

# *Effects of L1 orthography and L1 phonology on L2 English pronunciation*

Article

Accepted Version

Mairano, P., Bassetti, B., Sokolovic-Perovic, M. and Cerni, T. (2018) Effects of L1 orthography and L1 phonology on L2 English pronunciation. *Revue Française de Linguistique Appliquée*, XXIII (1). pp. 45-57. ISSN 1386-1204 Available at <https://centaur.reading.ac.uk/77792/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://www.rfla-journal.org/en/content/past-issues/2018-1>

Publisher: Publications Linguistiques

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading



## Effects of L1 orthography and L1 phonology on L2 English pronunciation

Paolo Mairano, University of Warwick (UK)

Bene Bassetti, University of Warwick (UK)

Mirjana Sokolović-Perović, University of Reading (UK)

Tania Cerni, University of Warwick (UK)

### Abstract

Recent research has revealed the effects of orthography on the pronunciation of consonant durations in the L2 English of L1 Italian speakers (e.g. the [p] in *floppy* being pronounced as longer than in *copy*). In this paper we compared this orthographic effect with an orthography-independent effect of L1 phonology, namely VOT. We measured closure durations and VOT for plosives produced by 30 learners of L2 English in Italy, 30 Italian late bilingual speakers of L2 English living in the UK, and 30 native English speakers. While VOT values produced by late bilinguals differed significantly from those produced by learners, closure durations were similar across the two groups. Additionally, L1 Italian VOT values proved that late bilinguals adapted VOT in L2 English by a larger extent than learners. It appears that the effects of orthography on L2 consonant duration can be more resistant to naturalistic L2 exposure than orthography-independent effects of L1 phonology.

Keywords: pronunciation, orthography, gemination, VOT.

### Résumé

Des recherches récentes ont dévoilé les effets de l'orthographe sur la durée consonantique en anglais L2 de locuteurs natifs italiens (le [p] étant réalisé plus long pour *floppy* que pour *copy*). Nous avons comparé cet effet orthographique avec un effet de la phonologie de la L1 indépendant de l'orthographe, le *Voice Onset Time* (VOT). La tenue et le VOT d'occlusives dans des mots-cibles ont été mesurés chez 30 apprenants d'anglais L2 en Italie, 30 bilingues tardifs italiens résidant au Royaume Uni et 30 locuteurs natifs d'anglais. Si les valeurs de VOT produites par les bilingues se différencient de manière significative par rapport à celles des apprenants, les valeurs de tenue restent comparables dans les deux groupes. En outre, le taux d'adaptation du VOT en anglais L2 est plus important chez les bilingues tardifs que chez les apprenants. Cela prouve que les effets orthographiques sur la longueur consonantique peuvent être plus résistants à l'exposition de la L2 que les effets de la phonologie de la L1 indépendants de l'orthographe.

Mots clés : prononciation, orthographe, gémination, VOT.

### 1. Introduction

Several studies have investigated the effects of first language (L1) phonology on second language (L2) speech production. Only recently have researchers started investigating the effects of orthographic forms on L2 pronunciation. In this study, we investigate if L2 naturalistic exposure can have a different effect on an orthography-induced timing phenomenon (consonantal length) and on an orthography-independent one (Voice Onset Time).

#### 1.1 Effects of orthography on L2 speech production

Adult learners often acquire second languages through a mix of written and oral input, and in some cases written input may even be predominant (Bassetti 2008). The written form of words is far less subject to variability than their oral form, thereby providing a firm anchor for learners. It has been shown that written input facilitates the learning of the phonological form of words in a novel language (Erdener & Burnham 2005) and the perceptual discrimination of L2 phonological contrasts (Escudero, Hayes-Harb & Mitterer 2008; Escudero & Wanrooij 2010).

It is therefore not surprising that orthographic forms can affect the way learners pronounce a second language. Spelling does not always have an effect (Simon, Cambless & Alves 2010), and can sometimes even lead to non-native-like productions (Bassetti, Hayes-Harb & Escudero 2015). Previous research in this field has documented the effects of orthography on the L2 speech production of learners of a number of languages. For example, Zampini (1994) reported that L1 English speakers of L2 Spanish pronounced [b] or [v] in correspondence with letters <b> and <v> respectively, although [v] is not the native-like realization of <v> (<b> and <v> in Spanish both represent the phoneme /b/, pronounced as

[b] or [β] according to the phonological context). It appears that L1 English speakers pronounced [b] and [v] following the grapheme-to-phoneme correspondences of their L1. Another orthographic effect was revealed by Bassetti (2017), who found that L1 Italian speakers of L2 English produce longer consonants when they are spelled with double consonant letters (e.g., in *kitty* vs *city*), because double consonant letters represent long consonants in their L1 orthography. These findings suggest that orthography can lead L2 speakers to produce a phonological category which does not exist in the target language. The following section reports recent studies about the effect of orthography on consonantal length in L2 English.

### 1.2 Effects of orthography on consonantal length

Italian has a phonological opposition between geminate (long) consonants and singleton (short) consonants, as in *fato* ([*fato*] ‘fate’) versus *fatto* ([*fato*] ‘fact’). Italian geminate consonants are about twice as long as their singleton counterparts (Esposito & Di Benedetto 1999). Geminates have a high functional load in Italian, where it occurs in word-internal (mainly intervocalic) position for 15 out of 23 Italian consonants (Bertinetto & Loporcaro 2005). Five other consonants are intrinsically geminate, i.e. they tend to be realized as long in all contexts that allow gemination. The opposition between singletons and geminates is systematically reflected in the orthography, with double consonant letters representing geminates and single consonant letters representing singletons.

On the other hand, gemination is not a feature of English. Consonants may be realized as long at the boundary between two morphemes that end and start with the same sound, such as *top pick* or *unnamed* (‘fake gemination’, Oh & Redford 2012), and in Southern Welsh English after a stressed vowel, as in *city* (Kaye, 2005), but these are not phonological geminates. Double consonant letters do not represent consonantal length (e.g. *manner*).

Bassetti (2017) reported that L1 Italian learners of L2 English tend to produce consonants as longer if they are represented with double consonant letters than when they are represented with single consonant letters (e.g. in *floppy* vs *copy*), although the magnitude of the effect is smaller than reported for their L1: consonants spelled with double consonant letters were on average 1.66 times as long as consonants spelled with single consonant letters. This effect was present both when learners saw the written form of the target words in a reading aloud task, and when they did not see the written form of the target words in a delayed word repetition task. Bassetti and colleagues (under review) showed that this can lead to the production of minimal pairs distinguished by a long or short consonant in L2 English (e.g. a long [n] in *Finnish* and a short [n] in *finish*), and is also present in phonological contexts where Italian phonotactic restrictions would not license it, such as in word-final position (e.g. *ad* - *add*). Moreover, the authors demonstrated that this contrast is produced not only by instructed learners, but also by Italian residents in the UK, suggesting that this orthographic effect has a certain resistance to prolonged naturalistic L2 exposure. Similar effects of spelling on consonantal length were found in L1 Japanese speakers of L2 English (Sokolović-Perović & *al.*, under review).

### 1.3 Effects of L1 on VOT production in L2

The effects of L1 on L2 speech production are well documented (Colantoni & *al.* 2015), and, in particular, a number of researchers found effects of L1 on L2 Voice Onset Time (VOT). VOT is the interval of time between the release of a plosive and the start of vocal fold vibration. Lisker and Abramson (1964) first proposed VOT as an acoustic correlate of voicing distinctions across languages. When the periodic signal starts before the release of a plosive (‘voicing lead’) the VOT has a negative value, and vice versa when the periodic signal starts after the release (‘voicing lag’) the VOT has a positive value. Thus, voiced plosives generally have negative VOT values, voiceless unaspirated plosives have values around or slightly above 0 (short-lag VOT), and voiceless aspirated plosives have high positive values (long-lag VOT). Lisker and Abramson (1964) reported VOT values for a number of languages, showing for example that initial /p/, /t/, /k/ are realized with long-lag VOT in American English and short-lag VOT in Puerto Rican Spanish. Since then, VOT has been used in hundreds of studies, and compared across many languages (e.g. Cho & Ladefoged 1999). It has become an important measure of linguistic variation and has been shown to characterise the speech of dialectal, regional and social groups (among many, Ryalls & *al.* 1997; Silva 2006; Docherty & *al.* 2011; Oh 2011), pathological speech (Auzou & *al.* 2000), heritage speakers (Nagy 2015), and contact varieties (Vietti & *al.* 2018), and

has been used to investigate the perception patterns of new-borns (e.g. Eimas & *al.* 1971) and children (e.g. Aslin & *al.* 1981).

Research on VOT values in L2 speakers revealed that the L1 affects L2 VOT values in the L2 speech production of learners and consecutive bilinguals (Caramazza & *al.* 1973; Flege 1991; Flege & *al.* 1995; Splendido 2016), as well as simultaneous bilinguals (Kehoe & *al.* 2004; MacLeod & Stoel-Gammon 2005; Kupisch & *al.* 2014). Looking in particular at Italian speakers of L2 English, Flege & *al.* (1995) reported that native Italian speakers of L2 English have different VOT values compared with native Canadian English speakers. Italian has short-lag VOT for voiceless plosives (Vagges & *al.*, 1978, reported values for [p], [t], [k] of 12, 17, 50 ms respectively), whereas English has long-lag VOT (Lisker & Abramson, 1964:394, reported 58, 70, 80 ms for [p], [t], [k] respectively; Flege & *al.*, 1995:15, reported 57, 78, 77 ms for Canadian speakers). Flege & *al.* (1995) found that, despite great variability, Italian L1 speakers of L2 English who had arrived in Canada before the age of 15 were more likely to produce native-like VOT values, while speakers who had arrived in Canada after the age of 15 tended to produce VOT values that were intermediate between L1 Italian values and L1 English values.

#### 1.4 The present study

This study aimed at testing whether orthography-induced effects on L2 speech production are more resistant to naturalistic exposure than orthography-independent effects. It has been suggested that lengthy naturalistic exposure to a foreign language may result in more nativelike speech production (Trofimovich & Baker, 2006). However, this may not apply (or may apply to a lesser extent) to those features of L2 speech production that are caused by orthographic representations. The reason for this would be that such features are reinforced by continuous exposure to written representations, although they are not present in native spoken input. In other words, speakers with naturalistic exposure to the L2 are exposed to both written and oral input, and the effect of written input may prevail because of its immutable nature, as discussed above. Exposure to written input would have consequences for orthography-induced effects, such as consonant lengthening in Italian speakers of L2 English, but not for orthography-independent effects induced by phonological and phonetic patterns of the L1. The latter effects would only undergo the consequences of exposure to oral input. In order to investigate this question, we compared an orthography-induced timing phenomenon (consonant duration) and an orthography-independent one (VOT) in L2 speakers with and without naturalistic exposure. We selected consonant duration and VOT because there is consistent evidence in the literature that the former is affected by orthographic representations in L1 Italian speakers of L2 English (Bassetti 2017, Bassetti & *al.*, under review), while the latter has long been used as an acoustic cue to foreign accent in English and other second languages (e.g. Flege & *al.*, 1995). Crucially, gemination and VOT are both timing phenomena, which makes them more easily comparable from the acoustic and articulatory points of view. Although the two phenomena are obviously different (VOT reflects the relative timing of glottal and supraglottal events in plosive production, while gemination represents the relative timing of the onset and the release of the hold phase for a plosive), their comparison allows us to measure an orthography-induced phenomenon and an orthography-independent one that share the same underlying dimension – time. In order to minimise the effects of other variables, we measured closure duration and VOT in the same target consonant in the same word, for instance comparing the VOT and duration of [p] in *propose* and *oppose*. Furthermore, we measured closure duration and VOT in L2 speakers' production of the corresponding Italian consonants in cognates, for instance [p] in Italian words *propose* and *oppose*. This provided a baseline measure of gemination and VOT in the participants' L1, against which we measured their L2 production.

In order to investigate the effects of naturalistic exposure, we compared instructed learners, who had little or no experience of living abroad, and late bilinguals who had been living in an English-speaking country. If naturalistic exposure reduces both the effects of L1 VOT patterns and of orthography, bilinguals – who have been exposed to years of native spoken input – should have more native-like VOTs and consonant durations than instructed learners. If, however, orthographic effects are more resistant to naturalistic exposure than phonological effects, bilinguals should produce the same consonant with more native-like VOT values than learners, but should produce consonant durations comparable to those of learners. This may be because continuous exposure to written input would reinforce the

presence of orthographic effects in L2 English, interfering with the native spoken input where there are no differences in consonant duration.

## 2. Methodology

### 2.1 Participants

There were three groups (all  $n = 30$ ): Italian instructed learners of L2 English living in Italy, Italian late bilinguals living in the UK, and native English speakers. No participant reported visual, reading or language difficulties.

The learners group was composed of Italian instructed learners of L2 English living in Italy (age:  $M = 18$  years,  $SD = 0.5$ ; 18 were males). They were attending the fourth year in one of three high schools in Rome, where they studied English language and literature as a compulsory subject. Most of them had had little exposure to native speakers: their current English teachers were native Italian speakers, and 18 of them had never been to an English-speaking country; the median length of their stay in an English-speaking country for the other 12 participants was 1.4 months (range: 10-60 days). All but one reported considering a native-like pronunciation of English as very important or important. Participants were slightly keener on a British English than an American English pronunciation (on a 7-point scale where 7 = 'very much',  $M_{British} = 4.9$ ,  $SD = 1.4$ ;  $M_{American} = 4.6$ ,  $SD = 1.2$ ), and they used British textbooks for grammar and civilization.

The bilingual group were instructed late bilinguals (L1 Italian, L2 English) living in the UK and recruited in London (age  $M = 32.5$  yrs,  $SD = 7.3$ ; 10 were males). Entry requirements for participants in this group were the following: (a) participants had to originally come from Central or Southern regions of Italy (excluding Sardinia), where gemination is well documented (Giordano & Savy 2012); (b) participants had lived in Italy at least until completion of their high school degree; (c) participants had been living in the UK for six years and seven months on average (range: 3-24 years). Before moving to the UK, they had studied English in Italy for a median of eight years (range: 3-13). Nineteen out of 24 respondents reported considering a native-like pronunciation as important or very important (six answers were missing). Participants of this group reported spending more time reading and listening to English than Italian (Reading:  $Mdn_{English} = 15$  hrs per week, range: 1-84,  $Mdn_{Italian} = 5$  hrs per week, range: 1-15; Listening:  $Mdn_{English} = 14$ , range: 0-84,  $Mdn_{Italian} = 4$  hours, range: 0-50), but similar amounts of speaking in the two languages ( $Mdn_{English} = 25$ , range: 5-70,  $Mdn_{Italian} = 20$ , range: 2-80).

The native English group was composed of native speakers of British English (age:  $M = 30.8$ ,  $SD = 13.5$ ; 18 were males). They were undergraduate students or professionals, recruited in London or the Midlands. Their self-reported native English variety was mostly RP or Southern British English ( $n = 15$  and 10 respectively; three participants reported their variety as Yorkshire English and two as Midlands English).

Participation in the study was voluntary and compensated. For underage participants, a signed consent form was obtained from a parent or guardian.

### 2.2 Materials

We used six English words and six matching Italian words (Table 1). The target sounds were the three voiceless plosives /p/, /t/, /k/. Within each word pair, the same target sound was spelled with one consonant letter (henceforth <C>) in one word, and with double consonant letters (<CC>) in the other one.

Table 1. Orthographic forms and phonetic transcriptions of target words. Underlining of target consonants added for clarity.

| Target consonant | Consonant spelling | Language    |           |            |          |
|------------------|--------------------|-------------|-----------|------------|----------|
|                  |                    | English     |           | Italian    |          |
| /p/              | <C>                | [prə'pʰəuz] | propose   | [pro'poze] | propose  |
|                  | <CC>               | [ə'pʰəuz]   | oppose    | [o'p:ose]  | oppose   |
| /t/              | <C>                | [ə'tʰæksiə] | ataxia    | [a'taviko] | atavico  |
|                  | <CC>               | [ə'tʰækinj] | attacking | [a't:ak:i] | attacchi |
| /k/              | <C>                | [ə'kʰju:t]  | acute     | [a'kute]   | acute    |
|                  | <CC>               | [ə'kʰju:z]  | accuse    | [a'ku:ze]  | accuse   |

Words were selected so that we could measure consonant lengthening and VOT on the same target segment. The target plosives always appeared in pre-stressed position, where long-lag VOT is expected in English (Docherty 1992), but never in word-initial position, where gemination is not allowed by the phonotactic restrictions of Italian (Bertinetto & Lopoparo 2005). Moreover, we controlled as many variables as possible in order to minimize potential confounds: vowels preceding and following the target sound were orthographically and phonetically the same within each word pair (e.g. orthographically *propose* - *oppose*, phonetically [prə'pəʊz] - [ə'pəʊz]); stress fell on the second syllable on all words. Finally, in order to match the orthographic form and phonological form of the English and Italian words as much as possible, five of the six words were cognates (the English word *ataxia* was matched with the non-cognate *atavico* because the Italian cognate *atassia* has a different stress pattern). We did not avoid cognate words since growing evidence suggests that orthography-induced gemination affects all words, with consonant lengthening only slightly longer for some English words with Italian cognates containing a geminate (Bassetti & *al.* under review, while Bassetti 2017 found no differences between cognates and non-cognates).

### 2.3 Task and procedure

Participants were tested individually as part of a wider protocol designed to study the effects of orthography on L2 production, perception and phonological awareness. Equipment and testing conditions are reported in Bassetti & *al.* (under review). Recordings were digitized and saved in .wav format at 44 kHz.

All participants performed a word reading aloud task in English. We followed the same elicitation procedure as in Study 1 in Bassetti (2017): participants received a printed list of sentences, with one word underlined within each sentence. This was done in order to present target words in a natural context and to help participants recognise them. Sentences were extracted from the *British National Corpus Online Service*. Participants produced each target word three times within the frame sentence ‘I say \_\_\_ and pause’.

Bilinguals and learners then performed an Italian word reading aloud task. Italian words were used to measure closure and VOT in the participants’ native language, to be later compared with their L2 English realisations. Sentences were extracted from the *COLFIS* corpus (Bertinetto & *al.* 2005), and the frame sentence was ‘Dico \_\_\_ senza pausa’ (‘I say \_\_\_ without pauses’). The semantic content of the sentences was chosen in order to minimise the production of pauses after the target words. All participants performed the L2 task first, followed by the L1 task, in order to avoid a prime of the dominant language (Italian) on the L2 (English).

### 2.4 Analysis

For each target consonant we measured closure duration and VOT. Since closure duration is the main acoustic cue for distinguishing singleton and geminate plosives in Italian (Esposito & Di Benedetto 1999), we measured the duration of the hold phase, from the end of the formant structure of the preceding vowel, to (but not including) the onset of the release of the plosive. VOT was measured as the interval between the release of the plosive and the first noticeable periodic cycle. Measurements were performed in Praat (Boersma & Weenink 2016) manually by a trained phonetician (the first author). The realizations of 106 consonants (i.e. 3.9% of the data) had to be excluded from the analysis because of hesitations, deviant pronunciations in the case of L2 speakers (e.g. ['ækjʊ:t] instead of [ə'kjʊ:t] for *acute*), or because they had been unintentionally omitted by participants. In total, we measured closure duration and VOT for 2,594 consonants.

Statistical analyses were performed using *R* 3.4.2 (R Core Team 2017). We built linear mixed-effects models with the *lme4* 4.1 package (Bates & *al.* 2015) and used the *multcomp* 1.4 package (Hothorn & *al.* 2008) for post-hoc tests.

## 3. Results

### 3.1 Closure duration

We calculated closure duration ratio for each participant for each word pair, by dividing the closure duration of the target plosive in each <CC> word (e.g. [k] in ‘accuse’) by the

duration of the target plosive in its corresponding <C> word (e.g. [k] in ‘acute’). A ratio of one indicates no difference in duration between <CC> and <C> sounds, whereas a ratio of two indicates that the <CC> sound is realized twice as long as the <C> sound. Using ratios rather than raw durations helps to normalize with respect to speech rate across participants. Table 2 shows mean closure duration ratios.

Table 2. Mean closure duration ratios by group (learners, bilinguals, native English) and language of testing (English, Italian; SDs in brackets).

| Group          | Language    |             |
|----------------|-------------|-------------|
|                | English     | Italian     |
| Learners       | 1.43 (0.40) | 2.03 (0.50) |
| Bilinguals     | 1.32 (0.43) | 1.88 (0.50) |
| Native English | 1.02 (0.20) | N/A         |

### 3.1.1 Closure duration ratios in English

As Table 2 shows, the mean ratio of native English speakers is 1.02, as they produce <C> and <CC> sounds with comparable durations. On the other hand, the two L2 groups have mean ratios of 1.43 and 1.32 (learners and bilinguals respectively) when speaking English, showing that they produced <CC> sounds as longer than <C> sounds. Results were entered into a linear mixed-effects model with closure duration ratio in English as the dependent variable, group (learner, bilingual, native English) as a fixed effect, and word pair and participant as random effects:  $ratio \sim group + (1|participant) + (1|wordPair)$ . Model reduction and a likelihood ratio test showed an effect of group ( $\chi^2(2) = 41.947, p < 0.001$ ). Pairwise post-hoc tests with Tukey correction revealed that ratios were smaller for native English speakers than for learners ( $z = -6.986, p < 0.001$ ) and bilinguals ( $z = -5.107, p < 0.001$ ), but ratios for bilinguals and learners did not differ significantly ( $z = 1.880, p = 0.144$ ).

### 3.1.2 L2 speakers’ closure duration ratios in English and Italian

Figure 1 shows L2 speakers’ mean closure duration ratios in L1 Italian productions and in L2 English productions. To compare ratios in L2 English and L1 Italian, and to compare learners and bilinguals, we built a linear mixed-effects model with closure duration ratio as the dependent variable, group (learner, bilingual) and language (English, Italian) as fixed effects, and word pair and participant as random effects:  $ratio \sim group * language + (1+language|participant) + (1|wordPair)$ . Model reductions and likelihood ratio tests revealed a main effect of language ( $\chi^2(2) = 84.669, p < 0.001$ ). There was no significant effect of group ( $\chi^2(2) = 4.955, p = 0.084$ ) or interaction ( $\chi^2(1) = 0.0268, p = 0.605$ ). The absence of interaction shows that both bilinguals and learners have smaller ratios in L2 English than in L1 Italian, and the effect of language is comparable between the two groups. Post-hoc pairwise comparisons with Tukey correction revealed ratios in Italian were larger than in English for both bilinguals ( $z = 9.282, p < 0.001$ ) and learners ( $z = 9.898, p < 0.001$ ), and that there were no differences between the two groups in either Italian ( $z = -1.933, p = 0.208$ ) or English ( $z = -1.755, p = 0.288$ ).

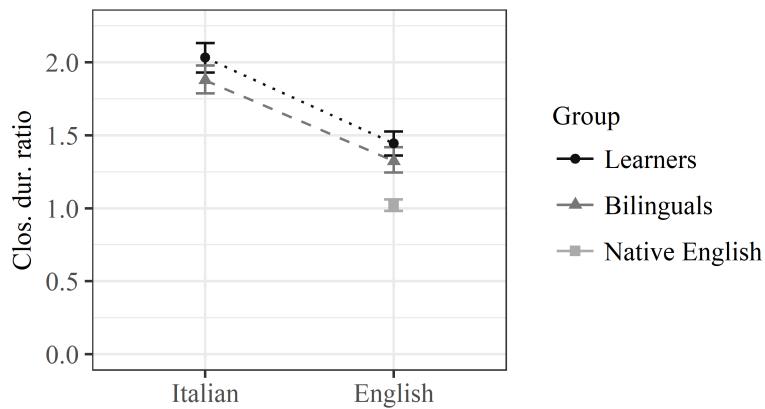


Figure 1. Mean closure duration ratios by group (learners, bilinguals and native English) and language of testing (English, Italian). Error bars indicate 0.95 confidence intervals.

### 3.2 VOT

Table 3 shows mean VOT durations in ms for the learner, bilingual and native English groups.

Table 3. Mean VOT durations for /p/, /t/, /k/ (in ms; *SD* in brackets) by group (learners, bilinguals, native English) and language (English, Italian).

| Group      | English |         |          |         | Italian |        |         |         |
|------------|---------|---------|----------|---------|---------|--------|---------|---------|
|            | /p/     | /t/     | /k/      | overall | /p/     | /t/    | /k/     | overall |
| Learners   | 29 (12) | 29 (16) | 77 (16)  | 46 (27) | 24 (11) | 18 (4) | 50 (15) | 31 (18) |
| Bilinguals | 43 (20) | 36 (22) | 92 (21)  | 57 (32) | 28 (9)  | 19 (8) | 57 (14) | 35 (20) |
| Ntv. Eng.  | 62 (15) | 79 (21) | 106 (24) | 82 (27) | /       | /      | /       | /       |

#### 3.2.1 VOT in English

As Table 3 shows, the native English group produced the longest VOTs, the learner group produced the shortest ones, and the bilingual group produced intermediate ones. We analysed results using a linear mixed-effects model with VOT as the dependent variable, group (learner, bilingual, native English) as the fixed effect, and word and participant as random effects. We included random slopes for word as per the following formula:  $VOT \sim group + (I|participant) + (I|word)$ . Model reduction and a likelihood ratio test showed an effect of group ( $\chi^2(2) = 69.807, p < 0.001$ ). Pairwise comparisons with Tukey correction revealed that native English speakers produced longer VOTs than both bilinguals ( $z = 6.771, p < .001$ ) and learners ( $z = 9.952, p < .001$ ), and bilinguals produced longer VOTs than learners ( $z = 3.183, p = .004$ ).

#### 3.2.2 L2 speakers' VOT in English and Italian

Figure 2 shows L2 speakers' mean VOTs in L1 Italian productions and in L2 English productions. To compare VOTs in L2 English and L1 Italian, and to compare learners and bilinguals, we built a linear mixed-effects model with VOT as dependent variable, group (learner, bilingual) and language of testing (English, Italian) as fixed effects, and word pair and participant as random effects:  $VOT \sim group * language + (I+language|participant) + (I|word)$ . Model reductions and likelihood ratio tests revealed effects of language ( $\chi^2(2) = 84.423, p < 0.001$ ) and group ( $\chi^2(2) = 12.174, p = 0.002$ ), and an interaction ( $\chi^2(1) = 6.954, p = 0.008$ ). A post-hoc analysis with Tukey correction revealed that all pairwise differences across groups and languages were significant ( $z = -2.519, p = 0.050$  for learners and bilinguals in Italian;  $z = -3.514, p = 0.002$  for learners and bilinguals in English;  $p < 0.001$  for all others). The interaction between group and language shows that both bilinguals and learners increase their VOT durations when speaking English compared with Italian, but the increase is much larger among bilinguals than among learners.

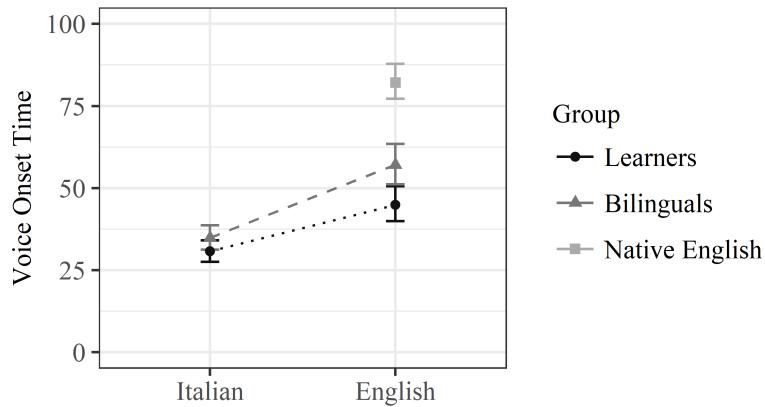


Figure 2. Mean VOT durations (in ms) by group (learners, bilinguals, native English) and language of testing (English, Italian). Error bars indicate 0.95 confidence intervals.

### 3.3 VOT and closure duration

Finally, we investigated the relationship between the extent to which participants adapted VOT values and closure duration ratios when producing English word pairs and Italian word pairs. In order to do so, for each participant we calculated a  $\Delta$ VOT as the difference between their Italian and English VOT values, and a  $\Delta$ ClosureDurationRatio as the difference between their ratios in Italian and English. There was no correlation between  $\Delta$ VOT and  $\Delta$ ClosureDurationRatio for either bilinguals ( $r = -0.06, p = 0.759$ ) or learners ( $r = -0.16, p = 0.408$ ).

## 4. Discussion

This study aimed at testing whether an orthography-induced pronunciation feature could be more resistant to naturalistic L2 exposure than an orthography-independent one. To this aim, we compared two effects of L1 Italian affecting the timing of L2 English consonants: the effect of consonant spelling (<C> vs <CC>) on closure duration, and the effect of L1 on VOT productions in the L2. We compared a group of instructed learners and a group of late bilinguals living in an English-speaking environment. We found both effects of L1 orthography on L2 gemination and effects of L1 VOT patterns on L2 VOT. Crucially, we found that bilinguals with stay abroad experience produced more native-like VOTs compared with instructed learners, but closure durations in L2 English did not differ across the two groups. Below we discuss first the effects of L1 orthography on closure duration, then the effects of L1 on VOT, and finally the comparison of the two.

L1 orthography affected L2 speakers' consonant duration in English, in line with previous findings. Both learners and late bilinguals produced the same English consonant as longer when it was spelled with double letters than when it was spelled with a single letter, whereas spelling did not affect native speakers' consonant duration, as found in previous studies (Bassetti 2017; Bassetti & *al.*, under review). The learners' mean closure duration ratio of 1.43 was smaller than the ratios found by Bassetti (2017) with similar participants (1.66 in a word reading task, and 1.70 in two word repetition tasks), but both the learners' and the bilinguals' ratios were similar to the 1.38-1.39 ratios reported by Bassetti and colleagues (under review). While differences across studies may be due to factors such as different selection of words or individual variations among participants, there is now a consistent body of evidence from different types of L2 speakers, different tasks and different words confirming that Italian speakers of L2 English produce English double-letter sounds as long consonants. Closure duration ratios were smaller in L2 English than in L1 Italian (see also Bassetti & *al.*, under review). This may be due to the interference between orthographic input (double letters) and spoken input (native speakers' productions where consonants are of similar duration in words spelled with single or double letters).

L1 affected VOT realizations in L2 English, as both learners and bilinguals produced shorter VOT values than native speakers. These results confirmed findings from adult Italian immigrants in Canada reported by Flege and colleagues (1995), as well as findings of other studies on L2 speech where L1 has long-lag VOT and L2 has short-lag VOT (e.g., Caramazza & *al.* 1973 for L1 French speakers of L2 English; Flege 1991 for L1 Spanish

speakers of L2 English). Crucially, we found that bilinguals produced longer VOTs in L2 English than learners, suggesting that exposure to the L2 results in a larger VOT adaptation. Additionally, bilinguals produced longer VOT values than learners also in L1 Italian, although the difference between groups is smaller than in L2 English: such differences in L1 VOT values may depend on different dialectal backgrounds (Loporcaro 2009) between the two groups (Roman variety for learners, various Central and Southern varieties for bilinguals), or may reflect a shift of L1 VOTs in the bilingual group as a result of naturalistic L2 exposure (as demonstrated by Major 1992; Sancier & Fowler 1997).

The main finding of this study is that, although bilinguals adapted VOTs and produced longer VOTs than learners with no naturalistic L2 exposure, they were affected by orthography as much as learners: both groups produced comparable closure durations for consonants spelled with double letters. The analysis of consonantal lengthening in L2 English vs L1 Italian revealed no interaction between groups and languages: both learners and bilinguals decreased the magnitude of <CC> lengthening when speaking L2 English compared with L1 Italian. Despite overall smaller closure duration ratios for bilinguals than for learners in both L1 Italian and L2 English (perhaps attributable to age and/or dialectal differences across groups), the two groups behaved similarly in both languages. In contrast, the interaction of language and group was significant for VOT: both groups produced longer VOTs in L2 English than L1 Italian, but bilinguals adapted their VOT by a larger extent than learners. Again, there were small differences in VOT between the two groups which may be attributed to age or dialectal differences; yet, the interaction of group and language is evidence of a larger VOT adaptation in L2 English by bilinguals than learners.

We can conclude that orthographic effects on L2 pronunciation can not only withstand L2 exposure, but can even be more resistant to naturalistic L2 exposure than the effects of L1 phonetics and phonology. We argue that this happens because orthography-induced phenomena are reinforced by continuous exposure to written input and are therefore more likely to resist the influence of contrasting oral input, compared to phenomena that are rooted in L1 phonetic and phonological patterns. L2 speakers with naturalistic exposure to the L2 are exposed both to oral and written input. However, when there is no conflict between the two types of input (as in the case of long-lag VOT) continuous exposure to the oral form may be beneficial and result in more native-like productions, as shown by many studies (see 1.3). By contrast, when written and oral input are incongruent (from the point of view of the L2 speaker, as in the case of consonantal length), the beneficial effects of exposure to L2 oral input may be reduced. This would explain our finding that L1 Italian late bilinguals adapt VOT more than they adapt gemination.

Future studies may investigate a larger number of target words, speakers with different language combinations, and other correlates of foreign accent, such as vowel formants. In particular, it may be worth exploring spontaneous and semi-spontaneous speech in learner corpora, which have already been used to investigate acoustic correlates of foreign accent (see Díaz-Negrillo, Ballier & Thompson 2013). Finally, it may be interesting to compare orthographic effects vs effects of L1 phonology on the intelligibility of L2 pronunciation. In the case of the present study, consonant lengthening may affect intelligibility less than non-native-like VOT, because native speakers of languages without gemination are less sensitive to consonantal duration than speakers of languages with phonological gemination (Hallé & Ridouane 2011), whereas non-native-like VOT may lead to incorrect categorizations (for instance, short-lag /t/ can be confused with /d/).

### Acknowledgements

The data for this study were collected within a project funded by a Leverhulme Trust Research Grant [grant number: RPG 2013 180] awarded to B. Bassetti and J. Masterson. The authors are grateful to all students and teachers who took part, and to the two anonymous reviewers who provided useful comments and suggestions.

### Bibliography

Aslin, R. N., Pisoni, D. B., Hennessy, B. L. & *al.* (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child development*, 52(4), 1135-1145.

Auzou, P., Ozsancak, C., Morris, R. J. & *al.* (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: a review. *Clinical Linguistics & Phonetics*, 14(2), 131-150.

Bassetti, B. (2008). Orthographic input and second language phonology. In Piske, T. & Young-Scholten, M. (eds.), *Input Matters in SLA*. Clevedon, UK: Multilingual Matters (pp. 191-206).

Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(11), 1835-1842.

Bassetti, B., Hayes-Harb, R. & Escudero, P. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, 36(1), 1-6.

Bassetti, B., Sokolović-Perović, M., Mairano, P. & al. (under review). Orthography-induced length contrasts in the second language phonological systems of experienced speakers of English as a Second Language: Evidence from minimal pairs.

Bates, D., Maechler, M., Bolker, B. & al. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Bertinetto, P. M., Burani, C., Laudanna, A. & al. (2005). Corpus e lessico di frequenza dell’italiano scritto (CoLFIS). Retrieved from <http://linguistica.sns.it/CoLFIS/Home.htm>

Bertinetto, P. M. & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, 35(2), 131-151.

Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer (Version 6.0.19). Retrieved from <http://www.praat.org/>

British National Corpus Online service. Retrieved from <http://corpora.lancs.ac.uk/BNCweb/>

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B. & al. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *The Journal of the Acoustical Society of America*, 54(2), 421-428.

Cho, T. & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of phonetics*, 27(2), 207-229.

Colantoni, L., Steele, J. & Escudero, P. (2015). *Second language speech*. Cambridge University Press.

Díaz-Negrillo, A., Ballier, N., & Thompson, P. (Eds.). (2013). *Automatic treatment and analysis of learner corpus data*. Amsterdam, Philadelphia: John Benjamins Publishing Company.

Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris.

Docherty, G. J., Watt, D., Llamas, C. & al. (2011). Variation in voice onset time along the Scottish-English border. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 591-594), 17-21 August 2011, Hong Kong.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. & al. (1971). Speech perception in infants. *Science*, 171(3968), 303-306.

Erdener, V. D. & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, 55(2), 191-228.

Escudero, P., Hayes-Harb, R. & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345-360.

Escudero, P. & Wanrooij, K. (2010). Learning the phonological forms of new words: Effects of orthographic and auditory input. *Language and Speech*, 53(3), 361-381.

Esposito, A. & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America*, 106(4), 2051-2062.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *The Journal of the Acoustical Society of America*, 89(1), 395-411.

Flege, J. E., Munro, M. J. & MacKay, I. R. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, 16(1), 1-26.

Giordano, R. & Savy, R. (2012). Sulla consonantizzazione del consonantismo dell’italiano: Consonanti geminate, rafforzate e fricative alveolari in contesto intervocalico. In P. Bianchi, N. De Blasi, C. De Caprio & al. (Eds.), *La variazione nell’italiano e nella sua storia: Varietà e varianti linguistiche e testuali* (pp. 431-446). Firenze, Italy: Franco Cesati Editore.

Hallé, P. & Ridouane, R. (2011). French listeners’ deafness to Tashlhiyt Berber/bi/-/bbi. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 811-814), 17-21 August 2011, Hong Kong (China).

Hothorn, T., Bretz, F. & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346-363.

Kaye, A. S. (2005). Gemination in English. *English Today*, 21(2), 43-55.

Kehoe, M. M., Lleó, C. & Rakow, M. (2004). Voice onset time in bilingual German-Spanish children. *Bilingualism: Language and Cognition*, 7(1), 71-88.

Kupisch, T., Lein, T., Barton D. & al. (2014). Acquisition outcomes across domains in adult simultaneous bilinguals with French as weaker and stronger language. *Journal of French Language Studies*, 24(3), 347-376.

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

Loporcaro, M. (2009). *Profilo linguistico dei dialetti italiani*. Bari: Edizioni Laterza.

MacLeod, A. A. & Stoel-Gammon, C. (2005). Are bilinguals different? What VOT tells us about simultaneous bilinguals. *Journal of Multilingual Communication Disorders*, 3(2), 118-127.

Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, 76, 190-208.

Nagy, N. (2015). A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, 309-327.

Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics*, 39(1), 59-67.

Oh, G. E. & Redford, M. A. (2012). The production and phonetic representation of fake geminates in English. *Journal of phonetics*, 40(1), 82-91.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>

Ryalls, J., Zipprer, A. & Baldauff, P. (1997). A preliminary investigation of the effects of gender and race on voice onset time. *Journal of Speech, Language, and Hearing Research*, 40(3), 642-645.

Sancier, M. L. & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421-436.

Silva, D. J. (2006). Variation in voice onset time for Korean stops: A case for recent sound change. *Korean Linguistics*, 13(1), 1-16.

Simon, E., Chambliss, D., & Alves, U. K. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language sciences*, 32(3), 380-394.

Sokolović-Perović, M., Bassetti, B. & Dillon, S. (under review). Effects of English orthographic forms on speech production in Japanese speakers of English as a Second Language.

Splendido, F. (2016). Phonetic-phonological development of early second language French Support for a child second language (cL2) mode of acquisition in phonology. *Revue française de linguistique appliquée*, 21(2), 19-32.

Trofimovich, P. & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28(1), 1-30.

Vagges, K., Ferrero, F. E., Caldognetto, E. M. & al. (1978). Some acoustic characteristics of Italian consonants. *Journal of Italian Linguistics*, 3, 69-85.

Viotti, A., Alber, B. & Vogt, B. (2018). Initial laryngeal neutralization in Tyrolean, *Phonology*, vol. 35, pp. 79-114.

Zampini, M. L. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania*, 77(3), 470-481.

Corresponding author: Paolo Mairano  
paolomairano@gmail.com  
Laboratorio di Fonetica “Arturo Genre”  
Via Sant’Ottavio 20, 10124, Torino, Italy.