Mitterer, H. et al. (2021). *Xjenza Online*, 9 Special Issue: 162–172.

Xjenza Online: Science Journal of the Malta Chamber of Scientists
www.xjenza.org
DOI: 10.7423/XJENZA.2021.3.09

Research Article

The role of orthography in learning a second language: Evidence from Maltese English

H. Mitterer*
Department of Cognitive Science, Faculty of Media and Knowledge Sciences, University of Malta, Msida, Malta

Abstract. Research has indicated that the acquisition of a second language (L2), in particular of its phonology, is influenced by orthography. For instance, Bassetti (2017) found that Italian learners of English produce the /p/ in words with a double letter ⟨p⟩ (such as *pepper*) with a longer [p] than the /p/ in words with a single letter ⟨p⟩ (such as *weapon*). This indicates that Italian learners are influenced by their first language (L1) orthography-to-phonology rules, where a phonological quantity contrast between short and long consonants is cued as such in orthography. We tested whether this pattern is due to a focus on orthography in most formal L2 education by testing Maltese learners of English. Just as Italian learners, Maltese learners have a quantity distinction in their native language that is coded by single versus double letters. However, unlike Italian learners, the English L2 is used spontaneously outside the classroom, so that acquisition is based less on orthography. The results show that Maltese learners do not make a quantity distinction in English words with single versus double letters. This indicates that earlier results are due to the focus on orthography in formal education rather than an automatic use of orthography in speech processing.

1 The role of orthography in learning a second language: Evidence from Maltese English

The relation between spoken and written language is a curious one. Even though spoken language precedes written language both in human evolution and individual development, our thinking about language is strongly based on written language. For instance, most native speakers of English will say that English has five to six vowels, depending on whether 'y' should be counted. However, English has more than 10 different vowels, the exact number depending on the version of English. Going beyond such effects on phonological (un-)awareness, there are claims that learning a written language influences the structure of the system of spoken language (Kolinsky et al., 2021; Morais, 2021; Pattamadilok et al., 2009; Ziegler et al., 1998). In a highly cited paper, Dehaene et al. (2010) investigated brain responses to speech in different groups of adults in Brazil who did not learn to read as children due to socio-economic reasons. Some of them had remained illiterate while others had recently learned to read in adulthood. These groups are hence similar in their socio-economic background, but one has recently learned to read. Dehaene et al. (2010) found that the response to speech in left temporal areas, even when presented by itself without orthography, was stronger in the literate group, suggesting that learning to read changes networks that are used for speech perception. While these conclusions are contested for the processing of the native language (Cutler et al., 2012; Hervais-Adelman et al., 2021; Mitterer et al., 2015), an influence of orthography may be less surprising in the acquisition of a second language. More often than not, learning a second language in formal education relies strongly on the written modality. This may be partly due to the fact that written language provides an additional mnemonic that facilitates learning (for a recent review, see Hayes-Harb et al. (2021)). Moreover, students may very well capitalize on the written modality, as it is the modality which mostly determines their grades. It is hence not surprising that orthography plays a role in L2 speech processing. These influences are, however, not always facilitatory (Hayes-Harb et al., 2021). For instance, English learners of German may produce word-final stops, which are always unvoiced in German due to final devoicing (/hund/ → [hunt]), with differences in voicing depending on the orthography. That is, they produced voiced stops for words that end on graphemes that usually indicate a voiced

Correspondence to: H. Mitterer (holger.mitterer@um.edu.mt)

stop (i.e., ⟨b⟩,⟨d⟩, and ⟨g⟩, see Hayes-Harb et al. (2018) and Young-Scholten (2002)). While the German words *Rat* (Engl., 'advice') and *Rad* ('bike') are homophonous in German (or nearly homophonous, see Roettger et al. (2014)), English learners apparently use their grapheme-phoneme correspondences from their L1 and make a difference between these words, since, in English, a ⟨d⟩ is typically produced as a /d/ and the contrast between /t/ and /d/ is functional in all positions (though the contrast may be neutralized in spontaneous speech in some positions, see Pitt et al. (2011)). Similar negative effects have been found for Brazilian learners of English, who pronounce the "silent" ⟨e⟩ in words such as *bone*, and English learners of Spanish, who associate the letter ⟨b⟩ only with the stop, but not the approximant allophone of the phoneme /b/ in Spanish (Shea, 2017). Another negative transfer of L1 orthography to L2 production has been reported in series of papers by Bassetti and colleagues for Italian learners of English. Italian makes quantity contrasts for consonants, so that there is a singleton /t/ and a geminate /t:/ (e.g., *seta* 'silk' vs. *setta* 'sect'). Acoustically, the difference between singletons (/t/ as in *seta*) and geminates (/t:/ as in *setta*) is mostly signalled in terms of duration (though there are some exceptions to this rule, see Mitterer (2018)), and the difference is made in orthography with single versus double letters. Bassetti (2017) tested whether the difference between a single and double letter also influences the pronunciation of English words by Italian learners, even though English does not have quantity contrast and consonant letter doubling mostly signals the quality of the preceding vowel (so that a double letter is unlikely after a tense vowel or a diphthong). However, Italian learners might use their L1 orthography-to-phonology rules when learning English and produce the word *pepper* with a longer medial /p/ than the word *weapon*. This is indeed what Bassetti (2017) found. Importantly, this was tested under two conditions, one in which the word productions were triggered by a picture prompt that included the written word and one in which the written word was absent. In both conditions, Italian learners produced longer consonants for double letters than for single letters, indicating that the effect even occurs in the absence of any orthographic input at the time of testing. A second study (Bassetti et al., 2018) replicated this pattern with homophones (e.g., *finish* vs. *Finnish*) and also found an effect on English vowels depending on whether they were spelled with a digraph or not (e.g., *seen* vs. *scene*). The latter effect, with a lengthening of about 10%, was, however, considerably smaller than the effect for consonants, with a lengthening of about 30 to 50% when the geminate would have been phonotactically legal in Italian (see also Bassetti et al. (2020)). A similar, albeit smaller effect was found in Japanese (Sokolović-Perović et al., 2020), in which the geminate is indicated by one abstract grapheme that can be combined with any consonant (Sadakata et al., 2014). Importantly, there are (at least) two competing explanations for such effects. One explanation argues that this is simply another consequence of the automatic activation of orthographic representations during spoken-language processing, as argued for the L1 (e.g., Pattamadilok et al. (2009)). Another explanation focusses more on the learning process, in which learners are exposed to orthographic forms which may then influence phonological representations. This account gains credibility when considering the finding that reading a novel word (in the L1) leads to a phonological representation for that word (Bakker et al., 2014). The two accounts differ in that the influence is automatic according to the first account while it is considered mediated by learning in the second. One way to address this issue is to test whether the effect is moderated by other variables, because a hallmark of automaticity is that it is not strongly moderated by third variables. In this context, Bassetti and colleagues (Bassetti et al., 2020; Bassetti et al., 2018) investigated whether the orthographic influence is moderated by proficiency, comparing intermediate learners in Italy with (late) bilinguals living in the UK. In a first study (Bassetti et al., 2018), both groups performed similarly, with double letters being pronounced 38 to 39% longer than single letters, but in a later study with a larger sample size (80 intermediate learners compared to 80 highly proficient bilinguals, Bassetti et al., 2020), the bilingual group pronounced the double letters 33% longer while the learners produced them 58% longer than the single letters. However, even highly proficient bilinguals still showed an effect. This suggests that the orthographic effect is quite stable and due to an automatic activation of orthography. One important aspect of the bilinguals tested in those studies (Bassetti et al., 2020; Bassetti et al., 2018), however, is that their acquisition was based on high-school experience, in which orthography tends to take centre-stage. That is, the bilingual group learned English in Italian high schools before moving to England at the age of 18 (or later). Young-Scholten (2002), who only tested three English L1 participants learning German, argued that individual differences in orthographic influences on speech production in her sample arose due to exposure to written German. However, with only three participants, this cannot be considered more than a post-hoc suggestion rather than a decisive data point. Therefore, in the current study, we test whether an influence of orthography effect is found in a group of learners for which acquisition is accompanied by spontaneous interactions in the L2

from very early on. Such a sample can be found on Malta, where both English and Maltese are official languages, but Maltese is clearly the language that most learners learn as their first language (Vella, 2013). Even though some claim that English is becoming the primary language (Thusat et al., 2009), a recent survey by the National Council for the Maltese Language found that 97% of the respondents say that Maltese is their primary language (Il-Kunsill Nazzjonali tal-Ilsien Malti, 2021). Maltese is a Semitic language with many lexical borrowings from Italian and English through language contact, but still retaining a non-concatenative root morphology for Semitic words. Moreover, Maltese also distinguishes singleton and geminate consonants, which are signalled in orthography by a contrast between single and double letters (e.g., *dafak* /dafak/ Engl. 'to laugh' vs. *daffak* /daf:ak/ Engl. 'to make somebody laugh'). With regard to the use of single versus double letters, Maltese learners of English are therefore in a similar situation as Italian learners. However, a difference arises in how English is acquired. Maltese learners are exposed to English when they start going to school at the age of 5. The situation may be compared to the situation in Catalunya, where Spanish dominant learners are exposed to Catalan from kindergarten age onwards, which already leads to limitations on L2 phonological acquisition (Pallier et al., 1997). Nevertheless, since some speakers in Malta will use English spontaneously, learners are exposed to spoken language in a naturalistic setting from early on. This makes the situation clearly distinct from the Italian-English late bilinguals in the studies of Bassetti and colleagues (Bassetti et al., 2020; Bassetti et al., 2018). We therefore asked the question whether the effect of orthography as reported by Bassetti and colleagues for Italian learners also arises for Maltese learners of English. In doing so, we replicated the procedure from Bassetti (2017) with a delayed word repetition task with no orthographic prompt. In this task, participant first hear and repeat a phrase (e.g.,"a pot and a kettle", see figure 1 for details). They then hear the same phrase with a missing word (e.g.,"a pot and a...") and have to say three times which word is missing (i.e.,"the word kettle is missing").

2 Method

2.1 Participants

Forty native speakers of Maltese participants recruited from the student population at the University of Malta took part in the study. They reported no hearing impairments or visual difficulties. The data from five participants were excluded because they failed to produce a sufficient number of valid responses. Four participants produced the wrong target sentence and replied with "The missing words is ..." rather than "The word ... is missing". This is problematic because durations are then strongly influenced by utterance-final lengthening. One participant was rejected because only 60% of the utterances were correct, while all other participants scored over 70% correct responses. In the final data set, data from thirty-five participants (nineteen males and sixteen females; age: $M = 23.5$, $SD = 8.9$) were used. In a questionnaire given before the experiment, native speakers of Maltese reported English and Maltese, spoken and written, proficiency use on an average of 3.69 ($SD = 0.16$) on a scale from 1 (not proficient) to 4 (very proficient).

2.2 Materials

We used the same 18 items used in Bassetti (2017) and added 18 items, which were identified with a lexical search for pairs with the same vowel-consonant-vowel (VCV) sequence. As consonants, we looked for voiceless stops because those allow relatively straightforward segmentation in the acoustic signal and consequent measurement of duration. We looked for pairs in which the consonant was once written with a double letter and once as a single letter. Appendix A provides a list of all items, in which there are six pairs each for /p/, /t/, and /k/, including their frequency ('zipf'-scaled, see Heuven et al. (2014) in spoken and written language). Within each pair, the words had the same number of syllables, a comparable frequency in both spoken and written forms, and the critical consonants appearing in the same phonological environment, that is, in the same syllable and consequently same position with regard to stress and the same surrounding phonemes. Note that in Maltese, geminates are less phonologically restricted and can occur before glides (e.g., *nettjar*, Engl., 'netting'), so that pairs such as *acute-accuse* provide an environment where a geminate is legal in the participants' L1. For each target, we generated (or reused from Bassetti (2017)) a phrase in which the target was somewhat predictable (e.g., for *pepper* "oil, vinegar; salt and pepper"). These phrases were recorded by a female native speaker of Maltese English in two versions, a full version and a version with the target word missing.

2.3 Procedure

Participants first filled in an informed-consent form and then a short questionnaire about their use of English and Maltese at various stages in their life (0-6, 6-12, 12-17, adulthood). Then they were familiarized with the production task with two practice items outside the booth. When they indicated to have understood the procedure, they entered the sound-attenuated booth and performed the production task. In the booth, they first completed another test trial and then moved on to the 36 experi-

mental items.

On each trial (see figure 1) they first heard a phrase while seeing a related picture. They then were prompted to count backwards from five to one before repeating the phrase. The same counting procedure was used by Bassetti (2017) to reduce any influence from hearing the target word during exposure on the pronunciation in the test phrase. After counting backwards, they were prompted to repeat the phrase. Their performance was monitored by an experimenter who then decided whether their repetition was correct. If not, the exposure part was repeated (up to 5 times). After having repeated the phrase correctly, they again saw the picture and heard the phrase with one word missing. They were then prompted three times to name the missing word in the sentence frame "the word ... is missing". The order of the 36 items was randomized with the constraint that if one item of a pair appeared in the first half, the other appeared in the second half to minimize any demand characteristic. After the 36 experimental trials, participants did a dictation task, during which they heard the target words in a random order and had to write them down. An experimental session lasted about 25 minutes.

2.4 Apparatus and Equipment

The experiment was designed in PsychoPy v3.0.7 (Peirce, 2007). The participants were seated in a sound-proof booth, viewing the task through a monitor which was connected to the main computer, a standard PC, outside the experimental booth. The experimenter could listen to the participants' responses and switch between trials using the space bar. All auditory tasks, such as response listening and voice recording, were done through Scarlett Studio 2i2 set, 2nd generation, including a headset, a microphone, and an audio interface that digitized the responses before storing them on the computer.

2.5 Data coding and analysis

The resulting sound files were automatically aligned using the Webmaus tool (Strunk et al., 2014). The resulting segmentations were hand corrected. While it was initially planned to measure closure duration (following Bassetti (2017)), initial attempts at coding revealed that Maltese speakers often produced a lenited fricative version of the phoneme /t/ (see, e.g., Mitterer et al. (2006), for a similar allophonic variation in Dutch). This made it impossible to measure closure duration (see figure 2, where the closure duration would be zero, since no full closure is ever attained). Therefore, the duration of the obstruent was measured from either the onset of frication or closure to the onset of voicing for the following vowel. The resulting duration data were analysed using linear-mixed effect models in R4.0.4 (Team, 2020) using the package lmerTest (Kuznetsova et al., 2020). We used letter quantity as a treatment coded predictor, with singleton mapped on the intercept. Participant and Pair (i.e., the paired items, such as *weapon-pepper*) were used as random effects. As described below, we explored the effect of using various random-effect structures on the outcome of the analysis and also present some ANOVAs to further elucidate the effect of random slopes on the outcomes. All data and analysis files are available here: https://osf.io/dzvfb/.

3 Results

For the 35 participants in the final data set, with 3780 data points, 385 (10.2%) were rejected because the participant did not provide a correct response, or the response was too slow to be recorded in the recording time window [1] . Another 192 trials (5.6%) were rejected because the participants failed to provide the correct response in the dictation task. For the remaining 3203 observations, the mean consonant duration was 101ms for single letter words and 105ms for double letter words. Figure 3 shows the mean duration for single- and double-letters for participants and items with the grand means and within-subject standard errors (Morey, 2008) as error bars. There is a small lengthening effect that is relatively stable over participants but not over items. This interestingly is exactly the situation described by Clark (1973) when arguing for the need to take item variability into account. This is currently achieved using linear mixed-effects models (Baayen et al., 2008). As it turns out, the results also show the need for random slopes when using such models. Table 1 shows the results of the models with and without random slopes. These results indicate that the inclusion of random slopes is necessary here. Without a random slope of quantity over pair, there seems to be a highly reliable effect of letter quantity. This analysis is in fact similar to a by-participant ANOVA on participant means. Indeed, such an analysis also provides an estimate that would indicate a significant effect of letter quantity ($F(1, 34) = 77.1$, $p < .001$). However, once a random slope for the quantity effect over item-pairs, no effect is observed anymore (just as in an by-item ANOVA, $F(1, 33) = 1.11$, $p = 0.306$). This indicates that the effect is highly variable over items, but not over participants. Given that there is no significant effect, it is important to consider what kind of effects are unlikely given the current data. The effect of a double letter is estimated at

[1] Pretests had shown that the PsychoPy version used at time of testing was not storing recordings correctly if those were ended by a button press (e.g., by the experimenter). Therefore, a fixed recording time window of 3s was chosen that was long enough to contain the majority of the responses without making the procedure too slow for the participants.

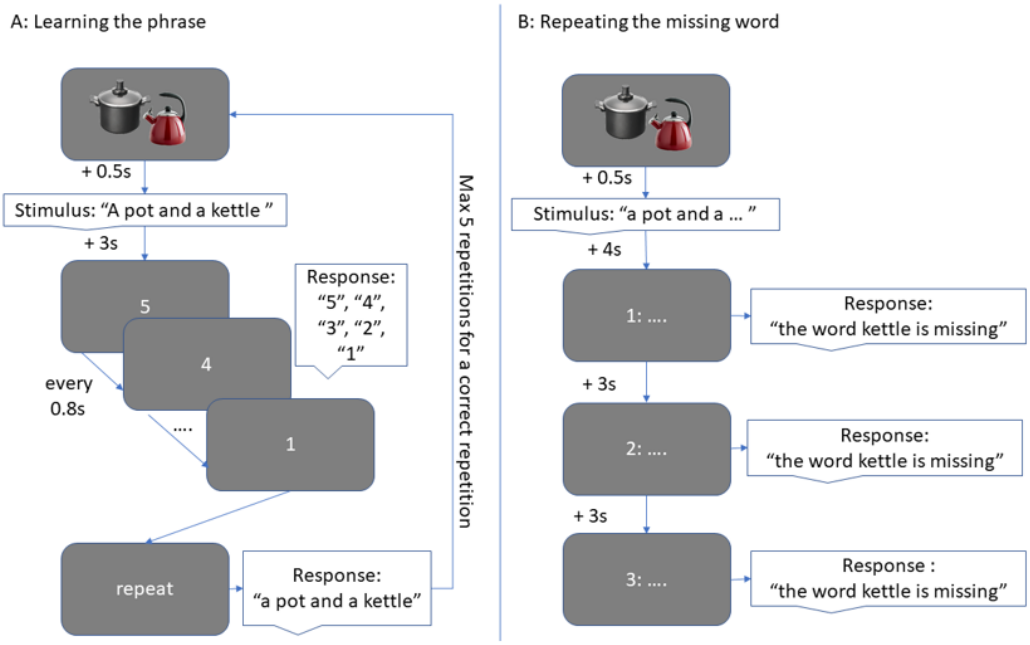

Figure 1: Schematic representation of a trial with a learning phase (left panel) and the test phrase (right panel).

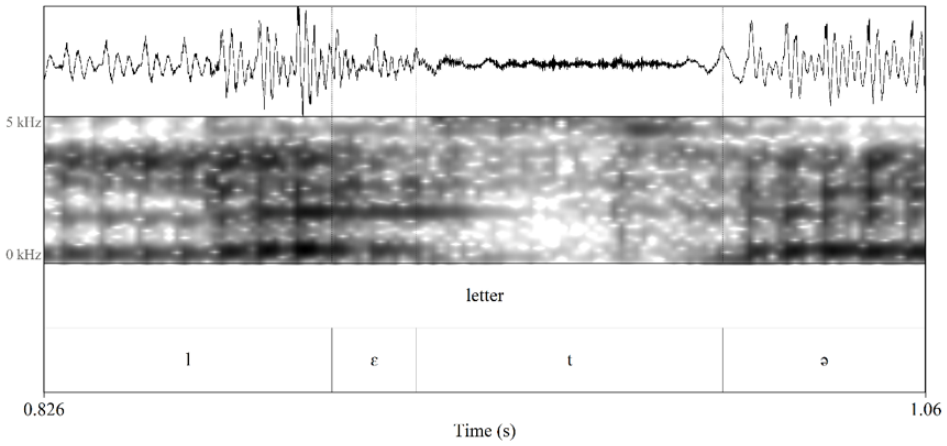

Figure 2: Example of a coded utterance with oscillogram and spectrogram for the item letter, in which no full closure is achieved for the /t/.

3.5ms with a standard error of 3. With an estimate of 18 degrees of freedom, the critical *t*-value is 2.11, meaning that the strongest elongation by a double letter in the confidence interval is 9.8ms. This is roughly 10% increase in duration (singletons are on average 101ms long), which is much less than the elongations of about 30-50% of elongation reported in the studies by Bassetti and colleagues (Bassetti, 2017; Bassetti et al., 2020). That indicates that the failure to find an effect is not due to a lack of power.

The variability by items was further investigated by examining the random slopes. Figure 4 provides a histogram of the item random slope for letter quantity, showing two items with a strong effect, it turns out that the two items that generate a clear effect in the expected direction are *broccoli* vs. *crocodile* and *opposed* vs. *proposed*. Importantly, in both cases, the double-letter word has a cognate in Maltese. With these items removed, the mean consonant duration for single- and double-letter items are nearly identical (single letters: 101.7ms, double letters: 102.0ms). It is worthwhile to consider whether the effect is moderated by speaker properties. We tested three potential variables, the amount of English spoken within the first 6 years of life, the self-estimated proficiency in written and spoken English. Amongst themselves, these were only moderately correlated (early English-spoken proficiency: $r = 0.31$, early English-written proficiency $r = 0.24$, spoken proficiency – written proficiency: $r = 0.54$), so that there are no issues with collinearity. The variables were normalized (i.e., *z*-scored) for this analysis (as suggested by Baayen et al. (2008)), so that the regression weight for letter quantity is representative for the values observed in the study. Table 2 shows the outcome of the analysis, with no moderation of the effect of quantity by proficiency.

4 Discussion

The current data show that Maltese speakers of English do not routinely use their L1 orthography-to-phonology rules when speaking English, unlike the Italian listeners in the studies of Bassetti and colleagues (Bassetti, 2017; Bassetti et al., 2020). As indicated above, the lack of an effect is not due to a lack of power. In the earlier studies, effects of 30-50% elongation were reported. Such effect would have been clearly significant with the current design, the confidence interval observed in this study makes effects of more than 10% unlikely. How do we explain the difference between the results with Maltese learners here and those of Bassetti et al. (Bassetti et al., 2020; Sokolović-Perović et al., 2020)? One obvious explanation would be that the Maltese group is more proficient in English than the Italian learners tested by Bassetti and colleagues. However, this explanation has problems accounting for some aspects of the data. First of all, while Bassetti et al. (2020) found effects of proficiency, the contrast between participant groups in their study was drastic, comparing high-school children at around 17 years of age, who were "studying English for 3 hours a week as a compulsory school subject, using British English textbooks" (Bassetti et al., 2020) with Italian-English bilinguals that had lived in England for on average more than six years. This massive difference in language experience lowered the effect of orthography by only about one third (from 53% to 33% longer consonants for double letters). Based on this finding, it is difficult to see how much proficiency would be needed to push this effect down to zero. Moreover, within the Maltese group, there were some proficiency differences, but those did not influence the results. As such, it seems that Maltese English is acquired as a community language, and that even the more Maltese-dominant children get "pulled along" by those who use English more regularly (the use of English outside formal settings is highly variable in the Maltese community, see Vella (2013)). Since proficiency differences are hence unlikely to explain the differences between the results with Maltese and Italian learners, the best remaining explanation is that effects of L1 orthography on L2 phonology may be dependent on how the L2 is acquired. If acquisition is initially strongly based on the written modality, orthographic influences of the L1 are clearly observed and persist even when learners are immersed in the L2 for many years (Bassetti, 2017; Bassetti et al., 2020; Bassetti et al., 2018). However, if language acquisition early on includes spontaneous usage, as in Malta, such effects are apparently not observed, at least not consistently across items. As a consequence, the current data indicate that effects of orthography are not due to an automatic activation of orthographic knowledge, since this account would predict that the effect should also be found in Maltese learners of English. Instead, it is more likely that the effect comes about as early word learning for Italian learners is highly dependent on reading experience, which is not necessarily the case for Maltese learners. The current data, therefore, support the assumption that the processing of spoken language can proceed without the automatic activation of orthography (Hervais-Adelman et al., 2021; Mitterer et al., 2015; Viebahn et al., 2018). This does not mean that learning to read does not influence the processing of spoken language (Mishra et al., 2012), but that those influences may be mediated by learning rather than by an automatic activation of orthography during speech processing. In this context, it is useful to note that item effects were strong in the current study but negligible in the study of Bassetti and colleagues, with

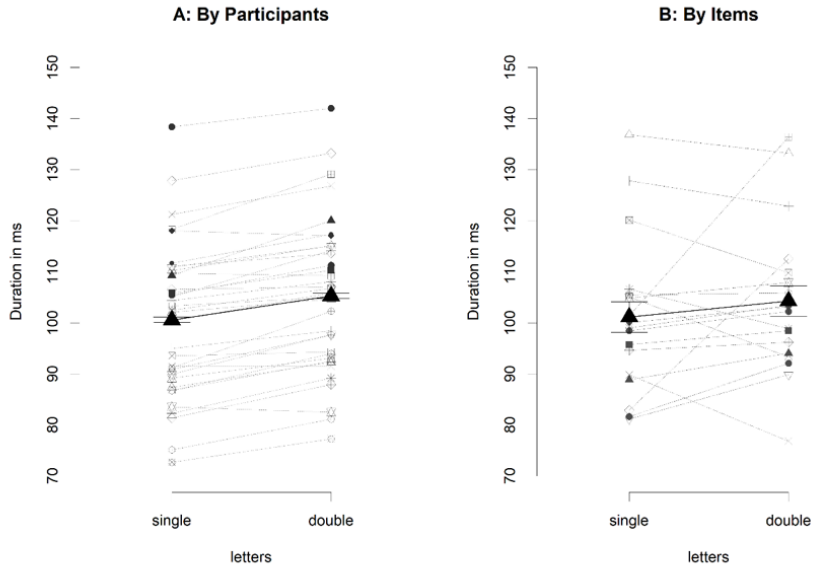

Figure 3: Mean duration for single-letter and double-letter words by participants (Panel A) and items (Panel B). The larger triangles represent the grand mean and the respective standard error for within-participant (or within-item) designs following Morey (2008).

Model	$B_{letters}$(SE)	t	df	p	Model comparison
Random intercepts only	4.1 (0.6)	7.226	3152	$< .001$	
+Random slope by item	3.5 (3.0)	1.173	18	0.256	$X^2(2) = 421.0$, $p < .00$
+Random slope by speaker	3.5 (3.0)	1.162	18	0.261	$X^2(2) = 0.8$, $p = 0.0670$

Table 1: Effect of the number of letters in different linear mixed-effects models with different random-effect structures.

	B(SE)	t(df)	p
Intercept	101.309 (4.279)	23.674 (34)	< 0.001
has2letters	3.459 (2.977)	1.162 (18)	0.26
ProficiencySpeaking	2.639 (2.832)	0.932 (35)	0.358
ProficiencyWriting	-1.707 (2.895)	-0.589 (35)	0.559
earlyEnglishUsage	1.856 (2.573)	0.721 (35)	0.476
has2letters : ProficiencySpeaking	-0.513 (0.66)	-0.777 (34)	0.443
has2letters : ProficiencyWriting	0.962 (0.687)	1.401 (36)	0.17
has2letters : earlyEnglishUsage	-0.725 (0.603)	-1.202 (35)	0.237

Table 2: Output of the linear mixed-effects model with English proficiency measures as co-variates.

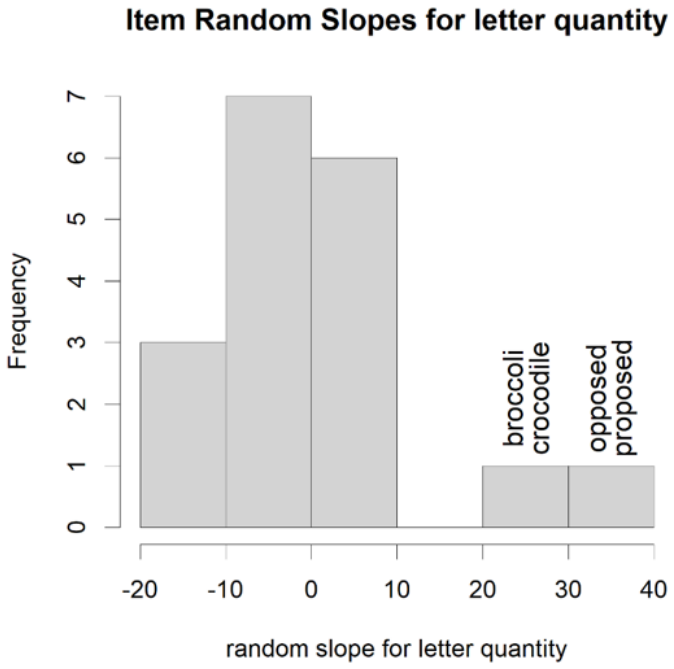

Figure 4: Histogram of the random slopes of the letter-quantity effects for the 18 item pairs.

item and participant-analysis always agreeing with each other. Moreover, random slopes in Bassetti et al. (2020) over item pairs were taken out of the model since they were not supported by the data. This indicates that the large item variability found here may also be specific to the Maltese situation. There are two things to consider in this context. First, the data structure here replicates closely those "made-up" data structures in publications that argue that participant and item variability must be taken into account (Barr et al., 2013; Clark, 1973). On those made-up data sets, one item usually has a strong effect while all others do not. Because each participant is exposed to this item, participants' means are likely to reflect an effect and all participants show an effect quite consistently. While such scenarios may seem unlikely and the examples provided by, for instance, Clark (1973) may seem exaggerated, the current study shows that such examples are in fact realistic and underscore the need to consider both participant and item variability in a statistical analysis (Westfall et al., 2014). Secondly, we find an effect on two items for which the double letter parts of the pairs have cognates in Maltese. This effect might potentially be explained by current theories on lexical selection in bilinguals. There is a general consensus that words are activated non-selectively in bilingual speakers (Dijkstra et al., 2002; Lauro et al., 2017), that is, when trying to name an apple, a Maltese-English bilingual are likely to also activate *tuffieħa*, the respective word in Maltese. It is likely that this co-activation may then influence production, especially as the claim that bilingualism leads to an advantage in conflict monitoring remains controversial (de Bruin et al., 2015; Duñabeitia et al., 2014; Lowe et al., 2021). Given that this is a post-hoc observation, however, we need to be careful to draw strong conclusions, but these data at least suggest that the influence of the L1 on the L2 may be partly moderated by cognate status. This is an avenue for further research. To summarize, the current study tested the influence of L1 orthography on the phonological processing of the L2 in a situation in which the acquisition of the L2 is accompanied by spontaneous usage outside the classroom from early on. In contrast to earlier reports that L1 orthography influences the acquisition of an L2, no such effect was found here. It is therefore argued that orthographic influences in L2 acquisition are not automatic but may be tied to the type of acquisition of the L2.

References

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects

for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Bakker, I., Takashima, A., van Hell, J. G., Janzen, G. & McQueen, J. M. (2014). Competition from unseen or unheard novel words: Lexical consolidation across modalities. *Journal of Memory and Language*, *73*, 116–130.

Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

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., Mairano, P., Masterson, J. & Cerni, T. (2020). Effects of orthographic forms on second language speech production and phonological awareness, with consideration of speaker-level predictors. *Language Learning*, *70*(4), 1218–1256.

Bassetti, B., Sokolović-Perović, M., Mairano, P. & Cerni, T. (2018). Orthography-induced length contrasts in the second language phonological systems of l2 speakers of English: Evidence from minimal pairs. *Language and Speech*, *61*(4), 577–597.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 335–359.

Cutler, A. & Davis, C. (2012). An orthographic effect in phoneme processing, and its limitations. *Frontiers in Psychology, 0*.

de Bruin, A., Treccani, B. & Della Sala, S. (2015). Cognitive advantage in bilingualism: An example of publication bias? *Psychological Science*, *26*(1), 99–107.

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J. & Cohen, L. (2010). How learning to read changes the cortical networks for vision and language. *Science*, *330*(6009), 1359–1364.

Dijkstra, T. & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197.

Duñabeitia, J. A., Hernández, J. A., Antón, E., Macizo, P., Estévez, A., Fuentes, L. J. & Carreiras, M. (2014). *The inhibitory advantage in bilingual children revisited* (No. *3*).

Hayes-Harb, R. & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297–326.

Hayes-Harb, R., Brown, K. & Smith, B. L. (2018). Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and Speech*, *61*(4), 547–564.

Hervais-Adelman, A., Kumar, U., Mishra, R. K., Tripathi, V. N., Guleria, A., Singh, J. P. & Huettig, F. (2021). How does literacy affect speech processing? not by enhancing cortical responses to speech, but by promoting connectivity of acoustic-phonetic and graphomotor cortices. *BioRxiv*.

Heuven, W. J. B. v., Mandera, P., Keuleers, E. & Brysbaert, M. (2014). *Subtlex-UK: A new and improved word frequency database for British English* (Quarterly Journal of Experimental Psychology).

Il-Kunsill Nazzjonali tal-Ilsien Malti. (2021). *Il-qagħda tal-ilsien Malti: Stħarriġ nazzjonali 2021*. University of Malta.

Kolinsky, R., Navas, A. L., de Paula, V., F., R. d. B., N., d. M. B., L., B. S. & Serniclaes, W. (2021). The impact of alphabetic literacy on the perception of speech sounds. *Cognition*, *213*, 10468.

Kuznetsova, A., Brockhoff, P. B., Christensen, R. H. B. & Jensen, S. P. (2020). *lmerTest: Tests in Linear Mixed Effects Models (3.1–3)*.

Lauro, J. & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, *92*, 217–233.

Lowe, C. J., Cho, I., Goldsmith, S. F. & Morton, J. B. (2021). The bilingual advantage in children's executive functioning is not related to language status: A meta-analytic review. *Psychological Science*, *32*(7), 1115–1146.

Mishra, R. K., Singh, N., Pandey, A. & Huettig, F. (2012). Spoken language-mediated anticipatory eye-movements are modulated by reading ability—evidence from Indian low and high literates. *Journal of Eye Movement Research*, *5*.

Mitterer, H. (2018). Not all geminates are created equal: Evidence from Maltese glottal consonants. *Journal of Phonetics*, *66*, 28–44.

Mitterer, H. & Ernestus, M. (2006). Listeners recover /t/s that speakers lenite: Evidence from /t/-lenition in dutch. *Journal of Phonetics*, *34*, 73–103.

Mitterer, H. & Reinisch, E. (2015). Letters don't matter: No effect of orthography on the perception of conversational speech. *Journal of Memory and Language*, *85*, 116–134.

Morais, J. (2021). The phoneme: A conceptual heritage from alphabetic literacy. *Cognition*, *213*, 10474.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to cousineau. *Tutorials in Quantitative Methods for Psychology*, *4*, 61–64.

Pallier, C., Bosch, L. & Sebastian-Gallés, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, *64*(3), B9–B17.

Pattamadilok, C., Perre, L., Dufau, S. & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience*, *21*(1), 169–179.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Pitt, M. A., Dilley, L. & Tat, M. (2011). Exploring the role of exposure frequency in recognizing pronunciation variants. *Journal of Phonetics*, *39*(3), 304–311.

Roettger, T. B., Winter, B., Grawunder, S., Kirby, J. & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, *43*, 11–25.

Sadakata, M., Shingai, M., Sulpizio, S., Brandmeyer, A. & Sekiyama, K. (2014). Language specific listening of Japanese geminate consonants: A cross-linguistic study. *Frontiers in Psychology*, *5*, 1422.

Shea, C. (2017). L1 English / l2 Spanish: Orthography–phonology activation without contrasts. *Second Language Research*, *33*(2), 207–232.

Sokolović-Perović, M., Bassetti, B. & Dillon, S. (2020). English orthographic forms affect l2 English speech production in native users of a non-alphabetic writing system. *Bilingualism: Language and Cognition*, *23*(3), 591–601.

Strunk, J., Schiel, F. & Seifart, F. (2014). Untrained forced alignment of transcriptions and audio for language documentation corpora using webmaus. In K. C. N. Calzolari, T. Declerck, H. Loftsson, B. Maegaard, J. Mariani, A. Moreno, J. Odijk & S. Piperidis (Eds.), *Proceedings of the ninth international conference on language resources and evaluation* (pp. 3940–3947). European Language Resources Association (ELRA).

Team, R. C. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Thusat, J., Anderson, E., Davis, S., Ferris, M., Javed, A., Laughlin, A., McFarland, C., Sangsiri, R., Sinclair, J., Vastalo, V., Whelan, W. & Wrubel, J. (2009). Maltese English and the nativization phase of the dynamic model. *English Today*, *25*(2), 25–32.

Vella, A. (2013). Languages and language varieties in Malta. *International Journal of Bilingual Education and Bilingualism*, *16*(5), 532–552.

Viebahn, M. C., McQueen, J. M., Ernestus, M., Frauenfelder, U. H. & B"urki, A. (2018). How much does orthography influence the processing of reduced word forms? evidence from novel-word learning about French schwa deletion. *Quarterly Journal of Experimental Psychology*, *71*, 2378–2394.

Westfall, J., Kenny, D. A. & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology*, *143*(5), 2020–2045.

Young-Scholten, M. (2002). Orthographic input in l2 phonological development. In T. P. P. Burmester & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of henning wode* (pp. 263–279). Wissenschaftlicher Verlag Trier.

Ziegler, J. C. & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review*, *5*(4), 683–689.

Appendices

		Frequency in			
Item	**Sentence**	**All BBC programs**	**Preschool programs**	**School-kids programs**	**BNC (written)**
Acute	We both felt an acute pain.	3.66	2.23	3.42	4.36
Accuse	I don't want to accuse anyone.	3.58	2.23	3.26	3.56
Document	Please look at this document here.	4.18	2.53	3.19	4.72
Occupy	How do you occupy your time?	3.55	2.23	3.07	4.02
Nicaragua	I have friends from Peru and Nicaragua.	2.58	2.23	2.47	3.68
Piccadilly	An expensive shop near Piccadilly Circus.	3.29	2.23	2.34	3.50
Weapon	This is a very ancient weapon.	4.29	2.93	4.59	4.29
pepper	Salt and pepper, oil and vinegar.	4.38	4.68	4.21	3.98
Rapidly	Nowadays the world is rapidly changing.	4.01	2.53	3.51	4.66
Happily	A group of happily married couples.	4.11	4.41	4.18	4.25
Copy	Could I please have a copy?	4.38	4.66	4.62	4.78
Floppy	Take a CD or a floppy.	3.31	4.18	3.72	3.71
Latin	She studies Greek and Latin poetry.	4.20	3.57	4.26	4.44
Chatting	She is chatting on the phone.	4.03	3.97	4.00	3.79
City	They both work as city lawyers.	5.40	5.30	5.03	5.36
Kitty	My god, a Hello Kitty room.	4.02	4.23	4.40	3.59
Vitamins	This drink contains vitamins and sugar.	3.35	3.41	3.65	3.68
Littering	No littering, take your litter home.	2.83	2.53	3.21	2.64
Sweater	Jeans and a sweater	3.24	3.63	3.38	3.78
Letter	number and letter	4.85	4.87	4.82	5.13
Broccoli	carrots, peas, and broccoli	3.71	4.09	4.05	3.13
Crocodile	He went to see the Nile crocodile	3.95	4.86	4.52	3.41
Apple	Banana and apple	4.58	5.08	4.62	4.42
Chapel	There was a church and a chapel	4.06	2.23	3.28	4.33
Kettle	A pot and a kettle	4.02	4.37	4.05	3.96
Metal	He used wood and metal	4.63	4.67	4.69	4.66
Mitten	In winter, wear scarf, mitten, and hat	2.43	3.19	2.77	2
Britain	It's rainy in Britain	2.56	3.19	3.01	2.59
Piccolo	He played trombone, piccolo, and trumpet	3.06	2.23	2.34	3.26
Nicotine	Caffeine and nicotine	2.57	2.71	3.37	2.04
Raccoon	A trashcan with raccoon alarm	2.84	3.66	3.26	3
Cocoon	First larva then cocoon	4.25	2.53	2.82	4.92
Proposed	Last Saturday, he proposed	4.3	2.23	3.47	4.42
Opposed	His argument was opposed	4.88	4.83	4.83	4.81
Proper	His attire was proper	4.25	3.41	3.85	4.27
Copper	There was silver, copper, and gold	4.66	4.72	4.61	4.9

Note: Frequencies are zipf-scaled (see Heuven et al. (2014)).

Table 3: The items used in the experiment.