IMMEDIATE PHONETIC INTERFERENCE IN CODE-SWITCHING AND INTERPRETING

Šárka Šimáčková and Václav Jonáš Podlipský

Department of English and American Studies, Faculty of Arts, Palacký University Olomouc, Czech Republic
sarka.simackova@upol.cz, vaclav.j.podlipsky@upol.cz

ABSTRACT

VOT of English voiceless stops produced by L1-dominant Czech-English bilinguals is examined in light of three hypotheses: (1) switching languages induces an immediate increase in L1 interference, (2) speakers experienced with switching languages show less immediate interference, and (3) interpreting induces greater interference than code-switching. Eighteen bilinguals, trained vs beginning interpreters, produced short sentences in a modified delayed repetition / translation task under three conditions: L2-only, code-switching into L2, and interpreting into L2. An effect of short-term L1 interference was observed: the speakers produced shorter, i.e. more L1-like, VOTs in code-switching and interpreting than in the L2-only condition, though the VOT reduction interacted with the place of articulation differently for speakers with more and less native-like VOTs. The effect was weaker for trained interpreters, giving some support to the second hypothesis. The type of the bilingual task (code-switching vs interpreting) did not affect VOT.

Keywords: L1-L2 interaction, L1 interference, code-switching, interpreting, VOT production.

1. INTRODUCTION

The extent of L1 phonetic interference in a learner's L2 is affected by the amount L1 use [9, 16, 26, 28]. Recent studies have examined immediate L1 use as a source of short-term interference in the context of bilingual code-switching (CS) [1, 2, 5, 11, 29]. These studies differentiate between long-term interference on the one hand, i.e. differences in competence between a monolingual and a bilingual, and interference as a performance phenomenon [12] which takes place in real time and can momentarily increase. This short-term effect arises when a bilingual speaker performs in the bilingual language mode [14], communicating with (or listening to) other bilinguals and alternating between languages.

One issue pursued in studies of both long-term and short-term interlanguage interactions is their direction. It has been well established that L1-to-L2 is not the only pattern of influence, see review in [1]. The direction and magnitude of interlanguage interaction depends on the pattern of L1 and L2 acquisition and use, and the resulting language dominance in a bilingual. In studies of immediate interlanguage interactions due to CS (which almost exclusively measure VOT), a variety of interaction patterns have emerged. Grosjean and Miller [15] reported no effect of CS for French-English bilinguals. Bullock et al. [5], found a unidirectional L1-to-L2 influence for Spanish learners of English and L2-to-L1 influence for English learners of Spanish, in both cases before or at the site of the switch. In Simonet [29] early Catalan-Spanish bilinguals showed effects of CS (on the vowel height of two Catalan back vowels), whether they were Spanish-dominant or Catalan-dominant. For English intermediate learners of Spanish, González-López [11] observed both L1-to-L2 and L2-to-L1 effects on positive VOT, albeit patterning differently across *p*, *t*, and *k*. In Antoniou et al. [1], early L2-dominant Greek-English bilinguals showed a unidirectional influence of L1 on L2. Finally, Balukas and Koops [2], examining spontaneous CS, found influence of L1 Spanish on L2 English, but not vice versa.

For comparability with previous research, the present study also selects VOT of voiceless stops as a measure of immediate L1 interference. We assess the influence of the bilinguals' L1 Czech, where voiceless stops have short positive VOT [22], on their L2 English. The participants in our study are clearly L1-dominant but highly L2-proficient bilinguals who learned their L2 in foreign-language (FL) settings. Although achieving phonetic accuracy in L2, even at high levels of overall proficiency, may seem difficult for such FL learners, they are not completely constrained by their L1 phonology, see e.g. [23]. Therefore, we expect Czech learners to produce longer VOTs in L2 English voiceless stops than it is typical for Czech stops. We further predict that VOTs in their monolingual-mode English utterances will be longer than VOTs in Czech-to-English code-switched utterances.

As pointed out in [2], most studies of immediate interlanguage interactions do not address participants' experience with CS, although bilinguals may differ widely in this respect [24]. Bilinguals who report frequent switching between languages in daily life have been found to switch faster both between languages and between non-linguistic tasks, showing an advantage in executive control [27]. Differences between bilinguals in executive control, more specifically in their inhibitory skills, can predict how well their long-term L1 and L2 phonological representations are separated [19, 20]. Our question is whether inhibitory skills enhanced by extensive experience with switching languages can lead to reduced immediate L1-to-L2 interference. We compare utterances produced by trained and beginning students of interpreting. We expect less L1 interference in bilingual-mode L2 productions by the trained interpreters than by the beginners.

Interpreting can be construed as a bilingual-mode performance requiring a specific kind of inhibitory skill [13]. Compared to CS, switching between languages during (even consecutive) interpreting involves increased cognitive demands. As a result of planning L2 production based on a memory of an utterance received in L1, inhibiting the L1 may be more difficult than in regular CS, at least for beginning interpreters. In addition, the cognitive strain is likely to interfere with interpreters' self-monitoring. Thus, we predict more immediate L1 interference during interpreting than regular CS.

2. METHOD

This study compared the VOT of voiceless stops in L2 productions of 12 English target words under 3 conditions: monolingual-mode English condition (EN), code-switching (CS), and interpreting (IN).

2.1. Participants

Twenty Czech female college students of English interpreting, aged 19-27, were recruited: 10 from the 1st year (Novices) and 10 from the 3rd year (Experts). All were advanced FL learners who had achieved level C1 proficiency in English according to CEFR [30]. The Experts had completed 5 interpreting courses, the Novices none yet.

2.2. Stimuli and procedure

In all three conditions, the target words were 6 *p*-initial and 6 *t*-initial monosyllabic English words, with a non-high vowel after the stop. In the EN and CS conditions, each target word was heard in two short English sentences (5-8 syllables), once initial and once as the final word. In the IN condition, the closest Czech equivalent of each target word was placed in two Czech sentences. Twenty-four sentences with the target words (2 places of articulation × 6 target words × 2 sentence positions) and 64 filler sentences were created per condition. While the target words were identical in the EN and CS conditions, the 88 sentences were different. In the IN condition, one half were translations of sentences from the EN condition and the other from the CS condition. Five native Czech speakers (3 females), aged 21-25, recorded the Czech sentences. The English sentences were recorded by 2 British and 3 American English speakers (2 females overall), aged 24-54, in the quote frame *I should say_*. Each speaker produced all (Czech or English) sentences, but only one realization of each sentence by different speakers was selected for the final stimulus set. The prompting questions (see below) were recorded by the same speakers. The English speakers' 24 EN-condition sentences were measured to obtain reference data.

We used a delayed repetition (EN, CS) and translation (IN) task. In EN, participants heard an English sentence (e.g. *We like the new pub.*) followed by the English prompt *What should you say?* said by a different person. The participant's response used the quote frame *I should say,_* (e.g. *I should say, "We like the new pub."*). In CS, an English sentence was followed by the Czech prompt *Co jsi slyšel?* 'What did you hear?'. Participants responded by using the quote frame *Slyšel jsem_* ('I heard') and then switching back into English (e.g. *Slyšel jsem, "Let's go to a pub."*). In the IN condition, a Czech sentence was followed by the prompt *Co teď řekneš?* ('What will you say?') and the participant's response was *Teď řeknu_* ('Now I will say') followed by a translation of what they had heard into English. In all conditions, the 88 sentences were randomized uniquely for each participant.

Data were collected individually in a sound booth in two sessions separated by at least 48 hours. The monolingual EN session was conducted exclusively in English by a Czech experimenter with a native-like English accent. The bilingual session consisted of the CS and IN conditions (order counterbalanced) separated by a 5-minute break. One of three research assistants conducted the session in Czech. The order of sessions was determined by participant availability (10 took the EN session first). Stimuli were presented via Sennheiser HD 280 pro headphones. Recordings were made using a Zoom H4n digital recorder with a 44.1 kHz sampling rate and 16 bit quantization.

2.3. Measurements

The first author and a phonetically trained research assistant used Praat [4] to manually label boundaries of sentences, target words, and positive VOT intervals using constant criteria, based on [21].

A few target words could not be analysed (4 in EN, 6 in CS, 6 in IN in total) because they were not correctly reproduced or did not appear in the participants' translations. Sentences with obvious disfluencies, as evaluated perceptually by the annotators, were excluded from computing the mean sentence duration (6 in EN, 20 in CS, and 17 in IN).

Studies of interference often use absolute VOT durations, e.g. [5, 11, 15]. This is problematic if a constant speaking tempo cannot be guaranteed across experimental conditions since VOT changes with speaking tempo [17]. Equally important are tempo changes within utterances. Local tempo is not controlled for in [5] where, compared to the monolingual production, English positive VOT was shorter before and at the site of the switch compared to the post-switch position. However, the post-switch position coincided with the end of an utterance, a site of final lengthening [25] and of pitch accent placement also associated with lengthening [18]. In order to control for speaking tempo we introduced a tempo-normalized VOT measure calculated as the ratio of VOT duration to the duration of the given word.

3. RESULTS

3.1. Bilinguals vs native speakers

Productions of *p* and *t* from the EN condition constituted baseline data. An inspection of participants' mean VOTs revealed a bimodal distribution, with the right peak (long VOTs) overlapping completely with the distribution of native speakers' data. Thus, an a posteriori grouping of the bilinguals was adopted: mean VOTs falling below 2 SDs [3, 10] of the native mean (60.8 ms, 1 SD = 3.6 ms) were considered short positive VOT. This divided the participants into 2 groups: LongLag ($n = 9$) and ShortLag ($n = 9$). The 2 bilingual groups and the native speakers were compared in terms of sentence durations, target word durations, and VOT. Means and standard deviations are summarized in Table 1.

A one-way ANOVA found no significant effect of Speaker group (LongLag, ShortLag, and Native English) on mean sentence durations ($p > .2$). Mean target word durations were submitted to a repeated-measure (RM) ANOVA with Speaker group as a between- and Position of target word (Sentence-initial and final) as a within-subject factor. Only Position yielded a significant effect ($F(1, 20) = 472.39$, $p < .001$), with words longer sentence-finally than initially. Taken together, the sentence and target-word duration results suggest that bilinguals did not speak more slowly than the native speakers (nor did the bilingual groups differ from each other). Since we measure tempo-normalized VOT, Position can be excluded as a factor from the subsequent analysis of VOT.

Normalized VOTs were submitted to a RM ANOVA with Speaker group as the between- and Place of articulation (*p, t*) as the within-subject factor. Unsurprisingly, Speaker group had a significant effect ($F[2, 20] = 32.734$, $p < .001$). As expected, Place of articulation ($F[1, 20] = 40.238$, $p < .001$) also affected VOT (shorter for *p* than for *t*). A significant Speaker group × Place interaction was also found ($F[2, 20] = 16.350$, $p < .001$): a post-hoc Tukey HSD test showed that the gap between *p* and *t* was bigger for ShortLag than for the other 2 groups.

Table 1: Mean sentence, target word, and VOT durations (in ms) SD are in parentheses.

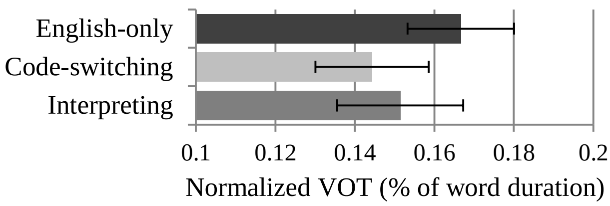
speaker group	sentence dur.	word dur.		VOT		
		initial	final	*p*	*t*	mean
ShortLag	**1279** (138)	**222** (34)	**345** (34)	**20** (10)	**43** (17)	**32** (12)
LongLag	**1337** (124)	**248** (36)	**378** (43)	**58** (13)	**71** (14)	**64** (13)
Native	**1196** (157)	**233** (36)	**362** (45)	**53** (6)	**68** (4)	**61** (4)

3.2. Bilinguals' productions across conditions

We first assessed changes in tempo at the sentence and target word level. A RM ANOVA with the between-subject factors Speaker group (LongLag, ShortLag) and Experience with interpreting (Expert, Novice) and the within-subject factor Condition (EN, CS, IN) was run on mean sentence durations. None of the 3 factors had a significant effect ($p > .25$). Thus, the participants did not change sentence durations depending on the condition. (Recall that disfluent sentences were excluded.) Mean word durations were submitted to another RM ANOVA, with the additional factor of Position of word (Initial, Final). Only the effects of Condition ($F[2, 28] = 6.20$, $p = .01$) and Position ($F[1, 14] = 153.92$, $p < .001$) and their interaction ($F(2, 28) = 5.51$, $p < .01$) were significant. A post-Hoc Tukey test revealed that during interpreting the final target words were slower than in the other 2 conditions ($p < .05$).

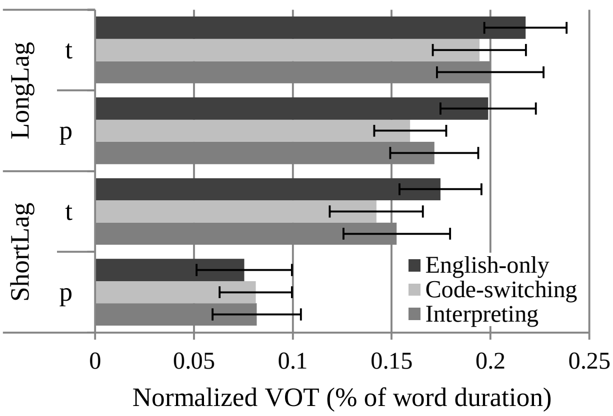
Turning to VOT, a 2 (Speaker group) × 2 (Experience with interpreting) × 3 (Condition) × 2 (Place of articul.) RM ANOVA was run on the mean normalized VOTs. Significant main effects of Speaker group ($F[1, 14] = 42.386$, $p < .001$) and Place ($F[1, 14] = 114.282$, $p < .001$) were again found (this time with data from all conditions). Importantly, as is apparent from Fig. 1., Condition also produced a significant effect ($F[2, 28] = 5.773$, $p < .01$). A post-hoc Tukey test found no difference between the CS and IN conditions ($p > .45$); VOTs were significantly higher in EN than in CS, and the EN vs IN difference approached significance (p = .061). There was no significant interaction between Experience with interpreting and Condition.

Figure 1: Bilinguals' mean normalized VOT by condition. Error bars are .95 confidence intervals.



There was also a significant interaction of Condition, Place and Speaker group ($F[2, 28] = 6.685$, $p < .01$). As seen in Fig. 2, in the ShortLag group, the VOTs of *p* were comparable across conditions and those of *t* were longer in EN than in CS ($p < .01$), while in the LongLag group VOTs were longer in EN than in CS ($p < .001$) and IN ($p < .05$) only for *p*, and *t* VOTs did not differ across conditions ($p > .07$).

Figure 2: Bilinguals' mean normalized VOT by Condition, Speaker group, and Place of articulation. Error bars are .95 confidence intervals.



To explore the factor of Experience with interpreting further, when a new RM ANOVA was run with Experts' data only, there was no significant effect of Condition ($p > .14$), whereas a separate RM ANOVA for Novices did find an effect of Condition ($F[2, 16] = 4.228$, $p = .0336$) with the expected decrease of VOT in the bilingual condition.

4. DISCUSSION

Considering long-term interference first, our results suggest that some late L1-dominant but advanced learners can learn to produce L2 voiceless stops in a native-like way, despite numerous reports of late L2 learners typically producing intermediate VOT, e.g. [6–8]. When speaking in the monolingual English mode, half of our 18 L1-Czech L2-English bilinguals (whom we labelled LongLag) produced average VOT values not below our reference native speakers' range. The other half (ShortLag) produced *p* with VOT typical for Czech, a short-lag language [22], and *t* with VOT values somewhat longer than in Czech. This difference in learning to aspirate *p* and *t* may be perceptually driven: the longer and higher-frequencies-involving aspiration noise of native English *t* may be more noticeable to learners and hence easier to acquire.

The main goal of this study was to test whether interference from L1 becomes temporarily stronger when late L1-dominant bilinguals switch from L1 into L2 than when they simply speak in L2. Our three production tasks required the bilinguals either (i) to report in L2 what they had just heard in L2, or (ii) to start an utterance in L1 and switch into L2 to report what they had heard in L2, or (iii) to start an utterance in L1 and switch into L2 to report what they had heard in L1. Overall results showed that bilingual productions of voiceless stops in the two latter, bilingual, tasks were more L1-like compared to the baseline productions in the L2-only task. Thus, short-term L1-to-L2 interference induced by performing in the bilingual mode was indeed found. This is in line with [1] who demonstrated a unidirectional L1-to-L2 influence for early L2-dominant bilinguals.

Another goal of our study was to compare immediate interference during interpreting and code-switching. We reasoned that interpreting requires a simultaneous activation of L1 and L2 in a way that code-switching does not. The speaker plans an utterance in the output L2 while referring to a message stored in the source L1. The results found no difference in VOT between the two bilingual tasks. If anything, the mean VOT rose slightly during interpreting (see Fig. 1). With a modification of the interpreting task removing the L1-to-L2 switch between the quote frame and the translation, a differential effect of code-switching and interpreting on immediate interference may be found.

Finally, we tested whether one's experience with alternating between languages is reflected in an improved ability to switch completely between L1 and L2 phonetic categories. Separate analyses of the beginner and trained interpreters' data found short-term interference only for beginners, which suggests that the role of experience is worth exploring further.

5. ACKNOWLEDGEMENTS

The work of the first author was supported by an internal research grant from the Faculty of Arts of Palacký University Olomouc.

6. REFERENCES

[1] Antoniou, M., Best, C. T., Tyler, M. D., Kroos, C. 2011. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *J. Phonetics* 39, 558–570.

[2] Balukas, C., Koops, C. 2014. Spanish-English bilingual voice onset time in spontaneous code-switching. *Int. J. Bilingual.* 20, 1–21.

[3] Birdsong, D. 2005. Nativelikeness and non-nativelikeness in L2A research. *IRAL – Int. Rev. Appl. Li.* 43, 319–328.

[4] Boersma, P., Weenink, D. 2014. Praat: doing phonetics by computer [Computer program]. Version 5.4., retrieved 4 October 2014, http://www.praat.org/.

[5] Bullock, B. E., Toribio, A. J., González, V., Dalola, A. 2006. Language dominance and performance outcomes in bilingual pronunciation. *Proc. 8th Generative Approaches to Second Language Acquisition* Banff, 9–16.

[6] Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., Carbone, E. 1973. The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *J. Acoust. Soc. Am. 54*, 421–428.

[7] Flege, J. E. 1991. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *J. Acoust. Soc. Am.* 89, 395–411.

[8] Flege, J. E., Eefting, W. 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Commun.* 6, 185–202.

[9] Flege, J. E., Frieda, E. M., Nozawa, T. 1997. Amount of native-language (L1) use affects the pronunciation of an L2. *J. Phonetics* 25, 169–186.

[10] Flege, J. E., Munro, M. J., MacKay, I. R. 1995. Factors affecting strength of perceived foreign accent in a second language. *J. Acoust. Soc. Am.* 97, 3125–3134.

[11] González-López, V. G. 2012. Spanish and English word-initial voiceless stop production in code-switched vs. monolingual structures. *Second Lang. Res.* 28, 243–263.

[12] Grosjean, F. 1985. Polyglot aphasics and language mixing: A comment on Perecman (1984). *Brain Lang.* 26, 349–355.

[13] Grosjean, J. 2001. The bilingual's language modes. In: Nicol, J. (ed.), *One Mind, Two Languages*. Oxford: Blackwell, 1–22.

[14] Grosjean, F. 2011. An attempt to isolate, and then differentiate, transfer and interference. *Int. J. Bilingual.* 16, 11–21.

[15] Grosjean, F., Miller, J. L. 1994. Going in and out of languages: An example of bilingual flexibility. *Psychol. Sci.* 5, 201–206.

[16] Guion, S. G., Flege, J. E., Loftin, J. D. 2000. The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *J. Phonetics* 28, 27–42.

[17] Kessinger, R. H., Blumstein, S. E. 1998. Effects of speaking rate on voice-onset time and vowel production: Some implications for perception studies. *J. Phonetics* 26, 117–128.

[18] Klatt, D. H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *J. Acoust. Soc. Am. 59*, 1208–1221.

[19] Lev-Ari, S., Peperkamp, S. 2013. Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *J. Phonetics* 41, 320–331.

[20] Lev-Ari, S., Peperkamp, S. 2014. The influence of inhibitory skill on phonological representations in production and perception. *J. Phonetics* 47, 36–46.

[21] Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20, 384–422.

[22] Machač, P. 2006. *Temporální a spektrální struktura českých explozív* [The temporal and spectral structure of Czech plosives]. PhD diss., Charles University, Prague.

[23] Mora, J. C., Rochdi, Y., Kivistö-de Souza, H. 2014. Mimicking accented speech as L2 phonological awareness. *Lang. Aware.* 23, 57–75.

[24] Myers-Scotton, C. 1993. Common and uncommon ground: Social and structural factors in codeswitching. Lang. Soc. 22, 475–503.

[25] Oller, D. K. 1973. The effect of position in utterance on speech segment duration in English. *J. Acoust. Soc. Am*. 54, 1235–1247.

[26] Piske, T., MacKay, I. R., Flege, J. E. 2001. Factors affecting degree of foreign accent in an L2: A review. *J. Phonetics* 29, 191–215.

[27] Prior, A., Gollan, T. H. 2011. Good language-switchers are good task-switchers: Evidence from Spanish-English and Mandarin–English bilinguals. *J. Int. Neuropsych. Soc.* 17, 682–691.

[28] Sancier, M. L., Fowler, C. A. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *J. Phonetics* 25, 421–436.

[29] Simonet, M. 2014. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *J. Phonetics* 43, 26–37.

[30] Verhelst, N., Van Avermaet, P., Takala, S., Figueras, N., North, B. 2009. *Common European Framework of Reference for Languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.