

Phonetic imitation in L2 speech: Immediate imitation of English consonant glottalization by speakers of Polish

Phonetic imitation is a ubiquitous process in speech production. Speakers have a strong tendency to imitate their interlocutors both in a native and a non-native language. It is especially important in acquiring non-native speech, because it allows forming new sound categories. In the current study we investigated whether and to what extent Polish learners of English are able to imitate t-glottalization observed especially in British English. A total of 25 Polish learners of English imitated English models' productions with t-glottalization that were subsequently compared to their default productions (pre-test) and post-exposure production (post-test). The results showed that the participants successfully imitated t-glottalization after the exposure to the model talker. The generalisation effect was limited in its magnitude in that only some of the non-imitated words had traces of glottalization. The results are discussed in terms of the differences in the implementation of glottalization in Polish and English and of how phonetic imitation informs second-language speech acquisition.

Keywords: phonetic imitation; glottalization; second-language speech; language acquisition

Subject classification codes: include these here if the journal requires them

1. Introduction

Phonetic imitation, also known as accommodation (Babel 2012), alignment (Trofimovich & Kennedy 2014), or convergence (Pardo et al. 2012) is a process whereby a speaker adjusts spectral and temporal properties of his/her speech towards those of an interlocutor or a model talker. Such 'tuning in' has been reported for a variety of acoustic features such as Voice Onset Time (Shockley et al. 2004; Nielsen 2011), vowel quality (Babel 2012; Pardo et al. 2010), allophonic realisations of /l/ (Honorof et al. 2011), as well as speaking rate, intensity, or long-term average spectra (Gregory & Webster 1996; Namy et al. 2002). Phonetic imitation appears to be yet

another subcomponent of a general capacity of humans to reproduce actions and intentions of others (Hauser 1996), considering the fact that alignment with a conversation partner also emerges for other aspects of language such as semantics (Katzir & Singh 2013), syntax (Gries 2005), vocabulary (Boghrati et al. 2016), and morphology (Dunstan 2010).

Phonetic imitation plays an important role in second-language speech research as it is considered to be a window into (1) which sounds and phonetic features of L2 are difficult for learners (imitation effectiveness may or may not be inversely correlated with difficulty); (2) to what extent L1 speech habits may be temporarily bypassed after exposure to a native model; (3) the degree of a carryover effect from test productions (direct exposure) to post-test productions. Most studies on phonetic imitation in L2 employed immediate imitation in laboratory conditions. It should be noted that what we call ‘imitation’ in this study is technically different from a similar concept known as ‘shadowing’. The review of the second-language speech literature reveals that an unambiguous distinction between imitation and shadowing is frequently missing and the two methodological terms tend to be used interchangeably. While imitation has a clear focus on form and accuracy with a marked volitional component (i.e., imitate the speaker as closely and as accurately as possible), shadowing relies on asking listeners to repeat what they hear as quickly as possible without explicitly or implicitly instructing them to imitate (Mitterer & Ernestus 2008; Mitterer & Müsseler 2013; Pardo et al. 2018). As a consequence, the target of shadowing is repetition of words or phrases together with its latency measured as the time between hearing a token and repeating it (Fowler et al. 2003; Marlsen-Wilson 1973, 1975; Peschke et al. 2009; Porter & Castellanos 1980; Porter & Lubker 1980). Previous research has shown that the observable degree of unintentional imitation emerges in shadowing, even if the

shadowers are not instructed to imitate the productions that they repeat (Babel 2010, 2012; Babel & Bulatov 2012; Fowler et al. 2003; Namy et al. 2002; Shockley et al. 2004), however the magnitude of convergence is less compared to explicit (instructed) imitation (Adank et al. 2010; Dufour & Nguyen 2013; Pardo et al. 2010). The studies in L2 speech imitation typically use precise instructions to imitate a model talker as faithfully as possible (Carignan 2018; Chen et al. 2023; Rojczyk 2012, 2013; Rojczyk et al. 2013; Rojczyk & Rallo Fabra 2023; Zając & Rojczyk 2014), and, as a result, fit into the methodological pattern of imitation rather than shadowing, because convergence is the target and not a mere by-product of repetition. Considering the methodological expectations in research on L2 speech (investigating to what extent L2 learners converge with a model native talker), it is therefore more terminologically precise to refer to L2 learners' productions after exposure to the model talker as immediate imitations rather than shadowing.

In a typical imitation paradigm, the participants commence with reading words presented in orthographic form (baseline condition), followed by immediate imitation during which they hear and repeat the same words after a model talker. The last stage consists of re-reading the target words (post-test) in order to assess the magnitude of carryover effect from exposure to a model talker. Previous research has demonstrated that immediate imitation after a model talker enables learners to produce phonetic features that are absent in their L1 and that have not yet been fully acquired in their L2 speech. This effect has been reported for phonetic-acoustic features such as VOT (Flege and Eeefing 1988), vowel duration (Zając & Rojczyk 2014; Podlipský & Šimáčková 2015), formant frequencies of vowels (Rojczyk 2013; Llompарт & Reinisch 2018; Jiang & Kennison 2022), the lack of release in stop consonants (Rojczyk et al. 2013), or suprasegmentals (Ulbrich 2021). In the current study, we contribute to L2 phonetic

imitation studies by investigating consonantal glottalization as another potential speech property that may emerge in immediate imitation of L2 speech.

2. Glottalization in English and Polish

Glottalisation is well attested in numerous varieties of English, where glottal gestures may occur as a syllable boundary marker resolving a vowel hiatus, as a marker of emphasis on vowel initial words or as low level variants of coda voiceless stops. Classic phonological descriptions of English divide the glottal gestures in syllable codas into two categories, termed ‘glottal replacement’ and ‘glottal reinforcement’ (glottalling and pre-glottalisation), where they either replace or accompany voiceless oral stops /p, t, k, tʃ/, the latter type of allophony being “increasingly typical of many types of British English” (Cruttenden 2014: 171).

Although glottalisation is typically symbolised in segmental terms i.e. as a glottal stop [ʔ] or a sequence of a glottal and an oral stop [$\widehat{ʔp}$, $\widehat{ʔt}$, $\widehat{ʔk}$, $\widehat{ʔtʃ}$], the realization of these events varies from glottal plosives to creaky or breathy voice (Redi & Shattuck-Hufnagel 2001; Ashby & Przedlacka 2014; Keating et al. 2015). In English speech canonical glottal plosives are hardly found at all, with glottalisation realised chiefly of a period of disturbed vocal fold vibration.¹ Glottal replacement is a frequent variant of coda /t/ in British English, though the contexts of its occurrence differ between varieties (for details see Cruttenden 2014), with non-standard accents being more permissive i.e. allowing both coda and intervocalic word medial glottalling (as in *water*, *butter*).

¹ See Docherty and Foulkes (1999) and Ashby and Przedlacka (2014) for illustrations of glottal events in non-SSBE varieties of English. For glottalization in American English, see for instance Seyfarth and Garellek (2020) or Kaźmierski (2020).

In Polish, on the other hand, glottalization appears to be a syllable-level process motivated by sustaining word and morpheme integrity (Schwartz 2012, 2017). This mechanism blocks any resyllabification which is relatively common in English. As a result phrases such as *pod oknem* ‘under the window’ and *do auta* ‘into the car’ will have pre-vocalic word-boundary glottalization marker [pot ʔoknɐm], [do ʔaũta] respectively. No consonantal glottal reinforcement or replacement are observed in Polish. Therefore, an interesting L2 learning scenario emerges in which Polish learners of English possess a phonetic mechanism of glottalization in their native language although its distribution is largely limited compared to the one attested in L2.

3. The current study

In the current study we formulated the following research questions:

RQ1: What is the magnitude of L2 imitation of t-glottalization in English by Polish learners after immediate exposure?

RQ2: Is there a significant difference between a baseline condition (pre-test) and a post-test condition (carryover effect after exposure)?

RQ3: Is there a significant difference between intervocalic and non-prevocalic contexts of t-glottalization after both immediate exposure and in the post-test?

3.1. Participants

A total of 25 Polish students of English (21 females and 4 males) were recruited to participate in the study. They were all volunteers who received a small financial compensation (approximately 5 Euro) for their time. They ranged in age between 18 and 22 years with the mean of 20. They were all second-year students at the Institute of Linguistics with English as major. Their self-reported proficiency in English ranged

from B2 to C1 in the Common European Framework of Reference for Languages (CERFL). Their proficiency had additionally been verified by annual proficiency examinations in written tests (writing, text comprehension, grammar) and speaking tests (fluency, pronunciation). None of the participants reported any speech or hearing disorders. They had all normal or corrected-to-normal vision.

3.2. Materials

The experimental material consisted of 48 words that were imitated (i.e., an auditory stimulus was presented before reading the word aloud). Besides 16 filler words (such as *ago, always, floor, nothing*), which lacked any /t/ segment and which were not analysed, the material included two sets of target words with /t/: 16 words with an intervocalic coda /t/ (VtV: *beautiful, better, butter, cutting, daughter, eat up, energetic, forty, getting, hotter, it all, lot of, patriotic, relative, sort of, water*) and 16 words with non-prevocalic coda /t/ (VtC or Vt#: *a lot, bat, bit, cat, cut, eight, feet, fit, football, hot, hot weather, sit down, start, straight, what, white*). In addition, there were four other words with coda /t/ that were not imitated (*meet, nut, letter, city*), i.e., they were presented only orthographically in the pre-test and post-test. The aim is to examine whether any learning effect extends beyond the words presented auditorily in the imitation task.

The sound stimuli were produced by a range of speakers of Southern British English, ensuring variability of voices and idiolects. Most of the recordings were taken from an earlier sociophonetic study (Przedlacka 2002) that focused on the realization of coda /t/. The words were verbal responses to lexical questions in a spoken interview undertaken in several schools. The participants were aged 14 to 16 and raised in the Home Counties and London, thus their speech was phonologically close to Standard Southern British English. Although they produced a variety of /t/ realizations, only the forms with

t-glottalling (t-replacement) were selected. In addition, several recordings with t-glottalling were taken from YouTube videos. The stimuli were selected so that they would be comparable to tokens selected from the field recordings made by Przedlacka (2002), i.e., the speakers spoke with a Cockney accent, the words were produced in isolation and not in sentences, and the quality of the recordings was similar. All recordings were normalized to an RMS of 70 dB and saved as mono files with 16 kHz sampling rate.

3.3. Procedure

The participants' productions during the experiment were recorded individually in a sound-proof booth in the Speech Processing Laboratory, University of Silesia in Katowice. The experiment was run using the presentation software DMDX (Forster & Forster 2003). The audio signal was registered using a Sennheiser HMD headset dynamic microphone fed by a USBEPre2 (Sound Devices) amplifier. The signal was captured at 44,100 Hz with 24-bit quantization. In the imitation task, model voices were presented aurally at a 70-dB listening level. The experiment was composed of three tasks preceded by a familiarisation stage. In the first task (pre-test), the participants were asked to read the words presented in an orthographic form on the screen. In the second task (test), the participants were exposed to both orthographic forms and model talker productions and were instructed to immediately imitate the speaker they had heard as faithfully as possible (for a discussion about how instructions affect imitation see Pardo et al. 2018). In the third task (post-test), the procedure was similar to the first task in that the participants read orthographic representations of the words. The order of lexical items in each task was randomised for each speaker individually. The words were displayed on a 17-inch monitor using a black 28-point font against a white background. Each word was presented for 2800 milliseconds in the pre-test and post-

test and for 4000 milliseconds in the test, which was longer in order to accommodate the duration of the sound stimulus. Altogether, the participants produced 152 words during the three tasks. The whole individual session lasted approximately 20 minutes.

3.4. Measurement criteria

Each produced token was coded in terms of its phonetic realization using a combination of auditory and acoustic analysis. It was classified as (1) *glottalized*, which included all types of glottalization (from a full glottal stop [ʔ] to various lenited variants), and which included both glottal replacement and glottal reinforcement of /t/ (see Fig. 1c and 1b); or as (2) *non-glottalized*, where /t/-related glottalization was not present (for example, if the speaker had a creaky voice throughout the word, this was not a sufficient reason to consider the target segment glottalized). The non-glottalized tokens included alveolar stops with an oral gesture (e.g., [t], [tʰ], [d], [ɾ], [ts] plus variants with no audible release; see Fig. 1a). In addition, the /t/ was sometimes elided, especially word-finally (see Fig. 1d). In such cases, there was no trace of the /t/ segment left (and *white* was thus indistinguishable from *why*), unlike in the no audible release type, where the /t/ segment affects formant transitions in the offset of the vowel. However, both are considered non-glottalized, as opposed to glottalized tokens. Only the latter category showed typical signs of glottalization in the waveform and spectrogram (irregular, creaky phonation around the target segment, sometimes accompanied by a sudden cessation of energy at the offset of the vowel). Missing responses and tokens with mispronunciations ($n = 30$) were omitted in the analyses.

Figure 1 here

3.5. Statistical analysis

The output of the analysis was a binary dependent variable (1 = glottalized production, 2 = non-glottalized production), which was analysed statistically using mixed-effects logistic regression (Bates et al. 2015; Winter 2020). The fixed and random effects are specified in the results section. In order to evaluate the statistical significance of the predictors and their interaction, likelihood-ratio tests were performed, comparing the constructed model with a reduced model lacking the relevant term. Subsequently, pairwise comparisons were done using the *emmeans* package (Lenth 2020), accounting for multiple tests with the Bonferroni correction.

4. Analysis and results

4.1. Test items

The 32 main target items were presented in all three tasks. A total of 2370 valid responses were collected (Table 1). The data are displayed in Figure 2, with glottalized tokens in black bars. Clearly, their number was very small in the baseline condition, and all such instances were in the non-prevocalic context (VtC or Vt#). However, the participants responded substantially to the exposure of t-glottaling in the imitation task, leading to an increase in the number of glottalized tokens in both conditions. This effect seems to weaken in the post-test task, and there were more glottalized responses in the non-prevocalic condition again. Trial order effects were not detected.

Task	Position	N	% glottalized	Lower CI	Upper CI
Pre-test	Intervocalic	400	0.0	0.0	0.9
Pre-test	Non-prevocalic	400	6.5	4.3	9.4
Test	Intervocalic	400	30.3	25.8	35.0
Test	Non-prevocalic	400	30.0	25.6	34.8
Post-test	Intervocalic	385	7.5	5.1	10.6
Post-test	Non-prevocalic	385	17.4	13.8	21.6

Table 1: The percentage of glottalized tokens in the three tasks and two conditions. 95% confidence intervals (based on binomial tests computed from the whole sample, see N) are provided.

Figure 2 here

To test the effect of task and position on the number of glottalized tokens, an LME logistic regression model was constructed with TASK (T1 [= reference level], T2, T3) and POSITION (interV [= reference level], non-preV) as fixed effects (including their interaction), and ITEM, PARTICIPANT, and by-participant slopes for task and position as random effects (by-item slopes could not be added due to convergence problems). Due to an absolute zero in one condition (T1, Intervocalic; see Tab 1.), statistical evaluation is problematic. Rather than dropping that condition from the model or keeping the absolute zero, which would limit the number of possible comparisons, we opted for setting one datapoint chosen at random to the other value (i.e., a glottalized item). Consequently, the full structure of the experimental design could be examined. Model summary is provided in Table 2 and the effect plot is shown in Figure 3.

The interaction between TASK and CONDITION proved to be significant ($\chi^2(2) = 34.4, p < 0.001$). Namely, Tukey post-hoc tests revealed that the difference between intervocalic and non-prevocalic items was evaluated as significant in T1 (odds ratio = 0.02, SE = 0.02, z ratio = -3.5, $p < 0.001$) and T3 (odds ratio = 0.15, SE = 0.09, z ratio = -3.2, $p = 0.002$) but not in T2 (odds ratio = 0.74, SE = 0.36, z ratio = -0.6, $p = 0.543$). . Furthermore, the imitation task (T2) was significantly different from both other tasks in the non-prevocalic condition (T1/T2: odds ratio = 0.13, SE = 0.06, z ratio = -4.8, $p < 0.001$; T2/T3: odds ratio = 3.25, SE = 1.03, z ratio = 3.7, $p < 0.001$), while the increase

from T1 to T3 was not evaluated as significant (odds ratio = 0.42, SE = 0.21, z ratio = -1.7, $p = 0.257$). In the intervocalic position, all three comparisons were statistically significant, including the T1 to T3 increase (T1/T2: odds ratio = 0.003, SE = 0.003, z ratio = -5.2, $p < 0.001$; T2/T3: odds ratio = 15.63, SE = 6.76, z ratio = 6.4, $p < 0.001$; T1/T3: odds ratio = 0.05, SE = 0.05, z ratio = -2.7, $p = 0.023$).

Parameter	Estimate	SE	z value	p value
Intercept [= reference level]	-7.54	1.19	-6.3	< 0.001
Task (T2)	5.84	1.12	5.2	< 0.001
Task (T3)	3.10	1.16	2.7	0.008
Position (Non-preV)	4.10	1.18	3.5	< 0.001
Task (T2) : Position (Non-preV)	-3.80	1.09	-3.5	< 0.001
Task (T3) : Position (Non-preV)	-2.23	1.11	-2.0	0.044

Table 2: Fixed effect parameters of the GLME model (logit values). The reference level is T1 intervocalic.

Figure 3 here

Two issues might be of concern here. It is obvious that the model’s predicted means (Fig. 3) are generally much lower than the observed means in Table 1. The source of this discrepancy seems to lie in the random effects component – the distribution of the Best Linear Unbiased Predictions (BLUPs) is non-centred, and their sum (for both participants and items) is thus not close to 0 but positive (probably due to unbalanced data). Therefore, adding the random effects to the linear predictor of the fixed effects would, contrary to expectations, increase the average prediction somewhat, which would be more in line with the observed percentages.²

² Interestingly, when we fit the data with a numeric dependent variable (0 or 1) using a standard lmer model with the same effect structure as before, the predicted means align with the

The second issue is the presence of zeros in the observed data, which was dealt with by setting one datapoint to 1 (i.e., a glottalized item). To test whether our model is robust, we conducted ten additional analyses in which the location of the alteration was varied at random within the 400 items of the T1, intervocalic condition. None of the ten models showed marked differences from the model reported here, suggesting it is not sensitive to the specific cell being altered. Averages of the parameters for the ten models are displayed in Table 3. Average post-hoc Tukey tables also looked very similar, with no changes in the significance of individual comparisons. Finally, we conducted an analysis with the absolute zero condition dropped from the model; as for the comparisons that were possible in this reduced design, the effects and p-values were again similar.

Parameter	Estimate	SE	z value	p value
Intercept [= reference level]	-7.34	1.17	-6.3	< 0.001
Task (T2)	5.68	1.11	5.1	< 0.001
Task (T3)	2.98	1.15	2.6	0.010
Position (Non-preV)	3.95	1.17	3.4	< 0.001
Task (T2) : Position (Non-preV)	-3.69	1.09	-3.4	< 0.001
Task (T3) : Position (Non-preV)	-2.17	1.10	-2.0	0.049

Table 3: Means of fixed effect parameters of ten GLME models (logit values) differing in the datapoint manipulation (see text). The reference level is T1 intervocalic.

observed means almost perfectly. P-values are generally higher now, and the significance of pairwise comparisons changes somewhat. There is now no significant difference between positions in T1 ($p = 0.081$), and the effects of task are altered for the comparison T1 / T3 (in intervocalic position, the difference is no longer significant with $p = 0.128$, while in non-prevocalic position, there is now a significant difference with $p = 0.006$).

4.2. Generalization items

The four non-imitated words (*meet, nut, letter, city*) were present only in the pre-test and post-test. Figure 4 reveals that the effect of exposure to t-glottalling had a very small generalization effect to new words. The few glottalized tokens in the post-test were constrained to the non-prevocalic position.

Figure 4 here

5. Discussion

This study examined the ability of Polish learners to imitate t-glottaling in English words. Exposure to stimuli with glottal realisations of coda /t/ in an imitation task had strong immediate effects, with a third of the tokens realised with glottal variants in both conditions. However, the retention of the effect was weak, with a substantial drop in the number of glottal variant tokens in the post-test task, this decrease being more extensive in the intervocalic position. Those results are in line with Šturm et al. (2022), where the non-prevocalic position was likewise associated with higher rates of glottalization in imitations by Czech learners. The relatively low level of the carry-over effect from imitation to post-imitation productions – which was nevertheless significant in the intervocalic position – seems to point to the separate mechanisms governing imitated speech production and non-imitated speech production. The former seems to be largely conditioned by purely imitative psychoacoustic reactions while the latter resorts to the learners' lexical and phonological representations.

A similar dissociation was found by Llompart & Reinisch (2019), who compared imitation with perceptual categorisation and word reading. The difference between the

study by Llompart & Reinisch (2019) and the current study is that in their study the participants first completed the word reading task followed by the imitation task in order to avoid any potential influences from imitation onto word reading. In our study the word reading task immediately preceded and followed imitation with the view of directly comparing the baseline productions (pre-imitative word reading) to the productions influenced by imitation (post-imitative word reading). Despite these methodological differences, we also found that the learners' productions quickly regained the features of their default baseline pre-imitative speech suggesting that, although imitation may be, to a certain extent, mediated by the interaction of phonetic properties of L1 and L2 sounds, it does not necessarily directly exploit the learners' productive phonological representations in L2. Such a decay of fine-grained phonetic detail may be predicted by limitations of the working memory (Baddeley 2010; Baddeley Hitch 1974), due to which imitated phonetic episodic traces are lost, resulting in a likely switch from purely phonetic imitative productions to default phonologically-coded productions (Chen et al. 2023; Hao & de Jong 2016; Rojczyk 2012).

Regarding the position of the target /t/ phoneme, the pre-test and post-test tasks yielded higher glottalization rates in the non-prevocalic than in the intervocalic condition. We may assume that in the baseline condition, the intervocalic position offers little room for the Polish speakers to glottalize spontaneously, especially in such a controlled task as reading words. However, the lower imitation performance in the VtV in the post-test and equal performance regarding VtV and Vt#/VtC during the imitation task contradicts our expectation based on language-internal considerations.

Our hypothesis was that the intervocalic position would produce higher rates of glottal productions because (1) Polish word-initial vowel glottalization excludes the non-prevocalic context and (2) intervocalic glottalisation is perceptually more

salient. Familiarity with Polish intervocalic glottalization was thus expected to facilitate the imitation in that position in English. This prediction does not find support in the current results. Glottalization was more successfully imitated in the contexts where it is absent in Polish than in those where it frequently occurs. This result may be due to the fact that in Polish intervocalic glottalization is used with a very different function, as a syllable structure, not an allophonic phenomenon, thus making L1 transfer more difficult.

Overall poor retention of glottalisation might be compounded by sociolinguistic factors. Studying English at degree level, our imitators are being exposed to standard varieties of English in the university classroom settings. The majority of those students tend to be highly motivated to use a standard accent variety as they move on to professional occupations (language instruction, translation and interpreting), where competence in a standard is implicitly required. Despite a steadily increasing presence of t-glottalling in SSBE (Standard Southern British English) in the non-prevocalic positions (Fabricius 2002), there is a history of negative attitudes towards glottal variants of English voiceless stops (Towarnicky 2022). Coupled with the fact that glottalisation is rarely explicitly trained and often a conservative view of phonetic features of current English speech, these factors contribute to a perception that the expected native-like English pronunciation has little glottalisation. Therefore, despite the presence of this feature in the model stimuli, it is not adopted by the imitators.

This is an interesting phenomenon that deserves further attention in terms of how the imitated feature may be suppressed or blocked as a result of incongruence with the imitators' image of correct speech in L2. This suggests that sociolinguistic and cultural factors may mediate imitative performance.

The pool of 25 participants and 32 target words did not behave uniformly; there was substantial variability across both participants and items. The glottalization rate across subjects in the imitation task ranged from 0% (participants completely immune to imitation) to 84% (participants imitating surprisingly well). The mean rate per participant was 30%. A similar scenario arose across items, with glottalization rates ranging from 4% to 64% (words like *football*, *beautiful*, *sit down*, *eight*, *cutting* in the lower range, *what*, *hot*, *water*, *cut*, *lot of* in the higher range). The length of the word seemed not to play a role as, for instance, *hot* and *hotter* were both in the higher range and, generally, both short and long items appeared at either end. Nevertheless, there was a tendency for open vowels to be associated with higher rates of glottalization. Specifically, the upper third of items contained /ɒ/, /ɔ:/, /ʌ/, /æ/ or /ɑ:/ before the alveolar plosive, whereas the lower third contained /ʊ/, /u:/, /ɪ/, /i:/, /eɪ/ or /aɪ/ (and /ə/ in *relative* and /ʌ/ in *cutting*, which have /ɪ/ in the following vowel). However, no systematic asymmetries regarding the position variable could be observed.

Future research will inform studies on phonetic imitations by incorporating other variables that may importantly impact the results such as recruiting learners of English whose L1 has little glottalization whatsoever, such as e.g., Spanish or Italian. Moreover, future research should compare imitations that are implicit versus those that are explicitly evoked by a clear instruction to sound like a model talker.

Data Availability Statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Acknowledgements

Research supported by the National Science Centre Poland grant *Phonetic imitation in a native and non-native language* (UMO-2019/35/B/HS2/02767) to the first author and by the funds granted under Research Excellence Initiative of the University of Silesia in Katowice.

Supplementary materials

OSF: <https://osf.io/p3zes/>

References

- Adank, Patti, Peter Hagoort, & Harold Bekkering. 2010. Imitation improves language comprehension. *Psychological Science* 21(12). 1903-1909.
- Ashby, Michael & Joanna Przedlacka. 2014. Measuring incompleteness: Acoustic correlates of glottal articulations. *Journal of the International Phonetic Association* 44. 283-296.
- Babel, Molly. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39. 437-456.
- Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40. 177-189.
- Babel, Molly & Dasha Bulatov. 2012. The role of fundamental frequency in phonetic accommodation. *Language and Speech* 55. 231-248.
- Baddeley, Alan. 2010. Working memory. *Current Biology* 20(4). R136-R140.
- Baddeley, Alan & Graham Hitch. 1974. Working memory. *Psychology of Learning and Motivation* 8. 47-89.
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1-48.

- Bogharti, Reihane, Joe Hoover, Kate M. Johnson, Justin Garten, & Morteza Dehghani. 2016. Syntax accommodation in social media conversation. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, 1463-1468, Cognitive Science Society.
- Carignan, Christopher. 2018. Using naïve listener imitations of native speaker productions to investigate mechanisms of listener-based sound change. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 9(1). 1-31.
- Chen, Juqiang, Catherine T. Best, & Mark Antoniou. 2023. Phonological and phonetic contributions to Thai-naïve Mandarin and Vietnamese speakers' imitation of Thai lexical tones: Effects of memory load and stimulus variability. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 14(1). 1-41.
- Cruttenden, Alan. 2014. *Gimson's pronunciation of English*. 8th edition. London: Routledge.
- Docherty, Gerard, J. & Paul Foulkes. 1999. Derby and Newcastle: Instrumental phonetics and variationist studies. In Paul Foulkes and Gerard J. Docherty (eds.), *Urban voices*, 47–71. London: Arnold.
- Dufour, Sophie & Noël Nguyen. 2013. How much imitation is there in a shadowing task? *Frontiers in Psychology* 4. Article 346. 1-7.
- Dunstan, Stephany B. 2010. Identities in transition: The use of AAVE grammatical features by Hispanic adolescents in two North Carolina communities. *American Speech* 85(2). 185-204.
- Fabricius, Anne. 2002. Ongoing change in modern RP: Evidence for the disappearing stigma of t-glottalling. *English World-Wide* 23(1). 115–136.

- Flege, James E. & Wieke Eefting. 1988. Imitation of VOT continuum by native speakers of English and Spanish: Evidence for phonetic category formation. *Journal of the Acoustical Society of America* 83(2). 729-740.
- Forster, Kenneth I. & Jonathan C. Forster. 2003. DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers* 35(1). 116-124.
- Fowler, Carol A., Julie M. Brown, Laura Sabadini, & Jeffrey Weihing. 2003. Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language* 49. 296-314.
- Gregory, Stanford W. & Stephen Webster. 1996. A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status predictions. *Journal of Personality and Social Psychology* 70. 1231-1240.
- Gries, Stefan, T. 2005. Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research* 34(4). 365-399.
- Hao, Yen-Chen & Kenneth de Jong. 2016. Imitation of second language sounds in relation to L2 perception and production. *Journal of Phonetics* (54). 151-168.
- Hauser, Marc D. 1996. *The evolution of communication*. Cambridge, MA: MIT Press.
- Honorof, Douglas N., Jeffrey Weihing, & Carol A. Fowler. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39. 18-38.
- Jiang, Fan & Shelia Kennison. 2022. The impact of L2 learners' belief about and interlocutor's English proficiency on L2 phonetic accommodation. *Journal of Psycholinguistic Research* 51. 217-234.
- Katzir, Roni & Raj Singh. 2013. A note on presupposition accommodation. *Semantics and Pragmatics* 6. 1-16.

- Kaźmierski, Kamil. 2020. Prevocalic t-glottaling across word boundaries in Midland American English. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1): 13. 1–23.
- Keating, Patricia, Marc Garellek & Jody Kreiman 2015. Acoustic properties of different kinds of creaky voice. In *Proceedings of the 18th International Congress of Phonetic Sciences* (paper 821), Glasgow. University of Glasgow.
- Lenth, Russell. 2020. emmeans: Estimated Marginal Means, aka Least-Squares Means. (1.5.2-1) [Computer software]. Available at: <https://CRAN.R-project.org/package=emmeans>
- Llompart, Miquel & Eva Reinisch. 2018. Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech* 62(3). 594-622.
- Marlsen-Wilson, William D. 1973. Linguistic structure and speech shadowing at very short latencies. *Nature* 244. 522-523.
- Marlsen-Wilson, William D. 1975. Sentence perception as an interactive parallel process. *Science* 189.226-228.
- Mitterer, Holger & Mirjam Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109. 168-173.
- Mitterer, Holger & Jochen Müsseler. 2013. Regional accent variation in the shadowing task: Evidence for a loose perception-action coupling in speech. *Attention, Perception and Psychophysics* 75(3). 557-575.
- Namy, Laura L., Lynne C. Nygaard, & Denise Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21. 422-432

- Nielsen, Kuniko. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2). 132-142.
- Pardo, Jennifer S., Isabel Cajori Jay, & Robert M. Krauss. 2010. Conversational role influences speech imitation. *Attention, Perception and Psychophysics* 72. 2254-2264.
- Pardo, Jennifer S., Rachel Gibbons, Alexandra Suppes, & Robert M. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40. 190-197.
- Pardo, Jennifer S., Adelya Urmanche, Sherilyn Wilman, Jaclyn Wiener, Nicholas Mason, Keagan Francis, Melanie Ward. 2018. A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics* 69. 1-11.
- Peschke, Claudia, Wolfram Ziegler, J. Kappes, & Annette Baumgaertner. 2009. Auditory-motor integration during fast repetition: The neuronal correlates of shadowing. *NeuroImage* 47(1). 392-402.
- Podlipský Václav J. & Šárka Šimáčková. 2015. Phonetic imitation is not conditioned by preservation of phonological contrast but by perceptual salience. In *Proceedings of the 18th International Congress of Phonetic Sciences*.
- Porter, Robert J. & F. Xavier Castellanos. 1980. Speech-production measures of speech perception: Rapid shadowing of VCV syllables. *Journal of the Acoustical Society of America* 67. 1349-1356.
- Porter, Robert J. & James F. Lubker. 1980. Rapid reproduction of vowel-vowel sequences: Evidence for a fast and direct acoustic-motoric linkage in speech. *Journal of Speech and Hearing Research* 23. 593-602.
- Przedlacka, Joanna. 2002. *Estuary English? A sociophonetic study of teenage speech in the Home Counties*. Frankfurt: Peter Lang.

- Redi, Laura & Stephanie Shattuck-Hufnagel. 2001. Variation in the realization of glottalization in normal speakers. *Journal of Phonetics* 29(4). 407–429.
- Rojczyk, Arkadiusz. 2012. Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EUROSLA 2012, Poznań Poland.
- Rojczyk, Arkadiusz. 2013. Phonetic imitation of L2 vowels in a rapid shadowing task. In John Levis & Kimberly LeVelle (eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, 66-76. Ames, IA: Iowa State University.
- Rojczyk, Arkadiusz, Andrzej Porzuczek & Marcin Bergier. 2013. Immediate and distracted imitation in second-language speech: Unreleased plosives in English. *Research in Language* 11. 3-18.
- Rojczyk, Arkadiusz & Lucrecia Rallo Fabra. 2023. Geminate rearticulation is psychoacoustically robust: Phonetic imitation of Polish geminates by Spanish-Catalan bilinguals. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS), Prague the Czech Republic*.
- Schwartz, Geoffrey. 2012. Initial glottalization and final devoicing in Polish English. *Research in Language* 10(2). 159-171.
- Schwartz, Geoffrey. 2017. The perception of *sandhi*-blocking in Polish vowel-initial words. *Folia Linguistica* 51(3). 671-693.
- Seyfarth, Scott. & Marc Garellek, M. 2020. Physical and phonological causes of coda /t/ glottalization in the mainstream American English of central Ohio. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1). 1–33.
- Shockley, Kevin, Laura Sabadini, & Carol A. Fowler. 2004. Imitation in shadowed words. *Perception and Psychophysics* 66(3). 422-429.

- Towarnicky, Dayle M. 2022. *Examining the linguistic ideology “Throaty sounds are bad for performers”: The history of negative attitudes towards glottal stops and laryngealization in English*. Doctoral Dissertation, City University of New York.
- Trofimovich, Pavel & Sara Kennedy. 2014. Interactive alignment between bilingual interlocutors, evidence from two information-exchange tasks. *Bilingualism, Language and Cognition* 17(4). 822-836.
- Ulbrich, Christiane. 2021. Phonetic accommodation on the segmental and suprasegmental level of speech in native – non-native collaborative tasks. *Language and Speech*, DOI: <https://doi.org/10.1177/00238309211050094>
- Šturm, Pavel, Joanna Przedlacka & Arkadiusz Rojczyk. 2022. Phonetic imitation of t-glottaling by Czech speakers of English. *Linguistica Pragensia* 32(1). 142-165.
- Winter, Bodo. 2020. *Statistics for Linguists: An Introduction Using R*. London: Routledge.
- Zajac, Magdalena & Arkadiusz Rojczyk. 2014. Imitation of English vowel duration upon exposure to native and non-native speech. *Poznan Studies in Contemporary Linguistics* 50(4). 495-514.

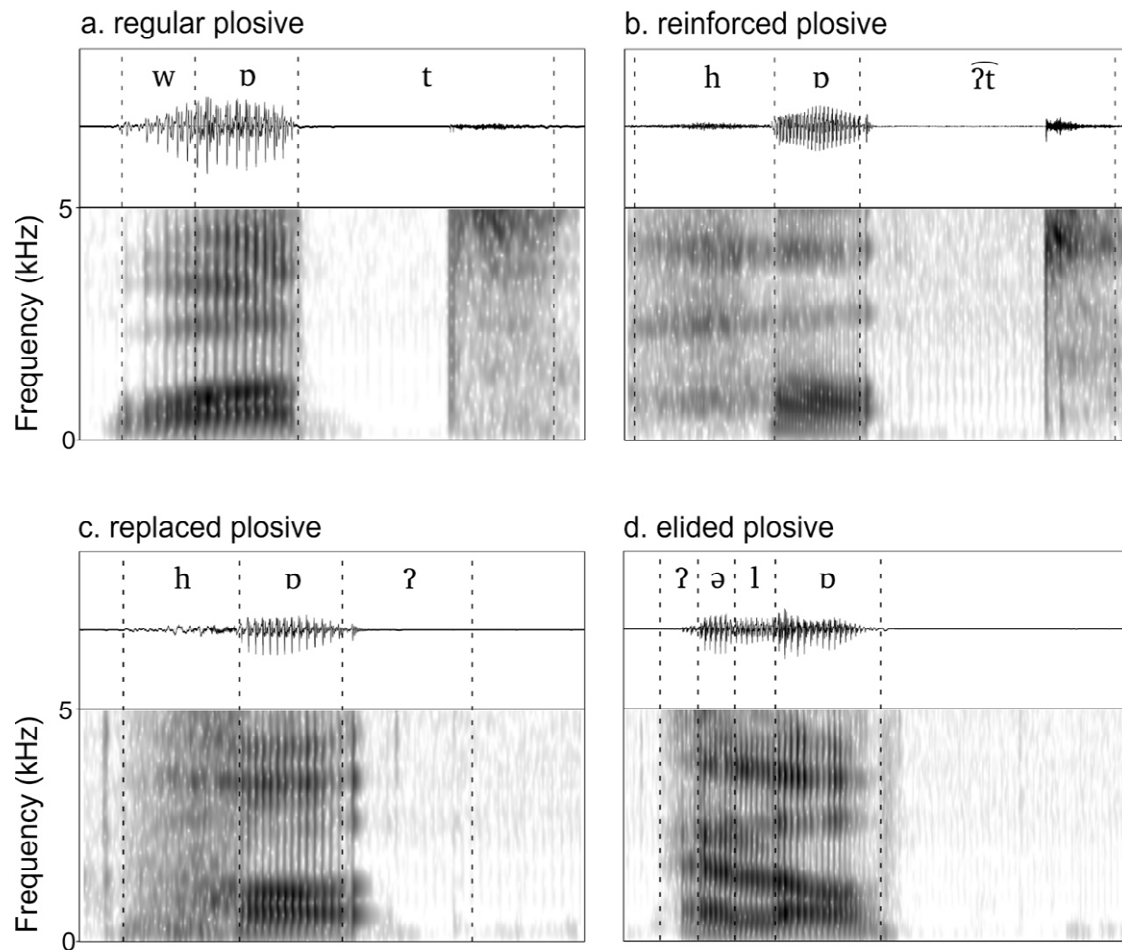


Figure 1: Examples of target plosives: **a)** regular plosive in *what*; **b)** glottally-reinforced plosive in *hot*; **c)** glottally-replaced plosive in *hot*; **d)** elided plosive in *a lot*. Only the cases in b) and c) are considered ‘glottalized’, i.e., successfully imitated in our analysis.

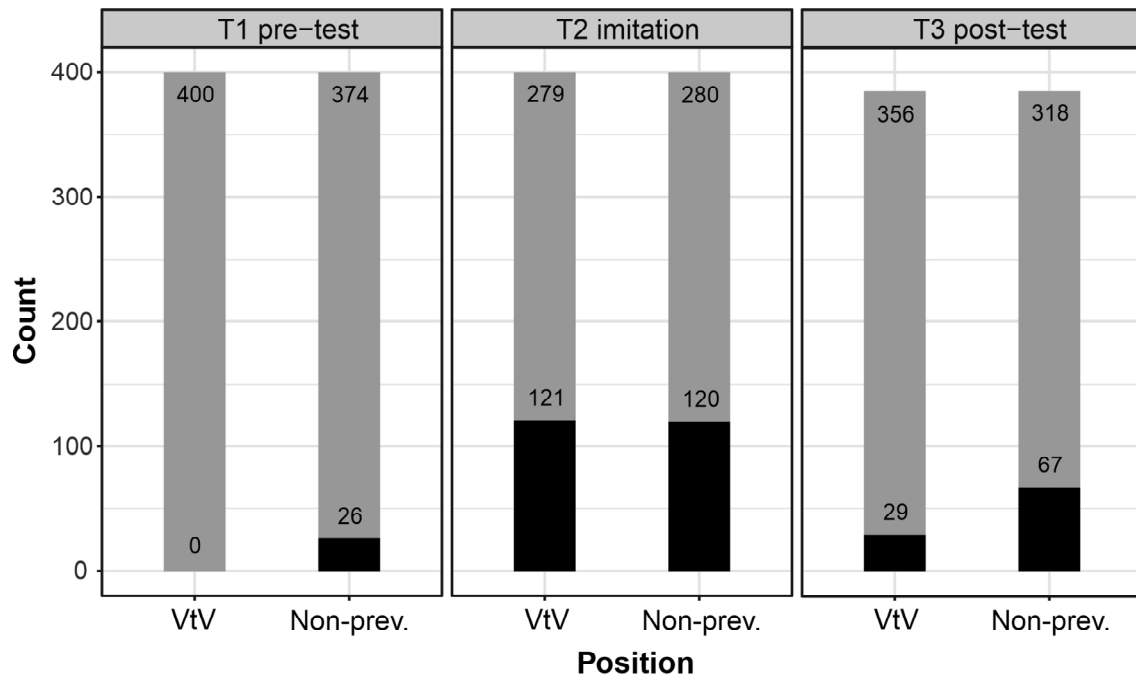


Figure 2: The number of glottalized (black) and non-glottalized (grey) tokens in the three tasks and two conditions.

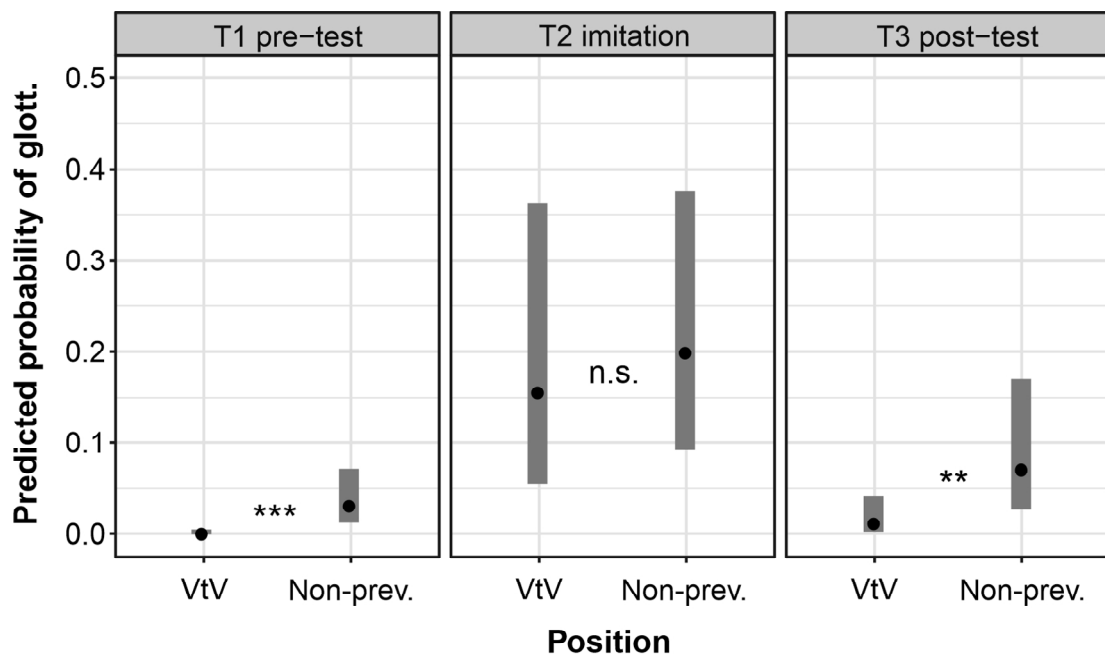


Figure 3: Predicted probability of glottalization as a function of task and position. The model allows for an interaction between the two effects.

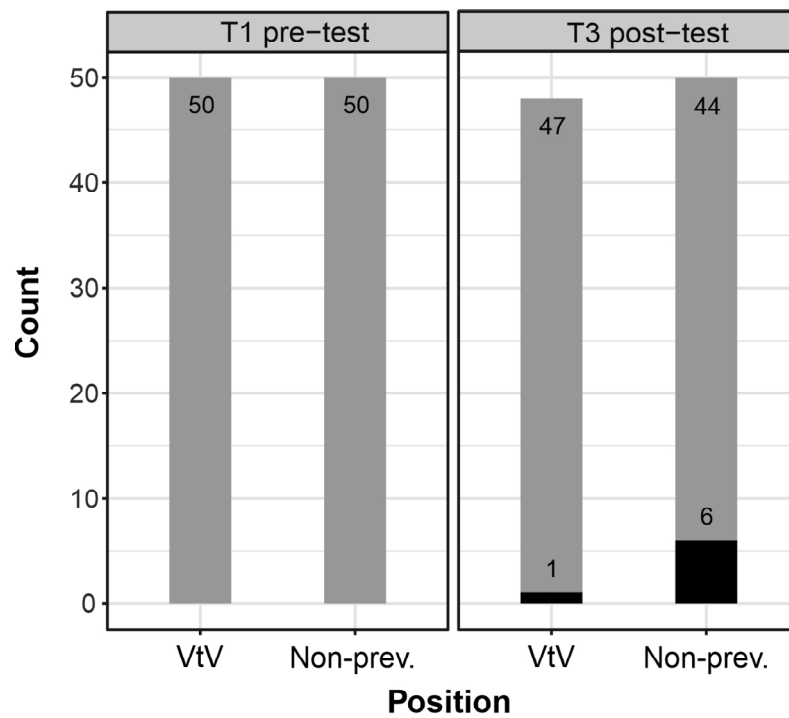


Figure 4: The number of glottalized (black) and non-glottalized (grey) tokens in the non-imitated tasks and two conditions. Only the words *meet*, *nut*, *letter*, *city* that were not presented auditorily.