

Phonetic imitation of VOT in L2 English: variation as a function of model talker

Magdalena Zajac
University of Lodz
zajac1234@gmail.com

ABSTRACT

The study focused on phonetic imitation in the speech of Polish learners of English, who were exposed to two pronunciation varieties: Polish-accented English and native English. The phonetic variables under investigation were voicing lag values in word-initial /p k/ and voicing lead values in word-initial /b g/. The stimuli were monosyllabic English words provided by two model talkers: a native speaker of English and a native speaker of Polish. It was found that the magnitude of phonetic imitation differed as a function of model talker. The results suggest that L2 phonetic imitation may be affected by attitudinal factors and the stage of acquisition of a given pronunciation feature.

Keywords: L2 phonetics, phonetic imitation, L2 English pronunciation, VOT

1. INTRODUCTION

Numerous previous studies indicate that speakers tend to adapt their linguistic behaviour according to who they are talking or listening to. The phenomenon has been investigated under various names and with the use of different methodological procedures. One major strand of research refers to the process as accommodation (or convergence) and studies it in spontaneous conversational interactions (e.g. [5], [7], [9], [15], [19], [21]). Another important body of work refers to the phenomenon as imitation and examines it in non-interactive settings in which the speakers are required to repeat pre-recorded single words (e.g. [1], [2], [8], [10], [17], [20], [26], [27]). The current study investigates pronunciation shifts with the use of a socially minimal, non-interactive experimental design; for this reason, the term phonetic imitation is applied. The term will be used to refer to a process in which a speaker becomes more similar-sounding to another individual when repeating speech stimuli provided by this person.

Previous research indicates that phonetic imitation is often unintentional and that the process may stem from an automatic reflex of the human brain (e.g. [6], [8], [132], [26]). Other studies imply that despite the socially minimal experimental setting, phonetic imitation may be conditioned by certain socio-psychological factors such as

personality traits, social group membership and speakers' attitude towards a given talker or social group (e.g. [1], [2], [18], [27]).

Phonetic imitation has been found to operate in L2 speech. More specifically, L2 learners of English have been found to imitate the following phonetic features: VOT duration in voiceless stops ([23]), spectral characteristics of the TRAP vowel ([24]), duration of release burst in stop sequences ([25]), vowel duration as a cue for consonant voicing ([28], [29]). Interestingly, the results of one of the studies showed that the magnitude of imitation in the learners' pronunciation differed as a function of the model talker (native vs. non-native) ([29]). It was found that the subjects conformed to native-speaker norms even when presented with non-native-like vowel renditions. The authors argued that the observed imitation pattern may have been affected by the participants' attitudes towards English pronunciation. However, the claim could not be verified since no attitudinal measures were applied in the study.

2. THE CURRENT STUDY

The general aim of the study was to examine L2 phonetic imitation upon exposure to native and non-native speech. More specifically, the goal was to investigate whether the magnitude of imitation varies as a function of the native/non-native status of the model talker and to determine whether imitation patterns are conditioned by learners' attitude towards L2 pronunciation. The study concentrated on the speech behaviour of Polish learners of English, who were exposed to two pronunciation varieties: Polish-accented English and native English.

2.1. Phonetic variables

The phonetic parameters selected for analysis include aspiration and pre-voicing in word-initial stops. Aspiration was operationalised as voicing lag values in initial /p t k/; pre-voicing was operationalised as voicing lead values in initial /b d g/. Word-initial stops have different phonetic realisations in the subjects' L1 and L2, i.e. /p t k/ are generally produced with long voicing lag in English and short voicing lag in Polish; /b d g/ are generally

realised with short voicing lag in English and voicing lead in Polish (e.g. [11], [12],[14], [16]).

2.2. Stimulus

The stimuli were 24 monosyllabic English with different word-initial stop consonants, arranged into voiced-voiceless minimal pairs followed by the same vowel: *bat-pat*, *bet-pet*, *bun-pun*, *bop-pop*, *Dutch-touch*, *dog-tog*, *dip-tip*, *Dan-tan*, *gap-cap*, *goat-coat*, *gut-cut*, *got-cot*. The words were recorded by a native speaker of Standard Southern British English and a native speaker of Polish. The model talkers were both male and of similar age (the English model talker was in his late twenties, the Polish model talker in his early thirties). The native Polish speaker was a qualified phonetician who imitated a heavy Polish accent for the purposes of the study. The model talkers used natural speaking tempo and falling intonation in each token.

As can be seen in Table 1, word-initial /p t k/ were realised as unaspirated by the Polish model talker and as aspirated by the English model talker.

Table 1: Mean voicing lag values in /p t k/ for the two model talkers; NNS – Polish model talker, NS – English model talker.

	/p/	/t/	/k/	overall
NNS	+13 ms	+14 ms	+27 ms	+18 ms
NS	+112 ms	+158 ms	+133 ms	+134 ms

As illustrated in Table 2, the Polish model talker produced word-initial /b d g/ with considerable amounts of pre-voicing; the English model talker consistently devoiced the three stops.

Table 2: Mean voicing lead values in /p t k/ for the two model talkers; NNS – Polish model talker, NS – English model talker

	/b/	/d/	/g/	overall
NNS	-161 ms	-169 ms	-149 ms	-160 ms
NS	0 ms	0 ms	0 ms	0 ms

2.3. Participants

Although a total of 38 participants were recorded for the purposes of the study, the number was reduced to 25 due to frequent misidentifications of the target words (see Section 2.5.). There were 16 females and 9 males; the age of the subjects ranged from 20 to 22 (M=20.5). The participants were all second-year students of English Studies, recruited from the University of Lodz. Their level of English proficiency ranged from upper-intermediate to

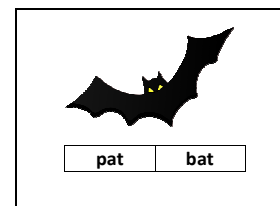
advanced and they all had long experience with learning English (M=14 years, SD=2.2, Min.=9 years, Max=15 years). At the time of the experiment, the subjects had completed three semesters of an English phonetics and phonology course; the course covered the realisation of English word-initial stops. None of the participants reported any speech or hearing disorders.

2.4. Experimental procedure

The experimental procedure used in the current study is based on pilot work on phonetic imitation in L2 speech. The study is part of a larger investigation, which included additional speech stimuli and experimental tasks that are not discussed in this paper.

In the first phase of the experiment (baseline task), the informants were presented with a self-running Power Point presentation; each slide included two English words (minimal pairs) and a picture. An example is provided in Fig.1.

Figure 1: Sample PowerPoint slide from the baseline task



The subjects' task was to decide which of the two words in presented in a given picture by reading it out loud.

The experiment included two imitation tasks in which the participants were given an exercise sheet with minimal pairs that contained the analysed tokens. In the 1st imitation task, the subjects listened to the stimuli provided by the Polish model talker; in the 2nd imitation task, they listened to the stimuli provided by the English model talker. The informants were instructed to identify the words produced by the two speakers by reading them out loud and marking them on the provided exercise sheet.

Following the production phase of the experimental procedure, the participants were required to complete a questionnaire designed to gauge their attitudes towards native and Polish-accented English. The subjects were asked to evaluate the pronunciation of the model talkers (with respect to variables such as perceived level of intelligence, education, attractiveness, etc.; Lickert-type scale was used). The questionnaire included a

self-report component concerned with the subjects' phonetic performance in the baseline and imitation tasks. The participants chose between three options: a) *I concentrated on my pronunciation and tried to sound native-like*, b) *I paid no attention to my pronunciation* or c) *other*. Another component of the questionnaire examined subjects' attitudes towards Polish-accented English, their opinion on the importance of using-native like pronunciation when speaking English and the importance of pronunciation as compared with other language skills (Licker-type scale). Additionally, the questionnaire included open-ended questions in which the participants could further comment on phonetic performance of the model talkers and their own speech behaviour.

2.5. Data analysis

Voicing lag and voicing lead were measured using Praat speech-analysis software package ([3]) by means of waveform and spectrographic display. Voicing lag in /p t k/ tokens was measured as the temporal span between the first peak of release burst and the onset of the first complete vibration of the vocal folds (e.g. [4], [16], [22]). Voicing lead in /b d g/ tokens was identified as the time interval represented by the voice bar (e.g. [16]).

Although a total number of 24 target words served as stimuli in the current study, a considerable number had to be excluded from statistical analysis due to frequent misidentifications in the imitation tasks. Examination of the answer sheets used in the two tasks revealed that the target words produced by the model talkers were often confused with their voiced or voiceless counterparts. Ultimately, the minimal pairs that were least frequently misidentified were selected for statistical analysis and include the following word sets: *pat-bat*, *pop-bop*, *cap-gap*, *cut-gut*. It was also necessary to reduce the number of participants from 38 to 25; productions of those subjects who committed the greatest amount of identification errors were excluded from analysis.

The relationship between task type (imitation of native vs. non-native speech) and the two phonetic parameters (aspiration and pre-voicing) was tested by conducting two-way repeated measures ANOVAs (one for each phonetic parameter).

3. RESULTS AND ANALYSIS

3.1. Questionnaire results

Questionnaire results revealed that the native speaker's pronunciation was rated higher than the

non-native speaker's pronunciation (M=20, SD=4.4 for the Polish model talker; M=30, SD=4.1 for the English model talker). The mean score achieved by the subjects in the component that examined attitudes towards Polish-accented English was 49 (SD=7.2; the maximum score in this part of the questionnaire was 65; achieving a score close to this number was expected to signify a strong bias against Polish-accented speech). When answering the open-ended questions, many participants pointed to the importance of using "correct", native-like pronunciation. Several subjects suggested that the Polish model talker's pronunciation was incorrect and/or that he mispronounced some of the analysed words. The pronunciation of the English model talker, on the other hand, was often referred to as "correct" or "proper".

In the self-report component of the questionnaire, the majority of the participants stated that they concentrated on their pronunciation and attempted to realise the analysed target words in a native-like manner. This option was selected by: 84% of the participants with respect to the baseline task, 68% with respect to the 1st imitation task, 92% with respect to the 2nd imitation task. Several participants declared that they imitated the non-native speaker in the 1st imitation task; all stated that they did it inadvertently.

The findings can be interpreted to mean that the subjects viewed the native speaker's pronunciation in a more positive light than the non-native speaker's pronunciation and generally exhibited a preference for native over Polish-accented English. The results also indicate that overall, the participants attempted to sound native-like when producing the target words.

3.2. The results of statistical analysis

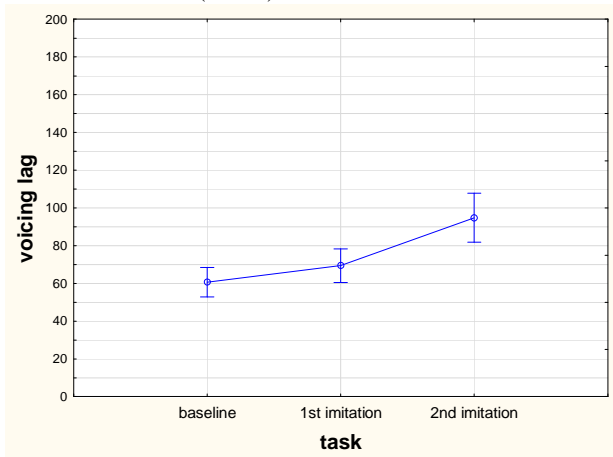
3.2.1. Aspiration

The main effect of task on voicing lag values was highly significant [$F(2, 48)=58.5, p<.001$] (see Fig. 3). Post hoc Bonferroni tests showed that the increase in duration between the baseline (M=61 ms, SD=3.8) and the 1st imitation task (M=69 ms, SD=4.3) was statistically significant [$p<.01$]. The increase in the 2nd imitation task (M=95 ms, SD=6.3) as compared with the baseline task was also found to be statistically significant [$p<.001$].

The findings imply that the subjects converged towards the native model talker and diverged from the non-native model talker by increasing the amount of aspiration in their productions.

4. DISCUSSION

Figure 3: Mean voicing lag values (ms) across different tasks (N=25).

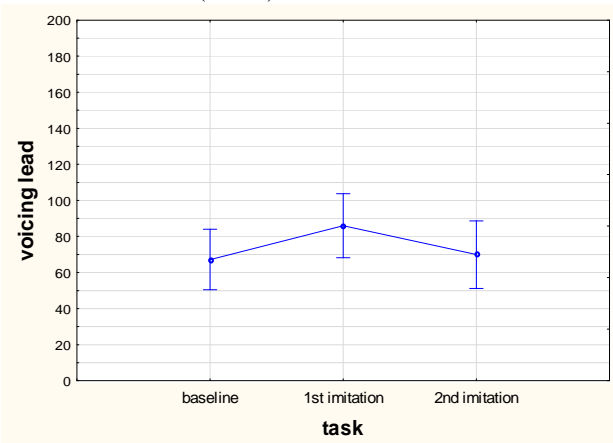


3.2.2. Pre-voicing

During statistical analysis, voicing lead values were converted into positive numbers in order to facilitate the interpretation of the results. The main effect of task on voicing lead was found to be significant [$F(2,48)=4.5$, $p<.05$] (see Fig. 4). Post hoc Bonferroni tests showed that the increase in duration between the baseline ($M=67$ ms $SD=8.1$) and the 1st imitation task ($M=86$ ms $SD=8.6$) was statistically significant [$p<.01$]. The difference in duration between the baseline and the 2nd imitation task ($M=70$ ms $SD=9.1$) was very small and was not found to be statistically significant.

The findings indicate that the participants converged towards the Polish speaker by increasing the amount of pre-voicing in their productions and maintained their default voicing lead values following exposure to the English speaker's pronunciation.

Figure 4: Mean voicing lead values (ms) across different tasks (N=25).



In the case of aspiration, the results indicate that the subjects converged towards the L2 norm following exposure to both native and non-native pronunciation. The pattern can be explained by a preference for native-like pronunciation and a desire to sound native-like on the part of the participants (as indicated by the questionnaire results). Additionally, the data show that the informants had already produced the analysed tokens as aspirated in the baseline task (i.e. in a native-like manner), which could imply that they had mastered this L2 pronunciation feature. Thus, it is possible that they were able to converge towards higher voicing lag values when listening to the native model talker and managed to overcome potential L1 interference when listening to the non-native model talker because they had succeeded in acquiring this feature of English pronunciation.

A pattern opposite to the one found for aspiration can be observed in the case of pre-voicing. The results suggest that the informants converged towards the non-native speaker but not towards the native speaker. The finding seems especially interesting in light of the fact that the informants exhibited a pro-native-pronunciation bias and mostly stated they wished to sound native-like. The subjects realised word-initial /b g/ with substantial amounts of pre-voicing in the baseline task (i.e. in accordance with L1 articulatory habits), which could suggest that they had not yet mastered a native-like realisation of the two English stops. Therefore, it seems possible that it was L1 transfer that prevented convergence on devoicing and facilitated the imitation of extended pre-voicing. These observations raise the possibility that the magnitude of imitation in L2 pronunciation may be conditioned by the stage of acquisition of a given pronunciation feature.

5. CONCLUSIONS

The results of the study suggest that the magnitude of phonetic imitation in the speech of Polish learners of English varied as a function of the native/non-native status of the model talker. The observed imitation pattern appears to be related to attitudinal factors; the findings imply that a preference for target-like pronunciation prompted the learners to converge towards native speech and diverge from foreign-accented speech. However, the effect was observed in the case of one of the investigated features only. The data raise the possibility that the influence of attitudinal factors on L2 phonetic

imitation may be conditioned the stage of acquisition of a given L2 pronunciation feature.

7. REFERENCES

- [1] Babel, M. 2009. Phonetic and social selectivity in speech accommodation. Unpublished PhD dissertation.
- [2] Babel, M. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39, 437–456.
- [3] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glott International* 10, 341-345.
- [4] Cole, J., Kim, H., Choi, H., Hasegawa-Johnson, M. 2007. Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics* 35, 180–209.
- [5] Coupland, N. 1984. Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language* 46, 49-70.
- [6] Delvaux, V., Soquet, A. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64, 145-73
- [7] Giles, H. 1973. Accent mobility: a model and some data. *Anthropological Linguistics* 15, 87-105.
- [8] Goldinger, S. 1998. Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review* 105, 251-279.
- [8] Gregory, S., Webster, S. 1996. A Nonverbal Signal in Voices of Interview Partners Effectively Predicts Communication Accommodation and Social Status Perceptions. *Journal of Personality and Social Psychology* 70, 1231-1240.
- [10] Honorof, D., Weihing, J., Fowler, C. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39, 18-38.
- [11] Keating, P. A. 1980. A phonetic study of voicing contrast in Polish. PhD dissertation, Brown University.
- [12] Keating, P. A., Mikoś, M. J., Ganong III, W. F. 1981. A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America* 70, 1261-1271.
- [13] Kim, M. 2011. Phonetic convergence after perceptual exposure to native and non-native speech: preliminary findings based on fine-grained acoustic-phonetic measurement. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1074-1077.
- [14] Kopczyński, A. 1977. *Polish and American English consonant phonemes: A contrastive study*. Warszawa: Państwowe Wydawnictwo Naukowe.
- [15] Lewandowski, N. 2012. Talent in nonnative phonetic convergence. Unpublished PhD dissertation. Universität Stuttgart.
- [16] Lisker, L. and Abramson, A. S. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 527-565
- [17] Mitterer, H., Ernestus, M. 2008. The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109, 168-173.
- [18] Namy, L., Nygaard, L., Sauerteig, D. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21, 422-432.
- [19] Natale, M. (1975) Convergence of Mean Vocal Intensity in Dyadic Communication as a Function of Social Desirability. *Journal of Personality and Social Psychology* 32, 790-804.
- [20] Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39, 132-142.
- [21] Pardo, J. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119, 2382–2393.
- [22] Rojczyk, A. 2010. *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- [23] Rojczyk, A. 2012. Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EuroSLA22 – 22nd Annual Conference of the European Second Language Association, Poznań, Poland, 5-8 September.
- [24] Rojczyk, A. 2013. Phonetic imitation of L2 vowels in a rapid shadowing task. In: J. Levis and K. LeVelle (eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*. Ames, IA: Iowa State University, 66-76.
- [25] Rojczyk, A., Porzuczek, A., Bergier, M. 2013. Immediate [23] and distracted imitation in second-language speech: Unreleased plosives in English. *Research in Language* 11, 3-18.
- [26] Shockley, K., Sabadini, L., Fowler, C. A. 2004. Imitation in shadowing words. *Perception & Psychophysics* 66, 422-429.
- [27] Yu, A. C., Abrego-Collier, C., Sonderegger, M. 2013. Phonetic imitation from an individual-difference perspective: Subjective attitude, personality and “autistic” traits. *PLoS ONE* 8, 1-13.
- [28] Zając, M. (2013) Phonetic imitation of vowel duration in L2 speech. *Research in Language* 11, 19-29.
- [29] Zając, M., Rojczyk, A. 2014. Imitation of English vowel duration upon exposure to native and non-native speech. *Poznan Studies in Contemporary Linguistics* 50, 495–514.