PERCEPTION-PRODUCTION MISMATCH IN L1 DRIFT

Jolanta Sypiańska[1] and Zuzanna Cal[2]

[1]University of Szczecin, Poland; [2]Adam Mickiewicz University, Poznań, Poland
jolanta.sypianska@usz.edu.pl, zuzanna.cal@amu.edu.pl

ABSTRACT

Most studies on the relationship between perception and production show an interdependency of the two modalities. We investigated whether the perception-production link is also present in L1 drift. L1 Polish/L2 English speakers were asked to decide whether they heard a voiced or voiceless stop in a two-alternative forced-choice (2AFC) task. The stimuli were obtained from VOT continua of word-initial (/p,b,t,d,k,g/). We expected that L1 speakers of a true voicing may experience L1 perceptual drift from the aspirating L2 English by placing the boundary between voiced and voiceless at higher VOT values. The results showed that the speakers perceived the voiceless-voiced boundary at 0ms. However, they produced a great amount of L1 aspirated stops and realized some L1 voiced stops either as voiceless or with limited prevoicing. We conclude that, in L1 drift situations, there may be a perception-production mismatch that defines L1 drift as an element of execution without influencing phonological representations.

Keywords: L1 drift, perception, production, aspirating language, true voicing language.

1. INTRODUCTION

There is a body of research that supports the interdependence of perception and production both in the native [1] and foreign languages [2,3]. The two modalities must share or have access to at least some part of the sounds' representation. Moreover, the perception-production link defines perception as a precondition to production. Speakers first possess the ability to recognise that the sound belongs to a particular category, only then can they produce the sound accurately. Widely accepted models of speech production in foreign language acquisition, such as the Speech Learning Model [4] or the Perceptual Assimilation Model [5], advocate the link by stressing that accurate production of an L2 sound is possible only when the learner establishes an L2 sound category that is sufficiently different from a similar L1 sound category. However, some studies have found a perception-production mismatch where either perception exceeded production [2,6] or production but not perception was on target with regard to native speaker norms [7].

This paper focuses on the perception and production of voiced-voiceless stops. The differences in voice onset time (VOT) that lead to differences in fortis and lenis stop realisation in true voicing languages versus aspirating languages have been a frequently discussed issue in L1 drift literature. In terms of perception, [8] found that L1 English, L2 French novice learners shifted their boundary between voiced and voiceless stops towards the more French-like values after intensive exposure to L2 French in L2 immersion. [7] compared L1 Spanish, L2 English and L1 English, L2 Spanish groups with different lengths of residence in the L2 country and showed that each group had less native-like L1 category boundaries. However, in an L1 setting, no statistically significant shift in the boundary between the voiced and voiceless bilabial stops was observed [9].

In production, [10] found that L1 Japanese, L2 English speakers prolong their L1 Japanese VOT in Japanese short-lag stops under the influence of L2 English long-lag stops. Interestingly, [11] describe an asymmetry in the susceptibility to L1 drift between the voiced and voiceless stop series in which the L1 Polish voiced stops showed more drift than the voiceless stops under the influence of L2 English. Also, [12] noticed that only voiced stops in both languages of two English-Czech simultaneous bilinguals drifted towards more negative, so more Czech-like values, after an extended stay in Czechia. There was no change in the voiceless stops for either language. However, in a study on L1 Polish senior learners of L2 English, it was mostly the voiceless series that underwent L1 drift and the speakers largely maintained prevoicing in the voiced series [13].

In terms of the perception-production link in L1 drift setting, [14] showed that the L1 English, L2 Spanish group with the greatest amount of L2 experience produced short VOT in L1 voiceless stops (more Spanish like) but placed the perceptual boundary for voiced and voiceless at higher VOT values in comparison to other groups. No statistically significant correlations between production and perception were found in other groups in the study. These results suggest that there may be no clear link between the two dimensions.

The perception-production link has been explored in an L1 drift setting only to a limited extent. Thus, the aim of the paper is to investigate L1 drift in perception and production of voiced-voiceless stops, an element that has previously been shown to undergo drift in both modalities but was mostly investigated separately. The first hypothesis is that L1 drift will be manifested in both modalities of L1 Polish fortis and lenis stops under the influence of L2 English. The second hypothesis is based on the perception-production link and states that the perception of L1 Polish stops will be reflected in L1 production. The manifestation of the link should be visible in a higher VOT category for fortis and lenis stops reflected in a tendency to produce aspirated voiceless fortis stops and voiceless lenis stops in L1 Polish.

2. METHOD

2.1. Participants

Twenty-four L1 Polish, L2 English speakers took part in the experiment including twenty-three females and one male participant. They were all Polish students of English studies at Adam Mickiewicz University in Poznań whose age ranged from 19 to 27. They were advanced learners of English and the L2 proficiency was measured with the Lexical Test for Advanced Learners of English (LexTale) [15]. Their self-reported average of English instruction per week was 20 hours. Most of them reported past exposure to German but currently claimed having no instruction or other forms of exposure in an aspirating language.

2.2. Procedure

The experiment consisted of four parts. First of all, the participants were asked to fill in a questionnaire, which included questions about their biodata as well as language learning history and language use. After that, they completed LexTale [15] as a measure of English proficiency. Having completed the test, they participated in the perception experiment and later took part in a recording session.

2.2.1. Perception task

The perception part of the experiment involved preparation of VOT continua based on one-syllable minimal pairs with word initial stop sounds (/p,b,t,d,k,g/). The continua ranged from -96 ms to +96 ms with 17 steps each of 12 ms and were prepared using one prevoiced and one aspirated stop produced by a trained phonetician with Polish as the native language. Three places of articulation namely bilabial, dental and velar were taken into consideration. The tokens were presented in one syllable words including par, bar, tam, dam, kas, gaz. Participants took part in a two-alternative forced-choice (2AFC) task programmed in the E-Prime 3.0 software [16]. After hearing a sound, they were asked to decide which word, out of two presented on a screen, they heard. The presented words differed only by voicing of the initial stop sounds. Each sound from the continua appeared twice in the experiment, thus the participants were presented with 102 tokens altogether.

2.2.2. Production task

In the production task, the participants were recorded reading the minimal-pair words used in the perception experiment with word-initial voiced and voiceless stops (par, bar, tam, dam, kas, gaz) mixed with filler words. The participants were presented with the words on a computer screen via a PowerPoint presentation and were asked to read words as they appeared. Finally, in order to check whether the participants themselves produce aspirated and voiceless short-lag stops in their L2, they were asked to produce English words with initial voiced and voiceless plosives in an analogical context to the L1 vocabulary used for the E-prime experiment (pat, bat, tan, Dan, cast, gas), also together with filler words. The recording session took place in a quiet room and the task was approximately 5 minutes long.

3. RESULTS

3.1. Perception

To calculate the location of perceptual boundaries in each place of articulation (PoA), the data was converted using a logistic regression function in SPSS [17] and the boundary location was computed with the use of the following formula: $-LN(b0)/LN(b1)$, where $b0$ corresponded to the constant and $b1$ to the slope of the function. Figure 1 presents mean values of boundary location three PoA - bilabial, coronal and velar.

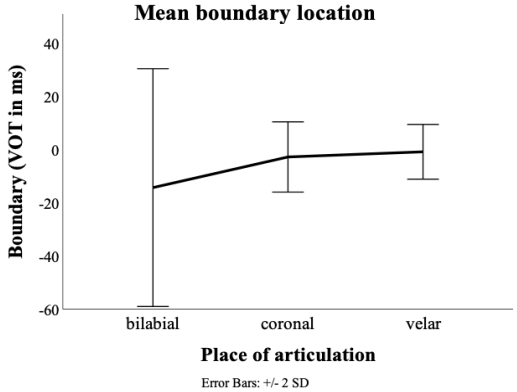

Figure 1. Mean perceptual boundary locations (in ms) for all three places of articulation

The bilabial continuum turned out to have the earliest boundary yielding on average -14.22 ms (SD = 22.38). In coronal stops the mean boundary location was -2.68 ms (SD = 6.61) and the velar continuum appeared to have a slightly higher boundary, that is -0.74 ms (SD = 5.15).

3.2. Production

The participants' stop realisations (N=283, 142 fortis sounds, 141 lenis sounds) were analysed in Praat by measuring VOT manually. The productions were then coded into five categories, two for fortis stops: aspirated and voiceless, and three for voiced stops: prevoiced, voiceless and mixed (Fig.2). The last category refers to tokens of voiced stops that include portions of voiceless signal at some point between the closure and the release as opposed to prevoiced tokens which maintain voicing throughout the duration of the closure. The results show that the speakers have acquired aspiration and use it as a main type of realisation for fortis L2 English stops (almost 100% correctness, Fig.3). Aspirated stops also dominate the L1 fortis stop production which is why more than two-thirds of the Polish fortis stops are produced incorrectly. Lenis stop production is more varied for both languages. In L1 Polish, the correct prevoiced realisation is present in 60% of tokens whereas the incorrect productions belong to the voiceless or mixed categories. Finally, prevoiced and mixed tokens in L2 English productions amount to approximately 50% of incorrect realisations (Fig.3.).

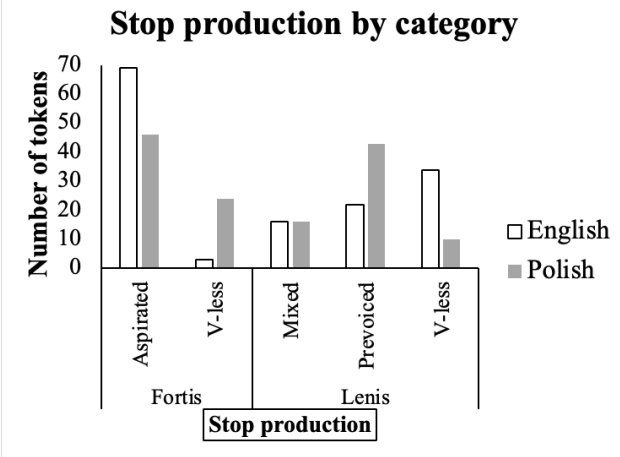

Figure 2. Number of tokens in Polish and English produced in each category for fortis stops: aspirated, voiceless, and lenis stops: mixed, prevoiced and voiceless

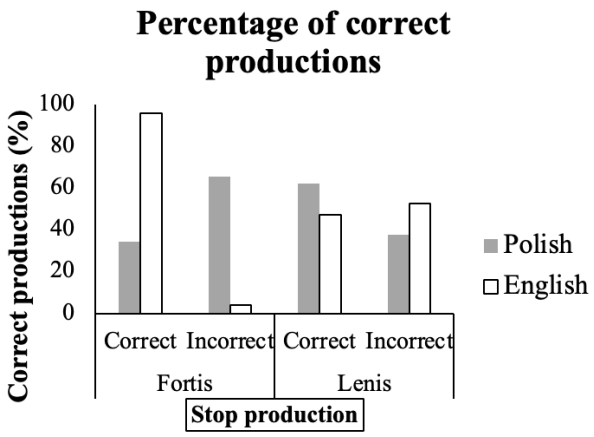

Figure 3. Percentage of correct fortis and lenis stop productions in Polish and English

3.3. Perception-production link

A linear regression was run to examine the link between perception and production. The boundary (dependent variable) was regressed on predicting categorical variables of fortis realisation (voiceless, aspirated) and lenis realisation (voiceless, prevoiced, mixed). The analysis was conducted separately for each continuum. Table 1 provides the results of the statistical analysis.

Dependent variable: Perceptual boundary (VOT in ms)

Independent variable	Coefficient (B)	Standard Error	t-Statistic	P-value
PoA: bilabial				
Fortis[a]				
Voiceless	7.909	11.553	.685	.504
Lenis[b]				
Voiceless	7.428	17.967	.413	.686
Mixed	7.979	17.967	.444	.664
PoA: coronal				
Fortis[a]				
Voiceless	-.785	3.309	-.237	.815
Lenis[b]				
Voiceless	7.985	6.815	1.172	.258
Mixed	3.564	3.713	.960	.351
PoA: velar				
Fortis[a]				
Voiceless	-1.307	2.564	-.510	.616
Lenis[b]				
Voiceless	2.467	2.907	.849	.408
Mixed	4.776	2.569	1.859	.080

[a] reference factor = Aspirated
[b] reference factor = Prevoiced

Table 1. Results of linear regression across places of articulation

No effect of fortis realisation was found in all three places of articulation, that is bilabial (F=.469, p=.504), coronal (F=.56, p=.815) and velar (F=.260 p=.616). Similarly, lenis realisation was not found to

be a predicting variable for the perceptual boundary location in all places of articulation, namely bilabial (F=.162, p=.825), coronal (F=1.031, p=.378), velar (F=1.728, p=.207).

4. DISCUSSION AND CONCLUSION

The perception data showed that the participants placed the perceptual boundary between voiced and voiceless coronal and velar stops in their L1 Polish at approximately 0ms that is within the same 12ms step of the VOT continuum. The boundary for the bilabial stop was placed one step earlier in the prevoicing part of the stop. These results indicate that the speakers did not experience perceptual drift which would be manifested in a category boundary at higher (positive) VOT values.

The production data from the L2 showed that the speakers have acquired aspiration in the L2, and at least partly, learned to suppress prevoicing in the L2 lenis stops. This gives a basis for cross-linguistic influence from the L2 to the L1. Furthermore, L1 production data revealed that aspiration was present in the majority of L1 tokens resulting in an error rate of more than 60%. However, lenis stops were produced with greater variation and only 60% of the tokens were correctly prevoiced.

The first hypothesis regarding L1 drift under the influence of the L2 in both L1 perception and production has only partly been corroborated. With a basis for cross-linguistic influence from the L2, it can be stated that L1 drift in Polish stemmed from L2 English. It can also be concluded that L1 drift in production was manifested in both voiced and voiceless stops but was greater in the later series. However, there was no perceptual L1 drift.

Lack of changes in the speakers' L1 perceptual boundary may be attributed to the fact that, unlike previous studies on perceptual drift in VOT [8,14], the current speakers resided in the L1 country before and at the moment of data gathering. Thus, they may not have had enough exposure to the L2 to trigger changes in L1 perception.

The results give further evidence of L1 drift in the form of prolonging L1 VOT in short-lag stops under the influence of L2 English long-lag stops as in [10,13]. However, they do not substantiate the asymmetry between voiced and voiceless series found in [11,12]. The differences in the results could be task-related as both [11,12] used a variation of an imitation protocol as opposed to words in isolation that are not previously heard or read as part of the study procedure.

Finally, the results of the regression revealed that neither fortis nor lenis realisations were statistically significant predictors of the perceptual boundary between voiced and voiceless stops. Although there was a tendency to produce aspirated voiceless fortis stops and voiceless lenis stops in L1 Polish, the category boundary was not placed at higher (positive) VOT values. Since our data exhibit drift in production but not in perception, we conclude that there is a mismatch between the two modalities. This partly confirms [7] whose results indicate no perception-production link apart from one group who showed L1 drift in production but not in perception. This result contradicts theories of speech perception in which perception precedes production, and accurate production is a consequence of correct perception.

All in all, we conclude that, in L1 drift situations, there may be a perception-production mismatch that could define L1 drift as an element of execution without influencing phonological representations.

6. REFERENCES

[1] Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Stockmann, E., Tiede, M. 2004. The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *Journal of. the Acoustical Society of America* 116, 2338–2344.

[2] Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., Yamada, T. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32, 233-250.

[3] Baker, W., Trofimovich, P. 2006. Perceptual paths to accurate production of L2 vowels: The role of individual differences. *International Review of Applied Linguistics* 44, 231-250.

[3] Best, C., Tyler, M. 2007. Nonnative and second-language speech perception. In: Bohn, O.-S., Munro, M. J. (eds), *Language experience in second language speech learning*. John Benjamins, 13–34.

[4] Flege, J. E. 1995. Second language speech learning: Theory, findings, and problems. In:. Strange, W. (ed), *Speech perception and linguistic experience: Issues in cross-linguistic research*. York Press, 233-277.

[5] Best, C., Tyler, M. 2007. Nonnative and second-language speech perception. In: Bohn, O.-S., Munro, M. J. (eds), *Language experience in second language speech learning*. John Benjamins, 13–34.

[6] Chan, A. Y. W. 2014. The perception and production of English speech sounds by Cantonese ESL learners in Hong Kong. *Linguistics* 52, 35-72.

[7] Darcy, I., Kruger, F. 2012. Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics* 40, 568-581.

[8] Tice, M., Woodley, M. 2012. Paguettes & bastries: Novice French learners show shifts in native phoneme boundaries. *LSA Annual Meeting Extended Abstracts* 3, 1-5.

[9] Osborne, D.M. 2016. "The acquisition of fine phonetic detail in a foreign language: perception and production of stops in L2 English and L1 Portuguese." Ph.D. dissertation, University of Arizona.

[10] Harada, T. 2003. L2 influence on L1 speech in the production of VOT. *Proc. 15th ICPHS* Barcelona, 1085-1088.

[11] Schwartz, G., Dzierla, J., Wojtkowiak, E. 2019. Laryngeal phonology and asymmetrical cross-language phonetic influence. In: Wrembel, M., Kiełkiewicz-Janowiak, A., Gąsiorowski, P. (eds), *New Approaches to the study of sound structure and speech: Interdisciplinary Work in Honour of Katarzyna Dziubalska-Kołaczyk*. Routledge, 316-325.

[12] Podlipský, V.J., Šimáčková, Š., Chládková, K. 2020. Phonetic drift reveals interconnected phonological representations in simultaneous bilinguals: A case study of English and Czech stop consonants. *International Journal of Bilingualism* 25, 483-832.

[13] Sypiańska, J. 2021. Production of voice onset time (VOT) by senior Polish learners of English. *Open Linguistics* 7, 316-330.

[14] Gorba, C. 2019. Bidirectional influence on L1 Spanish and L2 English stop perception: the role of L2 experience. *The Journal of the Acoustical Society of America* 145, 1-7.

[15] Lemhöfer, K., Broersma, M. 2012. Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods* 44, 325-343.

[16] Psychology Software Tools, Inc. [E-Prime 3.0]. 2016. Retrieved from https://support.pstnet.com/.

[17] IBM Corp. Released 2021. IBM SPSS Statistics for Macintosh, Version 28.0. Armonk, NY: IBM Corp.