PERCEPTION OF VOICING DISTINCTION IN SYLLABLE-INITIAL STOPS BY MULTILINGUAL SPEAKERS

Zuzanna Cal and Magdalena Wrembel

Adam Mickiewicz University, Poznań, Poland
zuzanna.cal@amu.edu.pl, magdala@amu.edu.pl

ABSTRACT

The objective of the current study is to explore how multilingual learners categorise fortis and lenis stops in the three languages they know. In particular, we aim to investigate whether patterns of categorisation of VOT continua are specific to a language and place of articulation (PoA), as well as whether trilinguals experience boundary shifts in any of the languages and PoA. The study was conducted on L1 Polish-L2 English-L3 Norwegian speakers (n=22, aged 20), who had just started studying Norwegian as part of intensive formal training. The experiment included preparation of three continua (/b-p/, /d-t/, /g-k/) in Polish, English and Norwegian and was run in PsychoPy. The participants listened to the words from the continua and decided whether they hear a voiced or voiceless sound at the beginning of each word. The results point to rather unique patterns of VOT categorisation across languages and PoA with possible L1-L2 and L1-L3 interference.

Keywords: speech perception, L3 phonology, multilingualism, VOT, third language acquisition

1. INTRODUCTION

Research on multilingual perception of speech has gained increasing interest in the recent years. Extensive investigations [1, 2, 3, 4, 5] have pointed to a potential multilingual advantage according to which prior linguistic experience has facilitative effects on the perception of novel L3 contrasts. What is more, recent research also pointed out to rather specific trends. For example, [6] found combined effect of L1, markedness and L2/L3 proficiency in multilingual speech perception of rhotics by children. Furthermore, in the study on perception of sibilant pairs, [7] showed that multilinguals assimilate some L3 sounds to both L1 and L2 categories, with a preference for the latter and that beginner L3 learners are likely to perceive subtle acoustic differences in novel phonological contrasts.

As far as VOT perception is concerned, a considerable amount of literature has been published over the last few decades [8, 9, 10, 11, 12, 13, 14]. In these studies, which assume, however, the perspective of second language acquisition, VOT has been shown to be susceptible to the influence of a number of factors, including the age of onset, L1/L2 status or individual differences. Recent research attested different patterns of perception while comparing monolinguals and bilinguals. For example, [13] found that monolingual speakers showed a consistent pattern while in bilinguals, the two languages were perceived in significantly different ways.

Although VOT perception in L2 acquisition has been widely researched, to the best of our knowledge, only a few studies have attempted to explore VOT perception from the multilingual perspective. [15] investigated the effect of language learning on the perception of VOT in L2/L3 and L1. The study was designed to tap into the influence of a newly acquired language on the previously known ones. The task involved identification of a /bi-pi/ continuum by two groups (bi- and trilingual) of language learners that differed in learning settings and experience. The results pointed to both regressive and progressive cross-linguistic influence. Another study that tackled VOT perception in multilingual learners [16] aimed to investigate the perception and production of word-initial stops by L1 Mandarin Chinese L2 English speakers learning Japanese or Russian as L3. The participants performed an identification task of /p, t, k, b, d, g/ in word-initial position of monosyllable or disyllabic words. The results pointed to the role of phonetic similarity in different stop categories between L1, L2, and L3 that might have contributed to learners' confusion in perception. What is more, higher accuracy scores were found in the perception of voiced stops than in voiceless ones in L3 Japanese/Russian. As voiced stops in the participants L3s differ from those in their L1, they might have formed a new category, and therefore, be more easily perceived by the learners, which is in line with Flege's Speech Learning Model [17].

All in all, research on the subject, though quite prominent, has been mostly restricted to limited comparisons of languages, places of articulation and types of speakers. This paper attempts to expand the existing body of literature by examining perception of word-initial stops by trilingual speakers of L1 Polish L2 English and L3 Norwegian. These three languages seem to be complementary with respect to the features of their stop consonants. Firstly, Polish is a true voicing language that has prevoicing in the voiced series of stops and short-lag VOT in voiceless stops [8]. English, on the other hand, is an aspirating language in which /b, d, g/ are partially voiced and /p, t, k/ are aspirated [18]. When it comes to Norwegian, the voiced stops are prevoiced, similarly to Polish ones, but its voiceless stops are aspirated and resemble those in English [19, 20].

2. METHOD

2.1. Aims and predictions

The aim of this study is to investigate how L1 Polish - L2 English - L3 Norwegian speakers categorise fortis and lenis stops in the three languages. The following research questions were asked in the study: (1) What are the patterns of VOT categorisation in multilinguals? Are they language- and PoA-specific? (2) What are the perceptual boundary locations for the perception of voiced and voiceless stops in all three languages? Do they point to potential sources of CLI?

There are two predictions regarding the outcomes of the study. In relation to the first RQ, it is hypothesised that multilingual advantage might trigger more language- and PoA- specific patterns of VOT categorisation. The second RQ considers perceptual boundary locations. We foresee two possible scenarios depending on the language status and/or linguistic similarities. First of all, due to existing similarities in the process of learning foreign languages [21], as well as intrinsic phonetic characteristics, i.e., aspiration being present in the voiceless series of stops in L2 English and L3 Norwegian, we may hypothesise that the participants experience boundary shifts towards the English values. In other words, the boundaries in L3 Norwegian might be found to be later in the continuum and yield similar values to those of L2 English, particularly for the fortis series. On the other hand, at the initial stages of L3 acquisition, and due to the existing similarities between Polish and Norwegian voiced stops, the participants might also rely heavily on their L1. In this sense, we hypothesise that the perceptual boundaries in L3 Norwegian might be shifted towards those of L1 Polish and appear earlier in the continua.

2.2. Participants

The participants included 19 L1 Polish speakers, who had been learning English as their L2 and had recently started learning L3 Norwegian extensively by attending Norwegian Studies programme at two Polish universities. There were 14 females and 5 males, whose mean age was 20 years old. Prior to the experiment the participants had had limited, almost none, contact with Norwegian. Their mean age of onset in English and Norwegian was 6.3 and 19.6 years respectively.

They were asked to fill in the Language History Questionnaire (LHQ) [22] in order to obtain information about their language learning history and use, including the age of onset, declared exposure to languages, length and type of instruction etc. The participants also completed proficiency placement tests in both of their foreign languages – English and Norwegian. Based on the grading criteria they were classified as intermediate (approximately B2 level) learners of L2 English and beginner (A1) learners of L3 Norwegian. Knowledge of any other foreign languages has been reported as minute and the participants claimed not to have used any of them for an extensive period of time.

2.3. Procedure and stimuli

The study design involved the preparation of VOT continua with the use of a Praat script [23] based on one-syllable minimal pairs with word initial stop sounds /p,b,t,d,k,g/. The token words included *pas-bas*, *tam-dam*, *kas-gaz* for Polish, *pen-ben*, *tan-Dan*, *cast-gas* for English and *par-bar*, *ta-da*, *kar-gard* for Norwegian and were recorded in carrier phrases by native speakers of the respective languages. Separate continua were created for each place of articulation and for each language. Each token from the continua differed from the next one by 10 ms. Table 1 shows specific values and the number of steps in each of the continua. The ranges were based on the values obtained from native speakers' recordings in all three languages.

	Polish	English	Norwegian
b-p	-90 - 30 ms (13 steps)	0 - 70 ms (8 steps)	-140 - 80 ms (23 steps)
d-t	-130 - 20 ms (16 steps)	0 - 90 ms (10 steps)	-130 - 90 ms (23 steps)
g-k	-80 - 60 ms (15 steps)	0 - 70 ms (8 steps)	-140 - 90 ms (24 steps)

Table 1: Ranges and number of steps of VOT continua (in ms) for the three languages.

The experiment was programmed in PsychoPy [24] and conducted in three separate language blocks on different days. In each session, the participants were presented with audio stimuli and a visual representation of two sounds respective to the continuum from which the sound was played. The task was to identify which of the presented sounds appeared at the beginning of each word by clicking an appropriate key on a computer keyboard. Every token was presented 3 times. The experiment was conducted in a quiet room and the participants were instructed to wear headphones. Each language block lasted around 7-10 minutes.

3. RESULTS

The obtained data included accuracy rates and reaction times (RT). First of all, the Pearson correlation was run to compare accuracy scores (%) for each step of the continuum with RT (ms). The statistical analysis was run in SPSS [25]. Moderate and significant negative correlation coefficients were obtained for all three continua in L1 Polish (/b-p/: r=-0.44, p=.035; /d-t/: r=-0.54, p=.008, /g-k/: r=-0.55, p=.005), see Table 2. Strong and significant negative correlation was found in the /b-p/ continuum in L2 English (r=-0.81, p=.016) and the /d-t/ continuum in L3 Norwegian (r=-0.72, p=.002). Table 2 presents correlation coefficients for all continua. All r-values turned out to be negative, except for the /g-k/ continuum in English, which yielded quite weak and positive correlation. All in all, the values show that the lower the accuracy scores, the higher RT, which means that the participants took longer to react when they were less confident about the answer.

Language	Continuum	r	p
Polish	b-p	-0.44	.035
	d-t	-0.54	.008
	g-k	-0.55	.005
English	b-p	-0.81	.016
	d-t	-0.38	.285
	g-k	0.23	.578
Norwegian	b-p	-0.54	.056
	d-t	-0.72	.002
	g-k	-0.47	.080

Table 2: The Pearson correlation coefficients for each continuum and language.

Figures 1-3 show accuracy rates for the identification of voiced plosive sounds in all three continua in the three languages. What can be observed in these charts is that there are some discrepancies across L1/L2/L3 languages, especially visible in the /b-p/ continuum, but also slightly in the /d-t/. However, in the /g-k/ continuum, the participants were more consistent in their answers across the languages.

To obtain the perceptual boundary location, the data was converted using a logistic regression and the boundary location was calculated with the use of the following formula: *-LN(b0)/LN(b1)*, where *b0* corresponded to the constant and *b1* to the obtained slope of the function. Table 3 provides mean perceptual boundaries and standard deviations for each continuum and language.

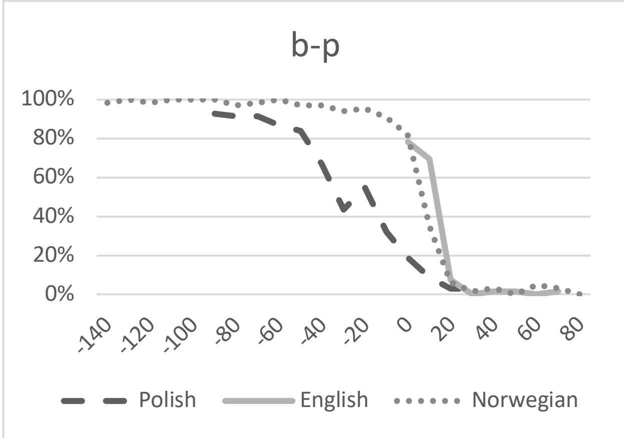

Figure 1: Accuracy rates for the identification of /b/ in L1 Polish, L2 English and L3 Norwegian.

Figure 2: Accuracy rates for the identification of /d/ in L1 Polish, L2 English and L3 Norwegian.

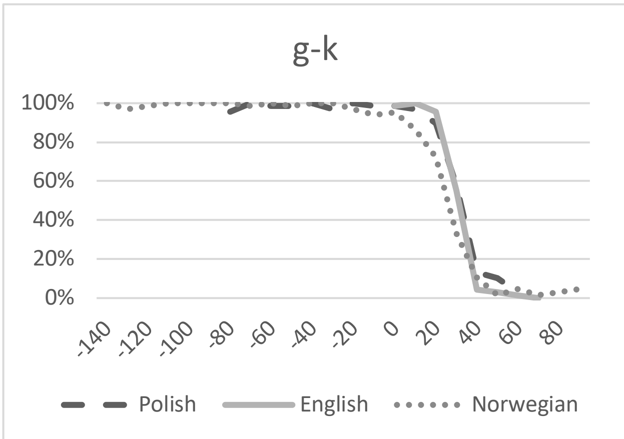

Figure 3: Accuracy rates for the identification of /g/ in L1 Polish, L2 English and L3 Norwegian.

The mean perceptual boundary location for the /b-p/ continuum was the earliest in Polish (-25.42 ms), followed by Norwegian (-2.13 ms) and English (14.65 ms). In /d-t/ continuum, the mean perceptual boundary turned out to be the lowest in Norwegian (10.11 ms), then Polish (14.05 ms) and, finally, English (35.97 ms). As far as /g-k/ continuum is concerned, the earliest perceptual boundary was found in Norwegian (18.77 ms), followed by Polish (29.87 ms) and English (30.82 ms), which yielded similar values.

Language	Continuum	Mean	Std. Deviation
Polish	b-p	-25.42	18.89
	d-t	14.05	15.82
	g-k	29.87	14.00
English	b-p	14.65	8.05
	d-t	35.97	6.52
	g-k	30.82	3.93
Norwegian	b-p	-2.13	8.20
	d-t	10.11	9.01
	g-k	18.77	11.54

Table 3: Mean perceptual boundary (in ms) in each continuum and language.

A Linear Mixed Model was run to investigate the differences in perceptual boundary locations with language (Polish, English, Norwegian), place of articulation (PoA; labial, coronal, velar), as well as interaction between language and PoA as fixed factors, and participant as a random effect. The results showed significant main effects of Language (F=43.878, p<.001), PoA (F=108.036, p<.001) and their interaction (F=18.822, p<.001).

Pair-wise comparisons between Languages and PoA pointed to the following results. In the /b-p/ continuum, statistically significant differences were found between all three languages, that is between L1 Polish and L2 English (t=-9.279, p<.001), L1 Polish and L3 Norwegian (t=-6.653, p<.001), as well as between L2 English and L3 Norwegian (t=3.800, p<.001). With regards to the /d-t/ continuum, statistically significant differences were shown between L1 Polish and L2 English (t=-2.954, p=.004), as well as L2 English and L3 Norwegian (t=7.613, p<.001). As far as the /g-k/ continuum is concerned, statistically significant differences were found between L1 Polish and L3 Norwegian (t=3.080, p=.003), but also between L2 English and L3 Norwegian (t=3.546, p=.001).

4. DISCUSSION AND CONCLUSION

The aim of this paper was to investigate VOT perception by multilingual learners. We intended to address research questions as to whether patterns of categorisation of VOT continua are specific to a language and PoA, as well as whether trilinguals experience boundary shifts in any of their languages and PoA.

The first prediction set out in this study was that multilingual advantage would trigger more language- and PoA-specific patterns of VOT categorisation. This prediction was, to a great extent, confirmed by our data. Statistical analysis pointed to significant differences in most of the cases. That is, perceptual boundaries in /b-p/ continuum turned out to be statistically different in all three languages. In /d-t/ continuum, the differences appeared between L1 and L2, as well as L2 and L3. Finally, in /g-k/ continuum significant differences were observed between L1 and L3, as well as L2 and L3. These obtained differences suggest that multilingual learners generally maintain language- and PoA- specific patterns of VOT categorisation. This finding might be an indication of a multilingual advantage, according to which, multilingual learners are more prone to perceive subtle linguistic contrasts [1, 2, 3, 4, 5].

The second prediction dealt with perceptual boundary locations. Two possible scenarios were hypothesised depending on the language status (L1/L2/L3) and inherent phonetic similarities. Firstly, it was predicted that the boundaries of L2 English and L3 Norwegian would be closer due to the existing similarities between voiceless stops, but also due to the process of foreign language learning [21]. On the other hand, it was predicted that the participants might also rely heavily on their L1 while learning an additional language [26], which, in turn, would contribute to L1 and L3 boundaries being closer together. We found confirmation of the second scenario only in /d-t/ continuum, where no statistically significant difference was found between L1 Polish and L3 Norwegian. This suggests possible interdependence between the two languages. What was also found is no statistically significant difference between L1 Polish and L2 English in /g-k/ continuum, which suggests possible, unexpected, interactions between the L1 and L2. No other traces of CLI were attested in the data. Again, pointing to the possible role of multilingual advantage and confirming the findings of [13].

Recapitulating, this paper explored categorisation of VOT continuum by multilingual learners of L1 Polish, L2 English, L3 Norwegian. It was found that the learners maintain patterns of VOT categorisation specific to a language and place of articulation. Possible, though limited interactions were attested between L1-L2 and L1-L3, in one of the three places of articulation, which suggests plausible, though rather weak L1 interference in the perception of VOT in both the L2 and L3. Ongoing comparisons with a longitudinal data collection design and a control group performance will allow for a more in-depth analyses and further conclusions.

5. ACKNOWLEDGEMENT

This research is supported by a project grant of the Polish National Science Centre (NCN) OPUS-19-HS (UMO-2020/37/B/HS2/00617) CLIMAD "Cross-linguistic influence in multilingualism across domains: Phonology and syntax".

6. REFERENCES

[1] Enomoto, K. 1994. L2 perceptual acquisition: The effect of multilingual linguistic experience on the perception of a "less novel" contrast. Edinburgh *Working Papers in Applied Linguistics* 5, 15–29.

[2] Tremblay, M.-C., Sabourin, L. 2012. Comparing behavioral discrimination and learning abilities in monolinguals, bilinguals and multilinguals. *J. Acoust. Soc. Am*. 132, 3465–3474.

[3] Antoniou, M., Liang, E., Ettlinger, M., Wong, P. 2015. The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition* 18, 683–695.

[4] Kopečková, R. 2015. Differences in the perception of English vowel sounds by child L2 and L3 learners. In: U. Gut, R. Fuchs, E.-M. Wunder (eds.), *Universal or diverse paths to English phonology*. Mouton de Gruyter, 71–90.

[5] Onishi, H. 2016. The effect of L2 experience on L3 perception. *Int. J. Multilingual*. 13, 459–475.

[6] Balas, A., Kopečková, R., Wrembel, M., & Warren, P. (2019). Perception of rhotics by Multilingual Children. *Proc. 19th ICPhS* Melbourne, 3725–3729.

[7] Wrembel, M., Marecka, M., & Kopečková, R. 2019. Extending perceptual assimilation model to L3 phonological acquisition. *Int. J. Multilingual.* 16, 513–533.

[8] Keating, P., Mikoś, M., Ganong III, W. 1981. A cross-language study of range of VOT in the perception of stop consonant voicing, *J. Acoust. Soc. Am*. 70, 1260–1271.

[9] Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *J. Phonetics* 15, 47–65.

[10] Rojczyk, A. 2010. *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*. Wydawnictwo Uniwersytetu Śląskiego.

[11] Stölten, K. 2013. "The Effects of Age of Onset on VOT in L2 Acquisition and L1 Attrition: A Study of the Speech Production and Perception of Advanced Spanish-Swedish Bilinguals," Ph.D. dissertation, Stockholm University.

[12] Kong, E. 2019. Individual differences in categorical perception: L1 English learners' L2 perception of Korean stops. *Phonetics and Speech Sciences* 11, 63–70.

[13] García-Sierra, A., Schifano, E., Duncan, G. M., Fish, M. S. 2021. An analysis of the perception of stop consonants in bilinguals and monolinguals in different phonetic contexts: A range-based language cueing approach. *Attention, perception & psychophysics* 83, 1878–1896.

[14] Pan, L., Ke, H., Styles, S. J. 2022. Early linguistic experience shapes bilingual adults' hearing for phonemes in both languages. *Scientific Reports* 12.

[15] Liu, Z., Gorba, C., Cebrian, J. 2019. Effects of learning an additional language on VOT perception. *Proc. 19th ICPhS Melbourne*, 260–254.

[16] Liu, J., & Lin, J. 2021. A cross-linguistic study of L3 phonological acquisition of stop contrasts. *SAGE Open* 11.

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

[18] Lisker, L., Abramson A. 1964. A cross-language study of voicing in initial stops, *Word* 20, 384–422.

[19] Ringen, C., Dommelen, W.A. 2013. Quantity and laryngeal contrasts in Norwegian. *J. Phonetics* 41, 479–490.

[20] Czarnecki, P. 2016. *The Phonology of Quantity in Icelandic and Norwegian*. Wydział Neofilologii UAM w Poznaniu.

[21] Bardel, C., & Falk, Y. 2007. The role of the second language in third language acquisition: the case of Germanic syntax. *Second Language Research* 23, 459–484.

[22] Li, P., Zhang, F., Yu, A., Zhao, X. 2020. Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition* 23, 938–944.

[23] Winn, M. B. 2020. Praat script.

[24] Peirce, J., Gray, J.R., Simpson, S., MacAskill, M., Höchenberger, Sogo, H., Kastman, E., Lindeløv, J.K. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51, 195–203.

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

[26] Hermas, A. 2010. Language acquisition as computational resetting: verb movement in L3 initial state. *Int. J. Multilingual* 7, 343–362.