



**Illusory vowels in Spanish-English sequential bilinguals:
Evidence that accurate L2 perception is neither necessary
nor sufficient for accurate L2 production**

Journal:	<i>Second Language Research</i>
Manuscript ID	SLR-18-0044.R3
Manuscript Type:	Original Article
Keywords:	phonotactic constraints, speech perception, speech production, Spanish, English, L2 use, grammatical proficiency, age of L2 acquisition, executive control
Abstract:	<p>Spanish native speakers are known to pronounce onset /sC/ clusters in English with a prothetic vowel, as in <i>esport</i> for <i>sport</i>, due to their native language phonotactic constraints. We assessed whether accurate production of e.g. <i>spi</i> instead of <i>espi</i>, was related to accurate perceptual discrimination of this contrast in L2 speech of Spanish-English sequential bilinguals. A same-different discrimination task in stimulus pairs such as <i>spi-espi</i> assessed speech perception and a phonemic verbal fluency task elicited speech production.</p> <p>Logistic mixed model regressions revealed significant differences in accuracy between the bilinguals and the English monolinguals, although some bilinguals performed within the monolingual range. For the production task, but not for the perception task, bilinguals with more exposure to English and greater grammatical knowledge of English performed significantly more accurately than those with less exposure and lower grammatical knowledge. There was no significant correlation between production accuracy and perception accuracy.</p> <p>Through examining phonotactic constraints, these results expand a growing body of research into single sounds which suggests dissociations between L2 perception and production. In contrast to predictions made by L2 speech models, the findings indicate that accurate L2 perception is neither necessary nor sufficient for accurate L2 production, and instead are interpreted to indicate that the two capacities recruit different executive control mechanisms and are acquired – at least to a certain extent - independently in L2 acquisition.</p>

1
2
3 Illusory vowels in Spanish-English sequential bilinguals: Evidence that accurate L2
4 perception is neither necessary nor sufficient for accurate L2 production
5
6

7 Esther de Leeuw¹, Linnaea Stockall¹,
8 Dimitra Lazaridou-Chatzigoga², Celia Gorba Masip³
9

10
11 ¹Queen Mary University of London
12

13 ²Humboldt-Universität zu Berlin and University of Cambridge
14

15 ³Universitat Autònoma de Barcelona
16
17
18
19

20 **Acknowledgements**
21

22 The authors are particularly grateful to Professor Juana Gil Fernández who welcomed
23 us to Madrid, where we collected most of the data at her *Laboratorio de Fonética* of
24 the University of Madrid, and to Professor Fernando Sánchez Miret at the University
25 of Salamanca. Without their help, we would never have been able to conduct this
26 study. We are also grateful to Annette Zhao for the data analysis, all of the
27 participants who took part in this study, and to the British Academy for funding the
28 project (BA grant SG112797: PI: EdeL; co-investigator: LS).
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Abstract**
4

5 Spanish native speakers are known to pronounce onset /sC/ clusters in English with a
6 prothetic vowel, as in *esport* for *sport*, due to their native language phonotactic
7 constraints. We assessed whether accurate production of e.g. *spi* instead of *espi*, was
8 related to accurate perceptual discrimination of this contrast in L2 speech of Spanish-
9 English sequential bilinguals. A same-different discrimination task in stimulus pairs
10 such as *spi-espi* assessed speech perception and a phonemic verbal fluency task elicited
11 speech production.
12
13
14
15
16
17
18
19
20

21 Logistic mixed model regressions revealed significant differences in accuracy
22 between the bilinguals and the English monolinguals, although some bilinguals
23 performed within the monolingual range. For the production task, but not for the
24 perception task, bilinguals with more exposure to English and greater grammatical
25 knowledge of English performed significantly more accurately than those with less
26 exposure and lower grammatical knowledge. There was no significant correlation
27 between production accuracy and perception accuracy.
28
29
30
31
32
33
34
35
36

37 Through examining phonotactic constraints, these results expand a growing
38 body of research into single sounds which suggests dissociations between L2
39 perception and production. In contrast to predictions made by L2 speech models, the
40 findings indicate that accurate L2 perception is neither necessary nor sufficient for
41 accurate L2 production, and instead are interpreted to indicate that the two capacities
42 recruit different executive control mechanisms and are acquired – at least to a certain
43 extent - independently in L2 acquisition.
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Introduction

While there is some research into bilinguals' perception and production of single sounds, there is little work investigating L2 acquisition of phonotactic constraints, and to our knowledge no work systematically investigating the relationship between perception and production of phonotactic constraints. Our study examined phonotactic constraints, i.e. language specific licit versus illicit sequences of sounds, in the speech perception and production of Spanish-English sequential bilinguals. As mentioned by Dupoux, Kakehi, Hirose, Pallier, & Mehler (1999), "in Spanish, /s/ + consonant clusters are always preceded by a vowel and we have informally heard reports by Spanish speakers who maintain that they hear the vowel [e] preceding English words that begin with an /sC/ cluster" (p. 1568). Accordingly, Spanish learners of English often mispronounce initial /sC/ clusters with a prothetic initial vowel (Hualde, 2005).

The main question of this study was whether accurately *perceiving* the /sC/ cluster in English (and therefore *not* perceiving an illusory preceding [e]) was associated with the accurate *production* of the /sC/ cluster (and therefore *not* pronouncing a preceding vowel in English speech where it is not required) within a group of Spanish-English sequential bilinguals. Furthermore, it was examined whether increased and earlier exposure to the L2, and improved L2 grammatical proficiency as measured through a C-Test (Klein-Braley, 1985; Raatz and Klein-Braley, 1981), would help to improve accurate perception and production of /sC/ clusters. Bilingual data were collected in Spain (Madrid and Salamanca, henceforth referred to as the Spain bilinguals), and in London, UK (henceforth referred to as the UK bilinguals). This study therefore aimed to shed light on how the "rules" of language specific sound sequences are implemented in L2 speech acquisition - when they violate rules in the L1 - and feeds

1
2
3 into a growing body of research which examines the relationship between L2 perception
4 and production.
5
6

7 *L2 speech acquisition models*

8
9
10 Often, when considering L2 speech acquisition, the premise is that problems accurately
11 *producing* L2 speech arise from difficulties in accurately *perceiving* L2 speech. It seems
12 logical that perception abilities would need to be in place *before* accurate production
13 abilities are possible, and that difficulties producing L2 speech would have a perceptual
14 basis, such that incorrect perception would lead to incorrect production (see e.g.
15 Escudero, 2005; and Llisterri, 1995, for an overview of early studies). For example,
16 according to the Speech Learning Model (SLM), it is claimed that those L2 sounds
17 which have a similar (although not identical) counterpart in the L1 are the most difficult
18 to learn to pronounce because a similar sound is more likely to be *perceived* within the
19 same category as that of the L1 (Flege, 1995; Flege, Schirru, & MacKay, 2003). A new
20 category, on the other hand, is more likely to be created for a dissimilar sound
21 (originally termed ‘new’ (Flege, 1987)), which is not perceived to be the same sound.
22 Long-term pronunciation problems are, according to the SLM, more likely in the case
23 of similar sounds than in dissimilar sounds as perceptual equivalence classification
24 prevents experienced L2 learners from producing similar, but not dissimilar sounds
25 (Flege, 1995). This notion of equivalence classification has strong parallels with Kuhl’s
26 Native Language Magnet Theory which suggests that the nearer an L2 sound is to an
27 L1 sound (specifically to its prototype), the more it will be perceptually assimilated to
28 this L1 sound (Kuhl, 2004).
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 The notion of perceptual similarity is also intrinsic to Best’s Perceptual
55 Assimilation Model (PAM) (1995) (see also PAM-2 with regard to L2 acquisition of
56 prosody, So and Best, 2010, 2014). According to this model, the similarity between L1
57
58
59
60

1
2
3 and L2 sounds is based on the perceived resemblance of articulatory gestures used to
4 produce L2 sounds, in comparison to those used to produce the closest L1 sound. As
5 such, perception entails the ability to detect articulatory properties of speech (i.e.,
6 tongue movement, vocal tract size, etc.), suggesting that speech perception and
7 production are aligned. This model has strong parallels with the Motor Theory of
8 speech perception, proposed by Liberman and Mattingly (1985), which claims that “the
9 objects of speech perception are the intended phonetic gestures of the speaker,
10 represented in the brain as invariant motor commands that call for movements of the
11 articulators through certain linguistically significant configurations” (p. 2) (see also
12 Fowler (1996) and Ohala (1996)).

13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

What unifies both the SLM and PAM is that perception is considered to be intrinsically linked to production, and that either perception precedes production (SLM), or that they develop in tandem (PAM). However, as will be discussed, there is a growing body of research examining the perception and production of individual L2 sounds which suggests dissociations between L2 perception and production. Moreover, languages differ not only because of differences in their respective phoneme inventories, but also in their realisation of phonotactic constraints. Indeed, very little is known regarding the acquisition of new permissible sound sequences in the L2 which violate L1 phonotactic constraints. The results from this study therefore inform both PAM and the SLM regarding the assertion that difficulties in producing L2 speech are linked to difficulties in perception.

Research into perception and production of L2 speech

Some previous research into the question of whether accurate perception leads to accurate production in L2 acquisition suggests that perception and production are

1
2
3 intertwined. For example, it has been found that non-native English speakers who were
4 considered to be experienced L2 speakers (who lived in the US for approximately 7
5 years) produced and perceived English vowels more accurately than relatively
6 inexperienced non-native participants (who lived in the US for under 1 year) (Flege,
7 Bohn, & Jang, 1997). Both production and perception accuracy appeared to depend on
8 the perceived relation between English vowels and vowels in the participants' specific
9 L1 inventory. Such findings provide support for the SLM in suggesting that perceptual
10 similarity between the L1 and L2 is linked to speech production in the L2.
11
12
13
14
15
16
17
18
19
20
21

22 However, other research has not substantiated a clear dependency between
23 perception and production of L2 speech, and it has been found that some learners are
24 able to produce differences between L2 sounds that they cannot perceive (Baker &
25 Trofimovich, 2006; Beach, Burnham, & Kitamura, 2001; Kartushina & Frauenfelder,
26 2014; Sheldon & Strange, 1982; Zampini, 1998). For example, early results from
27 Sheldon and Strange (1985) revealed that native Japanese speakers who had learned to
28 produce /l/ and /r/ appropriately still made perception errors in perceiving the contrast.
29 The authors summarised that “perceptual mastery of a foreign contrast does not
30 necessarily precede adult learners' ability to produce acceptable tokens of the
31 contrasting phonemes, and may, in fact, sometimes lag behind production mastery” (p.
32 254). Similar results from Zampini (1998) into perception and production of voiced and
33 voiceless plosives in Spanish by English native speakers confirmed that L2 production
34 may in some cases precede perception and that learners do not begin to adjust
35 perceptual boundaries until they have attained accurate production categories. Likewise,
36 research examining the production of the three-way voicing difference in Thai bilabial
37 stops by Greek-English bilinguals and English monolinguals has indicated that those
38 bilinguals who “exaggerated” the voicing difference in their Thai speech production
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 also best perceived those differences when listening (Beach et al., 2001). The study
4
5 concluded that production profiles are an important adjunct to the assessment of
6
7 bilingual speakers, and have important implications for the interface between
8
9 perception and production.
10

11
12 The current investigation expands on such findings by examining whether
13
14 accurately perceiving syllable onset /sC/ clusters in English (although such clusters are
15
16 not allowed in Spanish) would likewise help to *avoid* production of a prothetic vowel
17
18 before such clusters in Spanish-English sequential bilinguals, e.g. would those
19
20 bilinguals who *perceive* a difference between stimuli such as *spi* and *espi* be less likely
21
22 to *produce* a prothetic vowel in a phonemic verbal fluency task eliciting onset /sC/
23
24 clusters, whilst bearing in mind individual differences in English L2 exposure, age of
25
26 acquisition, and grammatical proficiency in English?
27
28
29

30
31 This question is particularly interesting when viewed in relation to executive
32
33 control, which has been linked to bilingualism and L2 acquisition (Bialystok, 2009,
34
35 2017; de Leeuw and Bogulski, 2016). Executive control processes are hypothesised to
36
37 supervise “the selection, initiation, execution, and termination” of multiple task
38
39 performance (Rubinstein et al., 2001: 763). Frequently postulated executive
40
41 mechanisms underlying the overall system are the (1) *shifting* of mental sets, (2)
42
43 *monitoring* and updating of working memory representations, and (3) *inhibiting*
44
45 competing stimuli (Miyake et al., 2000). It may be that speech production recruits
46
47 different executive control functions than speech perception, as the former is inherently
48
49 more active than the latter, i.e. when producing speech, the articulators must be co-
50
51 ordinated to match the intended output, whilst when perceiving speech, there is no
52
53 output. Different underlying executive control processes may have influenced the
54
55
56
57
58
59
60

1
2
3 findings from the aforementioned studies, and are therefore also considered to be
4
5 potentially relevant in relation to the present research.
6

7 8 **Research into perception of L2 phonotactic constraints** 9

10 In seminal research on the perception of illusory vowels by Japanese native speakers,
11
12 it was found in four experiments comparing French and Japanese hearers that
13
14 phonotactic properties of Japanese induced native Japanese listeners to perceive
15
16 “illusory” vowels inside consonant clusters in VCCV stimuli, where there were in fact
17
18 no vowels, as Japanese does not allow such word medial consonant clusters (e.g. *ebzo*
19
20 was perceived as *ebuzo*) (Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; see also
21
22 Harris, 1983; Itô and Mester, 1995).
23
24

25
26 In contrast, French native speakers had no difficulty perceiving this contrast
27
28 (French allows such consonant clusters, e.g. ‘observer’ /obzerve/), However, the French
29
30 native speakers had difficulties discriminating items that differed in vowel length (e.g.
31
32 *ebuzo* vs. *ebuuzo*), while Japanese native speakers had no difficulty discriminating such
33
34 stimuli (vowel length is contrastive in Japanese but not in French). The conclusion from
35
36 this series of studies was that models of speech perception should be revised to account
37
38 for language specific phonotactic constraints. However, although the research found
39
40 significant differences between the Japanese and French native speakers, it was not
41
42 further investigated whether, for example, *some* of the Japanese native speakers might
43
44 have been able to acquire the French phonotactic constraints, nor whether accurate
45
46 perception of the French phonotactic constraints (i.e. licit consonant clusters) may have
47
48 been associated with accurate *production* of the consonant clusters.
49
50
51
52

53
54 In more recent research, several populations of Japanese-Brazilian bilinguals
55
56 were tested in order to determine whether the perceived illusory vowel differed between
57
58 bilingual groups (Parlato-Oliveira, Christophe, Hirose, & Dupoux, 2010). The results
59
60

1
2
3 from an explicit metalinguistic task showed that, as expected based on the phonotactics
4 of each language, monolingual Japanese participants mainly perceived the stimuli as
5 containing the epenthetic vowel /u/ (e.g. in items like *ebna* and *ebuna*) while
6 monolingual Brazilian participants mainly perceived the epenthetic vowel /i/ (e.g. in
7 items like *ebna* and *ebina*). Interestingly, first-generation immigrants, who had
8 Brazilian Portuguese as an L2, appeared to behave like Japanese monolinguals,
9 suggesting that their immersion within a Brazilian-speaking country, which started in
10 adulthood, “did not induce them to significantly modify their phonological settings” (p.
11 3742). In sharp contrast, second-generation immigrants behaved like Brazilian
12 monolinguals and were more likely to perceive the epenthetic vowel /i/. Moreover, in
13 an implicit perceptual task, it was found that all three groups of bilinguals behaved
14 exactly like Brazilian monolinguals, “showing high confusability between *ebna* and
15 *ebina*, while the distinction between *ebna* and *ebuna* proved to be easy to for them” (p.
16 3745). Parlato-Oliveira et al. (2010) summarised that an effect of L2 exposure was most
17 clearly observed in the explicit task but not in the implicit task and suggested that L2
18 experience is more likely to affect explicit, or metalinguistic, perceptual tasks.

19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40 Another study, specifically examining Spanish–English sequential bilinguals,
41 similarly found differences between bilinguals with regard to their capacity to perceive
42 phonotactic constraints in the L2 (Carlson et al., 2016), noting that English not only
43 permits #sC, the prohibited sequence in Spanish, but it favours this sequence over
44 #VsC, the preferred Spanish repair. In both the identification and discrimination tasks,
45 the Spanish–English bilinguals exhibited weaker perceptual repair effects relative to
46 Spanish monolinguals, who came from Cuetos et al. (2011). This was true even for the
47 bilinguals who were considered to be dominant in Spanish; however, the perceptual
48 repair effects were least pronounced for English-dominant bilinguals, suggesting that
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 the acquisition of English influenced the perception of phonotactic constraints in
4 Spanish (Carlson et al., 2016). These results support related research which has
5
6 similarly found that /sC/ onset clusters in English activate Spanish phonotactic
7
8 constraints in Spanish–English bilinguals (Freeman et al., 2016) and most recent
9
10 research suggesting that even the perception of L1 Brazilian Portuguese phonotactic
11
12 constraints are modified by the acquisition of English as an L2 (Cabrelli et al., 2019).
13
14
15

16
17 However, in none of these studies was speech production investigated. Indeed,
18
19 as it is quite often accepted that perception provides the foundation for production, it
20
21 might be – potentially prematurely - interpreted that such results regarding perception
22
23 provide insight into what those same participants might produce in their L2 speech.
24
25 However, based on the previously discussed research which shows weak to not existent
26
27 relationships between L2 speech perception and production (see also Kartushina &
28
29 Frauenfelder, 2014), we cannot assume that such studies into the perception of L2 sound
30
31 sequences provide direct insight into the capacity of L2 learners to produce those same
32
33 L2 sound sequences.
34
35
36

37
38 In the present study, by focusing on a single phonotactic feature, we were able
39
40 to systematically examine the degree to which perception and production of the
41
42 prothetic vowel were associated with one another in Spanish native speakers with
43
44 English as an L2. The research by Hallé et al. (2008) confirms that the phenomenon of
45
46 an illusory prothetic vowel preceding sibilant consonant clusters in Spanish native
47
48 speakers is real, but the study does not actually employ a straightforward, simple,
49
50 auditory discrimination task along the lines of Dupoux et al. (1999), and, moreover,
51
52 Hallé et al.'s (2008) research does not compare perception and production in the same
53
54 experiment and same speakers, which is the aim of the current study.
55
56
57
58
59
60

Methodology

The study comprised two main tasks. In Task I, the perception of a prothetic vowel in minimal pairs was examined in the same two participant groups: English monolinguals, and Spanish-English bilinguals. In Task II, the production of prothetic vowels preceding /sC/ consonant clusters in onset position was assessed to determine whether there was an association between perceptual and production accuracy. Additionally, predictor variables were examined which might have played a role in determining production and perception in the sequential Spanish–English bilinguals, i.e. amount of daily exposure to L2, as self-assessed; age of L2 acquisition; and grammatical proficiency in the L2, as measured by a written C-Test (see Appendix I for questions and Appendix II for information). Grammatical proficiency was thought to be an interesting variable to examine because it may have been that bilinguals with higher grammatical knowledge were also more aware of the phonotactic rules of English, which allow /sC/ clusters in onset position (see Kivistö-de Souza, 2015 and Morales Pech and Izquierdo, 2011 for related research). This metalinguistic awareness may have improved their English speech production, and, potentially also their speech perception, if the two domains are related.

Participants

In total, 42 participants were examined. Eight participants were native English speakers, whose data were collected in London, United Kingdom. These participants listed no fluency in other languages, although some did have knowledge of other languages due to school education and vacations abroad. Seven participants were native Spanish speakers with English as an L2 who were recorded in London, United Kingdom. These

participants were either working in London, or were students on a year abroad, and had lived in the UK for between 3 and 16 months. The remaining 27 participants were also all native Spanish speakers whose data were collected in Spain. These participants reported that they had learned English as an L2 in Spain, and used English to varying extents in their daily lives. Many of the Spain bilinguals had also spent time in an English speaking country with the average time being 7 months, and both the United Kingdom and Spain bilinguals were therefore collapsed into one group. All participants were between 40 and 18 years of age and most were either university students, researchers, administrators, or academic faculty. All of the bilinguals considered Spanish to be their native language, and English to be the language in which they had the greatest proficiency after Spanish. The mean amount of daily English use of the bilinguals, which was self-assessed, was 30% with a standard deviation of 18%; the mean grammatical proficiency score was 38% with a standard deviation of 16%; and the mean age of English acquisition was 8 years of age with a standard deviation of 5.4 years.

	English monolinguals	Spanish-English bilinguals
Number of participants	8	34
Average age at time of experiment, standard deviation in brackets	23.6 (6.3)	27.15 (11.9)
Females : males	4:4	24:10

Table 1: Participant background information, see Appendix II.

It is relevant to note that more data were collected initially than included in the analysis. This is because a number of English monolinguals turned out to have high proficiencies in additional languages; and a number of participants in Spain cited that they were bilinguals in either Catalan or Galician, and sometimes that they did not consider themselves to be native speakers of Spanish, which we had not expected when

1
2
3 originally conducting the experiment. To ensure the reliability of our study, we
4
5 therefore excluded 19 participants from the analysis.
6
7
8
9

10 *Data collection procedure*

11
12 The same procedure was followed in both Madrid, Salamanca and London, and in each
13
14 setting the recordings were conducted in a sound attenuated room at respectively the
15
16 *Laboratorio de Fonética* of the *Centro de Ciencias Humanas y Sociales*, the University
17
18 of Salamanca, or the Phonetics Laboratory of Queen Mary University of London.
19
20 Recordings were conducted entirely in English in all cases in order to ensure an English
21
22 mode (Grosjean, 1998), i.e. this investigation tested L2 English perception and
23
24 production; however, particularly with regard to the perception task, as the tokens were
25
26 non-words, the participants' accuracy also provides insight into perceptual capacities
27
28 more generally (Carlson et al., 2016) and therefore the results also have ramifications
29
30 on the mechanisms of L1 perception (Cabrelli et al., 2019) and its plasticity in the
31
32 context of bilingualism (de Leeuw and Celata, 2019).
33
34
35
36
37

38 Participants entered into the sound attenuated room where the outline of the
39
40 procedure was described to them, and then filled in the participant consent form.
41
42 Thereafter, the research investigator (either Edel or DLC) filled in an adapted version
43
44 of the MPI Language Background Questionnaire (Gullberg and Indefrey, 2003), by
45
46 asking the participant the questions on the form (see Appendix I). In general, the
47
48 completion of this form took between 15 and 20 minutes. Participants then completed
49
50 the C-test, for which they were given a maximum of 10 minutes (see Appendix I), the
51
52 onset elicitation task (i.e. the phonemic verbal fluency task), and then the perceptual
53
54 discrimination task. Although the perceptual discrimination task was completed last to
55
56 ensure /sC/ clusters were not made salient to participants before the production task,
57
58
59
60

1
2
3 the perception task is presented initially in this manuscript, and the production task
4
5 thereafter. In total, data elicitation took no more than 45 minutes.
6
7
8
9

10 **Task 1: Perception Task Elicitation**

11
12 The purpose of this task was to determine the extent to which Spanish native speakers
13
14 with English as an L2 perceived a difference between stimuli such as *spi* and *espi* (see
15
16 Table 3 for the full set of stimuli). To assess this question, participants were invited to
17
18 take part in a same-different perceptual discrimination task (AX task). Participants wore
19
20 Sennheiser over ear headphones, and were presented with two non-words directly after
21
22 one another using DMDX software (Forster and Forster, 2003), which also recorded
23
24 participant responses, e.g. (1) *spi* and *spu* or (2) *spi* and *espi* or (3) *spi* and *spi* (see Table
25
26 2). For ease of exposition, we will continue to use these tokens as the model stimuli. In
27
28 line with what we know about Spanish (Hallé et al., 2008; Hualde, 2005), we expected
29
30 the Spanish-English bilinguals to perceive a difference between the stimuli in (2) less
31
32 accurately than English monolinguals, but we did not expect a difference between the
33
34 Spanish-English bilinguals and English monolinguals in (1) and (3) (see Table 2). We
35
36 did not expect to find a difference in Condition (3) because the items presented in the
37
38 AX discrimination task were the same in these trials, and we did not expect a difference
39
40 in Condition (1) because /i/ and /u/ are contrastive in both English (e.g. *beat* versus *boot*)
41
42 and Spanish (e.g. *sí* versus *su*). However, as English favours /sC/ onset clusters over
43
44 /VsC/ onsets (Carlson et al., 2016), we also expected that English monolinguals might
45
46 have difficulties in Condition 2, albeit less than the Spanish native speakers. There were
47
48 approximately twice as many stimuli pairs for Condition 3 as for the other two
49
50 conditions in order to ensure an equal number of same stimuli pairs (Condition 3) versus
51
52 different stimuli pairs (Conditions 1 & 2), see Table 2, and 197 trials in total.
53
54
55
56
57
58
59
60

	Contrast example	Nr. of trials	Prediction
Condition 1	<i>spi-spu</i>	49	No difference between Spanish–English bilinguals and English monolinguals
Condition 2	<i>spi-espi</i>	49	Spanish–English bilinguals will exhibit less accuracy than English monolinguals
Condition 3	<i>spi-spi</i>	99	No difference between Spanish–English bilinguals and English monolinguals

Table 2: Conditions in perception task and general predictions

Stimuli

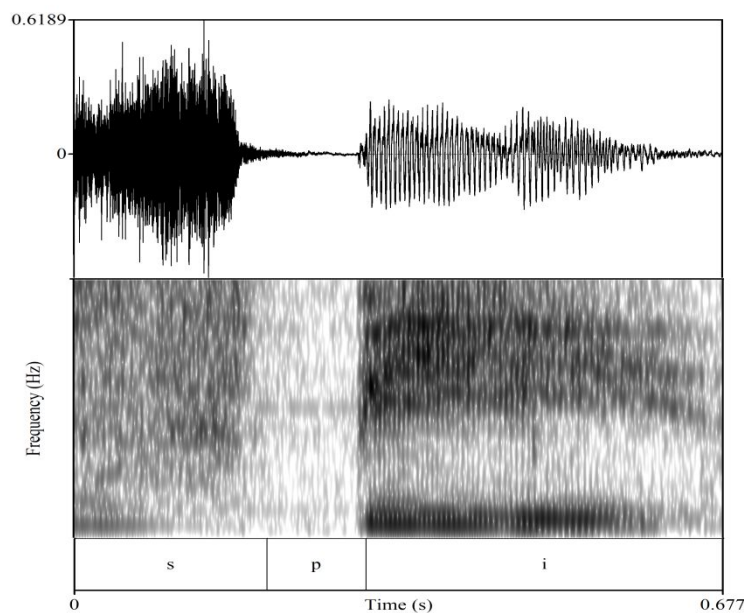
Stimuli were produced by the first author, who is a native speaker of English. The author produced all of the *espi* tokens first, listed in column b in Table 3.

<i>a. Edited token with no prothetic vowel</i>	<i>b. Original token with prothetic vowel</i>
<i>spi</i>	<i>espi</i>
<i>spu</i>	<i>espu</i>
<i>sli</i>	<i>esli</i>
<i>slu</i>	<i>eslu</i>
<i>ski</i>	<i>eski</i>
<i>sku</i>	<i>esku</i>
<i>smi</i>	<i>esmi</i>
<i>smu</i>	<i>esmu</i>
<i>sni</i>	<i>esni</i>
<i>snu</i>	<i>esnu</i>
<i>sti</i>	<i>esti</i>
<i>stu</i>	<i>estu</i>

Table 3: Tokens used in same-different discrimination task.

Thereafter, the initial vowel was edited from the stimulus, such that a new stimulus was created: *spi*. These consonant clusters (i.e. /sp/, /st/, /sk/, /sm/, /sn/, /sl/) were chosen because they are the possible /sC/ clusters in English, and therefore also the same onsets which were elicited in the phonemic verbal fluency task.

1
2
3 Delimiting the initial vowel from the following sibilant was not difficult based
4
5 on analysis of the spectrogram and waveform. The boundary was marked where the
6
7 periodicity of the vowel ended and the aperiodicity of the fricative commenced.
8
9 Therefore, these stimuli (a. *spi* and b. *espi*) were exactly the same except for the vowel
10
11 in b (see Figure 1 and 2 of /spi/ and /espi/). Note that it may have been that formant
12
13 remnants from the initial vowel remained in the sibilant due to coarticulation effects,
14
15 although all actual periodicity from the vowel was cut out (see Fig. 1 & 2), which might
16
17 potentially have led to more perceived instances of /e/ on the part of the bilinguals;
18
19 however, all bilinguals heard the same items, and therefore this would have affected
20
21 their discrimination of *spi* versus *espi* equally.
22
23
24
25
26



47 **Figure 1: Token of *spi* used in perception task.**
48
49
50
51
52
53
54
55
56
57
58
59
60

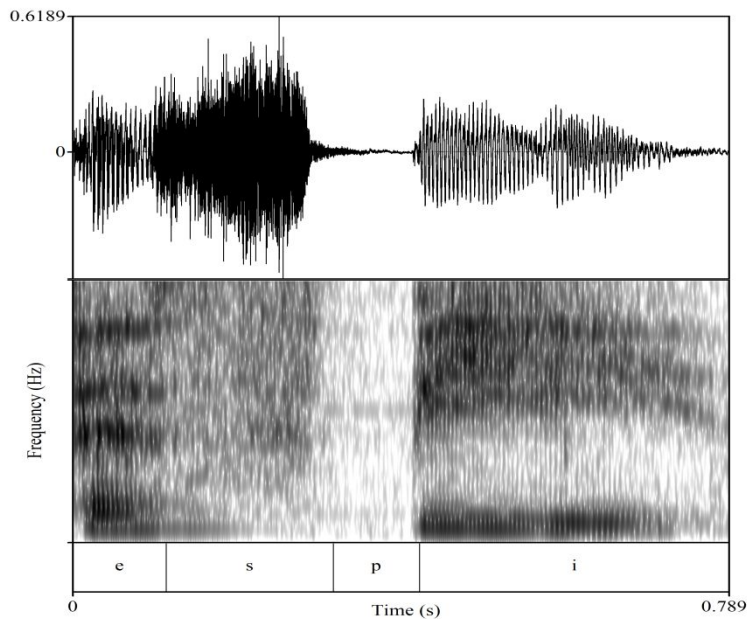
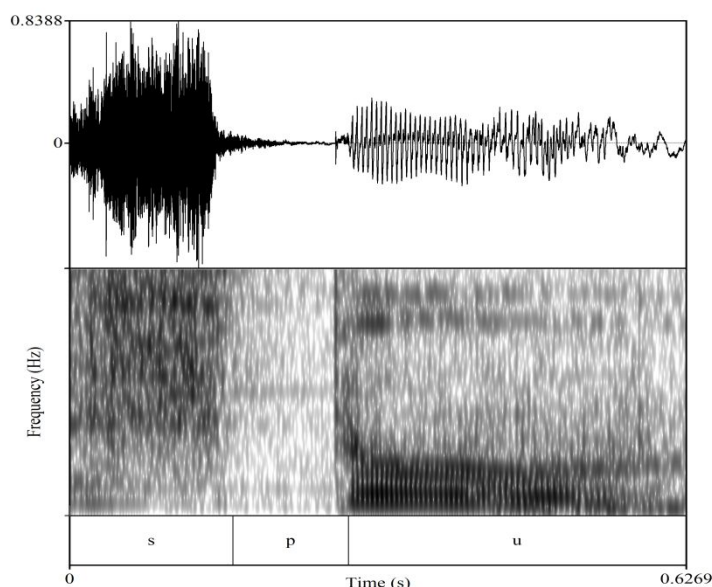


Figure 2: Token of *espi* used in perception task.

As this study examined the L2 acquisition of English, the initial vowel in *b* was the open-mid front unrounded vowel /ɛ/, which is present in the first syllable of English words like *escape* and *estate*. This vowel is also similar to Spanish /e/, and in some transcriptions has been considered to be an allophone of this phoneme (Tomás Navarro, 1918), although other transcriptions refute this (Martínez-Celdrán et al., 2003). This vowel was considered to be the most appropriate for the purpose of this study, i.e. in contrast to schwa, because (1) it is present in English words preceding /sC/ clusters, and hence the bilinguals would have been exposed to this sound sequence in their acquisition of English, and because (2) it is similar to Spanish /e/, and hence more likely to be identifiable on the part of the bilinguals than English schwa would have been. A final reason rationalising the choice of /ɛ/ was that (3) it was more naturalistic for the speaker to *produce*, as this vowel indeed occurs in English words, and therefore the recordings were likewise more likely to also *sound* naturalistic for the listeners. In essence, had we chosen schwa as the initial vowel, it may have been more likely that

1
2
3 the *quality* of the vowel impeded discrimination, rather than the *presence* of the vowel
4
5 itself. Sound files are online for interested readers.
6
7

8 The duration of the vowel approximated 80ms to 85ms, and was therefore
9
10 considered to be controlled yet naturalistic. As in Table 2, only high vowels followed
11
12 the /sC/ cluster as in stressed position these English vowels were thought to be more
13
14 similar to the Spanish high vowels than English low vowels to Spanish low vowels. For
15
16 example, in Spanish there is just one low vowel, /a/, whilst in English there are
17
18 numerous, e.g. depending on dialect /a/, /ɑ/ and /ɒ/ (Hualde, 2005). As such, the task
19
20 was considered to focus on the discrimination of the vowel prothesis in the original
21
22 token, and no additional challenges were present which might have additionally
23
24 encumbered the perception of the stimuli. Note that the speaker produced
25
26 monophthongal realisations of the vowels in the *espi* items with low second formant
27
28 frequencies (see Fig. 3 & 4), i.e. no GOOSE-fronting, which has been reported in new
29
30 varieties of British English (Cheshire et al., 2011; Mills, 2014).
31
32
33
34
35
36



57 **Figure 3: Token of *spu* used in perception task.**
58
59
60

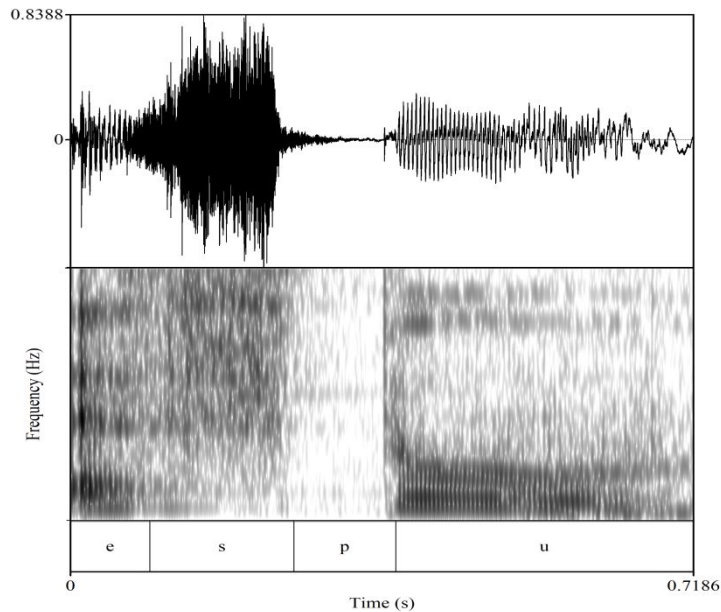


Figure 4: Token of *espu* used in perception task.

Procedure

The experiment comprised three conditions, which were presented randomly rather than in blocks. In each condition participants had to distinguish between the pairs of stimuli presented immediately after the other in conditions (1), (2) and (3). Participants were not specifically told what language the items would be in, but all of the instructions for this specific task were in English, and the entire experimental procedure was in English, as well as the questionnaire, so participants were considered to have been in an English mode (Grosjean, 1998). As such, it was thought that they would be more likely to discriminate *spi-espi* pairs than had the experiment been conducted entirely in Spanish.

Condition 1 contained stimulus pairs like *spi-spu*, which we expected to be relatively easy for both Spanish native speakers and English native speakers. Condition 2 was the target condition in which participants had to distinguish between *spi-espi* pairs. This condition was predicted to be more difficult for the Spanish native speakers with English as an L2 than for the English monolinguals. In Condition 3, participants

1
2
3 had to determine whether *spi-spi* type pairs were the same or different, and this, like
4
5 Condition 1, was expected to be relatively easy for both the Spanish-English bilinguals
6
7 and the English monolinguals. However, because all three conditions were presented
8
9 randomly to the participants, the task was on the whole expected to be more difficult
10
11 for the bilinguals than for the monolinguals, such that the Spanish-English bilinguals
12
13 may also have made more errors, or have had slower response times in Conditions 1
14
15 and 3.
16
17
18

19
20 Participants indicated their response using a handheld video game controller by
21
22 pressing the left button for same and the right button for different. Participants were
23
24 given 2500ms to complete their decision, and then the next stimulus pair was presented
25
26 automatically. A 2500ms duration was allowed because it worked well in the pilot
27
28 project as a duration that was long enough to allow responses from the full range of
29
30 participants, but fast enough to keep the pace of the experiment proceeding without
31
32 long lags, and to force a rapid, automatic response (rather than a more reflective slow
33
34 response). Note that Dupoux et al. 1999 exp3 used a 4 sec cut-off, but we found that
35
36 we lost less than 3% of trials with a 2500ms cut-off in the pilot (in the current
37
38 experiment, we lost 4.5% of all trials due to no response within the time window).
39
40
41

42
43 There was a short practice block, consisting of 6 trials: 3 different, 3 same.
44
45 Participants were given feedback after the practice block, and could ask questions if the
46
47 instructions were not clear, although this never happened. The timing ‘clock’ started
48
49 with the onset of the second wav file, so participants could press a button to indicate
50
51 ‘same’ or ‘different’ as soon as the second file began. There were only 2 trials where a
52
53 response was given before 300ms passed across the entire data set, so we included all
54
55 trials in the analysis. There was no beep between stimuli; participants just heard the two
56
57
58
59
60

1
2
3 sound files and then silence until they pressed the response button or 2500ms elapsed
4
5 with no response (following Dupoux et al. 1999).
6
7

8 In total, the perception task lasted 12 minutes, which included two short breaks
9
10 equally dispersed. Including the explanation of the perceptual discrimination task, the
11
12 entire procedure took 15 minutes.
13
14
15
16

17 **Task II: Prothetic Vowel Production Elicitation**

18
19 For the prothetic vowel production task, participants took part in a phonemic verbal
20
21 fluency task. In the standard version of this task, participants are asked to produce as
22
23 many unique words as possible which start with a given letter within a given timeframe
24
25 (Newcombe, 1969). The vast majority of studies that use this task use single letters as
26
27 cues, but there is at least one study that used double-letter cues (e.g. “fa”, “sm”) when
28
29 comparing Spanish-English bilinguals to English monolinguals (Sandoval et al., 2010).
30
31 In the standard test, the participant’s score is the number of unique words produced for
32
33 each given letter, but for our purposes, the accuracy score reflected accurate /sC/
34
35 pronunciation (i.e. no prothetic vowel) for each unique word.
36
37
38
39

40 The phonemic verbal fluency task, also called the letter fluency task, is a
41
42 standard task used in both normative and non-normative language assessments of native
43
44 and non-native languages, often implemented together with the semantic verbal fluency
45
46 task for which word categories are listed, e.g. words like *cat* and *dog* would be listed if
47
48 “animal” were given as a category (Grogan et al., 2009; Newcombe, 1969).
49
50 Interestingly, some research has shown that monolinguals and fluent bilinguals (in their
51
52 L2) score similarly on the phonemic verbal fluency task, i.e. they are able to list
53
54 approximately the same amount of words, but that monolinguals obtain higher scores
55
56 on the semantic verbal fluency task (Gollan et al., 2002; Portocarrero et al., 2007, but
57
58
59
60

1
2
3 see also Bialystok et al., 2008). For the purposes of our study, the bilinguals were
4 considered to be capable of the phonemic verbal fluency task, but that the task would
5 also be slightly more demanding than e.g. a word list reading task, for which we thought
6 that the full orthographic representation of the word would enable them to focus on
7 their pronunciation, and in this way not necessarily reflect what they would produce in
8 normal conversation.
9

10
11
12 In the instructions for the phonemic verbal fluency task, it was stated that they
13 would see one or two letters, like 'b' or 'gr' and should name as many English words
14 as possible which begin with the sounds which those letters spell. An example of 'pr'
15 was given, for which one should say 'print' but not 'pint' "because 'pint' does not begin
16 with 'pr'. Likewise, if they saw 'sk', it was instructed that they could say 'scream' or
17 'sky' since "they both begin with the 'sk' sounds", but that 'scene' would be incorrect.
18 Participants were given 10 seconds for each onset, and they were asked to only name
19 unrelated words, e.g. they should avoid naming both "print" and "printed". A short
20 amount of time of 10 seconds was given for each onset so that participants were not
21 able to reflect too much on their pronunciation. In total, 22 onsets were elicited, but
22 only the /sC/ clusters were of interest to the study at hand: /sp/, /st/, /sk/, /sl/, /sn/ and
23 /sm/. As there were six clusters, participants were given in total 1 minute to produce
24 the /sC/ clusters, which is the standard total time in a phonemic verbal fluency task
25 (Newcombe, 1969).
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 The phonemic verbal fluency task is often used to assess executive control
50 abilities, as participants need to retrieve words, which requires them to access their
51 mental lexicon, whilst focusing on the task, selecting words meeting certain constraints
52 and inhibiting repetition (Miyake et al., 2000). However, the amount of words produced
53 is considered to be dependent on not only executive control abilities, but also on e.g.
54
55
56
57
58
59
60

1
2
3 vocabulary size (Sauzéon et al., 2011). As such, this task was considered to be a more
4
5 demanding task than other potential production elicitation techniques, and more
6
7 representative of normal everyday speech production, which would likewise require
8
9 both monolinguals and bilinguals to access their mental lexicon whilst monitoring
10
11 pronunciation.
12
13

14
15 Moreover, the task was also considered to be potentially more cognitively
16
17 demanding and more reflective of “real world” language tasks than the previously
18
19 described perception task. In this way, it could be argued that the tasks were not
20
21 balanced, but we thought that it was interesting to investigate whether the “raw capacity”
22
23 to discriminate *spi-espi* in the perception task, unimpeded by other cognitive tasks,
24
25 would be shown to facilitate accurate production of /sC/ clusters in English. A crucial
26
27 question was therefore whether bilinguals who scored higher in the production task
28
29 would likewise also necessarily score higher in the perception task, if accurate
30
31 perception is indeed a requirement for accurate production.
32
33
34
35
36

37 *Acoustic analysis*

38
39 In order to determine whether a prothetic vowel was present in the Spanish-English
40
41 bilinguals’ production of word-initial consonant clusters beginning with /sC/ in English,
42
43 an acoustic analysis was conducted on recordings obtained from all participants taking
44
45 part in the production task.
46
47
48

49
50 Using Praat (Boersma and Weenink, 2010), a textgrid file was created which
51
52 included interval tiers (respectively, “word”, “vduration”) and one point tier (“notes”).
53
54 In the top “word” tier, real words which were produced by participants were input
55
56 directly while non-words were spelt out according to participants’ actual pronunciations
57
58 and were marked with an “X” at the end of the created word. The tier called “vduration”
59
60

1
2
3 was used to annotate the prothetic vowel, if there was one at all, and the point tier was
4
5 used to make comments on individual items, when considered necessary. To delimit
6
7 the words, the cursor to indicate the start of the word was placed at the closest 0-
8
9 crossing in the waveform where either phonation began from the prothetic vowel, or
10
11 where aperiodic noise began from the sibilant. The cursor to indicate the end of the
12
13 word was placed at the closest 0-crossing where there was considered to be a significant
14
15 drop in waveform amplitude.
16
17

18
19 Both the spectrogram and the waveform were used to determine the onset and
20
21 offset of the prothetic vowels preceding the onset of the sibilant. Where necessary,
22
23 Praat's recognition of a pitch contour (i.e. lack of recognition for the voiceless sibilant)
24
25 was also considered as an indicator of voicing in the case of the potential vowel. In
26
27 most cases, the onset and offset of the potential prothetic vowel were located at the
28
29 beginning and the end of periodicity where the waveform crossed the 0-axis (some
30
31 prothetic vowels started with a glottal stop, or some other form of plosion – and this
32
33 was also counted as part of the vowel). The presence of formants in the spectrogram
34
35 was also used as an indicator of vowel prothesis. However, in some cases, the prothetic
36
37 vowel was produced with creaky voice. In such cases, auditory perception of an initial
38
39 vowel was the main cue which was used to determine whether there was a prothetic
40
41 vowel. The offset of the vowel was determined by the onset of the following sibilant.
42
43 Determination of the sibilant was relatively uncontroversial; marked by the onset of a
44
45 high intensity band of aperiodic friction in the waveform and spectrogram. As was the
46
47 case for the onset of the potential prothetic vowel, this transition from vowel to sibilant
48
49 was marked at the nearest 0-crossing in the waveform. The duration of the prothetic
50
51 vowel was then measured in milliseconds (ms), and data regarding duration can be
52
53 found in the appendix; however, here due to space constraints, we only present binary
54
55
56
57
58
59
60

1
2
3 data of both the perception and production results (findings from continuous durational
4 analyses confirmed binary data).
5
6
7
8
9

10 **Statistical analysis**

11
12 For both the perception and production results, data were organised in CSV files using
13 Excel software. Thereafter, R software (R Core Team, 2017) was used for the
14 analyses and a series of binomial mixed-effects regression models were built for the
15 accuracy results of the perception and production tasks and a series of linear
16 regression models were built for the reaction time results of the perception task (using
17 the lme4 package in R; Bates, Maechler, Bolker, & Walker, 2015) for the examination
18 of the influence of fixed and random factors on the response. For the analysis of the
19 perception data, timed out responses (i.e. trials which were not discriminated within
20 2500ms) were excluded resulting in the exclusion of 371 trials out of 8274 trials. For
21 the analysis of the production data, 616 individual word tokens were analysed. We
22 present the perception results initially, then the production results, then the results
23 regarding perception as a predictor of speech production.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41

42 **Results**

43 **Task I: Perception**

44
45
46
47
48
49
50
51 The accuracy and reaction time results of the perception experiment are summarised
52 per group and condition in Table 4 and the binary accuracy results are visualised in
53 Figure 5, revealing that Spanish–English bilinguals performed least accurately in
54
55
56
57
58
59
60

condition 2 compared to conditions 1 and 3, and less accurately than the English monolinguals.

	<i>English monolinguals</i>		<i>Spanish-English bilinguals</i>	
	Percent of correct responses	Response times for correct responses	Percent of correct responses	Response times for correct responses
<i>Condition 1: spi-spu</i>	96.05 (19.5)	1175.4 (217.12)	95.17 (21.45)	1193.48 (252.26)
<i>Condition 2: spi-espi</i>	87.4 (33.23)	1240.89 (302.94)	62.29 (48.48)	1311.41 (384.38)
<i>Condition 3: spi-spi</i>	97.36 (16.04)	1197.19 (308.67)	93.11 (25.33)	1228.09 (341.35)

Table 4: Percentage of correct responses and response times for each group (standard deviations in brackets).

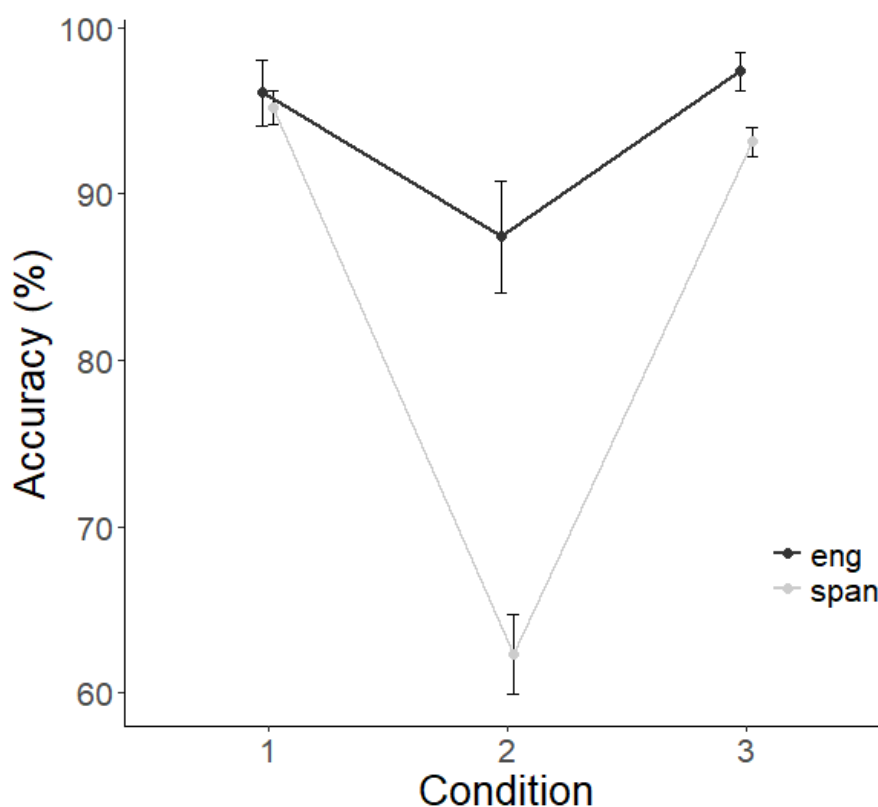


Figure 5: Perception accuracy in English monolinguals versus Spanish-English bilinguals.

For the data analysis of the accuracy results, the modelled response was accuracy (i.e. yes or no, whether the participant was accurate in the AX discrimination or not), and the fixed factors were language background (L1, English vs. Spanish) and condition

(C1: *spi-spu*, C2: *spi-espi*, C3: *spi-spi*). Random intercept was consonant, i.e. second consonant in the sC cluster (e.g. /p/ in /sp/) (participant was initially included but this model failed to converge). Models were manually stepped-down (using likelihood ratio tests) from maximal models containing all factors and possible interactions to the ‘best’ model that only contained significant predictors or predictors that participated in significant interactions (Barr et al., 2013). The best fitting model determined by step-down pairwise model comparison was $\text{acc} \sim \text{L1} * \text{condition} + (1 | \text{consonant})$. The model parameters are in Table 5.

	Estimate	Std. Error	z value	Pr (> z)
(Intercept=L1=English, Cond=1)	3.1955	0.2689	11.884	< 2e-16 ***
L1=Spanish, Cond=1	-0.2099	0.2888	-0.727	0.46741
L1=English, Cond=2	-1.2577	0.3060	-4.111	3.95e-05 ***
L1=English, Cond=3	0.4174	0.3484	1.198	0.23084
L1=Span x Cond=2	-1.2300	0.3318	-3.707	0.00021 ***
L1=Span x Cond=3	0.7967	0.3742	-2.129	0.03325 *

Table 5: Estimates, standard errors, z values and p values of the best fitting model for the perception experiment.

As displayed in Table 5, the main finding was that, as predicted, the magnitude of the difference in the Spanish speakers between condition 1 and condition 2 was greater than the difference in the English speakers between condition 1 and condition 2 ($\beta = -1.2300$, $z = -3.707$, $p < 0.001$), although responses to condition 2 were also less accurate than for condition 1 for English speakers ($\beta = -1.2577$, $z = -4.111$, $p < 0.0001$). Furthermore, the magnitude of the difference in the Spanish speakers between condition 1 and condition 3 was also greater than the difference in the

1
2
3 English speakers between condition 1 and condition 3 ($\beta = 0.7967, z = -2.129, p <$
4
5 0.05). No other comparisons were significant.

6
7
8 In a second analysis, we investigated possible differences among the Spanish–
9
10 English bilinguals in their responses to the critical condition 2 (e.g. *spi-espi*) using the
11
12 predictor variables of percentage of English daily use, age of English acquisition, and
13
14 C-Test proficiency score, i.e. grammatical score, which were scaled using the scale
15
16 function in R. However, none of these predictors proved to be significant ($ps > 0.05$).

17
18
19 For the reaction time results, in our first analysis, the modelled response was
20
21 log-transformed RT and the fixed factors were language background (L1, English vs.
22
23 Spanish) and condition (C1: *spi-spu*, C2: *spi-espi*, C3: *spi-spi*). Random intercepts
24
25 were participant and again consonant. None of the predictors or their interaction
26
27 proved to be significant ($ps > 0.05$). Nevertheless, the numerical RT pattern was
28
29 consistent with the accuracy results pattern with the Spanish–English bilinguals
30
31 having longer RTs in condition 2 (e.g. *spi-espi*).

32
33
34 In a second analysis, we investigated possible differences among the Spanish–
35
36 English bilinguals in their responses to the critical condition 2 (e.g. *spi-espi*) using the
37
38 predictor variables of percentage of English daily use, age of English acquisition, and
39
40 C-Test proficiency score, which were scaled using the scale function in R; however,
41
42 none of the predictors proved to be significant ($ps > 0.05$).

43 44 45 **Task II: Production**

46
47
48
49
50
51
52
53 Table 6 summarises the rates of prothesis, for each of the six cues of interest, for the
54
55 two participant groups and Figure 6 visualises the results.

<i>Cluster</i>	<i>English monolinguals (%)</i>	<i>Spanish–English bilinguals (%)</i>
sk	8.70 (28.8)	55.56 (50.0)
sl	11.1 (32.3)	55.13 (50.1)
sm	0.00 (0.0)	57.53 (49.8)
sn	6.67 (25.8)	52.86 (50.3)
sp	4.35 (20.9)	62.63 (48.63)
st	8.33 (28.2)	67.37 (47.1)

Table 6: Percentage of words produced with prothesis for each sC cue, by participant group (standard deviations in brackets).

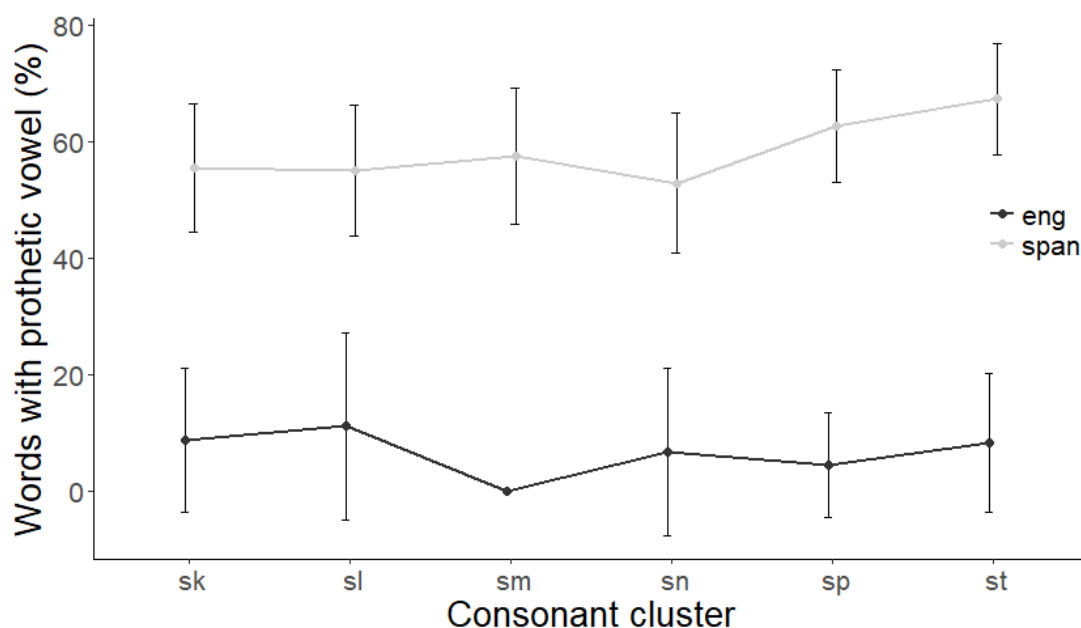


Figure 6: Production accuracy in English monolinguals versus Spanish-English bilinguals.

For the production experiment, the modelled response was whether a prothetic vowel was produced or not (i.e. prothesis), for each word produced, and the fixed factor was language background (L1 English vs. L1 Spanish). Random intercepts included participant and consonant cluster (i.e. /st/, /sp/, /sk/, /sm/, /sn/ and /sl/). The maximal

model motivated by our hypotheses was $\text{prothesis} \sim \text{L1} + (1 \mid \text{participant}) + (1 \mid \text{consonant.cluster})$.

	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>Pr(> z)</i>
(Intercept) (L1 = English)	-3.7248	0.8353	-4.459	8.23E-06 ***
L1 = Spanish	4.2509	0.886	4.798	1.61E-06 ***

Table 7: Estimates, standard errors, z values and p values of the best fitting model for the production experiment.

As displayed, for the production analysis, the main finding was that as predicted, there was a significant difference between the English monolinguals and the Spanish-English bilinguals ($\beta = 4.2509$, $z = 4.798$, $p < 0.00001$) with the bilinguals producing significantly more prothetic vowels than the monolinguals. To exemplify this difference, individual words (i.e. *speak*, *slain*, *smile*) produced by the Spanish-English bilinguals are displayed in figures 7-9.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

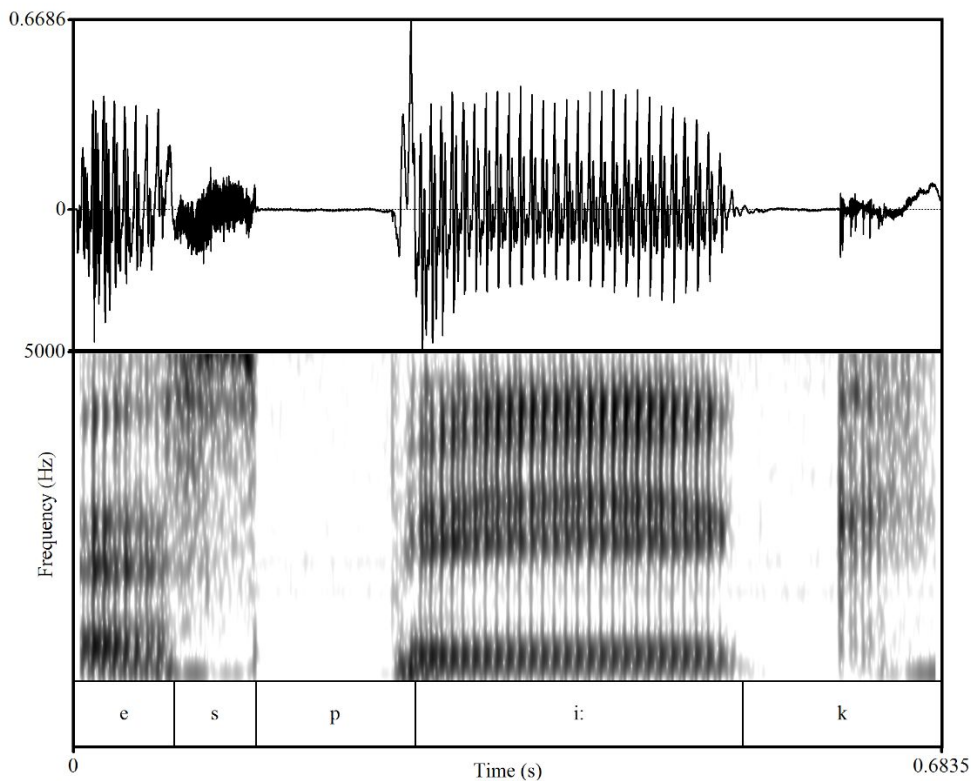


Figure 7: The word “speak” produced by a Spanish-English bilingual with prosthesis.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

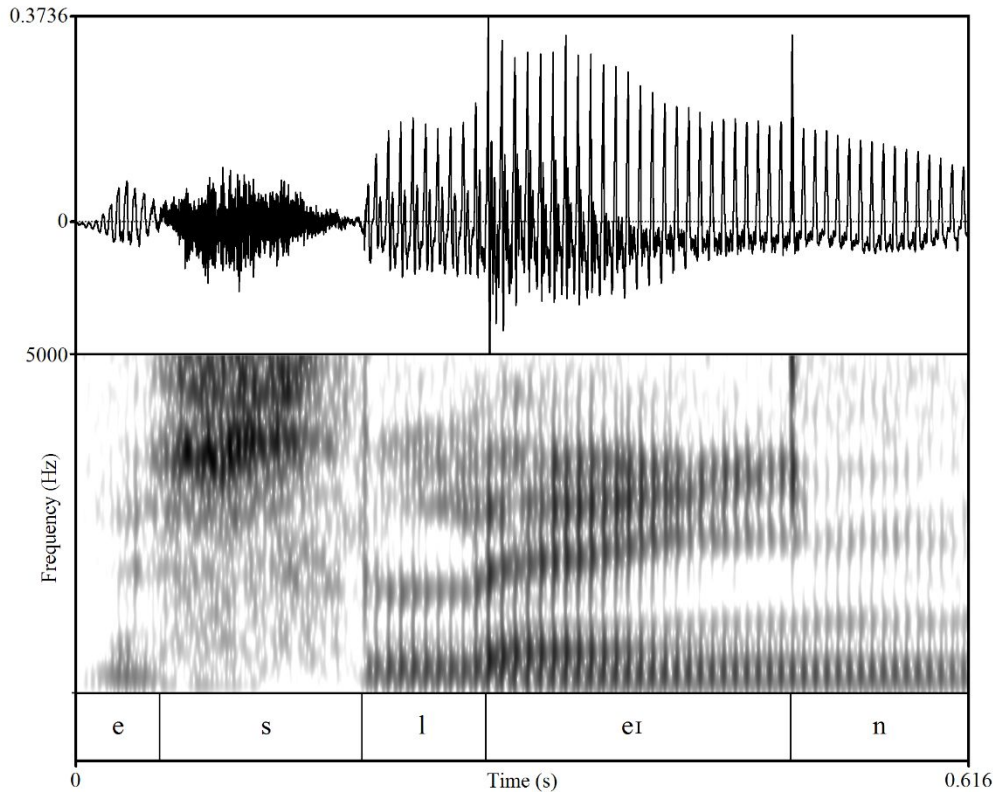


Figure 8: The word “slain” produced by a Spanish-English bilingual with prosthesis.

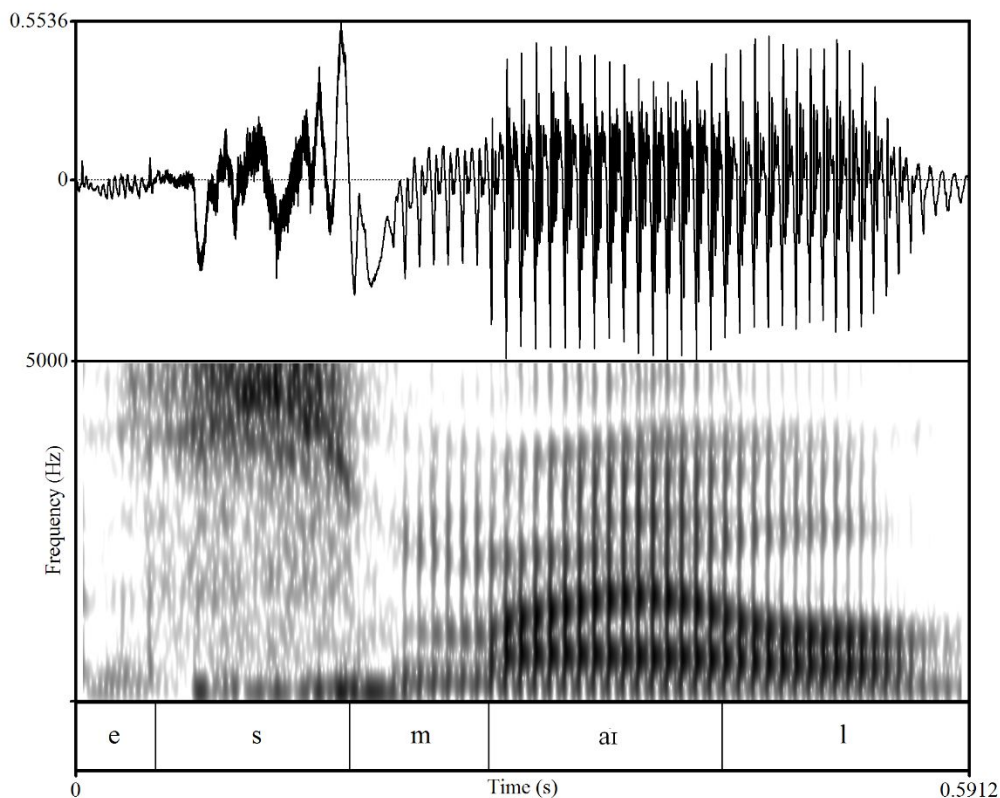


Figure 9: The word “smile” produced by a Spanish-English bilingual with prosthesis.

Perception as a predictor for production

For the next analysis stage, we calculated a mean perception accuracy score on the critical Condition 2 contrast (e.g. *spi-espi*) in the perception experiment for each bilingual speaker. This score was included as a possible predictor in this second analysis of the production results, along with percentage of English daily use, age of English acquisition, and grammatical proficiency score (C-test result) in order to examine whether these factors would be influential in determining accuracy of pronunciation (i.e. either producing the prosthetic vowel in the phonemic verbal fluency task or not). The maximal model tested was: prosthesis ~ Percent of English in Daily Use + Age of Acquisition + English Proficiency +

Condition2PerceptionAccuracy + (1 | participant) + (1 | consonant.cluster)¹. As before, all continuous predictor variables were scaled using the scale function in R. Models were manually stepped-down (using likelihood ratio tests) from maximal models containing all factors to the best fitting model that only contained significant predictors (Barr et al., 2013). The best fitting model determined by step-down pairwise model comparison was prothesis ~ Percent of English in Daily Use + English Proficiency + (1 | participant) + (1 | consonant.cluster). Neither age of English acquisition nor accuracy rate on condition 2 of the perception task was a significant predictor in the maximal model, and removing those terms from the maximal model did not decrease the model fit (as assessed by likelihood ratio tests) (all $ps > 0.05$). The parameters of the winning model are displayed in Table 8. As shown, both percentage of daily English use ($\beta = -0.5454$, $z = -2.523$, $p < 0.05$) and English proficiency ($\beta = -0.6648$, $z = -3.057$, $p < 0.01$) significantly influenced the amount of prothesis produced by the Spanish-English bilinguals. N. b. an analysis of duration of prothesis revealed similar overall findings.

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)				
Prothesis	0.03271	0.41492	0.079	0.93716
Daily Eng Use	-0.5454	0.2162	-2.523	0.01165*
Eng Proficiency	-0.6648	0.2174	-3.057	0.00223**

Table 8: Estimates, standard errors, z values and p values of the best fitting model for prothesis produced by the Spanish–English bilinguals.

Figures 10 and 11 plot the relationships between the two significant predictor variables (respectively English daily use and English grammatical proficiency) and

¹ Interaction terms were not included in the maximal model as there were not enough observations to allow for so many parameters.

the proportion of words produced with prothesis for each speaker, with the grey ribbon representing 95% confidence interval. As displayed in Figure 10, a higher amount of English use was correlated with less prothetic vowel productions, i.e. a more English-like pronunciation ($t = -2.1289$, $df = 32$, $p < 0.05$, $r = -0.35$). As displayed in Figure 11, a higher grammatical proficiency in English was correlated with less prothetic vowel productions, i.e. a more English-like pronunciation ($t = -2.5228$, $df = 32$, $p = 0.05$, $r = -0.41$).

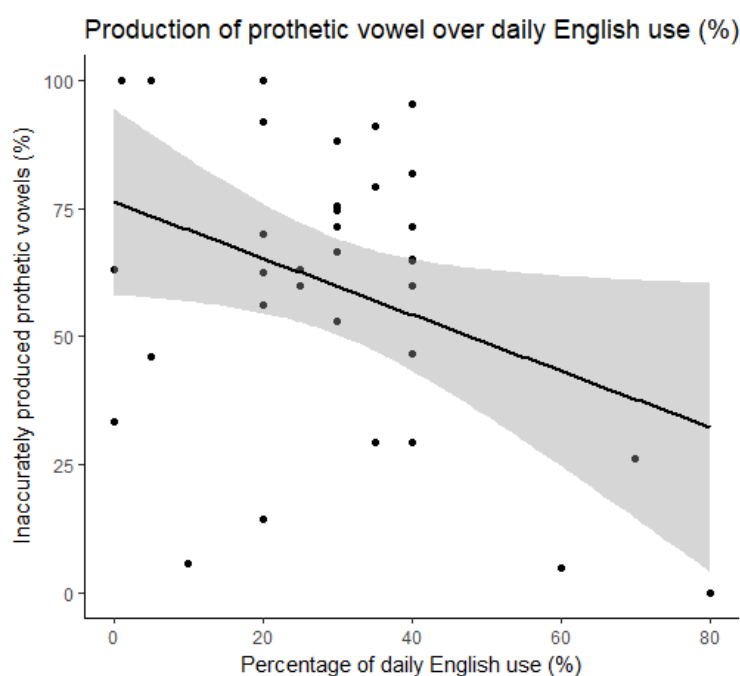
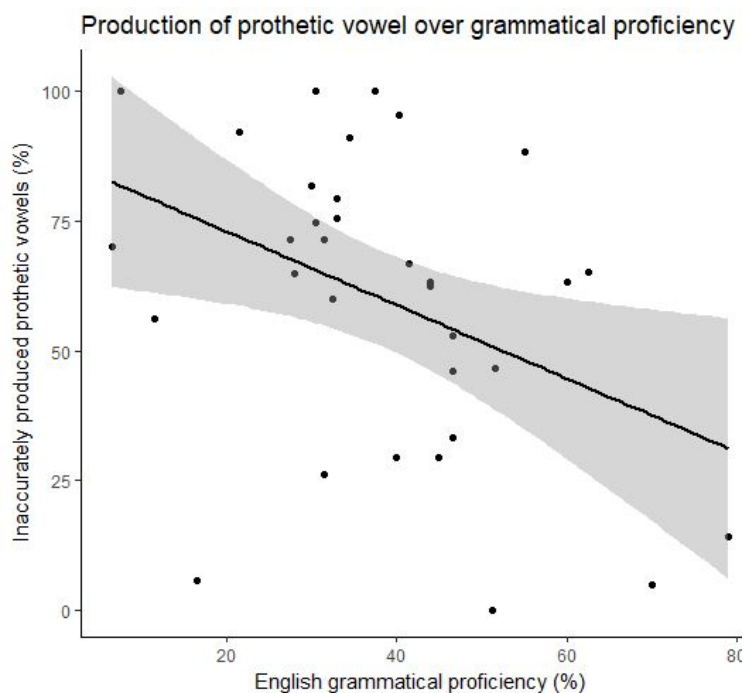
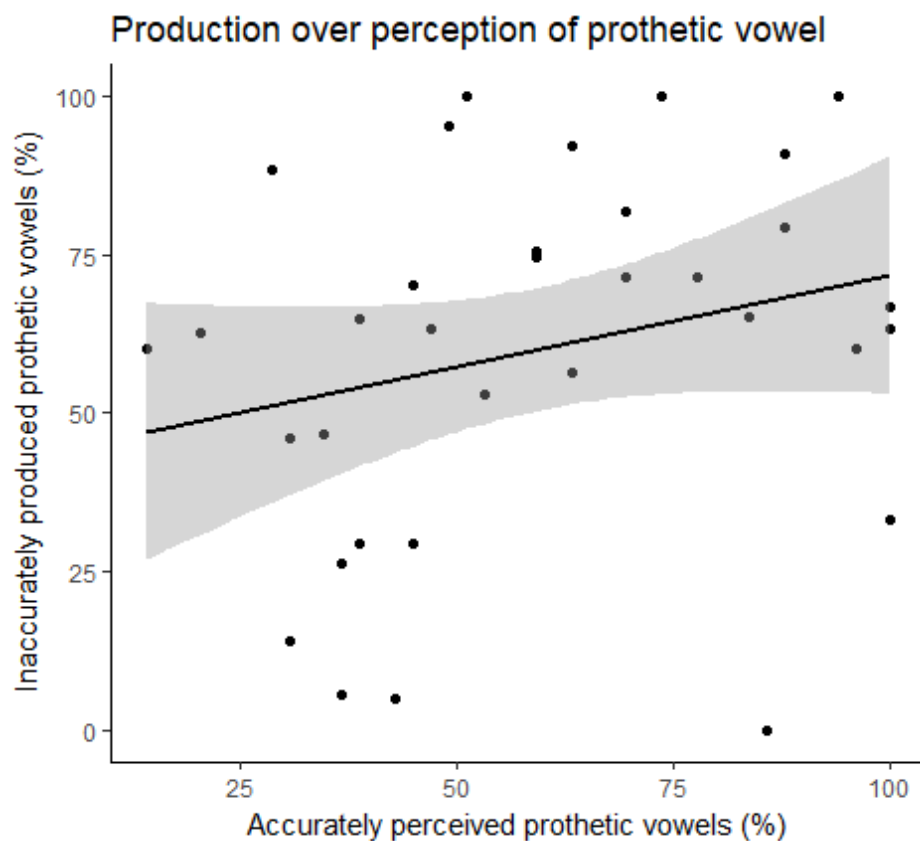


Figure 10: Production of prothetic vowel over daily English use (%).



25 **Figure 11: Production of prothetic vowel over grammatical proficiency C-test**
26 **result (%).**

27
28
29 Figure 12 plots the non-significant relationship between perception accuracy and
30 prothesis proportion ($t=1.4906$, $df=32$, $p=0.1458$). If the four quarters of the scatterplot
31 are inspected, it is possible to see that the bottom right-hand quarter is almost empty,
32 indicating that hardly anyone displayed accurate production and perception.
33 Alternatively, in the top left-hand corner, many bilinguals displayed perception
34 accuracy rates below chance level (i.e. beneath 50%), as well as high prothetic vowel
35 rates. In the top right-hand quarter, bilinguals are visible who showed perception
36 accuracy rates above chance, combined with high prothetic vowel rates, indicating that
37 accurate perception was not sufficient for accurate production. In the bottom left-hand
38 corner, bilinguals are visible who showed perception accuracy rates below chance,
39 combined with low prothetic vowel rates, indicating that perception was not necessary
40 for accurate production.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



31
32
33
34

Figure 12: Production over perception of prothetic vowel (non-significant relationship).

35
36

Discussion of findings

37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

In this study, speech perception was assessed through a same-different discrimination task in stimulus pairs such as *spi-espi* and speech production of vowel prothesis was assessed through a phonemic verbal fluency task in Spanish-English sequential bilinguals. The primary objective was to examine whether accurately perceiving the prothetic vowel preceding /sC/ clusters would likewise help to *not* produce this prothetic vowel in the L2. In general, the findings revealed that there was a significant difference between the monolingual English speakers and the Spanish-English bilinguals, and that there was no clear relationship between the perception and production of /sC/ clusters in onset position (which are restricted by Spanish phonotactic rules) by the Spanish-English bilinguals. Results indicated that the

1
2
3 bilinguals were less successful at discriminating the difference between tokens like *spi-*
4 *espi* than the English monolinguals, although, interestingly, the monolinguals were, like
5 the bilinguals, also more accurate in the *spi-spu* and *spi-spi* trials than in the *spi-espi*
6 trials. The reason the monolingual English speakers may have performed less
7 accurately on the *spi-espi* trials than the *spi-spu* and *spi-spi* trials is potentially because,
8 as observed by Carlson (2016), English favours #sC over #VsC. Nevertheless, the
9 English monolinguals performed significantly more accurately than the Spanish
10 bilinguals in discriminating *spi-espi* in comparison to the other two conditions.
11 Furthermore, in this population of sequential Spanish-English bilinguals accurately *not*
12 producing the prothetic vowel in the phonemic fluency task was *not* associated with
13 accurate perception of the prothetic vowel.
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 Moreover, although a greater amount of English daily use and improved
29 grammatical proficiency were associated with better production (i.e. *not* producing the
30 prothetic vowel), these same predictor variables did not lead to better discriminatory
31 abilities in the perception task. Therefore, the results suggest that with regard to this
32 particular phonotactic constraint, production appears to be modified by increased
33 exposure to the L2 and grammatical proficiency, whereas perception does not.
34
35
36
37
38
39
40
41

42 Moreover, the analysis of the individual bilinguals did not fully verify the
43 assertion that accurate speech perception would be a prerequisite for accurate speech
44 production. For example, some participants scored poorly on the production task, but
45 above chance level on the perception task; alternatively, many participants scored
46 highly on the production task, but below chance level on the perception task. Although
47 many participants differed with regard to their accuracy on the perception versus
48 production tasks, some sequential Spanish-English bilinguals achieved high accuracy
49 results on both the perception and production tasks. Alternatively, other participants
50
51
52
53
54
55
56
57
58
59
60

1
2
3 scored poorly on both tasks. These different profile patterns suggest that accurate L2
4
5 perception was neither necessary nor sufficient for accurate L2 production in the case
6
7 of these sequential Spanish-English bilinguals, but that, as previously noted, increased
8
9 exposure to English and improved grammatical proficiency could improve production
10
11 results.
12
13

14
15 Therefore, as in previous research into the relationship between speech
16
17 perception and production, no clear overall pattern was evident in terms of whether
18
19 perception or production necessarily “comes first” (Baker & Trofimovich, 2006; Beach,
20
21 Burnham, & Kitamura, 2001; Llisterri, 1995; Kartushina & Frauenfelder, 2014;
22
23 Sheldon & Strange, 1982; Zampini, 1998). With regard to increased L2 exposure, it
24
25 seems plausible that bilinguals exposed to more English would have been more likely
26
27 to develop English-like pronunciation, and, likewise, it seems plausible that bilinguals
28
29 with higher grammatical proficiency would likewise be more metalinguistically *aware*
30
31 that in English /sC/ clusters in onset position are licit, and that this knowledge and
32
33 exposure would enhance their L2 speech production, but not necessarily their
34
35 performance in the speeded perceptual discrimination task.
36
37
38
39

40
41 Why did the bilinguals seem to be able to produce English L2 speech without
42
43 the vowel in front of /sC/ clusters, but increased English use and improved English
44
45 grammatical proficiency did not seem to improve their perception of this contrast?
46
47 Firstly, in answering this question, it is important to remember that some participants
48
49 *did* perceive a difference but nevertheless produced /sC/ clusters in the Spanish manner
50
51 – so it would not be entirely correct to say that perceptual processes are always more
52
53 robust and pervasive as far as L1 influence in L2 processing than is speech production.
54
55 If that were the case then we would only have found evidence for enhanced speech
56
57 production, but not for speech perception. What is more likely is that either the tasks
58
59
60

1
2
3 assessed were actually quite different - and assessed speech perception and speech
4 production differently - or, alternatively, other predictor variables, not assessed within
5 this study, impact speech perception, rather than L2 use and L2 grammatical
6 proficiency, which proved to enhance speech production. For example, it may be that
7 some individuals simply *hear* differences between sounds better than others. We know
8 as well that some people suffer in musical terms from amusia, i.e. tone deafness (Powell,
9 2016). Similarly, someone who is less “apt” at hearing differences between sounds
10 might nonetheless be able to *not* produce the prothetic vowel in front of English /sC/
11 clusters with the metalinguistic knowledge that this sound combination is *possible* in
12 English, and, with practice, become quite good at consistently *not* producing the vowel.
13 Note that what is different about this speech production task is that accuracy was
14 reached when *nothing* was pronounced, rather than when a sound was produced more
15 “native-like”, e.g. such as a more similar vowel quality. Alternatively, someone who is
16 very “apt” at hearing differences between sounds (Powell, 2016) might essentially have
17 a talent for hearing differences, which would be revealed in the perception task, but,
18 when taking part in the production task, that same person might not perform particularly
19 highly if not focussed, or if he or she didn’t know the grammatical rule in English, or
20 had not internalised that rule through practice. The point here is that different factors
21 might lead to improved speech perception than lead to improved speech production,
22 and the current study might only have examined factors which lead to improved speech
23 production, i.e. no production.

24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52 It may also be that the striking interpersonal differences between bilingual
53 participants arose due to different individual approaches to the tasks. Parlato-Oliveira
54 et al. (2010) summarised in their work that an effect of L2 exposure was most clearly
55 observed in their explicit task, but not in the implicit task, and suggested that L2
56
57
58
59
60

1
2
3 experience is more likely to affect explicit, or metalinguistic, perceptual tasks. With
4 regard to our AX discrimination task, participants may have been metalinguistically
5 aware of the phonotactic constraint it assessed (note that grammatical knowledge was
6 correlated with production of the prothetic vowel), and the interpersonal differences
7 may have been a result of some participants simply being more aware of this
8 metalinguistic difference than others.
9

10
11
12 Furthermore, AX discrimination tasks require low-level perceptual judgments
13 on whether auditory input is the same or different. This is arguably easier to
14 discriminate than, for example, *whether* a vowel is present at the beginning of /sC/ or
15 *whether* the auditory input forms a word or non-word, and hence the perception task
16 may have assessed a rather low-level ability to acoustically discriminate, rather than the
17 prothetic vowel - as a grammatical constraint - as such. Therefore, the interpersonal
18 differences in the perception results may have been the result of some participants
19 simply acoustically discriminating better than others, with no direct insight into their
20 grammatical knowledge; however, the bilinguals were just as good as the English
21 monolinguals at perceiving the *spi-spu* contrast, whereas there was a significant
22 difference in their ability on the *spi-espi* task, so it cannot simply be the case that some
23 participants simply acoustically discriminated better than others.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 It may also be that some of the prothetic vowels were actually hesitation
47 markers, and not reflective of grammatical constraints. It has been found that in
48 spontaneous speech hesitation markers are often used and there are discrepant
49 interpretations regarding their function, e.g. whether “um” and “uh” are used to “hold
50 the floor” or whether they are symptomatic of cognitive processes on the part of the
51 speaker, or both (Clark and Fox Tree, 2002; de Leeuw, 2007; Fox Tree, 2001; Maclay
52 and Osgood, 1959; Shriberg and Lickley, 1993). The present methodology included all
53
54
55
56
57
58
59
60

1
2
3 vocalic utterances preceding the sibilant as prothetic, and thus, a future analysis may
4
5 consider including a read task as well, where hesitations are less likely.
6
7

8
9 It may also be, moreover, that the perception and production tasks recruited
10 different cognitive functions. As previously discussed, the phonemic verbal fluency
11 task is often used to assess executive control, as participants need to retrieve words,
12 which requires them to access their mental lexicon, whilst focusing on the task,
13 selecting words meeting certain constraints and inhibiting repetition (Miyake et al.,
14 2000). This particular production task of eliciting words which begin with /sC/ clusters
15 would require the Spanish L1 participants to shift, as they moved from one cluster to
16 the next, monitor, to ensure that they didn't repeat words, and to remember new words;
17 as well as *inhibit* their native Spanish language phonotactic constraint, as well as
18 previously named words. In contrast, it could be argued that the perception task might
19 have recruited a shifting mechanism, but that working memory and inhibition would
20 have been less involved, i.e. the participants only had two stimuli to remember.
21 Likewise, it may be more generally that the two capacities – L2 speech production and
22 L2 perception - recruit different executive control mechanisms to varying extents,
23 which could be an interesting research avenue to pursue in the future.
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 Essentially, the perception task in the current study was a 'can you do it' type
45 task, whilst for the production task the participants were required to think quickly
46 within a time constraint, which takes their focus off of their pronunciation. As such, in
47 the current experiment, the production task was cognitively more costly, and it therefore
48 seems possible that the relationship between production and perception might have
49 been more clear cut if the tasks had required similar cognitive demands. That said, if it
50 is accepted that the production task was more cognitively demanding than the
51 perception task, and accurate perception is indeed a prerequisite for accurate production,
52
53
54
55
56
57
58
59
60

1
2
3 we should nevertheless *not* have observed bilinguals who had high accuracy rates on
4 the production task, but low accuracy rates on the perception task (i.e. bottom left-hand
5 quarter in Figure 12). The fact that this quarter was indeed quite full of participants,
6
7 indicates that accurate L2 speech perception is not necessary for accurate L2 speech
8
9 production.
10
11
12
13
14

15 Perhaps it was not only the task which varied in terms of different executive
16 control mechanisms, or whether it was explicit or implicit (although this is of course
17 the case too), but that the individuals approached the same tasks differently. For
18 example, in the case of the perception task, participants may have differed in the amount
19 of attention they devoted to the task. Those participants who were very focused may
20 have performed more accurately than those who were less focused, regardless of how
21 proficient they were in their L2. Likewise, those participants who were more focussed
22 on their pronunciation in the production task might have produced no prothetic vowel,
23 whilst those participants who focussed more on the task of naming as many words as
24 possible could have more frequently produced a prothetic vowel. Again, this would
25 have little to do with their proficiency in English as such, but rather with how they
26 approached the task, and may help to explain the differences observed between
27 participants. Nonetheless, at face value, the results do not lend direct support to the
28 notion that perception precedes production in L2 acquisition.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 As such, the findings suggest that L2 speech models such as SLM and PAM,
49 which postulate to different extents that perception is linked to production, may need
50 revisiting. Indeed, some of the participants in the present study were able to *not* produce
51 the prothetic vowel (i.e. accurately, as in English), but performed below chance with
52 regard to the perception of the prothetic vowel in the same-different discrimination task.
53
54
55
56
57
58
59
60 The findings from the correlational analysis, which indicated no significant relationship

1
2
3 between perception and production, as well as in the mixed model analyses, which
4
5 likewise did not indicate that perception abilities predicted production, both suggest
6
7 that perception and production were not clearly related to one another with regard to
8
9 the tasks at hand.
10

11
12 In brief, the results confirm a growing body of research examining the
13
14 perception and production of individual L2 sounds which suggests dissociations
15
16 between L2 perception and production. All in all, the results reveal that it is possible
17
18 for sequential bilinguals to acquire knowledge of new permissible phoneme sequences,
19
20 which violate their L1 phonotactic constraints, but that there is a great amount of
21
22 interpersonal variability within regard to how the L2 acquisitional process unfolds.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- Baker W and Trofimovich P (2006) Perceptual paths to accurate production of L2 vowels: The role of individual differences. *IRAL – International Review of Applied Linguistics in Language Teaching* 44(3): 231–250.
- Barr D J, Levy R, Scheepers, C and Tily, HJ (2013) Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3): 255-278.
- Beach, E. F., Burnham, D. & Kitamura. C. (2001). Bilingualism and the relationship between perception and production: Greek/English bilinguals and Thai bilabial stops. *International Journal of Bilingualism*, 5(2), 221–235.
- Best CT (1995) A direct realist view of cross-language speech perception. In: Strange W (ed.) *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Baltimore: York Press, pp. 171–204.
- Bialystok E, Craik F and Luk G (2008) Cognitive control and lexical access in younger and older bilinguals. *Journal of experimental psychology. Learning, memory, and cognition* 34(4): 859–873.
- Boersma P and Weenink D (2010) *PRAAT*. University of Amsterdam. Available at: <http://www.praat.org>.
- Cabrelli J, Luque A and Finestrat-Martínez I (2019) Influence of L2 English phonotactics in L1 Brazilian Portuguese illusory vowel perception. *Journal of Phonetics* 73: 55–69.
- Carlson MT, Goldrick M, Blasingame M, et al. (2016) Navigating conflicting phonotactic constraints in bilingual speech perception. *Bilingualism: Language and Cognition* 19(5): 939–954.
- Cheshire J, Kerswill P, Fox S, et al. (2011) Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15(2): 151–196.
- Clark HH and Fox Tree JE (2002) Using uh and um in spontaneous speaking. *Cognition* 84(1): 73–111.
- Cuetos F, Hallé P, Dominguez A, et al. (2011) Perception of prothetic /e/ in #sC utterances: Gating data. In: *ICPhS XVII Conference Proceedings*, Hong Kong, 17 August 2011.
- de Leeuw E and Bogulski CA (2016) Frequent L2 language use enhances executive control in bilinguals. *Bilingualism: Language and Cognition* 19(5): 907–913.
- de Leeuw E (2007) Hesitation markers in English, German, and Dutch. *Journal of Germanic Linguistics* 19(2): 85–114.

- 1
2
3 de Leeuw E and Celata C (2019) Plasticity of native phonetic and phonological
4 domains in the context of bilingualism. *Journal of Phonetics* 75: 88–93.
5
6 Dupoux E, Kakehi K, Hirose Y, et al. (1999) Epenthetic vowels in Japanese: A
7 perceptual illusion? *Journal of Experimental Psychology: Human Perception*
8 *and Performance* 25(6). American Psychological Association: 1568.
9
10 Escudero P (2005) *Linguistic Perception and Second Language Acquisition:*
11 *Explaining the Attainment of Optimal Phonological Categorization.*
12 University of Utrecht, Utrecht.
13
14
15 Flege JE (1987) The production of ‘new’ and ‘similar’ phones in a foreign language:
16 Evidence for the effect of equivalence classification. *Journal of Phonetics* 15:
17 47–65.
18
19
20 Flege JE (1995) Second language speech learning: Theory, findings, and problems.
21 In: Strange W (ed.) *Speech Perception and Linguistic Experience: Theoretical*
22 *and Methodological Issues.* Maryland: York Press, pp. 233–277.
23
24 Flege JE, Bohn O-S and Jang S (1997) Effects of experience on non-native speakers’
25 production and perception of English vowels. *Journal of Phonetics* 25(4):
26 437–470.
27
28
29 Flege JE, Schirru C and MacKay IRA (2003) Interaction between the native and
30 second language phonetic subsystems. *Speech Communication* 40(4): 467–
31 491.
32
33 Forster KI and Forster JC (2003) DMDX: A Windows display program with
34 millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*
35 35(1): 116–124.
36
37
38 Fowler CA (1996) Listeners do hear sounds, not tongues. *The Journal of the*
39 *Acoustical Society of America* 99(3): 1730–1741.
40
41 Fox Tree JE (2001) Listeners’ uses of um and uh in speech comprehension. *Memory &*
42 *Cognition* 29(2): 320–326.
43
44 Freeman MR, Blumenfeld HK and Marian V (2016) Phonotactic constraints are
45 activated across languages in bilinguals. *Frontiers in Psychology* 7.
46
47 Gollan TH, Montoya RI and Werner GA (2002) Semantic and letter fluency in
48 Spanish-English bilinguals. *Neuropsychology* 16(4): 562–576.
49
50
51 Grogan A, Green DW, Ali N, et al. (2009) Structural correlates of semantic and
52 phonemic fluency ability in first and second languages. *Cerebral Cortex*
53 19(11): 2690–2698.
54
55
56 Grosjean F (1998) Studying bilinguals: Methodological and conceptual issues.
57 *Bilingualism: Language and Cognition* 1(02): 131–149.
58
59 Gullberg M and Indefrey P (2003) Language History Questionnaire. Developed in
60 The Dynamics of Multilingual Processing. Max Planck Institute for

- Psycholinguistics, Nijmegen. Available at:
<http://www.mpi.nl/departments/other-research/research-projects/the-dynamics-of-multilingual-processing/tools/Lang-Hist-Quest-Engl.pdf>.
- Hallé PA, Dominguez A, Cuetos F, et al. (2008) Phonological mediation in visual masked priming: Evidence from phonotactic repair. *Journal of Experimental Psychology: Human Perception and Performance* 34(1): 177–192.
- Hualde JI (2005) *The Sounds of Spanish*. Cambridge, UK. Available at:
<http://www.cambridge.org/gb/academic/subjects/languages-linguistics/semantics-and-pragmatics/sounds-spanish> (accessed 3 January 2018).
- Kartushina N and Frauenfelder UH (2014) On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology* 5.
- Kivistö-de Souza H (2015) *Phonological awareness and pronunciation in a second language*. PhD. Universitat de Barcelona, Barcelona. Available at:
<https://core.ac.uk/download/pdf/78632484.pdf>.
- Klein-Braley C (1985) A cloze-up on the C-Test: a study in the construct validation of authentic tests. *Language Testing* 2(1): 76–104.
- Kuhl PK (2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* 5(11): 831–843.
- Liberman AM and Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21(1): 1–36.
- Llisterri J (1995) Relationships between speech production and speech perception in a second language. In: *Proceedings of the 13th International Congress of Phonetic Sciences*, Stockholm, 1995.
- Maclay H and Osgood CE (1959) Hesitation Phenomena in Spontaneous English Speech. *WORD* 15(1): 19–44.
- Martínez-Celdrán E, Fernández-Planas AM and Carrera-Sabaté J (2003) Castilian Spanish. *Journal of the International Phonetic Association* 33(2): 255–259.
- Mills L (2014) *An articulatory study of GOOSE-fronting in varieties of English*. PhD. Queen Margaret University, Edinburgh. Available at:
<https://eresearch.qmu.ac.uk/handle/20.500.12289/7678> (accessed 4 September 2018).
- Miyake A, Friedman NP, Emerson MJ, et al. (2000) The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology* 41(1): 49–100.
- Morales Pech D and Izquierdo J (2011) L2 phonology learning among young adult learners of English: Effects of regular classroom-based instruction and L2

- 1
2
3 proficiency. *Revista Electrónica de Investigación Educativa* 13(1). Available
4 at: <http://redie.uabc.mx/vol12no2/contentsmoralesizquierdo.html>.
5
6
7 Newcombe F (1969) *Missile Wounds of the Brain. A Study of Psychological Deficits*.
8 Oxford, UK: Oxford University Press.
9
10 Ohala JJ (1996) Speech perception is hearing sounds, not tongues. *The Journal of the*
11 *Acoustical Society of America* 99(3): 1718–1725.
12
13 Parlato-Oliveira E, Christophe A, Hirose Y, et al. (2010) Plasticity of illusory vowel
14 perception in Brazilian-Japanese bilinguals. *The Journal of the Acoustical*
15 *Society of America* 127(6): 3738–3748.
16
17 Portocarrero JS, Burrig RG and Donovick PJ (2007) Vocabulary and verbal fluency
18 of bilingual and monolingual college students. *Archives of clinical*
19 *neuropsychology : The official journal of the National Academy of*
20 *Neuropsychologists* 22(3): 415–422.
21
22
23 Powell J (2016) *Why We Love Music: From Mozart to Metallica - The Emotional*
24 *Power of Beautiful Sounds*. London: John Murray. Available at:
25 [http://www.hodder.co.uk/HodderStoughton/books/detail.page?isbn=97814736](http://www.hodder.co.uk/HodderStoughton/books/detail.page?isbn=9781473613751)
26 [13751](http://www.hodder.co.uk/HodderStoughton/books/detail.page?isbn=9781473613751) (accessed 10 September 2018).
27
28
29 Raatz U and Klein-Braley C (1981) *The C-Test--A Modification of the Cloze*
30 *Procedure*. Not available separately; see FL 013 014.
31
32 Rubinstein JS, Meyer DE and Evans JE (2001) Executive control of cognitive
33 processes in task switching. *Journal of Experimental Psychology: Human*
34 *Perception and Performance* 27(4): 763–797.
35
36 Sandoval TC, Gollan TH, Ferreira, VS and Salmon, DP (2010). What causes the
37 bilingual disadvantage in verbal fluency: The dual-task analogy. *Bilingualism:*
38 *Language and Cognition*, 13: 231–252.
39
40
41 Sauzéon H, Raboutet C, Rodrigues J, et al. (2011) Verbal knowledge as a
42 compensation determinant of adult Age differences in verbal fluency tasks
43 over time. *Journal of Adult Development* 18(3): 144–154.
44
45 Sheldon A and Strange W (1982) The acquisition of /r/ and /l/ by Japanese learners of
46 English: Evidence that speech production can precede speech perception.
47 *Applied Psycholinguistics* 3(3): 243–261.
48
49 Shriberg EE and Lickley RJ (1993) Intonation of Clause-Internal Filled Pauses.
50 *Phonetica* 50(3): 172–179.
51
52
53 So CK and Best CT (2010) Cross-language perception of non-native tonal contrasts:
54 Effects of native phonological and phonetic influences. *Language and Speech*
55 53(2): 273–293.
56
57 So CK and Best CT (2014) Phonetic influences on English and French listeners’
58 assimilation of mandarin tones to native prosodic categories. *Studies in Second*
59 *Language Acquisition* 36(2): 195–221.
60

1
2
3 Tomás Navarro T (1918) *Manual de Pronunciación Española*. Reprinted in 1982.
4 Madrid: CSIC.
5

6
7 Zampini ML (1998) The relationship between the production and perception of L2
8 Spanish stops. *Texas Papers in Foreign Language Education*: 1–18.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Appendix I

Language use question regarding daily use

Consider your present amount of English and Spanish language use. Divide 100% into English and Spanish use, e.g. if you only speak English presently, write "English=100%; Spanish=0%". If you also use another language, or other languages, include them in your division of 100%.

Alternatively, if you speak both English and Spanish equally, write "English=50%; Spanish=50%". Divide according to your own estimates, considering all aspects of language use, e.g. both passive (watching television) and active (speaking with parents).

C-tests

On the next pages you will find 2 small texts in total. Each text contains gaps where parts of some words have been left out (no whole words are missing, though). Please try and fill in the gaps. In many cases there are several possibilities, so there are no right or wrong answers. Thank you very much for your help.

Text 1:

Two former US navy ships contaminated with chemicals were expected to arrive in the English Channel last night. The Maritime and Coastguard Agency sa _____ the ves _____, at the cen _____ of an enviro _____ row, we _____ being to _____ through the cha _____, before hea _____ up the east co _____ to Hartlepool. Pl _____ to dism _____ them in north-east England have been she _____ after being dee _____ to fl _____ international ru _____. Last we _____, the gover _____ said the ships could be sto _____ in Hartlepool before go _____ back acr _____ the Atlantic.

Two former US navy ships contaminated with chemicals were expected to arrive in the English Channel last night. The Maritime and Coastguard Agency says the vessels, at the centre of an environmental row, were being towed through the channel, before heading up the east coast to Hartlepool. Plans to dismantle them in north-east England have been shelved after being deemed to flaunt international rules. Last week, the government said the ships could be stowed/stored in Hartlepool before going back across the Atlantic.

Text 2:

Don't get me wrong. I love magazines. I've been addicted to them since my teenage years. There's some _____ about wom _____ magazine superfi _____ that I of _____ enjoy. But

1
2
3 oh boy _____, they are just _____ so, so
4 frustr_____ predictable. I reckon you could
5 cobble o_____ together very
6 eas_____ in five minutes. Take the
7 cover for example: the cover image, get a
8 head and shoulder shot of a smiling,
9 heavily make-uped and airbrushed
10 model (or optionally a famous person).

11 Don't get me wrong. I love magazines. I've been addicted to them since my teenage
12 years. There's something about women's magazine superficiality that I often enjoy. But
13 oh boy, they are just so, so frustratingly predictable. I reckon you could cobble one
14 together very easily in five minutes. Take the cover for example: the cover image, get
15 a head and shoulder shot of a smiling, heavily make-uped and airbrushed model (or
16 optionally a famous person).

Appendix II

Participant	Recorded in Spain (2) vs UK (1)	Amount of Daily English Use (%)	C-Test Result	Age of English Acquisition
1	2	25	32.5	4
2	2	20	44	6
3	2	30	55	6
4	2	20	79	7
5	2	5	46.5	6
6	1	40	51.5	5
7	2	70	31.5	6
8	1	10	16.5	30
9	2	40	28	11
10	2	35	45	8
11	1	60	70	7
12	2	40	40	1
13	1	20	6.5	10
14	2	25	60	3
15	1	40	40.25	5
16	2	20	37.5	8
17	2	30	46.5	8
18	2	30	30.5	10
19	1	30	33	7
20	2	20	21.5	6
21	2	20	11.5	7
22	2	40	27.5	7
23	2	40	30	3
24	2	1	7.5	5
25	2	30	31.5	5
26	2	40	62.5	7
27	1	80	51.25	7
28	2	35	34.5	8
29	2	35	33	6
30	2	5	30.5	6
31	2	40	32.5	10
32	2	0	46.5	10
33	2	0	44	11
34	2	30	41.5	4

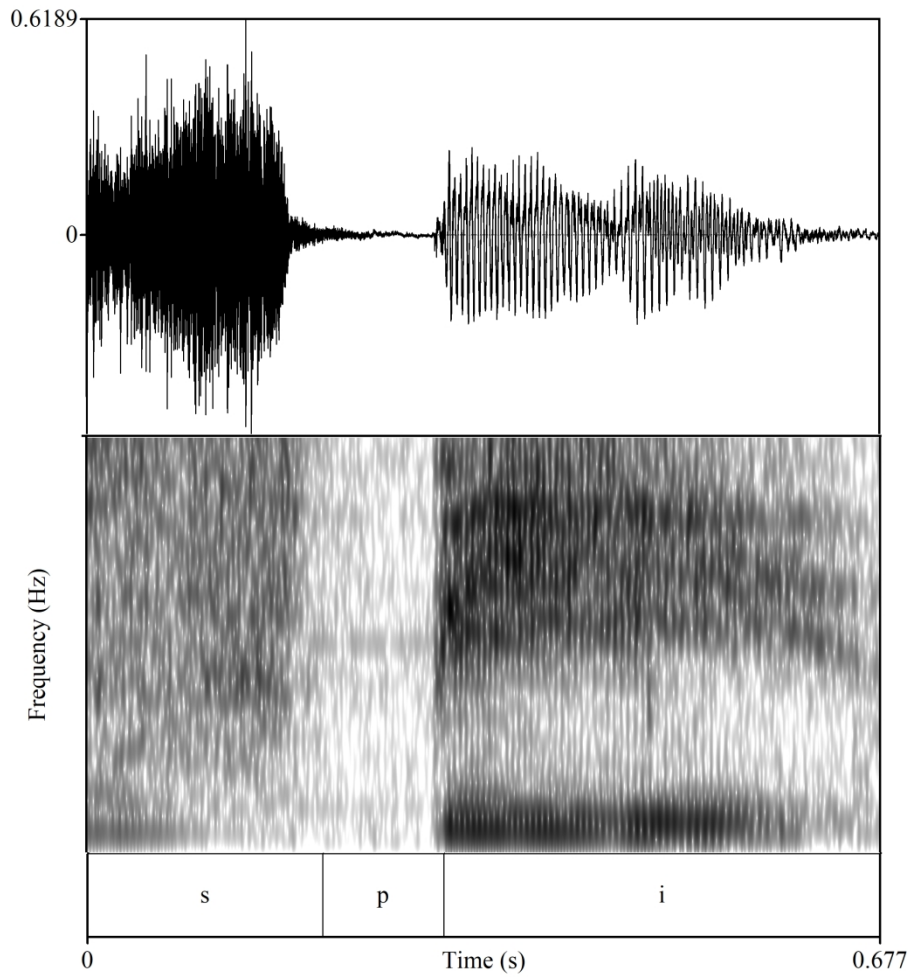


Figure 1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

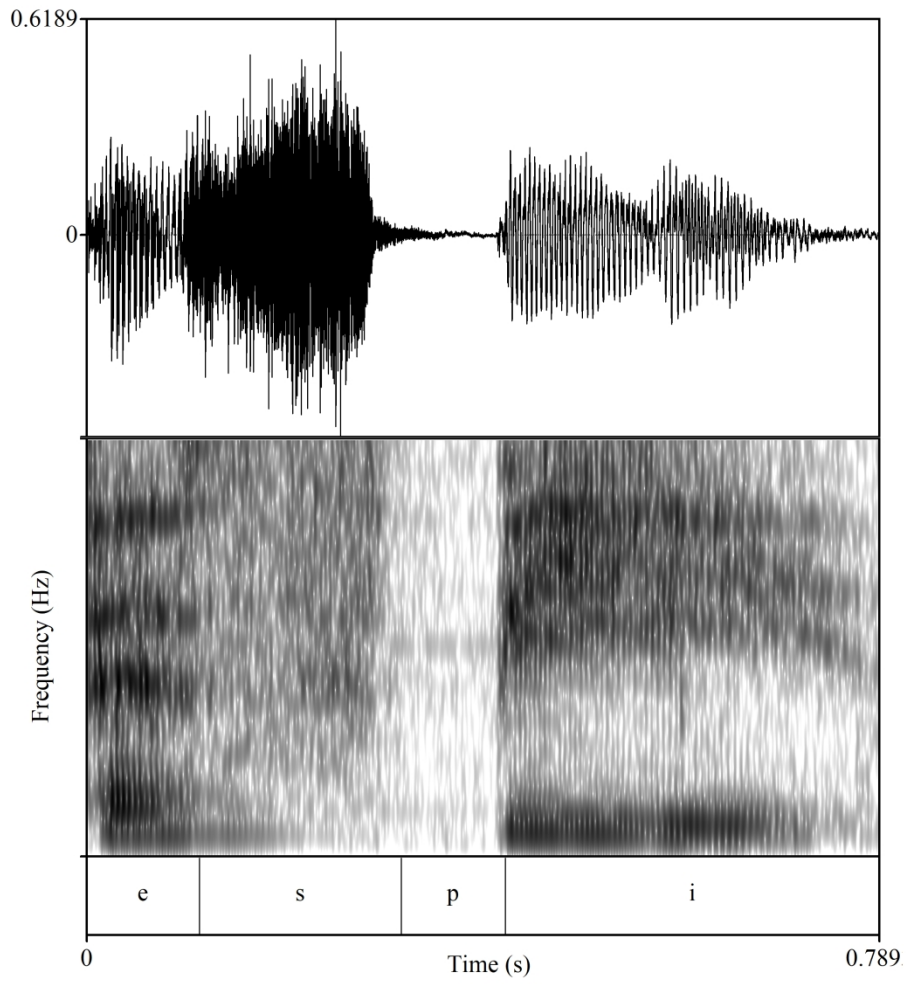


Figure 2

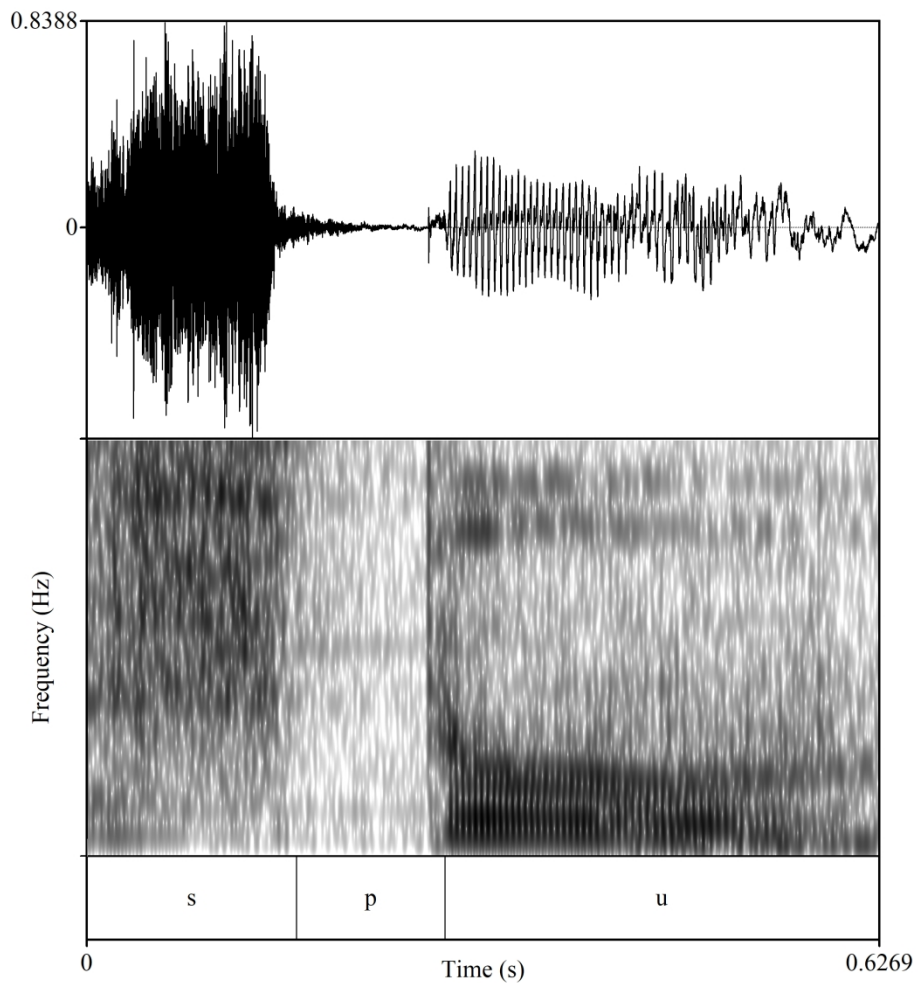


Figure 3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

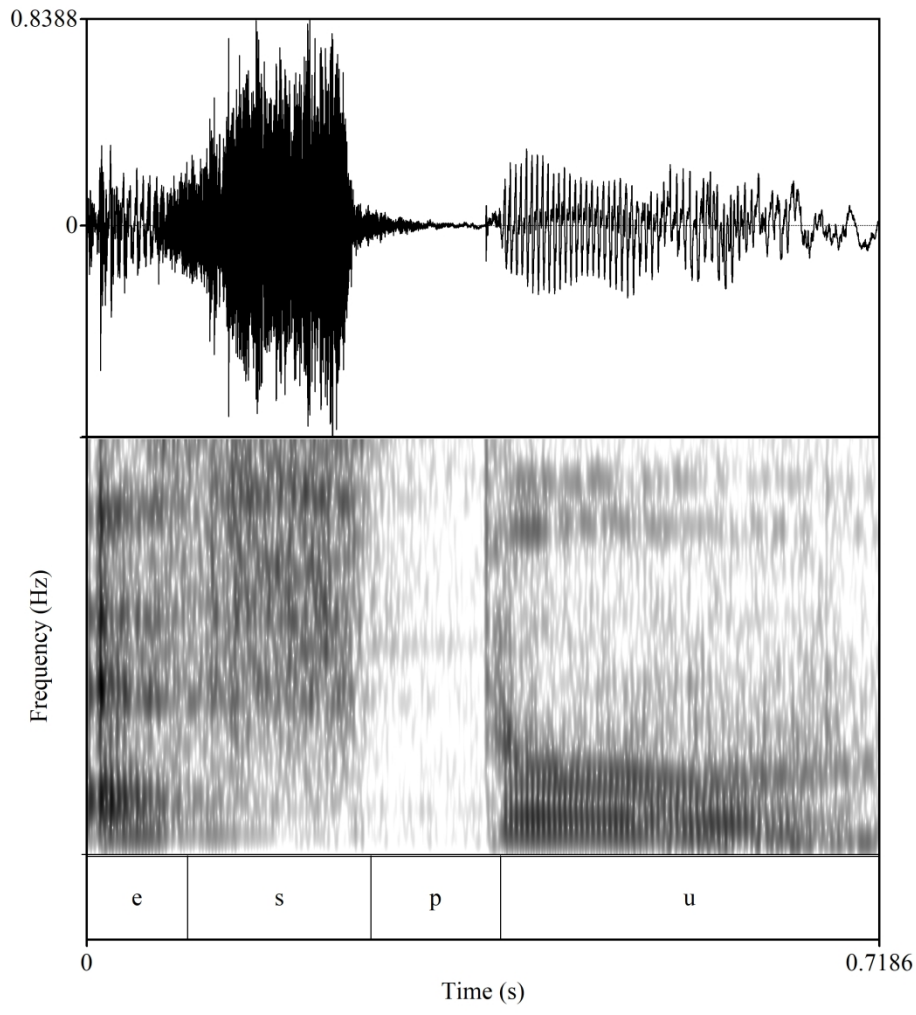


Figure 4

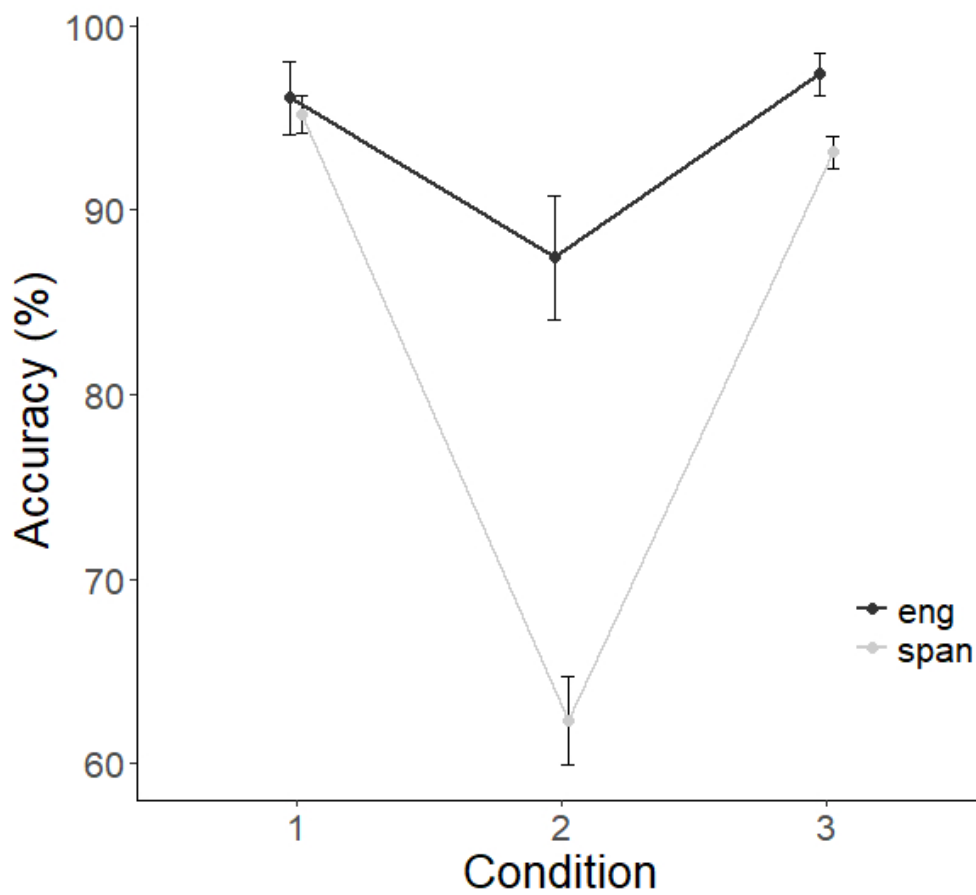


Figure 5

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

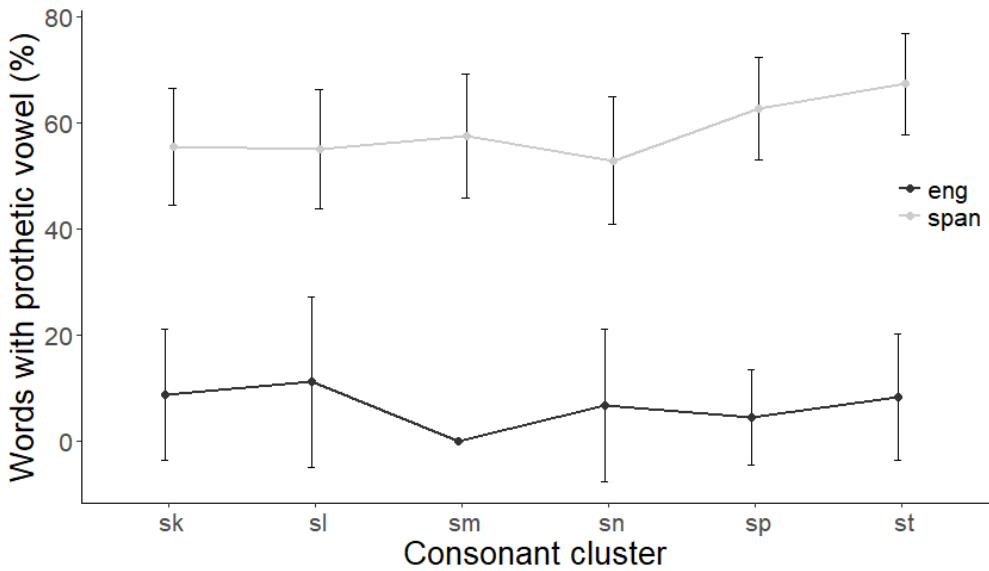


Figure 6

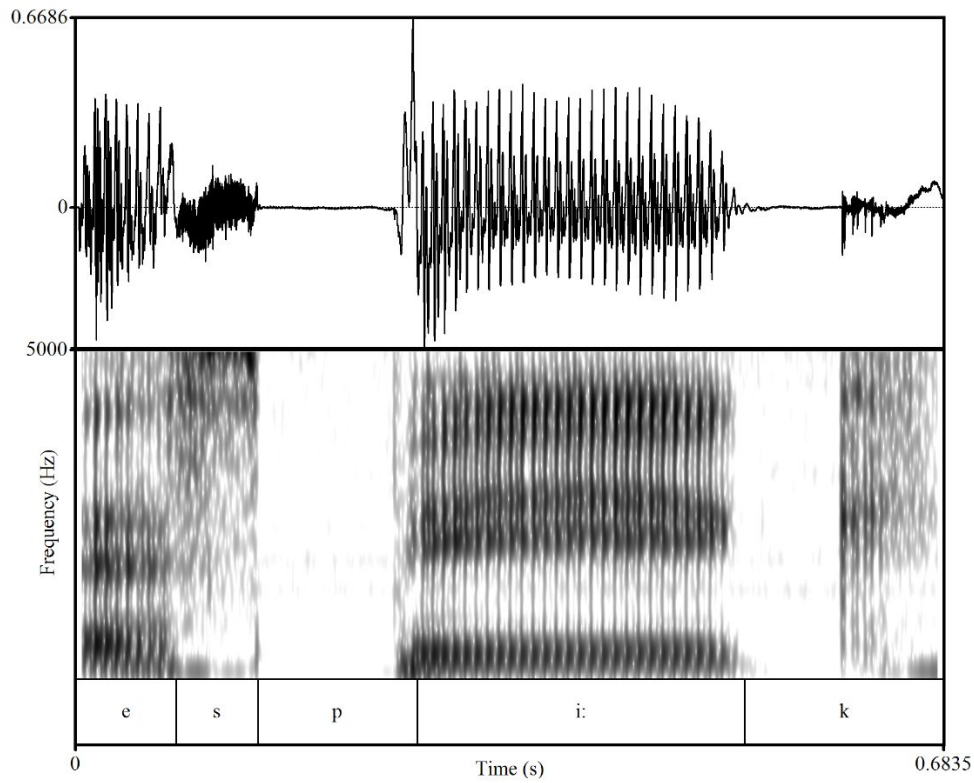


Figure 4: Example of a prothetic vowel, 'speak', produced by Participant 18.

Peer Review

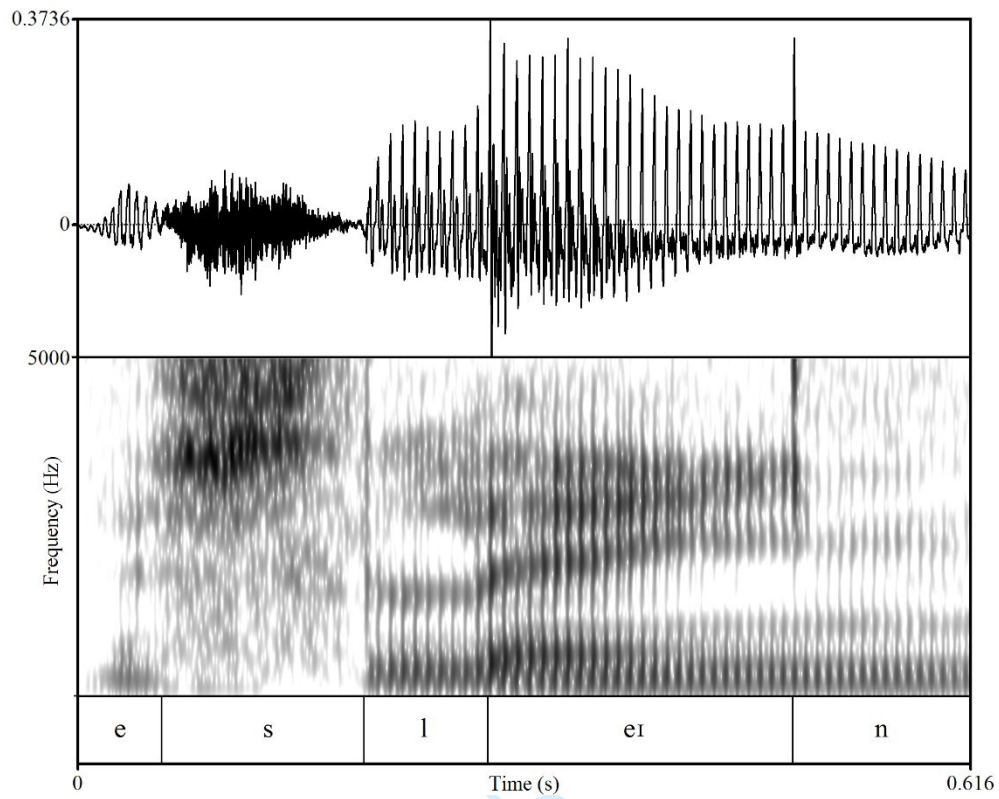


Figure 5: Example of a prothetic vowel, 'slain' produced by Participant 16.

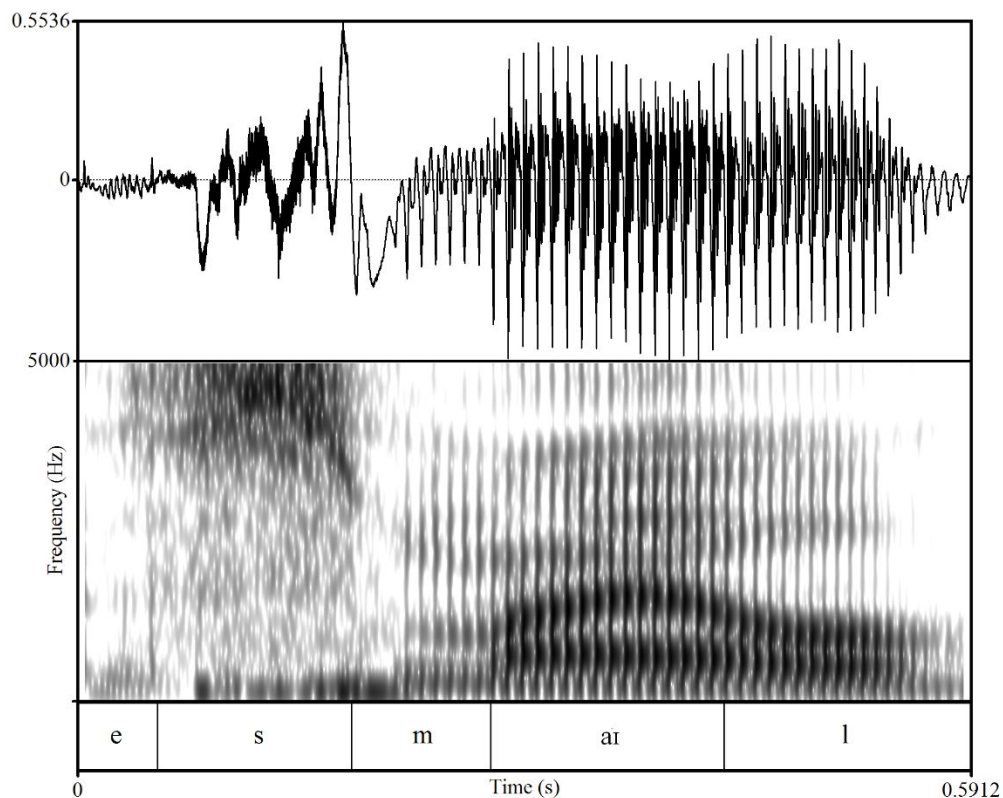


Figure 6: Example of a prothetic vowel, 'smile' produced by Participant 1.

Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

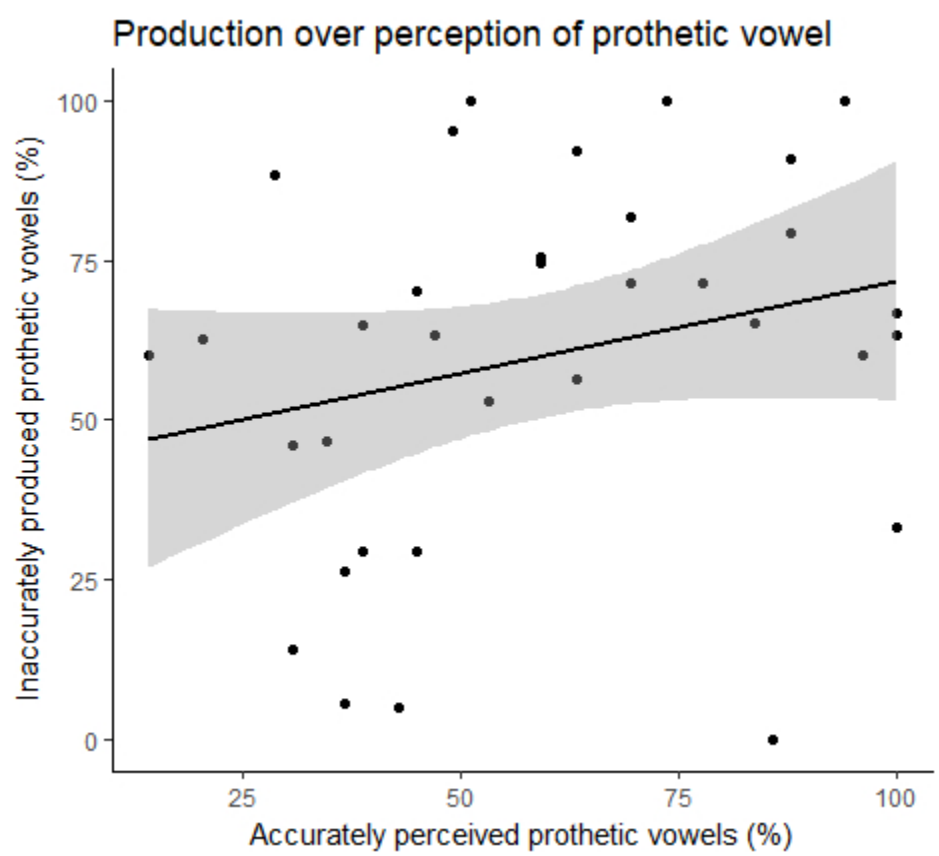


Figure 10

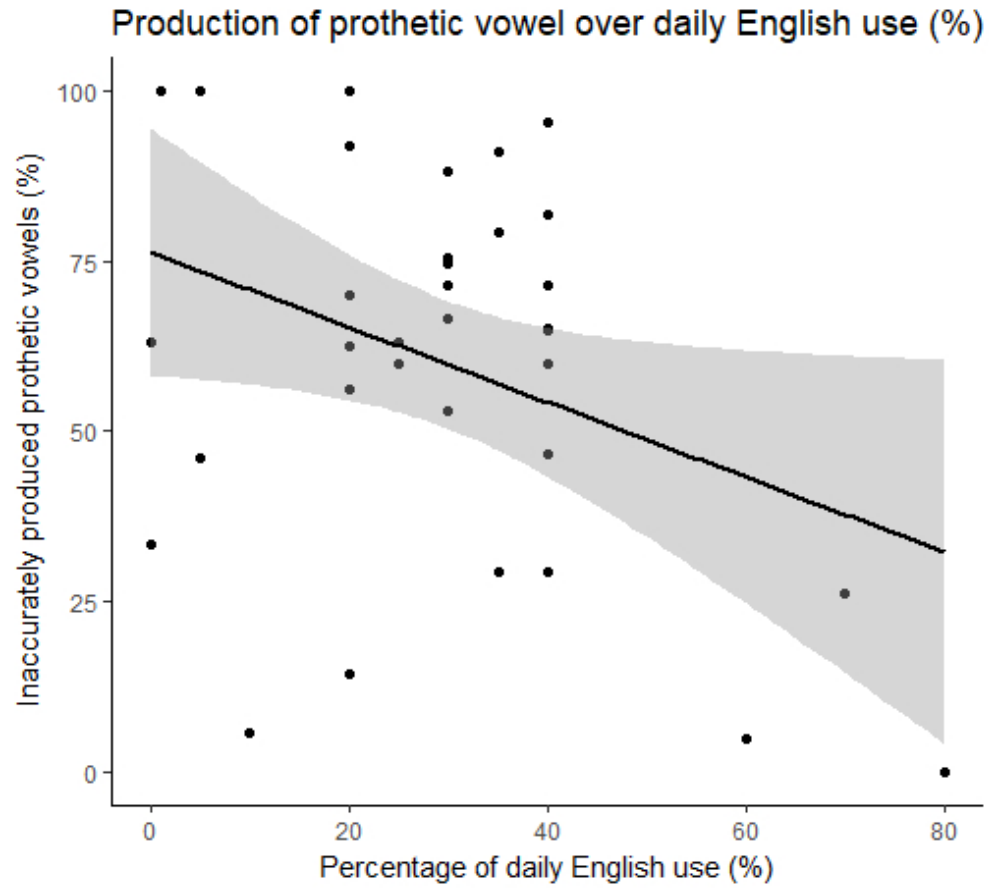


Figure 11

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

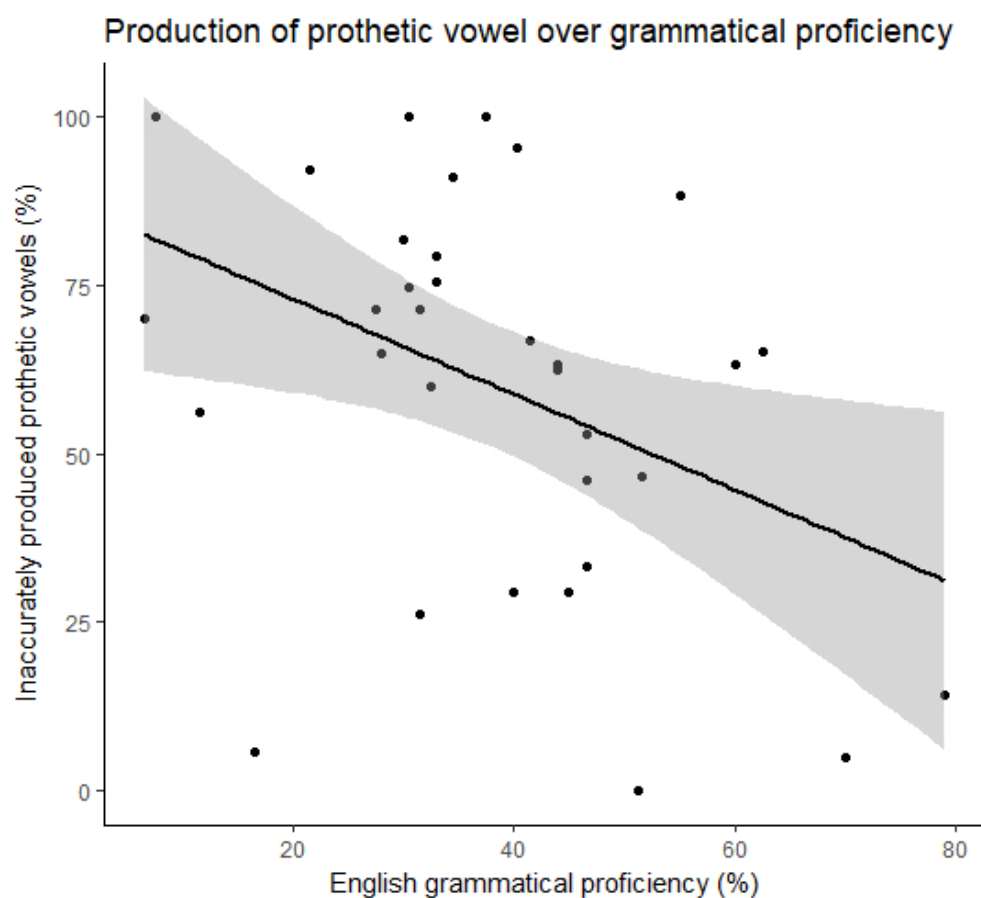


Figure 12