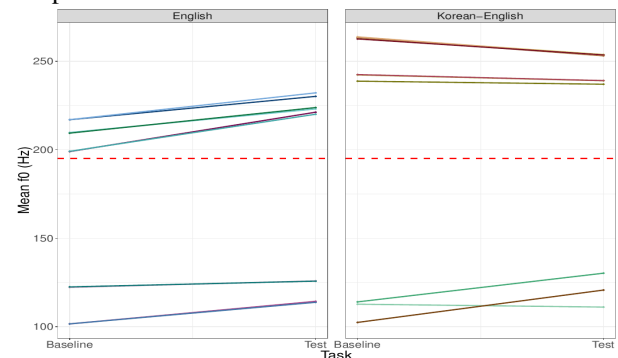


Fig. 2 shows individual participants' changes in f_0 , the red line indicating the f_0 of the model talker. The changes in f_0 are remarkably consistent within groups, both among those groups that converge and those that diverge. The hypothesis that participants who share an L1 with the model speaker would converge more is not supported. Although convergence is observed among both the L1- and the L2-English groups, we found more consistent convergence towards the model speaker among the L2 speakers than among the L1 speakers. With respect to gender, the results indicate a greater likelihood for male speakers to exhibit convergence than female speakers. Although there is a small difference in the size of male vs female sample (7 vs 13), this difference is not major and there is considerable consistency of patterns within groups.

Figure 2: Individual participant accommodation patterns



5. DISCUSSION

We tested for convergence towards the f_0 of a model speaker of American English, among monolingual native speakers of English Korean-English bilinguals. We tested the hypothesis that speakers who shared the native language of the model speaker would be more likely to converge than those who did not. We also assessed the effect of participant gender on the likelihood of convergence. The results did not support this hypothesis about the effect of language—the Korean-English group was, in fact, more consistent in showing convergence towards the model speaker than the monolingual English group. That is not to say that the English group demonstrated no convergence. Both language groups showed significant convergence towards the model speaker's f_0 . However, among the English monolinguals, only the male speakers converged, while the female speakers consistently diverged from the model.

In the literature on phonetic accommodation the effect of gender has not been entirely consistent. Some investigations, such as that of Pardo, report results showing the male speakers converge more than female speakers [15], whereas others, such as that of Namy, report that female speakers converge more [13]. Among investigations of f_0 in which consideration is given to participant gender, findings of Babel and Bulatov [1] are consistent with those presented here. In Babel and Bulatov's investigation, male and female native speakers of American English shadowed words produced by another male native speaker. In their unfiltered audio condition, male participants showed greater convergence towards the model speaker than female participants. While our female native-English participants showed divergence, the relative change in f_0 between genders is consistent, with male speakers showing clear convergence and female speakers showing either weaker convergence or divergence. Our findings for the English group are also consistent with Namy *et al.* [13] who found that listeners more consistently judged productions of male speakers to have converged towards an interlocutor than those of female speakers. We should note that Babel and Bulatov [1] measured the average f_0 over the course of word productions, whereas we measured f_0 onset, and their model speaker was male, whereas ours was female. Likewise, Namy's measures are listener's holistic judgments rather than acoustic measurements. Perhaps precisely because of these differences, the pattern across investigations of male participants converging reliably and

female participants showing variability in accommodation is striking. Among investigations of accommodation in which cross-language effects on f_0 are considered, Chang [4] found significant changes in the English f_0 and VOT values of native speakers of American English while they learned Korean from scratch in an immersion course. With regard to f_0 , Chang found significant changes among the female participants but no such changes among the male participants. This is the opposite pattern to the one reported here. Although language contact is a shared aspect between Chang and the present investigation, we should note that Chang focused on L2-induced changes in participants' L1, whereas our focus among the Korean-English bilingual group is on changes in the L2 that may be modulated by the L1. Thus, his reports of phonetic drift (accommodative effects of one of a speaker's language on another of their languages) need not necessarily match our effects of phonetic accommodation (within a single language of a speaker).

One related line of research that may offer an explanation for the divergence of the English female group is that of Tobin *et al.* [20]. In a somewhat different experimental paradigm, they found that the distance between VOTs of a speech stimulus and those of the participant affects the extent to which participants converge. The key here is that of phonetic distance. Considering Fig. 2, it is clear that, with one exception, those participants whose baseline f_0 is 50Hz or more from that of the model speaker (all but one of the male participants and all of the KoEn male participants) converge towards the model speaker's f_0 . Conversely, those participants whose baseline f_0 is very similar to (within 25 Hz of) that of the model diverge.

Summing up, speech accommodation is a phenomenon that is influenced by a wide variety of factors, including individual speaker differences, as well as social and linguistic factors. We are increasingly able to bring these under experimental control. Language differences did not hinder phonetic accommodation for non-native speakers in the present investigation (cf. [10]). In fact, gender played a greater role in modulating convergence in this investigation. Further, although there may be automatic elements to accommodation [14], given that speakers retain individual speech characteristics, it is not purely an automatic process of the language system. Various factors, including the distance between interlocutors' phonetic parameter values [20], seem to play a role in modulating presence and degree of accommodation.

6. REFERENCES

- [1] Babel, & B. D., M. 2012. The role of fundamental frequency in phonetic accommodation. *Language and Speech* 55(2), 231–248.
- [2] Babel, M., Bulatov, D. 2012. The Role of Fundamental Frequency in Phonetic Accommodation. *Language and Speech* 55(2), 231–248.
- [3] Bates, D. M., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67, 1–48.
- [4] Chang, C. B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics* 40(2), 249–268.
- [5] Flege, J. E., Eefting, W. 1988. Imitation of a VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *The Journal of the Acoustical Society of America* 83(2), 729–740.
- [6] Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2), 251–279.
- [7] Gregory, S. W., Dagan, K. a., Webster, S. 1997. Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perception of communication quality. *Journal of Nonverbal Behavior* 21(1), 23–43.
- [8] Gregory, S. W., Green, B. E., Carrothers, R. M., Dagan, K. A., Webster, S. W. 2001. Verifying the primacy of voice fundamental frequency in social status accommodation. *Language and Communication* 21(1), 37–60.
- [9] Kim, M. 2011. Phonetic convergence after perceptual exposure to native and nonnative speech: Preliminary findings based on fine-grained acoustic-phonetic measurement. *17th International Congress of Phonetic Sciences (ICPhS XVII)* (August), 1074–1077.
- [10] Kim, M., Horton, W. S., Bradlow, A. R. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology* 2(1), 125–156.
- [11] Kuberski, S. R., Tobin, S. J., Gafos, A. I. 2016. A landmark-based approach to automatic voice onset time estimation in stop-vowel sequences. *Signal and Information Processing (GlobalSIP), 2016 IEEE Global Conference on*. IEEE 60–64.
- [12] Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. 2017. lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software* 82(13).
- [13] Namy, L. L., Nygaard, L. C., Sauerteig, D. 2002. Gender Differences in Vocal Accommodation. *Journal of Language and Social Psychology* 21(4), 422–432.
- [14] Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2), 132–142.
- [15] Pardo, J. S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4), 2382–2393.
- [16] Silva, D. J. 2006. Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology* 23(02), 287–308.
- [17] Team, R. C., others, 2014. R: A language and environment for statistical computing.
- [18] Tobin, S. 2013. Phonetic accommodation in Spanish-English and Korean-English bilinguals. *Proceedings of Meetings on Acoustics* 19, 60087.
- [19] Tobin, S. 2015. A dynamic approach to phonetic change. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.
- [20] Tobin, S., Hullebus, M., Gafos, A. "In press". Immediate phonetic convergence in a cue-distractor paradigm. *The Journal of the Acoustical Society of America, Express Letters* 14X(Y).